

**AES3-1992 (r1997)**  
**Revision of AES3-1985**  
**Amendment 1-1997**  
**Amendment 2-1998**  
**Amendment 3-1999**  
**Amendment 4-1999**

# **AES Recommended practice for digital audio engineering — Serial transmission format for two- channel linearly represented digital audio data**

*Published by*  
**Audio Engineering Society, Inc.**  
Copyright © 1992-2000 by the Audio Engineering Society

## **Abstract**

The format provides for the serial digital transmission of two channels of periodically sampled and uniformly quantized audio signals on a single shielded twisted wire pair. The transmission rate is such that samples of audio data, one from each channel, are transmitted in time division multiplex in one sample period. Provision is made for the transmission of both user and interface related data as well as of timing related data, which may be used for editing and other purposes. It is expected that the format will be used to convey audio data that have been sampled at any of the sampling frequencies recognized by the AES5, *Recommended Practice for Professional Digital Audio Applications Employing Pulse-Code Modulation — Preferred Sampling Frequencies*.

An AES standard implies a consensus of those directly and materially affected by its scope and provisions and is intended as a guide to aid the manufacturer, the consumer, and the general public. The existence of an AES standard does not in any respect preclude anyone, whether or not he or she has approved the document, from manufacturing, marketing, purchasing, or using products, processes, or procedures not in agreement with the Standard. Prior to approval, all parties were provided opportunities to comment or object to any provision. Approval does not assume any liability to any patent owner, nor does it assume any obligation whatever to parties adopting the standards document. This document is subject to periodic review and users are cautioned to obtain the latest edition.

## Contents

Foreword .....	4
Corrigendum 2000-10-11 .....	4
Amendment 1-1997 .....	5
Amendment 2-1998 .....	6
Amendment 3-1999 .....	7
Amendment 4-1999 .....	12
1 Scope .....	14
2 Interface format .....	15
2.1 Terminology .....	15
2.1.1 sampling frequency .....	15
2.1.2 audio sample word .....	15
2.1.3 auxiliary sample bits .....	15
2.1.4 validity bit .....	15
2.1.5 channel status .....	15
2.1.6 user data .....	15
2.1.7 parity bit .....	15
2.1.8 preambles .....	15
2.1.9 subframe .....	15
2.1.10 frame .....	15
2.1.11 block .....	15
2.1.12 channel coding .....	16
2.1.13 unit interval (UI) .....	16
2.1.14 interface jitter .....	16
2.1.15 intrinsic jitter .....	16
2.1.16 jitter gain .....	16
2.1.17 frame rate .....	16
2.2 Structure of format .....	16
2.2.1 Subframe format .....	16
2.2.2 Frame format .....	17
2.3 Channel coding .....	18
2.4 Preambles .....	18
2.5 Validity bit .....	19
3 User data format .....	19
4 Channel status format .....	19
5 Interface format implementation .....	25
5.1 General .....	25
5.2 Transmitter .....	25
5.2.1 "Minimum" implementation of channel status .....	26
5.2.2 "Standard" implementation of channel status .....	26
5.2.3 "Enhanced" implementation of channel status .....	26
5.3 Receivers .....	26
6 Electrical requirements .....	26
6.1 General Characteristics .....	26
6.2 Line driver characteristics .....	27
6.2.1 Output impedance .....	27
6.2.2 Signal amplitude .....	27
6.2.3 Balance .....	27
6.2.4 Rise and fall times .....	27
6.2.5 Output interface jitter .....	27
6.2.5.1 Intrinsic jitter .....	27
6.2.5.2 Jitter gain .....	28
6.3 Line receiver characteristics .....	29
6.3.1 Terminating impedance .....	29

6.3.2 Maximum input signals .....	29
6.3.3 Minimum input signals .....	29
6.3.4 Receiver equalization .....	29
6.3.5 Common-mode rejection .....	30
6.3.6 Receiver jitter tolerance .....	30
6.4 Connectors .....	30
7 Normative References .....	31
ANNEX A Provision of additional, voice-quality channels via the digital audio interface .....	33
ANNEX B Generation of CRCC (byte 23) for channel status .....	34
Annex C Informative references .....	36

## Foreword

[This foreword is not a part of the *AES Recommended practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data*, AES3-1992.]

This document discusses the format and line protocols for a revision of the AES recommendation, originally published in 1985, for the serial transmission format for linearly represented digital audio data over conventional shielded twisted-pair conductors, of up to at least 100 m in length, without equalization. The organization and style of the revised document are patterned after portions of International Electrotechnical Commission (IEC) Publication 958 and International Radio Consultative Committee (CCIR) Recommendation 647, which are well known in the international technical community.

It has been six years since AES3-1985 was adopted as a standard, and much experience with equipment, installations, and applications in professional audio and broadcasting has been accumulated. AES3 has been widely accepted as the primary means of transmitting digital audio for two-channel and multichannel (by combinations of connections) professional and broadcast studio use. Another standard, AES10, has been recently adopted for multichannel use and will in the future provide a more efficient means of transmission for a large number of channels (56). AES10 is based on and designed to be compatible with AES3 for transmitted data, the terminology associated with the data, and the intended use of the data. Also AES11, *Synchronization of digital audio equipment in studio operations*, refers to a signal conforming to this revised form of AES3. Applications and uses of one or two channels of digital audio as supported by this revised version of AES3 will remain important and numerous.

The revision is intended to simplify and clarify language, improve electrical performance, minimize confusion with the IEC Publication 958 "consumer use" specification, allocate certain previously reserved bits to new applications, and improve compatibility by improving uniformity of transmitter implementation in regard to validity, user, channel status, and parity bits. To further facilitate adoption of this standard for the diverse applications and conditions for which it is intended, a separate engineering guideline document — not part of this standard — is in preparation.

AES3 has been under constant review since the standard was issued, and the present document reflects the collective experience and opinions of many users, manufacturers, and organizations familiar with equipment or systems employing AES3. Experience includes operation in locations such as large broadcast centers, small recording studios, and field operations. This revision was written in close cooperation with the European Broadcasting Union (EBU). At the time it was written, the AES Working Group on Digital Input/Output Interfacing included the following individuals who contributed to this standard: T. Attenborough, B. Blüthgen, S. Busby, D. Bush, R. Cabot, R. Caine, C. Cellier, S. Culnane, A. Fasbender, R. Finger, B. Fletcher, B. Foster, N. Gilchrist, T. Griffiths, R. Hankinson, S. Herla, R. Hoffner, B. Hogan, T. Holman, Y. Ishida, C. Jenkins, T. Jensen, A. Jubb, A. Komly, R. Lagadec, P. Lidbetter, B. Locanthi, S. Lyman, L. Moller, A. Mornington-West, C. Musialik, J. Nunn, D. Queen, C. Sanchez, J. Schuster, T. Setogawa, T. Shelton, S. Shibata, A. Swanson, A. Viallevieille, D. Walstra, J. Wilkinson, and P. Wilton.

Robert A. Finger, *Chairman*  
Working Group SC-2-2 Digital Input/Output Interfacing  
1991 March

### Corrigendum 2000-10-11

in annex B, figure B.1, lower half marked CRCC, at the right start of the dashed line annotated with  $x^8$ , a missing zero was added;

in clause 4, byte 1, bits 0-3, bit state 1111, "not be used" was corrected to "not to be used";

in clause 4, byte 1, bits 4-7, last line, "not be used" was corrected to "not to be used";

in annex C, SMPTE, "and television Engineers" was corrected to "and Television Engineers"

in Amendment 4-1999, "and television Engineers" was corrected to "and Television Engineers"

### Corrigendum 2001-05-25

in figure 6 the missing connection for the cable shield to the line driver has been replaced

This printing of AES3-1992 incorporates Amendment 1-1997, Amendment 2-1998, Amendment 3-1999, and Amendment 4-1999 as shown in the following text. It has been repaginated accordingly but has not been updated to current AES style. All clause, table, and figure numbering has been retained. The American National Standards Institute version of this standard has not been amended and remains available as S4.40-1992.

#### **Amendment 1-1997**

[Amend AES3-1992, *AES Recommended practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data* by adding new subclauses 2.1.13 to 2.1.16. In the next revision of AES3, these subclauses will become part of clause 3, according to current AES style.]

##### **2.1.13**

##### **unit interval(UI)**

Shortest nominal time interval in the coding scheme.

NOTE – There are 128 UI in a sample frame.

##### **2.1.14**

##### **interface jitter**

Deviation in timing of interface data transitions (zero crossings) when measured with respect to an ideal clock.

##### **2.1.15**

##### **intrinsic jitter**

Output interface jitter of a device that is either free-running or is synchronized to a jitter-free reference.

##### **2.1.16**

##### **jitter gain**

Ratio, expressed in decibels, of the amplitude of jitter at the synchronization input of a device to the resultant jitter at the output of the device.

NOTE – This definition excludes the effect of intrinsic jitter.

[Further amend AES3-1992, *AES Recommended practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data* by deleting old subclause 6.2.5 and adding new subclause 6.2.5 with figures 9 and 10. In the next revision of AES3, the subclauses and figures may be renumbered according to current AES style. Deletions are shown by strikethrough.]

#### **~~6.2.5 Data jitter~~**

~~Data transitions shall occur within  $\pm 20$  ns of an ideal jitter-free clock measured at the half-voltage points.~~

~~NOTE – This specification applies only to the signal after channel coding. Tighter specifications apply to the audio sample clock.~~

#### **6.2.5 Output interface jitter**

Jitter at the output of a device shall be measured as the sum of the jitter intrinsic to the device and jitter being passed through from the timing reference of the device.

### 6.2.5.1 Intrinsic jitter

The peak value of the intrinsic jitter at the output of the interface, measured at all the transition zero crossings shall be less than 0.025 UI when measured with the intrinsic-jitter measurement filter.

NOTE 1 – This requirement applies both when the equipment is locked to an effectively jitter-free timing reference (which may be a modulated digital audio signal) and when the equipment is free-running.

NOTE 2 – The intrinsic-jitter measurement-filter characteristic is shown in figure 9. It shows a minimum-phase high-pass filter with 3 dB attenuation at 700 Hz, a first order roll-off to 70 Hz and with a pass-band gain of unity.

**Figure 9. Intrinsic-jitter measurement-filter characteristic**

### 6.2.5.2 Jitter gain

The sinusoidal jitter gain from any timing reference input to the signal output shall be less than 2 dB at all frequencies.

