SMPTE REGISTERED DISCLOSURE DOCUMENT

RDD 6-2008

for Television — Description and Guide to the Use of the Dolby® E Audio Metadata Serial Bitstream

The attached document is a Registered Disclosure Document prepared by the proponent identified below. It has been examined by the appropriate SMPTE Technology Committee and is believed to contain adequate information to satisfy the objectives defined in the Scope, and to be technically consistent.

This document is NOT a Standard, Recommended Practice or Engineering Guideline, and does NOT imply a finding or representation of the Society.

Errors in this document should be reported to the proponent identified below, with a copy to eng@smpte.org.

All other inquiries in respect of this document, including inquiries as to intellectual property requirements that may be attached to use of the disclosed technology, should be addressed to the proponent identified below.

Every attempt has been made to ensure that the information contained in this document is accurate. Dolby Laboratories is, however, not responsible for the action, effect, consequences, etc. of any non-Dolby implementation of any device(s) based on this document. If any errors, inconsistencies, or any points that require clarification are found, contact Dolby Laboratories at the address below as soon as possible.

Proponent contact information:

Dolby Laboratories Inc.
San Francisco Broadcast Group
100 Potrero Avenue
San Francisco CA 94103-4813

tvaudio@dolby.com

Copyright © 2008 by THE SOCIETY OF
MOTION PICTURE AND TELEVISION ENGINEERS
3 Barker Avenue, White Plains, NY 10601
(914) 761-1100

Approved March 12, 2008
Corrections to Tables 7-12, 7-13 and 7-14 made on March 6, 2009

Table of Contents

1 Scope

Dolby Laboratories has developed a serial communication protocol that is used to transfer audio metadata between various products manufactured by Dolby. This document describes the types and function of the various metadata fields carried in the serial bitstream so that others interested in communicating with various Dolby® E and Dolby Digital audio codecs and associated equipment may do so.

Each metadata element is classified as either a Professional or Consumer type of metadata. The Reserved elements are meaningless to the user but must not be changed, and must be passed on to preserve the structure and timing of the bitstream. Each of the Professional or Consumer elements is further classified as either a Read Only or as a Read & Write element. Its function is described, and permissible operations on those elements are discussed.

2 Related Documents

TIA-485-A 1998, Electrical Characteristics of Generators and Receivers for use in Balanced Digital Multipoint Systems

SMPTE 318M-1999, Television and Audio — Synchronization of 59.94- or 50-Hz Related Video and Audio Systems in Analog and Digital Areas — Reference Signals

SMPTE RP 168-2002, Definition of Vertical Interval Switching Point for Synchronous Video Switching

3 Introduction

Dolby Digital and Dolby E are both data-rate reduction technologies that use metadata. Metadata can be multiplexed into the Dolby Digital or Dolby E bitstream, to describe the encoded audio and convey information that precisely controls downstream encoders and decoders. Metadata can also be carried separately as a serial data stream between Dolby E and/or Dolby Digital equipment, or third party equipment.

Dolby Digital is a *transmission bitstream* (sometimes called an *emission bitstream*) intended for delivery to the consumer at home through a medium such as DTV or DVD. It consists of a single encoded program of up to six channels of audio described by one stream of Consumer metadata. The consumer's Dolby Digital decoder reproduces the program audio according to the metadata values set by the program creator, and according to settings for speaker configuration, bass management, and dynamic range that are chosen by the consumer to match his specific home theater equipment and environmental conditions.

Dolby E is a Professional audio encoding technology capable of carrying up to eight channels of encoded audio and metadata, used primarily for broadcast applications, such as program contribution and distribution. Each program is discrete, with its own Consumer metadata multiplexed into the Dolby E stream. The Dolby E bitstream also carries a set of Professional metadata elements, used to describe the programs being carried, some other items useful in the Professional domain and to identify the Consumer metadata segments. Some additional metadata elements, marked as "Reserved", are used to automatically configure a Dolby E decoder or a Dolby Digital encoder, maintain bitstream format, and so forth. These elements do not carry any applicable Consumer or Professional information, but must be passed on unaltered to the next device in the chain if the system is to function correctly. The description of these Reserved elements is limited to their size, so the rest of the stream can be parsed.

Metadata is first inserted during program creation or mastering, and is carried through the contribution and distribution stages of a broadcast plant, where it leaves the Dolby E environment and is transferred to the Dolby Digital encoder for transmission to the home, where it controls the Dolby Digital decoder. The same ideas apply to the metadata accompanying the Dolby Digital information on a DVD, or for other applications.

This document describes the protocol, timing, and other essential characteristics of the serial bitstream used to transfer audio metadata between various Dolby manufactured products. This information will allow users of the metadata to locate, recover, and interpret the audio metadata information carried in the serial bitstream.

Each of the Metadata elements is classified as either a Professional or Consumer data type. Professional metadata is used primarily by Dolby E systems to control or describe the “E” stream itself (program configuration, channel gain, etc.). Professional metadata will rarely, if ever, be passed on to consumer devices. The Consumer metadata, on the other hand, is generally used to control consumer decoders, so includes items like dynamic range control. Some is classified as Read Only, as it provides useful information, but should not be changed. Other items may be read and modified, so are labeled as Read and Write, with rules to follow if they are modified.

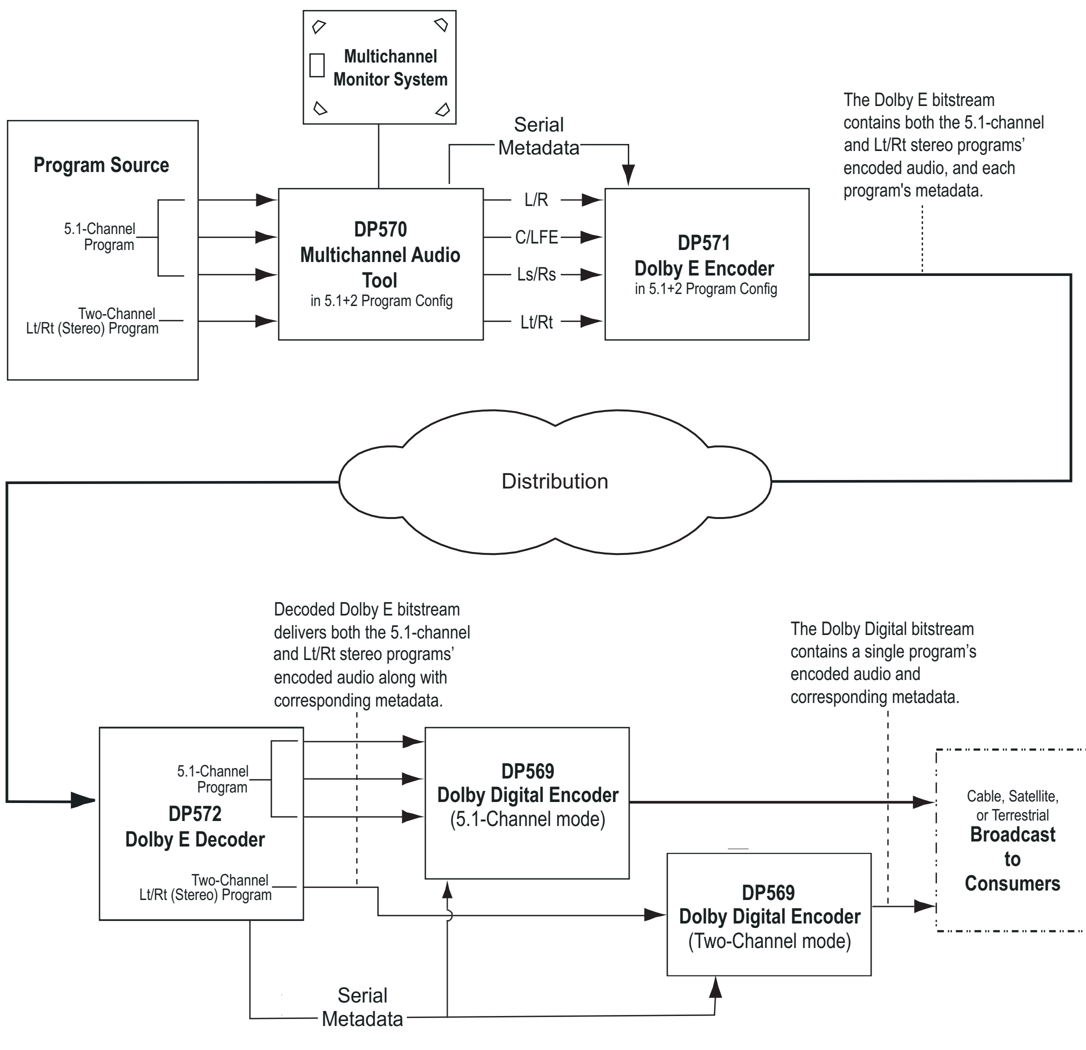


Figure 3-1 – Metadata Flow from Production to Consumer

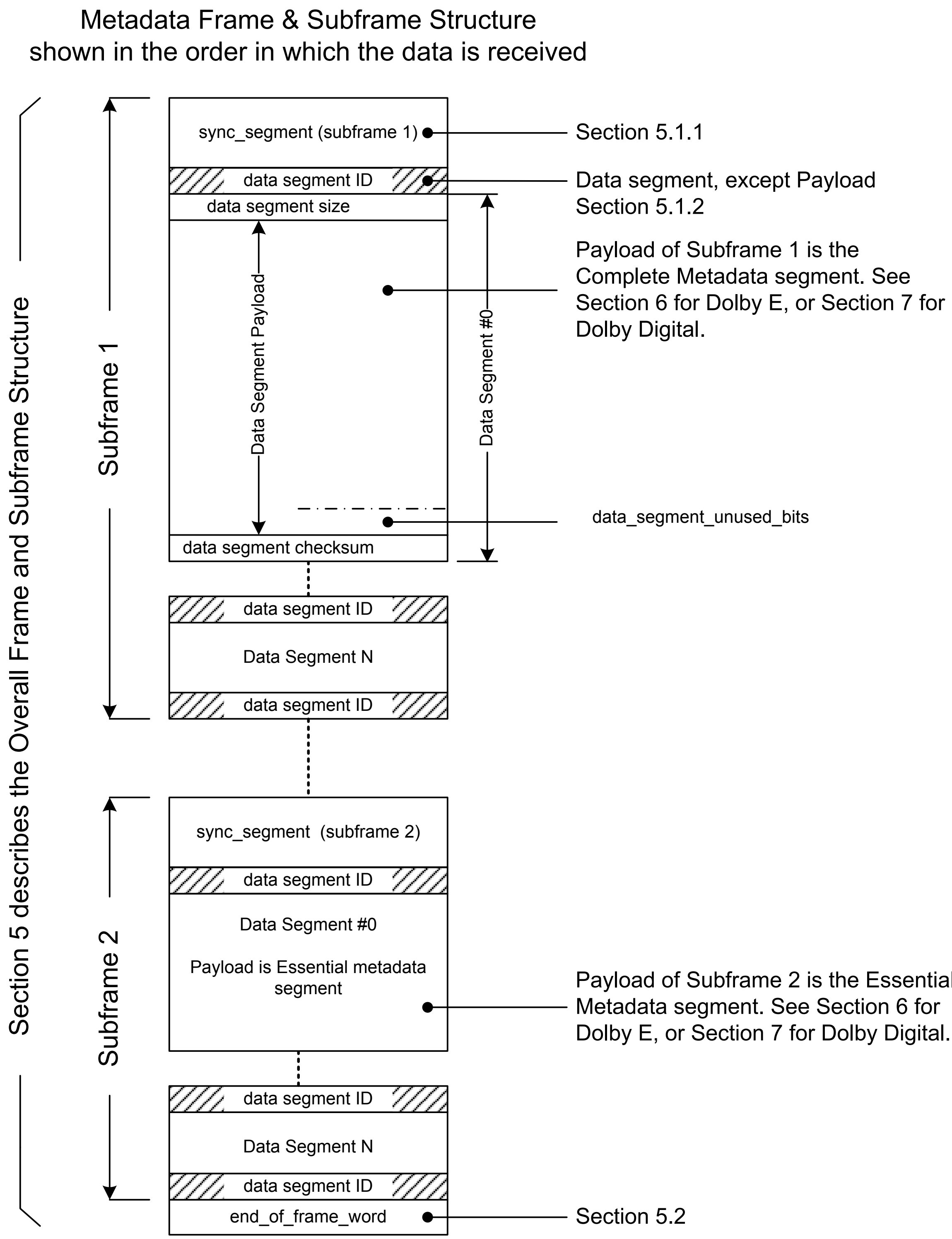


Figure 3-2 – Metadata Subframe Field Content Descriptions

4 Serial Metadata Stream Structure and Timing on a Serial Data Link

Audio metadata on a serial data link shall be transmitted as an intermittent, asynchronous serial data stream running at 115,200 baud, with 1 Start bit, 8 data bits, 1 Stop bit and no parity bit, at TIA/EIA-485-A (formerly RS-485) electrical levels. Bursts of serial data occur at the related 24-, 25- or 30-Hz video frame rates, and shall be synchronized with the associated video signal as shown in Figure 4-2. The transmission of the first and second metadata subframes shall be timed to avoid the areas around the vertical interval switch points of the associated video signal to avoid corrupting the metadata stream when it is switched synchronously with the video signal. The duration and timing of these forbidden intervals allows some tolerance in the relative timing of the video and metadata streams. This is detailed in Secton 4.2.

