Problems of Upward and Downward Compatibility in Multichannel Stereo Systems

3404 (J-6)

Michael A. Gerzon
Technical Consultant
Oxford, UK

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

Problems of Upward and Downward Compatibility in Multichannel Stereo Systems

Michael A. Gerzon

Technical Consultant, 57 Juxon St., Oxford OX2 6 DJ, U.K.

Abstract

This paper reviews, in the context of multispeaker stereo systems using up to five loudspeakers, the problem of determining the best upward and downward compatibility matrixing for other reproduction modes. Published HDTV compatibility matrices are shown to take an inadequate account of three factors: (i) psychoacoustic quality of localisation (ii) the effect of using different panpot laws, and (iii) the problem of cascadability, i.e. the losses causes by repeated conversions along a long production chain between different reproduction modes.

0. INTRODUCTION

In references [1] and [2], Meares and Theile report on recent work in standardising HDTV stereo systems using anything up to eight transmission channels, including various different proposals for "upconversion" and "downconversion" between different reproduction modes using different numbers of loudspeakers. In three recent papers, [3] to [5], the author reported on detailed work on the problem for the case of frontal-stage stereo systems using from 1 to 5 loudspeakers across a frontal stereo stage. Since, especially in refs. [4] and [5], this work was highly technical and mathematical, it is one aim of this paper to report on the conclusions in a much more concrete manner, and with reference to other proposals.

There are three areas in which the proposals under discussion by Meares [1] and Theile [2] present problems, largely because older ideas of "compatibility" between different reproduction modes no longer prove adequate in an environment where, as reported by Meares [1], up to ten reproduction modes are being considered. Studies by the author have indicated that, when Ambisonic and with-height reproduction modes are also considered, future reproduction hierarchies for HDTV may have to consider seriously up to about 30 different modes, all of which must be rendered mutually compatible both at the production stage and at the reproduction stage.

In this paper, for simplicity, we confine our study to the frontal stage reproduction case reported in [3] to [6], although in future papers we shall study much more complex surround-sound and Ambisonic hierarchies.

The three kinds of problems with prior approaches lie in the following areas. First, it is necessary, especially in "upconversion" from one number of stereo speakers to a greater number of stereo speakers that the reproduced effect via the larger number of speakers have optimal psychoacoustic quality for a wide variety of source material. Second, especially when more than two speakers are used, there exist several different possible ways of panning sounds between the three or more loudspeakers, not just "pairwise panning" between adjacent pairs of speakers, and it is important that any upward or downward conversion matrix should work well for all these different panpot laws [6]. Third, there is the problem of cascadability, discussed in [3], but not considered in [1] and [2], whereby the effect of repeated up- and down-conversion of signals between different reproduction modes in a long broadcast production chain must be controlled. It is unacceptable if a few cascaded conversion stages result in a poor or unbalanced sound, even if one conversion, on its own, results in acceptable "compatibility"

The problem of cascadability, or multistage compatibility, is one that in other contexts is of great concern to broadcasters, since broadcast signals are often relayed, rebroadcast and reprocessed many times down a long production chain. Yet the problem of cascadability seems barely to have been considered in the multichannel audio context, possibly because hitherto, with mono/2-channel stereo, cascadability is trivial and automatic. This is no longer automatic in the case of more complex hierarchies.

In this paper, we shall show, by means of simple examples, that some prior art apparently "compatible" up- and down-conversion matrices are not cascadable, i.e. do not have multistage compatibility. This problem applies whether the preferred transmission modes use the so-called "compatibility matrixing" approach in which additional information is represented by added transmission channels, or the "downward mixing" approach in which speaker feed signals for the most complex reproduction mode are transmitted and simpler modes are mixed down from these signals.

1. CASCADABILITY

Figure 1 shows frontal-stage stereo loudspeaker layouts using one to five loudspeakers (We adopt the convention of refering to mono as "one-speaker stereo" to avoid making exceptions for the mono case). We assume that all loudspeakers in reference or monitoring speaker layouts are all at the same distance from a listener in the ideal stereo seat, and that the angle between adjacent speakers in any given speaker layout is the same as for any other adjacent pair. The notations for the different speaker feeds in each layout is as shown in figure 1.

The problem of cascadability can be illustrated by reference to the problem of converting between 3-speaker stereo and 4-speaker stereo. In ref. [1], the following up-conversion from 3 to 4 loudspeakers is suggested:

$$L_4 = L_3$$
$$L_5 = R_5 = 0.7071\ C_3$$
$$R_4 = R_3 \qquad (1)$$

so that the outer speaker feeds remain as in the 3 speaker case, but the centre speaker feed is split among the 2 central speakers of the 4-speaker layout so as to reproduce with the same total energy. Apart from the author [3-5], there seems to be a general concensus that equ. (1) is a desirable 4×3 upconversion matrix.

The naive down conversion matrix from 4 speakers to 3 speakers:

$$L_3 = L_4$$
$$C_3 = 0.7071(L_5 + R_5)$$
$$R_3 = R_4 \qquad (2)$$

has also been proposed [1,2], but suffers from the disadvantage that as sounds are panned from left to right across the 4-speaker stage, they dwell at the fixed centre position of the 3-speaker stage for one-third of the time, resulting in an unnatural build-up of sounds at the central position. While proposal (2) has generally been rejected [1,2] as a 3×4 downconversion matrix for this reason, it does have the property that if, as shown in figure 2, one upconverts a 3-channel signal to 4 speakers via equ. (1), and then downconverts it back to 3 speakers via equ. (2), the final 3-speaker feeds are the same as what one started with.

Meares has [1,2] suggested a more sophisticated 3×4 downconversion matrix that largely prevents the centre-dwell problem for panned 4-speaker sounds, by panning the L_5 and R_5 speaker feeds into the 3-channel stage at about ⅓-left and ⅓-right, giving either

$$L_3 = L_4 + 0.4472\ L_5$$
$$C_3 = 0.8944\ (L_5 + R_5)$$
$$R_3 = R_4 + 0.4472\ R_5 \qquad (3)$$

or

$$L_3 = L_4 + 0.5000\ L_5$$
$$C_3 = 0.8660\ (L_5 + R_5)$$
$$R_3 = R_4 + 0.5000\ R_5 \qquad (4)$$

However, while 3×4 downconversion matrices of the forms equ. (3) or (4) improve the stereo effect for simply pairwise panned 4-channel signals via 3 speakers, they create problems if fed with 4×3 upconverted signals of the form of equ. (1). For example, if the 3×4 downconversion matrix of equ. (3) follows the 4×3 upconversion matrix of equ. (1) (as shown in fig. 2), one gets a final 3-speaker feed

$$L_3' = L_3 + 0.3162\ C_3$$
$$C_3' = 1.2649\ C_3$$
$$R_3' = R_3 + 0.3162\ C_3\ , \qquad (5)$$

which increases the relative level of the C_3 signals by 2.55 dB, as well as causing it to crosstalk (at a relative level of - 12.04 dB) onto each of the two outer speakers.

