SMPTE REGISTERED DISCLOSURE DOCUMENT

Format for Non-PCM Audio and Data in AES3 — Dolby-E® Data Type

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.

This document defines the Dolby-E mapping onto AES-3 according to SMPTE ST 337 and SMPTE ST 339, and to be a proper reference in SMPTE ST 338 for code point 28.

Proponent contact information:

Jeffrey Riedmiller
Dolby Laboratories
100 Potrero Avenue
San Francisco, CA 94103
USA

Email: JCR@dolby.com

Approved
March 27, 2015

Table of Contents

Page

1 Scope

This Registered Disclosure Document (RDD) describes how Dolby® E data rate reduced (non-PCM) audio streams are packed into an AES3 data stream, following the methods described by SMPTE ST 337. The Dolby E data type is identified by a data type number listed in SMPTE ST 338 and carried in the burst information word of the data burst preamble. The RDD describes how the Dolby E data is packed into the AES3 data stream.

Note: "Dolby", and the double -D symbol are trademarks of Dolby Laboratories.

2 Normative References

The following standards contain provisions which, through reference in this text, constitute provisions of this recommended practice. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this recommended practice are encouraged to investigate the possibility of applying the most recent edition of the standards indicated below.

Dolby Laboratories, “Dolby E Professional Encoder Development Manual”

SMPTE ST 337:2015, Format for Non-PCM Audio and Data in an AES3 Serial Digital Audio Interface

SMPTE ST 339:2015, Format for Non-PCM Audio and Data in AES3 — Generic Data Types

3 Definitions and Acronyms

3.1 Definitions

3.1.1 Latency

Delay time of an external audio decoder to decode a Dolby E data burst, defined as the sum of two values; the receiving delay time and the decoding delay time.

3.1.2 Fractional frame rates

Dolby E supports two fractional audio frame rates. These frame rates are written in shorthand notation, as specified in Table 1.

Table 1 – Shorthand notation for fractional frame rates

Fractional Dolby E audio frame rate (fps)	Shorthand version
$24 \times {}^{1\,000}/_{1\,001}$	23.976
$30 \times {}^{1\,000}/_{1\,001}$	29.97

Note: Dolby E is usable with video systems running at other frame rates, such as 60/1.001 fps, by using the methods documented in SMPTE RDD 19.

3.2 Acronyms

3.2.1 fps

frames per second

3.2.2 PCM

Pulse Code Modulation

4 Dolby E

4.1 Overview

Dolby E coded audio, as defined in the Dolby E Professional Encoder Development Manual, shall be transported in an AES3 data stream as a series of Data Bursts. Each Data Burst shall start with a Burst Preamble as defined by SMPTE ST 337, containing information about the Burst Payload, which shall follow the Burst Preamble. The Burst Payload shall consist of a Dolby E frame. The Burst Payload shall be followed by enough padding words (which shall be PCM zeros, or digital silence) to make the resulting Data Burst duration exactly match the duration in samples of baseband (PCM) audio that the Dolby E coded audio represents.

The resulting Data Bursts shall be placed in the audio sample word/aux data fields of AES3 subframes at regular intervals in either the frame or subframe mode (see SMPTE ST 337, Section 5). Data Bursts shall be placed in the AES3 transport, using either 16, 20, or 24 bits of the available data space.

A single Dolby E Frame shall form the Burst Payload, as shown in Figure 1.

