EBU

OPERATING EUROVISION AND EURORADIO

TR 049

CONVERSION TECHNIQUES FOR MULTICHANNEL AUDIO FORMATS

ORIGINAL PUBLICATION: MARCH 2002

(REPLACES BPN 042)

Geneva
June 2019

Status of this document, June 2019

This Technical Report contains the content of BPN 042, an EBU members-only document originally published in March 2002. Nothing, other than typos and corrections to English usage, has been changed from BPN 042. As such the text is of historical relevance, certain statements in it must be interpreted as being pertinent in March 2002 and it does not represent current-day practices and technology. Interest has been shown in the subject matter for reference purposes by non-members and the EBU Technical Committee deemed republication as a TR acceptable. The document will not be developed further; Next Generation Audio[1] (NGA) systems have now substantially replaced the technology dealt with herein.

Original Summary

Multichannel audio is not new but recently has come to prominence with the public, following the introduction of several new broadcasting systems (DAB, DVB) and recording media (DVD) that are able to deliver multichannel audio to the home.

These delivery systems use the universal multichannel audio format described in the ITU-R Recommendation BS.775 [1] which have been recommended by EBU for broadcasting [2]. This format provides three front channels and two surround channels and is referred to as "3/2" format. If an optional low frequency channel is used for effects, the format is usually dubbed "5.1". For convenience we will generally refer to this format as the "5.1 format" even if the LFE channel is not used.

However, as shown in a recent EBU Report [3], several other multichannel audio formats have been, and some still are, in use. Existing material made in these formats may need to be converted into the target formats (3/2 or 5.1 respectively). This will be especially true during the early period of introduction of multichannel audio services enough programme material may not available in the 5.1 format.

This document is a short guide how to convert multichannel programme material from an existing format to the new standards used for broadcasting. Conversion may involve changing the signal format, the coding format, or to aligning other characteristics which may be different to those of the target 5.1 format.

The document also give advice on how to deal with compatibility between multichannel audio material and existing audio formats, such as two-channel stereophony or Dolby Surround[2], that are expected to remain in uses for some years.

A bibliography lists important sources and relevant international standards.

A front-page note was added to BPN 042 in 2012, as follows:

“Whilst much of the information in this report is still valid, some details, such as the use of 640 kbit/s for emission codecs, is not representative of later practices.”

[1] See, for example, Recommendations ITU-R BS.2076 and ITU-R BS.2088, EBU Tech 3388, Tech 3392 and TR 045.

[2] "Dolby Stereo" was the original name of the Dolby surround sound system designed for use in the cinema. This system itself is still used in many cinematic productions but is now referred to as just " Dolby".

The "Dolby Surround" system was developed from "Dolby Stereo" for use in consumer formats.

Contents

Conversion Techniques for Multichannel Audio Formats (March 2002)

EBU Committee	First Issued	Revised	Re-issued
TC	2002 (as BPN)		2019 (as TR)

Keywords: Mono, Stereo, Multichannel, Quadraphony, upmix, downmix, conversion.

1. Introduction

1.1 General

For many years it has been widely recognised that conventional stereophony, using two loudspeaker channels, has serious limitations and that additional channels are desirable to provide improved stereophonic images.

The universal 3/2-stereo format, according to ITU-R Recommendation [1], supplements the conventional left and right stereo channels with an additional centre front channel and two surround channels. This format is generally referred to as the 3/2 format. This offers enhanced quality to the stereophonic presentation, not only for applications with accompanying pictures but also for audio-only programmes. This format has been recommended by the EBU for new broadcasting services, for instance DAB and DVB (see also EBU Recommendation R96 [2]). An optional addition to this format is a low frequency front channel, which can be used to provide special effects and enhancements. This variation is generally known as the 5.1 format.

However, as shown in the recent EBU Report BPN 021 [3] several other different multichannel audio formats have been and still are, available on the market. Programme material produced in these formats sometimes needs to be converted to the recommended target format (3/2 or 5.1 respectively). This will particularly be true during the early period of introduction of multichannel audio services, especially if not enough programme material is available in the recommended formats.

This document is a short guide on how to convert multichannel programme material of different available formats to fit in with new broadcasting possibilities such as, for instance, DAB and DVB.

Additionally, the document deals with the compatibility problems between material produced in the 3/2 or 5.1 formats and the existing audio formats used by broadcasters, such as two-channel stereophony or Dolby Surround.

1.2 Definitions

Format conversion

Conversion of programme material from a source multichannel format with different channel allocations to the recommended 3/2 (5.1) format. This may involve up- or down-conversion depending the number of audio channels in the source programme.

Up-conversion

Depending on the context, up-conversion can mean one of three processes:

1. A linear process to create signals for the multichannel format without adding any changes to the sound image, for instance when a stereo-, quadraphonic or matrix surround production is going to be broadcast through a 3/2 multichannel system.

2. The generation of entirely artificial multichannel signals from a system with fewer channels (for instance a conventional 2-channel stereophonic pair) which was not intended to contain multichannel information.

3. A complete remix of the programme material from an original multitrack recording.

Some organisations have many 4 channel “quadraphonic” recordings available where up-conversion to 5.0 may be worthwhile.

Trials have also been made with up-conversion from 2-channel matrixed surround (e.g. Dolby surround) to the 3/2 format. This kind of up-conversion is not recommended because there may be an automatic down-conversion process at the receiving end of the transmitting chain, which can seriously distort the intended sound image.

Down-conversion

Down-conversion is needed when a multichannel audio production is to be broadcast through a mono, stereo or matrix surround delivery system.

Most existing production and delivery systems use 2-channel stereophonic distribution, recording, routing and monitoring systems. There is a need for some method of conversion from 5.0 or 5.1 to a compatible format. Ideally, such a conversion should allow existing equipment to process the converted signals transparently, as though they were 2-channel stereophonic signals. This includes monitoring the signals at least for content and basic quality. Inevitably, there will be a more than one such 2-channel signal.

An important aspect is downward compatibility of 3/2 music recordings with 2-channel stereo systems. The down-mix equations and coefficients according to BS.775 [1], can be used but it is not possible to guarantee that the resulting two-channel down-mix will automatically be as aesthetically satisfying as the original 3/2-stereo version. Ideally, the down-mix should sound the same as an appropriate conventional two-channel recording originally mixed in 2/0-stereo from the same set of microphone signals. It has been found that in practice the automatic down-mix will not always provide optimum quality regarding a number of parameters such as reverberance balance, loudness balance, perception of depth, etc.

Codec conversion

Change of a given digitally encoded source bit-stream to another target codec system. It includes conversion of encoded bit-streams between different multichannel coding- or transmission systems such as MPEG-2 BC (backwards compatible to MPEG-1 [4]), MPEG-2 AAC, or Dolby Digital (AC-3) and of course this also includes conversion of metadata. Codec conversion can normally be done only by decoding the source to a non-coded format and re-encoding into the target format. Sometimes a specific transcoder is available for a specific conversion.

1.3 Characteristics of conversion processes

To convert material from an existing format different to the 3/2 format, it may be necessary to convert some or all of the following characteristics:

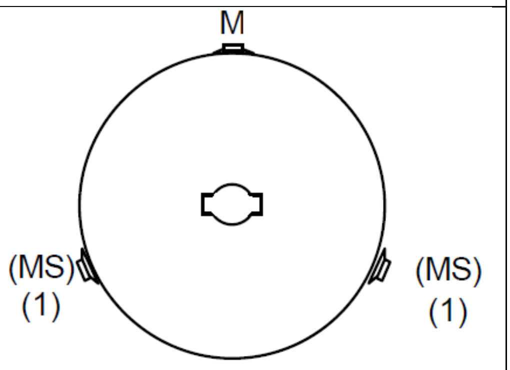
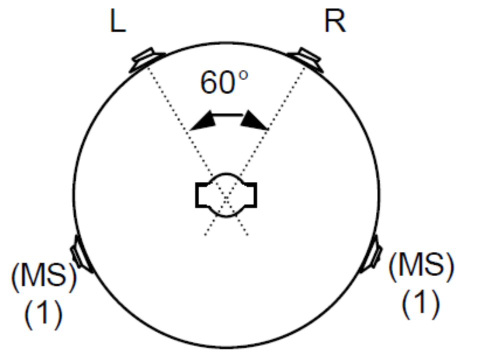
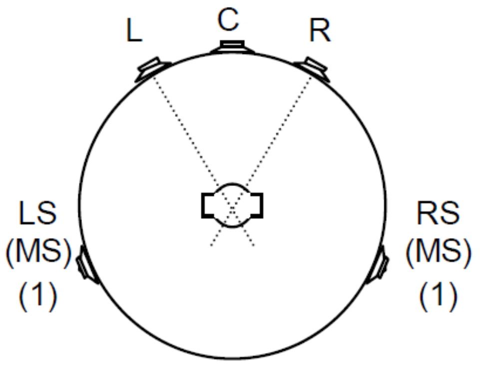
- **Channel assignment** (if the existing material has a different channel or track assignment to the target format)

