



ATSC

ADVANCED TELEVISION
SYSTEMS COMMITTEE

ATSC Digital Television Standard, Part 5 – AC-3 Audio System Characteristics

Doc. A/53 Part 5:2014
28 August 2014

Advanced Television Systems Committee
1776 K Street, N.W.
Washington, D.C. 20006
202-872-9160

The Advanced Television Systems Committee, Inc., is an international, non-profit organization developing voluntary standards for digital television. The ATSC member organizations represent the broadcast, broadcast equipment, motion picture, consumer electronics, computer, cable, satellite, and semiconductor industries.

Specifically, ATSC is working to coordinate television standards among different communications media focusing on digital television, interactive systems, and broadband multimedia communications. ATSC is also developing digital television implementation strategies and presenting educational seminars on the ATSC standards.

ATSC was formed in 1982 by the member organizations of the Joint Committee on InterSociety Coordination (JCIC): the Electronic Industries Association (EIA), the Institute of Electrical and Electronic Engineers (IEEE), the National Association of Broadcasters (NAB), the National Cable Telecommunications Association (NCTA), and the Society of Motion Picture and Television Engineers (SMPTE). Currently, there are approximately 120 members representing the broadcast, broadcast equipment, motion picture, consumer electronics, computer, cable, satellite, and semiconductor industries.

ATSC Digital TV Standards include digital high definition television (HDTV), standard definition television (SDTV), data broadcasting, multichannel surround-sound audio, and satellite direct-to-home broadcasting.

Note: The user's attention is called to the possibility that compliance with this standard may require use of an invention covered by patent rights. By publication of this standard, no position is taken with respect to the validity of this claim or of any patent rights in connection therewith. One or more patent holders have, however, filed a statement regarding the terms on which such patent holder(s) may be willing to grant a license under these rights to individuals or entities desiring to obtain such a license. Details may be obtained from the ATSC Secretary and the patent holder.

Revision History

A/53 Part 5:2007 – Initial version of A/53 Part 5 approved. Previously Part 6 was Annex B of A/53.	3 January 2007
A/53 Part 5:2010 – First revision of A/53 Part 5 approved.	6 July 2010
A/53 Part 5:2014 – Second revision of A/53 Part 5 approved.	28 August 2014

Table of Contents

1.	SCOPE	4
2.	REFERENCES	4
2.1	Normative References	4
2.2	Informative References	4
3.	COMPLIANCE NOTATION	4
3.1	Treatment of Syntactic Elements	5
3.2	Symbols, Abbreviations, and Mathematical Operators	5
4.	SYSTEM OVERVIEW (INFORMATIVE).....	5
5.	CONSTRAINTS ON AC-3 CODING	6
5.1	Constraints on AC-3 coding	6
5.2	Sampling Frequency	6
5.3	Bit Rate	7
5.4	Audio Coding Modes	7
5.5	Dialog Level	7
5.6	Dynamic Range Compression	7
5.6.1	Dynamic Range Compression – “Line mode”	7
5.6.2	Dynamic Range Compression – “RF Mode”	7
5.7	STD Audio Buffer Size	7
6.	INTRODUCTION TO AC-3 AUDIO SERVICES.....	8
6.1	Options for Providing Video Description	8
7.	CONSTRAINTS ASSOCIATED WITH SPECIFIC AUDIO SERVICE TYPES	8
7.1	Summary of Audio Service Types	8
7.1.1	Complete Main Audio Service (CM)	8
7.1.2	Visually Impaired (VI)	8
7.1.3	Hearing Impaired (HI)	9

Index of Figures

Figure 4.1 Audio subsystem in the digital television system.	6
---	----------

ATSC Digital Television Standard, Part 5: AC-3 Audio System Characteristics

1. SCOPE

This Part of A/53 establishes a set of constraints on A/52 [1] for, and normative provisions for, the coding of audio as AC-3.

2. REFERENCES

All standards are subject to revision. Users of this Standard are cautioned that newer editions might or might not be compatible.

2.1 Normative References

The following documents, in whole or in part, as referenced in this document, contain specific provisions that are to be followed strictly in order to implement a provision of this standard.

- [1] ATSC: “Digital Audio Compression (AC-3, E-AC-3),” Doc. A/52:2012, Advanced Television Systems Committee, Washington, D.C., 17 December 2012.
- [2] ITU: “Algorithms to measure audio programme loudness and true-peak audio level”, Recommendation ITU-R BS.1770-3, International Telecommunications Union, Geneva, 2012.