NOTE – If jitter attenuation is provided and it is such that the sinusoidal jitter gain falls below the jitter transfer function mask of figure 10 then the equipment specification should state that the equipment jitter attenuation is within this specification. The mask imposes no additional limit on low-frequency jitter gain. The limit starts at the input-jitter frequency of 500 Hz where it is 0 dB, and falls to –6 dB at and above 1 kHz.

**Figure 10. Jitter transfer-function mask**

[Further amend AES3-1992, *AES Recommended practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data* by adding new subclause 6.3.6 and figure 11. In the next revision of AES3, this subclause and figure may be renumbered according to current AES style.]

### 6.3.6 Receiver jitter tolerance

An interface data receiver should correctly decode an incoming data stream with any sinusoidal jitter defined by the jitter tolerance template of figure 11.

NOTE – The template requires a jitter tolerance of 0.25 UI peak-to-peak at high frequencies, increasing with the inverse of frequency below 8 kHz to level off at 10 UI peak-to-peak below 200 Hz.

**Figure 11. Jitter tolerance template**

### Amendment 2-1998

[Amend AES3-1992, *AES Recommended practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data* by revising the byte 0 table of clause 4.

In the next revision of AES3, this clause will become clause 5, according to current AES style and the table will be reformatted. Changes are underlined and deletions are shown by strikethrough.]

bit 1 0 ~~Normal audio mode.~~  
1 ~~Nonaudio mode.~~

bit 1 0 Audio sample word represents linear PCM samples.  
1 Audio sample word used for purposes other than linear PCM samples.

bit 5 ~~+~~ Source sampling frequency unlocked.  
0 ~~Default and source sampling frequency locked.~~

bit 5 0 Default. Lock condition not indicated.  
1 Source sampling frequency unlocked.

[Amend AES3-1992, *AES Recommended practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data* by adding a state to the bits 4-7 {corrected 1998-06-12, was 0-3} listing in the byte 1 table of clause 4. In the next revision of AES3 this clause will become clause 5 according to current AES style and the table will be reformatted. The addition is underlined.

NOTE – This addition references a clause of an International Electrotechnical Commission (IEC) standard at the stage of committee draft for voting at the time of first publication of this amendment. This amendment is adopted pending approval of the revised IEC standard. Following IEC approval of 60958-3, this note will be deleted.]

0 1 0 0 User data conforms to the general user data format defined in IEC 60958-3

### **Amendment 3-1999**

[Amend AES3-1992, *AES Recommended practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data* by modifying the second and third paragraphs of clause 1 Scope. Changes are underlined and deletions are shown by ~~striethrough~~.]

It is expected that the format will be used to convey audio data that have been sampled at any of the sampling frequencies recognized by the *AES5 Recommended Practice for Professional Digital Audio Applications Employing Pulse-Code Modulation-Preferred Sampling Frequencies*. Note that conformance with this interface specification does not require equipment to utilise these rates. Also the capability of the interface to indicate other sample rates does not imply that it is recommended that equipment supports these rates. To eliminate doubt equipment specifications should define supported sampling frequencies.

The format is intended for use with shielded twisted-pair cable of conventional design over distances of up to 100 m without transmission equalization or any special equalization at the receiver and at frame rates of up to 50 kHz. Longer cable lengths and higher frame rates may be used, but with a rapidly increasing requirement for care in cable selection and possible receiver equalization or the use of active repeaters, or both.

[Amend AES3-1992, *AES Recommended practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data* by adding new subclause 2.1.17. In the next revision of AES3, this subclause will become part of clause 3, according to current AES style.]

#### **2.1.17 frame rate**

Rate of transmission of frames on the interface.

[Amend AES3-1992, *AES Recommended practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data* subclause 2.2.2. In the next revision of AES3, this subclause will be renumbered according to current AES style]

#### **2.2.2 Frame format**

A frame is uniquely composed of two subframes (see figure 2). Except where otherwise specified the rate of transmission of frames corresponds exactly to the source sampling frequency.

[Amend AES3-1992, *AES Recommended practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data* by adding two notes to clause 4 at the end of the specification for byte 0. In the next revision of AES3, this clause will be renumbered according to current AES style]

NOTE 1 – The significance of byte 0, bit 0 is...

NOTE 2 – The indication of sampling frequency, or the use of one of the sampling frequencies that can be indicated in this byte, is not a requirement for operation of the interface. The 00 state of bits 6-7 may be used if the transmitter does not support the indication of sampling frequency, the sampling frequency is unknown, or the sample frequency is not one of those that can be indicated in this byte. In the latter case for some sampling frequencies byte 4 may be used to indicate the correct value.

NOTE 3 – When byte 1, bits 1-3 indicate single channel double sampling frequency mode then the sampling frequency of the audio signal is twice that indicated by bits 6-7 of byte 0.



[Amend AES3-1992, *AES Recommended practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data* by modifying the state coding for byte 1 bits 0-3 in clause 4. In the next revision of AES3, this clause will be renumbered according to current AES style]

Byte 1

bits 0-3 Encoded channel mode

bit 0 1 2 3

state	0 0 0 0	Mode not indicated. Receiver default to two-channel mode. Manual override is enabled.
	0 0 0 1	Two channel mode. Manual override is disabled.
	0 0 1 0	Single channel mode. (Monophonic). Manual override is disabled.
	0 0 1 1	Primary/secondary mode. (subframe 1 is primary). Manual override is disabled.
	0 1 0 0	Stereophonic mode. (channel 1 is left channel). Manual override is disabled.
	0 1 0 1	Reserved for user defined applications.
	0 1 1 0	Reserved for user defined applications.
	0 1 1 1	<u>Single channel double sampling frequency mode. Sub-frames 1 and 2 carry successive samples of the same signal. The sampling frequency of the signal is double the frame rate, and is double the sampling frequency indicated in byte 0 (but not double the rate indicated in byte 4, if that is used). Manual override is disabled. Vector to byte 3 for channel identification.</u>
	1 0 0 0	<u>Single channel double sampling frequency mode - stereo mode left. Sub-frames 1 and 2 carry successive samples of the same signal. The sampling frequency of the signal is double the frame rate, and is double the sampling frequency indicated in byte 0 (but not double the rate indicated in byte 4, if that is used). Manual override is disabled.</u>
	1 0 0 1	<u>Single channel double sampling frequency mode - stereo mode right. Sub-frames 1 and 2 carry successive samples of the same signal. The sampling frequency of the signal is double the frame rate, and is double the sampling frequency indicated in byte 0 (but not double the rate indicated in byte 4, if that is used). Manual override is disabled.</u>
	1 1 1 1	<u>Multichannel mode. Vector to byte 3 for channel identification.</u> <del>Reserved for future applications.</del>

All other states of bits 0-3 are reserved and are not to be used until further defined.

[Amend AES3-1992, *AES Recommended practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data* by modifying the state coding and notes for byte 3 in clause 4. In the next revision of AES3, this clause will be renumbered according to current AES style.]

Byte

bits 0-7 ~~Vectored target byte from byte 1~~

bit ~~Reserved for future use as multichannel function description. These bits are set to logic 0 at present~~

state

bit 7	0	Undefined multichannel mode (default).
	1	<u>Defined multichannel modes.</u>

The definition of the remaining bit states depends on the state of bit 7

bits 0-6 Channel number (when byte 3 bit 7 is 0).

The channel number is the value of the byte (with bit 0 as the least significant bit) plus one.

bits 4-6 Multichannel mode (when byte 3 bit 7 is 1).

bit 4 5 6

	0 0 0	<u>Multichannel mode 0. The channel number is defined by bits 0-3 of this byte.</u>
	1 0 0	<u>Multichannel mode 1. The channel number is defined by bits 0-3 of this byte.</u>
	0 1 0	<u>Multichannel mode 2. The channel number is defined by bits 0-3 of this byte.</u>
	1 1 0	<u>Multichannel mode 3. The channel number is defined by bits 0-3 of this byte.</u>

1 1 1 User defined multichannel mode. The channel number is defined by bits 0-3 of this byte.

All other states of bits 4-6 are reserved and are not to be used until further defined.

bits 0-3 Channel number (when byte 3 bit 7 is 1).

The channel number is one plus the numeric value of value of these bits taken as a binary number (with bit 3 as the most significant bit).

NOTE 1 - The defined multichannel modes identify mappings between channel numbers and function. (The standard mappings are under consideration. Some mappings may involve groupings of up to 32 channels by combining two modes.)

NOTE 2 - For compatibility with equipment that is only sensitive to the channel status data in one subframe the channel carried by subframe 2 may indicate the same channel number as channel 1. In that case it is implicit that the second channel has a number one higher than the channel of subframe 1 except in single channel double sampling frequency mode.

NOTE 3 - When bit 7 is 1 the 4 bit channel number can be mapped to the channel numbering in bits 20-23 of the consumer mode channel status defined in IEC 60958-3. In this case channel A of consumer mode maps to channel 2, channel B maps to channel 3 and so on.

[Amend AES3-1992, *AES Recommended practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data* by modifying the state coding for byte 4 bits 2-7 in clause 4. In the next revision of AES3, this clause will be renumbered according to current AES style.]

bits 2-7 Reserved and set to logic 0 until further defined.

bit 2 Reserved.

bits 3-6 Sampling frequency.

bit 6 5 4 3

<u>state</u>	<u>0 0 0 0</u>	<u>Not indicated (default).</u>
	<u>0 0 0 1</u>	<u>24 kHz</u>
	<u>0 0 1 0</u>	<u>96 kHz</u>
	<u>0 0 1 1</u>	<u>192 kHz</u>
	<u>0 1 0 0</u>	<u>Reserved.</u>
	<u>0 1 0 1</u>	<u>Reserved.</u>
	<u>0 1 1 0</u>	<u>Reserved.</u>
	<u>0 1 1 1</u>	<u>Reserved.</u>
	<u>1 0 0 0</u>	<u>reserved (for vectoring).</u>
	<u>1 0 0 1</u>	<u>22.05 kHz</u>
	<u>1 0 1 0</u>	<u>88.2 kHz</u>
	<u>1 0 1 1</u>	<u>176.4 kHz</u>
	<u>1 1 0 0</u>	<u>Reserved.</u>
	<u>1 1 0 1</u>	<u>Reserved.</u>
	<u>1 1 1 0</u>	<u>Reserved.</u>
	<u>1 1 1 1</u>	<u>User defined .</u>

bit 7 Sampling frequency scaling flag.