- **Level alignment** (if the existing material has a different level alignment for one or more signals to the target format)
- **Format conversion** (if the existing material has a different channel format – for example 2/0, 2/1, 2/2, 3/1 - to the 3/2 target format, or if, for example, a recording device is not able to store all the signals on its available recording tracks)
- **Codec conversion** (if the existing material uses a different coding or compression scheme to the target format)

1.4 Compatible multichannel input/output formats for radio and television

Recommendation ITU-R BS.775, which defines the universal multichannel format 3/2, also defines a grouping of hierarchical, downwards compatible multichannel formats and corresponding loudspeaker arrangements. These are designed to be used for radio and television, see Table 1.

Table 1: Hierarchy of compatible multichannel audio formats acc. to ITU-R BS.775-1 [1]

<table>
<tr><th>System</th><th>Channels</th><th>Code</th><th>Loudspeaker Arrangement</th></tr>
<tr><td>Mono channel</td><td>M</td><td>1/0</td><td></td></tr>
<tr><td>Mono plus mono surround</td><td>M/MS</td><td>1/1</td><td>
</td></tr>
<tr><td>Two channel stereo</td><td>L/R</td><td>2/0</td><td rowspan="3">
</td></tr>
<tr><td>Two channels plus 1 surround</td><td>L/R/MS</td><td>2/1</td></tr>
<tr><td>Two channels plus 2 surrounds</td><td>L/R/LS/RS</td><td>2/2</td></tr>
<tr><td>Three channel stereo</td><td>L/C/R</td><td>3/0</td><td rowspan="3">
</td></tr>
<tr><td>Three channels plus 1 surround</td><td>L/C/R/MS</td><td>3/1</td></tr>
<tr><td>Three channels plus 2 surrounds</td><td>L/C/R/LS/RS</td><td>3/2</td></tr>
</table>

(1) In the case of mono surround, MS, the signals feeding LS and RS should preferably be de-correlated.

The following audio formats are commonly in use in Radio and TV:

- 1/0 (monophonic)
- 2/0 (two-channel stereophonic)
- 3/1 (multichannel format, transmitted as a matrixed 2-channel signal, for instance, in Dolby Surround)
- 3/2 (discrete universal multichannel format)

Other formats may be encountered on other sources of programme material, such as cinema films or DVD. (see EBU report BPN 021 [3]).

1.5 Down-compatibility requirements

Reception and reproduction of TV and radio programmes in multichannel audio format will be possible for a minority of the customers only, whilst the others will need a high quality two-channel or monophonic signal. This will be true not only during the introduction period but also, in general, in the long term.

Therefore, every production, recording and transmission of a multichannel audio programme, with or without accompanying pictures, should try to provide the optimum downward compatibility with the existing 2/0 and 1/0 audio formats (possibly also with matrixed Dolby Surround signals), see also EBU Recommendation R96 [2].

Special consideration must be given to the effects of any automatic down-mix via the matrix procedure in the encoder (e.g. MPEG-2 BC) or via the automatic down-mix in the consumer receiving equipment (e.g. Dolby Digital). Therefore, it is very important to listen to these down-mixed versions during production or postproduction.

It is also essential to point out that all automatic down-mix procedures that are currently known do not include the LFE channel. For more details see sections 2.4 and 2.5.

2. Format conversion

2.1 Channel assignment

A list of commonly used designations for the signals or channels in a multichannel environment is given in Table 2.

Table 2: Designations and Abbreviations for Audio Channels/Signals in a Multichannel Environment (conforms to ITU-R BR.1384 [5], EBU R91 [6] and SMPTE RP200 [7])

Audio Channel/Signal	Abbreviation
Left Front	L
Centre Front	C
Right Front	R
Left Surround	LS
Right Surround	RS
Low Frequency Effects/Enhancement	LFE
Mono Surround	MS
Mono Surround at a –3 dB signal level	MS (-3 dB)
Left Total	Lt

Right Total	Rt
Stereo Left (compatible version)	L_0
Stereo Right (compatible version)	R_0
Monophonic	M
Freely useable	F
Unassigned / Unused	U

Table 3 shows the track allocations for recording multichannel stereophonic sound programme material that have been recommended by the ITU-R [5], the EBU [6] and SMPTE [8]. It is also useful to use these allocations for other interfaces for multichannel systems such as the output channels of the mixing console or file formats.

Table 3: Track assignments for multichannel audio media [ITU-R, EBU, SMPTE, etc.]

<table>
<tr><th>Track</th><th>ITU-R, EBU</th><th colspan="3">SMPTE Standard Assignment A</th><th>SMPTE Standard Assignment B</th></tr>
<tr><td>1</td><td>L</td><td colspan="3">L</td><td>L</td></tr>
<tr><td>2</td><td>R</td><td colspan="3">R</td><td>R</td></tr>
<tr><td>3</td><td>C</td><td colspan="3">C</td><td>C</td></tr>
<tr><td>4</td><td>LFE</td><td colspan="3">LFE</td><td>LFE</td></tr>
<tr><td>5</td><td>LS or MS (-3 dB)</td><td colspan="3">LS[1)]</td><td>LS[1)]</td></tr>
<tr><td>6</td><td>RS or MS (-3 dB)</td><td colspan="3">RS[1)]</td><td>RS[1)]</td></tr>
<tr><td>7</td><td>A[3)]</td><td>Lt</td><td rowspan="2">or</td><td>Lo</td><td>F[2)]</td></tr>
<tr><td>8</td><td>B[3)]</td><td>Rt</td><td>Ro</td><td>F[2)]</td></tr>
</table>

Notes on Table 3:

1) *In the case of programmes with a monophonic surround channel, the MS (-3 dB) monophonic surround signal can be placed on both tracks 5 and 6. This allows a programme with a single surround channel to be treated as a programme with two surround channels. The MS (-3 dB) signal will be reproduced on both the LS and RS loudspeakers, with a relative level of –3 dB with respect to the front channels. The combined power into the room will be the correct relative level of 0 dB (SMPTE).*

2) *In assignment B, tracks 7 and 8 may be used freely for any purpose. (SMPTE)*

3) *If relevant; otherwise these tracks can be used freely. (ITU-R, EBU)*

According to the ITU-R and EBU, tracks 7 and 8 may be used for the left/right channels (A/B) of a 2/0 format (two-channel stereo) signal for comparisons and/or for the mandatory supplement of the audio part of the DVD-Video; or for the additional signals LC/RC of a 5/2/1 (7.1) format.

A further note in SMPTE 320M states:

"If the main audio program requires less than six tracks, the use of any of the unneeded tracks for other purposes is not compliant with this Standard. While nonconforming assignments can be expected to occur in practice, such usage is outside the scope and specification of the Standard. It is recommended that track assignments which are not according to this standard remain as consistent as possible with Table 3, so as to minimise the need to re-patch studio configurations. For example, if a 2/2 program (L, R, LS, RS) and a 2/0 program (Lt, Rt) are all placed in the first 6 tracks, the 2/2 program should occupy tracks 1, 2, 5, and 6 respectively, and the 2/0 program should occupy tracks 3 and 4."

It should be noted that some of the existing multichannel audio formats used with cinema films do not meet these SMPTE requirements

2.2 Audio levels in a multichannel environment

2.2.1 Digital coding levels

The digital coding level should make the most efficient use of the available dynamic range. The use of the dynamic range is controlled by two factors:

1. The way in which the audio levels are measured (the metering method).
2. The difference between the full digital modulation level, (0 dBFS), and the modulation that occurs when the maximum operating level is reached, using the metering method defined under point 1. This difference is usually called “headroom”.

