AES3-2003
Revision of AES3-1992

REVISED
AES standard for digital audio — Digital input-output interfacing — Serial transmission format for two-channel linearly represented digital audio data

Published by
Audio Engineering Society, Inc.

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.*

2003-09-09 printing

Contents

Foreword

[This foreword is not a part of the PROPOSED DRAFT REVISED *AES standard on digital audio — Digital input-output interfacing — Serial transmission format for two-channel linearly represented digital audio data, DRAFT* AES3-xxxx.]

Foreword to second edition

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. Bluethgen, 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.

R. A. Finger, chair
SC-02-02 Working Group on Digital Input-Output Interfacing
1991 March

Foreword to third edition

This revision of AES3 was prepared under project AES3-R by a writing group of the SC-02-02 Working Group on Input-Output Interfacing of the SC-02 Subcommittee on Digital Audio.

This revision is intended to consolidate the four amendments to the second edition into a single text with minimum impact on the normative sense of the original. Clause and figure numbering has been updated accordingly. Minor normative changes have been made to define a User Bit Management state for metadata and to update "Electrical requirements, General characteristics" to follow a more generalised model adopted by IEC 60958-4 in 2002. Minor editorial changes have also been made, including the tabulation of the Channel Status Data specification.

At the time it was written, the AES Working Group on Digital Input-Output Interfacing included the following individuals who contributed to this standard: J. Dunn, J. Brown, R. Cabot, R. Caine, C. Chambers, R. Chinn, I. Dennis, C. Dinneen, A. Eckhart, R. Finger, J. Fujimori, M. Furukawa, C. Gaunt, J. Grant, S. Harris, S. Herla, W. Krafft, S. Lyman, A. Mason, H. Nakashima, J. Nunn, W. Oxford, J. Paul, D. Queen, J Schmidt, S. Scott, P. Skirrow, Y. Sohma, C. Travis, M. Yonge, J. Yoshio.

NOTE In AES standards documents, sentences containing the verb "shall" are requirements for compliance with the standard. Sentences containing the verb "should" are strong suggestions (recommendations). Sentences giving permission use the verb "may." Sentences expressing a possibility use the verb "can."

REVISED

AES standard for digital audio — Digital input-output interfacing — 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. 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 *AES recommended practice for digital audio engineering -- Synchronization of digital audio equipment in studio operations*. An engineering guideline document to accompany this interface specification has been published as AES-2id *AES information document for digital audio engineering -- Guidelines for the use of the AES3 interface*.

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

2 Normative references

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

AES11, *AES recommended practice for digital audio engineering—Synchronization of digital audio equipment in studio operations,* Audio Engineering Society, New York, NY, USA .

AES18, *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.

ITU-T Recommendation V.11: *Electrical characteristics for balanced double-current interchange circuits operating at data signalling rates up to 10 Mbit/s*, International Telecommunication Union, Geneva, Switzerland..

IEC 60268-12, *Sound system equipment, part 12: Application of connectors for broadcast and similar use*, International Electrotechnical Commission, Geneva, Switzerland.

IEC 60958-3, *Digital audio interface - Part 3: Consumer applications*, International Electrotechnical Commission, Geneva, Switzerland.

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

3 Definitions and abbreviations

3.1
sampling frequency
frequency of the samples representing an audio signal

> NOTE When more than one audio signal is transmitted through the same interface, the sampling frequencies are identical.

3.2
audio sample word
amplitude of a digital audio sample

> NOTE 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.

3.3
auxiliary sample bits
four 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

3.4
validity bit
bit indicating whether the audio sample bits in the same subframe are suitable for conversion to an analog audio signal

3.5
channel status
bits carrying, in a fixed format derived from the block (see 3.11), information associated with each audio channel which is decodable by any interface user

3.6
user data
channel provided to carry any other information

3.7

parity bit
bit provided to permit the detection of an odd number of errors resulting from malfunctions in the interface

3.8
preambles
specific patterns used for synchronization. See 4.3.

3.9
subframe
fixed structure used to carry the information described in 4.1.1 and 4.1.2

3.10
frame
sequence of two successive and associated subframes, see 4.1.2

3.11
block
group of 192 consecutive frames

NOTE The start of a block is designated by a special subframe preamble. See 4.3.

3.12
channel coding
coding describing the method by which the binary digits are represented for transmission through the interface

3.13
unit interval
UI
shortest nominal time interval in the coding scheme

NOTE There are 128 UI in a sample frame.

3.14
interface jitter
deviation in timing of interface data transitions (zero crossings) when measured with respect to an ideal clock

3.15
intrinsic jitter
output interface jitter of a device that is either free-running or is synchronized to a jitter-free reference

3.16
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.

3.17
frame rate
rate of transmission of frames

4 Interface format

4.1 Structure of format

4.1.1 Subframe format

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

Time slots 0 to 3, the preambles, carry one of the three permitted preambles. See 4.1.2 and 4.3; also see figure 2.

Time slots 4 to 27, the 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, the validity bit, carries the validity bit associated with the audio sample word. See 4.4.