4.1 Structure on an Asynchronous Serial Link

Each metadata burst or frame shall carry the audio metadata information related to a specific group of audio samples which are, in turn, related in time to a specific video frame or pair of frames. (Figure 4-2 shows the relationship between metadata, interlaced and progressive frames). Metadata frames shall be subdivided into a first and second subframe, each of which shall be prefixed by a specific sync sequence. The number of bytes in each subframe depends on the number and configuration of the associated rate reduced audio programs.

Figure 4-1 shows the general structure of the serial Metadata stream. The duration of the subframes depends on the number and format of the associated audio programs, but the repetition period of the complete metadata frame shall be equal to the frame period of the associated interlaced video signal operating at or below 30 Hz, or to two frame periods of a progressively scanned video signal operating at greater than 30 Hz. The top part of Figure 4-1 shows the structure of the serial Metadata stream when the duration of Metadata Subframe 1 is less than one video field (or one Progressive frame). The bottom part of Figure 4-1 shows the structure when the duration of Metadata Subframe 1 is greater than one video field (or one Progressive frame).

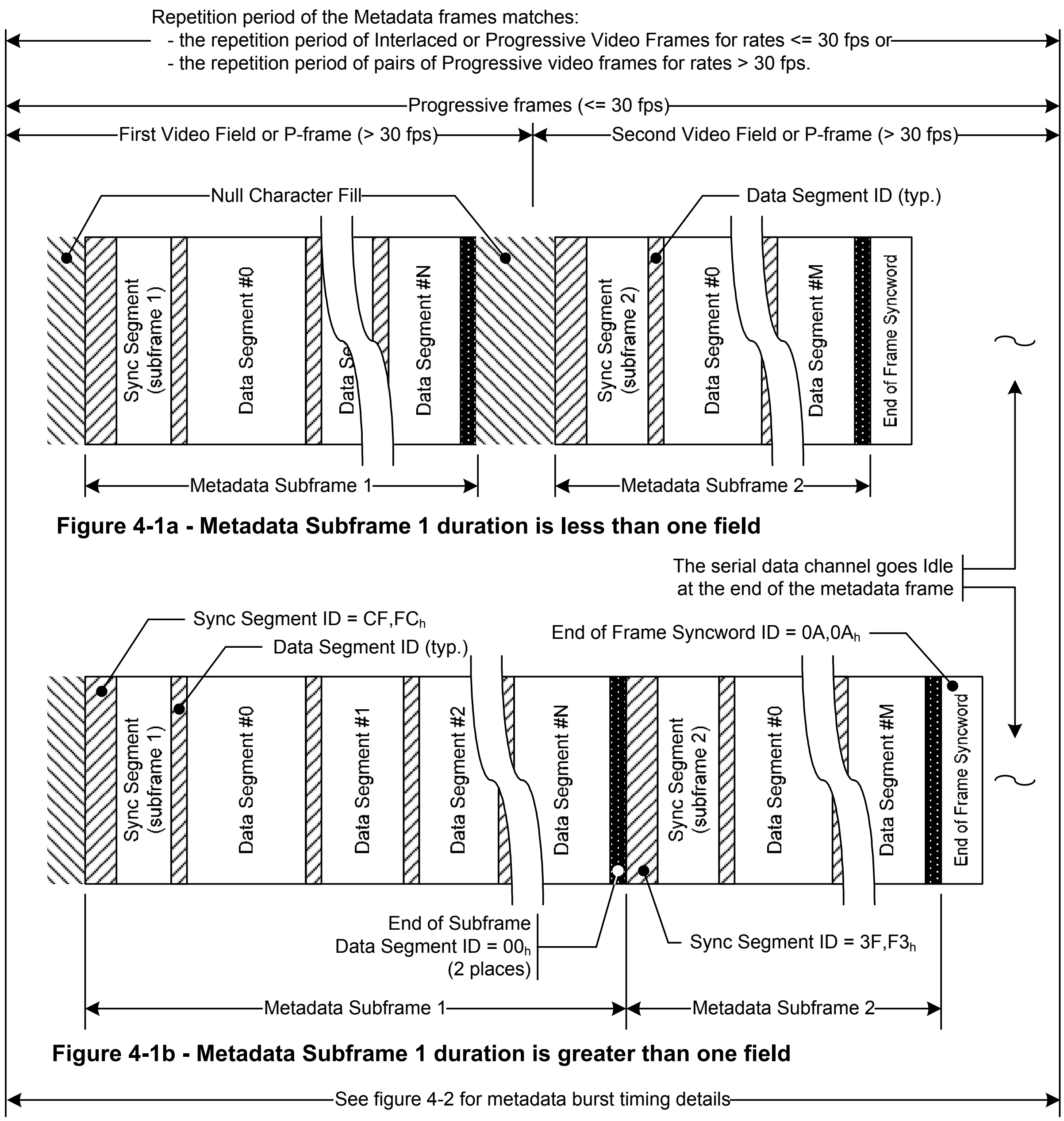


Figure 4-1 – Structure of a Metadata Frame in the Asynchronous Serial Stream

4.2 Metadata Frame Timing Requirements

The serial metadata stream is broken into two subframes, each of which is timed with respect to the associated video signal to avoid having a vertical interval switch point fall within either subframe.

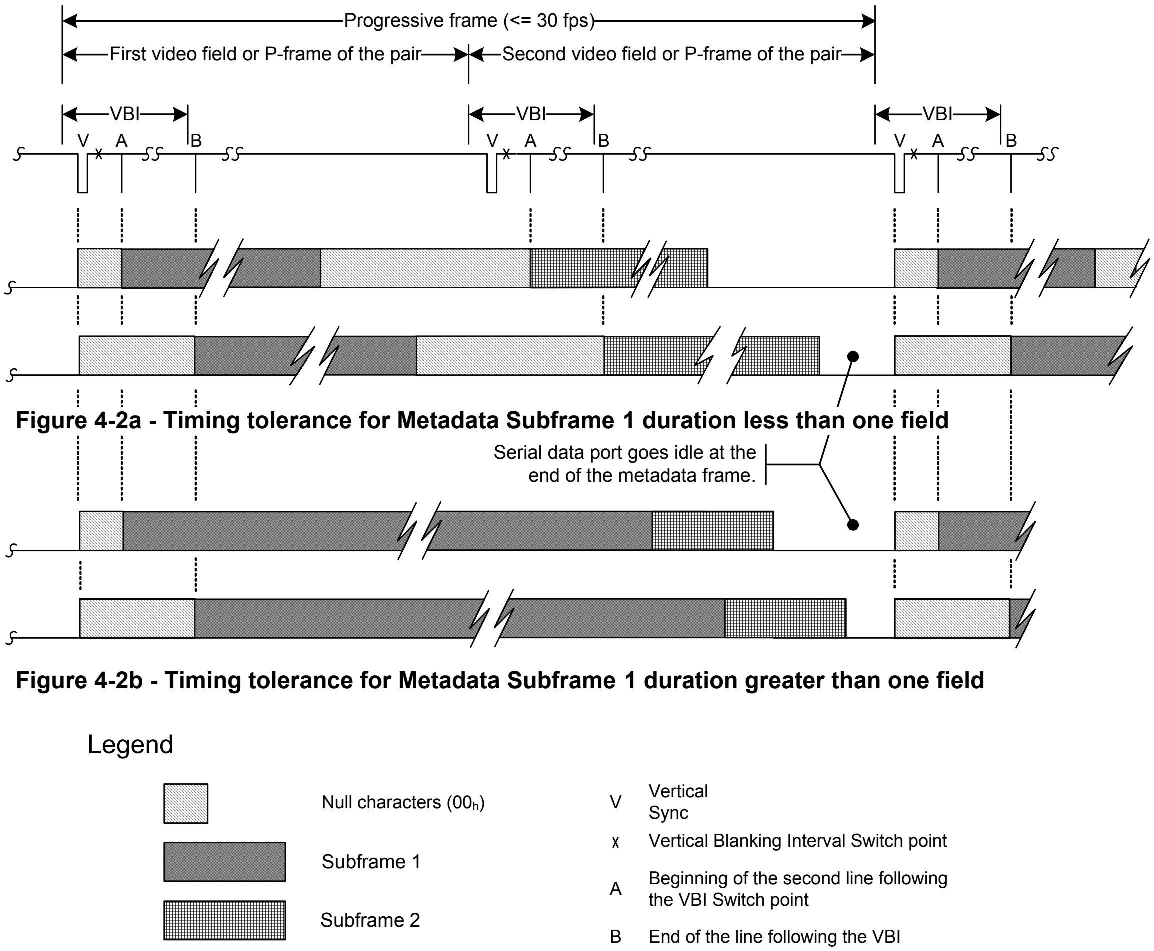


Figure 4-2a - Timing tolerance for Metadata Subframe 1 duration less than one field

Figure 4-2b - Timing tolerance for Metadata Subframe 1 duration greater than one field

Figure 4-2 – Serial Metadata Stream Timing

Metadata subframe 1 shall begin no earlier than the beginning of the second line following the vertical interval switch point and no later than the end of the line following the vertical blanking interval of the associated video signal (see Figures 4-2a and b).

If metadata subframe 1 ends before the usual start point of metadata subframe 2, then Null characters (00_h) shall be transmitted until metadata subframe 2 begins.

Metadata subframe 2 shall begin no earlier than the beginning of the second line following the vertical interval switch point and no later than the end of the line following the vertical blanking interval of the associated video signal, as shown in Figures 4-2a and b, unless the duration of subframe 1 is greater than that of video Field 1 or the first P frame, in which case subframe 2 shall begin when subframe 1 is finished, as shown in the lower parts of Figure 4-2 a and b.

The asynchronous serial data channel shall go idle at the end of the second metadata subframe.

The signal alignment and Vertical Interval switching point relationships between both interlaced or progressively scanned 1125, 750, 625 and 525-line systems are defined by SMPTE RP 168 (Definition of Vertical Interval Switching Point for Synchronous Video Switching).

For progressive video formats operating at rates greater than 30 frames per second, the first video frame of each pair of progressively scanned frames is the frame that is most closely aligned to the first field of the interlaced reference signal. All the timing details shall be referenced to the first frame of the progressive pair.

Note: Most television facility infrastructure is unaware of the frame pair alignment for progressive video formats operating at rates greater than 30 frames per second. As a consequence, switching video may result in a disruption to the metadata subframe structure.

All text and diagrams in this section apply equally to video systems operating at integer and {integer/1.001} frame rates. The basic rate and timing of the recovered serial metadata stream shall match that of the associated video signal.

The serial metadata port should[1] output a series of Null (00_h) characters coincidentally with the Vertical sync period of the first field (or first progressive frame) and continue until the beginning of metadata subframe 1. All receivers shall accept serial metadata streams with or without Null characters prepended to metadata subframe 1.

[1] NOTE that many serial metadata streams begin with a series of Null characters. These begin coincidentally with the vertical sync pulse of the associated video signal and continue until the first metadata subframe begins. These Null characters have been used as an aid in timing the beginning of the metadata stream, as a run in sequence for the metadata stream, and can be used as an "extended sync word" to help avoid misdetection of the subframe 1 sync_segment.

5 Metadata Frame Structure

The following pseudo code details the data fields and the order of arrival of information in a metadata frame. This pseudo code is roughly based on the C language syntax, but is simplified for ease of reading. For fields that are larger than 1 bit, the order of the bits is always MSB first (or MSB at the left, and LSB to the right).

Table 5-1 – Metadata Frame Syntax

Syntax	word size (bits)
metadata_frame()	
{	
metadata_subframe() /* **first metadata subframe */**	
metadata_subframe() /* **second metadata subframe */**	
end_of_frame_sync_word..	16
} /*end of Metadata frame */	

5.1 Metadata Subframe Field Descriptions

The first metadata subframe begins with a unique sync segment, followed by a number of data segments, as shown in Table 5-2.

Table 5-2 – Subframe Syntax

Syntax	word size (bits)
metadata_subframe()	
{	
sync_segment()	
while(1)	
{	
data_segment_id..	8
if(data_segment_id==0)	
{	
break;	
}	
else	
{	
data_segment()	
}	
}	
} /*end of metadata subframe */	

5.1.1 sync_segment Field Descriptions

Table 5-3 – Sync Segment Syntax

Syntax	word size (bits)
sync_segment()	
{	
start_of_subframe_sync_word	16
revision_id	4
originator_id	8
originator_address	16
frame_count	16
Reserved	12
}	

start_of_subframe_sync_word **Professional** **Read Only**

Word size: 16 bits
Valid range: See Table 5-4

The start_of_subframe_sync_word is a 16-bit field used to indicate the beginning of a metadata subframe, allowing devices to easily synchronize to any of the subframes in the metadata bitstream. The sync word for the first metadata subframe also indicates the beginning of the metadata frame, as shown in Table 5-4.

Table 5-4 – Valid Start of Subframe Sync Words

Subframe sync_words	Value
Sync_word (first subframe)	0xCFFC
Sync_word (second subframe)	0x3FF3

revision_id **Professional** **Read Only**

Word size: 4 bits
Valid range: 0-15

The revision_id field indicates the revision level of the metadata subframe syntax. If a metadata receiver is presented with a revision id higher than it recognizes, it should abandon further parsing of the bitstream for the remaining metadata subframe. For the current version of the metadata frame syntax, this field must be set to 0.

originator_id **Professional** **Read and Write**

Word size: 8 bits
Valid range: See Table 5-5

The originator_id field (Table 5-5) may identify the manufacturer of the of the device (or software) that created the metadata frame. This allows manufacturer specific data to be carried in the following “originator_address” field, when this field is not set to 0. If the originator_id field is set to 0, then the originator_address field must also be set to 0.

