AES10-2003
Revision of AES10-1991

AES Recommended Practice for Digital Audio Engineering — Serial Multichannel Audio Digital Interface (MADI)

Published by

Audio Engineering Society, Inc.

Abstract

This standard describes the data organization for a multichannel audio digital interface. It includes a bit-level description, features in common with the AES3 two-channel format, and the data rates required for its utilization. The specification provides for the serial digital transmission of 32, 56, or 64 channels of linearly represented digital audio data at a common sampling frequency within the range 32 kHz to 96 kHz, having a resolution of up to 24 bits per channel. The format makes possible the transmission and reception of the complete 28-bit channel word (excluding preamble) as specified in AES3, providing for the validity, user, channel status, and parity information allowable under that standard. The transmission format is of the asynchronous simplex type and is specified for a single 75-Ω coaxial cable point-to-point interconnection or the use of fibre-optic cables.

Contents

Foreword

[This forward is not part of the *AES Recommended Practice for Digital Audio Engineering—Serial Multichannel Audio Digital Interface (MADI)*, AES10-xxxx.]

Foreword to original edition

This document was prepared as a result of a desire by four manufacturers of digital audio equipment to produce a common interface for serial transfer of digital multichannel audio in recording and broadcast studio applications. The benefit of such an interface is the simplification of multichannel digital audio equipment interconnection, commensurate with the obvious possibilities offered by the nature of the digitized audio signal. Throughout the development of the interface, the following intentions have been adhered to as faithfully as possible:

- The documentation produced by the group shall enter the public domain as soon as feasibility of the interface is established.
- The interface itself shall be simple to engineer and use.
- The cost and simplicity of the interface shall be such that the benefits of its use shall be easily justifiable.
- The interface shall not depend on the existence of hardware or software the rights to which are owned by any one or more members of the group.

The following individuals have contributed to the document: P. Eastty, T. Fujisawa, C. Jenkins, A. Jubb, P. Lidbetter, R. Salter, D. Ward, and J. Wilkinson.

R. A. Finger, chair
AES Standards Committee Working Group on Digital Input/Output Interfacing
1989 September

At the time of approval of this document for publication, the AES Standards Committee had the following membership: Y. Abe, J. S. Brawley, R. C. Cabot, P. D' Antonio, D. Eger (Chair), R. A. Finger, I. Joel, W. Hogan, T. F. Holman, M. Klasco, D. L. Klepper, B. N. Locanthi, J. P. Nunn, D. Queen (Secretary), T. Roseberry, W. T. Shelton, W. D. Storm, T. Telesky, H. Tendeloo, and F. E. Toole.

Foreword to second edition

This revision recognizes other uses to which the interface has been put, notably distributed routing and hence the increase in channels to the maximum of 64 at 48 kHz, the introduction of 96-kHz sampling in digital audio origination and the use of data transmission in the carrier system.

In this edition, figures 3 through 8 of AES10-1991 have been renumbered to figures 5 through 10. Clauses 2 though 6 of AES10-1991 have been renumbered to clauses 3 through 7, according to IEC guidelines.

This edition has been written by a writing group of SC-02-02. Contributors include R. Caine, C. Travis, R. Silfvast, and others.

J. Dunn, chair
R. A. Finger, vice-chair
SC-02-02 Working Group on Digital Input/Output Interfacing
2002-09-23

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

AES Recommended Practice for Digital Audio Engineering — Serial Multichannel Audio Digital Interface (MADI)

0 Patents

The Audio Engineering Society draws attention to patents concerning the functionality described in annex B. Those contemplating making use of this functionality should consider that any such use of sync symbols may be subject to patents. Applicable patents include GB2276796, US5487067 and JP7015458.

The AES holds no position concerning the evidence, validity and scope of these patent rights.

Attention is drawn to the possibility that some of the elements of this AES standard may be the subject of patent rights other than those identified herein. AES shall not be held responsible for identifying any or all such patent rights.

