International Telecommunication Union

ITU-R
Radiocommunication Sector of ITU

Report ITU-R BS.2466-1
(09/2022)

Guidelines for the use of the ITU-R ADM Renderer

BS Series

Broadcasting service (sound)

Foreword

The role of the Radiocommunication Sector is to ensure the rational, equitable, efficient and economical use of the radio-frequency spectrum by all radiocommunication services, including satellite services, and carry out studies without limit of frequency range on the basis of which Recommendations are adopted.

The regulatory and policy functions of the Radiocommunication Sector are performed by World and Regional Radiocommunication Conferences and Radiocommunication Assemblies supported by Study Groups.

Policy on Intellectual Property Right (IPR)

ITU-R policy on IPR is described in the Common Patent Policy for ITU-T/ITU-R/ISO/IEC referenced in Resolution ITU-R 1. Forms to be used for the submission of patent statements and licensing declarations by patent holders are available from http://www.itu.int/ITU-R/go/patents/en where the Guidelines for Implementation of the Common Patent Policy for ITU-T/ITU-R/ISO/IEC and the ITU-R patent information database can also be found.

<table>
<tr><td colspan="2" align="center">**Series of ITU-R Reports**</td></tr>
<tr><td colspan="2" align="center">(Also available online at http://www.itu.int/publ/R-REP/en)</td></tr>
<tr><td>**Series**</td><td align="center">**Title**</td></tr>
<tr><td>**BO**</td><td>Satellite delivery</td></tr>
<tr><td>**BR**</td><td>Recording for production, archival and play-out; film for television</td></tr>
<tr><td>**BS**</td><td>**Broadcasting service (sound)**</td></tr>
<tr><td>**BT**</td><td>Broadcasting service (television)</td></tr>
<tr><td>**F**</td><td>Fixed service</td></tr>
<tr><td>**M**</td><td>Mobile, radiodetermination, amateur and related satellite services</td></tr>
<tr><td>**P**</td><td>Radiowave propagation</td></tr>
<tr><td>**RA**</td><td>Radio astronomy</td></tr>
<tr><td>**RS**</td><td>Remote sensing systems</td></tr>
<tr><td>**S**</td><td>Fixed-satellite service</td></tr>
<tr><td>**SA**</td><td>Space applications and meteorology</td></tr>
<tr><td>**SF**</td><td>Frequency sharing and coordination between fixed-satellite and fixed service systems</td></tr>
<tr><td>**SM**</td><td>Spectrum management</td></tr>
</table>

Note: This ITU-R Report was approved in English by the Study Group under the procedure detailed in Resolution ITU-R 1.

Electronic Publication
Geneva, 2022

© ITU 2022

REPORT ITU-R BS.2466-1

Guidelines for the use of the ITU-R ADM Renderer

(2019-2022)

1 Overview of ITU-R ADM Renderer

1.1 Introduction

Advanced Sound Systems[1] (AdvSS) technology[2] is a new challenge for the broadcast and broadband industries. Users are strongly encouraged to share their experiences with the ITU-R to help create the most valuable documentation for AdvSS. In particular the experiences of content creators and system developer are very much welcomed.

This Report is intended to provide guidelines for the use of the ITU-R ADM Renderer, which is specified in Recommendation ITU-R BS.2127, and which is a complementary component of the ITU-R Advanced Sound System, specified in Recommendation ITU-R BS.2051 and ADM metadata specified in Recommendation ITU-R BS.2076.

The ADM Renderer can be seen as part of the evolution, or an extension, of mixers and down-mixers that are often used today, where the principal of allowing creative control of the listening experience is continued or extended into the environment of AdvSS. The ADM Renderer is also a bridge between today's audio production and the new world of AdvSS.

It is expected that future sound systems used for audio-visual media will include the ITU-R's AdvSS This technology is intended to provide an improved sound experience that can be more 'immersive' for users, and/or provide additional audio signals to give users a more personalised experience. AdvSS require new signal formats to be used, and they may require a greater number of listening-room loudspeaker positions or specialised sound bars (or possibly if used, binaural headphones) than is needed for existing stereo and 3/2 multichannel sound systems specified in Recommendation ITU-R BS.775. The ADM Renderer and its functionality specified in Recommendation ITU-R BS.2127 is based upon known methods that are currently being deployed in production/delivery and reproduction systems for AdvSS/NGA today.

NOTE – In this Report the term 'ADM Renderer' is used to mean the ITU-R ADM Renderer.

1.2 Audio elements and their uses

The ADM Renderer in this case generates the required loudspeaker signals from 'audio elements' that are input to it via a delivery platform. These 'audio elements' are combinations of a particular sound and the required signalling ('metadata') that define how the sound should be reproduced by the receiver. There are three types of audio elements.

1   Channel-based

2   Object-based

3   Scene-based.

Each of these has unique attributes:

---

[1]   Recommendation ITU-R BS.2051 – Advanced sound system for programme production.

[2]   Sometimes known as Next Generation Audio, of which AdvSS is one application.

–   Channel-based – the format essentially used for today's audio: mono, stereo, and multichannel sound systems, where signals correspond to positions of reproduction loudspeakers, and can be fed directly to them.

–   Object-based – where sound 'objects' are recorded, and the mixing and panning information is recorded with them, but not executed until the sound is reproduced.

–   Scene-based – where sound is recorded in standard directional patterns that can be transformed to any desired reproduction geometry. This is also known in its basic form as 'Ambisonics' recording, though a more developed form, Higher Order Ambisonics (HOA), is also used.

The ADM Renderer converts the audio elements to audio signals to feed for loudspeaker configurations or headphones. It may need to address any or all three types of audio elements. The ADM Renderer can be implemented in hardware or software, noting that there may be an impact on the performance and functionality in each case.