If the signals are controlled with a peak programme meter, PPM, in accordance with IEC 268 [9], a headroom of 9 dB provides the best use of the dynamic range. This is also recommended in EBU R68 [10]. It should be noted, that other bodies have defined different amounts of headroom, for instance SMPTE RP200 [7] has 11 dB.

2.2.2 Level alignment between channels

Unfortunately, there is a difference between relative levels of the channels recommended in ITU-R BS.775 [1], ITU-R BS.1116 [11], EBU R22 [12] on the one hand and SMPTE RP200 [7] on the other.

The ITU and EBU use the same relative levels for every main loudspeaker/channel (L, C, R, LS, RS), considering the loudspeaker arrangement as a universal 5-channel set-up

The SMPTE recommends that the level from the LS and RS loudspeakers should be reduced by 3 dB in the case of independent surround signals.

The SMPTE recommendation may be intended for cinemas aligned for the older Dolby Stereo 3/1 system to make the overall sound pressure level of all the surround loudspeakers equal to the level from one of the front channels. This does not, however, seem to be appropriate for a universal discrete 5-channel system as it results in a 3 dB lower sound level for the LS and the RS loudspeaker channels. Such a difference will cause not only a minor loss of the overall sound pressure level in a 3/2 loudspeaker arrangement (about 1.5 to 2 dB), but also a more serious distortion of the front-to-back balance of the sound image, if recordings of different sources are reproduced by the same reproduction arrangement.

Therefore, it would be better to incorporate this level difference into the matrix decoder itself, or to take it into consideration for the recording of the 3/1 signals in a decoded format on a certain recording medium. This latter way is specified both in EBU R91 [6] and Recommendation ITU-R BR.1384 [4].

2.2.3 Reproduction sound pressure level

Recommendation ITU-R BS.1116 [11] and EBU R22 [12] both specify a reference listening level for reference test conditions. Using broadband noise at the maximum operating level as a source signal, all channels should produce an equal sound pressure level at the reference listening position. If the maximum reproduction level is L_{max}, then each channel should be adjusted to give a sound pressure level (RMS, IEC/A-weighted, slow) at the reference listening position of:

$$L_{ref} = L_{max} - 10 \log n \text{ dB(A)}$$

where n is the number of reproduction channels in the total set-up.

In general, it should be recognised that operating loudness level that is used in practice will depend on several additional parameters which cannot be completely specified, such as:

- The kind or nature of the programme material and its individual dynamic range.
- The relative level of the recording to the reference level and how it was monitored (peak/true peak/VU metering, with or without headroom vs full scale).
- The measuring signal (broadband/narrow-band) and the measuring method (weighting).
- Last but not least, the individual need or taste of the listener.

Therefore, in many cases, the actual operating level which is set may differ from the reference level by as much as ±10 dB. In practice, L_{max} will be chosen in accordance with the listeners preferences, but a good average value is 85 dB(A). The value of L_{max} will depend on the programme content, spectral content and loudness, and on the peak-to-average ratio of the programme.

2.3 Conversion of 3/2 audio signals for recording on 4-channel systems

Most existing installations use 2-channel stereophonic systems for distribution, recording, routing and monitoring. In addition, many recorders, especially in television, have 4 audio tracks, usually based on two AES/EBU digital interfaces. It would be useful if 5.0 or 5.1 multichannel signals could be converted into a compatible format that could exploit these existing installations and equipment.

The aim of such a conversion would be to allow existing equipment to process the converted signals transparently, as though they were 2-channel stereophonic signals. This would also allow the signals to be monitored at least to check for the content, levels and basic quality. Inevitably, more than one such 2-channel signal will be used to carry the multichannel signal.

2.3.1 Proposed WARP format

The Warp Corporation has proposed a system that condenses the 5.1 signals of the multichannel format into two compatible AES/EBU digital bit streams [13]. Three of the four available channels are used for the front audio signals, which are entirely unaffected by the conversion process. The two surround signals are combined by halving their bandwidths and sample rates and interleaving the two sets of samples. The LFE channel, if present, may be carried by low-bandwidth modulation of the lower significant bits of the combined surround signal. Metadata can also be carried by similar modulation.

AES/EBU stream	A		B	
Sub-frame	1	2	1	2
Audio channel	L	R	C	LS* + RS* + LFE**

In this way, the AES/EBU stream A appears to be a normal L/R stereophonic pair and can be monitored as such. It should be noted however, that this is **not** the compatible stereo signal, $L_0 + R_0$, because it does not contain any information from the centre or the surround signals.

The full-bandwidth channel C carried in AES/EBU stream B can be monitored as one of a pair of normal AES/EBU stereo signals. The signal in the second channel of stream B is a close approximation of a half-bandwidth combination of the surround sound signals and can be monitored as if it were a normal AES/EBU stereo signal.

Although, this is a straightforward approach, it does not provide a “normal” two-channel signal in a studio without additional signal processing. Moreover, it assumes the information in the surround channels is ambience so that the system is not suitable for multichannel signals that contain discrete or independent sound content in the surround channels.

2.3.2 Proposals using compressed signals (SMPTE)

A different approach is to map low bit rate encoded versions of the 5.1 signals into one or more AES/EBU digital stream.

SMPTE specify a method to carry various types of non-PCM audio data in the payload of AES/EBU bit streams. The formatting of ATSC A/52 (AC-3) data has already been published as SMPTE 340M [14]. The use of MPEG-1 and MPEG-2 as well as Dolby-E encoded data is foreseen in a similar way [15], [16], [17]. The channel status of the AES/EBU stream should signal “non-audio use”, so the signal should not produce any output from a D-to-A converter. The audio content of bit stream, however, cannot be monitored directly but first must be decoded into separate signals for each channel.

The MPEG-1, MPEG-2 and AC-3 audio compression systems are standardised, well-tested and understood coding schemes. Dolby-E, however, is a relatively new coding scheme (although it may be close to AC-3) and it has not currently been submitted to any internationally controlled quality tests.

All these proposals use processes that remove information from the signals. Because the coding process probably occurs in an early stage of the production process, any post-processing would involve cascading of compression stages and this may adversely influence audio quality.

In theory, a lossless coding system would be more appropriate for this purpose but such a system has not yet been implemented.

2.4 Backwards compatible multichannel audio coding

This section is based on an internal EBU document [18] that provides guidance for broadcasters and service providers on how to implement the MPEG-2 Layer II [19] audio coding, which is the International Standard recommended for multichannel audio in television in ITU-R BT.1196 [20].

2.4.1 Coding configurations

The MPEG Audio standard is a generic standard, which can be adapted in several ways to particular applications. The coding configuration can be chosen depending on the input material. Among the options available are some that cover stereo/multichannel compatibility and several for both fixed and variable bit rates.

An MPEG-2 multichannel encoder will automatically select an appropriate coding option for the encoded bit stream in accordance with the properties of the source material. In some cases, there are still options left open to the content provider. It should be borne in mind that for contribution and distribution applications, none of the composite coding modes should be applied in the encoder.

Two-channel stereo material

In this case the coding will be according to MPEG-1 Audio, Layer II. The use of joint stereo coding techniques (also known as intensity stereo) is advised to lower the required bit rate for a given audio quality.

Two-channel Dolby Surround material

In this case the coding will also be according to MPEG-1 Audio, Layer II. To avoid loss of the surround information, joint stereo coding techniques should be avoided, or used with special care. Some encoders have a special ‘Pro-Logic’ setting that restricts the use of joint stereo coding to those cases where the sound image would not be affected by Pro-Logic decoding.

5.1-channel material

In this case the coding will be according to MPEG-2 Audio, Layer II. This coding of a MPEG-2 multichannel bit stream is designed to be compliant with MPEG-1 so that an MPEG-1 stereo decoder can also decode the bit-stream. A matrixing technique is used, so that the decoded two-channel audio is a down-mix of the multichannel original. The down-mixing properties depend upon the matrix procedure selected (see next section).