2.2 Informative References

The following documents contain information that may be helpful in applying this Standard.

- [3] ATSC: “ATSC Digital Television Standard, Part 3 – Service Multiplex and Transport Subsystem Characteristics,” Doc. A/53, Part 3:2013, Advanced Television Systems Committee, Washington, D.C., 7 August 2013.
- [4] AES: “AES Recommended practice for digital audio engineering—Serial transmission format for two-channel linearly represented digital audio data,” Doc. AES3-2009, Audio Engineering Society, New York, N.Y., 2009. (This document is now a 4 part document).
- [5] ANSI: “Specification for Sound Level Meters,” Doc. ANSI S1.4-1983 (R 2001) with Amd.S1.4A-1995, American National Standards Institute, Washington, D.C.
- [6] ATSC: “Digital Television Standard, Part 1 – Digital Television System,” Doc. A/53, Part 1:2013, Advanced Television Systems Committee, Washington, D.C., 7 August 2013.
- [7] ATSC: “Recommended Practice - Techniques for Establishing and Maintaining Audio Loudness for Digital Television,” Doc. A/85:2013, Advanced Television Systems Committee, Washington, D.C., 12 March 2013.
- [8] SMPTE: “ST 272:2004 Television – Formatting AES/EBU Audio and Auxiliary Data into Digital Video Ancillary Data Space,” Society of Motion Picture and Television Engineers, White Plains, NY.
- [9] SMPTE: “ST 299-1:2009 24-Bit Digital Audio Format for SMPTE 292 Bit-Serial Interface,” Society of Motion Picture and Television Engineers, White Plains, NY.

3. COMPLIANCE NOTATION

This section defines compliance terms for use by this document:

shall – This word indicates specific provisions that are to be followed strictly (no deviation is permitted).

shall not – This phrase indicates specific provisions that are absolutely prohibited.

should – This word indicates that a certain course of action is preferred but not necessarily required.

should not – This phrase means a certain possibility or course of action is undesirable but not prohibited.

3.1 Treatment of Syntactic Elements

This document contains symbolic references to syntactic elements used in the audio, video, and transport coding subsystems. These references are typographically distinguished by the use of a different font (e.g., *restricted*), may contain the underscore character (e.g., `sequence_end_code`) and may consist of character strings that are not English words (e.g., `dynrng`).

The word “dialog” may be spelled as “dialogue” in other ATSC documents. The difference in spelling does not indicate any difference in meaning.

3.2 Symbols, Abbreviations, and Mathematical Operators

The symbols, terms, and mathematical operators used herein are as found in Section 3.4 of ATSC A/53 Part 1 [6].

4. SYSTEM OVERVIEW (INFORMATIVE)

As illustrated in Figure 4.1, the audio subsystem comprises the audio encoding/decoding function and resides between the audio inputs/outputs and the transport subsystem. The audio encoder(s) is (are) responsible for generating the audio elementary stream(s) which are encoded representations of the baseband audio input signals. At the receiver, the audio subsystem is responsible for decoding the audio elementary stream(s) back into baseband audio. The elements of the system covered by this part are indicated in Figure 4.1.