Time slot 29, the user data bit, carries 1 bit of the user data channel associated with the audio channel transmitted in the same subframe. See clause 5.

Time slot 30, the channel status bit, carries 1 bit of the channel status information associated with the audio channel transmitted in the same subframe. See clause 6.

Time slot 31, the 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.

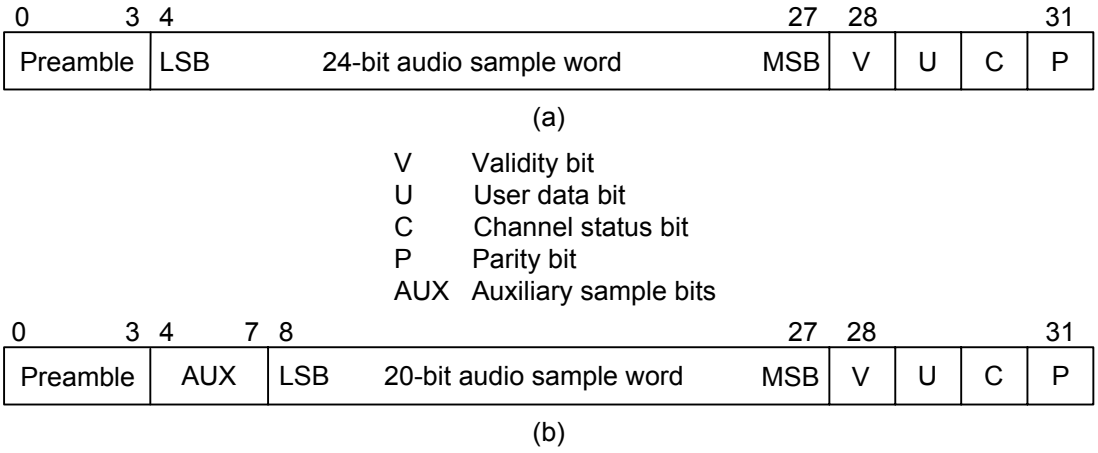


Figure 1 — Subframe format

4.1.2 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. Examples include:

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.

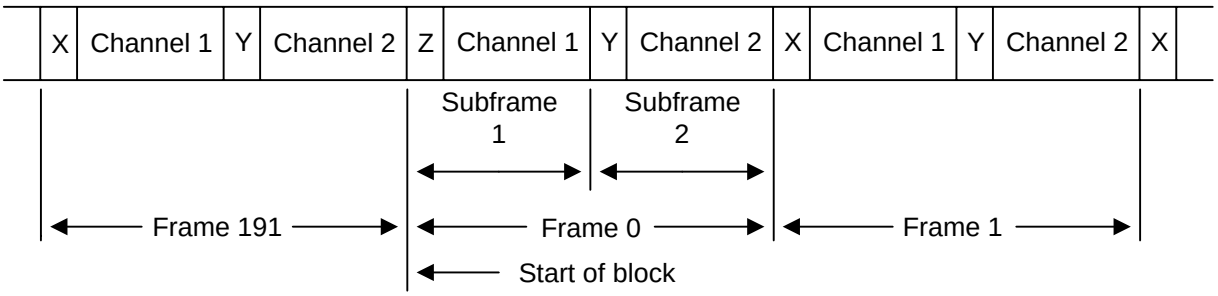


Figure 2 — Frame format

4.2 Channel coding

To minimize the direct-current (d.c.) 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.

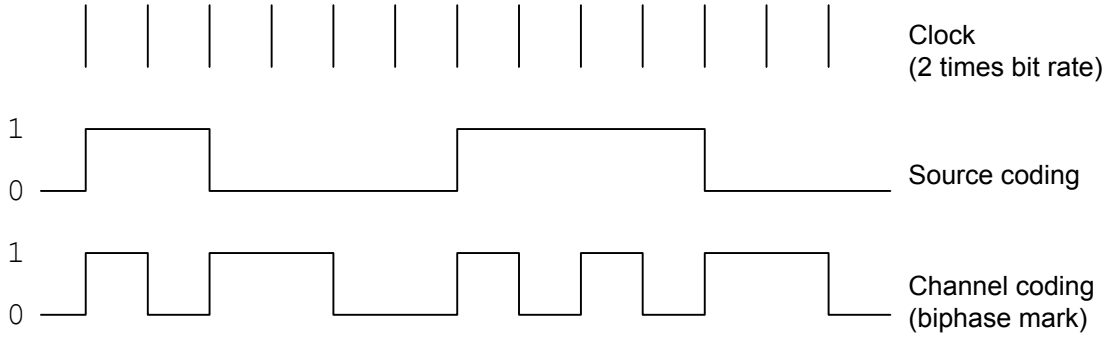


Figure 3 — Channel coding

4.3 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:

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

Like biphase code, these preambles are d.c. 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 4.1.1. 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.