The originator_id field values are maintained for Dolby by the SMPTE Registration Authority. Contact Dolby at tvaudio@dolby.com to request a new **originator_id** value.

(NOTE – Table 5-5 is up to date to 2007 08 22)

Table 5-5 – Allowable Values for originator_id

originator_id	Manufacturer identity
0	unspecified
1 -32	Dolby Laboratories
33- 42	Harris Broadcast Communications (Leitch)
43 - 255	Reserved for future registrations

originator_address **Professional** **Read and Write**

Word size: 16

Valid range: All values (see below)

The originator_address field has been re-defined to allow metadata suppliers to include what amounts to private data in the metadata stream.

If the metadata supplier has placed a valid registered originator_id in the originator_id field, then the originator_address field may carry any desired data .

If the originator_id field is set to 0, then the originator_address field must also be set to 0.

frame_count **Professional** **Read and Write**

Word size: 16

Valid range: All possible values

The frame_count field is an unsigned integer that increments for each metadata frame. After the highest unsigned integer value has been reached, this field will wrap back to all zeros. Note that this value is used for determining discontinuities between metadata frames and it must always be incremented by one after each metadata frame.

Reserved **Professional** **Read and Write**

Word size: 12

Valid range: 0

This field is reserved and must not be changed, and must be passed on intact to preserve the structure and timing of the bitstream.

5.1.2 data_segment Field Descriptions

All Data Segments wrap the payload data in the following fields. The data_segment_id field describes the payload data type, and is also used to signal the end of a subframe.

data_segment_id **Professional** **Read Only**

Word size: 8

Valid range: See Table 5–6

The data_segment_id indicates the type of data being carried in the payload of the following data_segment. See Table 5-6.

Table 5-6 – Allowable Values for data_segment_id

data_segment_id	Data Segment type
0	None (Signals the end of a subframe)
1	Dolby E Complete Metadata Data Segment
2	Dolby E Essential Metadata Data Segment
3	Dolby Digital Complete Program Metadata Data Segment with Extended BSI support
4	Dolby Digital Essential Program Metadata Data Segment, with Extended BSI support
5	Dolby Digital Complete Program Metadata Data Segment with No Extended BSI support
6	Dolby Digital Essential Program Metadata Data Segment with No Extended BSI support
7–255	Reserved

If the data_segment_id field is not = 0, then the following fields wrap the actual payload data.

Table 5-7 – Data Segment Syntax

Syntax	**word size (bits)**
data_segment()	
{	
data_segment_size	8
data_segment_payload()	
data_segment_unused_bits	variable
data_segment_checksum	8
}	

data_segment_size **Pro and Cons** **Read and Write**

Word size: 8

Valid range: 0–255 (0 is interpreted as 256 bytes)

This field specifies the number of bytes in the data segment payload including the **data_segment_unused_bits**. Note that since the data segment payload does not necessarily contain an integer number of bytes, this field represents the least integer greater than or equal to the number of bytes in the segment payload.

data_segment_size = (n/8)

rounded up to the nearest integer, where n represents the number of bits in the data_segment_payload.

data_segment_payload() **See Sections 6 and 7**

The data_segment_payload() field contains a variable number of bits that carry different types of metadata information identified by the data_segment_id. Sections 6 and 7 describe the syntax for each of the data_segments defined in Table 5-6.

data_segment_unused_bits **Pro and Cons** **Read and Write**

Word size: Variable (0–7)

Valid range: All bits must be set to zero

This variable length field is used to force each data segment to have an integer number of bytes. All the bits in this field should be set to zero. The number of bits in this field can be computed as follows:

number of bits in data_segment_unused_bits field = (m*8) – n

,

where m is the number of bytes in the data_segment_size field and n is the number of bits in the data_segment_payload field.

data_segment_checksum **Pro and Cons** **Read and Write**

Word size: 8

Valid range: All values

This field is a 2's complement checksum for the data segment. It is calculated over the data_segment_size, the data_segment_payload(), and the data_segment_unused_bits fields. The checksum is initialized with the value of the data_segment_size byte, then bytes of the data_segment_payload are added sequentially from the beginning of the data_segment_payload. Any carry bits generated are thrown away by ANDing the totals with 0xFF.

The pseudo code for the operation is:

```
/* initialize the checksum to the data segment size */
checksum = data_segment_size;
/* perform checksum on the data segment payload including unused bits */
for (j = 0; j < data_segment_size; j++)
{
        data = data_segment_payload[j];
        checksum = (checksum + data) & 0xFF;
}
/* take the 2's complement of the running checksum */
checksum = ((~checksum) + 1) & 0xFF.
```

5.2 Metadata Frame Field Description

end_of_frame_sync_word **Professional and Consumer** **Read Only**

Word size: 16
Value: 0x0A0A

The end_of_frame_sync_word is used to signal the end of the metadata frame.

6 Dolby E Data Segment Payload Structure

The payload carries several types of information. Each is identified by the data_segment_id byte shown in Table 6-1. The data words within the payload do not always fall on byte boundaries, so data-Segment_unused_bits are added to the end of the payload to bring the entire payload length to an integer number of bytes.

6.1 Dolby E Complete Metadata Data Segment

6.1.1 Syntax Descriptions

Table 6-1 shows Dolby E complete metadata segment field descriptions.

Table 6-1 – Dolby E Complete Metadata Data Segment Field Descriptions

Syntax	word size (bits)
DolbyE_Complete_Metadata_Data_Segment()	
{	
program_config	6
frame_rate_code	4
pitch_shift_code	12
SMPTE_time_code	64
Reserved	8
for (pgm = 0; pgm < program_count; pgm++) /*program_count derived from program_config*/	
{	
description_text[pgm]	8
Reserved[pgm]	2
}	
for (ch = 0; ch < max_channels; ch++) /*max_channels derived from program_config*/	
{	
Reserved[ch]	4
Reserved[ch]	1
}	
for (ch = 0; ch < max_channels; ch++) /*max_channels derived from program_config*/	
{	
Reserved[ch]	10
Reserved[ch]	10
}	
}	

6.1.2 Field Descriptions

program_config **Professional** **Read Only**

Word size: 6
Valid range: 0–23, values 24 to 63 are Reserved

This frame element indicates how the programs are packed into the Dolby E frame, as shown in Table 6-2. The program configuration field defines the number of separate programs in the frame, and the number of separate channels in each program. It also identifies where in the sequence of metadata fields, the field associated with a specific program or channel within a program will be found, by specifying the order in which the program and channel data is packed into the Dolby E Complete data segment.

Fields, such as the description_text, have a [pgm] suffix to indicate that they apply to a program as a whole. Other elements that only apply to individual channels use a [ch] suffix to indicate this. The system used to organize the metadata elements, shown in Table 6-2, is to start with those that apply to the Left (L) channels for all the programs, in program sequence, followed by all the “C” or Center (also used for mono signals) channels in program sequence, followed by the Left Surround (Ls), Right (R), Low Frequency Effects (LFE) and finally the Right Surround (Rs) channels, again in program order, with the program count beginning at zero.

NOTE that the Channel Sequence shown in Table 6-2 does not apply to the audio inputs and outputs of Dolby E encoders and decoders. It is only used to locate specific parameters within the serial metadata bit stream.

Table 6-2 – Program Configuration

Program Config	Program Count	Channel Count	Program Sequence	Channel Sequence
0	2	8	5.1 + 2	0L, 0C, 0Ls, 1L, 0R, 0LFE, 0Rs, 1R
1	3	8	5.1 + 1 + 1	0L, 0C, 0Ls, 1C, 0R, 0LFE, 0Rs, 2C
2	2	8	4 + 4	0L, 0C, 1L, 1C, 0R, 0S, 1R, 1S
3	3	8	4 + 2 + 2	0L, 0C, 1L, 2L, 0R, 0S, 1R, 2R
4	4	8	4 + 2 + 1 + 1	0L, 0C, 1L, 2C, 0R, 0S, 1R, 3C
5	5	8	4 + 1 + 1 + 1 + 1	0L, 0C, 1C, 3C, 0R, 0S, 2C, 4C
6	4	8	2 + 2 + 2 + 2	0L, 1L, 2L, 3L, 0R, 1R, 2R, 3R
7	5	8	2 + 2 + 2 + 1 + 1	0L, 1L, 2L, 3C, 0R, 1R, 2R, 4C
8	6	8	2 + 2 + 1 + 1 + 1 + 1	0L, 1L, 2C, 4C, 0R, 1R, 3C, 5C
9	7	8	2 + 1 + 1 + 1 + 1 + 1 + 1	0L, 1C, 3C, 5C, 0R, 2C, 4C, 6C
10	8	8	1 + 1 + 1 + 1 + 1 + 1 + 1 + 1	0C, 2C, 4C, 6C, 1C, 3C, 5C, 7C
11	1	6	5.1	0L, 0C, 0Ls, 0R, 0LFE, 0Rs
12	2	6	4 + 2	0L, 0C, 1L, 0R, 0S, 1R
13	3	6	4 + 1 + 1	0L, 0C, 1C, 0R, 0S, 2C
14	3	6	2 + 2 + 2	0L, 1L, 2L, 0R, 1R, 2R
15	4	6	2 + 2 + 1 + 1	0L, 1L, 2C, 0R, 1R, 3C
16	5	6	2 + 1 + 1 + 1 + 1	0L, 1C, 3C, 0R, 2C, 4C
17	6	6	1 + 1 + 1 + 1 + 1 + 1	0C, 2C, 4C, 1C, 3C, 5C
18	1	4	4	0L, 0C, 0R, 0S
19	2	4	2 + 2	0L, 1L, 0R, 1R
20	3	4	2 + 1 + 1	0L, 1C, 0R, 2C
21	4	4	1 + 1 + 1 + 1	0C, 2C, 1C, 3C
22	1	8	7.1	0L,0C,0Ls,0BSL,0R,0LFE,0Rs,0BSR
23	1	8	7.1 Screen	0L,0C,0Ls,0Le,0R,0LFE,0Rs,0Re
24-63	reserved	reserved	reserved	reserved

frame_rate_code **Professional** **Read and Write**

Word size: 4
Valid range: See Table 6-3

This frame element indicates the frame rate of the video reference signal that the device producing the metadata stream is locked to, as shown in Table 6-3.

Table 6-3 – Frame Rate Code

Frame Rate Code	Frame Rate
1	24000/1001 ~ 23.98 Hz (film normalized to NTSC)
2	24 Hz (film rate)
3	25 Hz (PAL frame rate)
4	30000/1001 ~ 29.97 Hz (NTSC frame rate)
5	30 Hz
0, 6-15	reserved

pitch_shift_code **Professional** **Read and Write**

Word size: 12

Valid Range: -2048 to 2047

This field indicates the amount of pitch shift between the original playback speed and the current playback speed of the audio subframe. With pitch_shift_code = 0, the current playback speed is equal to the original playback speed, and therefore indicates no pitch shift. The pitch_shift_code is expressed in a logarithmic scale ranging from pitch_shift_code = -2048 (current playback speed is 25% of the original playback speed) to pitch_shift_code = 2047 (current playback speed is 399.73 % of the original playback speed). The equations below describe how to compute the pitch_shift_code.

pitch_shift_code = $1024 \text{x} \log_2$(pitch_shift_ratio),

where pitch_shift_ratio is computed as:

$$pitch\_shift\_ratio = \frac{current\_audio\_subframe\_playback\_speed}{original\_audio\_subframe\_playback\_speed}$$

SMPTE_time_code **Professional** **Read and Write**

Word size: 64

Valid range: See Table 6-4

This frame element contains information relating to the SMPTE longitudinal timecode associated with the Dolby E frame.

Table 6-4 – SMPTE Timecode

SMPTE Timecode Byte	MSB			Bit Number				LSB
	7	6	5	4	3	2	1	0
byte 0	[63]	[62]	[61]	[60]	[55]	[54]	[53]	[52]
byte 1	[59]	[58]	H20	H10	H8	H4	H2	H1
byte 2	[47]	[46]	[45]	[44]	[39]	[38]	[37]	[36]
byte 3	[43]	M40	M20	M10	M8	M4	M2	M1
byte 4	[31]	[30]	[29]	[28]	[23]	[22]	[21]	[20]
byte 5	[27]	S40	S20	S10	S8	S4	S2	S1
byte 6	[15]	[14]	[13]	[12]	[7]	[6]	[5]	[4]
byte 7	[11]	DF	F20	F10	F8	F4	F2	F1

In Table 6-4, the notation [N] corresponds to the Nth bit of the SMPTE timecode word. Some of these values are identified explicitly for ease of reference. The values corresponding to BCD-coded hour, minutes, seconds, and frames are indicated with Hn, Mn, Sn, and Fn notation. The drop frame flag is indicated as DF.

It is possible that the SMPTE timecode field may not contain valid timecode information. In these cases, the time code will be flagged as invalid by setting the BCD-coded hours field to the illegal value of 0x3F.

Reserved **Professional** **Read and Write**

Word size: 8
Valid range: 0

This field is reserved for future use and must be set to 0 by Dolby E encoders, and ignored by Dolby E decoders. It must be passed on intact to preserve the structure and timing of the bitstream.

description_text [pgm] **Professional** **Read and Write**

Word size: 8
Valid range: 0x00, 0x02, 0x03, 0x20–0x7E

This frame element is an ASCII-formatted byte that is part of a multi-character text description of the associated program. The value 0x02 (STX) indicates that the next character is the start of the text description, while the value 0x03 (ETX) indicates that the previous character was the end of the text description. The value 0x00 (NUL) indicates that no text description information is present in the current frame. Values between 0x20 and 0x7E convey the corresponding ASCII text character. All other values are reserved and should be discarded if received.