0 No scaling (default).

1 Sampling frequency is 1/1.001 times that indicated by byte 4 bits 3-6, or by byte 0 bits 6-7.

NOTE 1 - The sampling frequency indicated in byte 4 is not dependent on the channel mode indicated in byte 1.

NOTE 2 – The indication of sampling frequency, or the use of one of the sampling frequencies that can be indicated in this byte, is not a requirement for operation of the interface. The '0000' state of bits 3-6 may be used if the transmitter does not support the indication of sampling frequency in this byte, the sampling frequency is unknown, or the sample frequency is not one of those that can be indicated in this byte. In the later case for some sampling frequencies byte 0 may be used to indicate the correct value.

NOTE 3 – The reserved states of bits 3-6 of byte 4 are intended for later definition such that bit 6 is set to define rates related to 44.1 kHz (except for state 1000) and clear to defined rates related to 48 kHz. They should not be used until further defined.

[Amend AES3-1992, *AES Recommended practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data* by adding a sentence to the first paragraph and modifying the final paragraph of subclause 6.1. In the next revision of AES3, this subclause will be renumbered according to current AES style]

The frequency range used to qualify the interface electrical parameters is dependent on the maximum data rate supported. The upper frequency is 128 times the maximum frame rate.

The interconnecting cable shall be balanced and screened (shielded) with a nominal characteristic impedance of 110  $\Omega$  at frequencies from ~~0.1~~ 100 kHz to ~~6.0 MHz~~ 128 times the maximum frame rate.

[Amend AES3-1992, *AES Recommended practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data* by modifying subclause 6.2.1. In the next revision of AES3, this subclause will be renumbered according to current AES style.]

### **6.2.1 Output impedance**

The line driver shall have a balanced output with an internal impedance of 110  $\Omega \pm 20\%$ , at frequencies from ~~0.1~~ 100 kHz to ~~6.0 MHz~~ 128 times the maximum frame rate.

[Amend AES3-1992, *AES Recommended practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data* by modifying subclause 6.2.3. In the next revision of AES3, this subclause will be renumbered according to current AES style.]

### **6.2.3 Balance**

Any common-mode component at the output terminals shall be more than 30 dB below the signal at frequencies from dc to ~~6 MHz~~ 128 times the maximum frame rate.

[Amend AES3-1992, *AES Recommended practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data* by modifying subclause 6.3.1. In the

next revision of AES3, this subclause will be renumbered according to current AES style.]

### 6.3.1 Terminating impedance

The receiver shall present an essentially resistive impedance of  $110\ \Omega \pm 20\%$  to the interconnecting cable over the frequency band from 0.1 MHz to 6.0 MHz 128 times the maximum frame rate when measured across the input terminals. The application ...

[Amend AES3-1992, *AES Recommended practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data* by modifying subclause 6.3.4. In the next revision of AES3, this subclause will be renumbered according to current AES style.]

### 6.3.4 Receiver equalization

Optional equalization can be applied in the receiver to enable an interconnecting cable longer than 100 m to be used. A suggested frequency equalization characteristic for operation at frame rates of 48 kHz is shown in figure 8. The receiver shall meet the characteristics specified in 6.3.2 and 6.3.3.

[Amend AES3-1992, *AES Recommended practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data* by modifying the caption for Figure 8.]

Figure 8. Suggested equalizing characteristic for a receiver operating at 48 kHz frame rate.

### Amendment 4-1999

[Amend AES3-1992, *AES Recommended practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data* by revising the byte 2 table of clause 4. In the next revision of AES3, this clause will become clause 5, according to current AES style and the table will be reformatted. Changes and additions are underlined and deletions are shown by ~~striketrough~~.]

Byte 2

bits 6-7 ~~Reserved and are set to logic 0 until further defined.~~

bits 6-7 Indication of alignment level.

bit 6 7

0 0 Alignment level not indicated

0 1 Alignment to SMPTE RP155 (alignment level is 20 dB below maximum code).

1 0 Alignment to EBU R68 (alignment level is 18.06 dB below maximum code).

1 1 Reserved for future use.

[Amend AES3-1992, *AES Recommended practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data* by adding two references to a new informative annex C]

**Annex C**  
**(Informative)**

SMPTE Recommended Practice RP155-1997 *Audio levels for digital audio records on digital television tape recorders*. Society of Motion Picture and Television Engineers. New York, NY, US.

EBU Technical Recommendation R68-1992 *Alignment level in digital audio production equipment and in digital audio recorders*. European Broadcasting Union, Geneva, Switzerland.

**Revision of AES3-1985****Amendment 1-1997****Amendment 2-1998****Amendment 3-1999****Amendment 4-1999**

**AES Recommended practice  
for digital audio engineering —  
Serial transmission format for two-channel linearly  
represented digital audio data**

**1 Scope**

This document specifies a recommended interface for the serial digital transmission of two channels of periodically sampled and linearly represented digital audio data from one transmitter to one receiver.

It is expected that the format will be used to convey audio data that have been sampled at any of the sampling frequencies recognized by the AES5 *Recommended Practice for Professional Digital Audio Applications Employing Pulse-Code Modulation — Preferred Sampling Frequencies*. Note that conformance with this interface specification does not require equipment to utilise these rates. Also the capability of the interface to indicate other sample rates does not imply that it is recommended that equipment supports these rates. To eliminate doubt equipment specifications should define supported sampling frequencies.

The format is intended for use with shielded twisted-pair cable of conventional design over distances of up to 100 m without transmission equalization or any special equalization at the receiver and at frame rates of up to 50 kHz. Longer cable lengths and higher frame rates may be used, but with a rapidly increasing requirement for care in cable selection and possible receiver equalization or the use of active repeaters, or both.

The document does not cover connection to any common carrier equipment, nor does it specifically address any questions about the synchronizing of large systems, although by its nature the format permits easy synchronization of receiving devices to the transmitting device.

Specific synchronization issues are covered in AES11.

In this interface specification, mention is made of an interface for consumer use. The two interfaces are not identical.

An engineering guideline document to accompany this interface specification is in preparation.

## **2 Interface format**

### **2.1 Terminology**

#### **2.1.1 sampling frequency**

Frequency of the samples representing an audio signal. When more than one audio signal is transmitted through the same interface, the sampling frequencies are identical.

#### **2.1.2 audio sample word**

Amplitude of a digital audio sample. Representation is linear in 2's complement binary form. Positive numbers correspond to positive analog voltages at the input of the analog-to-digital converter (ADC). The number of bits per word can be specified from 16 to 24 in two coding ranges (less than or equal to 20 bits and less than or equal to 24 bits).

#### **2.1.3 auxiliary sample bits**

4 least significant bits (LSBs) which can be assigned as auxiliary sample bits and used for auxiliary information when the number of audio sample bits is less than or equal to 20.

#### **2.1.4 validity bit**

Bit indicating whether the audio sample bits in the subframe (time slots 4 to 27 or 8 to 27, depending on the audio word length as described in 2.2.1) are suitable for conversion to an analog audio signal.

#### **2.1.5 channel status**

Bits carrying, in a fixed format derived from the block (see 2.1.11), information associated with each audio channel which is decodable by any interface user.

#### **2.1.6 user data**

Channel provided to carry any other information.

#### **2.1.7 parity bit**

Bit provided to permit the detection of an odd number of errors resulting from malfunctions in the interface.

#### **2.1.8 preambles**

Specific patterns used for synchronization. There are three different preambles (see 2.4).

#### **2.1.9 subframe**

Fixed structure used to carry the information described in 2.1.1 to 2.1.8 (see 2.2.1 and 2.2.2).

#### **2.1.10 frame**

Sequence of two successive and associated subframes.

#### **2.1.11 block**

Group of 192 consecutive frames. The start of a block is designated by a special subframe preamble (See 2.4).

**2.1.12 channel coding**

Coding describing the method by which the binary digits are represented for transmission through the interface.

**2.1.13 unit interval (UI)**

Shortest nominal time interval in the coding scheme.

NOTE – There are 128 UI in a sample frame.

**2.1.14 interface jitter**

Deviation in timing of interface data transitions (zero crossings) when measured with respect to an ideal clock.

**2.1.15 intrinsic jitter**

Output interface jitter of a device that is either free-running or is synchronized to a jitter-free reference.

**2.1.16 jitter gain**

Ratio, expressed in decibels, of the amplitude of jitter at the synchronization input of a device to the resultant jitter at the output of the device.

NOTE – This definition excludes the effect of intrinsic jitter.

**2.1.17 frame rate**

Rate of transmission of frames on the interface.

**2.2 Structure of format****2.2.1 Subframe format**

Each subframe is divided into 32 time slots, numbered from 0 to 31 (see figure 1).

Time slots 0 to 3 (preambles) carry one of the three permitted preambles (see 2.2.2 and 2.4; also see figure 2).

Time slots 4 to 27 (audio sample word) carry the audio sample word in linear 2's complement representation. The most significant bit (MSB) is carried by time slot 27.

When a 24-bit coding range is used, the LSB is in time slot 4 [see figure 1(a)].

When a 20-bit coding range is sufficient, time slots 8 to 27 carry the audio sample word with the LSB in time slot 8. Time slots 4 to 7 may be used for other applications. Under these circumstances, the bits in time slots 4 to 7 are designated auxiliary sample bits [see figure 1(b)].

If the source provides fewer bits than the interface allows (either 20 or 24), the unused LSBs are set to logic 0.

Time slot 28 (validity bit) carries the validity bit associated with the audio sample word (see 2.5).

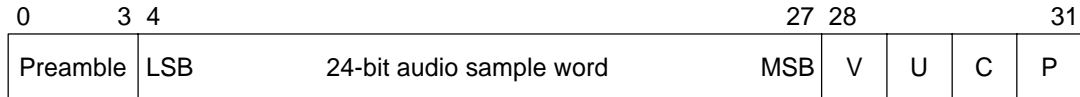
Time slot 29 (user data bit) carries 1 bit of the user data channel associated with the audio channel transmitted in the same subframe (see Section 3).