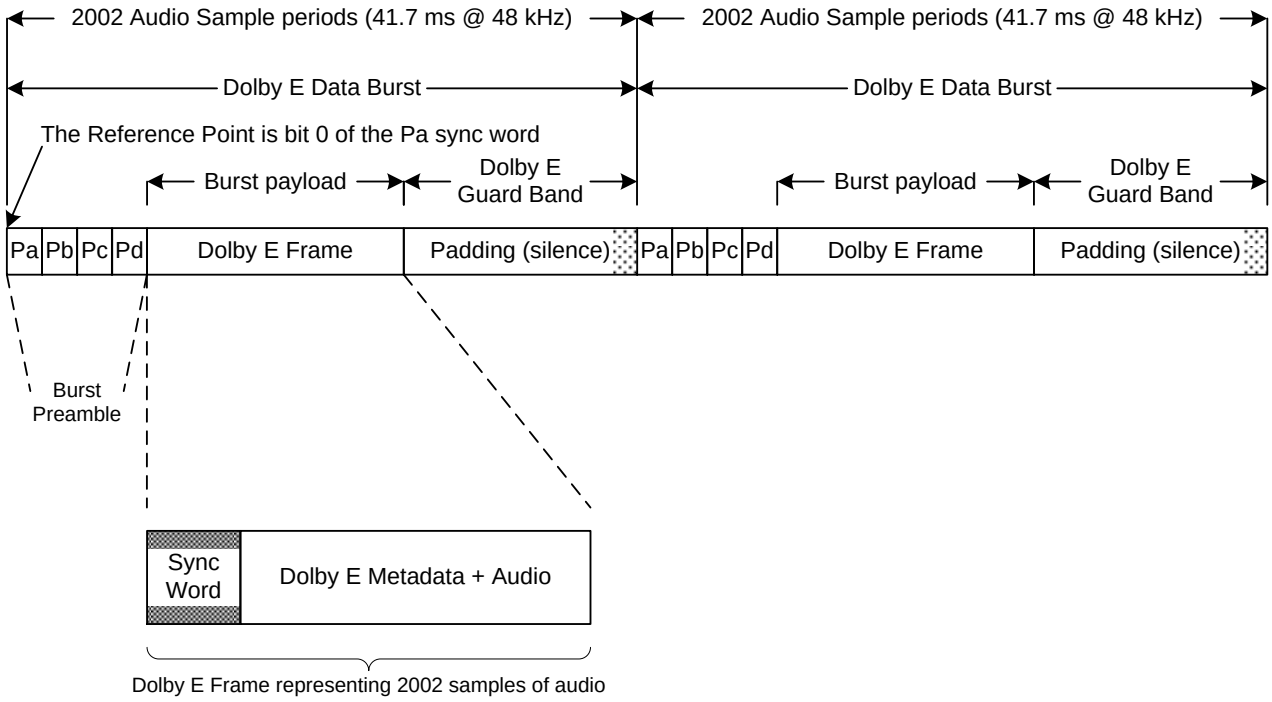

Figure 1 – Structure of a Dolby E Data Burst (Dolby E frame rate = 23.976 fps)

4.2 burst_preamble

The Pc word (burst_info value) of the burst_preamble carries the data_type identifier, the data_type_dependent and the data_stream_number information (see SMPTE ST 337, Table 7).

4.3 data_type identifier

The data_type identifier shall be set to 28.

4.4 data_type_dependent

The data_type_dependent bits shall be set to 0.

4.5 data_stream_number

The data_stream_number shall be set to 0h.

4.6 Dolby E burst_payload

A Dolby E bit stream consists of a sequence of Dolby E frames. The Dolby E burst-payload shall consist of a single Dolby E frame. The length of the Dolby E data-burst will depend on the encoded bit rate (which determines the Dolby E frame length).

4.7 Dolby E Sampling Frequency

When Dolby E data are conveyed by the AES3 interface, the AES3 frame frequency shall be equal to the sampling frequency of the Dolby E encoded audio. Bits 24-27 of the channel status word shall indicate the sampling frequency.

4.8 Dolby E Reference Point

The reference point of a Dolby E data burst is defined as bit 0 of the burst_preamble, i.e. the start of the data burst (Pa = sync word 1).

4.9 Dolby E Standard Repetition Rate

Dolby E data bursts are nominally placed in the AES3 interface such that the reference points of consecutive data bursts occur at a standard repetition rate. The standard repetition rate is defined as the number of AES3 frames that correspond to one video frame of an associated video reference signal corresponding to the Dolby E encoded frame rate. Table 2 shows the relationship between the Dolby E frame rate and the number of AES3 frames that correspond to each frame rate at a sampling frequency of 48 kHz.