Reserved [pgm] **Professional** **Read and Write**

Word size: 2
Valid range: 0

This field is reserved and must not be changed, and must be passed on intact to preserve the structure and timing of the bitstream.

Reserved [ch] **Professional** **Read and Write**

Word size: 4
Valid range: 0

This field is reserved and must not be changed, and must be passed on intact to preserve the structure and timing of the bitstream.

Reserved [ch] **Professional** **Read and Write**

Word size: 1
Valid range: 0

This field is reserved and must not be changed, and must be passed on intact to preserve the structure and timing of the bitstream.

Reserved [ch] **Professional** **Read and Write**

Word size: 10
Valid range: 0x3C0

This field is reserved and must not be changed, and must be passed on intact to preserve the structure and timing of the bitstream.

Reserved [ch] **Professional** **Read and Write**

Word size: 10
Valid range: 0x3C0

This field is reserved and must not be changed, and must be passed on intact to preserve the structure and timing of the bitstream.

6.2 Dolby E Essential Metadata Data Segment

This data segment contains essential Dolby E metadata parameters for a particular Dolby E frame. The information contained in this data segment is a subset of the information contained in the Dolby E Complete Metadata Data Segment. Refer to Section 6.1 for a description of the **program_config** field.

6.2.1 Syntax Description

Table 6-5 describes the syntax of each Dolby E Essential metadata data segment.

Table 6-5 – Syntax Description

Syntax	word size (bits)
DolbyE_Essential_Metadata_Data_Segment()	
{	
program_config	6
frame_rate_code	4
pitch_shift_code	12
Reserved	8
}	

6.2.2 Field Descriptions

program_config **Professional** **Read Only**

Word size: 6

Valid range: 0–21 and 63

This frame element indicates how the programs are packed into the Dolby E frame, as shown in Table 6-2. See section 6.1.2 for a complete description.

frame_rate_code **Professional** **Read and Write**

Word size: 4

Valid range: See Table 6-3

This frame element indicates the frame rate of the video reference signal that the device producing the metadata stream is locked to. See Section 6.1.2.

pitch_shift_code **Professional** **Read and Write**

Word size: 12

Valid range: –2048 to 2047

This field indicates the amount of pitch shift between the original playback speed and the current playback speed of the audio subframe. With pitch_shift_code = 0, the current playback speed is equal to the original playback speed, and therefore indicates no pitch shift. The pitch_shift_code is expressed in a logarithmic scale ranging from pitch_shift_code = -2048 (current playback speed is 25% of the original playback speed) to pitch_shift_code = 2047 (current playback speed is 399.73 % of the original playback speed). The equations below describe how to compute the pitch_shift_code.

pitch_shift_code = 1024xlog$_2$(pitch_shift_ratio),

where pitch_shift_ratio is computed as:

$$pitch\_shift\_ratio = \frac{current\_audio\_subframe\_playback\_speed}{original\_audio\_subframe\_playback\_speed}$$

Reserved	**Professional**	**Read and Write**

Word size: 8

Valid range: 0

This field is reserved for future use and must be set to 0 by Dolby E encoders, and ignored by Dolby E decoders. It must be passed on intact to preserve the structure and timing of the bitstream.

7 Dolby Digital Data Segment Payload Structure

The payload carries several types of information. Each is identified by the data_segment_id byte shown in Table 5-6.

7.1 Dolby Digital Complete Program Metadata Data Segment with Extended BSI Support

This data segment contains Dolby Digital bitstream parameters for each program, including the extended bitstream information (bsi) parameters defined in ATSC specification A/52B.

The Extended bitstream syntax may be implemented by some AC-3 encoders and interpreted by some AC-3 decoders. It is not necessary for all decoders to be aware of this extended syntax in order to properly reconstruct an audio sound field; however those decoders that are aware of this syntax will be able to take advantage of the new system features provided by the extensions.

The extended syntax redefines two 14-bit fields that were originally identified as ac3_timecod1 and ac3_timecod2 in ATSC specification A/52B, but were not being used. The extended bitstream syntax is signaled by a data_segment_id = 0x03.

7.1.1 Syntax Description

Table 7-1 – Syntax Description

Syntax	word size (bits)
DolbyD_Complete_Program_Metadata_XBSI_Payload()	
{	
program_id [pgm]	5
ac3_datarate [pgm]	5
ac3_bsmod [pgm]	3
ac3_acmod [pgm]	3
ac3_cmixlev [pgm]	2
ac3_surmixlev [pgm]	2
ac3_dsurmod [pgm]	2
ac3_lfeon [pgm]	1
ac3_dialnorm [pgm]	5
ac3_langcode [pgm]	1
ac3_langcod [pgm]	8
ac3_audprodie [pgm]	1
ac3_mixlevel [pgm]	5
ac3_roomtyp [pgm]	2
ac3_copyrightb [pgm]	1
ac3_origbs [pgm]	1
ac3_xbsi1e [pgm]	1
ac3_dmixmod [pgm]	2
ac3_ltrtcmixlev [pgm]	3
ac3_ltrtsurmixlev [pgm]	3
ac3_lorocmixlev [pgm]	3
ac3_lorosurmixlev [pgm]	3
ac3_xbsi2e [pgm]	1

Syntax	word size (bits)
ac3_dsurexmod [pgm]	2
ac3_dheadphonmod [pgm]	2
ac3_adconvtyp [pgm]	1
Reserved [pgm]	8
Reserved [pgm]	1
ac3_hpfon [pgm]	1
ac3_bwlpfon [pgm]	1
ac3_lfelpfon [pgm]	1
ac3_sur90on [pgm]	1
ac3_suratton [pgm]	1
ac3_rfpremphon [pgm]	1
ac3_compre [pgm]	1
ac3_compr1 [pgm]	8
ac3_dynrnge [pgm]	1
ac3_dynrng1 [pgm]	8
ac3_dynrng2 [pgm]	8
ac3_dynrng3 [pgm]	8
ac3_dynrng4 [pgm]	8
Reserved [pgm]	1
}	

7.1.2 Field Descriptions

program_id [pgm] **Professional** **Read Only**

Word size: 5
Valid range: 0–7 (Other values are reserved.)

This field identifies which program within a Dolby E frame the data segment payload applies to. The program_id field points to a specific program, based on the fixed order of programs specified by the Dolby E program_config field (see section 6.1.2).

Example: program_config = 7, (2+2+2+1+1) and program_id = 2, which means that the Dolby Digital_Complete_Program_Metadata_Payload() refers to the third stereo program for this specific program_config.

ac3_datarate[pgm] **Professional** **Read and Write**

Word size: 5
Valid range: See table 7-2 below

This frame element indicates the data rate that should be used to encode the AC-3 bitstream associated with the specified program, as shown in the table below.

NOTE – The standard data rate for ATSC transmission in the US is 384 kbps. Use of any other rate for this application will cause problems.

Table 7-2 – AC-3 Program Data Rate

data rate	data rate in kbps	data rate	data rate in kbps
0	32 kbps	11	224 kbps
1	40 kbps	12	256 kbps
2	48 kbps	13	320 kbps
3	56 kbps	14	384 kbps
4	64 kbps	15	448 kbps
5	80 kbps	16	512 kbps
6	96 kbps	17	576 kbps
7	112 kbps	18	640 kbps
8	128 kbps	19-30	reserved
9	160 kbps	31	not specified
10	192 kbps		

ac3_bsmod [pgm] **Consumer** **Read and Write**

Word size: 3
Valid range: 0–7 (All values)

The Bitstream mode code indicates the type of program service that is being carried as defined in Table 7-3.

Table 7-3 – AC-3 Bitstream Mode

bsmod	acmod	Type of Service
000	any	main audio service: complete main (CM)
001	any	main audio service: music and effects (ME)
010	any	associated service: visually impaired (VI)
011	any	associated service: hearing impaired (HI)
100	any	associated service: dialogue (D)
101	any	associated service: commentary (C)
110	any	associated service: emergency (E)
111	001	associated service: voice over (VO)
111	010 to 111	main audio service: karaoke

ac3_acmod [pgm] **Consumer** **Read and Write**

Word size: 3
Valid range: 0–7 (All values)

This 3-bit code, shown in Table 7-4, indicates which of the full bandwidth main service channels are in use. If the msb of acmod is 1, then the surround channels are in use and the ac3_surmixlev parameter follows later in the bitstream. If the msb of acmod is 0, the surround channels are not in use and the ac3_surmixlev is omitted from the bitstream. If the lsb of acmod is 0, the center channel is not in use. If the lsb of acmod is a 1, the center channel is in use.

NOTE ATSC A/52B metadata includes a parameter “nchans” that indicates if the reduced bandwidth LFE channel is in use. See also ac3_lfeon.

NOTE that the Channel Ordering shown below does not apply to the audio inputs and outputs of Dolby Digital encoders and decoders.

NOTE – ac3-acmod = 000 is not allowed in Dolby E streams or in external metadata, so is a reserved value.

Table 7-4 – AC-3 Audio Coding Mode

ac3-acmod	Audio Coding Mode	Channel Ordering
000	reserved	n/a
001	1/0	C
010	2/0	L, R
011	3/0	L,C,R
100	2/1	L, R, S
101	3/1	L, C, R, S
110	2/2	L, R, Ls, Rs
111	3/2	L, C, R, Ls, Rs

ac3_cmixlev [pgm] **Consumer** **Read and Write**

Word size: 2
Valid range: 0–4 (All values)

When three front channels are in use, this 2-bit code, shown in Table 7-5, indicates the nominal down mix level of the center channel with respect to the left and right channels. If cmixlev is set to the reserved code, decoders should still reproduce audio. The intermediate value of cmixlev (–4.5 dB) may be used in this case.

Table 7-5 – Center Mix Level

ac3_cmixlev	Center downmix level
00	0.707 (–3.0 dB)
01	0.595 (–4.5 dB)
10	0.500 (–6.0 dB)
11	reserved

ac3_surmixlev [pgm] **Consumer** **Read and Write**

Word size: 2
Valid range: 0–4 (All values)

If surround channels are in use, this 2-bit code, shown in Table 7-6, indicates the nominal down mix level of the surround channels. If surmixlev is set to the reserved code, the decoder should still reproduce audio. The intermediate value of surmixlev (–6 dB) may be used in this case.

Table 7-6 – Surround Mix Level

ac3_surmixlev	Surround downmix level
00	0.707 (–3.0 dB)
01	0.500 (–6.0 dB)
10	0
11	reserved

ac3_dsurmod [pgm] **Consumer** **Read and Write**

Word size: 2

Valid range: 0–4 (All values)

When operating in the two channel mode, this 2-bit code, as shown in Table 7-7 indicates whether or not the program has been encoded in Dolby Surround. This information is not used by the AC-3 decoder, but may be used by other portions of the audio reproduction equipment. If ac3_dsurmod is set to the reserved code, the decoder should still reproduce audio. The reserved code may be interpreted as "not indicated".

Table 7-7 – Dolby Surround Mode

dsurmod	Indication
'00'	not indicated
'01'	NOT Dolby Surround encoded
'10'	Dolby Surround encoded
'11'	reserved

ac3_lfeon [pgm] **Consumer** **Read and Write**

Word size: 1

Valid range: 0 and 1

This bit has a value of 1 if the LFE (low-frequency effects or subwoofer) channel is on, and a value of 0 if the LFE channel is off.

ac3_dialnorm [pgm] **Consumer** **Read and Write**

Word size: 5

Valid range: 1–31 (0 is interpreted as 31)

When audio from different sources is reproduced, the apparent loudness of the dialog element, which generally serves as a reference for the loudness of all the other program elements, frequently varies from source to source. These sources might be different program segments during a broadcast (i.e., the movie vs. a commercial message) different broadcast channels, or different media (disc vs. tape). The dialnorm parameter indicates the mean level of the dialog during the program, relative to 0 dBFS (or full-scale digital level).

The 5-bit dialnorm value is interpreted as an unsigned integer that indicates how many dB the subjective dialogue level is below full scale.

The dialnorm parameter is used by the section of the sound reproduction system responsible for setting the reproduction level. Listeners generally set the system volume control to reproduce dialog at their preferred loudness. The value of dialnorm can then be used to adjust the gain of the reproduction system so that the dialog loudness of all the programs is reproduced at 31 dB below full scale. This compensates for the different dialog levels from program to program, with the result that the loudness remains constant, at the listener's desired level, from program to program.

As an example, the level of a program whose dialnorm value is 25 will be reduced by 6 dB so that the dialog will be reproduced at 31 dB below full scale. Similarly, the level of a program with a dialnorm of 17 will be attenuated by 14 dB. See the annex for a discussion of dialog normalization, dynamic range control, etc.

ac3_langcode [pgm] Consumer Read and Write

Word size: 1

Valid range: 0 and 1

If the "language code exists" bit is a 1, the following 8 bits represent a language code. If this bit is a 0, the language of the audio service is not indicated.

NOTE - The ATSC Standard A/52B, Digital Audio Compression Standard (AC-3, E-AC-3) Revision B, 14 June 2005 no longer uses the ac3_langcod parameter to indicate the program language. For ATSC DTV applications, ac3_langcode shall be set to "0" (not applicable).

ac3_langcod [pgm] Consumer Read and Write

Word size: 8

Valid range See Table 7-8 for valid and reserved values

This frame element indicates the language of the audio service of the AC-3 bitstream associated with the specified program. Note that this element is present in the Dolby E frame regardless of the value of the "language code exists" flag.