The ADM Renderer may avoid the need for multiple version masters for different encoding/transmission systems as well as playback systems. It has the potential to save time, cost, and/or enhance the audience experience, depending on the system configuration.

As content creators and sound engineers adopt advanced sound system technologies, renderers will become a necessary utility in production, postproduction and for delivery to broadcasters. Versions of the renderer will also be part of domestic receiving equipment.

The basic principles of AdvSS are outlined in Recommendation ITU-R BS.2051. The ADM Renderer is specified in Recommendation ITU-R BS.2127, which provides a reference interpretation of the ADM metadata specified in Recommendation ITU-R BS.2076.

1.3   Renderer implementations

As noted in § 1.2, a renderer could be implemented either as dedicated hardware or as software. The cost, size and wiring would therefore be similar to other audio processing equipment, depending on the design and capabilities of the unit.

Depending on the application, renderers may be simple to use, or they may require more sophisticated control and configuration. Renderers that are very simple to use, or transparent to the user, would be desirable in a broadcast or broadband receiver. Renderers with more sophisticated control and configuration options may be needed to meet a sound engineer's production requirements, and to arrange the format required to be delivered to the broadcast or broadband system.

The data rate of the audio element metadata is usually a small fraction of the total audio data rate, but an overhead need to be allowed with streaming or broadcast to accommodate it.

Renderers are necessary in media receivers for consumer use and will be included in appropriate future home entertainment systems. To encourage suppliers to offer equipment with the best listening experience in home entertainment systems, it can be valuable that there is competition for quality/cost in the home systems that different companies provide. Thus the specifications and flexibility of the renderer used in the design of the domestic receiver should be a matter for the manufacturers themselves, provided they can interpret appropriately the audio elements.

Programme makers and those who monitor and check delivery however do necessarily need a renderer that can be made available by multiple manufacturers. This will enable uniformity in the delivery of the rendered format, and the experience the user will have with their delivered media content, given that it is delivered in the correct format.

Programme making can be needed in 'real-time' or following postproduction in 'non-real-time'. Renderers will be needed to cope with both of these circumstances. Currently, a software

implementation of the ADM Renderer is available (link to the software given in Recommendation ITU-R BS.2127) which can be used for non-real-time production.

The ADM Renderer specified in Recommendation ITU-R BS.2127 can thus be used for programme production, programme monitoring, and where needed for subjective evaluations (given other elements of the methodology are standardised). It may also serve as a guide for renderers in domestic equipment.

1.4   The options and usage for the ADM Renderer

The object sound signals delivered for AdvSS are described using positional metadata which uses either spherical coordinates to indicate the sound's relative direction from the listener or Cartesian coordinates to indicate the sound's location in the room. The ADM Renderer can decode and interpret both types of coordinates, and if needed, can convert between them, although such conversion cannot by nature be exact.

This conversion is not invertible, so conversion back and forth should be avoided. The user of the ADM Renderer can make a choice based on how the content was produced, and the likely media user's environment. The ADM Renderer specification also includes information on an alternative Point Source Panner should it be required.

Having a single standardised production renderer will bring greater consistency to the process of AdvSS content production and delivery to the end user. It should increase production levels, and thus eventually result in lower costs and more widely used AdvSS. The alternative would be for programme makers to buy proprietary equipment which, although satisfactory, may restrict the programme maker into using that kind of equipment.

Using the ADM Renderer, the content provider will know that, if they buy end user equipment that conforms to the ITU-R Recommendation, they will be able to cope with all eventualities for coordinate systems and will be able to experience what the listener will experience with his/her programme. Content makers will be 'future proofing' their investment in production equipment, which will, following wider use and volume production, become less expensive.

The interfaces and electrical connections should be straightforward as the 'input' is essentially the baseband decoded received signal, audio and metadata. The audio 'output' would be wired or wireless connections to the loudspeakers.

The user will have to select the option of Coordinate system, and if available in their equipment, the Point Source Panner. The user will need to input the details of the loudspeaker arrangement in use unless this is done automatically by the equipment. How much of the configuration is automated will depend on how the equipment is designed and implemented, but in principle operation should be relatively simple. Operator training and familiarisation equipment manuals will be needed.

The media production team will need training and practice in content production using AdvSS. It will open up new and exciting media experiences for sound producers, viewers and listeners, and challenge the creativity of content production teams. The media world will need to plan the migration to the new audio environment, coordinating the production of programmes with the availability of suitable reception systems. It may be appropriate, for example, to plan the migration based on major steps in delivery systems, such as combining it with the use of UHDTV.

The ADM Renderer specification was created to serve the media community, and to encourage and accelerate the use of AdvSS. It will serve a need for a standardised system for use in programme production and monitoring and, if needed, quality evaluations. It also will help equipment manufacturers with the design of renderers for consumer equipment.

The reproduction of AdvSS in the home may be expected to use a loudspeaker set up which applies loudspeaker positions from those specified in Recommendation ITU-R BS.2051. The loudspeaker

positions are input to the renderer to provide the correct inputs and co-ordinates to achieve the programme maker's / sound engineer's intent. Specially designed sound bars may function by creating faux loudspeaker positions.