Table 2 – Dolby E standard repetition rate for each frame rate

Dolby E Frame Rate (fps)	Dolby E Standard Repetition Rate
23.976	2,002 AES3 frames
24	2,000 AES3 frames
25	1,920 AES3 frames
29.97	1,601.6 AES3 frames
30	1,600 AES3 frames

At a frame rate of 29.97 fps frame rate, the defined repetition rate is not an integer number of AES3 frames due to the relationship of AES3 frames to video frames. In practice the spacing between data bursts will vary between 1,601 and 1,602 samples. For Dolby E bit streams that have a frame rate of 29.97 fps, it is important to maintain precise alignment with the corresponding video over 8,008 AES3 frames, which corresponds to five video frames at 29.97 fps. This is generally accomplished by repeating a decoded PCM output sample count of 1,602/1,601/1,602/1,601/1,602 every five frame periods.

4.10 Dolby E Standard Decode Latency

The Dolby E reference decode latency is defined as the time equivalent to one Dolby E frame period at the Dolby E frame rate. This means that a reference decoder would output the first PCM sample encoded in a Dolby E frame exactly one Dolby E frame period after the first bit of the frame is received by the decoder. Table 3 shows the Dolby E standard decode latency for each frame rate.

Table 3 – Dolby E standard decode latency for each frame rate

Dolby E Frame Rate (fps)	Dolby E standard decode latency	
23.976	2,002 AES3 frames	41.7 ms
24	2,000 AES3 frames	41.67 ms
25	1,920 AES3 frames	40 ms
29.97	1,601.6 AES3 frames	33.37 ms
30	1,600 AES3 frames	33.33 ms

4.11 Dolby E Reference Position

The Reference Position of a Burst Payload is defined by the relationship of the decoded audio to the associated video signal. A Burst Payload is in the Reference Position when the decoded audio from that Burst Payload is in sync with the associated video. The Dolby E reference position is defined in relation to the video frame that corresponds to the Dolby E-encoded audio samples.

A Dolby E data burst is defined as occupying the reference position in the AES3 stream if at least one of the following requirements is met:

1. **Ideal reference position:** The ideal reference position occurs in the AES3 stream at the first AES3 frame after the start of the defined video line of a video reference signal that corresponds to the Dolby E frame. The ideal Dolby E reference line position is specified in the “Ideal Position ±80µS” column of Table 4.

2. **Acceptable reference position:** the Dolby E data burst is considered to be in an acceptable reference position if the reference position is located ±80 microseconds from the ideal position. The acceptable reference positions are specified in the “Ideal Position -80µS” and “Ideal Position +80µS” column of Table 4.

3. **Safe reference position:** the Dolby E data burst is considered to be in a safe position (meaning no Dolby E frame data will be corrupted at a switch) if the reference point is within the following range:

    a. Earliest valid position: the reference point is immediately after the defined video switch point, as specified in the Earliest Valid Position column of Table 4.

    b. Latest valid position: the reference point ensures that the Dolby E frame data ends before the end of the video frame, as specified in the Latest Valid Position column of Table 4.