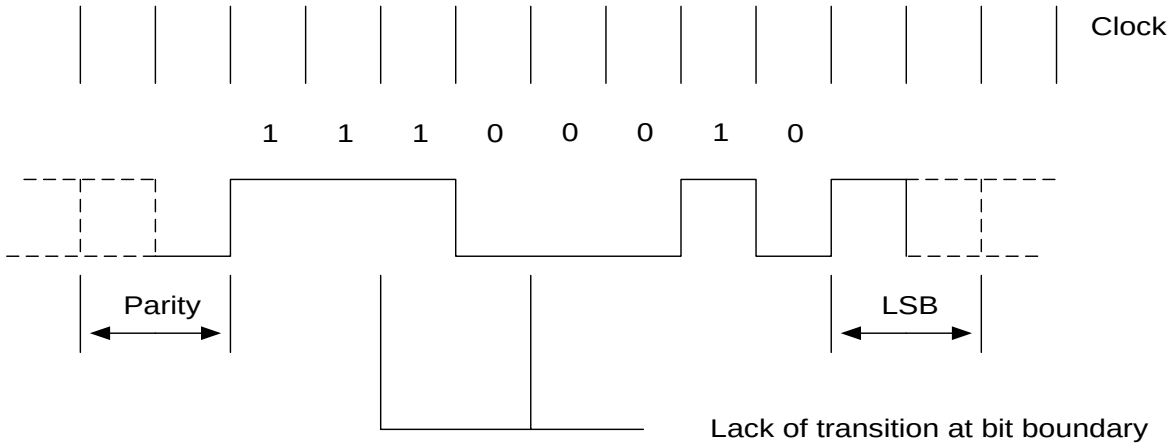


Figure 4 — Preamble X (11100010)

4.4 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.

5 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 to 7.

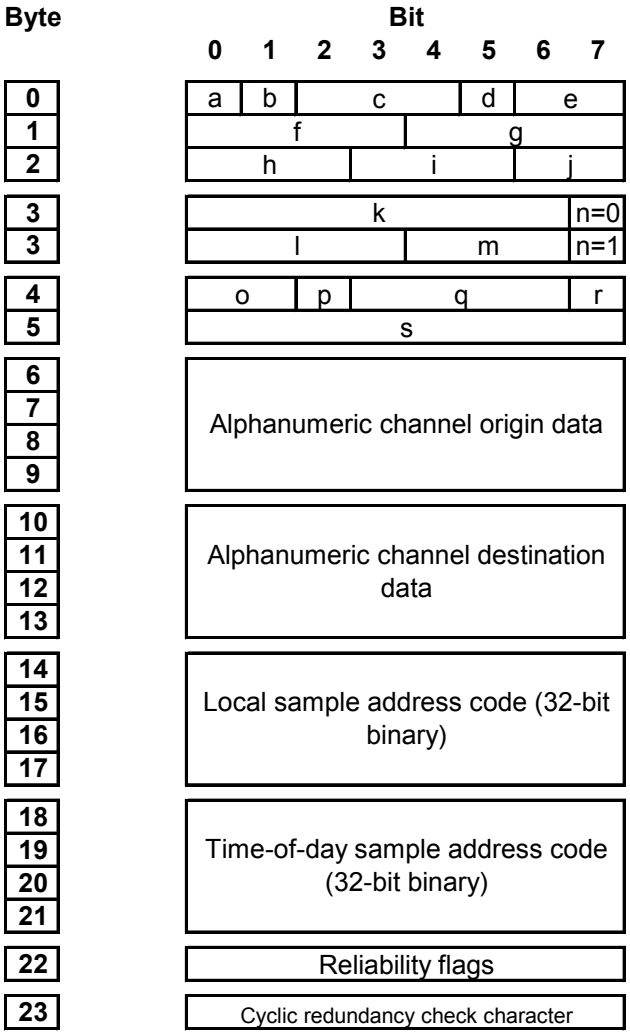
The default value of the user data bit is logic 0.

6 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.

The specific organization follows, wherein the suffix 0 designates the first byte or bit. Where multiple bit states represent a counting number, tables are arranged with most significant bit (MSB) first, except where noted as LSB first.



Key:

a	use of channel status block	**j**	indication of alignment level
b	linear PCM identification	**k**	channel number
c	audio signal pre-emphasis	**l**	channel number
d	lock indication	**m**	multichannel mode number
e	sampling frequency	**n**	multichanel mode
f	channel mode	**o**	digital audio reference signal
g	user bits management	**p**	reserved but undefined
h	use of auxiliary sample bits	**q**	sampling frequency
i	source word length	**r**	sampling frequency scaling flag
		s	reserved but undefined

Figure 5 - Channel status data format

Byte 0

bit	0	Use of channel status block
state	0	Consumer use of channel status block (see note).
	1	Professional use of channel status block.

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

bits	2 3 4	Audio signal emphasis
states	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 μs + 15 μs emphasis. Receiver manual override is disabled.
	1 1 1	International Telecommunication Union (ITU-T) J.17 emphasis (with 6,5-dB insertion loss at 800 Hz). Receiver manual override is disabled
	All other states of bits 2 to 4 are reserved and are not to be used until further defined.	