Time slot 30 (channel status bit) carries 1 bit of the channel status information associated with the audio channel transmitted in the same subframe (see Section 4).

Time slot 31 (parity bit) carries a parity bit such that time slots 4 to 31 inclusive will carry an even number of ones and an even number of zeros (even parity).

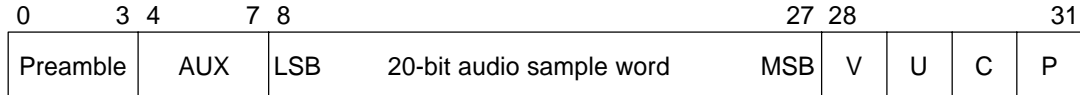
NOTE – The preambles have even parity as an explicit property.





(a)

V    Validity bit  
 U    User data bit  
 C    Channel status bit  
 P    Parity bit  
 AUX   Auxiliary sample bits



(b)

**Figure 1. Subframe format.**

### 2.2.2 Frame format

A frame is uniquely composed of two subframes (see figure 2). Except where otherwise specified the rate of transmission of frames corresponds exactly to the source sampling frequency.

The first subframe normally starts with preamble "X." However, the preamble changes to preamble "Z" once every 192 frames. This defines the block structure used to organize the channel status information. The second subframe always starts with preamble "Y."

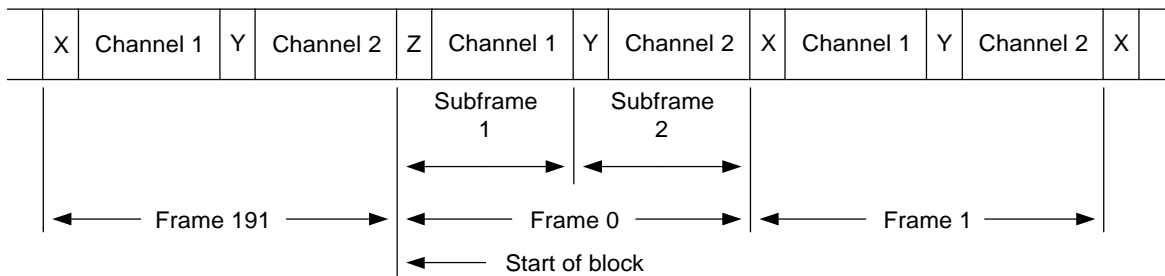
The modes of transmission are signaled by setting bits 0 to 3 of byte 1 of channel status.

*Two-channel mode:* In two-channel mode, the samples from both channels are transmitted in consecutive subframes. Channel 1 is in subframe 1, and channel 2 is in subframe 2.

*Stereophonic mode:* In stereophonic mode, the interface is used to transmit stereophonic audio in which the two channels are presumed to have been simultaneously sampled. The left, or "A," channel is in subframe 1, and the right, or "B," channel is in subframe 2.

*Single-channel mode (monophonic):* In monophonic mode, the transmitted bit rate remains at the normal two-channel rate and the audio sample word is placed in subframe 1. Time slots 4 to 31 of subframe 2 either carry the bits identical to subframe 1 or are set to logic 0. A receiver normally defaults to channel 1 unless manual override is provided.

*Primary/secondary mode:* In some applications requiring two channels where one of the channels is the main or primary channel while the other is a secondary channel, the primary channel is in subframe 1, and the secondary channel is in subframe 2.

**Figure 2. Frame format.**

2.3 Channel coding

To minimize the direct-current (dc) component on the transmission line, to facilitate clock recovery from the data stream, and to make the interface insensitive to the polarity of connections, time slots 4 to 31 are encoded in biphase-mark.

Each bit to be transmitted is represented by a symbol comprising two consecutive binary states. The first state of a symbol is always different from the second state of the previous symbol. The second state of the symbol is identical to the first if the bit to be transmitted is logic 0. However, it is different if the bit is logic 1 (see figure 3).

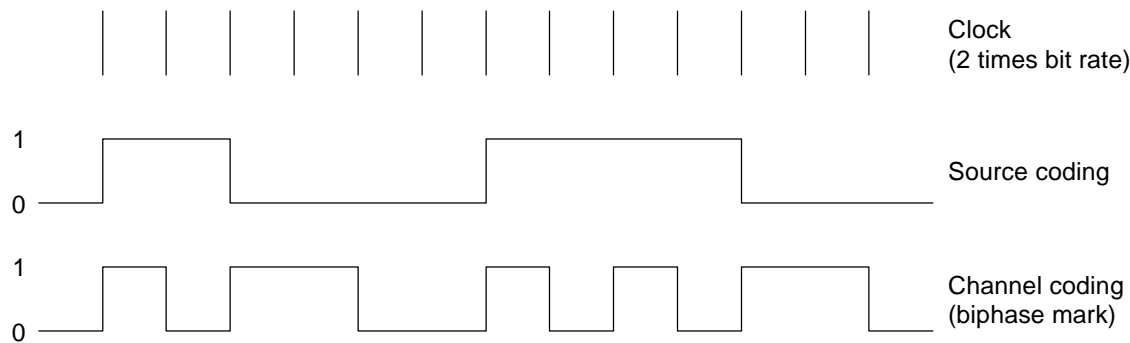


Figure 3. Channel coding.

2.4 Preambles

Preambles are specific patterns providing synchronization and identification of the subframes and blocks.

To achieve synchronization within one sampling period and to make this process completely reliable, these patterns violate the biphase-mark code rules, thereby avoiding the possibility of data imitating the preambles.

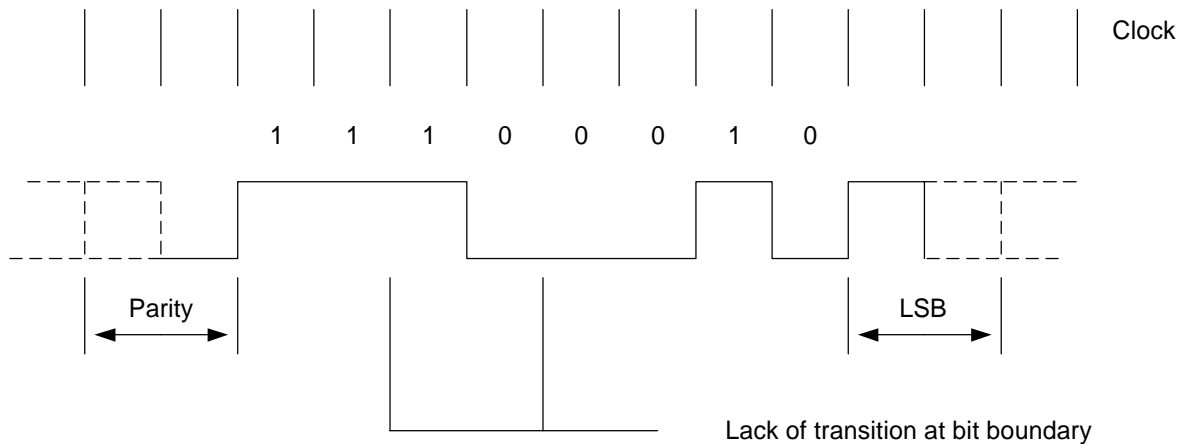
A set of three preambles is used. These preambles are transmitted in the time allocated to four time slots at the start of each subframe (time slots 0 to 3), and are represented by eight successive states. The first state of the preamble is always different from the second state of the previous symbol (representing the parity bit). Depending on this state the preambles are:

Preceding state	0	1	
	Channel coding		
"X"	11100010	00011101	Subframe 1
"Y"	11100100	00011011	Subframe 2
"Z"	11101000	00010111	Subframe 1 and block start

Like biphase code, these preambles are dc free and provide clock recovery. They differ in at least two states from any valid biphase sequence.

Figure 4 represents preamble "X."

NOTE – Owing to the even-parity bit in time slot 31, all preambles will start with a transition in the same direction (see 2.2.1). Thus only one of these sets of preambles will, in practice, be transmitted through the interface. However, it is necessary for either set to be decodable because a polarity reversal might occur in the connection.



**Figure 4. Preamble X (11100010).**

## 2.5 Validity bit

The validity bit is logic 0 if the audio sample word is suitable for conversion to an analog audio signal, and it is logic 1 if it is not.

There is no default state for the validity bit.

## 3 User data format

User data bits may be used in any way desired by the user.

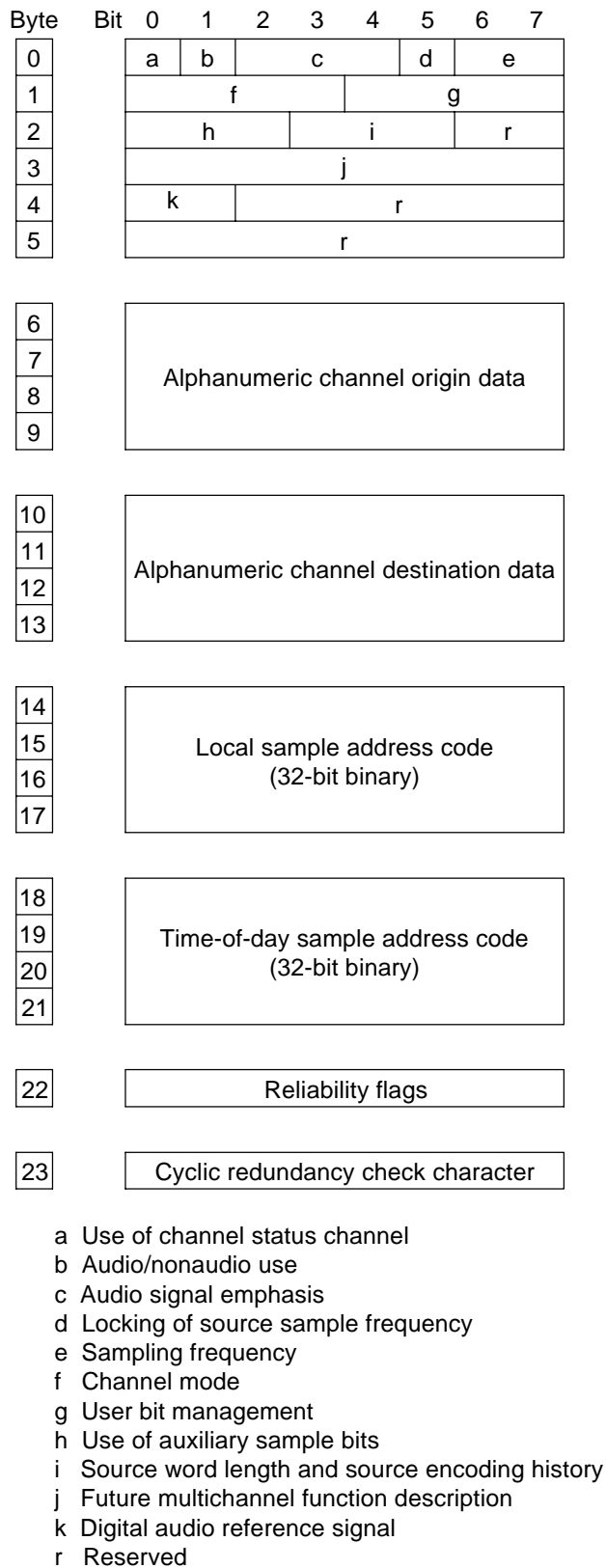
Possible formats for the user data channel are indicated by the channel status byte 1, bits 4–7.