Table 4 – Dolby E reference position per video format

Video Format	Frame Rate (fps)	Total Number of Lines	Earliest Valid Position		Ideal Position						Latest Valid Position	
					-80µS		±80µS		+80µS			
			Line	µS	Line	µS	Line	µS	Line	µS	Line	µS
625/50/i	25	625	8	450	11	650	12	730	13	810	30	1860
525/59.94/i	29.97	525	12	510	13	610	14	690	16	770	26	1400
Dolby Black/23.98	23.98	693	10	390	18	890	20	970	21	1050	63	3610
Dolby Black/24	24	625	9	390	17	890	18	970	19	1050	58	3610
Dolby Black/30	30	500	11	510	13	610	14	690	15	770	26	1530
1920x1080/60/i	30	1125	18	510	21	610	24	690	26	770	52	1530
1920x1080/59.94/i	29.97	1125	18	510	21	610	24	690	26	770	48	1400
1920x1080/50/i	25	1125	13	450	19	650	21	730	23	810	53	1860
1920x1080/48/i	24	1125	11	390	25	890	27	970	29	1050	98	3610
1920x1080/47.97/i	23.98	1125	11	390	25	890	27	970	29	1050	98	3610
1920x1080/60/p	30	1125	35	510	42	610	47	690	52	770	104	1530
1920x1080/59.94/p	29.97	1125	35	510	42	610	47	690	52	770	95	1400
1920x1080/50/p	25	1125	26	450	37	650	42	730	46	810	105	1860
1920x1080/30/p	30	1125	18	510	21	610	24	690	26	770	52	1530
1920x1080/29.97/p	29.97	1125	18	510	21	610	24	690	26	770	48	1400
1920x1080/25/p	25	1125	13	450	19	650	21	730	23	810	53	1860
1920x1080/24/p	24	1125	11	390	25	890	27	970	29	1050	98	3610
1920x1080/23.98/p	23.98	1125	11	390	25	890	27	970	29	1050	98	3610
1280x720/60/p	30	750	23	510	28	610	32	690	35	770	69	1530
1280x720/59.94/p	29.97	750	23	510	28	610	32	690	35	770	63	1400
1280x720/50/p	25	750	17	450	25	650	28	730	31	810	70	1860
1280x720/30/p	30	750	12	510	14	610	16	690	18	770	35	1530
1280x720/29.97/p	29.97	750	12	510	14	610	16	690	18	770	32	1400
1280x720/25/p	25	750	9	450	13	650	14	730	16	810	35	1860
1280x720/24/p	24	750	8	390	17	890	18	970	19	1050	65	3610
1280x720/23.98/p	23.98	750	8	390	17	890	18	970	19	1050	65	3610
µS = Microseconds												

The Dolby E data burst shall be coincident in time with the video frame that corresponds to the audio samples coded within the Dolby E frame.

Note: Requirements 1 and 2 reference a defined video reference signal. Requirement 3 references a content video signal that may be the video reference signal. If not, it is assumed that the Dolby E burst is locked to the video reference signal. See SMPTE RDD 19 for additional guidance regarding use of Dolby E with video systems operating at frame rates higher than 30 Hz.

Note: Requirements 1 and 2 meet the basic Dolby E phase synchronization requirement for AES3 transport applications. Requirement 4 defines the reference location for lip sync.

4.12 Dolby E Guard Band

To allow a video stream to be cut-edited or switched while ensuring that the accompanying Dolby E data remains error-free, a period of AES3 null data precedes the Dolby E data burst. This null data is referred to as the Dolby E guard band. The guard-band location and duration is chosen so that Dolby E data is not present on the AES3 interface during video lines in which switches or edits will occur, and so that some delay (a few lines) can be tolerated in the audio path.

The portion of the guard band that precedes the Dolby E data burst begins at the AES3 sample that is aligned to the SMPTE RP 168 reference point, and ends at the start of the Dolby E data burst at the Dolby E reference position (see Table 4). The size and location of the guard band are determined by the synchronization requirements specified in Section 5.7, and the size of the encoded Dolby E frame. The total duration of the Dolby E guard band for each frame rate is shown in Table 5.

Table 5 – Dolby E guard band duration

Dolby E Frame Rate (fps)	Total guard band duration (AES3 frames)
23.976	100
24	100
25	96
29.97	80
30	80

Figure 2 shows the alignment of a Dolby E data burst from a 25 fps Dolby E stream, and a 625 line, 25 fps video stream. In this case, the Dolby E reference position is at the Ideal Position as specified in Table 4.