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

.

bits	6 7	Sampling frequency
states	0 0	Sampling frequency not indicated. Receiver default to interface frame rate 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 60958-3 consumer use can be identified, and a receiver conforming only to IEC 60958-3 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 to 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 to 3 indicate single channel double sampling frequency mode then the sampling frequency of the audio signal is twice that indicated by bits 6 to 7 of byte 0.

Byte 1

bits	0 1 2 3	Channel mode
states	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	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 to 3 are reserved and are not to be used until further defined.	

bits	4 5 6 7	User bits management
states	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.
	0 1 0 1	Reserved for metadata
	All other states of bits 4 to 7 are reserved and are not to be used until further defined.	

Byte 2

bits	0 1 2	Use of auxiliary sample bit
states	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 to 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 4 5	Encoded audio sample word length of transmitted signal. See notes 2, 3, and 4	
		Audio sample word length if maximum length is 24 bits as indicated by bits 0 to 2 above.	Audio sample word length if maximum length is 20 bits as indicated by bits 0 to 2 above.
states	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 to 5 are reserved and are not to be used until further defined.		

bits	6 7	Indication of alignment level
states	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 to 5 indicates that the number of active bits within the 20-bit 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 automatic set.

NOTE 3 The nondefault states of bits 3 to 5 indicate the number of bits within the 20-bit 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 to 5, the MSB is in time slot 27 of the transmitted subframe as specified in 4.1.1.

Byte 3

bit	7	Multichannel mode
state	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 to 6	Channel number, when byte 3 bit 7 is 0.
value	The channel number is the numeric value of the byte, with bit 0 as the least significant bit, plus one.	

OR,

bits	4 5 6	Multichannel mode, when byte 3 bit 7 is 1.
states *note: LSB first*	0 0 0	Multichannel mode 0. The channel number is defined by bits 0 to 3 of this byte.
	1 0 0	Multichannel mode 1. The channel number is defined by bits 0 to 3 of this byte.
	0 1 0	Multichannel mode 2. The channel number is defined by bits 0 to 3 of this byte.
	1 1 0	Multichannel mode 3. The channel number is defined by bits 0 to 3 of this byte.
	1 1 1	User-defined multichannel mode. The channel number is defined by bits 0 to 3 of this byte.
	All other states of bits 4 to 6 are reserved and are not to be used until further defined.	

bits	0 to 3	Channel number, when byte 3 bit 7 is 1
value	The channel number is the numeric value of these four bits, with bit 0 as the least significant bit, plus one.	

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 to 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
states	0 0	Not a reference signal (default).
	0 1	Grade 1 reference signal – see AES11
	1 0	Grade 2 reference signal – see AES11
	1 1	Reserved and not used until further defined.

bit	2	Reserved

bits	3 4 5 6	Sampling frequency
states *note: LSB first*	0 0 0 0	Not indicated (default).
	1 0 0 0	24 kHz
	0 1 0 0	96 kHz
	1 1 0 0	192 kHz
	0 0 1 0	Reserved.
	1 0 1 0	Reserved.
	0 1 1 0	Reserved.
	1 1 1 0	Reserved.
	0 0 0 1	Reserved for vectoring.
	1 0 0 1	22,05 kHz
	0 1 0 1	88,2 kHz
	1 1 0 1	176,4 kHz
	0 0 1 1	Reserved.
	1 0 1 1	Reserved.
	0 1 1 1	Reserved.
	1 1 1 1	User defined.

bit	7	Sampling frequency scaling flag
state	0	No scaling (default)
	1	Sampling frequency is 1/1,001 times that indicated by byte 4 bits 3 to 6, or by byte 0 bits 6 to 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 to 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 to 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 to 7	Reserved
value	Set to logic 0 until further defined	

Bytes 6 to 9

bits	0 to 7	Alphanumeric channel origin data.
value (each byte)	7-bit data with no parity bit complying with ISO 646, International Reference Version (IRV). LSBs are transmitted first with logic 0 in bit 7. First character in message is byte 6. Nonprinted control characters, codes 01_{16} to $1F_{16}$ and $7F_{16}$, are not permitted. Default value is logic 0 (code 00_{16}).	

Note: ISO 646, IRV, is commonly identified as 7-bit ASCII

Bytes 10 to 13

bits	0 to 7	Alphanumeric channel destination data.
value (each byte)	7-bit data with no parity bit complying with ISO 646, International Reference Version (IRV). LSBs are transmitted first with logic 0 in bit 7. First character in message is byte 6. Nonprinted control characters, codes 01_{16} to $1F_{16}$ and $7F_{16}$, are not permitted. Default value is logic 0 (code 00_{16}).	

Bytes 14 to 17

bits	0 to 7	Local sample address code
value (each byte)	32-bit binary value representing first sample of current block. LSBs are transmitted first. Default value is logic 0.	