2.4.2 MPEG-2 matrix procedures

Four matrix procedures are defined in MPEG-2 to derive a compatible two-channel stereo signal from the received 5-channel multiplex. The chosen procedure is signalled in the bit stream and can be interpreted by a MPEG-2 decoder.

matrix procedure '00'

Matrix equations:

$$L_0 = L + 0.707*C + 0.707*LS$$

$$R_0 = R + 0.707*C + 0.707*RS$$

L_0 is a downmix of *L* (at 0 dB relative level), *LS* (at -3 dB relative level) and *C* (at -3 dB relative level). R_0 is a downmix of *R* (at 0 dB relative level), *RS* (at -3 dB relative level) and *C* (at -3 dB relative level).

When this matrix procedure is used, then an MPEG-1 decoder will yield a stereo compatible signal, suitable for two-channel reproduction.

Matrix procedure '01'

Matrix equations:

$$L_0 = L + 0.707*C + 0.5*LS$$

$$R_0 = R + 0.707*C + 0.5*RS$$

L_0 is a down-mix of *L* (at 0 dB relative level), *LS* (at -6 dB relative level) and *C* (at -3 dB relative level).

R_0 is a down-mix of *R* (at 0 dB relative level), *RS* (at -6 dB relative level) and *C* (at -3 dB relative level).

When this matrix procedure is used, then an MPEG-1 decoder will yield a stereo compatible signal, suited for two-channel reproduction. The primary difference between this matrix procedure and matrix procedure '00' is the relative level of the *LS* and *RS* signals in the down-mix. Selection between matrix procedures '00' and '01' may be done on an artistic basis, depending on the characteristics of the encoded material.

Matrix procedure '10'

Matrix equations:

$$Lt = L + 0.707*C - 0.707*jS$$

$$Rt = R + 0.707*C + 0.707*jS$$

jS is a down-mix of *LS* and *RS* as follows:

$$jS = LS + RS$$

where the "*j*" indicates a constant phase difference of 90 degrees of *jS* relative to *L*, *R* and *C*.

Lt is a downmix of *L*, *C* and phase-shifted versions of *LS* and *RS*.

Rt is a downmix of *R*, *C* and phase shifted versions of *LS* and *RS*.

Down-mixing is performed in such a way that a Dolby Surround compatible signal is obtained. As a result, an MPEG-1 decoder will yield a two-channel signal which can be fed to, and decoded by, a Dolby Pro-Logic decoder. Alternatively, the signal from the MPEG-1 decoder can be reproduced "as-is" in a two-channel fashion. In this case, the result will be artistically different from when matrix procedures '00' or '01' have been used, due to the way the surround signals are processed.

The encoding process in matrix modes 00, 01 and 10 are illustrated in the figures 1a and 1b.

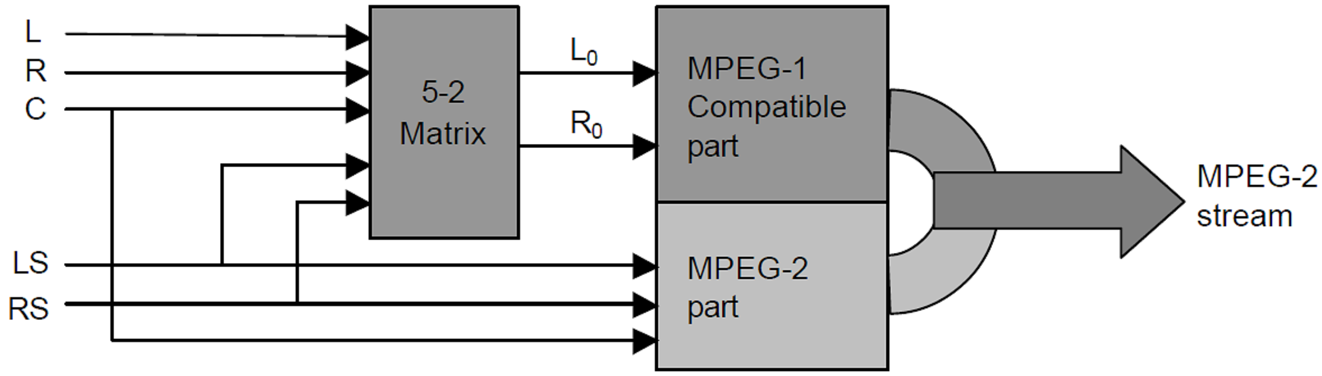


Figure 1a: The encoder in matrix modes 00 and 01

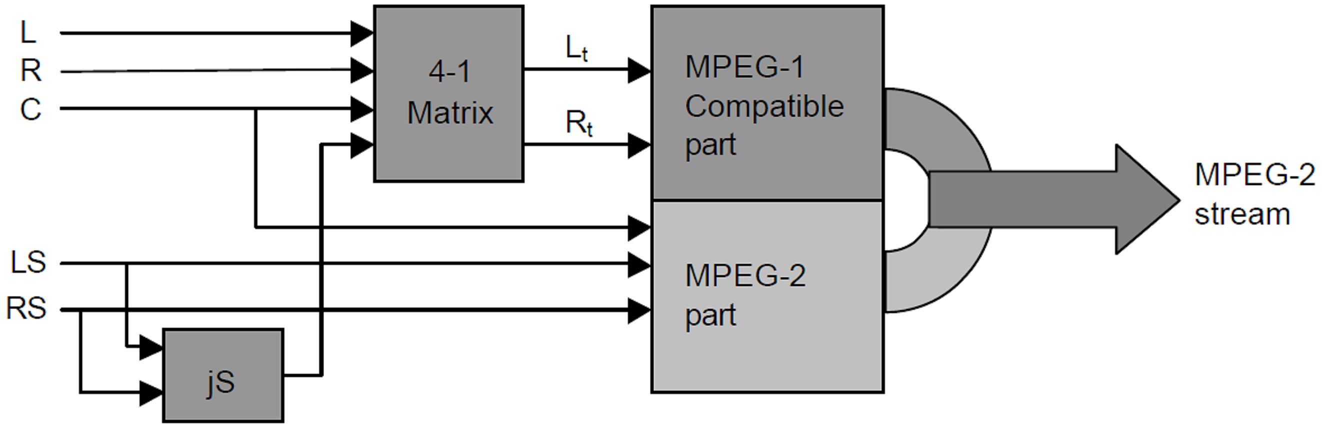


Figure 1b: The encoder in matrix mode 10 (Dolby)

MPEG decoding is shown schematically in Figure 2.

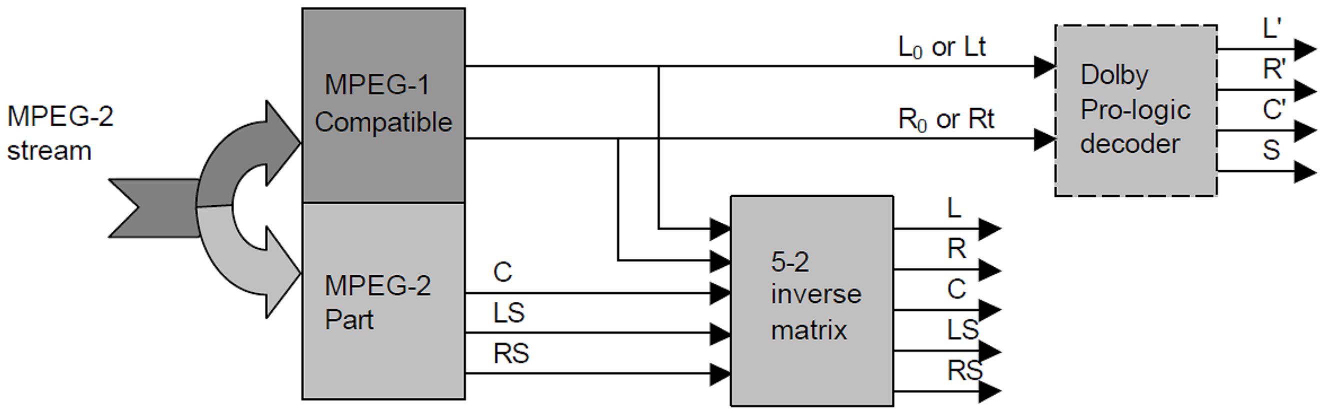


Figure 2: The decoder in matrix mode

It should be noted that, as far as multichannel decoding is concerned, all the matrix modes produce the same result at the multichannel output.