The fact that upconversion followed by downconversion changes the original 3-speaker sound to such a degree is quite serious, since it means that programme material mixed for 3-speaker stereo cannot be mixed with material prepared for 4-speaker stereo via an upconversion matrix such as shown in figure 3 without altered results if downconverted again later for 3-speaker reproduction.

Repeat up- and down-conversions alter the sound even more seriously. For example, if an original 3-speaker sound is subject to the 4×3 upconversion and the 3×4 downconversion, using equs. (1) and (3), shown in fig. 2 twice, then the result is given by

$$L_3'' = L_3 + 0.7162\ C_3$$
$$C_3'' = 1.6000\ C_3$$
$$R_3'' = R_3 + 0.7162\ C_3\ , \qquad (6)$$

which increases the relative level of C_3 signals by a massive 5.55 dB, and causing the centre image to crossfeed onto each of the two outer speakers at a level of only - 6.98 dB. It is clear that using the 4×3 upconversion of equ. (1) with the downconversion of equ. (3) is totally unsatisfactory in a complex production environment where cascading is likely to occur. The use of equ. (4) to replace equ. (3) shows similar, and only slightly less bad, effects.

Suppose now that one seeks to find a 3×4 downconversion matrix that recovers the original L_3, C_3 and R_3 from equ. (1). Assuming left/right symmetry, such a 3×4 downconversion matrix has the general form

$$L_3' = L_4 + k(L_5-R_5)$$
$$C_3' = 0.7071(L_5+R_5)$$
$$R_3' = R_4 + k(R_5-L_5) \qquad (7)$$

for an arbitrary chosen constant k. Unfortunately, no value of k gives an adequate downconversion as sounds are panned from left to right across the 4-speaker stage. Although k = 0.15 gives some reduction of the centre-dwell effect at low frequencies for a central listener, the dwell effect remains at higher frequencies or for noncentral listeners, due to the left/right symmetry with which the energy of an L_5 sound is fed into 3 speakers by equ. (7). Thus we conclude that cascadability requires the use of a 4×3 upconversion matrix differing from that of equ. (1). Thus we need to consider a less naive approach to upconversion and downconversion.

2. UP- AND DOWN-CONVERSION CRITERIA

The above worked examples illustrate that one must be careful in formulating criteria for upconversion and downconversion matrices. The cascadability problem can be thought of as arising due to the use of more that one panpot law for 4-speakers stereo. Previous studies [1,7] have been based largely on the assumption that multispeaker stereo uses pairwise panning, i.e. constant power amplitude panning between adjacent pairs of speakers. This assumption should be abandoned for several reasons.

First, it is known that in some cases, e.g. Ambisonic reproduction of due side images [8,9] , pairwise panning does not give optimum phantom image quality, and in [6] and [3], it is suggested that even for frontal stage stereo, better panpot laws can be devised. This will be dealt with in detail in ref. [10].

However, other panpot laws occur naturally in any case, since the effect of pairwise panning 3-speaker stereo and then upconverting it to 4-speaker stereo is to produce a 4-speaker stereo signal <u>not</u> using a 4-speaker pairwise panning law. The cascadability problem partly arises because this panpot law is poor, due to the inadequate design of the upconversion matrix, and partly because the downconversion matrix was not adequately designed to cope with likely non-pairwise 4-speaker panning laws.

In ref. [4], methods were given for designing upconversion matrices that have the property of preserving localisation quality not just for pairwise panning panpot laws, but for a wide variety of other possible panpot laws. It is certainly not obvious <u>a priori</u> that such upconversion matrices exist that work over a wide variety of panpot laws, and it is perhaps a remarkable empirical discovery that the design procedure for upmatrices in ref. [4] preserve localisation quality well for a wide range of possible panpot laws. Whatever the imperfections of the stereo localisation theory used in [4], it does provide a tool for detailed investigation of the performance of upmatrices and downmatrices for a wide variety of panning laws.

For frontal stage stereo, it is considered that all upconversion matrices should preserve the total reproduced energy of all possible stereo signals, for reasons (some fairly obvious, some not) detailed in ref. [4]. It can be shown mathematically that this means that the columns of the upconversion matrices should be orthogonal nit-length vectors. In particular, this implies in general that some of the matrix coefficients will be negative. Previous upmatrix and downmatrix proposals (see for example Meares [1,7]) have generally only considered the use of positive coefficients, but in all realistic cases, cascadability requires the use of some negative coefficients.

The use of negative coefficients in upconversion and downconversion matrices is, in the experience of the author, often misunderstood. While it is generally accepted that negative coefficients can widen stereo images for central listeners at frequencies below around 700 Hz, it is generally believed that such negative coefficients can play little useful role for listeners well away from the stereo seat or at high audio frequencies. While it is perfectly true that antiphase speaker feeds have relatively little beneficial effect on localisation well away from the stereo seat, especially at higher frequencies, this overlooks another important effect.

The reason why negative matrix coefficients are so useful lies in the fact that the stereo signals being matrixed are not independent signals, but rather have strong in-phase components in common between two or more signals. The effect of using conversion matrices involving some negative coefficients is that the negative coefficients help to cancel out some of these originally in-phase common components in the final speaker feeds, giving a lower and better-distributed crosstalk pattern than would be the case had conversion matrices with only nonnegative coefficients been used.

This role of negative conversion matrix coefficients in cancelling out unwanted crosstalk for panned signals was what made the optimisation of upconversion matrices in ref. [4] possible. In fact, the psychoacoustic theory used in [4] assumed that at high frequencies or for very noncentral listeners, antiphase signals from speakers made no difference in localisation - and the occurence of negative coefficients in the matrices is mainly for cancellation purposes.

The prior art described by Meares and Theile [1,2,7] not only has required that only nonnegative upconversion and downconversion coefficients be used, but a second assumption is that the outer speaker signals feeds of a stereo arrangement should be fed after conversion only to the corresponding outer speakers, and to no others. We have already seen that this prevents cascadability in the 4×3 upconversion case. It is true that any crosstalk from the outer speakers narrows subsequent reproduction, especially for off-centre listeners, but stringent application of this no-crosstalk requirement for outer speakers causes considerable degradation of other stereo imaging requirements for phantom images. In the work of [4], a certain amount of outer-speaker crosstalk, and a corresponding narrowing of the reproduction stage, was accepted as a necessary price for improved imaging quality elsewhere, and the ability to cascade conversions freely.

It is obviously desirable that any crosstalk coefficients of conversion matrices for outer speaker feeds should be kept as small as possible to minimise loss of stereo width, and in particular, crosstalk coefficients to the opposite side of the stereo stage should be particularly small, and the design methods of [4] did indeed ensure this.

The cascadability requirement for up- and down-conversion matrices was dealt with in considerable mathematical detail in ref. [5], but mainly in the context of a "compatibility matrixing" approach which adds one new transmission channel for each extra speaker channel. However, the results of that paper are not confined to the "compatibility matrixing" approach, but can also be applied to the "downward mixing" approach.

