

SMPTE RECOMMENDED PRACTICE

Requirements for Equipment  
Compatibility with Non-PCM  
AES3 Streams



Table of Contents	Page
Foreword .....	2
Introduction .....	2
1 Scope .....	3
2 Conformance Notation .....	3
3 Normative References .....	3
4 Specifications .....	3
4.1 Data Integrity .....	4
4.2 AES3 Word Length .....	4
4.3 Channel Status Data .....	4
Annex A Additional Equipment Behavior (Informative) .....	5
A.1 Manual Select Option .....	5
A.2 Word Length Truncation .....	5
A.3 Signal Processing .....	5
A.4 Cut at Edit or Switching Point .....	5
A.5 Embedded Audio (Smpte 272M, SMPTE 299M) .....	6
A.6 Level Metering .....	6
A.7 Digital-to-Analog Conversion .....	6

## Foreword

SMPTE (the Society of Motion Picture and Television Engineers) is an internationally-recognized standards developing organization. Headquartered and incorporated in the United States of America, SMPTE has members in over 80 countries on six continents. SMPTE's Engineering Documents, including Standards, Recommended Practices and Engineering Guidelines, are prepared by SMPTE's Technology Committees. Participation in these Committees is open to all with a bona fide interest in their work. SMPTE cooperates closely with other standards-developing organizations, including ISO, IEC and ITU.

SMPTE Engineering Documents are drafted in accordance with the rules given in Part XIII of its Administrative Practices.

SMPTE Recommended Practice RP 2005 was prepared by Technology Committee A29.

## Introduction

This section is entirely informative and does not form an integral part of this Engineering Document.

The AES3 signal format is generally used to carry two channels of linear PCM audio. The AES3 specification allows the signal to carry other data in place of audio data. In this document, such other data will be referred to as "non-PCM data". SMPTE 337M describes general formatting requirements when carrying non-PCM audio and data in an AES3 digital audio bitstream. When equipment handles AES3 PCM audio, some types of processing of the AES3 payload are often acceptable or even desirable. Examples are sample rate conversion to synchronize the PCM sample rate to a video reference, or dithering when reducing PCM word length. In order for a non-PCM AES3 signal carrying data to successfully traverse an AES3 plant, all equipment in the signal path should deliver the AES3 data payload exactly as received, and not perform any operations on the data which change any of the bit patterns. Equipment which is compliant with this Recommended Practice can be used to handle non-PCM AES3 streams such as those defined in SMPTE 337M and its associated family of standards.

AES3 streams are composed of a sequence of frames. Each frame consists of two interleaved subframes, one for the Channel 1 data and the second for the Channel 2 data. Each subframe consists of a synchronization preamble, a 24-bit data payload, a Validity bit, a User bit, a Channel Status bit and a Parity bit. The payload can carry either linearly coded PCM audio or SMPTE 337M type data.

A sequence of 192 subframes makes up a channel status block of 192 bits (or 24 bytes) containing channel status data. The first bit of channel status data (byte 0, bit 0) indicates Professional use of the data when set to logical one, or consumer use when set to logical zero. The second bit (byte 0, bit 1) indicates whether the contents of the subframe contain linear PCM audio (when set to logical zero) or "Nonaudio" (data) when set to logical one.

The specifications below are intended to ensure that the data contents of a non-PCM AES3 stream<sup>1</sup> are not corrupted by any processes which may be acceptable for PCM signals, and which may be found in many types of equipment. The specifications below are necessary, but not necessarily sufficient, to guarantee successful operation with non-PCM AES3 streams. Annex A provides several examples and explanations to support the requirements of this document.

---

<sup>1</sup> AES3 calls non-PCM data "audio sample word(s) used for purposes other than linear PCM samples".

## 1 Scope

This Recommended Practice contains specifications and other information that will allow equipment to transport, record, or otherwise convey without corruption AES3 signals containing non-PCM data, including data streams formatted according to SMPTE 337M.

## 2 Conformance Notation

Normative text is text that describes elements of the design that are indispensable or contains the conformance language keywords: "shall", "should", or "may". Informative text is text that is potentially helpful to the user, but not indispensable, and can be removed, changed, or added editorially without affecting interoperability. Informative text does not contain any conformance keywords.

All text in this document is, by default, normative, except: the Introduction, any section explicitly labeled as "Informative" or individual paragraphs that start with "Note:"

The keywords "shall" and "shall not" indicate requirements strictly to be followed in order to conform to the document and from which no deviation is permitted.