FIGURE 1

Simplified architecture and user controls overview
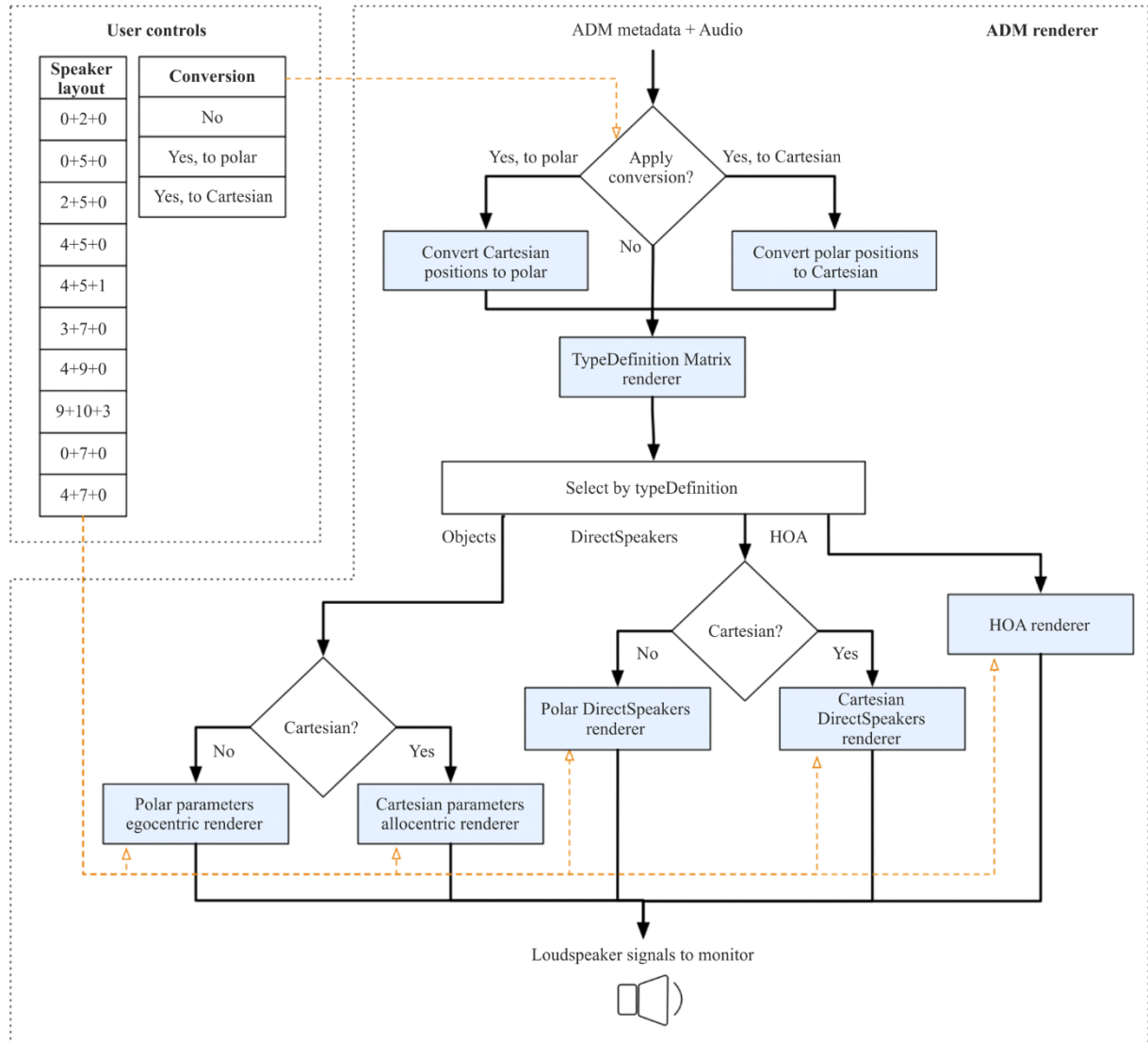
Report BS.2466-01

The supported target loudspeaker layouts include 0+2+0, 0+5+0, 2+5+0, 4+5+0, 4+5+1, 3+7+0, 4+9+0, 9+10+3, 0+7+0, and 4+7+0 systems (per Recommendation ITU-R BS.2051). The ADM input file may contain audio description for one DirectSpeakers layout, e.g. 4+9+0, while the specified target may be a different layout, e.g. 0+5+0. If the target layout includes fewer loudspeakers than the number of loudspeakers in the DirectSpeakers configuration, the ADM Renderer acts as a down-mixer. However, if the target layout includes more loudspeakers, the audio signal associated with each DirectSpeaker channel will be directed to the nearest target loudspeaker (i.e. no up-mixing of the audio content).

The ADM Renderer supports a number of optional parameters that allow the specification of, among other things, the real loudspeaker positions and screen dimensions of a reproduction setup.

The informative part of the specification of an alternative Point Source Panner provides a reference for an ego-centric renderer that results in certain desirable characteristics under some circumstances. For example, a limited leakage of the back elevated sound sources to the front loudspeakers, and front elevated sources with wider smooth transitions. The placement of these alternative virtual loudspeaker positions and their fold-down coefficients are based on by-ear optimizations and reflect the behaviour of some deployed renderers.

1.5 Allocentric frame of reference (FoR) and loudspeaker mapping

A reference frame is a means of representing the locations of entities in space. There are many ways to classify reference frames, but one fundamental consideration is the distinction between allocentric (or environmental) and egocentric (observer) frame of reference (FoR).

1   An allocentric frame of reference encodes object location using reference locations and directions based on other objects in the environment.

2   An egocentric frame of reference encodes object location relative to the position (location and orientation) of the observer or 'self'.

An egocentric reference is commonly used for the study and description of perception; the underlying physiological and neurological processes of acquisition and coding most directly relate to the egocentric reference. An allocentric reference is generally used for scene representations that are independent of a single observer position, and when the relationship between elements in the environment is of interest.