If we use $R_{n_2 n_1}$ to denote the conversion matrix from n_1-speaker stereo to n_2-speaker stereo, then we can symbolically write the cascadability requirement as follows:

If $n_2 \geqslant \min(n_1, n_3)$, then :

$$R_{n_3 n_2} R_{n_2 n_1} = R_{n_3 n_1} \; . \tag{8a}$$

and if $n_2 \geqslant n_1$, then :

$$R_{n_1 n_2} \; R_{n_2 n_1} = I_{n_1 n_1} \tag{8b}$$

where I_{nn} is the $n \times n$ identity matrix.

It is possible to deduce mathematically, as theorems deduced from equs. (8), that the effect of repeatedly cascading conversion matrices for various intermediate values n of speaker channels is equivalent to just one conversion from the initial number of speaker channels to the smallest intermediate number of speaker channels, followed by a second conversion from that smallest number up to the output number of speaker channels. Thus, if the cascadability requirements of equs. (8) are satisfied, the reproduction results can never be worse than those caused by the smallest "bottleneck number" n_B of speaker channels at intermediate stages in the production chain, with no other source of cascade losses (assuming otherwise perfect audio quality).

A second result that can easily be proved from equs. (8) is that for $n_2 \leqslant n_1$, $R_{n_1 n_2} R_{n_2 n_1}$ is idempotent, i.e.

$$(R_{n_1 n_2} R_{n_2 n_1})^2 = R_{n_1 n_2} R_{n_2 n_1} \quad . \tag{9}$$

The problem of designing conversion matrices that satisfy the cascadability rules of equ. (8) is not a hard problem to solve provided that one is prepared to use mathematical arguments from matrix algebra, but is not "elementary" without such mathematical aids. Given a desired choice of the upconversion matrices $R_{n+1\,n}$ converting from n-speaker stereo to (n+1)-speaker stereo, ref. [5] figure 10 can be shown to determine the general form of all the conversion matrices $R_{n_2 n_1}$ satisfying the cascadability rules of equs. (8), where the conversion matrix is the result of encoding n_1-speaker stereo into transmission channel signals and of decoding them back again into n_2 speaker signals.

Thus, whether or not the explicit transmission channels discussed in ref. [5] are used, the methods of ref. [5] can be used to generate cascadable conversion matrices.

In this paper, we shall discuss only those special solutions generated by the "orthogonal matrix" solutions discussed in ref. [5]. Although, for given upconversion matrices, it was shown in ref. [5] that not all solutions need be orthogonal, it has been found that the nonorthogonal solutions generally have worse downward compatibility than orthogonal solutions. This is fortunate, since the orthogonal solutions are also much simpler to describe.

3. AN ORTHOGONAL CONVERSION HIERARCHY

It was shown in ref. [4] that for reasonable speaker layouts, the following upconversion matrices are subjectively exceptionally good performers, giving substantially optimal preservation of the originally intended stereo effect via a larger number of speakers.

<u>3 × 2 upconversion matrix</u> R_{32}

This case involves, for best subjective results, the use of a frequency-dependent conversion matrix as follows:

$$\begin{pmatrix} L_3 \\ C_3 \\ R_3 \end{pmatrix} = \begin{pmatrix} 0.8536 + 0.0607\,A & -0.1464 + 0.0607\,A \\ 0.5000 - 0.0858\,A & 0.5000 - 0.0858\,A \\ -0.1464 + 0.0607\,A & 0.8536 + 0.0607\,A \end{pmatrix} \begin{pmatrix} L_2 \\ R_2 \end{pmatrix} \qquad (10)$$

where A is an all-pass network gain having gain -1 below 5 kHz and +1 above 5 kHz. Putting A = 0 gives a reasonable frequency-independent upconversion matrix, although not as good as the frequency-dependent case.

<u>4 × 3 upconversion matrix</u> R_{43}

$$\begin{pmatrix} L_4 \\ L_5 \\ R_5 \\ R_4 \end{pmatrix} = \begin{pmatrix} 0.9303 & -0.1297 & 0.0527 \\ 0.3314 & 0.6951 & -0.1479 \\ -0.1479 & 0.6951 & 0.3314 \\ 0.0527 & -0.1297 & 0.9303 \end{pmatrix} \begin{pmatrix} L_3 \\ C_3 \\ R_3 \end{pmatrix} \qquad (11)$$

<u>5 × 4 upconversion matrix</u> R_{54}

$$\begin{pmatrix} L_6 \\ L_7 \\ C_5 \\ R_7 \\ R_6 \end{pmatrix} = \begin{pmatrix} 0.9535 & -0.1084 & 0.0590 & -0.0324 \\ 0.2533 & 0.7870 & -0.1989 & 0.0859 \\ -0.1349 & 0.5708 & 0.5708 & -0.1349 \\ 0.0859 & -0.1989 & 0.7870 & 0.2533 \\ -0.0324 & 0.0590 & -0.1084 & 0.9535 \end{pmatrix} \begin{pmatrix} L_4 \\ L_5 \\ R_5 \\ R_4 \end{pmatrix} \qquad (12)$$

<u>Other upconversion matrices</u>

Other upconversion matrices are preferably formed by cascading the above three matrices. This yields the following "composite" upconversion matrices.

<u>4 × 2 upconversion matrix</u> R_{42}

$$\begin{pmatrix} L_4 \\ L_5 \\ R_5 \\ R_4 \end{pmatrix} = \begin{pmatrix} 0.7215 + 0.0708\,A & -0.1561 + 0.0708\,A \\ 0.6521 - 0.0485\,A & 0.1728 - 0.0485\,A \\ 0.1728 - 0.0485\,A & 0.6521 - 0.0485\,A \\ -0.1561 + 0.0708\,A & 0.7215 + 0.0708\,A \end{pmatrix} \begin{pmatrix} L_2 \\ R_2 \end{pmatrix} \qquad (13)$$

where as before, A is preferably an all-pass with gain -1 below 5 kHz and gain +1 above 5 kHz, or where A = 0 in the frequency independent case.

<u>5 × 2 upconversion matrix</u> R_{52}

$$\begin{bmatrix} L_6 \\ L_7 \\ C_5 \\ R_7 \\ R_6 \end{bmatrix} = \begin{bmatrix} 0.6325 + 0.0676\,A & -0.1525 + 0.0676\,A \\ 0.6482 - 0.0045\,A & 0.0287 - 0.0045\,A \\ 0.3945 - 0.0745\,A & 0.3945 - 0.0745\,A \\ 0.0287 - 0.0045\,A & 0.6482 - 0.0045\,A \\ -0.1525 + 0.0676\,A & 0.6325 + 0.0676\,A \end{bmatrix} \begin{bmatrix} L_2 \\ R_2 \end{bmatrix} \tag{14}$$

where as before A is an all-pass with gain -1 below 5 kHz and gain +1 above 5 kHz, or where A = 0 in the frequency-independent case.