The Dolby Pro-logic decoding is optional. If it is not present, the output of the MPEG-1 compatible part will be L_0, R_0 or Lt, Rt, depending on the matrix procedure chosen in the encoder.

Matrix procedure ‘11’

Matrix equations:

L_0 equals L.

R_0 equals R.

The multichannel signal is encoded in a non-matrixed fashion. An MPEG-1 decoder would reproduce only the *L* and *R* signals from the original set of multichannel signals.

Matrix procedure ‘11’ is intended for multichannel services that do not require two-channel compatibility. For example, after a transitional period when most of the receivers have become MPEG-2 Audio Layer II compliant and thus compatibility to MPEG-1 Audio will no longer be required, or in the case where an optimised 2-channel Dolby Surround soundtrack is separately available to a discrete 5.1 channel soundtrack.

Matrix procedures in production

Matrix procedures ‘00’, ‘01’ or ‘10’ should be used for broadcast emission only and should be avoided in production.

Other matrix procedures

Different down-mixes to those specified by matrix procedures ‘00’, ‘01’ or ‘10’, e.g. a continuously variable down-mix for different programme content, are technically feasible, but are subjects for further study.

The different down-mix versions can be signalled to the decoder in the ancillary data field to allow the selection of the appropriate compatibility matrix.

2.4.3 Input/output relationships for different MPEG decoders

Different kinds of receiver are expected to be used by the public. These will be equipped (internally or externally) with different types of decoder as shown in Table 4.

Table 4: Types of decoder fitted to Receivers

Decoder	Typically applied in receiver type
MPEG-1	portable / mobile / low end
MPEG-2 + down-mixing	stationary / mobile / standard
MPEG-2	stationary / mobile / high end or external

Because the MPEG-2 encoding-decoding process preserves the integrity of the channels, the original signals, *L*, *R*, *C*, *LS* and *RS* are available in an MPEG-2 decoder.

Therefore, an MPEG-2 decoder with down-mixing capability will enable the consumer to change the down-mix properties according to his/her needs. The decoder can, for example, construct a Dolby Surround compatible stereo signal independent of which down-mix matrix was selected for the bit stream.

Table 5 shows the output audio signal of different MPEG-Audio decoders when decoding a bit stream with a given combination of coding properties.

Table 5: Output audio signals from MPEG -Audio decoders for different MPEG bit-streams

Input signal	Encoding	Decoder	Output signal
Stereo	MPEG-1	MPEG-1	Stereo
		MPEG-2 down-mix	Stereo
		MPEG-2	Stereo
Dolby Surround	MPEG-1	MPEG-1	Dolby Surround
		MPEG-2 down-mix	Dolby Surround
		MPEG-2	Dolby Surround
5.1 discrete	MPEG-2, matrix mode ‘00’, ‘01’	MPEG-1	Stereo
		MPEG-2 down-mix	Dolby Surround or Stereo
		MPEG-2	5.1 discrete
	MPEG-2, matrix mode ‘10’	MPEG-1	Dolby Surround
		MPEG-2 down-mix	Dolby Surround or Stereo
		MPEG-2	5.1 discrete
Dolby Surround and 5.1 discrete	*Dolby Surround*: MPEG-1	MPEG-1	Dolby Surround
	5.1: MPEG-2, matrix mode ‘11’	MPEG-2 down-mix	Dolby Surround or Stereo
		MPEG-2	5.1 discrete

2.4.4 Compatibility with Dolby Surround

If compatibility with Dolby Surround is required, the matrix mode ‘10’ should be used. Figure 2 shows the typical output configuration for MPEG-1 and MPEG-2 Audio decoders.

2.4.5 MPEG bit rates

The MPEG-2 Layer II Audio coding standard allows bit rates from 32 to 1066 kbit/s to be used, including the 15 bit rates, defined by MPEG-1.

5.1 multichannel audio signals

The bit rates recommended for contribution and distribution are valid for the non-matrixed mode. The bit rate recommended for emission is valid for matrixed mode.

For contribution, a bit rate of 860 kbit/s should be used for the 5.1 format to fulfil the audio quality levels described in the requirements document for multichannel low bit rate audio coding. A bit rate of up to 213 kbit/s per channel for cascading and/or post-processing stages may be used.

For distribution, a bit rate of 640 kbit/s should be used to fulfil the audio quality levels, described in the requirements document for multichannel low bit rate audio coding.

For emission, a bit rate of 640 kbit/s should be used to fulfil the audio quality levels, described in the requirements document for multichannel low bit rate audio coding.

Compatible stereo signal

With the bit rate of 640 kbit/s, used for emission, the audio quality level of the compatible stereo signal fulfils the requirement of the EBU and ITU-R for broadcast quality of stereo audio signals.

2.4.6 Low Frequency Enhancement channel

The MPEG-2 Audio Layer II multichannel audio system provides an optional Low Frequency Enhancement/Extension (LFE) channel with a frequency range up to 120 Hz, according to Recommendation ITU-R BS.775 [1]. The purpose of this optional channel is to enable listeners to extend the low frequency content of the reproduced programme in terms of both frequency and level. It is intended to carry high amplitude, low frequency sound effects, which will be heard through special subwoofer loudspeakers. This will avoid the problem of the low-frequency content of the programme using up all the available headroom of the main channels in the digital domain. At the same time these high-level low-frequency signals need not be handled by the main loudspeakers, thereby avoiding overloading. The LFE channel is not intended to be used for the low frequency content of the stereo (front) part of the multichannel sound presentation. The main channels carry the normal low frequency sounds and they should be adequate on their own if the listener does not choose to or cannot use the LFE signal. Note that the LFE channel is not used in the MPEG matrix procedures mentioned in section 2.4.3. It is also not considered in any of the conversion procedures described above.

2.4.7 Multilingual service

In addition to the main programme, which may be 5.1 multichannel, separate multilingual channels are available in the MPEG system. These can be single channels for commentary or dialogue, an alternative two channel stereo presentation or an alternative multichannel presentation. The possibilities are given below:

Number of channels:

Up to 7

Coding options:

Either full rate or half rate coding.

Bit rate:

According to the Report on the subjective testing of coders at low sampling frequencies (ISO/IEC JTC1/SC29/WG11, N0848, Singapore, Nov. 1994) with half-rate coding, a bit rate of 56 kbit/s provides the quality required for speech. Half rate coding limits the bandwidth to about 11.5 kHz, which has no significant influence on the quality of speech signals, compared to a full bandwidth.

Multilingual services:

The language(s) can be signalled in the ancillary data. Synchronisation is not required between the multilingual services and the main programme.

This type of multilingual service is not backwards compatible to MPEG-1 Audio. However, if backwards compatibility is required for multilingual services, at least an MPEG-1 Audio two-channel decoder with half-sampling frequency decoding capability, defined by MPEG-2, should be used.

Note: As no other codec system supports the multilingual feature, a conversion to other coding schemes may not be possible.

2.4.8 Non-Backwards compatible multichannel audio coding (Dolby Digital, DTS)

Non-backwards compatible multichannel coding schemes such as Dolby Digital or DTS, can be used in the existing DVB and DVD audio standards. These systems produce a stream of audio frames in which a stereo signal does not exist as a separate part. The channels in these systems should not be considered to be completely independent because the encoding uses mutual masking between adjacent channels. As with MPEG-2, the number of channels encoded in one stream in the Dolby AC-3 (Dolby Digital) system can be freely selected.

Non-backwards compatible encoding methods allow a more precise calculation of the masking properties of the signals, which can result in a lower average bit rate for the same audio quality.

2.5 Cascading

Cascading is the occurrence of multiple decoding / encoding-processes of the same programme signal along a transmission chain. This can be a result of the structure or other technical requirements of the chain.

Cascading is unavoidable in a complete audio and television broadcasting chain, including programme contribution and distribution links. As mentioned above, a down-mix matrix should only be applied at the emission stage.

MPEG-2 Audio has already been shown to be very robust against cascading. The formal test results have been obtained from tests performed in 1995. A chain of 6 coding stages was tested. However, the consequences of cascading between MPEG-2 Audio and other multichannel audio coding systems, mainly those used in cinema or video applications, are not yet known.