1 Scope

This standard describes the data organization and electrical characteristics for a multichannel audio digital interface (MADI). It includes a bit-level description, features in common with the two-channel format of the AES3, *AES Recommended Practice for Digital Audio Engineering — Serial Transmission Format for Linearly Represented Digital Audio Data*, and the data rates required for its utilization. The specification provides for the serial digital transmission over coaxial or fibre-optic lines of 28, 56, or 64 channels of linearly represented digital data at a common sampling frequency within the range of 32 kHz to 96 kHz having a resolution of up to 24 bits per channel. Only single-point to single-point interconnections from one transmitter to one receiver are supported.

2 Normative references

AES3-1992 (r1997) *AES Recommended Practice for Digital Audio Engineering -- Serial transmission format for two-channel linearly represented digital audio data*.

AES11-1997 *AES Recommended practice for digital audio engineering – Synchronization of digital audio equipment in studio operations*.

ISO/IEC 9314-3 (1990-12) *Information processing systems - Fibre Distributed Data Interface (FDDI) - Part 3: Physical Layer Medium Dependent (PMD)*. Geneva CH: International Electrotechnical Commission.

IEC 60169-8 (1978-01) *Radio-frequency connectors. Part 8: R.F. coaxial connectors with inner diameter of outer conductor 6.5 mm (0.256 in) with bayonet lock - Characteristic impedance 50 ohms (Type BNC)*. Geneva CH: International Electrotechnical Commission.

3 Definitions and abbreviations

3.1
audio sample data
audio signal that has been periodically sampled, quantized, and digitally represented in 2's complement form

3.2
channel
set of audio sample data related to one signal accompanied by other data bits transmitted in any one period of the source sampling frequency

3.3
two-channel format
bit, block, and subframe structure (less preambles) of the AES3 serial transmission format for linearly represented digital audio data

3.4
frame
sequence of 64 or less (typically 56 or 28) subframes designated using numbers 0 to 63, each carrying audio sample and related data that are transmitted in one sample period, with the start of a frame beginning with the first bit of subframe 0

3.5
link
connection between a single serial multichannel digital audio transmitter and a single multichannel digital audio receiver

3.6
sync symbol
decoder synchronization symbol

3.7
MADI
multichannel audio digital interface

3.8
NRZI
Transmission code as described in ISO 9314. Note: the abbreviation NRZI has also been used for other related encoding schemes.

4 Format

This specification provides for the serial digital transmission over coaxial or fibre-optic lines of 28, 56, or 64 channels of linearly represented digital data at a common sampling frequency within the range of 32 kHz to 96 kHz having a resolution of up to 24 bits per channel. See figure 1.

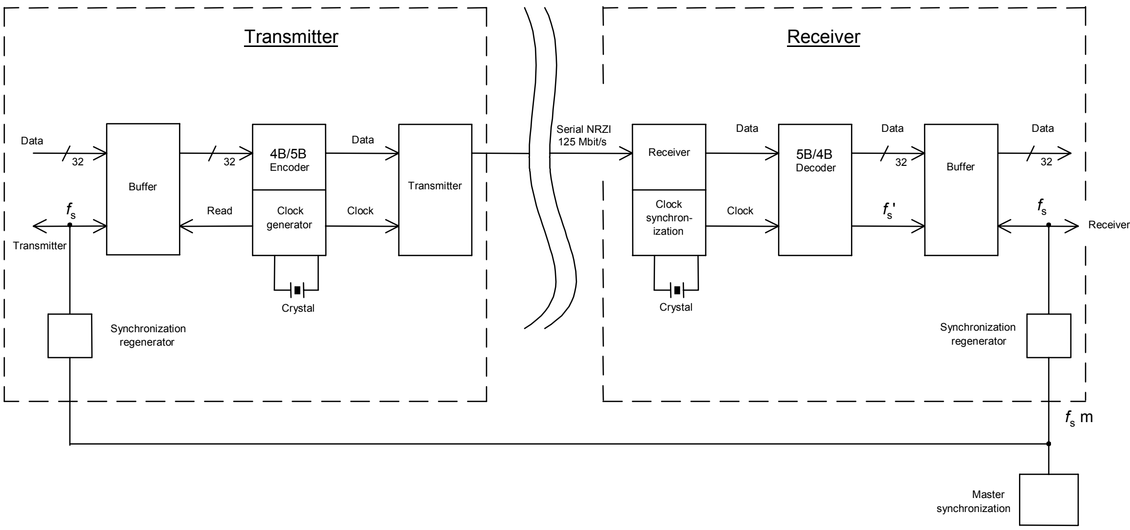

Figure 1 — Diagram of MADI

NOTE As sample rate changes non-return-to-zero inverted (NRZI) data rate stays constant; transmitter and receiver are asynchronous. Sampling frequencies (f_s) are 32 kHz to 96 kHz.