<u>5 × 3 upconversion matrix</u> R_{53}

$$\begin{bmatrix} L_6 \\ L_7 \\ C_5 \\ R_7 \\ R_6 \end{bmatrix} = \begin{bmatrix} 0.8407 & -0.1538 & 0.0557 \\ 0.5304 & 0.3648 & -0.0891 \\ -0.0279 & 0.8285 & -0.0279 \\ -0.0891 & 0.3648 & 0.5304 \\ 0.0557 & -0.1538 & 0.8407 \end{bmatrix} \begin{bmatrix} L_3 \\ C_3 \\ R_3 \end{bmatrix} \tag{15}$$

<u>Downconversion matrices</u>

The downconversion matrices for this case are obtained by putting A = 0 in the above and taking the matrix transpose (i.e. turning rows into columns and vice-versa). We warn that this "transpose property" is special to the orthogonal hierarchy case, and does not generalise. Thus we get the following downconversion matrices.

<u>2 × 3 downconversion matrix</u> R_{23}

$$\begin{bmatrix} L_2 \\ R_2 \end{bmatrix} = \begin{bmatrix} 0.8536 & 0.5000 & -0.1464 \\ -0.1464 & 0.5000 & 0.8536 \end{bmatrix} \begin{bmatrix} L_3 \\ C_3 \\ R_3 \end{bmatrix} \tag{16}$$

<u>3 × 4 downconversion matrix</u> R_{34}

$$\begin{bmatrix} L_3 \\ C_3 \\ R_3 \end{bmatrix} = \begin{bmatrix} 0.9303 & 0.3314 & -0.1479 & 0.0527 \\ -0.1297 & 0.6951 & 0.6951 & -0.1297 \\ 0.0527 & -0.1479 & 0.3314 & 0.9303 \end{bmatrix} \begin{bmatrix} L_4 \\ L_5 \\ R_5 \\ R_4 \end{bmatrix} \tag{17}$$

<u>4 × 5 downconversion matrix</u> R_{45}

$$\begin{bmatrix} L_4 \\ L_5 \\ R_5 \\ R_4 \end{bmatrix} = \begin{bmatrix} 0.9535 & 0.2533 & -0.1340 & 0.0859 & -0.0324 \\ -0.1084 & 0.7870 & 0.5708 & -0.1989 & 0.0590 \\ 0.0590 & -0.1989 & 0.5708 & 0.7870 & -0.1084 \\ -0.0324 & 0.0859 & -0.1349 & 0.2533 & 0.9535 \end{bmatrix} \begin{bmatrix} L_6 \\ L_7 \\ C_5 \\ R_7 \\ R_6 \end{bmatrix} \tag{18}$$

<u>2 × 4 downconversion matrix</u> R_{24}

$$\begin{bmatrix} L_2 \\ R_2 \end{bmatrix} = \begin{bmatrix} 0.7215 & 0.6521 & 0.1728 & -0.1561 \\ -0.1561 & 0.1728 & 0.6521 & 0.7215 \end{bmatrix} \begin{bmatrix} L_4 \\ L_5 \\ R_5 \\ R_4 \end{bmatrix} \qquad (19)$$

<u>2 × 5 downconversion matrix</u> R_{25}

$$\begin{bmatrix} L_2 \\ R_2 \end{bmatrix} = \begin{bmatrix} 0.6325 & 0.6482 & 0.3945 & 0.0287 & -0.1525 \\ -0.1525 & 0.0287 & 0.3945 & 0.6482 & 0.6325 \end{bmatrix} \begin{bmatrix} L_6 \\ L_7 \\ C_5 \\ R_7 \\ R_6 \end{bmatrix} \qquad (20)$$

<u>3 × 5 downconversion matrix</u> R_{35}

$$\begin{bmatrix} L_3 \\ C_3 \\ R_3 \end{bmatrix} = \begin{bmatrix} 0.8407 & 0.5304 & -0.0279 & -0.0891 & 0.0557 \\ -0.1538 & 0.3648 & 0.8285 & 0.3648 & -0.1538 \\ 0.0557 & -0.0891 & -0.0279 & 0.5304 & 0.8407 \end{bmatrix} \begin{bmatrix} L_6 \\ L_7 \\ C_5 \\ R_7 \\ R_6 \end{bmatrix} \qquad (21)$$

<u>monophonic downconversion</u> R_{1n} ($n = 2$ to 5)

$$C_1 = 0.7071\, L_2 + 0.7071\, R_2 \qquad (22a)$$

$$C_1 = 0.5000\, L_3 + 0.7071\, C_3 + 0.5000\, R_3 \qquad (22b)$$

$$C_1 = 0.3998\, L_4 + 0.5832\, L_5 + 0.5832\, R_5 + 0.3998\, R_4 \qquad (22c)$$

$$C_1 = 0.3394\, L_6 + 0.4786\, L_7 + 0.5579\, C_5 + 0.4786\, R_7 + 0.3394\, R_6 \qquad (22d)$$

<u>Selected down/up-conversion matrices</u>

<u>3 to 2 to 3 conversion</u> $R_{32}R_{23}$

$$\begin{bmatrix} L_3' \\ C_3' \\ R_3' \end{bmatrix} = \begin{bmatrix} 0.7500 + 0.0429\,A & 0.3536 + 0.0607\,A & -0.2500 + 0.0429\,A \\ 0.3536 - 0.0607\,A & 0.5000 - 0.0858\,A & 0.3536 - 0.0607\,A \\ -0.2500 + 0.0429\,A & 0.3536 + 0.0607\,A & 0.7500 + 0.0429\,A \end{bmatrix} \begin{bmatrix} L_3 \\ C_3 \\ R_3 \end{bmatrix}$$

<u>4 to 3 to 4 conversion</u> $R_{43}R_{34}$

$$\begin{bmatrix} L_4' \\ L_5' \\ R_5' \\ R_4' \end{bmatrix} = \begin{bmatrix} 0.8851 & 0.2103 & -0.2103 & -0.1149 \\ 0.2103 & 0.6149 & 0.3851 & -0.2103 \\ -0.2103 & 0.3851 & 0.6149 & 0.2103 \\ -0.1149 & -0.2103 & 0.2103 & 0.8851 \end{bmatrix} \begin{bmatrix} L_4 \\ L_5 \\ R_5 \\ R_4 \end{bmatrix} \qquad (24)$$

The above conversion matrices are optimised according to the specific values of decoder parameters ϕ, ϕ', ϕ_3, ϕ_D, ϕ_4, ϕ_5, (a,b,c) given in refs. [4] and [5]. Slightly different values, associated with different speaker layouts, will give marginally different equations above, but in all cases, coefficients will differ only a little from those given here.

in all cases, the coefficients will differ only a little from those given in equs. (10) to (24).

It will be seen that the number of up- and down-conversion matrices to be considered is quite large, and obviously one could not guarantee the cascadability property without using a mathematical design theory such as that associated with figure 10 of ref. [5], as the number of explicit possibilities that need to be investigated is large. The fact that the upconversion matrices are designed, by the theory of ref. [4] to automatically preserve both level-balance and stereo directional localisation quality for all signals greatly reduces the overall degree of checking of compatibility, which can be confined to study of the 10 downconversion matrices (16) to (22). While none of these always give perfect downcompatibility, none give unreasonable results with normal multichannel stereo source signals. Initial studies of nonorthogonal hierarchies (as discussed in ref. [5]) with the same upconversion matrices suggest that significant departures from orthogonality invariably give a worse balance of downward compatibility properties, either as regards level-balance or as regards stereo localisation effect.