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

Bytes 18 to 21

bits	0 to 7	Time-of-day sample address code
value (each byte)	32-bit binary value representing first sample of current block. 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 (that is, 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

bits	0 to 7	Reliability flags
bits	Flags used to identify whether the information carried by the channel status data is reliable. If reliable, the appropriate bits are set to logic 0 (default); if unreliable, the bits are set to logic 1.	
	0 to 3	Reserved and are set to logic 0 until further defined.
	4	Bytes 0 to 5.
	5	Bytes 6 to 13.
	6	Bytes 14 to 17.
	7	Bytes 18 to 21.

Byte 23

bits	0 to 7	Channel status data cyclic redundancy check character (CRCC).
value	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 7.2.1	

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.

7 Interface format implementation

7.1 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.

7.2 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 7.2.1, 7.2.2, and 7.2.3.

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 6 shall remain unchanged.

7.2.1 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, are implemented as required by an application, the interface transmitter shall be classified as a minimum implementation of channel status. See 7.2.2.

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 to 7 might cause improper operation in receiving devices not supporting manual override or auto set capabilities.

7.2.2 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 7.2.1 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.

7.2.3 Enhanced implementation of channel status

In addition to conforming to the requirements described in 7.2.2 for the standard implementation, the enhanced implementation shall provide further capabilities.

7.3 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.

8 Electrical requirements

8.1 General characteristics

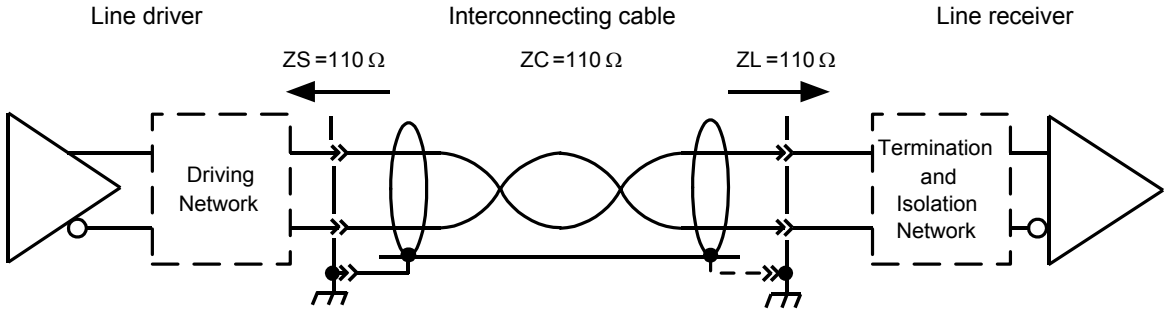
The electrical parameters of the interface are based on those defined in ITU-T recommendation V.11 which allow transmission of balanced-voltage digital signals up to a few hundred meters in length.

A circuit conforming to the general configuration shown in figure 6 may be used.

Although equalization may be used at the receiver, there shall be no equalization before transmission.

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 Ω at frequencies from 100 kHz to 128 times the maximum frame rate.



Note 1: Holding closer tolerances for the characteristic impedance of the cable, and for the driving and terminating impedances, can increase the cable lengths for reliable transmission and for higher data rates.

Note 2: Closer tolerances for the balance of the driving impedance, the terminating impedance, and for the cable itself can reduce both electromagnetic susceptibility and emissions.

Note 3: Using cable having lower loss at higher frequencies can improve the reliability of transmission for greater distances and higher data rates.

Figure 6 – Simplified example of the configuration of the circuit (balanced)

8.2 Line driver characteristics

8.2.1 Output impedance

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

8.2.2 Signal amplitude

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

8.2.3 Balance

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

8.2.4 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-Ω 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.

8.2.5 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.

8.2.5.1 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 jitter may be strongly asymmetric in character and the deviation from the ideal timing should meet the specification in either direction.

NOTE 2 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 3 The intrinsic-jitter measurement-filter characteristic is shown in figure 7. 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.

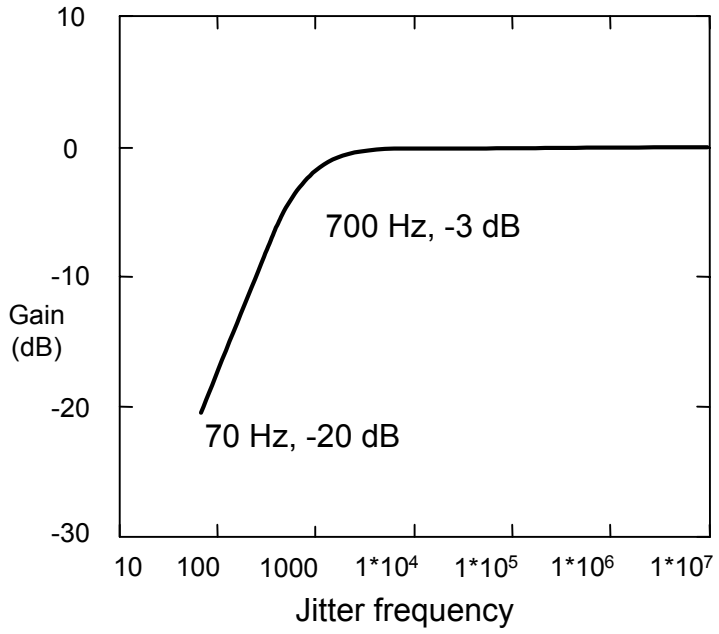


Figure7 — Intrinsic-jitter measurement-filter characteristic

8.2.5.2 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 8 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.