The keywords "should" and "should not" indicate that, among several possibilities, one is recommended as particularly suitable, without mentioning or excluding others; or that a certain course of action is preferred but not necessarily required; or that (in the negative form) a certain possibility or course of action is deprecated but not prohibited.

The keywords "may" and "need not" indicate courses of action permissible within the limits of the document.

The keyword "reserved" indicates a provision that is not defined at this time, shall not be used, and may be defined in the future. The keyword "forbidden" indicates "reserved" and in addition indicates that the provision will never be defined in the future.

A conformant implementation according to this document is one that includes all mandatory provisions ("shall") and, if implemented, all recommended provisions ("should") as described. A conformant implementation need not implement optional provisions ("may") and need not implement them as described.

Unless otherwise specified the order of precedence of the types of normative information in this document shall be as follows. Normative prose shall be the authoritative definition. Tables shall be next, followed by formal languages, then figures, and then any other language forms.

## 3 Normative References

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

AES3-2003, AES Standard for Digital Audio — Digital Input-Output Interfacing — Serial Transmission Format for Two-Channel Linearly Represented Digital Audio Data

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

SMPTE 299M-2004, Television — 24-Bit Digital Audio Format for SMPTE 292 Bit-Serial Interface

SMPTE 337M-2000, Television — Format for Non-PCM Audio and Data in an AES3 Serial Digital Audio Interface

## 4 Specifications

An AES3 signal carrying non-PCM data in Professional applications is identified by the states of bits 0 and 1 of byte 0 of the Channel Status block (see § 4.3). The following specifications shall apply to either or both subframes of an AES3 signal used to carry non-PCM data<sup>2</sup>.

### 4.1 Data Integrity

Equipment shall not modify the data or perform any kind of signal processing that will alter the data contained in the AES3 24-bit subframe payloads, subject to § 4.2.

### 4.2 AES3 Word Length

Equipment which cannot pass the entire 24-bit payload contained in an AES3 subframe shall clearly specify the word lengths supported (usually 16 or 20 bits). The word length of the AES3 subframe contents shall not be truncated.

### 4.3 Channel Status Data

Bit 0 of byte 0 of the Channel Status data shall be set to logical one to indicate Professional use of the signal. Bit 1 of byte 0 of the AES3 Channel Status data shall be set to logical zero for channels which contain linear PCM audio, and shall be set to logical one when the channel contains non-PCM data.

---

<sup>2</sup> In some cases, AES3 channel status byte 0, bit 1 may not a reliable indicator of the presence of non-PCM data. See Annex A.1 for further information.

## **Annex A (Informative)**

### **Additional Equipment Behavior**

#### **A.1 Manual Select Option**

In some cases, equipment may receive AES3 signals from existing equipment not adhering to this recommended practice such that channel status bit 1 of byte 0 is set to logical zero indicating PCM audio even though the AES3 signal contains non-PCM data. When possible, it is recommended that equipment adhering to this recommended practice allow a manual select option to override the status of channel status bit 1 of byte 0 on received bit streams. It is also desirable for equipment to automatically detect the presence of non-PCM data independently of channel status bit 1 of byte 0. A description of a possible auto detection algorithm is provided in SMPTE 337M Appendix A.

#### **A.2 Word Length Truncation**

Many devices (such as VTRs or other storage devices) may have audio data paths restricted to less than 24 bit audio word lengths. Such devices may still be used to convey AES3 streams containing non-PCM data if the non-PCM data occupies less than the full 24 bits available in the AES3 subframe. In particular, SMPTE 337M supports 16- and 20-bit data modes which are compatible with devices limited to 16- and 20-bit audio word lengths. In these devices, the AES3 word length should be truncated such that the remaining audio bits are not altered in any way. For instance, devices restricted to 16-bit word lengths should remove the least significant 8 bits of the 24-bit audio data word (4 bits from the aux data field and 4 least significant bits from the audio data field) without altering the remaining 16 bits.

Users should note that truncation of linear PCM sample words is often done by adding dither at the point of truncation. This dither will affect the bit patterns of the linear PCM words, and will change the data patterns of non-PCM data, therefore dither should not be applied when truncating audio words containing non-PCM data. It should also be noted that internal conversions, such as fixed to floating point conversions, may take place in some equipment. If such conversions are performed on words containing non-PCM data, care should be taken that these conversions do not perform rounding or any other numeric processes that may alter the non-PCM data.