The complexity of results with even such a simple hierarchy as the five frontal stage stereo systems suggests that more complex surround-sound hierarchies require great care in design, and the extension of the methods of refs. [4] and [5] to the surround-sound case will be described in future papers.

4. DOWNMIXING VERSUS COMPATIBILITY MATRIXING

We now discuss the use of downmixing in an actual transmission system. First, there is the question of whether a fixed matrixing for up- and down-conversion need be used at all. Particularly with digital transmission systems, there is no difficulty in altering matrixing coefficients so as to optimise the up- and down-conversion results according to the specific programme material being used and differently for different transmission and reception modes.

While at first sight this seems attractive, it results in almost insuperable production problems. First, the recording or mixdown engineer must make choices on what up- and down-matrixing must be used with a specific mixdown, and with the number of possible transmission and reception modes involved, this becomes equivalent to doing a major piece of engineering design for every mixdown. In practice if transmitted mixdown coefficients were to be used, probably a very small number of predefined possibilities known to be reasonable would be used, and the mixdown engineer would have familiarised him or herself with these by previous experience.

Secondly, even if the availability of several up- and down-conversion matrices proved to be desirable in order to optimise specific types of program material, this ignores the fact that in audio production, it will often be necessary to mix material from several different sources. If different sources were optimised for different conversion matrices, one then has to try and find a new set of conversion matrices that works well on the mixed material, i.e. that is reasonable for each component of the mix. In general, neither of the conversion matrix choices made for the individual components will work well on the others, since the conversion matrices will have been optimised for a particular sound. One may even find that, if one of the sound presentations being mixed together has been mixed down specifically for good results via a first choice of conversion matrices, it may sound dreadful via what would otherwise be a good general-purpose choice of conversion matrices.

Thus the need to mix material from different sources suggests that a common fixed set of conversion matrices should be used. That way, providing the original mixdown engineers did their job competently, material from any source can be mixed together without any need to recheck compatibility via all conversion matrices all over again.

Additionally, it is difficult to ensure cascadability in a long production chain using different numbers of channels at different points using an assignable conversion matrix approach, since one also needs to specify precisely how a downconverted signal should be upconverted at a later stage of the production chain for all possible down/up conversions. This means transmitting a whole designed hierarchy of matrices, including down/up conversion matrices, along the entire length of the production chain. It is unreasonable to expect mixdown engineers to be able to design complete hierarchical conversion systems for every mixdown, given the kind of complexity that we have seen is involved above!

Thus, in practice, one is forced to the conclusion that a practical system design must use a fixed cascadable hierarchically designed set of upconversion and downconversion matrices.

Given that a fixed set of upconversion and downconversion matrices must be used, the next question is what approach to transmission of the signals should be used. For example, should transmission involve explicit speaker feed signals, with a different assignation of channels for every speaker layout, or should they involve derived matrixed signals in such a manner that each increase by one in the number of loudspeakers involves simply adding one extra channel to the previous transmission channel signals. The first approach requires the use of downconversion for other speaker layouts, and is termed [1] the "downmixing" approach, and for a fixed reproduction speaker layout requires the use of a different downconversion matrix for every possible transmission mode. The second approach, termed "compatibility matrixing" [1], is by contrast simple in that only a single reproduction matrix is required for all transmission modes, including ones that were not standardised when the reciving equipment was built.

Thus the downmixing approach inherently inhibits future technological development, since a decoder cannot incorporate a downconversion matrix for a future transmission mode that has not yet been standardised, unless one designs the decoder to receive an arbitrary large number of transmission channels and also transmits explicit downconversion matrixing coefficients for all currently standardised reception modes. Even if this is done, technological development will be inhibited by the need to define new coefficients to support newly introduced reproduction modes even with existing broadcasts.

By contrast, the compatibility matrixing approach need only transmit a flag indicating the nature of each transmission channel, with the matrixing in the receiver simply feeding those channels into a fixed decoding matrix in a standardised way.

One of the few advantages of the downmixing approach is that it can be designed to ensure maximal masking of error artefacts in audio data compression systems. Such an approach is a poor substitute to designing audio data compression systems that minimise such error audibility in the first place; this has already been discussed in ref. [11], where it was shown that it is possible to design such systems without the risk of "stereo unmasking" of error artefacts. We shall discuss the design of data compression systems to minimise such stereo unmasking effects below.

The compatibility matrixing approach for frontal stage stereo systems simply transmits the first n of signals T_1 = M, T_2 = S, T_3 = T, T_4, T_5 for n-speaker stereo source material, and uses a fixed decoding matrix operating from the first n_2 transmission channels for reproduction using n_2 speakers. The actual matrices for encoding into transmission signals and decoding from transmission signals were detailed in refs. [3] and [5], and the construction of these encoding and decoding matrices to ensure that the upconversion and downconversion automatically satisfy the requirements of this paper were detailed in refs. [4] and [5].

Apart from switching off unused transmission channels, receivers need contain only one fixed matrix for its speaker layout, whatever source material is transmitted. One can optionally add frequency-dependence of the form described in equs. (10), (13) or (14) if it is known that only two channels are being received, in order to optimise the reproduced subjective results. This should never be necessary if three or more channels are being received, since the frequency-dependence for 2-channel source materials can be incorporated into the T-channel, for example as in figs 4 or 6 of [3], by the broadcaster as a part of the mixdown production process.

While the use of compatibility matrixing is simple and straightforward if the system hierarchy is designed to ensure compatibility and cascadability, all matrixing can exaggerate the subjective effect of errors in the transmission channels. While [1,2] this has become of concern with audio data compression systems, the problem is an old one. For example, the B-format representation of Soundfield signals [8], [9] was chosen rather than a speaker-feed type format for Ambisonics because it was found that the subjective effect of noise reduction system mistracking was about four times more disturbing, in terms of perceived image movement effects, in speaker feed than in B format representation.

There are two kinds of errors introduced by audio data compression systems [11]. One is noise-like errors, not cross-correlated with the wanted signals, and the other is gain-modulation type errors, which alter the stereo positioning of signal components moment by moment. The two kinds of errors cause different kinds of problems.

Enhanced stereo reproduction systems, whether of the frontal stage or surround-sound kind, are capable of better subjective quality due to providing the ears with more realistic information, and so are more revealing of error artefacts. A rule of thumb is that gain-modulation artefacts become about three to six times more audible than for two speaker stereo (depending on factors to be discussed below), and noise-like errors may have to be reduced by 10 to 15 dB if they are reproduced from a different direction to the wanted signal.

Some data compression systems, particularly those being standardised by ISO, involve suppression of some allegedly masked signal components, which can cause severe image movement effects if the signal components of some transmission channels are being gated while others are preserved. Systems like Dolby AC-2 [12] or AptX 100 [13] that do not gate out any audio band will tend to suffer from this effect less, although the author has shown [11] that even non-gated bands quantised at low bit rates still have significant amplitude modulation that may cause image movement effects of several degrees.