The allocentric coordinate system and its representation in Recommendation ITU-R BS.2076

The allocentric FoR uses Cartesian coordinates (x, y, z), shown in Fig. 2a, where the x-axis represents the 'left/right' dimension, the y-axis represents the 'front/rear' dimension, and the z axis the 'up/down' dimension. The coordinates (0, 0, 0) are located at the absolute centre of the allocentric rendering cube, shown in Fig. 2b, and the values on all axes range from -1 to 1, where $y = -1$ represents full rear or 'rear-most' location and $y = 1$ represents 'front-most'; $x = -1$ represents 'left-most', and $x = 1$ represents 'right-most'; and $z = -1$ represents 'down-most', and $z = 1$ represents 'up-most'.
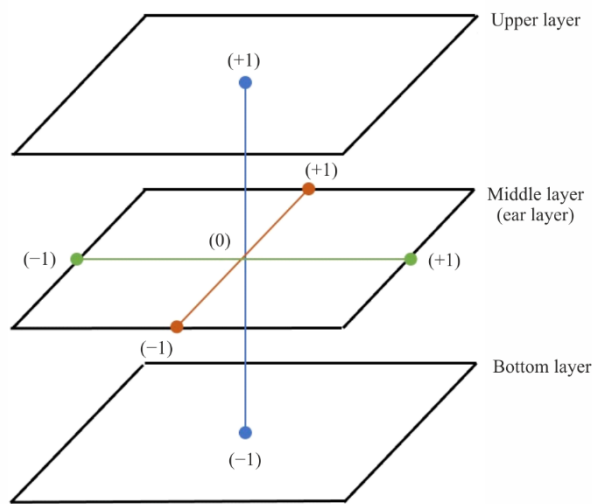
FIGURE 2

Allocentric coordinate system representation

a) Cartesian coordinate axes   b) Allocentric cube

Report BS.2466-02
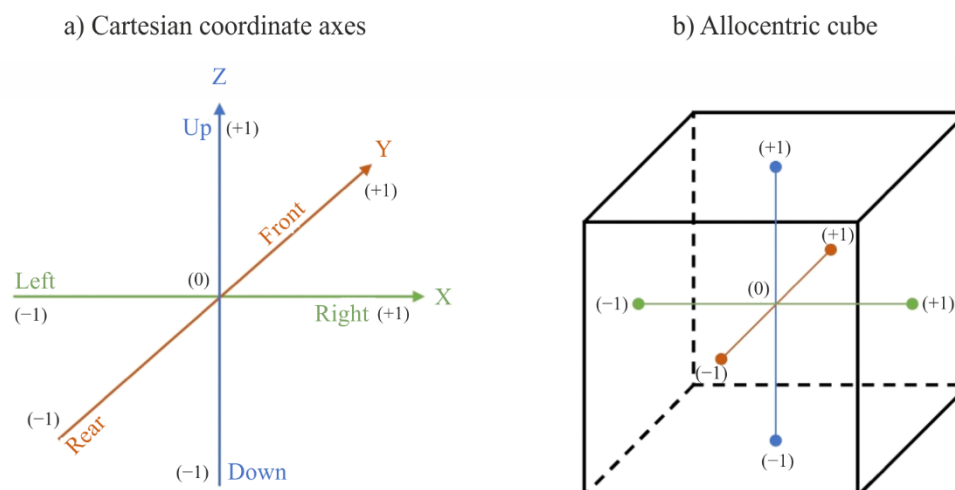
Recommendation ITU-R BS.2051 loudspeaker mapping in the allocentric cube

The FoR and the coordinate system described above is used to define sound source location. For an allocentric immersive sound format that utilizes both audio objects (sound location defined by coordinates) and audio channels (location defined by target loudspeaker), it is necessary to define loudspeaker locations within the object frame of reference in order to preserve the audio objects and audio channels relative positions. In fact, the loudspeaker locations provide the reference locations which define the allocentric FoR within the playback environment. Recommendation ITU-R BS.2051 defines loudspeakers over three layers: upper, middle and lower layers. In the allocentric rendering cube, loudspeakers are mapped to one of these layers where $z = -1$ for the 'down-most' or bottom layer, $z = 0$ for the middle layer (ear level), and $z = 1$ for the 'up-most' or upper layer, as shown in Fig. 3.
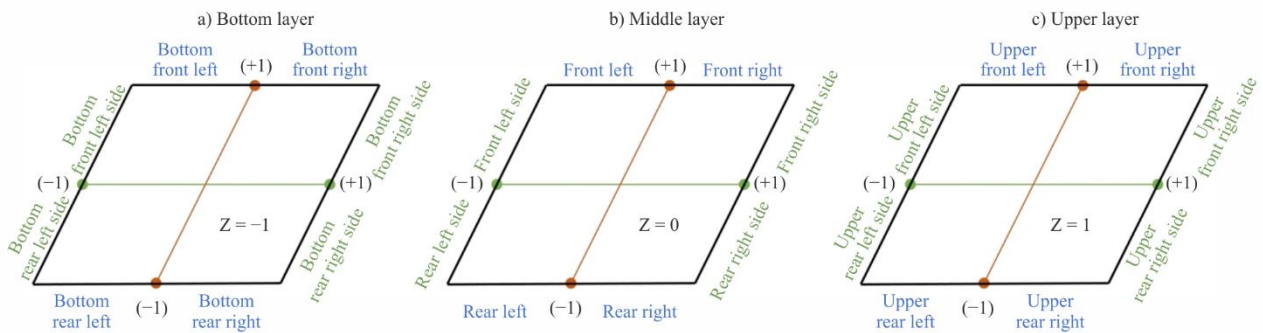
FIGURE 3

Allocentric layers

Report BS.2466-03

Figures 4a, 4b and 4c show the spatial intent of each layer. It is with this intent that loudspeakers for each system can be mapped. This is essential to understanding allocentric rendering and loudspeaker mapping.

FIGURE 4

Allocentric layers' spatial intent

Report BS.2466-04

The allocentric cube represents the intended sound scene, where audio elements are rendered within the spatial capabilities of each of the Recommendation ITU-R BS.2051 sound systems (A through J) and their associated loudspeakers.

Table 1 shows a first example of the mapping process for sound system C in Recommendation ITU-R BS.2051.