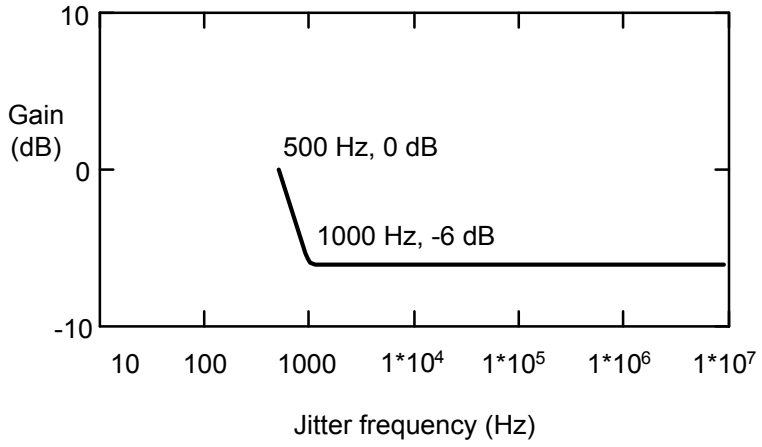


Figure 8 — Jitter transfer-function mask

8.3 Line receiver characteristics

8.3.1 Terminating impedance

The receiver shall present an essentially resistive impedance of 110 Ω with a tolerance of 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.

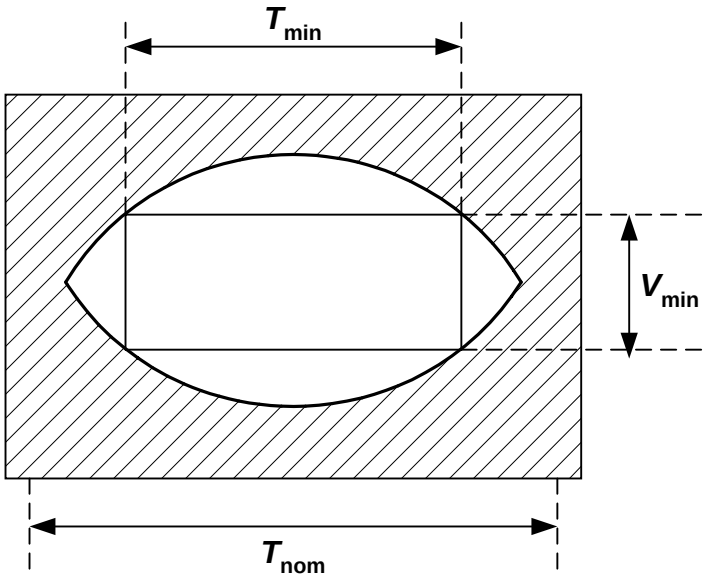
8.3.2 Maximum input signals

The receiver shall correctly interpret the data when connected directly to a line driver working between the extreme voltage limits specified in 8.2.2.

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

8.3.3 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_{nom}. See figure 9.



T_{min}: 0.5 × T_{nom}

V_{min}: 200 mV

T_{nom}: one-half the biphase symbol period

Figure 9 — Eye diagram

8.3.4 Receiver equalization

Equalization may 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 10. The receiver shall meet the requirements specified in 8.3.2 and 8.3.3.

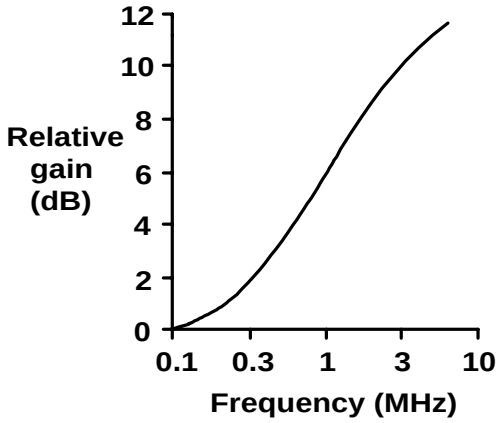


Figure 10 — Suggested equalizing characteristic for a receiver operating at 48 kHz frame rate

8.3.5 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 d.c. to 20 kHz.

8.3.6 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.