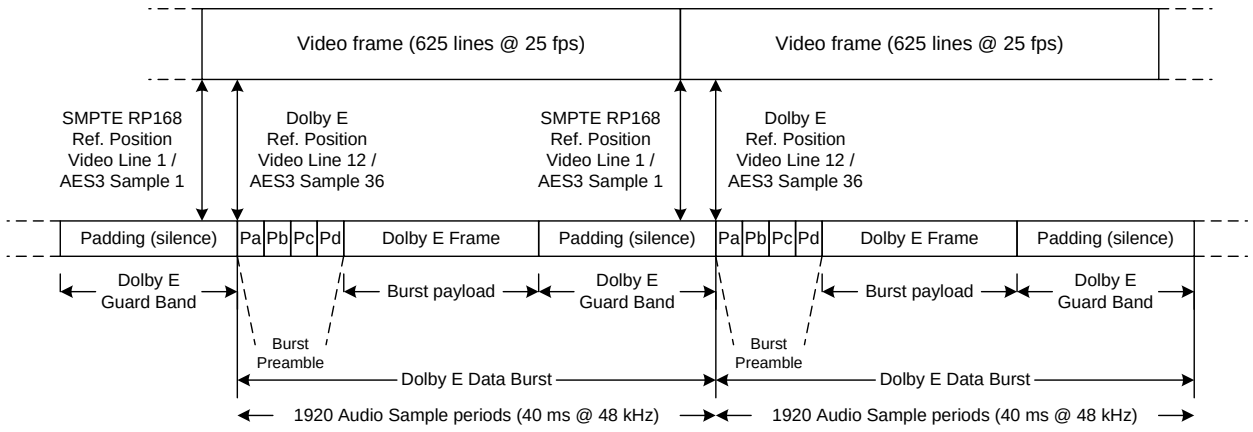

Figure 2 – 25 fps Dolby E data burst alignment with 625 line video @ 25 fps

Note: The Dolby E guard band applies to Dolby E data bursts only. Other data burst types may appear during the Dolby E guard-band interval. For instance, captioning data, timecode, or other types of low-bit-rate data may be carried in the space not used by Dolby E. In general, this data has a higher chance of being corrupted during a switch or cut edit.

4.13 Use of Pause Data Bursts between Dolby E Data Bursts

When a stream gap in a Dolby E stream is filled by a sequence of pause data-bursts, the Pa of the first pause data-burst shall be located one frame repetition rate following the Pa of the previous Dolby E data burst. The pause data-bursts shall be transmitted with a repetition rate of 3 AES3 frames, unless other repetition rates are necessary to precisely fill the stream gap (whose length may not be a multiple of 3 AES3 frames), subject to the zero padding requirements of SMPTE ST 337, Section 7.3.

When it is possible to control the length of the gap in a Dolby E data stream, it is recommended that the gap length be an integer multiple of the Dolby E audio frame duration (2 002, 2 000, 1 920, etc. AES3 frames), again respecting SMPTE ST 337, Section 7.3. This allows Dolby E decoders to optimize the gap concealment process.

The sequence of Pause data bursts may be interrupted to allow other data bursts to be multiplexed into the AES3 transport. See SMPTE ST 339 for details of the Pause data type and its use.

Annex A Bibliography (Informative)

Note: All references in this document to other SMPTE documents use the current numbering style (e.g. SMPTE ST 272:2004) although, during a transitional phase, the document as published (printed or PDF) may bear an older designation (such as SMPTE 272M-2004). Documents with the same root number (e.g. 272) and publication year (e.g. 2004) are functionally identical.

AES3-2009, AES Standard for Digital Audio Engineering — Serial Transmission Format for Two-Channel Linearly Represented Digital Audio Data

SMPTE ST 12-1:2014, Time and Control Code

SMPTE ST 272:2004, Television — Formatting AES/EBU Audio and Auxiliary Data into Digital Video Ancillary Data Space

SMPTE ST 291-1:2011, Ancillary Data Packet and Space Formatting

SMPTE ST 299-1:2009, 24-Bit Digital Audio Format for SMPTE ST 292 Bit-Serial Interface

SMPTE ST 302:2007, Television — Mapping of AES3 Data into an MPEG-2 Transport Stream

SMPTE ST 338:-2015, Format for Non-PCM Audio and Data in AES3 — Data Types

SMPTE RDD 19:2011, Guidelines on the Use of Dolby® E with Video Signals at Frame Rates Greater than 30 Hz

SMPTE RP 168:2009, Definition of Vertical Interval Switching Point for Synchronous Video Switching