4.1 Frame format

Each frame consists of n channels, which are numbered from 0 to n – 1. The channels shall be consecutive within the frame, starting with channel 0 as shown in figures 2, 3, and 4.

NOTE 1 Synchronization symbols not shown.

NOTE 2 The period of each pattern is shown for the 48 kHz sampling frequency. It can be longer for lower frequencies and can vary with varispeed operation.

MADI subframe	0	1	2	3	4					54	55	0
Audio channel	Ch 0	Ch 1	Ch 2	Ch 3						Ch 54	Ch 55	Ch 0
Sample number	n	n	n	n						n	n	n + 1
AES3 subframe	A	B	A	B						A	B	A

← 20,8 µs →

Figure 2 — 48 kHz with 56 channels working

MADI subframe	0	1	2	3	4			52	53	54	55	0
Audio channel	Ch 0	Ch 0	Ch 1	Ch 1	Ch 2			Ch 26	Ch 26	Ch 27	Ch 27	Ch 0
Sample number	n	n + 1	n	n + 1	n			n	n + 1	n	n + 1	n + 2
AES3 subframe	A	B	A	B	A			A	B	A	B	A

← 20,8 μs →

Figure 3 — 96 kHz with 28 channels working (legacy pattern, as formed by legacy encoder fed by 28 single-channel double sampling-frequency mode signals per AES3)

MADI subframe	0	1	2		26	27	0	1		26	27	0
Audio channel	Ch 0	Ch 1	Ch 2		Ch 26	Ch 27	Ch 0	Ch 1		Ch 26	Ch 27	Ch 0
Sample number	n	n	n		n	n	n + 1	n + 1		n + 1	n + 1	n + 2
AES3 subframe	A	B	A		A	B	A	B		A	B	A

← 20,8 μs →

Figure 4 — 96 kHz with 28 channels working (96 kHz frame pattern, using 10,4 μs framing, one channel-zero flag per frame)

4.2 Channel format

Each channel consists of 32 bits, of which 24 are allocated to audio or to other data as defined by the audio/nonaudio status flag. A further 4 bits represent the validity (V), user (U), status (C), and parity (P) bits of the two-channel AES3 interface, with a further 4 bits allocated for mode identification. In this manner, the two-channel format of AES3 is preserved. The channel format is shown in figure 5.

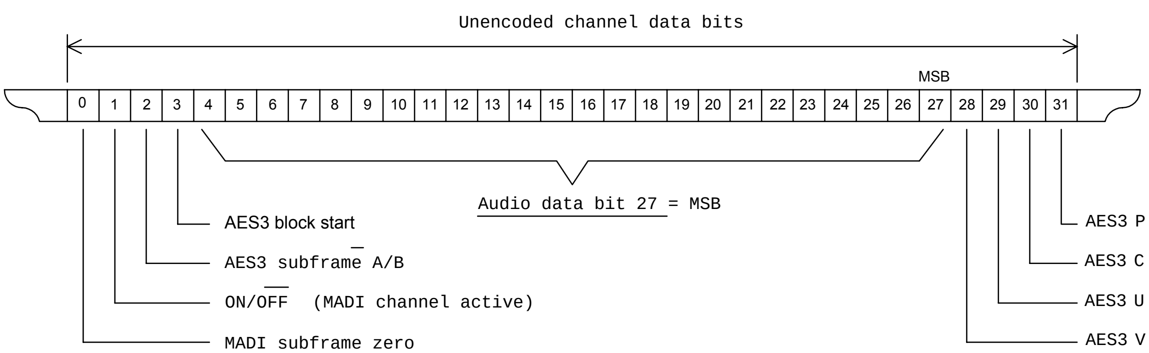

Figure 5 — Channel data format

4.2.1 Mode bits

The mode bits provide for frame synchronization, for block start per AES3, for identification of the A and B subframes also present in AES3, and for active/inactive status per channel.

