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 AES3-1985 two-channel format, and the data rates required for its utilization. The specification provides for the serial digital transmission of 56 channels of linearly represented digital audio data at a common sampling frequency within the range 32kHz to 48kHz (±12.5%), 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 AES3-1985, 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-ohm coaxial cable point-to-point interconnection. The use of fiber-optic medium is also possible, although not yet specified.

Foreword

[This forward is not part of the AES Recommended Practice for Digital Audio Engineering—Serial Multichannel Audio Digital Interface (MADI), AES10-1991.]

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.

ROBERT A. FINGER, *Chairman*
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: Yoshi-Haru Abe, James S. Brawley, Richard C. Cabot, Peter D’ Antonio, Donald Eger (Chair), Robert A. Finger, Irving Joel, William Hogan, Tomlinson F. Holman, Mike Klasco, David L. Klepper, Bart N. Locanthi, J. P. Nunn, Daniel Queen (Secretary), Tom Roseberry, W. T. Shelton, William D. Storm, Ted Telesky, Han Tendeloo, and Floyd E. Toole.

The American National Standards Institute version of this standard has not been reprinted and remains available as ANSI S4.43-1991.

Contents

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

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 AES-1985 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 fiber-optic lines of 56 channels of linearly represented digital data at a common sampling frequency within the range of 32kHz to 48kHz (±12.5%) having a resolution of up to 24 bits per channel (Fig. 1). Only single-point to single-point interconnections from one transmitter to one receiver are supported.

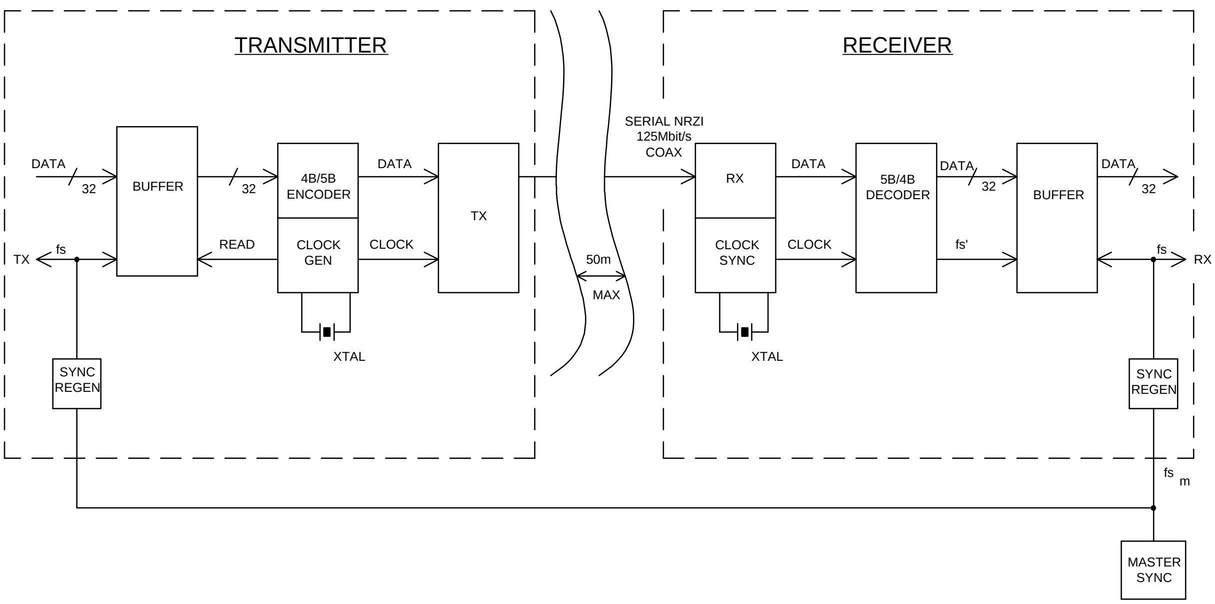

Fig. 1. Diagram of MADI. (*Note*: As sample rate changes NRZI data rate stays constant; TX and RX are asynchronous; samples rate 32 kHz to 48 kHz, ±12.5%.)

2 Terminology

2.1 Audio Sample Data

An audio signal that has been periodically sampled, quantized, and digitally represented in 2's complement form.

2.2 Channel

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

2.3 Two-Channel Format

The bit, block, and subframe structure (less preambles) of the AES3-1985 (ANSI S4.40-1985) serial transmission format for linearly represented digital audio data.

2.4 Frame

A sequence of 56 channels numbered 0 to 55, each carrying audio sample and related data that are transmitted in one sample period. The start of a frame begins with the first bit of channel 0.

2.5 Link

A connection between a single serial multichannel digital audio transmitter and a single multichannel digital audio receiver.