NOTE - The ATSC Standard A/52B, Digital Audio Compression Standard (AC-3, E-AC-3) Revision B, 14 June 2005 no longer uses the ac3_langcod parameter to indicate the program language. For ATSC DTV applications, ac3_langcod shall be set to "0" (not applicable).

Table 7-8 – Language Code Table

langcod	language	langcod	language	langcod	language	langcod	language
0x00	unknown/not applicable	0x20	Polish	0x40	background sound/clean feed	0x60	Moldavian
0x01	Albanian	0x21	Portugues e	0x41		0x61	Malaysian
0x02	Breton	0x22	Romanian	0x42		0x62	Malagasay
0x03	Catalan	0x23	Romansh	0x43		0x63	Macedonian
0x04	Croatian	0x24	Serbian	0x44		0x64	Laotian
0x05	Welsh	0x25	Slovak	0x45	Zulu	0x65	Korean
0x06	Czech	0x26	Slovene	0x46	Vietnamese	0x66	Khmer
0x07	Danish	0x27	Finnish	0x47	Uzbek	0x67	Kazakh
0x08	German	0x28	Swedish	0x48	Urdu	0x68	Kannada
0x09	English	0x29	Turkish	0x49	Ukrainian	0x69	Japanese
0x0A	Spanish	0x2A	Flemish	0x4A	Thai	0x6A	Indonesian
0x0B	Esperanto	0x2B	Walloon	0x4B	Telugu	0x6B	Hindi
0x0C	Estonian	0x2C		0x4C	Tatar	0x6C	Hebrew
0x0D	Basque	0x2D		0x4D	Tamil	0x6D	Hausa
0x0E	Faroese	0x2E		0x4E	Tadzhik	0x6E	Gurani
0x0F	French	0x2F		0x4F	Swahili	0x6F	Gujurati
0x10	Frisian	0x30	reserved for nat'l assignment	0x50	Sranan Tongo	0x70	Greek
0x11	Irish	0x31	"	0x51	Somali	0x71	Georgian
0x12	Gaelic	0x32	"	0x52	Sinhalese	0x72	Fulani
0x13	Galician	0x33	"	0x53	Shona	0x73	Dari
0x14	Icelandic	0x34	"	0x54	Serbo-Croat	0x74	Churash
0x15	Italian	0x35	"	0x55	Ruthenian	0x75	Chinese
0x16	Lappish	0x36	"	0x56	Russian	0x76	Burmese
0x17	Latin	0x37	"	0x57	Quechua	0x77	Bulgarian
0x18	Latvian	0x38	"	0x58	Pustu	0x78	Bengali
0x19	Luxembourgi an	0x39	"	0x59	Punjabi	0x79	Belorussian
0x1A	Lithuanian	0x3A	"	0x5A	Persian	0x7A	Bambora
0x1B	Hungarian	0x3B	"	0x5B	Papamiento	0x7B	Azerbijani
0x1C	Maltese	0x3C	"	0x5C	Oriya	0x7C	Assamese
0x1D	Dutch	0x3D	"	0x5D	Nepali	0x7D	Armenian
0x1E	Norwegian	0x3E	"	0x5E	Ndebele	0x7E	Arabic
0x1F	Occitan	0x3F	"	0x5F	Marathi	0x7F	Amharic

ac3_audprodie [pgm] **Consumer** **Read and Write**

Word Size: 3
Valid range: 0 and 1

If the "audio production info exists" bit is a 1, the mix level and roomtyp fields exist, giving information about the audio production environment used in making the program. If the "audio production info exists" bit is a 0, the mix level and roomtyp fields exist in the bitstream, but have no meaning.

ac3_mixlevel [pgm] **Consumer** **Read and Write**

Word size: 5
Valid range: 0–31

The "mixing level" code indicates the absolute acoustic sound pressure level of an individual program during the final audio mixing session. The 5-bit code represents a value in the range 0 to 31. The peak mixing level is 80 plus the value of mixlevel dB SPL, or 80 to 111 dB SPL. The peak mixing level is the acoustic level of a sine wave in a single channel whose peaks reach 100 percent in the PCM representation. The absolute SPL value is typically measured by means of pink noise with an RMS value of 20 or 30 dB below the peak RMS sine wave level. The value of mixlevel is not typically used within the AC-3 decoder, but may be used by other parts of the audio reproduction equipment. Note that this element is present in the Dolby E frame regardless of the value of the audio production information exists flag.

ac3_roomtyp [pgm] **Consumer** **Read and Write**

Word size: 2
Valid range: All values

The room type code, shown in Table 7-9 indicates the relative size and monitor frequency response curve of the mixing room used for the final audio mixing session. The value of roomtyp is not typically used by the AC-3 decoder, but may be used by other parts of the audio reproduction equipment. If roomtyp is set to the reserved code (which may be interpreted as "not indicated") the decoder should still reproduce audio. Note that this element is present in the Dolby E frame regardless of the value of the audio production information exists flag.

Table 7-9 –Room Type

Roomtyp	Type of Mixing Room
'00'	not indicated
'01'	large room, X curve monitor
'10'	small room, flat monitor
'11'	reserved

ac3_copyrightb [pgm] **Consumer** **Read and Write**

Word size: 1
Valid range: 0 and 1

If this bit has a value of 1, the information in the bitstream is indicated as protected by copyright. It has a value of 0 if the information is not indicated as protected.

ac3_origbs [pgm] **Consumer** **Read and Write**

Word size: 1
Valid range: 0 and 1

The "original bitstream" bit has a value of 1 if this is an original bitstream. This bit has a value of 0 if this is a copy of another bitstream.

ac3_xbsi1e [pgm] **Consumer** **Read and Write**

Word size: 1
Valid range: 0 and 1

If the “Extended bitstream information #1 exists” bit is a 1, the following 14 bits contain the following extended bitstream information, rather than the time code 1 data in the “no extended bitstream information” version. See the Dolby E metadata protocol layout chart.

ac3_dmixmod [pgm] **Consumer** **Read and Write**

Word size: 2
Valid range: See Table 7-10

The Preferred stereo downmix mode code, as shown in Table 7-10, indicates the type of stereo downmix preferred by the mastering engineer. This information may be used by the AC-3 decoder to automatically configure the type of stereo downmix, but may also be overridden or ignored. If dmixmod is set to the reserved code, the decoder should still reproduce audio. The reserved code may be interpreted as “not indicated”.

Table 7-10 – Preferred Stereo Downmix Mode

dmixmod	Indication
00	Not indicated
01	Lt/Rt downmix preferred
10	Lo/Ro downmix preferred
11	Reserved

NOTE: The meaning of this field is only defined as described if the audio coding mode is 3/0, 2/1, 3/1, 2/2, or 3/2. If the audio coding mode is 1+1, 1/0, or 2/0 then the meaning of this field is Reserved.

ac3_ltrtcmixlev [pgm] **Consumer** **Read and Write**

Word size: 3
Valid range: See Table 7-11

The “Lt/Rt center mix level” 3-bit code, shown in Table 7-11, indicates the nominal down mix level of the center channel with respect to the left and right channels in a Lt/Rt downmix.

Table 7-11 – Lt/Rt Center Mix Level

ltrtcmixlev	clev
000	1.414 (+3.0 dB)
001	1.189 (+1.5 dB)
010	1.000 (0.0 dB)
011	0.841 (–1.5 dB)
100	0.707 (–3.0 dB)
101	0.595 (–4.5 dB)
110	0.500 (–6.0 dB)
111	0.000 (–inf dB)

NOTE: The meaning of this field is only defined as described if the audio coding mode is 3/0, 3/1 or 3/2. If the audio coding mode is 1+1, 1/0, 2/0, 2/1, or 2/2 then the meaning of this field is Reserved.

ac3_ltrtsurmixlev [pgm] **Consumer** **Read and Write**

Word size: 3

Valid range: See Table 7-12

This 3-bit code, shown in Table 7-12, indicates the nominal down mix level of the surround channels with respect to the left and right channels in a Lt/Rt downmix.

Table 7-12 – Lt/Rt Surround Mix Level

ltrtsurmixlev	slev
000	1.414 (+3.0 dB) (Reserved)
001	1.189 (+1.5 dB) (Reserved)
010	1.000 (0.0 dB) (Reserved)
011	0.841 (–1.5 dB)
100	0.707 (–3.0 dB)
101	0.595 (–4.5 dB)
110	0.500 (–6.0 dB)
111	0.000 (–inf dB)

NOTE: The meaning of this field is only defined as described if the audio coding mode is 2/1, 3/1, 2/2 or 3/2. If the audio coding mode is 1+1, 1/0, 2/0 or 3/0 then the meaning of this field is Reserved.
ATSC Standard A/52B, Digital Audio Compression Standard (AC-3, E-AC-3) Revision B, 14 June 2005 states that the values '000', '001' and '010' are Reserved. It further states that for ATSC DTV applications, the decoder shall use a value of 0.841 for slev if one of the reserved values is received.

ac3_lorocmixlev [pgm] **Consumer** **Read and Write**

Word size: 3

Valid range: See Table 7-13

The "Lo/Ro center mix level" 3-bit code, shown in Table 7-13, indicates the nominal down mix level of the center channel with respect to the left and right channels in a Lo/Ro downmix.

Table 7-13 – Lo/Ro Center Mix Level

lorocmixlev	clev
000	1.414 (+3.0 dB)
001	1.189 (+1.5 dB)
010	1.000 (0.0 dB)
011	0.841 (–1.5 dB)
100	0.707 (–3.0 dB)
101	0.595 (–4.5 dB)
110	0.500 (–6.0 dB)
111	0.000 (–inf dB)

NOTE : The meaning of this field is only defined as described if the audio coding mode is 3/0, 3/1 or 3/2. If the audio coding mode is 1+1, 1/0, 2/0, 2/1 or 2/2 then the meaning of this field is Reserved..Values 0x000 to 0x010 inclusive are Reserved.

ac3_lorosurmixlev [pgm] **Consumer** **Read and Write**

Word size: 3

Valid range: See Table 7-14

The "Lo/Ro surround mix level" 3-bit code, shown in Table 7-14, indicates the nominal down mix level of the center channel with respect to the left and right channels in a Lo/Ro downmix.

Table 7-14 – Lo/Ro Surround Mix Level

lorosurmixlev	slev
000	1.414 (+3.0 dB) (Reserved)
001	1.189 (+1.5 dB) (Reserved)
010	1.000 (0.0 dB) (Reserved)
011	0.841 (–1.5 dB)
100	0.707 (–3.0 dB)
101	0.595 (–4.5 dB)
110	0.500 (–6.0 dB)
111	0.000 (–inf dB)

NOTE: The meaning of this field is only defined as described if the audio coding mode is 2/1, 3/1, 2/2 or 3/2. If the audio coding mode is 1+1, 1/0, 2/0 or 3/0 then the meaning of this field is Reserved.
ATSC Standard A/52B, Digital Audio Compression Standard (AC-3, E-AC-3) Revision B, 14 June 2005 states that the values ‘000’, ‘001’ and ‘010’ are Reserved. It further states that for ATSC DTV applications, the decoder shall use a value of 0.841 for slev if one of the reserved values is received.

ac3_xbsi2e [pgm] **Consumer** **Read and Write**

Word size: 1
Valid range: 0 and 1

If the "Extended bitstream information #2 exists" bit is a 1, the next 14 bits contain the following extended bitstream information, rather than the timecode 2 data in the “no extended bitstream information” version. See the Dolby E metadata protocol layout chart.

ac3_dsurexmod [pgm] **Consumer** **Read and Write**

Word size: 2
Valid range: See Table 7-15

The Dolby Surround EX™ mode code, as shown in Table 7-15, indicates whether or not the program has been encoded in Dolby Surround EX. This information is not used by the AC-3 decoder, but may be used by other portions of the audio reproduction equipment. If dsurexmod is set to the reserved code, the decoder should still reproduce audio. The reserved code may be interpreted as “not indicated”.

Table 7-15 – Dolby Surround EX Mode

ac3_dsurexmod	Indication
00	Not indicated
01	Not Dolby Surround EX encoded
10	Dolby Surround EX encoded
11	Reserved

NOTE: The meaning of this field is only defined as described if the audio coding mode is 2/2 or 3/2. If the audio coding mode is 1+1, 1/0, 2/0, 3/0, 2/1 or 3/1 then the meaning of this field is Reserved.

ac3_dheadphonmod [pgm] **Consumer** **Read and Write**

Word size: 2
Valid range: See Table 7-16

The “Dolby Headphone mode” code, as shown in Table 7-16, indicates whether or not the program has been Dolby Headphone-encoded. This information is not used by the AC-3 decoder, but may be used by other portions of the audio reproduction equipment. If ac3_dheadphonmod is set to the reserved code, the decoder should still reproduce audio. The reserved code may be interpreted as “not indicated”.

Table 7-16 – Dolby Headphone Mode

dheadphonmod	Indication
00	Not indicated
01	Not Dolby Headphone encoded
10	Dolby Headphone encoded
11	Reserved

NOTE: The meaning of this field is only defined as described if the audio coding mode is 2/0. If the audio coding mode is 1+1, 1/0, 3/0, 2/1, 3/1, 2/2 or 3/2 then the meaning of this field is Reserved.

ac3_adconvtyp [pgm] **Consumer** **Read and Write**

Word size: 1

Valid range: 0 and 1

The "A/D converter type" code, as shown Table 7-17, indicates the type of A/D converter technology used to capture the PCM audio. This information is not used by the AC-3 decoder, but may be used by other portions of the audio reproduction equipment. If the type of A/D converter used is not known, the "Standard" setting should be chosen.