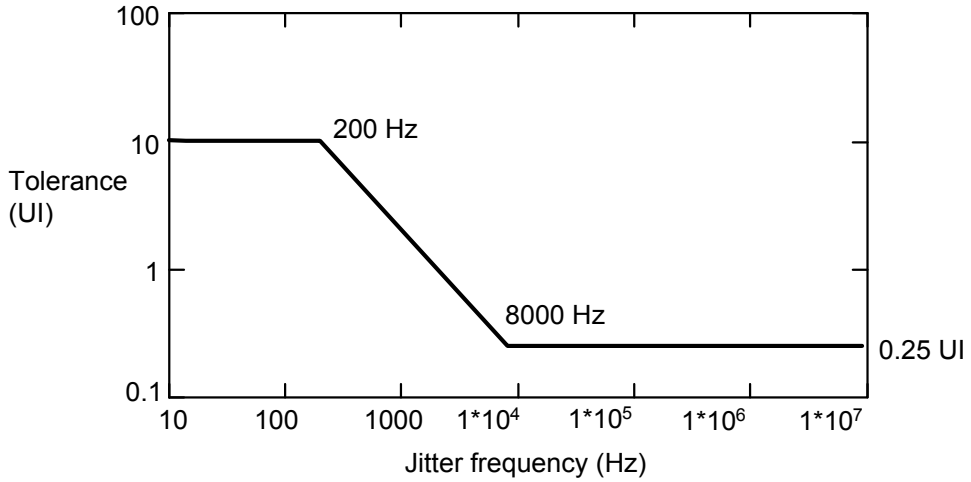


Figure 11 — Jitter tolerance template

8.4 Connectors
The standard connector for both outputs and inputs shall be the circular latching three-pin connector described in IEC 60268-12. Note: 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 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.

Annex A

(informative)

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.

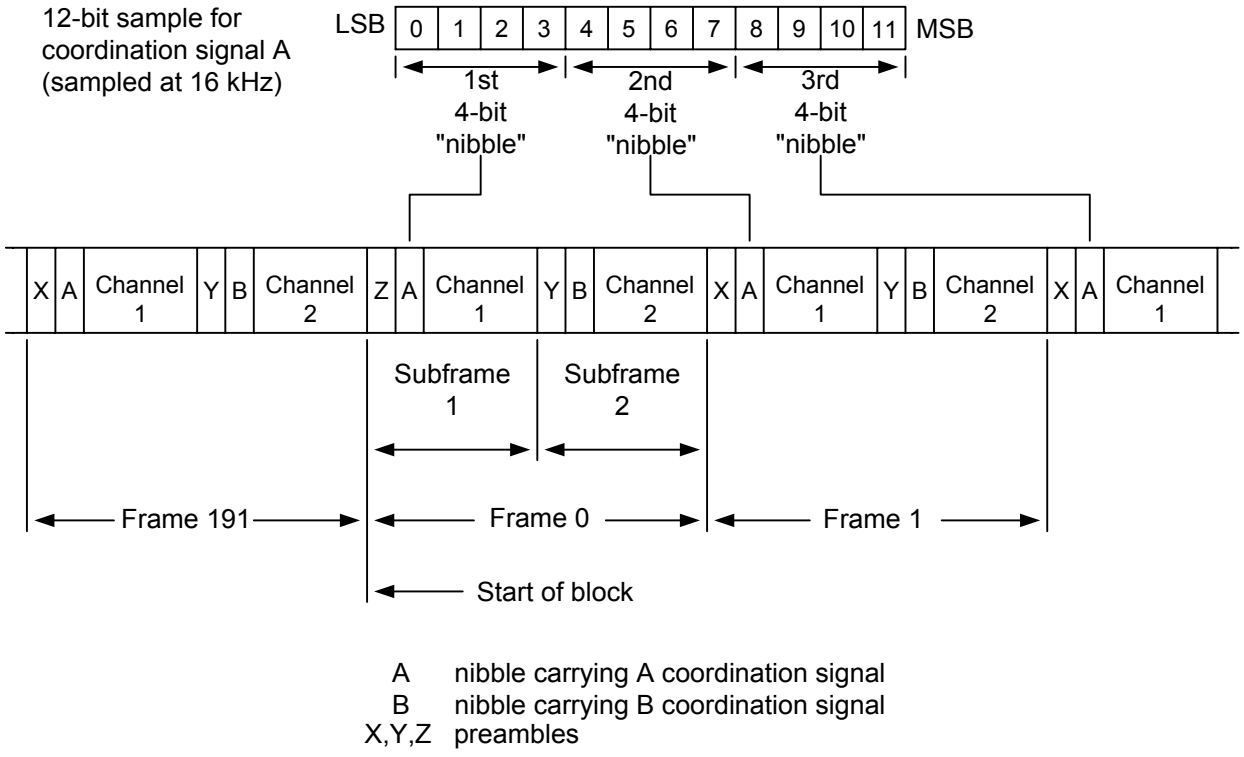


Figure A.1 — Frame and block structure

Annex B

(informative)

Generation of CRCC (byte 23) for channel status

The channel status block format of 192 bits includes a cyclic redundancy check (CRC) 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. l. The initial condition of all stages is logic 1.