The use of matrixing that minimises side-to-side movement of images helps to minimise the effect of gain modulation effects, and this suggests that matrixing that avoids image movements for central sounds may be desirable - i.e. the use of sum-and-difference type modes in which the interchange of left signals with their right mirror image counterparts causes transmission signals either to be unchanged (signals of "sum" type) or to be polarity inverted (signals of "difference" type). The transmission systems suggested in refs. [3] and [5], using transmission signals M (mono), S (stereo difference), T, T_4 and T_5 are of this desirable form, and so will tend to subjectively minimise image movement effects due to amplitude modulation.

In particular, the orthogonal M,S,T,T_4,T_5 transmission hierarchy of ref. [5] conveys the predominant localisation information via the first two transmission signals M and S, and the remaining signals T, T_4 and T_5 have progressively less effect on localisation, so that amplitude gain errors on these signals have a relatively marginal effect on localisation.

Thus the combination of these two effects makes the transmission hierarchy proposed in ref. [5] particularly tolerant of coding system amplitude gain errors insofar as image movement effects tend to be minimised provided that sufficient bits be allocated to the M and S signals to give good results via 2-speaker stereo.

While specific compatibility matrixing proposals help to minimise image movement effects due to amplitude modulation errors in audio data compression codecs, the effect of noiselike codec errors needs to be examined separately, since they in effect act as independent sound sources which will be "directionally unmasked" if they come from different directions to the wanted signals, or if they emerge from speakers that are otherwise substantially unactivated by wanted signals.

5. DIRECTIONAL UNMASKING

The only completely satisfactory solution to the problem of noiselike codec errors being unmasked due to being in different directions to wanted signals is to design data compression codecs to be such that the direction of noiselike coding errors is aligned with the direction of wanted signals. That this is possible was noted in ref. [11], but the methods suggested there are incompatible with currently proposed data compression methods. Another possible approach will be described in ref. [14].

However, in the case that noiselike codec errors are not directionally aligned with the wanted signal, it is still possible to reduce the effect of directional unmasking by a careful choice of transmission matrices.

Some transmission matrices will cause a marked degradation of reproduced signal-to-noise ratio, by reproducing error signal directions with a larger total energy gain than the wanted signal. This effect will not occur if the transmission decoding matrix is orthogonal or energy preserving, since (by definition) such matrices preserve the reproduced energy of all signals, and hence preserve signal-to-noise ratio. Nonorthogonal matrices can degrade signal-to-noise ratio for some signals. The preferred transmission systems in refs. [3] and [5] are based on orthogonal matrices, whereas those discussed in refs. [1,2,7] are not. As we shall see, this can cause a marked worsening of reproduced error energy, as well as increased directional unmasking.

The noiselike errors in data compression codecs are very different from the constant-level backgound noise occuring in traditional transmission systems, since the noise itself is signal-dependent, so that reproduced signal-to-noise ratio calculations need to be done in a signal-dependent way.

To facilitate such calculations, we can adopt a simplified model for the noiselike errors produced in data compression codecs. For the purposes of calculation, we shall model the noiselike errors produced in each decoder subband as having an energy in each codec transmission channel that is a fixed multiple k of the signal energy passing through that codec channel in that subband, and further assume that the error signals in different codec channels are mutually uncorrelated. The "constant" k may depend on the bit allocations for that subband, and the signal energies in other subbands, but we shall model it as being identical in different channels for the same subband - a model that may not give results too far from a realistic situation.

Consider two 3-channel stereo transmission systems, one an orthogonal-matrix transmission system proposed in ref. [5] and the other a system proposed in refs. [1], [2] and [7]. The first system encodes L_3, C_3 and R_3 into 3 transmission signals M, S, T given by

$$\begin{pmatrix} M \\ S \\ T \end{pmatrix} = \begin{pmatrix} 0.5000 & 0.7071 & 0.5000 \\ 0.7071 & 0.0000 & -0.7071 \\ 0.5000 & -0.7071 & 0.0000 \end{pmatrix} \begin{pmatrix} L_3 \\ C_3 \\ R_3 \end{pmatrix} \qquad (25a)$$

and decodes L_3, C_3, R_3 from the transmission signals M,S,T by

$$\begin{pmatrix} L_3 \\ C_3 \\ R_3 \end{pmatrix} = \begin{pmatrix} 0.5000 & 0.7071 & 0.5000 \\ 0.7071 & 0.0000 & -0.7071 \\ 0.5000 & -0.7071 & 0.5000 \end{pmatrix} \begin{pmatrix} M \\ S \\ T \end{pmatrix} \qquad (25b)$$

both of which matrices are orthogonal. The second system encodes L_3, C_3 and R_3 into 3 transmission signals L,R,T given by

$$\begin{pmatrix} L \\ R \\ T \end{pmatrix} = \begin{pmatrix} 1.0000 & 0.7071 & 0.0000 \\ 0.0000 & 0.7071 & 1.0000 \\ 0.0000 & 0.7071 & 0.0000 \end{pmatrix} \begin{pmatrix} L_3 \\ C_3 \\ R_3 \end{pmatrix} \qquad (26a)$$

and decodes L_3, C_3 and R_3 from the transmission signals L,R,T by

$$\begin{pmatrix} L_3 \\ C_3 \\ R_3 \end{pmatrix} = \begin{pmatrix} 1.0000 & 0.0000 & -1.0000 \\ 0.0000 & 0.0000 & 1.4142 \\ 0.0000 & 1.0000 & -1.0000 \end{pmatrix} \begin{pmatrix} L \\ R \\ T \end{pmatrix}. \qquad (26b)$$

Consider a signal with unit energy assigned just to the C_3 loudspeaker with no sound assigned to the other 2 speakers - such as will be the case for central on-screen dialogue for HDTV productions. Then for the matrix encoding of equ. (25), the respective noiselike codec error energy in the M, S and T signals are

$$\tfrac{1}{2}k,\ 0 \text{ and } \tfrac{1}{2}k \qquad (25c)$$

which are, by assumption, uncorrelated, so that after decoding by the matrix of equ. (25b), the noiselike error energies in the L_3, C_3 and R_3 channels are respectively

$$\tfrac{1}{4}k,\ \tfrac{1}{2}k \text{ and } \tfrac{1}{4}k\ . \qquad (25d)$$

For the encoding matrix of equ. (26a), however, the noiselike codec error energy in the L,R and T transmission channels are respectively

$$\tfrac{1}{2}k,\ \tfrac{1}{2}k \text{ and } \tfrac{1}{2}k \qquad (26c)$$

which by assumption are uncorrelated, so that after decoding by the matrix of equ. (26b), the noiselike error energies in the L_3, C_3 and R_3 channels are respectively

$$k,\ k \text{ and } k\ , \qquad (26d)$$

Comparing the codec error energies (25d) for the system of equs. (25) with the codec error energies (26d) for the ststem of equs. (26), we see two things:

(i) The matrixing of equs. (26) used with central signals causes a threefold increase in the total reproduced codec error energy as compared with the orthogonal matrixing of equs. (25) in each subband, and

(ii) the total of the directionally unmasked reproduced codec error energy, i.e. energy not from the C_3 speaker containing the wanted sound is increased fourfold by the system of equs. (26) as compared to the matrixing system of equs. (25).