Table 7-17 – A/D Converter Type

ac3_adconvtyp	Indication
0	Standard
1	HDCD

Reserved [pgm] **Professional** **Read and Write**

Word size: 8

Valid range: 0

This field is reserved and must not be changed, and must be passed on intact to preserve the structure and timing of the bitstream.

Reserved [pgm] **Professional** **Read and Write**

Word size: 1

Valid range: 0

This field is reserved and must not be changed, and must be passed on intact to preserve the structure and timing of the bitstream.

ac3_hpfon [pgm] **Professional** **Read and Write**

Word size: 1

Valid range: 0 and 1

This frame element indicates whether or not the DC blocking 3 Hz highpass filter is applied to the main input channels in the Dolby Digital encoder. It is used to remove DC offsets in the program audio and would only be switched off in exceptional circumstances.

0 indicates that the filter is disabled and 1 indicates that the filter is enabled.

ac3_bwlpfon [pgm] **Professional** **Read and Write**

Word size: 1

Valid range: 0 and 1

This parameter determines whether a lowpass filter is applied to the main input channels of a Dolby Digital encoder prior to encoding. The filter removes high frequency signals that are not encoded. At the suitable data rates this filter operates above 20 kHz. In all cases it prevents aliasing on decoding and is normally switched on. This parameter is not passed to the consumer decoder.

0 indicates that the filter is disabled and 1 indicates that the filter is enabled.

ac3_lfelpfon [pgm] **Professional** **Read and Write**

Word size: 1

Valid range: 0 and 1

This frame element determines whether a 120 Hz lowpass filter (used to remove frequencies above 120 Hz that would cause aliasing when decoded) is applied to the LFE channel input of a Dolby Digital encoder prior to encoding. It is ignored if the LFE channel is disabled. This filter should only be switched off if the audio to be encoded is known to have no signal above 120 Hz. This parameter is not sent to the consumer decoder.

0 indicates that the filter is disabled and 1 indicates that the filter is enabled.

ac3_sur90on [pgm] **Professional** **Read and Write**

Word size: 1

Valid range: 0 (disabled), 1 (enabled)

This frame element causes the Dolby Digital encoder to apply a 90-degree phase shift to the surround channels. This allows a Dolby Digital decoder to create a Lt/Rt downmix simply. For most material the phase shift has a minimal impact when the Dolby Digital program is decoded to 5.1 channels, but provides a Lt/Rt output that can be Pro Logic® decoded to L, C, R, and S, if desired. However, for some phase-critical material (such as music) this phase shift is audible when listening in 5.1 channels. Likewise some material down mixes to a satisfactory Lt/Rt signal without needing this phase shift. It is therefore important to balance the needs of the 5.1 mix and the Lt/Rt downmix for each program. This parameter is not sent to the consumer decoder.

ac3_suratton [pgm] **Professional** **Read and Write**

Word size: 1

Valid range: 0 (disabled), 1 (enabled)

The Surround 3 dB Attenuation frame element determines whether the surround channel(s) are attenuated 3 dB before encoding. It is used to compensate for the level of the surround signal(s) originating from a theatrical mixing room (dubbing stage) and those originating from mixing rooms used for Television and DVDs. For compatibility with older film formats, theatrical mixing rooms set the surround playback channel sensitivity (and thus the reproduced SPL) 3 dB lower than the sensitivity (thus reproduced SPL) of the front channels. The consequence is that the surround signal levels on tapes produced in a dubbing stage are 3 dB higher than those produced in a consumer mixing room. Therefore, to convert to a consumer mix from a theatrical mix, it is necessary to reduce the surround levels by 3 dB by enabling this frame element.

ac3_rfpremphon [pgm] **Professional** **Read and Write**

Word size: 1

Valid range: 0 (disabled), 1 (enabled)

This frame element is designed to protect against over modulation when a decoded Dolby Digital bitstream is RF modulated. When enabled ("1") the Dolby Digital encoder includes pre-emphasis in its calculations for RF Mode compression. The parameter has no effect when decoding using Line Mode compression. Except in rare cases, this parameter should be disabled.

ac3_compre [pgm] **Professional** **Read and Write**

Word size: 1

Valid range: 0 and 1

Each program carried in the Dolby E stream has either a compression profile name or a compression (compr) word associated with it. If ac3_compre is set to 0, the ac3_compr1 word indicates which dynamic range compression profile should be used by the AC-3 encoder, according to Table 7-17.

If this bit is a 1, the following 8 bits represent a compression control word. See Annex A for details of how the compr (and dynrng) gain values are expressed.

ac3_compr1 [pgm] **Professional and Consumer** **Read and Write**

Word size: 8

Valid range: 0–255 and see Table 7-18

This field indicates the RF compression word of the AC-3 bitstream associated with the specified program. ac3_compr1 is associated with the first half of the Dolby E frame, since it is carried in the Complete program metadata segment which applies to the audio in the first half of the Dolby E frame. Note that this field is always present in the Dolby E frame regardless of the value of the ac3_compre (RF compression word exists) flag as it carries either the compression profile name or a compression word.

The compr element allows the program provider (or broadcaster) to implement a large dynamic range reduction (heavy compression) in a way which assures that a monophonic downmix will not exceed a certain peak level. The heavily compressed audio program may be desirable for certain listening situations such as movie delivery to a hotel room, or to an airline seat. The peak level limitation is useful when, for instance, a monophonic downmix will feed an RF modulator and over modulation must be avoided.

If ac3_compre is set to 0, ac3_compr1 indicates which RF compression profile should be used by the AC-3 encoder, according to Table 7-18.

Table 7-18 – RF Compression Profile

ac3_compr1	RF Compression Profile
0	none
1	Film, Standard
2	Film, Light
3	Music, Standard
4	Music, Light
5	Speech
6	6–255 Reserved

See Annex A for details of how the compr and dynrng gain values are expressed.

ac3_dynrnge [pgm] **Professional** **Read and Write**

Word size: 1

Valid range: 0 and 1

See the ac3_dynrng1 - ac3_dynrng4 [pgm] notes below.

ac3_dynrng1 - ac3_dynrng4 [pgm] **Professional and Consumer** **Read and Write**

Word size: 8

Valid range: 0–255 and Annex A or Table 7-18

Each program carried in the Dolby E stream has either compression profile names or a stream of dynamic range control (dynrng) words associated with it. If ac3_dynrnge is set to 0, the ac3_dynrng1 word indicates which dynamic range compression profile should be used by the AC-3 encoder, according to Table 7-18.

Even though the compression profile normally remains constant over the duration of a program, compression profile names associated with each program in the Dolby E stream are carried by both the Complete and Essential metadata segments. The compression profile name carried in the Complete metadata segment is associated with the audio encoded in the first half of the Dolby E frame. The second compression profile name is carried in the Essential metadata segment, and associated with the audio encoded in the last half of the Dolby E frame.

If ac3_dynrnge is set to 1, then each metadata frame carries eight dynrng words per program that will be multiplexed into the Dolby Digital data stream and used by the Dolby Digital decoder to control the dynamic range of the reproduced program. Four of the eight words are carried in the Complete Metadata segment, and the remaining four words are carried in the Essential metadata segment. The first four words are associated with the audio encoded in the first half of the Dolby E frame and the last four words are associated with the audio encoded in the last half of the Dolby E frame.

See Annex A for details of how the compr and dynrng gain values are expressed.

Reserved [pgm] **Professional** **Read and Write**

Word size: 1

Valid range: 0

This field is reserved and must not be changed, and must be passed on intact to preserve the structure and timing of the bitstream.

7.2 Dolby Digital Complete Program Metadata Data Segment — No Extended BSI Support

This data segment contains Dolby Digital bitstream parameters for each program. This is the version of the bitstream syntax that uses the two 14-bit fields called ac3_timecod1 and ac3_timecod2 (along with their "exists" flags ac3_timecod1e and ac3_timecod2e) as shown in Table 7-19 and described in ATSC specification A/52B. All the other fields in the data segment are unaffected and as defined in Section 7.1. The Dolby E metadata protocol layout chart shows the differences explicitly.

7.2.1 Syntax Description

Table 7-19 – Dolby Digital Complete Program Metadata Segment without Extended Bitstream Information Syntax Description

Syntax	word size (bits)
DolbyD_Complete_Program_Metadata_NoXBSI_Payload()	
{	
program_id [pgm]	5
Reserved [pgm]	5
ac3_bsmod [pgm]	3
ac3_acmod [pgm]	3
ac3_cmixlev [pgm]	2
ac3_surmixlev [pgm]	2
ac3_dsurmod [pgm]	2
ac3_lfeon [pgm]	1
ac3_dialnorm [pgm]	5
ac3_langcode [pgm]	1
ac3_langcod [pgm]	8
ac3_audprodie [pgm]	1
ac3_mixlevel [pgm]	5
ac3_roomtyp [pgm]	2
ac3_copyrightb [pgm]	1
ac3_origbs [pgm]	1
ac3_timecod1e [pgm]	1
ac3_timecod1 [pgm]	14
ac3_timecod2e [pgm]	1
ac3_timecod2 [pgm]	14
ac3_hpfon [pgm]	1
ac3_bwlpfon [pgm]	1
ac3_lfelpfon [pgm]	1
ac3_sur90on [pgm]	1
ac3_suratton [pgm]	1
ac3_rfpremphon [pgm]	1
ac3_compre [pgm]	1
ac3_compr1 [pgm]	8
ac3_dynrnge [pgm]	1
ac3_dynrng1 [pgm]	8
ac3_dynrng2 [pgm]	8
ac3_dynrng3 [pgm]	8
ac3_dynrng4 [pgm]	8
Reserved [pgm]	1
}	

7.2.2 Field Descriptions

Only the timecode exists fields and the timecode fields themselves are described here. The rest of the frame elements are the same as described in Section 7.1.

ac3_timecod1e [pgm] and **Professional** **Read and Write**

ac3_timecod2e [pgm] **Professional** **Read and Write**

Word size: 1
Valid range: 0 and 1 (See Table 7-20)

These frame elements, taken together, indicate, as shown in Table 7-20, whether time codes follow in the bitstream. Since only the high resolution portion of the timecode is needed for fine synchronization, the 28-bit timecode is broken into two 14 bit halves. The low resolution first half represents the code in 8 second increments up to 24 hours. The high resolution second half represents the code in 1/64th frame increments up to 8 seconds.

Table 7-20 – Timecode Exists

ac3_timecod1e, timecod2e	Time code(s) present
0,0	not present
0,1	first half (14 bits) present
1,0	second half (14 bits) present
1,1	both halves (28 bits) present

timecod1 [pgm] **Professional** **Read and Write**

Word size: 14
Valid range: See below

This frame element represents the first half (or low resolution part) of the timecode associated with the program signal. The first 5 bits of this 14-bit field represent the time in hours, with valid values of 0–23. The next 6 bits represent the time in minutes, with valid values of 0–59. The final 3 bits represents the time in 8-second increments, with valid values of 0–7 (representing 0, 8, 16, ... 56 seconds).

timecod2 [pgm] **Professional** **Read and Write**

Word size: 14
Valid range: See below

This frame element represents the second half (or high resolution part) of the timecode associated with the program signal. The first 3 bits of this 14-bit field represent the time in seconds, with valid values from 0–7 (representing 0–7 seconds). The next 5 bits represents the time in frames, with valid values from 0–29. The final 6 bits represents fractions of 1/64 of a frame, with valid values from 0–63.

7.3 Dolby Digital Essential Program Metadata Data Segment

This data segment contains essential Dolby Digital metadata parameters for a single program. Some of the information contained in this data segment is a subset of the information contained in the Dolby Digital Complete Metadata Data Segment, and some is unique.

Different data_segment_ids (4 and 6) are used to indicate “with extended bsi support” and “no extended bsi support” respectively. This was done to allow future flexibility, but at present the contents of the Dolby Digital Essential Program Metadata Data Segment is the same in both cases.

7.3.1 Syntax Descriptions

Table 7-21 – Dolby D Essential Program Meta Data Payload

Syntax	word size (bits)
DolbyD_Essential_Program_Metadata_Payload()	
{	
program_id [pgm]	5
ac3_datarate [pgm]	5
ac3_bsmod [pgm]	3
ac3_acmod [pgm]	3
ac3_lfeon [pgm]	1
ac3_dialnorm [pgm]	5
ac3_compre [pgm]	1
ac3_compr2 [pgm]	8
ac3_dynrnge [pgm]	1
ac3_dynrng5 [pgm]	8
ac3_dynrng6 [pgm]	8
ac3_dynrng7 [pgm]	8
ac3_dynrng8 [pgm]	8
}	

Table 7-21 describes the syntax of each Dolby D Essential program metadata data segment.

7.3.2 Field Descriptions

program_id [pgm] **Professional** **Read and Write**

Word size: 5

Valid range: 0–7 (Other values are reserved.)

This field identifies which program within a Dolby E frame the data segment payload applies to. The program_id field points to a specific program, based on the fixed order of programs specified by the Dolby E program_config field.

Sections 6.1 with Table 6-2, Program Configuration, and the program_id description in Section 7.1.2 provide a full explanation and an example of how the program_id is used.

ac3_datarate[pgm] **Professional** **Read and Write**

Word size: 5

Valid range: See table below

This frame element indicates the data rate that should be used to encode the AC-3 bitstream associated with the specified program, as shown in Table 7-22 below.

.