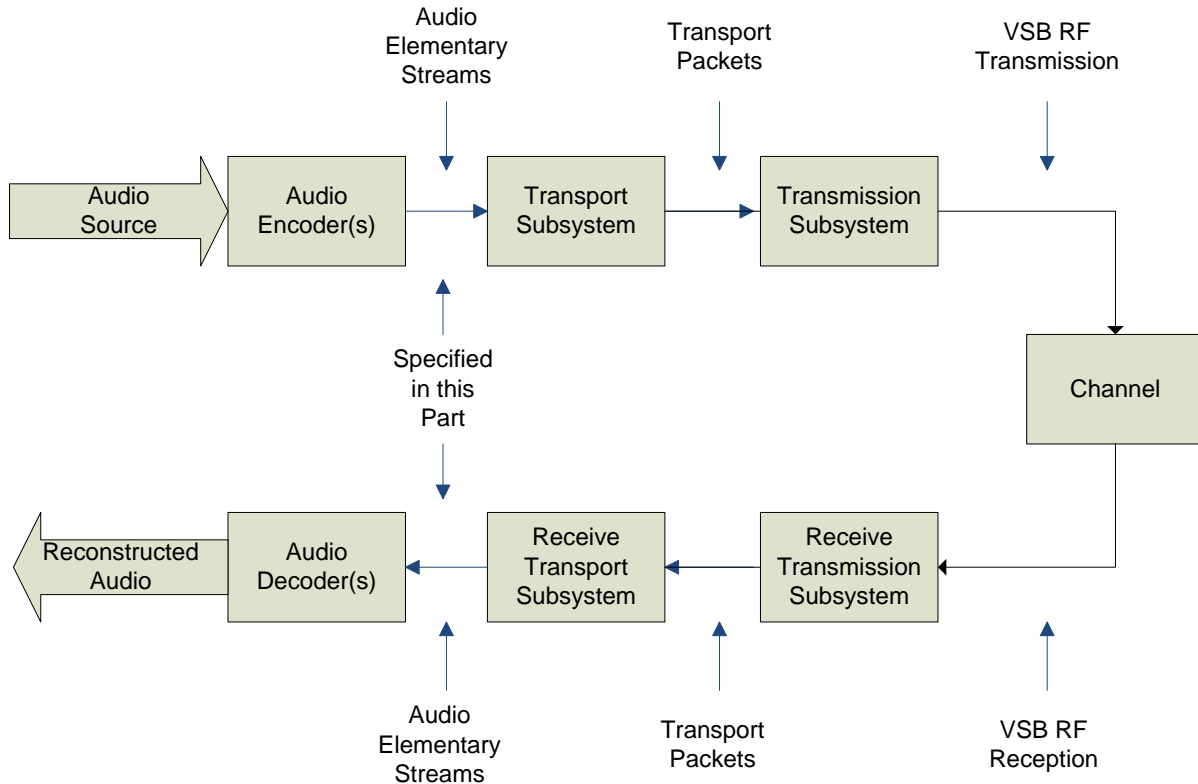


Figure 4.1 Audio subsystem in the digital television system.

5. CONSTRAINTS ON AC-3 CODING

This Section comprises the normative constraints for Elementary Stream (ES) components carrying audio encoded with the AC-3 codec. The coding of AC-3 audio streams shall conform to the “Digital Audio Compression (AC-3, E-AC-3) Standard,” A/52 [1], subject to the constraints outlined in this document.

5.1 Constraints on AC-3 coding

The AC-3 audio elementary streams shall comply with the syntax and semantics as specified in ATSC Doc. A/52 [1] (excluding the non-normative annexes and Annexes E, F, and G). Constraints on the coding of AC-3 syntactical elements shall be as constrained hereinafter.

AC-3 streams shall be constrained as follows:

- The value of *fscod* shall be ‘00’ (which indicates an audio sampling rate of 48 kHz).
- The value of *bsid* shall be ‘01000’ (which indicates the *stream_type* is 0x81).
- The value of *acmod* shall be \geq ‘001’ (which excludes 1+1 audio coding mode).
- The maximum data rate of an AC-3 main audio service stream shall be less than or equal to 448 kb/s.

5.2 Sampling Frequency

The system conveys digital audio sampled at a frequency of 48 kHz, which shall be locked to the 27 MHz system clock. The 48 kHz audio sampling clock is defined as:

$$48 \text{ kHz audio sample rate} = (2 \div 1125) \times (27 \text{ MHz system clock})$$

If analog signal inputs are employed, the A/D converters should sample at 48 kHz. If digital inputs are employed, the input sampling rate shall be 48 kHz, or the audio encoder shall contain sampling rate converters that convert the sampling rate to 48 kHz. Digital inputs may be either AES3 [4] signals on a cable or embedded in the corresponding video per SMPTE ST 272 [8] or SMPTE ST 299-1 [9].

5.3 Bit Rate

A main audio service, or an associated audio service which is a complete service (containing all necessary program elements) shall be encoded at a bit rate less than or equal to 448 kbps.

5.4 Audio Coding Modes

Audio services shall be encoded using any of the audio coding modes specified in A/52 [1], with the exception of the 1+1 mode. The value of *acmod* in the AC-3 bit stream shall have a value in the range of 1–7, with the value 0 prohibited.

5.5 Dialog Level

The value of the *dialnorm* parameter in the AC-3 elementary bit stream shall indicate the loudness¹ of the encoded audio content (typically of the average spoken dialog) using LKFS units. LKFS and its loudness measurement algorithm are specified in ITU-R Recommendation BS.1770, Annex 1 [2]². (Receivers use the value of *dialnorm* to adjust the reproduced audio level to normalize the loudness.)