TABLE 1

Sound system C allocentric loudspeaker mapping

Reference location	SP Label	x	y	z
Centre location in front	M+000	0	1	0
Left-most location in front	M+030	1	1	0
Right-most location in front	M-030	−1	1	0
Left-most location in rear	M+110	1	−1	0
Right-most location in rear	M−110	−1	−1	0
Left-most location in upper-front	U+030	1	1	1
Right-most location in upper-front	U−030	−1	1	1

Following the same approach with sound system J results in the mapping shown in Table 2.

TABLE 2

Sound system J allocentric loudspeaker mapping

Reference location	SP Label	x	y	z
Centre location in front	M+000	0	1	0
Left-most location in front	M+030	1	1	0
Right-most location in front	M-030	−1	1	0
Side-Left location delimiting front from rear	M+90	1	0	0
Side-Right location delimiting front from rear	M−90	−1	0	0
Left-most location in rear	M+135	1	−1	0
Right-most location in rear	M−135	−1	−1	0
Left-most location in upper-front	U+045	1	1	1
Right-most location in upper-front	U−045	−1	1	1
Left-most location in upper-rear	U+135	1	−1	1
Right-most location in upper-rear	U−135	−1	−1	1

When comparing the two Tables, one will notice that loudspeakers M+110/M−110 in Table 1 have the same coordinates as loudspeakers M+135/M−135 in Table 2. This is fine since in Sound system J the Rear Left and Rear Right loudspeakers are the most rear loudspeakers, while in Sound system C, it is the Left Surround and the Right Surround loudspeakers that are the most rear loudspeakers.

The same allocentric mapping principle is applied to all sound systems and their respective loudspeakers. The resulting sets of Cartesian coordinates for each system is then integrated within the allocentric renderer.

The reference location's coordinates shown in Tables 1 and 2 can be helpful to understand how allocentric rendering works, but it is important to recall that the cartesian loudspeaker mapping described here are reference positions used for the sole purpose of creating an allocentric frame of reference within the renderer. These reference positions are internal to the renderer and do not directly correspond to the physical loudspeaker locations.

Recommendation ITU-R BS.2051 define the loudspeakers' physical *nominal* locations with allowable tolerance to enable the design and installation of compatible sound playback systems. The loudspeakers form the reference locations for the Cartesian-based allocentric FoR and the *actual* loudspeaker locations implicitly translate the positions within the allocentric FoR into the physical acoustic space. The cylindrical loudspeaker locations specified in Table 1 of Recommendation ITU-R BS.2051 are perfectly compatible with the Cartesian allocentric frame of reference for sound scene description. Independently from having multiple listeners or speakers outside nominal positions, the allocentric frame of reference preserve the content's spatial intent within the space defined by the loudspeaker boundaries, essentially acoustically warping the rendering cube to fit this space.

2 Current challenges, opportunities, and suggestions for the ITU ADM Renderer (IAR)

Opportunities

As previously explained, a 'Renderer', in the context of AdvSS is the software or hardware system that converts the available AdvSS/NGA audio elements (the combination of sounds and the metadata that define the apparent positions of the sounds in the eventual audience listening environment) to loudspeaker feeds.

The Renderer is the tool by which the programme producer can monitor and 'tune' the available audio elements, so that the audience with appropriate equipment experiences the 'sound-stage' that the programme producer would like them to have.

The loudspeaker configuration used needs to be fed into the Renderer, which adapts its output loudspeaker feeds to suit it. In this way, the programme producer's monitoring experience, via his or her renderer, can be like the audience's listening experience, whatever the user's loudspeaker configuration may eventually be.

The IAR was intended to be a common renderer that could be used by programme producers throughout the world and could be arranged to provide an experience like that of the listener, whatever home renderer, loudspeaker layout, and coordinate system were being used by the audience.

Renderers for home systems are and will be developed by the manufacturers of home listening equipment. For commercial reasons, they may include different proprietary elements that can make the audio experience more attractive and saleable to users, while still using standard forms of input signals. Thus, it is unlikely that a totally common home renderer system will be used throughout the world.

However, having a common renderer, the IAR, for monitoring and tuning AdvSS/NGA for programme production will be very valuable. It will allow programme makers to tailor and adapt their output to their market, while enjoying the economies of scale of a single worldwide standard and avoiding any drawbacks of a captive production equipment markets. Having a common renderer system for production will encourage and enable the growth of AdvSS/NGA. An absence of a common renderer could slow down the worldwide roll out of AdvSS/NGA. The IAR will also provide a useful benchmark and guide for home equipment manufacturers and enable subjective evaluation results to be fairly compared.

NHK have developed an IAR compliant renderer and an IAR based mixing console that support ADM for file-based production and S-ADM for live production. They have been used for trials in Japan (see Report ITU-R BS.2159).

Challenges

The content production and delivery environment need to operate both with pre-recorded and live content. The ADM Metadata format and Open Renderer software has been available for some years for processing pre-recorded content. Processing software is available for download and use.

Metadata formats that allow real time operation (S-ADM) has not been available until recently. Open software for the IAR using S-ADM is not yet available.

There is no commercially currently available renderer equipment based on the IAR in Europe. Trials of AdvSS/NGA in Europe have only been possible using one of the commercially available proprietary production renderers. While none of them is equipped with the ITU ADM renderer, one was considered sufficiently close to be used in 2020 for the monitoring of a live sport productions of the Roland Garros Tennis tournament with 3D ambience plus commentaries as sound objects.

A shortcoming of the basic arrangement was the absence of an S-ADM generator. A further company has however launched a software package to generate S-ADM. This has been used together with one of the available proprietary renderers. There were some inconveniences in doing so. It was found necessary to manually setup the renderer to monitor the result, and manually setup the generator to create the S-ADM stream.

Thus, at least in Europe, the current situation regarding the availability of IAR equipment does not exploit the potential value of the IAR. It is hoped that production equipment manufacturers can be encouraged to develop IAR equipment. This will encourage the use and take up of AdvSS/NGA. This will benefit the public and media producers.