2.6 Codec conversion

Because there is no clear relationship between the encoded data in compatible and the non-compatible coding systems, direct conversion between them is unlikely. Instead, codec conversion will involve decoding the available channels to a linear representation, and re-encoding them.

2.7 Transcoding

If an audio signal has already been previously coded and decoded, any following encoding will in general benefit from intelligent transcoding. This process depends on knowing and using the coding decisions (scale factors and quantization per sub-band in the case of MPEG encoding) made by the previous encoder, as well as knowing the position of the data-frames in time. When the frame structure or the sub-band structures used by the subsequent coding algorithms are different from the preceding one, intelligent transcoding becomes very complicated. In the case of MPEG-BC multichannel encoding there is an additional complicating factor; the algorithm used to determine the quantization and the scale factors depends on the matrix mode that was in use. Probably, transcoding is only feasible if the same matrix mode is used again.

In addition to the above problems, there is also a difference in the frame-structure between MPEG-2 data and Dolby AC-3, which makes transcoding between these systems very difficult.

3. Up-conversion (Up-mixing)

3.1 2-channel to a higher number of channels)

Historically, conventional 2-channel stereophonic reproduction has mainly been a means of conveying discrete phantom image localisation and is often a highly limited compromise solution to the problem of conveying a more general spatial impression to the listener. Up conversion assumes that a conventional 2-channel stereophonic signal may (by chance) include spatial information that can be extracted to provide additional signals to drive a multiplicity of spaced loudspeakers. In the 1970s some experiments were performed using signals derived from the difference signal L – R. This signal was usually reproduced via 2 loudspeakers in anti-phase, as L – R at the left-hand side, and R – L at the right-hand side. This was known as the Hafler matrix. It was frequently found that additional low-pass filtering of the signal was necessary, and additional delays were helpful to keep the front images in front.

Very few formal tests of up-conversion algorithms have been described in published work [30]. One recent investigation into the subjective performance of several up-conversion algorithms was carried out as part of the EUREKA Project 1653, called MEDUSA [21]. The aim was to study methods by which sound reproduction could be enhanced for consumers using multichannel techniques. It included formal listening tests carried out according to the requirements of Recommendation ITU-R BS.1116 [11].

The different algorithms tested were:

1. A digital surround implementation with stereo surrounds but apparently without steering; delayed and filtered rear channels; optimised for music reproduction.
2. An analogue matrix process involving a dynamic centre channel and steered stereo rear channels; low-pass filtered rear channels but no additional rear channel delay; optimised for music applications.
3. A digital algorithm involving stereo rear channels and a degree of logic -controlled steering; delayed and filtered rear channels, optimised for music applications.
4. A four-channel (LCRS) spatialization algorithm, with mono surround split to both rear loudspeakers; designed as an alternative to Dolby Surround for use with unmatrixed material in consumer products.

The results of these tests can be summarised as follows:

1. Front images - all processors were clearly judged to make the front images poorer than that of the two-channel original material by at least one grade. Some of the processors had only small effects, but one showed very marked reductions in front image quality. The different processors had different effects on the front image quality such as, variously affecting apparent width, sharpness and location.
2. Spatial impression - overall, the processors made small improvements, ranging from essentially zero up to +1 grade, but the mean results disguised the fact that listeners responded very differently to the spatial effects produced by the different processors. Some graded the spatial impression of the processed versions much higher than the unprocessed; others graded them much lower. This was attributed partly to the difficulty the listeners had distinguishing between the quality and quantity of the spatial impression. Individual subjects tended to be self-consistent.
3. Preference - overall, the subjects showed a strong preference for the original, unprocessed material. This was more marked for some processors than others, but it was shown in every case. The results depended significantly on the type of programme. The preferences also depended strongly on the processors; in one case, the processed version was preferred about 50% of the time.

Overall, the tests showed that synthesising the surround signals from a two-channel stereophonic input was generally not preferred by the majority of expert listeners, but that at least one processor could create an overall impression that was preferred by nearly half of the subjects.

The inevitable conclusion is that there may be a place for up-conversion techniques in the listener's own environment. However, because of the overall lack of preference for the processed signal, it seems clear that techniques based on state-of-the-art algorithms today could not reasonably be applied generally - except, perhaps, for some types of programme material. Even then, probably only half of the audience would consider the process beneficial. This conclusion might be modified in future, if improved algorithms for up-conversion become available.

More recent publications [22] or [23] describe some other conversion techniques from stereo to five-channel sound systems which have not yet been tested in detail.

3.2 4-channel to 5-channel

Some broadcasters and recording studios (for instance, Hungarian Radio [24] or Denon) have many experimental recordings using 4-channel formats in their archives. It seems that it would be much more straightforward to up-convert this type of material to 3/2 format than any material in 2-channel format. A simple way to generate the missing centre signal could be to use a Dolby Surround decoder;

but the results should be checked carefully. Alternatively, some users propose the transmission of existing 4-channel programmes via the 3/2 environment as it is.

3.3 Virtual panning

Virtual surround processing systems can provide simple conversion methods between different stereo and surround signal formats. Such systems have been introduced recently in the form of tools integrated into digital mixing consoles or in effects processors or as separate computer applications. Virtual panning takes account of the physical principles and psychoacoustics of human hearing, to exploits important sound localisation properties. [25], [26], [27], [28].

To position a sound source in the virtual space, the systems generate several individual reflections for each channel, usually by means of DSP processors. These reflections are distributed to the surround channels as a function of the selected pan position. Some form of overall, pseudo-randomised 'ambience' can also be added, if required. The algorithms are usually based on a mixture of geometric and stochastic models, which can be defined by a few "perceptual" control parameters. The potential for changing the characteristic parameters describing the room and the position of the sound sources in the virtual environment is also very important. The sound character can be varied from cold, non-absorbing halls to warm, strongly-absorbing living rooms and the amplitude of all the virtual reflections, and consequently the entire ambience effect, can be boosted or suppressed. The advantage to the user is that different multichannel output formats can easily be obtained while the backward compatibility is guaranteed.

Note: All these virtual panning methods rely on clean and separate (i.e. not virtual) source signals being available. If they were to be used with existing pre-mixed signals (such as 2-channel stereophonic), each channel will be treated as a discrete single sound source. Therefore, conversion from existing 2-channel to multichannel recordings seems rather difficult and would need careful checking of the quality of the results.

Therefore, these methods are primarily suited for down-mixing of multitrack-recorded material than for conversion of already-mixed recordings from one to another audio format.

4. Conversion of data between different multichannel coding or transmission systems

Digital delivery systems can contain a variety of data signals. Some can be used to enhance the audio signal. For instance, DAB transmissions can include Dynamic Range Control information (DRC) and the indication of music or speech (Music/Speech flags). These are carried in the Programme Associated Data channel, PAD, as data synchronous to the audio programme. PAD can also carry other programme related but non-synchronous information (e.g. radio-text, HTML pages, JPEG-coded still pictures, Dynamic Labels, Universal Product Code/European Article Number (UPC/EAN). etc.

These signals need to be transferred or converted when the audio signals themselves are converted.

4.1 Dynamic Range Control

It has long been recognised that many listeners in less than perfect environments find it is impractical to make use of the full dynamic range which may be carried by a digital audio signal. With the help of the Dynamic Range Control (DRC) signal the receiver can reduce the dynamic range of the audio signal. The purpose of this is either to adapt the dynamic range of the audio signal to listening in a noisy environment or to adapt an audio source that has a too high dynamic range (typical movie sound tracks) for domestic listening.

A method of dynamic range control has already been defined for Digital Audio Broadcasting (DAB), see EN 300 401 [29]. In this case the audio signal is transmitted with full dynamic range and an additional control data is carried in the coded bit stream. In the receiver, the regenerated DRC data may be used to control the audio gain, i.e. to compress the dynamic range of the audio signal to match it to the requirements of the listener to improve audibility of the multichannel audio signal in difficult conditions.

DRC data may be also be incorporated as Programme Associated Data (PAD) in the MPEG-2 Layer II ancillary data field [30]. A detailed specification of DRC signalling for multichannel audio for DAB is currently under consideration in the EUREKA-147 DAB project.

4.2 Dolby metadata

Both the Dolby Digital and Dolby-E multichannel audio systems contain metadata used to describe the encoded audio signals. This metadata is carried within the bit streams and conveys information used to control downstream decoders and/or encoders. Metadata allows the content provider to control how the original programme material is reproduced in the home.