> NOTE AES-10id recommends the use of the `NotA/B` flag in a multichannel function. It is possible that a receiver can meet this realization, and should be able to default to alternate `A/B` operation if the multichannel function is presented.

4.2.2 Audio data representation

In the audio mode, the 24-bit format is represented linearly in 2's complement form, with the most significant bit (MSB) transmitted last. All unused audio bits within a channel shall be set to zero, with the V, U, C, and P bits set to default values, as defined by the AES3 two channel format.

4.2.3 Active channels
All active channels shall be consecutive, starting at channel zero. The channel active bit shall be set to 1 within each active channel.

4.2.4 Inactive channels
All inactive channels shall have all bits set to zero, including the channel active bit. Inactive channels shall always have a higher channel number than the highest numbered active channel.

4.2.5 Bit description
See tables 1 and 2.

Table 1 — Bit description

Bit	Name	Description	Sense
0	MADI subframe 0	Frame synchronization bit	1 = true
1	MADI channel active	Channel active bit	1 = true
2	AES3 subframe A/B	AES3 subframe identifier	1 = B
3	AES3 block start	First frame of AES3 block	1 = true
4-27	AES3 data bits	(bit 27 shall be MSB)	
28	AES3 V	Validity bit	0 = valid
29	AES3 U	User bit	true to AES3
30	AES3 C	Channel Status bit	true to AES3
31	AES3 P	Parity bit (excludes bits 0-3)	Even

Table 2 — Bits 2 - 3 compatibility with AES3

Bit 2	Bit 3	Two-Channel Form	Description
0	0	Form 2	A subframe
0	1	Form 1	A subframe status block start
1	0	Form 3	B subframe
1	1	Form 4*	B subframe status block start

*Does not conform to AES3.

4.3 Transmission format

The channels shall be transmitted serially by polarity-free coding, as defined in ISO 9314 (according to annex A) using a 4-bit to 5-bit encoding format, known as 4B5B, (according to annex B).

4.3.1 Encoding scheme
For the purposes of encoding, the 32-bit channel data shall be broken down into 8 words of 4 bits each, as shown in table 3.

Table 3 — 32-bit channel data

Word	Channel data bit
0	0123
1	4567
2	89 . .
3	. . . .
4	. . . .
5	. . . .
6	. . . .
7	. . 31

Each 4-bit word shall be encoded into a 5-bit word using the 4B5B coding scheme shown in table 4.

Table 4 — 5-bit word coding

4-Bit data	5-Bit encoded data
0000	11110
0001	01001
0010	10100
0011	10101
0100	01010
0101	01011
0110	01110
0111	01111
1000	10010
1001	10011
1010	10110
1011	10111
1100	11010
1101	11011
1110	11100
1111	11101

Each 5-bit encoded word shall be transmitted from the left, as defined in table 5.

Table 5 — 5-bit word transmission

Word	Channel link bit
0	01234
1	56789
2	. . . . .
3	. . . . .
4	. . . . .
5	. . . . .
6	. . . . .
7	. . . 39

This scheme enables a low direct-current (d.c.) bias to be maintained on the link. Although the link signal is nearly d.c. free, the audio signal may contain d.c. Figure 6 shows the link transmission format for one channel. Annex A illustrates the encoding process for a single-channel word.

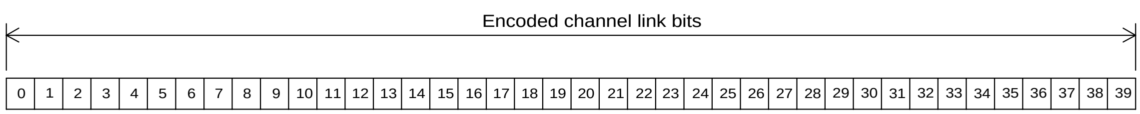

Figure 6 — Channel link format

4.3.2 4B5B synchronization symbol (sync symbol)

A 4B5B sync symbol shall be inserted into the data stream at least once per frame period to ensure transmitter and receiver synchronization of the 4B5B decoder in the receiver. Sufficient 4B5B sync symbols shall be inserted by interleaving with the encoded data words to fill the total link capacity. The 4B5B sync symbol is transmitted from the left. The 4B5B sync symbol may only be inserted at 40-bit channel boundaries, but may be repeated between channels or during the idle period after the last channel has been transmitted in each frame capacity, or both. The order placement of 4B5B sync symbols is not specified. Some examples of permissible positions of the 4B5B sync symbol are shown in figure 7.