Suggestions for future production tools that may include IAR

The IAR for production may be included in various tools. Suggested functionalities of these tools are listed below for the benefit of potential manufacturers:

- At least 16 audio PCM inputs, including 1 input metadata track (usually on channel 16) – Format: Audio over IP and/or MADI [3]in

- At least 24 audio PCM outputs – format: Audio over IP and/or MADI out

- Metadata Track authoring (read/write) and generation from pre-defined or user-registered metadata sets

- Authoring of multiple Metadata Presets (1 preset = 1 listening experience proposed by the sound engineer to the listener, i.e. English version, Stadium only, Audio described French Version, etc.)

- Live loudness measurements of each labelled audio element (i.e. 2.0, 5.1, 3D ambiences, commentaries, dialogues, etc.)

- Live loudness measurements of each preset

- Dedicated monitoring output to feed the monitoring loudspeakers with respect of loudspeaker layouts specified in Recommendation ITU-R BS.2051 (stereo, 5.1/0+5+0, 5.1+2/2+5+0, 5.1+4/4+5+0, 7.1+4/4+7+0, 22.2/9+10+3) with a dedicated AdvSS/NGA renderer

- Dedicated monitoring output for binaural services

---

[3] MADI or AES 10 interface format specified in Recommendation ITU-R BS.1873.

-           Fixed additional renderings: i.e. "I need a fixed stereo rendering of preset 1 to feed my legacy encoder"

3   How to use the software

3.1   General remarks

The ADM Renderer interprets the ADM metadata and the associated audio signals as prescribed in Recommendation ITU-R BS.2076 and generates the audio PCM waveforms for a specific loudspeaker system/layout or headphones feed. All reproduction systems/layouts specified in Recommendation ITU-R BS.2051 are supported.

The guidance of how an ADM metadata file should be generated in the first place by considering in particular which of its metadata should be used, how it should be used, and under what circumstances. It is envisaged that the choice and preparation of the ADM metadata, as well as the selection of its attributes/values, will be guided by a verification of how such metadata affects the sound generated by the ADM Renderer.

Consequently, tools would be expected that allow a content creator to "tune" the ADM metadata based on the output that the ADM Renderer generates. The ADM format provides for a rich set of audio-description features, not all of which may be utilized in every ADM content file. A choice of some metadata, e.g. selection of polar vs. Cartesian coordinates, may be guided by preferences outside the scope of this guide.

Given an ADM/BW64 file, the simplest way to generate the output audio signals is to call the ADM Renderer as follows:

```
itu_ADM_renderer -s target_system input_ADM_file.wav output_file.w
```

The ADM Renderer reference implementation renders ADM content from BW64 files according to the specification. It is written in the Python programming language and uses the numpy and scipy libraries for numerical computations.

These attributes make the reference implementation mostly unsuitable for use in situations where real-time or streaming rendering is required, or when the ADM Renderer is to be integrated into another tool.

Although in some circumstances the reference implementation may be used in production, the reference implementation primarily exists to aid in testing and standardization of the ADM Renderer.

In addition to the 'Python'[4] reference implementation, other ADM Renderer implementations are being developed (for example an implementation intended for real-time use, written in C++) and are expected to be available soon.

3.2   Implementation of renderers for use in production

When a renderer is required to run in real-time or be embedded within another system, new implementations of the standard will need to be developed to meet these requirements. These should be based on the text of the requirements, using the reference implementation for additional guidance on implementation, and for conformance testing.

Other implementations of the ADM Renderer do not have to use precisely the same methods and structure as the reference implementations. Of course, achieving the desired end result is the most important thing.

---

[4] Recommendation ITU-R BS.2127-0, § 1, includes the URL of the reference implementation.

Other implementations should be tested against the reference implementation. This can be done by:

– Testing that the output of the reference implementation and the new implementation are the same with the available test files. Note that many test files currently available were produced using systems which only support a small subset of ADM metadata, so will probably not result in full coverage of all cases.

– Testing that individual components of the systems result in the same (or close enough) outputs for the same inputs. For example, given the same position, the point-source panniers of two implementations should produce identical gains.

It is advisable to test both random cases (which are easy to generate and ensure a good level of base coverage), and cases designed to exercise all paths in both implementations (often called unit tests).

Test cases for the reference implementation are available in the test directories of the reference implementation. These test cases may be a good starting point, but if the methods used in the implementation being tested are significantly different, new test cases may need to be devised to test all paths.

In general, developers should take an adversarial approach to testing – tests should be designed to show that the systems are different if possible, rather than to show that they are the same in common cases.

3.3 Notes on ADM metadata

3.3.1 Size of audioBlockFormat

The position of audio objects in ADM metadata are described by two positions of start time and end time of the *audioBlockFormat*. Because the ADM Renderer does not interpolate position metadata, when the position of start time differs from that of end time, the output signals might be different depending on loudspeaker layouts. Therefore, the *audioBlockFormats* should be separated into small blocks, especially when the position of the audio object passes through typical loudspeaker's position.