Table 7 -22 — AC-3 Data Rate

data rate	data rate in kbps	data rate	data rate in kbps
0	32 kbps	11	224 kbps
1	40 kbps	12	256 kbps
2	48 kbps	13	320 kbps
3	56 kbps	14	384 kbps
4	64 kbps	15	448 kbps
5	80 kbps	16	512 kbps
6	96 kbps	17	576 kbps
7	112 kbps	18	640 kbps
8	128 kbps	19-30	reserved
9	160 kbps	31	not specified
10	192 kbps		

The following five metadata elements are identical to the same items listed in Table 7-1 and described in Section 7.1.2:

ac3_bsmod [pgm]
ac3_acmod [pgm]
ac3_lfeon [pgm]
ac3_dialnorm [pgm]
ac3_compre [pgm]

ac3_compr2 **Pro and Cons** **Read and Write**

Word size: 8

Valid range: 0–255 (See Table 7-23)

This field indicates the RF compression word of the AC-3 bitstream associated with the specified program. ac3_compr2 is associated with the second half of the Dolby E frame. Note that this field is always present in the Dolby E frame regardless of the value of the ac3_compre (RF compression word exists) flag as it carries either the compression profile name or a compression word.

The compr element allows the program provider (or broadcaster) to implement a large dynamic range reduction (heavy compression) in a way which assures that a monophonic downmix will not exceed a certain peak level. The heavily compressed audio program may be desirable for certain listening situations such as movie delivery to a hotel room, or to an airline seat. The peak level limitation is useful when, for instance, a monophonic downmix will feed an RF modulator and over modulation must be avoided.

If ac3_compre is set to 0, ac3_compr2 indicates which RF compression profile should be used by the AC-3 encoder, according to Table 7-23.

Table 7-23 – RF Compression Profile

ac3_compr2	RF compression profile
0	none
1	Film, Standard
2	Film, Light
3	Music, Standard
4	Music, Light
5	Speech
6	6–255 Reserved

See Annex A for details of how the compr and dynrng gain values are expressed.

ac3_dynrng5 to ac3_dynrng8 [pgm] **Pro and Cons** **Read and Write**

Word size: 8

Valid range: 0–255 and Annex A or Table 7-23

Each program carried in the Dolby E stream has either compression profile names or a stream of dynamic range control (dynrng) words associated with it. If ac3_dynrnge is set to 0, the ac3_dynrng5 word indicates which dynamic range compression profile should be used by the AC-3 encoder, according to the table shown above for ac3_compr2.

Even though the compression profile normally remains constant over the duration of a program, compression profile names associated with each program in the Dolby E stream are carried by both the Complete and Essential metadata segments. The compression profile name carried in the Complete metadata segment is associated with the audio encoded in the first half of the Dolby E frame. The second compression profile name is carried in the Essential metadata segment, and associated with the audio encoded in the last half of the Dolby E frame.

If ac3_dynrnge is set to 1, then each metadata frame carries eight dynrng words per program that will be multiplexed into the Dolby Digital data stream and used by the Dolby Digital decoder to control the dynamic range of the reproduced program.

Four of the eight words are carried in the Complete Metadata segment, and the remaining four words are carried in the Essential metadata segment. The first four words are associated with the audio encoded in the first half of the Dolby E frame and the last four words are associated with the audio encoded in the last half of the Dolby E frame.

See Annex A for details of how the compr and dynrng gain values are expressed.

Annex A:
Dynamic Range Control

A.1 Dialogue Normalization

The AC-3 syntax provides elements that allow the encoded bitstream to satisfy listeners in many different situations. The dialnorm element allows for uniform reproduction of spoken dialogue when decoding any AC-3 bitstream.

A.1.1 Overview

When audio from different sources is reproduced, the apparent loudness often varies from source to source. The different sources of audio might be different program segments during a broadcast (i.e., the movie vs. a commercial message); different broadcast channels; or different media (disc vs. tape). The AC-3 coding technology solves this problem by explicitly coding an indication of loudness into the AC-3 bitstream.

The subjective level of normal spoken dialogue is used as a reference. The 5-bit dialogue normalization word which is contained in BSI, dialnorm, is an indication of the subjective loudness of normal spoken dialogue compared to digital 100 percent. The 5-bit value is interpreted as an unsigned integer (most significant bit first) with a range of possible values from 1 to 31. The unsigned integer indicates the headroom in dB above the subjective dialogue level. This value can also be interpreted as an indication of how many dB the subjective dialogue level is below digital 100 percent.

The dialnorm value is not directly used by the AC-3 decoder. Rather, the value is used by the section of the sound reproduction system responsible for setting the reproduction volume, e.g. the system volume control. The system volume control is generally set based on listener input as to the desired loudness, or sound pressure level (SPL). The listener adjusts a volume control which generally directly adjusts the reproduction system gain. With AC-3 and the dialnorm value, the reproduction system gain becomes a function of both the listeners desired reproduction sound pressure level for dialogue, and the dialnorm value which indicates the level of dialogue in the audio signal. The listener is thus able to reliably set the volume level of dialogue, and the subjective level of dialogue will remain uniform no matter which AC-3 program is decoded.

A.1.2 Example

The listener adjusts the volume control to 67 dB. (With AC-3 dialogue normalization, it is possible to calibrate a system volume control directly in sound pressure level, and the indication will be accurate for any AC-3 encoded audio source). A high quality entertainment program is being received, and the AC-3 bitstream indicates that dialogue level is 25 dB below 100 percent digital level. The reproduction system automatically sets the reproduction system gain so that full scale digital signals reproduce at a sound pressure level of 92 dB. The spoken dialogue (down 25 dB) will thus reproduce at 67 dB SPL.

The broadcast program cuts to a commercial message, which has dialogue level at –15 dB with respect to 100 percent digital level. The system level gain automatically drops, so that digital 100 percent is now reproduced at 82 dB SPL. The dialogue of the commercial (down 15 dB) reproduces at a 67 dB SPL, as desired.

In order for the dialogue normalization system to work, the dialnorm value must be communicated from the AC-3 decoder to the system gain controller so that dialnorm can interact with the listener adjusted volume control. If the volume control function for a system is performed as a digital multiply inside the AC-3 decoder, then the listener selected volume setting must be communicated into the AC-3 decoder. The listener selected volume setting and the dialnorm value must be brought together and combined in order to adjust the final reproduction system gain.

Adjustment of the system volume control is not an AC-3 function. The AC-3 bitstream simply conveys useful information which allows the system volume control to be implemented in a way which automatically removes undesirable level variations between program sources. It is mandatory that the dialnorm value and the user selected volume setting both be used to set the reproduction system gain.

A.2 Dynamic Range Compression

The dynrng element allows the program provider to implement subjectively pleasing dynamic range reduction for most of the intended audience, while allowing individual members of the audience the option to experience more (or all) of the original dynamic range.

A.2.1 Overview

A consistent problem in the delivery of audio programming is that different members of the audience wish to enjoy different amounts of dynamic range. Original high quality programming (such as feature films) are typically mixed with quite a wide dynamic range. Using dialogue as a reference, loud sounds like explosions are often 20 dB or more louder, and faint sounds like leaves rustling may be 50 dB quieter. In many listening situations it is objectionable to allow the sound to become very loud, and thus the loudest sounds must be compressed downwards in level. Similarly, in many listening situations the very quiet sounds would be inaudible, and must be brought upwards in level to be heard. Since most of the audience will benefit from a limited program dynamic range, soundtracks which have been mixed with a wide dynamic range are generally compressed: the dynamic range is reduced by bringing down the level of the loud sounds and bringing up the level of the quiet sounds. While this satisfies the needs of much of the audience, it removes the ability of some in the audience to experience the original sound program in its intended form. The AC-3 audio coding technology solves this conflict by allowing dynamic range control values to be placed into the AC-3 bitstream.

The dynamic range control values, dynrng, indicate a gain change to be applied in the decoder in order to implement dynamic range compression. Each dynrng value can indicate a gain change of ±24 dB. The sequence of dynrng values are a compression control signal. An AC-3 encoder (or a bitstream processor) will generate the sequence of dynrng values. Each value is used by the AC-3 decoder to alter the gain of one or more audio blocks. The dynrng values typically indicate gain reduction during the loudest signal passages, and gain increases during the quiet passages. For the listener, it is desirable to bring the loudest sounds down in level towards dialogue level, and the quiet sounds up in level, again towards dialogue level. Sounds which are at the same loudness as the normal spoken dialogue will typically not have their gain changed.

The compression is actually applied to the audio in the AC-3 decoder. The encoded audio has full dynamic range. It is permissible for the AC-3 decoder to (optionally, under listener control) ignore the dynrng values in the bitstream. This will result in the full dynamic range of the audio.

A.2.2 Example

A feature film soundtrack is encoded into AC-3. The original program mix has dialogue level at –25 dB. Explosions reach full scale peak level of 0 dB. Some quiet sounds which are intended to be heard by all listeners are 50 dB below dialogue level (or –75 dB). A compression control signal (sequence of dynrng values) is generated by the AC-3 encoder. During those portions of the audio program where the audio level is higher than dialogue level the dynrng values indicate negative gain, or gain reduction. For full scale 0 dB signals (the loudest explosions), gain reduction of –15 dB is encoded into dynrng. For very quiet signals, a gain increase of 20 dB is encoded into dynrng.being reproduced. It is also permissible (again under listener control) for the decoder to use some fraction of the dynrng control value, and to use a different fraction of positive or negative values. The AC-3 decoder can thus reproduce either fully compressed audio (as intended by the compression control circuit in the AC-3 encoder); full dynamic range audio; or audio with partially compressed dynamic range, with different amounts of compression for high level signals and low level signals.

A listener wishes to reproduce this soundtrack quietly so as not to disturb anyone, but wishes to hear all of the intended program content. The AC-3 decoder is allowed to reproduce the default, which is full compression. The listener adjusts dialogue level to 60 dB SPL. The explosions will only go as loud as 70 dB (they are 25 dB louder than dialogue but get –15 dB of gain applied), and the quiet sounds will reproduce at 30 dB SPL (20 dB of gain is applied to their original level of 50 dB below dialogue level). The reproduced dynamic range will be 70 dB – 30 dB. The listening situation changes, and the listener now wishes to raise the reproduction level of dialogue to 70 dB SPL, but still wishes to limit how loud the program plays. Quiet sounds may be allowed to play as quietly as before. The listener instructs the AC-3 decoder to continue using the dynrng values which indicate gain reduction, but to attenuate the values which indicate gain increases by a factor of 1/2. The explosions will still reproduce 10 dB above dialogue level, which is now 80 dB SPL. The quiet sounds are now increased in level by 20 dB / 2 = 10 dB. They will now be reproduced 40 dB below dialogue level, at 30 dB SPL. The reproduced dynamic range is now 80 dB – 30 dB = 50 dB.B = 40 dB.

Another listener wishes the full original dynamic range of the audio. This listener adjusts the reproduced dialogue level to 75 dB SPL, and instructs the AC-3 decoder to ignore the dynamic range control signal. For this listener the quiet sounds reproduce at 25 dB SPL, and the explosions hit 100 dB SPL. The reproduced dynamic range is 100 dB – 25 dB = 75 dB. This reproduction is exactly as intended by the original program producer.

In order for this dynamic range control method to be effective, it should be used by all program providers. Since all broadcasters wish to supply programming in the form that is most usable by their audience, nearly all broadcasters will apply dynamic range compression to any audio program which has a wide dynamic range. This compression is not reversible unless it is implemented by the technique embedded in AC-3. If broadcasters make use of the embedded AC-3 dynamic range control system, then listeners can have some control over their reproduced dynamic range. Broadcasters must be confident that the compression characteristic that they introduce into AC-3 will, by default, be heard by the listeners. Therefore, the AC-3 decoder shall, by default, implement the compression characteristic indicated by the dynrng values in the data stream. AC-3 decoders may optionally allow listener control over the use of the dynrng values, so that the listener may select full or partial dynamic range reproduction.

A.2.3 Detailed Implementation

The dynrng field in the AC-3 data stream is 8 bits in length. In the case that acmod = 0 (1+1 mode, or 2 completely independent channels) dynrng applies to the first channel (Ch1), and dynrng2 applies to the second channel (Ch2). While dynrng is described below, dynrng2 is handled identically. The dynrng value may be present in any audio block. When the value is not present, the value from the previous block is used, except for block 0. In the case of block 0, if a new value of dynrng is not present, then a value of '0000 0000' should be used. The most significant bit of dynrng (and of dynrng2) is transmitted first. The first three bits indicate gain changes in 6.02 dB increments which can be implemented with an arithmetic shift operation. The following five bits indicate linear gain changes, and require a 6-bit multiply. We will represent the 3 and 5 bit fields of dynrng as following:

$X_0\, X_1\, X_2\, .\, Y_3\, Y_4\, Y_5\, Y_6\, Y_7$

The meaning of the X values is most simply described by considering X to represent a 3-bit signed integer with values from –4 to 3. The gain indicated by X is then (X + 1) * 6.02 dB. Table A-1 shows this in detail.

The value of Y is a linear representation of a gain change of up to 6 dB. Y is considered to be an unsigned fractional integer, with a leading value of 1, or: 0.1Y3 Y4 Y5 Y6 Y7 (base 2). Y can represent values between 0.1111112 (or 63/64) and 0.1000002 (or 1/2). Thus, Y can represent gain changes from –0.14 dB to –6.02 dB. The combination of X and Y values allows dynrng to indicate gain changes from 24.08 – 0.14 = +23.95 dB, to –18.06 – 6.02 = –24.08 dB. The bit code of '0000 0000' indicates 0 dB (unity) gain.