The default 4B5B sync symbol shall be `11000 10001`. There are 32 synchronization symbols specified in FDDI. Other symbols may be used in order to carry, for example, control data not associated with any audio channel. Annex B outlines this function.

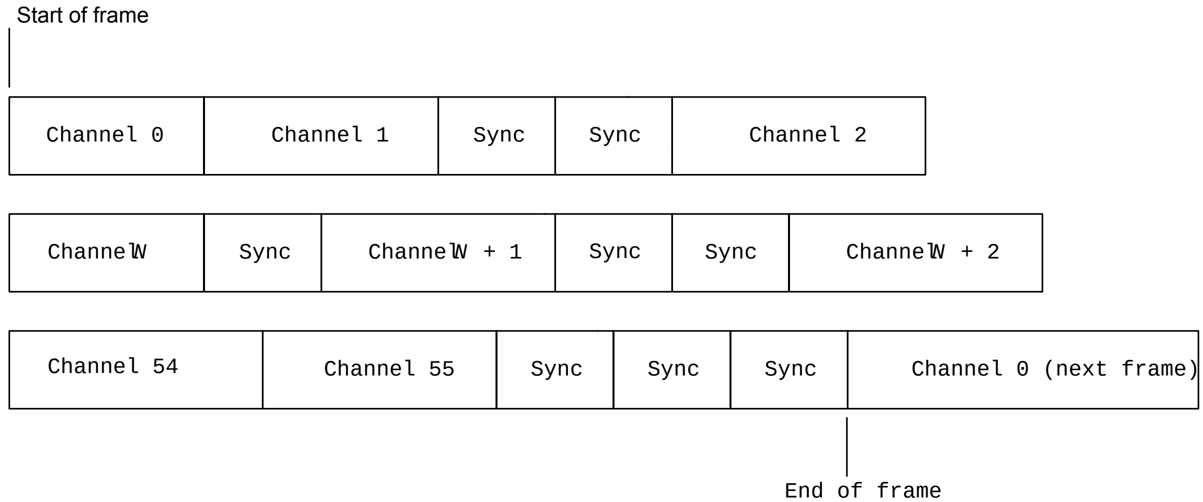

Figure 7 — Some permissible 4B5B sync symbol positions

4.3.3 Sequence of transmission

In any bit sequence, the left-hand symbol shall always represent the first in time.

5 Sampling frequency and data rates

5.1 Sampling frequency

The nominal sampling frequency at which the link operates shall be within one of three ranges:

a) 32 kHz to 48 kHz ± 12,5 %, 56 channels;
b) 32 kHz to 48 kHz nominal, 64 channels;
c) 64 kHz to 96 kHz ± 12,5 %, 28 channels.

Higher sampling frequencies may be accommodated (for example, 192 kHz) by using two or more channels per audio sample on the link.

NOTE 1 The provision of 56 channels at 48 kHz ± 12,5 % or of 28 channels at 96 kHz ± 12,5 % results in a maximum used data rate of 96,768 Mbit/s. 64 channels at 48 kHz or 32 channels at 96 kHz results in a maximum used data rate of 98,304 Mbit/s.

NOTE 2 The provision of 56 channels at 32 kHz ± 12,5 % results in a minimum used data rate of 50,176 Mbit/s.

5.2 Link transmission rate

The link transmission rate shall be 125 Mbit/s irrespective of the sampling frequency or number of active channels. See annex A.

5.3 Data transfer rate

The data transfer rate shall be 100 Mbit/s. The difference between the data transfer rate and the link transmission data rate is caused by the use of an encoding scheme. See 4.3.1.

6 Synchronization

This clause covers the sample synchronization of transmitters and receivers relative to a master synchronizing signal. It does not apply in the case of a master-slave connection only.