The default value of the user data bit is logic 0.

## 4 Channel status format

The channel status for each audio signal carries information associated with that audio signal, and thus it is possible for different channel status data to be carried in the two subframes of the digital audio signal. Examples of information to be carried in the channel status are: length of audio sample words, number of audio channels, sampling frequency, sample address code, alphanumeric source and destination codes, and emphasis.

Channel status information is organized in 192-bit blocks, subdivided into 24 bytes (see figure 5). The first bit of each block is carried in the frame with preamble "Z."



**Figure 5. Channel status data format.**

The specific organization follows, wherein the suffix 0 designates the first byte or bit.

*Byte 0*

bit 0	0	Consumer use of channel status block (see note).
	1	Professional use of channel status block.
bit 1	0	Audio sample word represents linear PCM samples.
	1	Audio sample word used for purposes other than linear PCM samples.
bits 2-4	Encoded audio signal emphasis.	
bit	2 3 4	
space	0 0 0	Emphasis not indicated. Receiver defaults to no emphasis with manual override enabled.
	1 0 0	No emphasis. Receiver manual override is disabled.
	1 1 0	50/15- $\mu$ s emphasis. Receiver manual override is disabled.
	1 1 1	International Telegraph and Telephone Consultative Committee (CCITT) J.17 emphasis (with 6.5-dB insertion loss at 800 Hz). Receiver manual override is disabled.
	All other states of bits 2-4 are reserved and are not to be used until further defined.	
bit 5	0	Default. Lock condition not indicated.
	1	Source sampling frequency unlocked.
bits 6-7	Encoded sampling frequency.	
bit	6 7	
state	0 0	Sampling frequency not indicated. Receiver default to 48 kHz and manual override or auto set is enabled.
	0 1	48-kHz sampling frequency. Manual override or auto set is disabled.
	1 0	44.1-kHz sampling frequency. Manual override or auto set is disabled.
	1 1	32-kHz sampling frequency. Manual override or auto set is disabled.

NOTE 1 – The significance of byte 0, bit 0 is such that a transmission from an interface conforming to IEC 958 "consumer use" can be identified, and a receiver conforming only to IEC 958 "consumer use" will correctly identify a transmission from a "professional-use" interface as defined in this standard. Connection of a "professional-use" transmitter with a "consumer-use" receiver or vice versa might result in unpredictable operation. Thus the following byte definitions only apply when bit 0 = logic 1 (professional use of the channel status block).

NOTE 2 – The indication of sampling frequency, or the use of one of the sampling frequencies that can be indicated in this byte, is not a requirement for operation of the interface. The 00 state of bits 6-7 may be used if the transmitter does not support the indication of sampling frequency, the sampling frequency is unknown, or the sample frequency is not one of those that can be indicated in this byte. In the latter case for some sampling frequencies byte 4 may be used to indicate the correct value.

NOTE 3 – When byte 1, bits 1-3 indicate single channel double sampling frequency mode then the sampling frequency of the audio signal is twice that indicated by bits 6-7 of byte 0.

*Byte 1*

bits 0-3	Encoded channel mode.	
bit	0 1 2 3	
state	0 0 0 0	Mode not indicated. Receiver default to two-channel mode. Manual override is enabled.
	0 0 0 1	Two-channel mode. Manual override is disabled.
	0 0 1 0	Single-channel mode (monophonic). Manual override is disabled.
	0 0 1 1	Primary/secondary mode (subframe 1 is primary). Manual override is disabled.
	0 1 0 0	Stereophonic mode (channel 1 is left channel). Manual override is disabled.
	0 1 0 1	Reserved for user-defined applications.
	0 1 1 0	Reserved for user-defined applications.

bits 4-7 bit state	0 1 1 1	Single channel double sampling frequency mode. Sub-frames 1 and 2 carry successive samples of the same signal. The sampling frequency of the signal is double the frame rate, and is double the sampling frequency indicated in byte 0 (but not double the rate indicated in byte 4, if that is used). Manual override is disabled. Vector to byte 3 for channel identification.
	1 0 0 0	Single channel double sampling frequency mode - stereo mode left. Sub-frames 1 and 2 carry successive samples of the same signal. The sampling frequency of the signal is double the frame rate, and is double the sampling frequency indicated in byte 0 (but not double the rate indicated in byte 4, if that is used). Manual override is disabled.
	1 0 0 1	Single channel double sampling frequency mode - stereo mode right. Sub-frames 1 and 2 carry successive samples of the same signal. The sampling frequency of the signal is double the frame rate, and is double the sampling frequency indicated in byte 0 (but not double the rate indicated in byte 4, if that is used). Manual override is disabled.
	1 1 1 1	Multichannel mode. Vector to byte 3 for channel identification.
	All other states of bits 0-3 are reserved and are not to be used until further defined.	
	Encoded user bits management.	
	4 5 6 7	
	0 0 0 0	Default, no user information is indicated.
	0 0 0 1	192-bit block structure. Preamble "Z" indicates the start of block.
	0 0 1 0	Reserved for the AES18 standard.
	0 0 1 1	User defined.
	0 1 0 0	User data conforms to the general user data format defined in IEC 60958-3.

All other states of bits 4-7 are reserved and are not to be used until further defined.

<i>Byte 2</i>		
bits 0-2 bit state	Encoded use of auxiliary sample bits.	
	0 1 2	
	0 0 0	Maximum audio sample word length is 20 bits (default). Use of auxiliary sample bits not defined.
	0 0 1	Maximum audio sample word length is 24 bits. Auxiliary sample bits are used for main audio sample data.
	0 1 0	Maximum audio sample word length is 20 bits. Auxiliary sample bits in this channel are used to carry a single coordination signal (see note 1).
	0 1 1	Reserved for user defined applications.
	All other states of bits 0-2 are reserved and are not to be used until further defined.	

NOTE 1 – The signal coding used for the coordination channel is described in Annex A.

bits 3-5 bit state	Encoded audio sample word length of transmitted signal (see notes 2, 3, and 4).			
	3 4 5	Audio sample word length if maximum length is 24 bits as indicated by bits 0-2 above.	Audio sample word length if maximum length is 20 bits as indicated by bits 0-2 above.	
	0 0 0	Word length not indicated (default).	Word length not indicated (default).	
	0 0 1	23 bits	19 bits	
	0 1 0	22 bits	18 bits	
	0 1 1	21 bits	17 bits	
	1 0 0	20 bits	16 bits	
	1 0 1	24 bits	20 bits	
	All other states of bits 3-5 are reserved and are not to be used until further defined.			
	Indication of alignment level.			
bits 6-7 bit	6 7			
	0 0	Alignment level not indicated		

0 1	Alignment to SMPTE RP155 (alignment level is 20 dB below maximum code).
1 0	Alignment to EBU R68 (alignment level is 18.06 dB below maximum code).
1 1	Reserved for future use.

NOTE 2 – The default state of bits 3-5 indicates that the number of active bits within the 20- or 24-bit coding range is not specified by the transmitter. The receiver should default to the maximum number of bits specified by the coding range and enable manual override or auto set.

NOTE 3 – The nondefault states of bits 3–5 indicate the number of bits within the 20- or 24-bit coding range which might be active. This is also an indirect expression of the number of LSBs that are certain to be inactive, which is equal to 20 or 24 minus the number corresponding to the bit state. The receiver should disable manual override and auto set for these bit states.

NOTE 4 – Irrespective of the audio sample word length as indicated by any of the states of bits 3–5, the MSB is in time slot 27 of the transmitted subframe as specified in 2.2.1.

#### Byte 3

bit 7	0	Undefined multichannel mode (default).
	1	Defined multichannel modes.
	The definition of the remaining bit states depends on the state of bit 7.	
bits 0-6	Channel number (when byte 3 bit 7 is 0).	
	The channel number is the value of the byte (with bit 0 as the least significant bit) plus one.	
bits 4-6	Multichannel mode (when byte 3 bit 7 is 1).	
bit	4 5 6	
	0 0 0	Multichannel mode 0. The channel number is defined by bits 0-3 of this byte.
	1 0 0	Multichannel mode 1. The channel number is defined by bits 0-3 of this byte.
	0 1 0	Multichannel mode 2. The channel number is defined by bits 0-3 of this byte.
	1 1 0	Multichannel mode 3. The channel number is defined by bits 0-3 of this byte.
	1 1 1	User defined multichannel mode. The channel number is defined by bits 0-3 of this byte.
	All other states of bits 4-6 are reserved and are not to be used until further defined.	
bits 0-3	Channel number (when byte 3 bit 7 is 1).	
	The channel number is one plus the numeric value of value of these bits taken as a binary number (with bit 3 as the most significant bit).	

NOTE 1 - The defined multichannel modes identify mappings between channel numbers and function. (The standard mappings are under consideration. Some mappings may involve groupings of up to 32 channels by combining two modes.)

NOTE 2 - For compatibility with equipment that is only sensitive to the channel status data in one subframe the channel carried by subframe 2 may indicate the same channel number as channel 1. In that case it is implicit that the second channel has a number one higher than the channel of subframe 1 except in single channel double sampling frequency mode.

NOTE 3 - When bit 7 is 1 the 4 bit channel number can be mapped to the channel numbering in bits 20-23 of the consumer mode channel status defined in IEC 60958-3. In this case channel A of consumer mode maps to channel 2, channel B maps to channel 3 and so on.