Thus the phenomenon reported by Theile [2] of increased audibility of codec errors with compatibility matrixing is not solely due to directional unmasking, but also due to the matrixing used (which incorporates the matrix of equs. (26) for sounds encoded into the frontal stage) increasing the total reproduced codec error energy. We have seen that use of an orthogonal matrix system reduces total reproduced codec error energy for the most important sound position, centre front, by 4.8 dB, and the directionally unmasked component of this error energy by 6 dB.

The method of analysing the performance of matrix transmission systems exemplified by the above calculations can be applied to general image positions for arbitrary encoding and decoding matrices, and the system of equs. (26) performs much better for hard left or hard right signals, but degrades as sounds move towards the centre.

A general theoretical result is that orthogonal matrix systems always give the minimum reproduced codec error energy, giving an energy signal to noise ratio of $1/k$, whereas nonorthogonal matrix systems may equal this performance for some sound positions, but may be considerably worse for others.

One conclusion is that the matrixing used with data compression systems should use nearly orthogonal matrices where possible, and in particular should minimise directional unmasking for the most important sound positions, i.e. near front centre. The compatibility matrixing considered in refs. [1], [2] and [7] performs particularly badly as we have seen.

6. CONCLUSIONS

In this paper, it has been shown that particular choices of upconversion and downconversion matrices between different frontal stage stereo reproduction modes that may appear "plausible" when considering isolated speaker feed signals cease to perform well when fed with panned multispeaker stereo signals of more complicated form, and in particular that an upconversion followed by a downconversion back to the original speaker format may give grossly altered reproduced results.

The desirability of all upconversions and downconversions being cascadable down a long broadcast production chain is considered, so that program material from different sources can be freely mixed without any more degradation than would be expected from the minimum number of channels in any intermediate stage of the chain. Current proposals for up- and down-conversion, such as those considered by Meares [1,7] and Theile [2], are shown not to be cascadable, which would lead to serious production problems as described in ref. [3].

This problem of noncascadable up- and down-conversions exists both in the "downmixing" and "compatibility matrixing" approaches, and it was shown that for cascadability of up- and down-conversion matrices, these matrices must:

(i) have some negative matrix coefficients, and

(ii) not feed extreme left and right speaker feeds to just the extreme speakers of the final reproduction layout, but must also crossblend to some of the inner speakers.

In both respects, cascadability violates implicit design constraints arbitrarily imposed in the systems of [1,2,7].

It was shown in ref. [4] that optimum preservation of the intended stereo effect for arbitrary panpot laws necessitates use of upconversion matrices that satisfy these requirement (i) and (ii) in any case, and in ref. [5], it was shown that a completely cascadable set of up- and down-conversion matrices could be designed systematically based on such psychoacoustically optimised upconversion matrices. In this paper, the explicit form of these up- and down-conversion matrices for one to 5 stereo loudspeakers has been given in equs. (10) to (24).

These up- and down-conversion matrices arise from an essentially unique orthogonal-matrix compatibility matrix method of transmission detailed in ref. [5], but can also be used with a downmixing transmission approach that requires cascadability.

The use of variable matrix coefficients in transmission was considered, but was shown to lead to considerable operational problems due to

(i) the need to be able to mix program material from several sources,

(ii) problems of cascadability, and

(iii) the sheer number of different up- and down-conversion matrices that have to be "designed" into the matrix coefficients transmitted.

Downmixing was also considered, and similarly found to be operationally inflexible when cascadability of different reproduction modes was considered, requiring elaborate switching of receivers for every possible transmitted reproduction mode. In particular, this would inhibit future upgrading to more elaborate reproduction modes (perhaps using more speakers, perhaps adding height or full 360^{o} directionality) since such modes would have in some way to be built into existing receivers in order to perform the requisite downmixing - so that receivers would have to cope with the maximum number of transmitted channels (say eight) that might conceivably be needed for future enhanced sound systems.

It was shown that the one operational problem that downmixing was designed to tackle - the audibility of data compression codec errors after matrixing, was in part due to choice of matrices that increase these errors, and that an orthogonal matrix transmission system, such as proposed in ref. [5] can reduce directionally unmasked errors for the most important sound directions by 6 dB as compared to other compatibility matrix systems being proposed.

It is not claimed that the use of orthogonal matrix coding and decoding on its own eliminates the unmasking of codec error signals, but that it does reduce these problems to a significant degree. Compatibility matrixing does become practical if used in conjunction with modified codecs that align the direction of codec errors with that of the wanted signal, as discussed in general terms in ref. [11], and in some detail in ref. [14]. It is thus suggested that downmixing approaches are unnecessary for use with audio data compression systems, and indeed in ref. [14] it is noted that downmixing still fails to mask directional codec errors for some surround-sound reproduction modes, so that a carefully-designed compatibility matrixing method may prove to be superior even in this regard.

Finally, it was noted that the up- and downconversion matrices arising from the work of refs. [3] to [5] and given as explicit equations (10) to (24) in this paper do not preserve the width of the stereo sound stage as a proportion of the total subtended angular width of the speaker layout, due to crosstalk of extreme speaker feeds in the up- and down-conversion. The alterations of width were shown to be relatively small in ref. [4], but this points to a danger of making inappropriate comparisons of sound stages before and after upconversion without making allowance for this by use of layouts with slightly different subtended angles.

In summary, it has been shown that prior-art proposals for matrixing and up- and down-conversion for multispeaker stereo systems have failed to take account of various operational requirements such as cascadability, use with a variety of multispeaker stereo panpot laws [10], localisation psychoacoustics of upconversion, minimisation of reproduced data compression codec error energy, operational flexibility and simplicity down long broadcast production chains, and potential for future upgradability to more elaborate reproduction modes.

The proposals of refs. [3] to [5], whose up- and down-conversion matrices are given explicitly in this paper, are proposed as the basis of a system design that takes account of these problems.

REFERENCES

[1] D. Meares, "High-Quality Sound for High Definition Television," Proc. AES 10th International Conf., London, pp. 163-177 (1991 Sept.)

[2] G. Theile, "HDTV Sound Systems: How Many Channels?", Proc. AES 10th International Conf., London, pp. 147-162 (1991 Sept.)

[3] M. Gerzon, "Production Approaches to Multispeaker Stereo," Studio Sound, to be published, probably 1991 Aptil/May

[4] M.A. Gerzon, "Optimal Reproduction Matrices for Multispeaker Stereo," Preprint 3180, 91st Audio Engineering Society Convention, New York, (1991 Oct.)

[5] M.A. Gerzon, "Hierarchical Transmission System for Multispeaker Stereo," Preprint 3199, 91st Audio Engineering Society Convention, New York, (1991 Oct.)