6.1 Sampling frequency synchronizing signal (synchronization signal)

Each transmitter and receiver shall be provided with an independently distributed master synchronizing signal. This signal shall be in accordance with AES11. Alternatively, a Society of Motion Picture and Television

Engineers or European Broadcasting Union video reference signal may be used in accordance with the timing reference point provisions of AES11.

A 96-kHz coded link shall be capable of being synchronized by either a 48-kHz or 96-kHz reference signal.

6.2 Sample timing

The link is not intended to carry sample timing information. The exact timing of connected equipment shall be controlled by the independently distributed master synchronizing signal, not by the MADI.

6.3 Transmitted frame start time

In order to maintain constant latency, the frame start time output from a transmitter shall be within ± 5 % of a sample period of the reference time defined by the transmitter's externally supplied master synchronizing signal.

6.4 Received frame start time

A receiver shall be able to correctly interpret a signal of any phase relative to the sample period of the externally supplied master synchronizing signal. Constant latency shall be maintained with a signal whose frame start time is within ± 25 % of a sample period of the reference time defined by the receiver's externally supplied master synchronizing signal.

7 Electrical characteristics

The transmission medium shall be either 75-Ω coaxial cable (see 7.1) or fibre-optic cable (see 7.2). For the purposes of transmission characterization, the data input to the encoder shall be replaced with a pseudorandom data generator having a sequence length of at least $2^{16} – 1$.

> NOTE The random data are applied prior to the 4-bit to 5-bit encoder in order to represent accurately those signals most likely to appear in normal transmission.

7.1 Coaxial cable

7.1.1 Transmitter

7.1.1.1 Line driver

The line driver shall have a single-ended output having an output impedance of 75 Ω ± 2 Ω. The connection between the emitter-coupled logic (ECL) signal transmitter, for example, and the coaxial cable may be achieved by the circuits shown in figure 8.

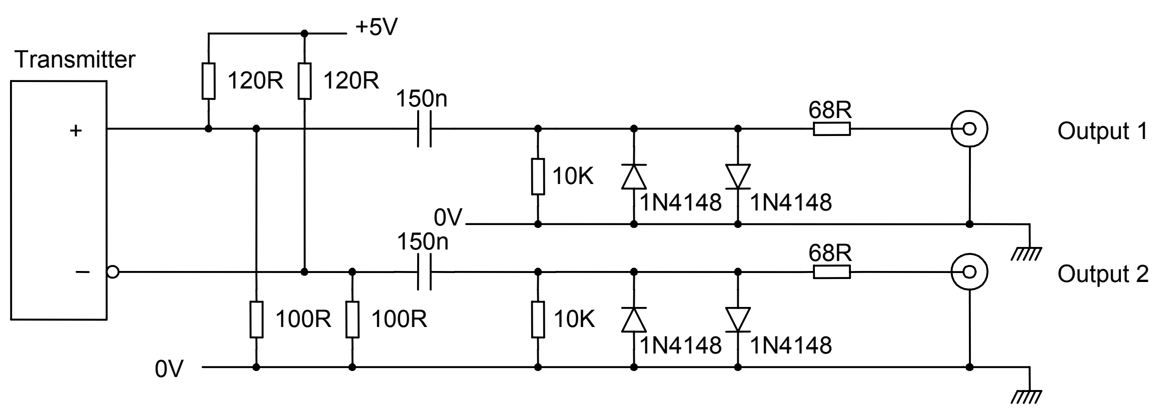

Figure 8 — MADI transmitter circuit buffer

7.1.1.2 Mean output

The average voltage of the output when terminated shall be 0 V ± 0,1 V with reference to the signal ground terminal.

7.1.1.3 Peak output

The peak-to-peak voltage of the output when terminated by a 75- Ω resistor shall be between 0,3 V and 0,6 V.

7.1.1.4 Rise and fall times

When the output is terminated by a 75- Ω resistor, the rise and fall times measured between the 20 % and 80 % amplitude points shall be no longer than 3 ns and no shorter than 1 ns, and the relative timing difference to the average of the amplitude points shall be no more than ± 0,5 ns.

7.1.2 Receiver

7.1.2.1 Eye pattern

The eye pattern represented by the characteristics of figure 9 shows the range of signals at the input terminals that should be decoded by a conformant receiver.