#### **A.3 Signal Processing**

Signal processing, by definition, will alter the data contained in the AES3 subframe payloads and destroy the data being carried. Any user adjustable gain controls on the equipment should be bypassed or locked to a value of exactly 1.0, and any signal processing nominally performed on linear PCM signals must be defeated. Examples of other kinds of signal processing functions that must be defeated are dithering, rounding, pitch correction, crossfades, fade in/outs, and any kind of fixed to floating point conversion processes that may alter the AES3 data. Sample rate conversion also changes the data, so must be avoided. This requirement may impose limitations on the usage of some non-PCM AES3 applications in certain areas where equipment which operates at different sample rates must be interconnected. It is the responsibility of equipment users to ensure that AES3 signals containing non-PCM data are properly synchronized to avoid corruption of that data.

#### **A.4 Cut at Edit or Switching Point**

Video tape recorders (VTR), video disk recorders (VDR), routers, switchers, or other devices, which edit or switch between two AES3 streams should perform an abrupt switch or butt splice between the two streams. Any PCM fade or crossfade operation (as is typically done with linear PCM signals to minimize clicks) must be defeated. The type of audio edit performed by a VTR or other recordable device must be set to a cut edit (or butt splice) equivalent to a crossfade time of zero.

## **A.5 Embedded Audio (SMPTE 272M, SMPTE 299M)**

In some applications, AES3 streams containing non-PCM data may be embedded in SDI signals according to SMPTE 272M or SMPTE 299M. Devices performing embedding or extraction (disembedding) of audio from SDI signals, or otherwise handling SDI signals containing non-PCM embedded data streams, shall follow the rules as described in this recommended practice; i.e., the contents of the non-PCM AES3 subframes shall not be altered. Note that as required in § 4, any sample rate conversion is to be disabled. In general, this implies operation at level "A", "B", or "C" (synchronous audio at 48 kHz) of SMPTE 272M.

In addition to the above requirements, SDI embedders and extractors should employ buffer management techniques that minimize alteration of the AES3 frame sequence (such as sample drops or repeats) due to buffer management or synchronization requirements. Embedders should maintain adequate buffer sizes to ensure that no AES3 frame slips occur once the operation of embedding an AES3 stream containing non-PCM data begins. Extractors should also maintain buffer sizes adequate to ensure that no AES3 frame slips occur once the operation of extracting an AES3 stream containing non-PCM data begins (assuming static conditions).

In the case in which a switch occurs in the video signal from which the audio is being extracted, any processing needed to manage the extractor buffer (such as AES3 frame drops or repeats) should be minimized. If required, it is desirable that buffer management be performed as close to the actual switch point as possible to minimize loss of data following the switch. If AES3 frame slips due to buffer management are required at points significantly later than the actual switch point (e.g., greater than 6 AES3 frames), it is desirable that extractors should implement "smart" buffer management techniques. In particular, it is desired that extractors be able to detect the presence of non-PCM data formatted according to SMPTE 337M, and to perform buffer management operations during "null data" periods so that no actual non-PCM data is lost or altered. An alternative technique is to perform buffer management operations during AES3 frames which are within 6 AES3 frames of the point defined by SMPTE RP 168 as the video frame or vertical interval switching point (e.g., video line 10 for NTSC).

In addition, it is recommended that buffer sizes within equipment performing SDI embedding and extraction be minimized as much as possible (within the constraints allowed) to reduce delay.

## **A.6 Level Metering**

Devices that include level meters for monitoring of signal levels of linear PCM audio contained within AES3 signals may encounter signals with non-PCM data. If possible, it is desirable that such devices should include a specific indication of the presence of non-PCM data, and not attempt to indicate level information, unless such devices are capable of decoding and properly representing the level audio signals encoded within non-PCM data streams. Users should be aware that many existing level meters may not provide this capability and will therefore display inaccurate level information, typically showing high level or clipped signals when non-PCM data is present; this behavior does, however, give the knowledgeable user an indication that the signal being metered is a data signal and not a PCM audio signal.

## **A.7 Digital-to-Analog Conversion**

Devices receiving AES3 streams containing non-PCM data should not convert data signals to an analog signal for the same reasons that make level metering impossible. The validity bit carried in each subframe of AES3 data is generally used to indicate whether or not the "audio sample word is suitable for conversion to an analog signal". The validity bit is set to logical one if the data is not suitable for conversion. Receiving devices should mute the output of any D/A converter receiving a data stream.

Note that the validity bit may be used to indicate the presence of non-audio data until the net complete channel status block is transmitted, but it should not be used as a substitute for channel status bit 1, byte 0.