Table A-1 – Meaning of 3 msb of dynrng

X1	X2	X3	Integer Value	Gain Indicated	Arithmetic Shifts
0	1	1	3	+24.08 dB	4 left
0	1	0	2	+18.06 dB	3 left
0	0	1	1	+12.04	2 left
0	0	0	0	+6.02	1 left
1	1	1	–1	0 dB	None
1	1	0	–2	–6.02 dB	1 right
1	0	1	–3	–12.04 dB	2 right
1	0	0	–4	–18.06 dB	3 right

Partial Compression

The dynrng value may be operated on in order to make it represent a gain change which is a fraction of the original value. In order to alter the amount of compression which will be applied, consider the dynrng to represent a signed fractional number, or

X0 . X1 X2 Y3 Y4 Y5 Y6 Y7

where X0 is the sign bit and X1 X2 Y3 Y4 Y5 Y6 Y7 are a 7-bit fraction. This 8-bit signed fractional number may be multiplied by a fraction indicating the fraction of the original compression to apply. If this value is multiplied by 1/2, then the compression range of ± 24 dB will be reduced to ±12 dB. After the multiplicative scaling, the 8-bit result is once again considered to be of the original form X0 X1 X2 .Y3 Y4 Y5 Y6 Y7 and used normally.

A.3 Heavy Compression; compr, compr2

The compr element allows the program provider (or broadcaster) to implement a large dynamic range reduction (heavy compression) in a way which assures that a monophonic downmix will not exceed a certain peak level. The heavily compressed audio program may be desirable for certain listening situations such as movie delivery to a hotel room, or to an airline seat. The peak level limitation is useful when, for instance, a monophonic downmix will feed an RF modulator and over modulation must be avoided.

A.3.1 Overview

Some products that decode the AC-3 bitstream will need to deliver the resulting audio via a link with very restricted dynamic range. One example is the case of a television signal decoder which must modulate the received picture and sound onto an RF channel in order to deliver a signal usable by a low cost television receiver. In this situation, it is necessary to restrict the maximum peak output level to a known value with respect to dialogue level, in order to prevent over modulation. Most of the time, the dynamic range control signal, dynrng, will produce adequate gain reduction so that the absolute peak level will be constrained. However, since the dynamic range control system is intended to implement a subjectively pleasing reduction in the range of perceived loudness, there is no assurance that it will control instantaneous signal peaks adequately to prevent over modulation.

In order to allow the decoded AC-3 signal to be constrained in peak level, a second control signal, compr, (compr2 for Ch2 in 1+1 mode) may be present in the AC-3 data stream. This control signal should be present in all bitstreams that are intended to be receivable by, for instance, a television set top decoder. The compr control signal is similar to the dynrng control signal in that it is used by the decoder to alter the reproduced audio level. The compr control signal has twice the control range as dynrng (± 48 dB compared to ± 24 dB) with 1/2 the resolution (0.5 dB vs. 0.25 dB). Also, since the compr control signal lives in BSI, it only has a time resolution of an AC-3 frame (32 ms) instead of a block (5.3 ms).

Products that require peak audio level to be constrained should use compr instead of dynrng when compr is present in BSI. Since most of the time the use of dynrng will prevent large peak levels, the AC-3 encoder may only need to insert compr occasionally; i.e., during those instants when the use of dynrng would lead to excessive peak level. If the decoder has been instructed to use compr, and compr is not present for a particular frame, then the dynrng control signal shall be used for that frame.

In some applications of AC-3, some receivers may wish to reproduce a very restricted dynamic range. In this case, the compr control signal may be present at all times. Then, the use of compr instead of dynrng will allow the reproduction of audio with very limited dynamic range. This might be useful, for instance, in the case of audio delivery to a hotel room or an airplane seat.

A.3.2 Detailed Implementation

The compr field in the AC-3 data stream is 8 bits in length. In the case that acmod = 0 (1+1 mode, or 2 completely independent channels) compr applies to the first channel (Ch1), and compr2 applies to the second channel (Ch2). While compr is described below (for Ch1), compr2 is handled identically (but for Ch2). The most significant bit is transmitted first. The first four bits indicate gain changes in 6.02 dB increments which can be implemented with an arithmetic shift operation. The following four bits indicate linear gain changes, and require a 5-bit multiply. We will represent the two 4-bit fields of compr as follows:

X0 . X1 X2 X3 Y4 Y5 Y6 Y7

The meaning of the X values is most simply described by considering X to represent a 4-bit signed integer with values from –8 to +7. The gain indicated by X is then (X + 1) * 6.02 dB. Table A-2 shows this in detail.

Table A-2 – Insert Table Caption

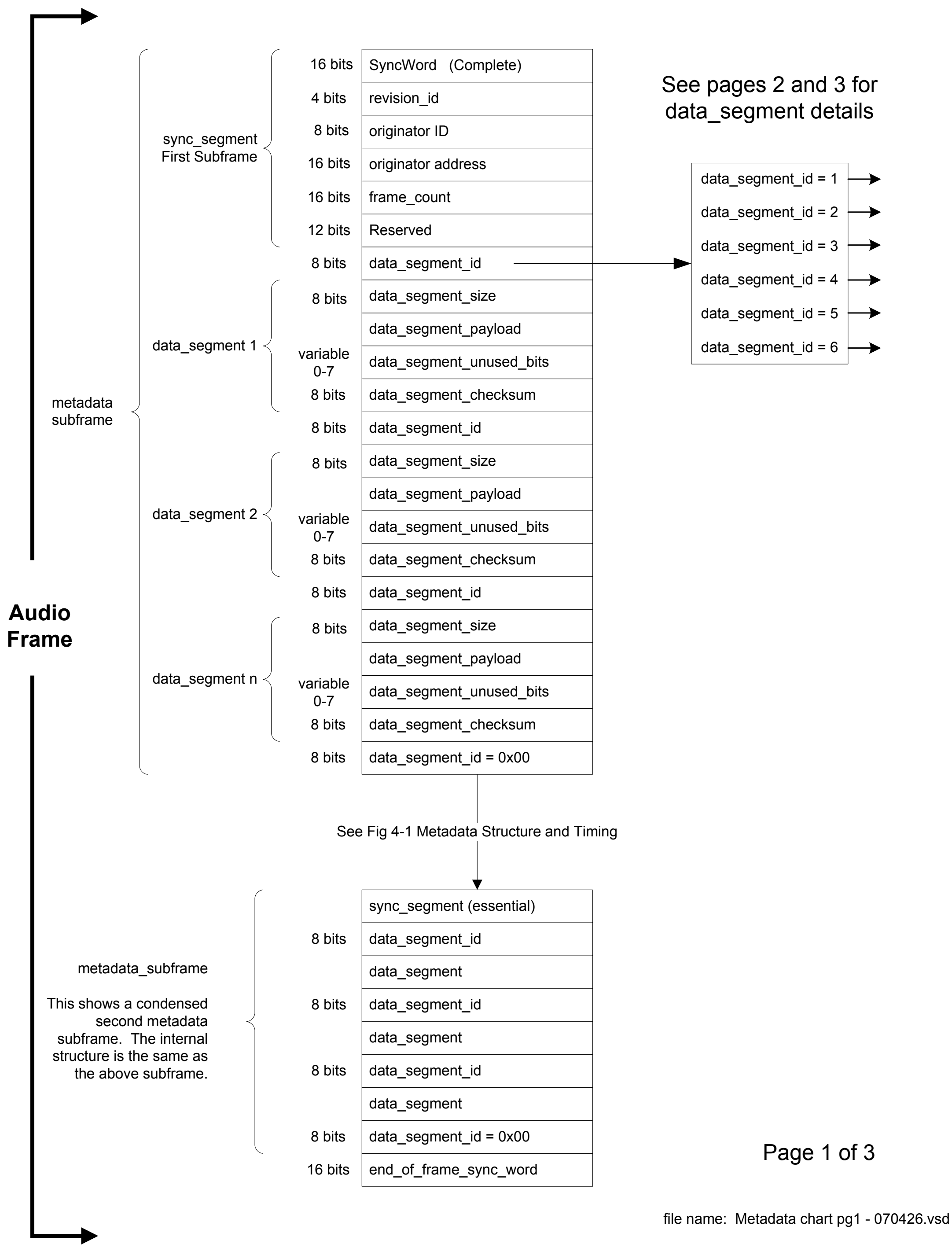
X_0	X_1	X_2	X_3	Integer Value	Gain Indicated	Arithmetic Shifts
0	1	1	1	7	+48.16 dB	8 left
0	1	1	0	6	+42.14 dB	7 left
0	1	0	1	5	+36.12 dB	6 left
0	1	0	0	4	+30.10 dB	5 left
0	0	1	1	3	+24.08	4 left
0	0	1	0	2	+18.06 dB	3 left
0	0	0	1	1	+12.04	2 left
0	0	0	0	0	+6.02 dB	1 left
1	1	1	1	–1	0 dB	None
1	1	1	0	–2	–6.02 dB	1 right
1	1	0	1	–3	–12.04 dB	2 right
1	1	0	0	–4	–18.06 dB	3 right
1	0	1	1	–5	–24.08 dB	4 right
1	0	1	0	–6	–30.10 dB	5 right
1	0	0	1	–7	–36.12 dB	6 right
1	0	0	0	–8	–42.14 dB	7 right

The value of Y is a linear representation of a gain change of up to –6 dB. Y is considered to be an unsigned fractional integer, with a leading value of 1, or: 0.1 Y4 Y5 Y6 Y7 (base 2). Y can represent values between 0.111112 (or 31/32) and 0.100002 (or 1/2). Thus, Y can represent gain changes from –0.28 dB to –6.02 dB. The combination of X and Y values allows compr to indicate gain changes from 48.16 – 0.28 = +47.89 dB, to –42.14 – 6.02 = –48.16 dB.

Annex B:
Dolby E Metadata Serial Protocol Chart

The following chart is provided to help the reader visualize the structure of the metadata protocol.

Appendix B: Dolby E MetaData Protocol Layout



Appendix B: Dolby E MetaData Protocol Layout

data_segment_id
values from page 1

data_segment_id = 1
data_segment_id = 2
data_segment_id = 3
data_segment_id = 4
data_segment_id = 5
data_segment_id = 6

1 - Dolby E Complete Metadata Data Segment

* Channel sequence will be derived from the current Program Configuration

	6 bits	program_config
	4 bits	frame_rate_code
	12 bits	pitch_shift_code
	64 bits	SMPTE_time_code
	8 bits	Reserved
produced once for every program present in the Dolby E stream	8 bits	description_text [pgm]
	2 bits	Reserved [pgm]
produced once for every channel present in the Dolby E stream*	4 bits	Reserved [ch]
	1 bit	Reservede [ch]
produced once for every channel present in the Dolby E stream*	10 bits	Reserved [ch]
	10 bits	Reserved [ch]

2 - Dolby E Essential Metadata Data Segment

6 bits	program_config
4 bits	frame_rate_code
12 bits	pitch_shift_code
8 bits	Reserved

4 & 6 - Dolby Digital Essential Program Metadata Data Segment

	5 bits	program_id
Each program in the Dolby E stream has an AC-3 data segment	5 bits	ac3_datarate
	3 bits	ac3_bsmod
	3 bits	ac3_acmod
	1 bit	ac3_lfeon
	5 bits	ac3_dialnorm
	1 bit	ac3_compre
	8 bits	ac3_compr2
	1 bit	ac3_dynrnge
	8 bits	ac3_dynrng5
	8 bits	ac3_dynrng6
	8 bits	ac3_dynrng7
	8 bits	ac3_dynrng8

Page 2 of 3

file name: Metadata chart pg 2 - 070426.vsd

Appendix B: Dolby E MetaData Protocol Layout

data_segment_id = 1
data_segment_id = 2
data_segment_id = 3
data_segment_id = 4
data_segment_id = 5
data_segment_id = 6

data_segment_id values from page 1

Page 3 of 3

5 - Dolby Digital Complete Program Metadata Data Segment No Extended BSI

5 bits	program_id
5 bits	ac3_datarate
3 bits	ac3_bsmod
3 bits	ac3_acmod
2 bits	ac3_cmixlev
2 bits	ac3_surmixlev
2 bits	ac3_dsurmod
1 bit	ac3_lfeon
5 bits	ac3_dialnorm
1 bit	ac3_langcode
8 bits	ac3_langcod
1 bit	ac3_audprodie
5 bits	ac3_mixlevel
2 bits	ac3_roomtyp
1 bit	ac3_copyrightb
1 bit	ac3_origbs
1 bit	ac3_timecod1e
14 bits	ac3_timecod1
1 bit	ac3_timecod2e
14 bits	ac3_timecod2
1 bit	ac3_hpfon
1 bit	ac3_bwlpfon
1 bit	ac3_lfelpfon
1 bit	ac3_sur90on
1 bit	ac3_suratton
1 bit	ac3_rfpremphon
1 bit	ac3_compre
8 bits	ac3_compr1
1 bit	ac3_dynrnge
8 bits	ac3_dynrng1
8 bits	ac3_dynrng2
8 bits	ac3_dynrng3
8 bits	ac3_dynrng4
1 bit	Reserved

Each program in the Dolby E stream has an AC-3 data segment

3 - Dolby Digital Complete Program Metadata Data Segment Extended BSI

Re-definition of these frame elements signaled by the data_segment_id. All other frame elements are as shown at left

1 bit	ac3_xbsi1e
2 bits	ac3_dmixmod
3 bits	ac3_ltrtcmixlev
3 bits	ac3_ltrtsurmixlev
3 bits	ac3_lorocmixlev
3 bits	ac3_lorosurmixlev
1 bit	ac3_xbsi2e
2 bits	ac3_dsurexmod2
2 bits	ac3_dheadphonmod
1 bit	ac3_adconvtyp
8 bits	Reserved
1 bit	Reserved

file name: Metadata chart pg 3 - 070426.vsd