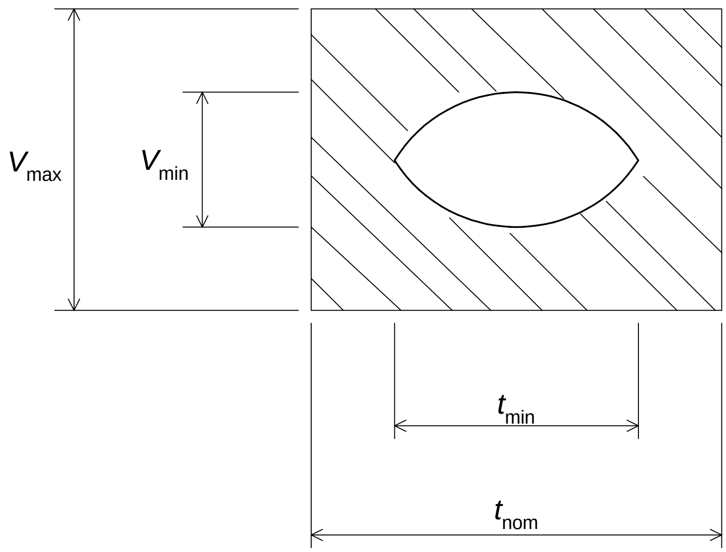

Figure 9 — Eye pattern diagram for maximum and minimum input signals: t_{nom} = 8 ns; t_{min} = 6 ns; V_{max} = 0,6 V; V_{min} = 0,15 V.

7.1.3 Cable

The coaxial cable shall have a 75- Ω ± 2 Ω characteristic impedance.

7.1.4 Connectors

BNC connectors to IEC 60169-8 shall be used throughout.

7.1.5 Interface circuit example

The connection between the coaxial cable medium and a balanced emitter-coupled logic (ECL) signal may be achieved by the circuit illustrated in figure 10.

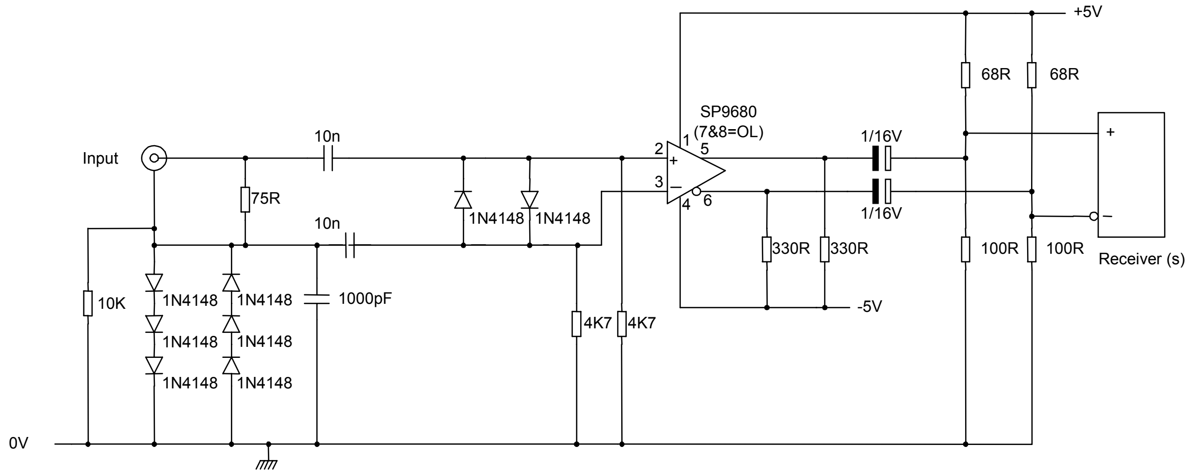

Figure 10 — MADI buffer circuits

7.1.6 Grounding

The coaxial cable shield shall be grounded at the transmitter. The coaxial cable shall be grounded to the receiver chassis at radio frequencies above 30 MHz.