FIGURE 5

Differences of played back loudspeakers depending on block size

Report BS.2466-05
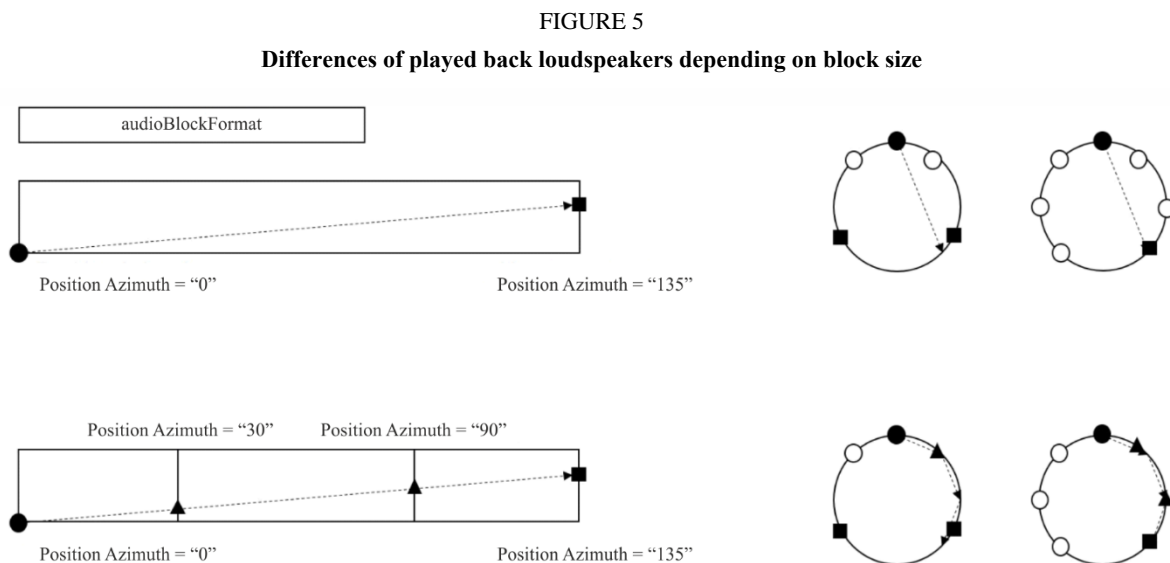
3.3.2   Conversion of position parameters between polar and cartesian coordinate systems

Section 10.1 of Recommendation ITU-R BS.2127 describes a method to convert position parameters between polar and cartesian coordinate systems. The conversion is based on the system D (4+5+0) layout, which was chosen primarily to ensure good conversion for content authored using 0+5+0.

However, there is one drawback to consider: the position of objects relative to SP Label M+135 and M−135 will be offset differently, depending on whether you convert from polar to cartesian, or from cartesian to polar coordinates.

The positions of cartesian objects authored on a system using M+135 and M−135 will be converted to positions further forward than their initially intended positions. Similarly, the positions of polar objects authored on a system using M+135 and M−135 will be converted to position further back than their initially intended positions.

As mentioned in § 10 of Recommendation ITU-R BS.2127-0, the results of conversion will not exactly match those without and should therefore be monitored. This is particularly true for content authored on systems using M+135 and M−135 where conversion back and forth between polar and Cartesian should be avoided.

3.3.3   Treatment of LFE channels

ADM can specify a frequency range of an *audioChannelFormat* using a *frequency* sub-element. If metadata indicating a *speakerLabel* which refers to an LFE channel is not present, IAR will assume it is an LFE channel if the *frequency* element is set to a cut-off frequency 200 Hz or less, as specified in § 8.2 of Recommendation ITU-R BS.2127-0. It should be noted that other Recommendations, such as ITU-R BS.775 and ITU-R BS.2051, specify the LFE cut-off frequency to be 120 Hz.

IAR does not apply any +10 dB offset (see Recommendation ITU-R BS.775) to the LFE channel to address playback calibration, with respect to the main channel, as this is done in the playback device.

FIGURE 6

Example of treatment of LFE channels
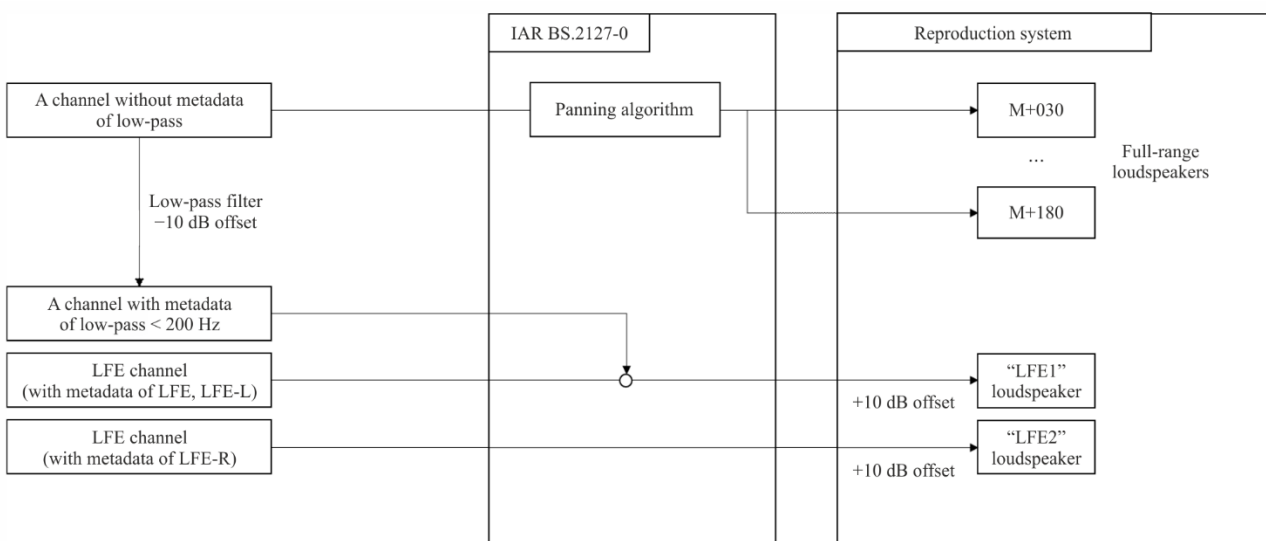
Report BS.2466-06

3.4   Interface with non-ADM systems

It should be possible to implement the ADM Renderer in systems like digital audio workstations, which may not support ADM as a native metadata format. Where possible the ADM Renderer should

still be used for metadata which might be turned into ADM, but may have to be adapted to work with non-ADM metadata formats. Where possible, this adaptation should be standardized to ensure interoperability with other systems.