#### Byte 4

bits 0-1	Digital audio reference signal (per AES11).	
bit	0 1	
state	0 0	Not a reference signal (default).
	0 1	Grade 1 reference signal.
	1 0	Grade 2 reference signal.
	1 1	Reserved and not used until further defined.
bit 2	Reserved.	
bits 3-6	Sampling frequency.	

bit	6 5 4 3
state	0 0 0 0 Not indicated (default).
	0 0 0 1 24 kHz
	0 0 1 0 96 kHz
	0 0 1 1 192 kHz
	0 1 0 0 Reserved.
	0 1 0 1 Reserved.
	0 1 1 0 Reserved.
	0 1 1 1 Reserved.
	1 0 0 0 Reserved (for vectoring).
	1 0 0 1 22.05 kHz
	1 0 1 0 88.2 kHz
	1 0 1 1 176.4 kHz
	1 1 0 0 Reserved.
	1 1 0 1 Reserved.
	1 1 1 0 Reserved.
	1 1 1 1 User defined .
bit 7	Sampling frequency scaling flag.
	0 No scaling (default).
	1 Sampling frequency is 1/1.001 times that indicated by byte 4 bits 3-6, or by byte 0 bits 6-7.

NOTE 1 – The sampling frequency indicated in byte 4 is not dependent on the channel mode indicated in byte 1.

NOTE 2 – The indication of sampling frequency, or the use of one of the sampling frequencies that can be indicated in this byte, is not a requirement for operation of the interface. The '0000' state of bits 3-6 may be used if the transmitter does not support the indication of sampling frequency in this byte, the sampling frequency is unknown, or the sample frequency is not one of those that can be indicated in this byte. In the later case for some sampling frequencies byte 0 may be used to indicate the correct value.

NOTE 3 – The reserved states of bits 3-6 of byte 4 are intended for later definition such that bit 6 is set to define rates related to 44.1 kHz (except for state 1000) and clear to defined rates related to 48 kHz. They should not be used until further defined.

#### Byte 5

bits 0-7 Reserved and are set to logic 0 until further defined.

#### Bytes 6-9

Alphanumeric channel origin data. First character in message is byte 6.  
 bits 0-7 7-bit International Organization for Standardization (ISO) 646,  
 (each American Standard Code for Information Interchange (ASCII),  
 byte) data with no parity bit. LSBs are transmitted first with logic 0 in bit 7.  
 Nonprinted control characters (codes 01 to 1F hex and 7F hex) are not permitted.  
 Default value is logic 0 (code 00 hex, ASCII null).

#### Bytes 10-13

Alphanumeric channel destination data.  
 First character in message is byte 10.  
 bits 0-7 7-bit ISO 646 (ASCII) data with no parity bit.  
 (each LSBs are transmitted first with logic 0 in bit 7.  
 byte) Nonprinted control characters (codes 01 to 1F hex and 7F hex)  
 are not permitted. Default value is logic 0 (code 00 hex, ASCII null).

#### Bytes 14-17

Local sample address code (32-bit binary with LSBs first).  
 Value is of first sample of current block.  
 bits 0-7



(each byte)                      LSBs are transmitted first. Default value is logic 0.

NOTE – This has the same function as a recording index counter.

#### *Bytes 18-21*

Time-of-day sample address code (32-bit binary with LSBs first). Value is of first sample of current block.

bits 0-7  
(each byte)                      LSBs are transmitted first. Default value is logic 0.

NOTE – This is the time of day laid down during the source encoding of the signal and remains unchanged during subsequent operations. A value of all zeros for the binary sample address code is, for transcoding to real time, or to time codes in particular, to be taken as midnight (i.e., 00 h, 00 min, 00 s, 00 frame). Transcoding of the binary number to any conventional time code requires accurate sample frequency information to provide a sample accurate time.

#### *Byte 22*

Flag used to identify whether the information carried by the channel status data is reliable. According to the following table, if reliable, the appropriate bits are set to logic 0 (default); if unreliable, the bits are set to logic 1.

bits 0-3                      Reserved and are set to logic 0 until further defined.  
bit 4                          Bytes 0 to 5.  
bit 5                          Bytes 6 to 13.  
bit 6                          Bytes 14 to 17.  
bit 7                          Bytes 18 to 21.

#### *Byte 23*

Channel status data cyclic redundancy check character (CRCC).

bits 0-7                      Generating polynomial is  $G(x) = x^8 + x^4 + x^3 + x^2 + 1$ . The CRCC conveys information to test valid reception of the entire channel status data block (bytes 0 to 22 inclusive). For serial implementations the initial condition of all ones should be used in generating the check bits with the LSB transmitted first. Default value is logic 0 for "minimum" implementation of channel status only (see 5.2.1).

NOTE – Annex B includes a diagram of the shift register circuit used to generate the code, two examples of channel status data, and the corresponding CRCC.

## **5 Interface format implementation**

### **5.1 General**

To promote compatible operation between items of equipment built to this specification it is necessary to establish which information bits and operational bits need to be encoded and sent by a transmitter and decoded by an interface receiver.

Documentation shall be provided describing the channel status features supported by the interface transmitters and receivers.

### **5.2 Transmitter**

Transmitters shall follow all the formatting and channel coding rules established in earlier sections of this specification including all notes therein. Along with the audio sample word, all transmitters shall correctly encode and transmit the validity bit, user bit, parity bit, and the three preambles. The channel status shall be encoded to one of the implementations given in 5.2.1, 5.2.2, and 5.2.3.

The following three implementations are defined: "minimum," "standard," and "enhanced." These terms are used to communicate in a simple manner the level of implementation of the interface transmitter involving the many features of channel status. Irrespective of the level of implementation, all reserved states of bits defined in Section 4 shall remain unchanged.

### 5.2.1 "Minimum" implementation of channel status

The "minimum" implementation represents the lowest level of implementation of the interface that meets the requirements of this specification document. In the "minimum" implementation, transmitters shall encode and transmit channel status byte 0 bit 0 with a state of logic 1 signifying "professional use of channel status block." All other channel status bits of byte 0 to byte 23 inclusive shall be transmitted with the default state of all logic 0's. In this circumstance, the receiver will adopt the default conditions specified in bytes 0 to 2.

If additional bytes of channel status (which do not fully comply with the "standard" implementation, see 5.2.2) are implemented as required by an application, the interface transmitter shall be classified as a "minimum" implementation of channel status.

It should be noted that the "minimum" implementation imposes severe operational restrictions on some receiving devices which may be connected to it. For example, receivers implementing byte 23 will normally show a cyclic redundancy check error when the default value of logic 0 is received as the CRCC. Also, reception of the default value for byte 0 bits 6-7 might cause improper operation in receiving devices not supporting manual override or auto set capabilities.

### 5.2.2 "Standard" implementation of channel status

The "standard" implementation provides a fundamental level of implementation which should prove sufficient for general applications in professional audio or broadcasting. In addition to conforming to the requirements described in 5.2.1 for the "minimum" implementation, a "standard" implementation interface transmitter shall correctly encode and transmit all channel status bits in byte 0, byte 1, byte 2, and byte 23 (CRCC) in the manner specified in this document.

### 5.2.3 "Enhanced" implementation of channel status

In addition to conforming to the requirements described in 5.2.2 for the "standard" implementation, the "enhanced" implementation shall provide further capabilities.

## 5.3 Receivers

Implementation in receivers is highly dependent on the application. Proper documentation shall be provided on the level of implementation of the interface receiver for decoding the transmitted information (validity, user, channel status, parity) and on whatever subsequent action is taken by the equipment of which it is a part.

## 6 Electrical requirements

### 6.1 General Characteristics

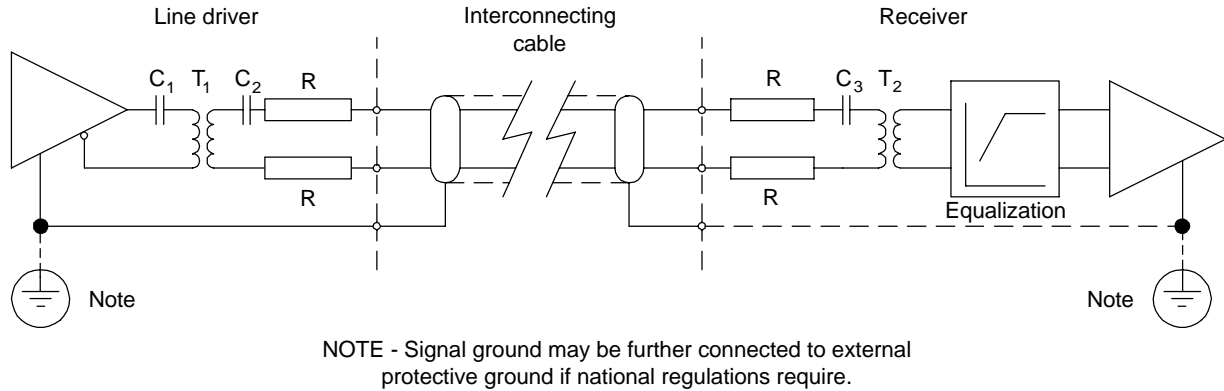
The electrical parameters of the interface are based on those defined in CCITT V.11 which allow transmission of balanced-voltage digital signals up to a few hundred meters distance. The frequency range used to qualify the interface electrical parameters is dependent on the maximum data rate supported. The upper frequency is 128 times the maximum frame rate.

In order to improve the balance of the transmitter or the receiver, or both, beyond that recommended by the CCITT, a circuit conforming to the general configuration shown in figure 6 may be used.

In this circuit, the series capacitors  $C_2$  and  $C_3$  isolate the transformers and prevent damage from connection to a source containing a dc voltage. In addition to achieving higher rejection of common-mode signals, the

transformers reduce grounding and electromagnetic interference (EMI) problems. Although equalization may be used at the receiver, no equalization before transmission shall be permitted.

The interconnecting cable shall be balanced and screened (shielded) with a nominal characteristic impedance of  $110\ \Omega$  at frequencies from 100 kHz to 128 times the maximum frame rate.



**Figure 6. General circuit configuration**

## 6.2 Line driver characteristics

### 6.2.1 Output impedance

The line driver shall have a balanced output with an internal impedance of  $110\ \Omega \pm 20\%$ , at frequencies from 100 kHz to 128 times the maximum frame rate when measured at the output terminals.