5.6 Dynamic Range Compression

Dynamic Range Compression (DRC) is an important feature of the AC-3 codec and is documented at length in ATSC A/85 [7] Section 9 and Annex F.

5.6.1 Dynamic Range Compression – “Line mode”

Each encoded audio block contains a dynamic range control word (*dynrng*) that is used by decoders (by default) to alter the level of the reproduced audio. The control words allow the decoded signal level to be increased or decreased by up to 24 dB.

5.6.2 Dynamic Range Compression – “RF Mode”

Each encoded audio frame contains a dynamic range control word (*compr*) that may be optionally used by decoders to render the audio with a very narrow dynamic range. The control words allow the decoded signal level to be increased or decreased by up to 48 dB.

5.7 STD Audio Buffer Size

The main audio buffer (*BS_n*, see A/52 Annex A [1]) shall be 2592 bytes.

¹ Methods to measure loudness and different modes for *dialnorm* management are explained in the ATSC Recommended Practice A/85 “Techniques for Establishing and Maintaining Audio Loudness for Digital Television” [7]; see particularly Section 5.2.

² A *dialnorm* value based on an “A” weighted integrated measurement (LAeq) (ANSI S1.4) [5] prior to the publication of this document need not be re-measured with BS.1770 [2].

6. INTRODUCTION TO AC-3 AUDIO SERVICES

As defined in A/52 [1], an AC-3 elementary stream contains the encoded representation of a single audio service. Multiple AC-3 elementary streams may be provided. There is a *complete main service* and there are various types of *associated services*.

The use of associated services Music and Effects (ME), Dialog (D), Voice Over (VO), Commentary (C), and Emergency (E) was deprecated in 2010.³

Associated services shall contain complete program mixes containing all audio program elements (dialog, music, effects, etc.) that are intended to be presented to a listener. This is indicated by the `full_svc` bit in the `AC-3_audio_stream_descriptor()` being set to a value of '1' (see A/52, Annex A [1]).

6.1 Options for Providing Video Description

The term "Video Description" is used for an audio service intended to improve the experience of people who are visually impaired, such as the insertion of audio narrated descriptions of a television program's key visual elements into natural pauses between the program's dialog. The term is used in statute and regulation. The abbreviation "VI" has been retained, as that is the label for this value of the `bsmod` field in A/52 [1].

7. CONSTRAINTS ASSOCIATED WITH SPECIFIC AUDIO SERVICE TYPES

This section provides constraints on specific audio service types when used in the ATSC system. Additional audio service types are defined in A/52 [1].

7.1 Summary of Audio Service Types

The following subsections document the constraints for adding an AC-3 audio elementary stream as a component of a digital television service having `service_type` values 2 or 3 (`service_type` is defined by A/53 Part 1 [6]). The definition of, and signaling of, such audio elementary stream components are found in A/53 Part 3 [3].

Note: Services having `service_type` values 2 or 3 are required by A/53 Part 3 [3] to contain at least one Complete Main (CM) audio service. Multiple CM audio services may be signaled.

7.1.1 Complete Main Audio Service (CM)

The CM type of audio service is signaled by using `bsmod` value '000' as described in A/52 [1]. The CM type of audio service shall contain a complete audio program (which typically includes dialog, music, silence, and effects). The CM service may contain from 1 to 5.1 audio channels. Audio in multiple languages may be provided by supplying multiple CM services, each in a different language.

7.1.2 Visually Impaired (VI)

The VI type of audio service is signaled by using `bsmod` value '010' as described in A/52 [1]. The VI type of audio service⁴ shall be constrained to be a complete program mix (which typically

³ Associated services that contain only a single program element were intended to be combined with the program elements from a main audio service by a receiver with two audio decoders and a mixer; such receivers were anticipated, but never materialized.

⁴ The term "Video Description" is sometimes used to describe such services.

includes dialog, music, silence, and effects) containing aspects designed to improve the experience of the visually impaired, such as the insertion of audio narrated descriptions of a television program's key visual elements into natural pauses between the program's dialog. The VI service may be coded using any number of channels (up to 5.1).

7.1.3 Hearing Impaired (HI)

The HI type of audio service is signaled by using bsmod value '011' as described in A/52 [1]. The HI associated service shall be a program mix (which typically includes dialog, music, silence, and effects) with enhanced intelligibility. The HI service may be coded using any number of channels (up to 5.1).

End of Part 5