[6] M. Gerzon, "Three Channels - The Future of Stereo?" Studio Sound, vol. 32 no. 6 pp. 112-125 (1990 June)

[7] D.J.Meares, "High Definition Sound for High Definition Television," Proc. AES 9th International Conf. "Television Sound Today and Tomorrow," Detroit Michigan (1991 Feb.) pp. 217-232

[8] M.A. Gerzon, "Ambisonics in Multichannel Broadcasting and Video," J. Audio Eng. Soc., vol. 33 no. 11, pp. 859-871 (1985 Nov.)

[9] M.A. Gerzon, "Ambisonics Part Two: Studio Techniques", Studio Sound, vo. 17 no. 8, pp. 24, 26, 28-30 (1975 Aug.)

[10] M.A. Gerzon, "Panpot Laws for Multispeaker Stereo", Preprint of the 92nd Audio Engineering Society Convention, Vienna (1992 March 24-27)

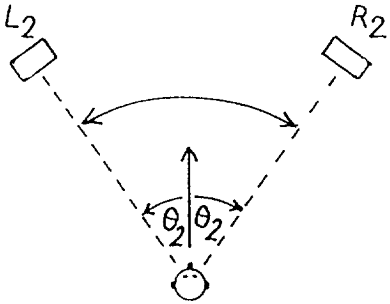
[11] M.A. Gerzon, "Problems of Error Masking in Audio Data Compression Systems," Preprint 3013 , 90th Audio Engineering Society Convention, Paris (1991 Feb.)

[12] L.D. Fielder & G.A. Davidson "AC-2: A Family of Low Complexity Transform-Based Music Coders," Proc. AES 10th International Conf., London (1991 Sept.) pp. 57-70

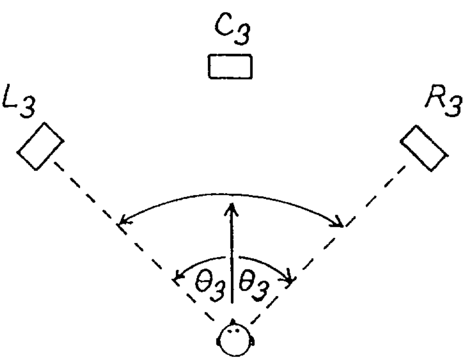
[13] M. Smyth & S. Smyth, "APT-X100: A Low-Delay, Low Bit-Rate Sub-Band ADPCM Audio Coder for Broadcasting," Proc. AES 10th International Conf., London (1991 Sept.) pp. 41-56

[14] M.A. Gerzon, "Directional Masking Coders for Multichannel Subband Audio Data Compression Systems," Preprint of the 92nd Audio Engineering Society Convention, Vienna (1992 March 24-7)

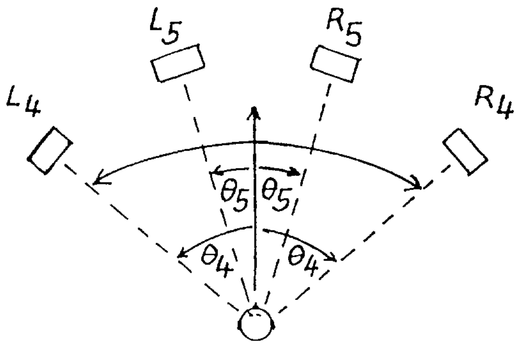
1a



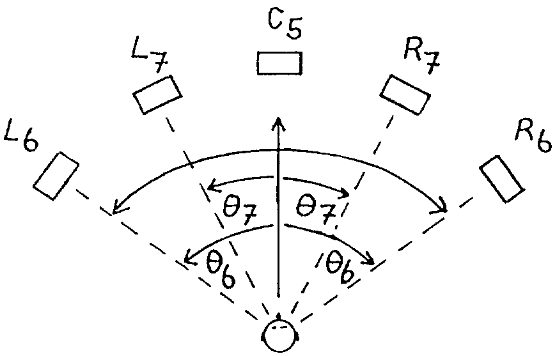
1b



1c



1d



1e

Figure 1. Loudspeaker layouts for front-stage stereo using in figs. 1a to 1e respectively from one to five loudspeakers indicating angles and speaker symbols.

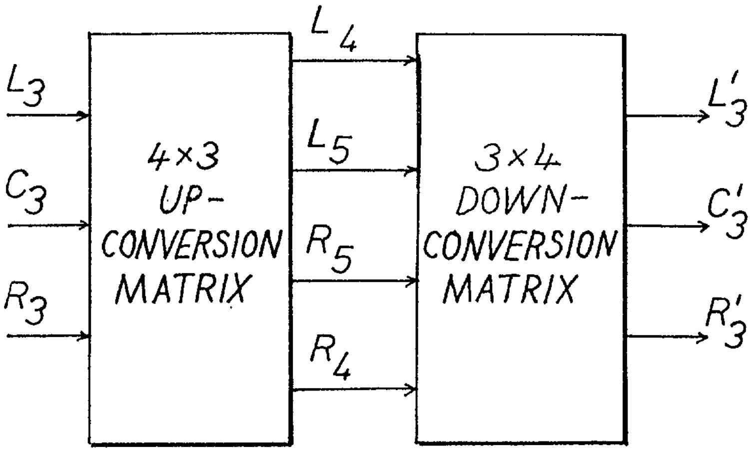


Figure 2. Effect of upconversion from 3 to 4 channels followed by downconversion from 4 to 3 channels.

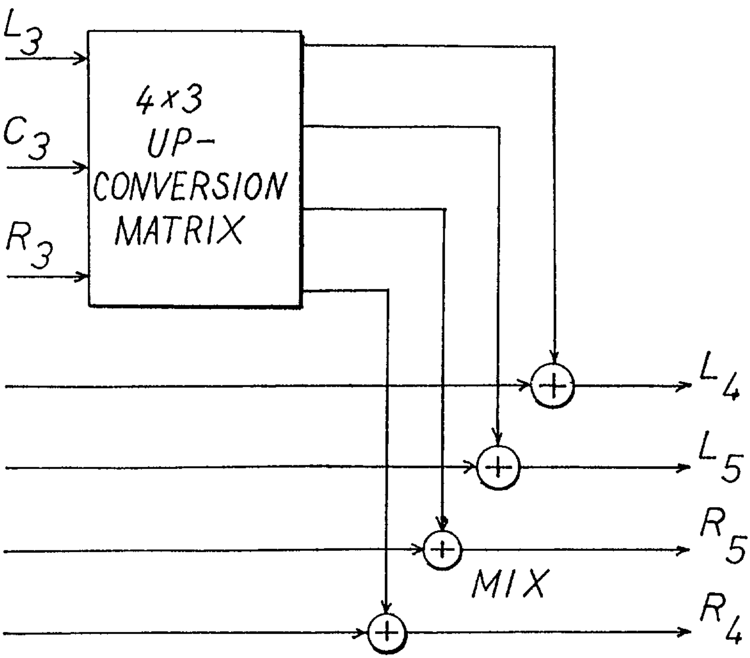


Figure 3. Mixing upconverted 3-channel material with 4-channel material.