In particular, the timing model of the ADM is quite different from the time-varying parameter models used in workstations, so more divergence is expected in, for example, gain interpolation than the calculation of gains themselves.

If live rendering is implemented in a product which supports ADM import or export, then the results of live rendering should be the same (or close as given above) in these situations:

– Content is rendered with the ADM Renderer reference software and imported then rendered live.

– Content is rendered live then exported and rendered using the ADM Renderer reference software.

This ensures that the system will still be interoperable with others which implement the ADM Renderer exactly.

3.5 Guidelines for subjective evaluations

Quality evaluation in the context of AdvSS is a new and more complex area. The designer or user of the AdvSS system may be interested in evaluating different facets of the reproduced sound. They might include 'sense of reality', 'realism', distortion, emotional responses, and other attributes.

Evaluations may serve to help design a system or help to make desired choices of parameter values or listening conditions. However, in all cases, the evaluation methodology used must be chosen to minimise unwanted bias, to be accurate, and to be reproduce-able. Test results from different evaluations cannot be fairly compared unless the same methodological conditions are used for all the tests. Achieving this parity of conditions in one of the reasons why the ADM Renderer should be used for subjective and objective quality evaluation.

The ADM Renderer should be used for appropriate kinds of quality evaluation where a reference renderer for ADM programme material is required or desired. This includes explicitly the usage of ADM content with subjective evaluation methodologies like Recommendations ITU-R BS.1534, ITU-R BS.1116 and ITU-R BS.1284, when rendering is required. It should be also applied for objective evaluation and measurements when ADM programme material is used.

While the ADM Renderer may serve as a benchmark/reference or an anchor renderer during the perceptual evaluation of other AdvSS systems, it should be noted that the term 'reference' does not preclude other renderers providing more desirable performance characteristics under certain circumstances over the ADM Renderer specified in Recommendation ITU-R BS.2127.

4 Explaining the purposes of the AdvSS/NGA

The ITU-R AdvSS/NGA system and the complementary emission systems (AC-4, DTS-UHD and MPEG-H 3D Audio) are intended to provide immersive sound system use with or without accompanying picture. They allow the content provider to offer a sound experience that can surround the user in three dimensions, is optimized by the content creator, and can be matched in the receiver to the home listening environment. This will be clearly an attractive option for future media consumers, most significantly with large screen viewing.

The AdvSS/NGA system is based on the delivery of a number of 'sound elements' to the users. These sound elements can be used for many purposes other than providing immersive sound.

Initial opinion soundings suggest also that an attractive early use of AdvSS would be to provide users with the ability to adjust the relative levels of 'sound elements', enabling for example further dialog intelligibility in a sound experience. There seems to a growing demand for such a facility. but it should be also recognized that allowing the user to reconfigure the sound balance should be a feature that is only allowed given the agreement of the content creator.

It may therefore be valuable to define AdvSS/NGA in terms of its dual role for personalisation and immersive sound. The personalisation aspects may prove 'saleable' to the public and industry initially.

5   Development of software implementations of the ADM Renderer

Until now the EBU ADM Renderer (EAR) specified in EBU Tech 3388, and the ITU ADM Renderer, were only available as a Python implementation. See previous section of this Report. This made it possible for developers to experiment with the capabilities of the renderer, but it also required original code to be developed from scratch for a practically useful implementation. The new 'libear', described below, intends to remove that need.

A first C++ library of the open EBU ADM Renderer (EAR) was released in 2019. However, at the time of writing this Report, work in all aspects of the new software implementations is not complete, and this should be born in mind when considering this section. The mentioned open-source software projects contain information about their current status.

The libear C++ library is available on GitHub. The software makes it much easier for creators of production solutions to implement AdvSS/NGA capabilities in their products.

The libear will natively interprets ADM metadata as specified in Recommendation ITU-R BS.2076. It is best used in conjunction with two related open-source C++ libraries, libbw64 and libadm, to complete the required set of functions for production solutions.

The ITU ADM Renderer plays a vital role in the object-based workflows given below.

–   Monitoring AdvSS/NGA (Next Generation Audio) productions throughout the whole workflow.

–   Pre-rendering AdvSS/NGA productions for legacy emission.

–   Pre-rendering parts of an AdvSS/NGA production if a reduction of complexity is necessary.

–   Using as the reference in AdvSS/NGA related listening tests.

Supporting all those applications, the aim is to make it as easy as possible to integrate the renderer into different applications. For real-time applications, a Python implementation is less than ideal; the core of the ITU ADM Renderer is implemented in C++.

The most complex part of the renderer implementation is the calculation of the gain values, which have to be applied to the input audio samples of each channel for channel-based audio, each audio object for object-based audio or each component for scene-based audio. Based on a set of metadata the calculation of the gain values can be done using the libear with a single function call.

Apart from the gain calculation, the library also provides basic digital signal processing functionality. This makes it easy to get started with the library, but depending on the targeted platform or application it can be useful to use a custom implementation, which is also possible. Another part of the reference implementation is the ability to process the ADM metadata and provide the right part of the metadata at the right time to the gain calculator. This is not part of this library, but can quite easily be implemented using the libadm.

The current release is a first version of the renderer library. It already supports channel-based, scene-based and object-based audio, but lacks support for some special parameters like screenEdgeLock or zoneExclusion.

The library will be useful in many different applications. For example, it could be wrapped into an audio plug-in for Digital Audio Workstations (DAW) or even more tightly integrated into a DAW. It could be integrated into a mixing desk monitoring section. With an ADM interface it could be efficiently integrated into a pre-render solution for a content management system. It could be integrated into a stand-alone monitoring system, which receives Serial ADM and renders in real time for different loudspeaker setups. There are many more applications.

The AdvSS system elements can also be used for a number of other services such as other languages, alternative commentaries, and services for those with reduced ability such as Audio Descriptions. These could also be an important part of the explanation of the use of AdvSS.