The Dolby E distribution system, designed for broadcast applications, such as programme origination and distribution, can carry up to eight encoded audio channels, arranged as any number of programmes from one single programme to eight individual programmes. Each discrete programme has its own metadata. Some of this metadata is used to automatically configure a downstream Dolby Digital encoder, while some is passed through to the consumer's Dolby Digital decoder.

Dolby Digital is used for consumer applications, such as transmission to the home (sometimes called an emission bit stream) or for DVD Authoring. The bit stream contains a single encoded programme of up to six channels described by common metadata. Dolby Digital metadata is a subset of the Dolby E metadata, containing only those parameters necessary for proper decoding by the consumer. The Dolby Digital decoder in the home processes the audio according to parameters in the metadata stream set by the programme creator as well as others (for example, bass management and dynamic range) set by the consumer to reflect his equipment and environment.

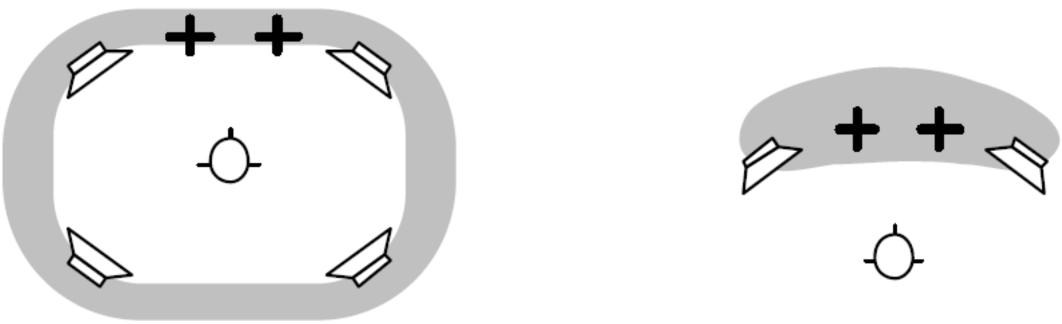
4.3 Conversion of Dolby metadata to PAD and vice-versa

Further work is needed on this subject.

5. Problems of artistic compatibility of quadraphonic recordings

In the 70s and 80s, several radio companies, for instance Hungarian Radio [24], made experimental quadraphonic recordings, mostly of radio drama. Because these recordings were broadcast in a 2-channel system, their compatibility with stereophonic reproduction was of primary importance during production.

Technically (i.e. concerning level and phase conditions), there were no problems converting these recordings to conventional 2-channel stereophony by matrixing, but from the artistic point of view the conversion was not always successful.



Quadraphonic reproduction **After Stereo matrixing**

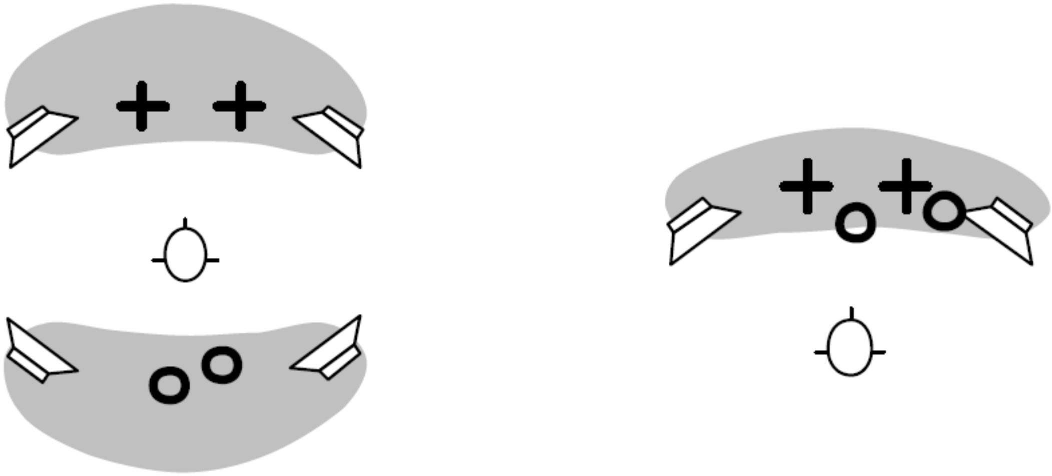
Figure 3

One of the sound stages used for radio drama was arranged similarly to those used for musical performances, where the dialogue was placed in the front, and the back and rear parts were only used for ambience (see Figure 3).

In this simple scheme there is no problem with stereo compatibility; the sound images between the rear loudspeakers, and those at the sides can easily be placed at the front.

If the sounds placed at the sides and the rear are relatively loud, however, mixing them with the frontal sound may impair the recognition of the speech. Therefore, the levels of the back channels were reduced by a few dBs when making the stereo down mix. A reduction in the level of the rear channels is especially advisable in the mix down if they contain very distinctive sounds.

A special situation arises when the front and the rear sets of channels are used to depict two different scenes or sound stages that occur simultaneously to tell two stories at the same time (see Figure 4).

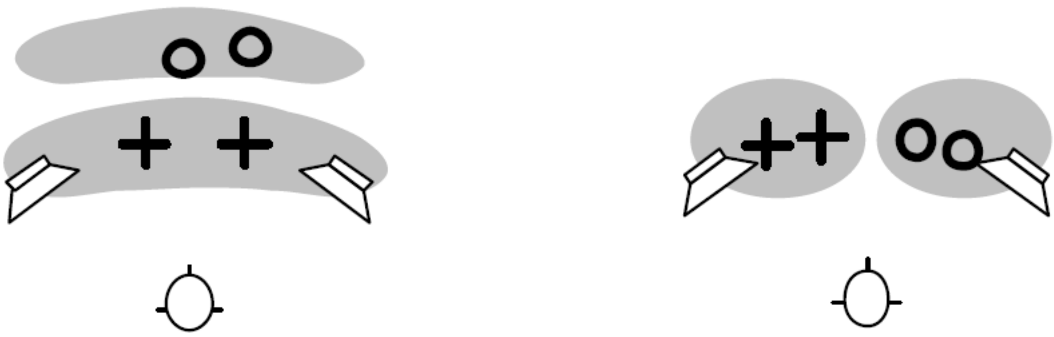


Quadraphonic reproduction with discrete rear stage **Confusion after stereo matrixing**

Figure 4

The first can be a real story and the other, behind the listener can be some vision of memory. Evidently, by mixing the rear channels with the front, this effect is lost, and confusion arises.

If the sound and the ambience of the front and rear scenes are very different, for instance one sounds very close, and the other very distant, then mixing them together can be acceptable (Figure 5).

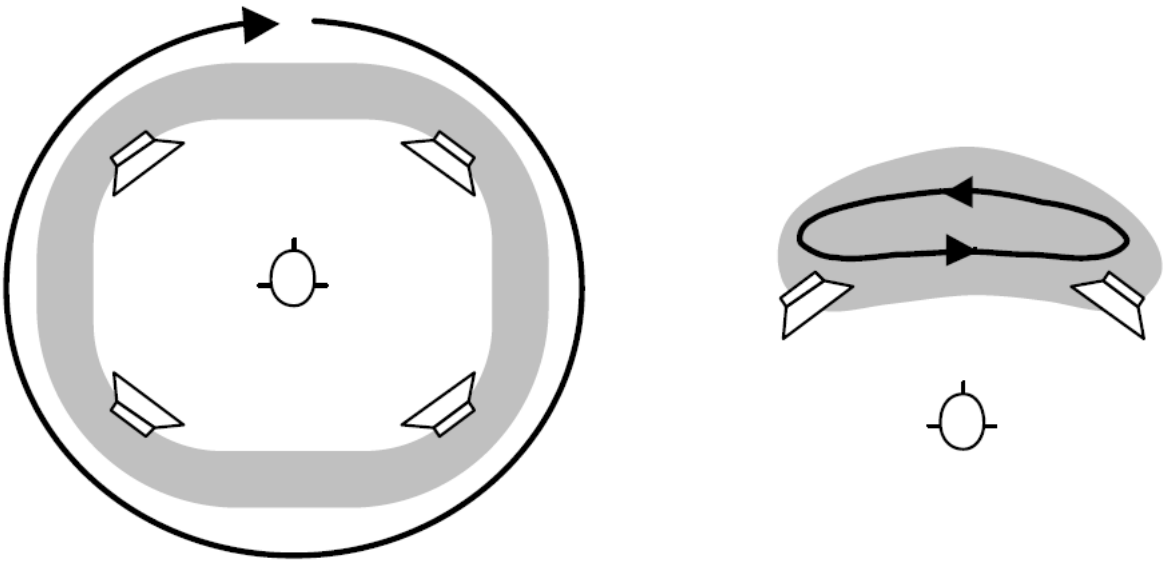


Less confusion if front and rear sounds are very different **Panning front and rear to different sides of stereo stage**

Figure 5

Putting the two scenes in the left and right stereo-channels may achieve less confusion but would result in ping-pong stereo, that had nothing to do with the essence of quadraphony.