### 6.2.2 Signal amplitude

The signal amplitude shall lie between 2 and 7 V peak to peak, when measured across a  $110\text{-}\Omega$  resistor connected to the output terminals, without any interconnecting cable present.

### 6.2.3 Balance

Any common-mode component at the output terminals shall be more than 30 dB below the signal at frequencies from dc to 128 times the maximum frame rate.

### 6.2.4 Rise and fall times

The rise and fall times, determined between the 10% and 90% amplitude points, shall be between 5 ns and 30 ns when measured across a  $110\text{-}\Omega$  resistor connected to the output terminals, without any interconnecting cable present.

NOTE – Operation toward the lower limit of 5 ns may improve the received signal eye pattern, but may increase EMI at the transmitter. Equipment must meet local regulations regarding EMI.

### 6.2.5 Output interface jitter

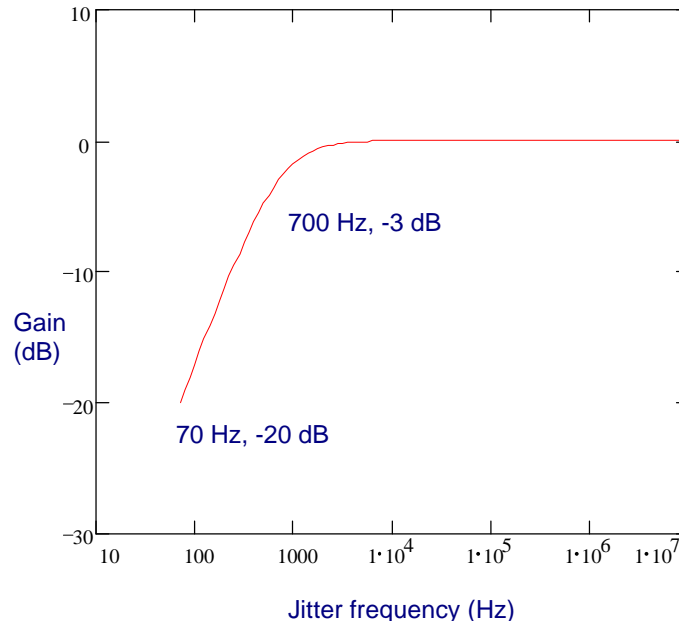
Jitter at the output of a device shall be measured as the sum of the jitter intrinsic to the device and jitter being passed through from the timing reference of the device.

#### 6.2.5.1 Intrinsic jitter

The peak value of the intrinsic jitter at the output of the interface, measured at all the transition zero crossings shall be less than 0.025 UI when measured with the intrinsic-jitter measurement filter.

NOTE 1 – This requirement applies both when the equipment is locked to an effectively jitter-free timing reference (which may be a modulated digital audio signal) and when the equipment is free-running.

NOTE 2 – The intrinsic-jitter measurement-filter characteristic is shown in figure 9. It shows a minimum-phase high-pass filter with 3 dB attenuation at 700 Hz, a first order roll-off to 70 Hz and with a pass-band gain of unity.

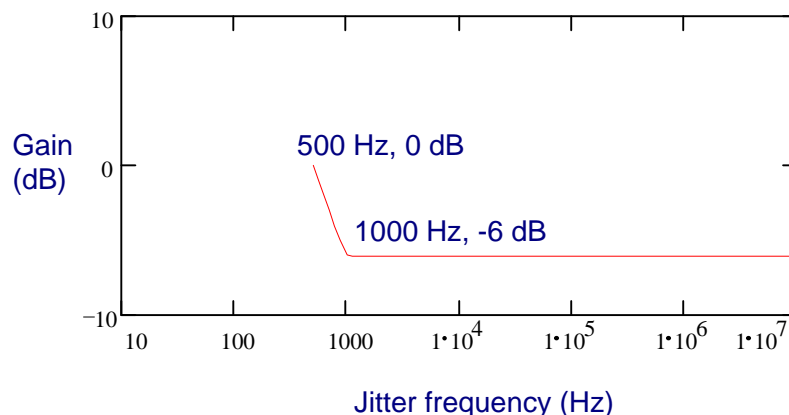


**Figure 9. Intrinsic-jitter measurement-filter characteristic.**

#### 6.2.5.2 Jitter gain

The sinusoidal jitter gain from any timing reference input to the signal output shall be less than 2 dB at all frequencies.

NOTE – If jitter attenuation is provided and it is such that the sinusoidal jitter gain falls below the jitter transfer function mask of figure 10 then the equipment specification should state that the equipment jitter attenuation is within this specification. The mask imposes no additional limit on low-frequency jitter gain. The limit starts at the input-jitter frequency of 500 Hz where it is 0 dB, and falls to -6 dB at and above 1 kHz.



**Figure 10. Jitter transfer-function mask.**

### 6.3 Line receiver characteristics

#### 6.3.1 Terminating impedance

The receiver shall present an essentially resistive impedance of  $110\ \Omega \pm 20\%$  to the interconnecting cable over the frequency band from 0.1 MHz to 128 times the maximum frame rate when measured across the input terminals. The application of more than one receiver to any one line might create transmission errors due to the resulting impedance mismatch.

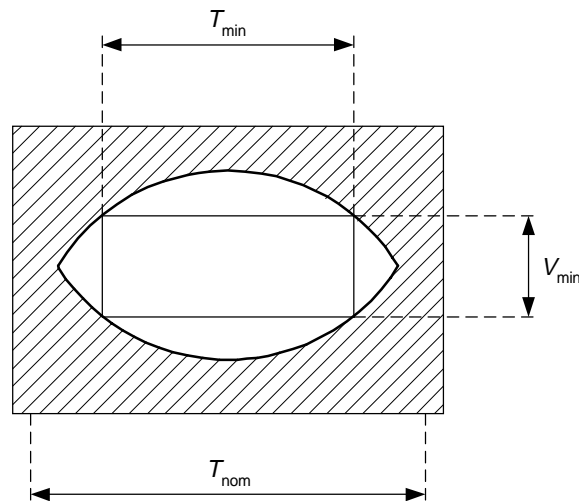
#### 6.3.2 Maximum input signals

The receiver shall correctly interpret the data when connected directly to a line driver working between the extreme voltage limits specified in 6.2.2.

NOTE – The AES3-1985 specification for line driver signal amplitude was 10 V peak to peak maximum.

#### 6.3.3 Minimum input signals

The receiver shall correctly sense the data when a random input signal produces the eye diagram characterized by a  $V_{\min}$  of 200 mV and  $T_{\min}$  of 50% of  $T_{\text{nom}}$  (see figure 7).



$$T_{\min} : 0.5 \times T_{\text{nom}}$$

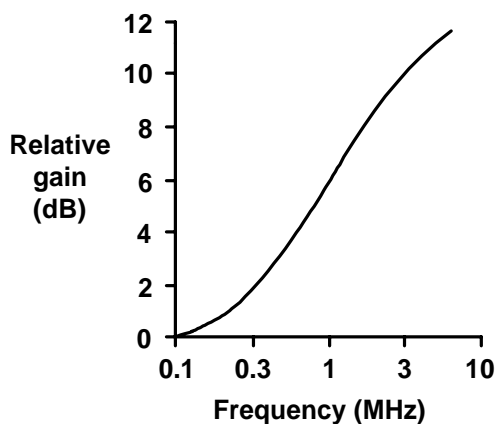
$$V_{\min} : 200\ \text{mV}$$

$$T_{\text{nom}} : \text{One-half the biphase symbol period}$$

**Figure 7. Eye diagram.**

#### 6.3.4 Receiver equalization

Optional equalization can be applied in the receiver to enable an interconnecting cable longer than 100 m to be used. A suggested frequency equalization characteristic for operation at frame rates of 48 kHz is shown in figure 8. The receiver shall meet the requirements specified in 6.3.2 and 6.3.3.



**Figure 8. Suggested equalizing characteristic for a receiver operating at 48 kHz frame rate.**

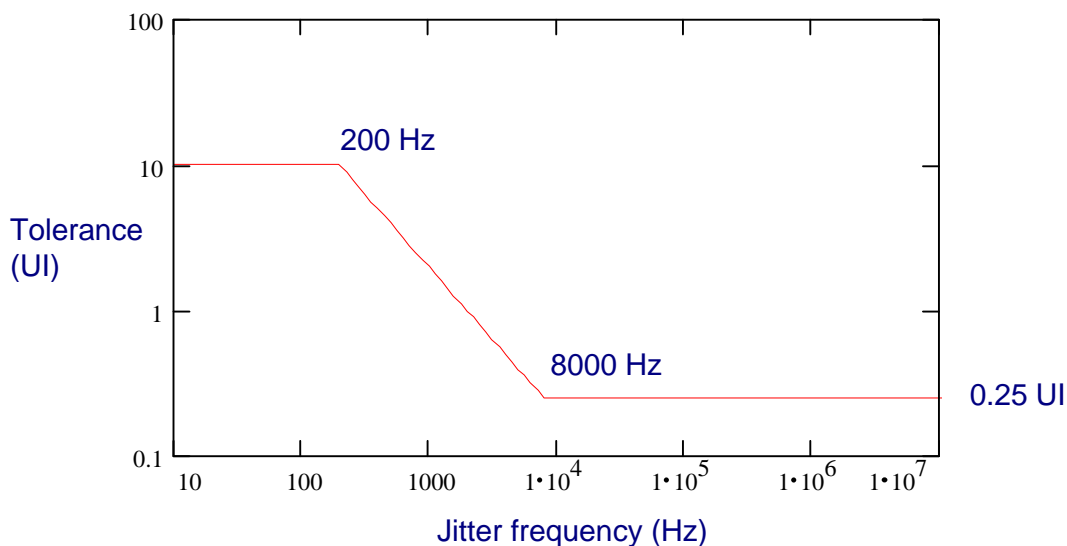
### 6.3.5 Common-mode rejection

There shall be no data errors introduced by the presence of a common-mode signal of up to 7 V peak at frequencies from dc to 20 kHz.

### 6.3.6 Receiver jitter tolerance

An interface data receiver should correctly decode an incoming data stream with any sinusoidal jitter defined by the jitter tolerance template of figure 11.