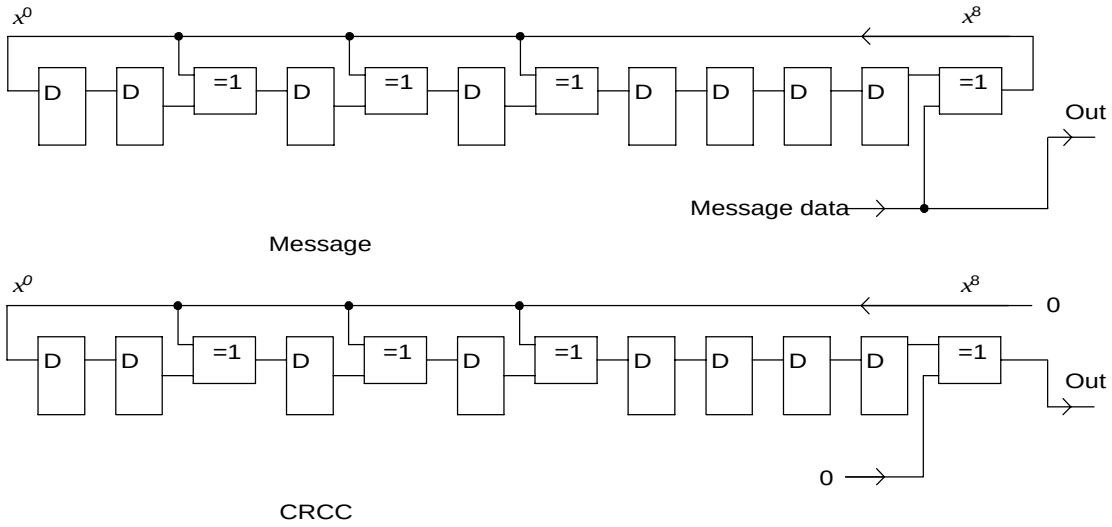


Figure B.1 — Flow diagram including exclusive-or gates

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:

Byte 23	Channel status data cyclic redundancy check character (CRCC)							
Bits	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Channel status bits	*184*	*185*	*186*	*187*	*188*	*189*	*190*	*191*
value	1	1	0	1	1	0	0	1

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:

Byte 23	**Channel status data cyclic redundancy check character (CRCC)**							
Bits	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Channel status bits	*184*	*185*	*186*	*187*	*188*	*189*	*190*	*191*
value	0	1	0	0	1	1	0	0

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

Annex C

(informative)

Informative references

AES-2id, *AES information document for digital audio engineering - Guidelines for the use of the AES3 interface*, Audio Engineering Society, New York, NY, USA.

AES-3id, *AES information document for digital audio engineering - Transmission of AES3 formatted data by unbalanced coaxial cable,* Audio Engineering Society, New York, NY, USA

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

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

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

IEC 60958-1, *Digital audio interface - Part 1: General*, International Electrotechnical Commission, Geneva, Switzerland.

IEC 60958-4, *Digital audio interface - Part 4: Professional applications*, International Electrotechnical Commission, Geneva, Switzerland.

ITU-R Recommendation BS647, *A digital audio interface for broadcasting studios*, International Telecommunication Union, Geneva, Switzerland.

ITU-T Recommendation J.17, *Pre-emphasis used on sound program circuits*, International Telecommunication Union, Geneva, Switzerland..

SMPTE Recommended Practice RP155-1997 *Reference level for digital audio systems.* Society of Motion Picture and Television Engineers. New York, NY, US.

Annex D

(informative)

AES3 signals on Structured Wiring

There is an increasing use of structured wiring (see EN50173 Information Technology - Generic Cabling Systems) for carrying AES3 signals. Notwithstanding Clause 8.1 of AES3-2003, this practice has been shown to be viable on Category 5 unscreened pairs, meeting EMI requirements, and offering transmission up to 400 metres overall unequalised, or 800 metres equalised, at 48kHz frame rate. (See AES preprint 3783, 96th convention, Amsterdam, February 1994 "Twisted-pair cables for AES/EBU Digital Audio Signals" D. G. Kirby, BBC R & DD, Table 3; later published in J. Audio Eng. Soc. Vol. 43, No. 3, 1995 March, pp. 137-146).

For a satisfactory outcome, this practice should only be used where the whole length of the connection uses Category 5 UTP (unscreened twisted pair) cable, including that which is installed and any flexible leads ("equipment cords"), and that the connection is unscreened for the whole length. Alternatively, the whole length of the connection should use Cat 5 STP (screened twisted pair), and the connection should be screened for the whole length.

Care should be taken in design of the interface to provide adequate balance on the twisted pair within the Category 5 cable.

Using RJ45 connectors, conventionally wired, current practice favours the use of pins 4 and 5 for AES3 signals (separating them from ATM signals on the same cable). Pins 3 and 6 are the preferred second pair. Note that, for full protection, the interface may have to withstand power voltages specified to support network equipment, and the use of transformers and blocking capacitors on the AES3 interface is strongly recommended.

Suggested practice: while the interface is by definition insensitive to polarity, for the purposes of constructing adaptors, XLR Pin 2 should be connected to RJ45 Pin 5 (or other odd-numbered pin), XLR Pin 3 should be connected to RJ45 Pin 4 (or even-numbered pin).