Sometimes, a sound source was recorded as if it were moving around the listener, as in Figure 6. This can be very effective in quadraphony; the movement itself was usually unreal or symbolic and was used to crown the dramatic suspense.



Circulating sound in quadraphony **Effect after stereo matrixing**

Figure 6

When mixed down to stereo, however, the quadraphonic circular motion is reduced to a simple oscillation from left to right and its dramatic effect was far less. Some useful effect can remain in stereo if the width of the scenes in the front and the back are very different, for instance because the channel separation is made far greater in the front than in the back.

The stereo mix became a thing in its own right, not really representing the quadraphonic original. This sort of movement around the listener cannot be replaced by anything in stereo.

Neither can lateral movements in the rear be converted into two-channel stereo. As an example, in the battle-scene in the opera 'Boris Godunov' by Pushkin, the two armies were situated one in front of the listener and the other behind him. The hustling atmosphere of the battle was well evoked by the lateral movements and by the movements of the pursuing and fleeing soldiers.

Moreover, in case of circular movement and at the movement of the lateral base, the compatibility is improved if the signal coming from the two rear channels are mixed to the front channels with reduced base. Sometimes, we used this method, too.

Generally, it can be stated that quadraphonic radio plays are not compatible with the two-channel stereo technique from an artistic point of view. The more the possibilities of quadraphony are exploited in the four-channel version, the less are the chances of compatibility in the stereo version.

6. References

[1] ITU R Recommendation BS.775 1 (1994): *Multichannel Stereophonic Sound System with and without accompanying Picture.*

[2] EBU Technical Recommendation R96–1999: *Formats for production and delivery of multichannel audio programmes.*

[3] EBU Technical Report BPN 021 (1999): *Multichannel Audio: Different reproduction and delivery formats available on the market.* (Available to EBU Members only)

[4] ISO/IEC: International Standard IS 11172 (1993) *Information technology – Coding of moving pictures and associated audio for digital storage media up to about 1.5 Mbit/s (MPEG 1). Part 3: Audio*

[5] ITU R Recommendation BR.1384 (1998) *Parameters for international exchange of multichannel sound recordings*

[6] EBU Technical Recommendation R91–1998: *Track allocations and recording levels for the exchange of multichannel audio signals*

[7] SMPTE RP 200 1999 *Relative and Absolute Sound Pressure Levels for Motion Picture Multichannel Sound Systems*

[8] SMPTE 320M 1999 *Television Channel Assignments and Levels on Multichannel Audio Media*

[9] IEC Publication 268 10 1998: *Peak programme level meter*

[10] EBU Technical Recommendation R68 1992 : *Alignment level in digital audio production equipment and in digital audio recorders*

[11] ITU R Recommendation BS.1116 1 (1997). *Methods for the subjective assessment of small impairments in audio systems including multichannel sound*

[12] EBU Recommendation R22 2000: *Listening conditions for the assessment of sound programme material; with Tech 3276 1997 (2nd edition) and Supplement 1 to Tech 3276 1997 (2nd edition): Listening conditions for the assessment of sound programme material: multichannel*

[13] DLE6400 *Digital Linear Encoder, Product information, The Warp Corporation Ltd, Cambridge, UK*

[14] SMPTE 340M 2000 *Television Format for Non PCM Audio and Data in an AES3 ATSC A/52 (AC 3) Data Type*

[15] SMPTE 337M 2000 *Television Format for Non PCM Audio and Data in an AES3 Serial Digital Audio Interface*

[16] SMPTE 338M 2000 *Television Format for Non PCM Audio and Data in an AES3 Data Types*

[17] SMPTE 339M 2000 *Television Format for Non PCM Audio and Data in an AES3 Generic Data Types*

[18] EBU doc, BMC 337 / B/CASE 047 (1998) *Appendix 6: Backwards compatible multi channel audio coding: Guidelines for the user*. (Available to EBU Members only)

[19] ISO/IEC: International Standard IS 13818 3 (1998) *Information technology – Coding of moving pictures and associated audio information – Part 3: Audio (MPEG 2)*

[20] ITU R Recommendation BT.1196 (1995): *Audio coding for Digital Terrestrial Television Broadcasting*

[21] Rumsey, F. Journal of the AES – 1999 Vol. 47, No. 7/8, *Controlled Subjective Assessments of Two to Five Channel Surround Sound Processing Algorithms.*

[22] Tuzman, A.K. AES 19th International Conference, Munich, 2001, *Conversion from stereo to multichannel using array processing techniques*

[23] Irwan, R. and Aarts, R.M. AES 19th International Conference, Munich, 2001, *A method to convert stereo to multichannel sound systems*

[24] Ujházy, L. (1999) *Compatibility Problems of Quadraphonic Radio Plays. Hungarian Radio, Budapest*

[25] U. Horbach: AES 103. Convention, New York, 1997, Preprint 4624, *New Techniques for the Production of Multichannel Sound.*

[26] V. Pulkki: Journal of the AES – June 1997, Vol.45, No. 6, *Virtual Sound Source Positioning Using Vector Base Amplitude Panning,*

[27] Robert Walker: AES 106th Convention, Munich, 1999, Preprint 4937, *A simple acoustic room model for Virtual Production.*

[28] Robert Walker: AES 108th Convention, Paris, 2000, Preprint 5138, *Approximation functions for Virtual Acoustic Modelling"*

[29] ETSI EN 300 401 (Edition 2 1998): *Radio broadcasting system; Digital Audio Broadcasting*

(DAB) to mobile, portable and fixed receivers

[30] Gerzon, M. A.: Various publications to the topic, see [21]

[31] Hoeg, W. Journal of the AES, -1998, Vol.46, No. 4, *Dynamic Range Control (DRC) for Multichannel Audio Systems*

7. Bibliography

[32] EBU Technical Report BPN019 (1999): *Report on the EBU Subjective Listening Tests of Multichannel Audio Codecs. (EBU Project B/CASE)* (Available to EBU Members only)

[33] ITU-T Recommendation J.52: *Digital Transmission of High-Quality Sound-Programme Signals using one, two or three 64 kbit/s Channels per Mono Signal (and up to six per Stereo Signal)*

[34] Meares, D. J. and Theile, G. Journal of the AES, April 1998 Vol. 48, No. 4, *Matrixed Surround Sound in an MPEG Digital World*

[35] ten Kate, W. (1994) AES 96th Convention, Amsterdam (Preprint 3792), *Compatibility Matrixing of Multi-channel Bit rate-Reduced Audio Signals*

[36] ITU- R Doc. 11/4 (1996): *'A Guide to Digital Terrestrial Television Broadcasting in the VHF/UHF Bands', section 3.7, 'Audio Compression and Coding'.*

8. Acknowledgements

This report is based on existing standards, publications and working documents of the EBU P/MCA, P/AFT and former B/CASE project groups. The main contributors to the report were:

Wolfgang Hoeg	T-Nova Berkom	Germany
Ulf Wüstenhagen	T-Nova Berkom	Germany
Gerhard Stoll	IRT	Germany
Milan Gazdik	CTV	Czech Republic
László Ujházy	MRT	Hungary
Tom Magchielse	NOS/NOB	Netherlands
Erik Lundbeck	SVT	Sweden
Vuslat Öztopçu	TRT	Turkey
Bob Walker	BBC	UK
John Emmett	UKITV	UK
Richard Chalmers	EBU	Switzerland