NOTE –The template requires a jitter tolerance of 0.25 UI peak-to-peak at high frequencies, increasing with the inverse of frequency below 8 kHz to level off at 10 UI peak-to-peak below 200 Hz.



**Figure 11. Jitter tolerance template.**

## 6.4 Connectors

The standard connector for both outputs and inputs shall be the circular latching three-pin connector described in IEC 268-12. (This type of connector is normally called XLR.)

An output connector fixed on an item of equipment shall use male pins with a female shell. The corresponding cable connector shall thus have female pins with a male shell.

An input connector fixed on an item of equipment shall use female pins with a male shell. The corresponding cable connector shall thus have male pins with a female shell. The pin usage shall be:

- Pin 1 Cable shield or signal earth;
- Pin 2 Signal;
- Pin 3 Signal.

(Note that the relative polarity of pins 2 and 3 is not important in the digital case.)

Equipment manufacturers should clearly label digital audio inputs and outputs as such, including the terms "digital audio input" or "digital audio output" as appropriate.

In such cases where panel space is limited and the function of the connector might be confused with an analog signal connector, the abbreviation DI or DO should be used to designate digital audio inputs and outputs, respectively.

## 7 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 the most recent editions of the Standards listed below should be obtained.]

AES5-1984, *AES Recommended Practice for Professional Digital Audio Applications Employing Pulse Code Modulation—Preferred Sampling Frequencies*, Audio Engineering Society, New York, NY, USA (1984).

AES10-1991, *AES Recommended Practice for Digital Audio Engineering—Serial Multichannel Audio Digital Interface (MADI)*, Audio Engineering Society, New York, NY, USA (1991).

AES11-1991, *AES Recommended Practice for Digital Audio Engineering—Synchronization of Digital Audio Equipment in Studio Operations*, Audio Engineering Society, New York, NY, USA (1991).

AES18-1992, *AES Recommended practice for digital audio engineering—Format for the user data channel of the AES digital audio interface*, Audio Engineering Society, New York, NY, USA (1992).

CCIR Recommendation 647, *A digital audio interface for broadcasting studios*, Green Book, vol. 10—part 1, International Radio Consultative Committee (CCIR), Dubrovnik, Yugoslavia (1986).

CCITT Recommendation J.17, *Pre-emphasis used on sound program circuits*, International Telegraph and Telephone Consultative Committee (1972).

CCITT Recommendation V.11: *Electrical characteristics for balanced double-current interchange circuits for general use with integrated circuit equipment in the field of data communications*, International Telegraph and Telephone Consultative Committee (1976,1980).

IEC 268-11, *Sound system equipment, part 11: Application of connectors for the interconnection of sound system components*, International Electrotechnical Commission, Geneva, Switzerland (1987).

IEC 268-12, *Sound system equipment, part 12: Application of connectors for broadcast and similar use*, International Electrotechnical Commission, Geneva, Switzerland (1987).

IEC 841, *Audio recording—PCM encoder/decoder system*, International Electrotechnical Commission, Geneva, Switzerland (1988).

IEC 908, *Compact disc digital audio system*, International Electrotechnical Commission, Geneva, Switzerland (1987).

IEC 958, *Digital audio interface*, International Electrotechnical Commission, Geneva, Switzerland (1989).

ISO 646: 1983, *Information processing—ISO 7-bit coded character set for information interchange*, International Organization for Standardization, Geneva, Switzerland (1983).



[The following Annexes are not a part of the *AES Recommended practice for digital audio engineering – Serial transmission format for two-channel linearly represented digital audio data*, AES3-1992. They are included for information purposes only.]

ANNEX A

Provision of additional, voice-quality channels via the digital audio interface

When a 20-bit coding range is sufficient for the audio signal, the 4 auxiliary sample bits may be used for a voice-quality coordination signal (talk back).

The voice-quality signal is sampled at exactly one-third of the sampling frequency for the main audio, coded uniformly with 12 bits per sample represented in 2's complement form. It is sent 4 bits at a time in the auxiliary sample bits of the interface subframes. One such signal may be sent in subframe 1 and another in subframe 2. The "Z" preamble at the start of each block is used as a frame alignment word for the voice-quality signals. The two subframes of frame 0 each contain the 4 LSBs of sample of their respective voice-quality signal, as shown in figure A.1. Figure A.1 also shows two voice-quality signals, one in each subframe.

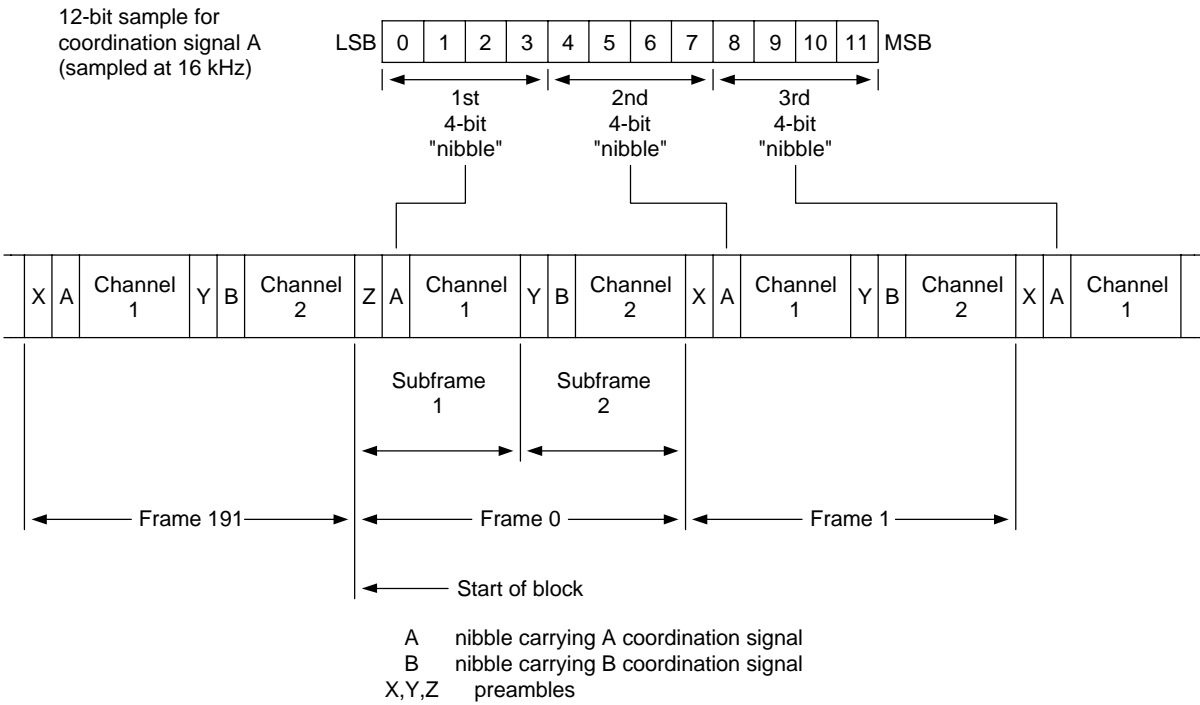


Figure A.1. Frame and block structure.

ANNEX B

Generation of CRCC (byte 23) for channel status

The channel status block format of 192 bits includes a CRC (cyclic redundancy check) code occupying the last 8 bits of the block (byte 23). The specification for the code is given by the generating polynomial:

$$G(x) = x^8 + x^4 + x^3 + x^2 + 1$$

An example of a hardware realization in the serial form is given in figure B. 1. The initial condition of all stages is logic 1.

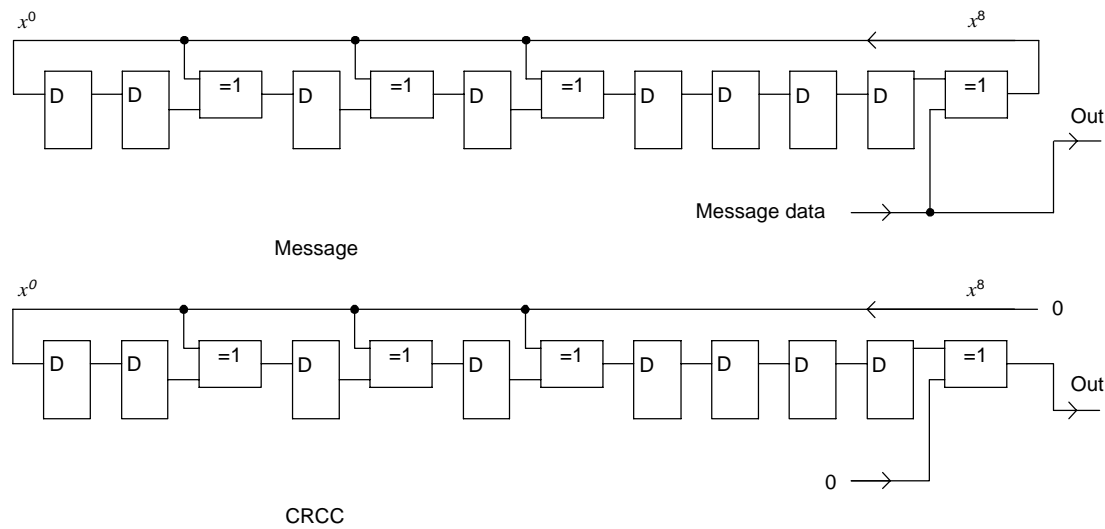


Figure B.1. Flow diagram including exclusive-or gates.

Two examples of channel status data and the resultant CRCC follow.

Example 1:

Byte	Bits set to logic 1
0	0 2 3 4 5
1	1
4	1

All other bits in channel status bytes 0 to 22 inclusive are set to logic 0:

CRCC Byte 23	1	1	0	1	1	0	0	1
Channel status bit:	184							191

Example 2:

Byte	Bits set to logic 1
0	0

All other bits in channel status bytes 0 to 22 inclusive are set to logic 0:

CRCC Byte23	0	1	0	0	1	1	0	0
Channel status bit:	184							191

No particular level of implementation should be taken as implied by the examples given.

**Annex C**

(Informative)

**Informative references**

SMPTE Recommended Practice RP155-1997 *Audio levels for digital audio records on digital television tape recorders*. Society of Motion Picture and Television Engineers. New York, NY, US.

EBU Technical Recommendation R68-1992 *Alignment level in digital audio production equipment and in digital audio recorders*. European Broadcasting Union, Geneva, Switzerland.