For the purpose of minimizing radio frequency emissions it is recommended that the connection be achieved by direct bonding of the coaxial cable body to the equipment chassis. At the receiver this may be achieved by capacitive bonding of the coaxial cable connector body to the receiver chassis. A suitable value of capacitor is 1000 pF. The capacitor should be a low-inductance type, having a sufficiently low impedance at all frequencies from 30 MHz to 500 MHz. The lead bonding lengths shall be kept as small as practical. This method prevents the possibility of audio frequency ground currents.

Note: Designers should note that specialised techniques, described in appropriate literature, are required in order that the interface meet international regulations for electromagnetic compatibility (EMC). Bonding the receiver coaxial outer to the enclosure at DC with a total 360 degree connection is preferred if other considerations do not preclude it.

7.2 Fibre optic interfacing

7.2.1 Fibre type

A fibre interface should be used as specified according to ISO/IEC 9314-3. It should be a graded-index fibre with a core diameter of 62,5 nm, nominal cladding diameter of 125 nm and a numerical aperture of 0,275, at a wavelength of 1300 nm. This specification can provide a range of up to 2 km.

7.2.2 Connectors

The ST1 connector listed on the databases page of www.aes.org/standards/ should be used. It is designed to be optically and mechanically compatible with the media interface connector (MIC) according to ISO/IEC 9314-3.

Annex A

(Informative)

Example of link encoding

Suppose the channel data is as follows:

	0		1			2		3
Bit:	0123	4567	8901	2345	6789	0123	4567	8901
Data:	1100	1010	0101	1111	0000	1100	0011	0000

These data words translate into the following:

Word	4-bit data	5-bit encoded data
0	1100	11010
1	1010	10110
2	0101	01011
3	1111	11101
4	0000	11110
5	1100	11010
6	0011	10101
7	0000	11110

The transmitted bit stream is thus:

	0		1		2		3	
Bit:	01234	56789	01234	56789	01234	56789	01234	56789
4B5B code :	11010	10110	01011	11101	11110	11010	10101	11110
Transmission code:	01001	10010	00110	10100	10101	10110	01100	10101

←——— Direction of transmission

NOTE 1: The tolerance of the link transmission rate of 125 Mbit/s as mentioned in 5.2 should be ± 100 ppm.

Annex B

(Normative)

Use of 4B5B sync symbols to carry channel-independent data

B.1 Introduction

This annex describes in outline a method of carrying control data in the MADI transport carrier independent of any particular audio channel. The transport sync symbol words inserted between audio data words can carry this control data by virtue of the fact that there are a number of forms of sync symbol, of which the default is that used by MADI systems complying with AES10. Four-bit nibbles are associated with 16 of the sync symbol forms, thus allowing data to be inserted in the available space. The default sync symbol described in 4.3.2 is associated with the binary value 0000.

A legacy MADI stream at full occupancy and highest permitted varispeed rate uses 96,768 Mbit/s, and a 64-channel 48 kHz stream uses 98,304 Mbit/s. Thus there will always be at least 1 Mbit/s for this data. This may need to be reduced to ensure that bit-stream synchronization is maintained.

B.2 Data insertion

B.2.1 Ordering

Default sync symbol words shall be transmitted at least as often as required, in order to guarantee correct data-recovery of the whole transport stream. Coded sync symbols shall be inserted as and when required, subject to the needs of audio data and the provision above.

B.2.2 Data coding

This is not specified in this standard, or in annex A.. A format relating to the high-level data link control protocol uses the look-up table below as an example.

Table B.1 — Data coding look-up table

Command number	Command symbol	Name of symbol	Function
0	11000 10001	JK	Sync
1	11111 11111	II	Not used
2	01101 01101	TT	Not used
3	01101 11001	TS	Not used
4	11111 00100	IH	SAL
5	01101 00111	TR	Not used
6	11001 00111	SR	Not used
7	11001 11001	SS	Not used
8	00100 00100	HH	HDLC 0
9	00100 11111	HI	HDLC 1
A	00100 00000	HQ	HDLC 2
B	00111 00111	RR	HDLC 3
C	00111 11001	RS	HDLC 4
D	00000 00100	QH	HDLC 5
E	00000 11111	QI	HDLC 6
F	00000 00000	QQ	HDLC 7

Annex C

(Informative)

Informative references

References for the 4B5B scheme are on the databases page of www.aes.org/standards/.