3 Format

3.1 Frame Format

Each frame consists of 56 channels, which are numbered from 0 to 55. The channels are consecutive within the frame, starting with channel 0. The frame format is shown in Fig. 2.

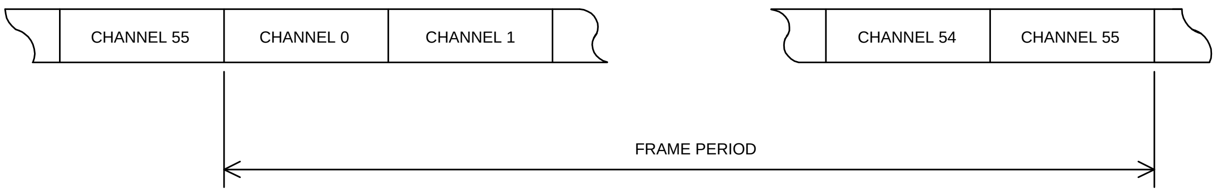

Fig.2. Frame format. (*Note*: Synchronization symbols not shown.)

3.2 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 AES3-1985 interface, with a further 4 bits allocated for mode identification. In this manner, the two-channel format of AES3-1985 is preserved. The channel format is shown in Fig. 3.

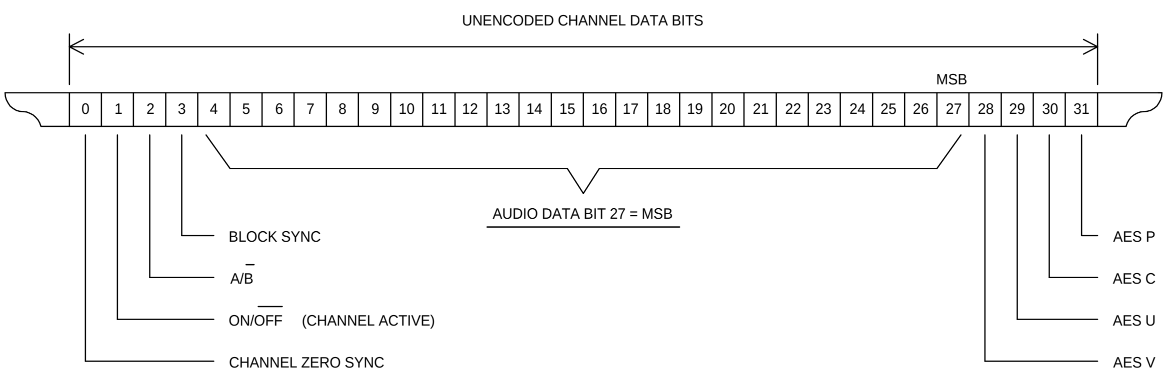

Fig. 3. Channel data format.

3.2.1 Mode Bits

The mode bits provide for frame synchronization, for block start per AES3-1985, for identification of the A and B subframes also present in AES3-1985, and for active/inactive status per channel.

3.2.2 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 AES3-1985 two channel format.

3.2.3 Active Channels

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

3.2.4 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.

3.2.5 Bit Description

See Tables 1 and 2.

Table 1

Bit	Name	Description	Sense
0	MADI channel 0	Frame synchronization bit	1 = true
1	MADI channel active	Channel active bit	1 = true
2	MADI channel A/B	AES3-1985A/B subframes	1 = B
3	MADI channel block sync	Channel block start	1 = true
4-27	AES data bits	(bit 27 shall be MSB)	
28	AES V	Validity bit	0 = true
29	AES U	User data bit	0 = true
30	AES C	Status data bit	0 = true
31	AES P	Parity bit (excludes bits 0-3)	Even

Table 2

Bits 2 - 3 compatibility with AES3-1985.

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

*Not per AES3-1985.

3.3 Transmission Format

The 56 channels are transmitted serially by non-return-to-zero inverted (NRZI) polarity-free coding, using a 4-bit to 5-bit encoding format.[1]

3.3.1 Encoding Scheme

For the purposes of encoding, the 32-bit channel data is broken down into 8 words of 4 bits each, as shown in Table 3.

Table 3

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

[1] As proposed in document 84-48 of American National Standards Institute accredited Standards Committee X3, Subcommittee T9.5.

Each 4-bit word is encoded into a 5-bit word, as shown in Table 4.

Table 4

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 is transmitted from the left, as defined in Table 5.

Table 5

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

This scheme enables a low direct-current (dc) bias to be maintained on the link. Although the link signal is nearly dc free, the audio signal may contain direct current. Fig. 4 shows the link transmission format for one channel. For the purposes of clarity, Appendix A illustrates the encoding process for a single-channel word.

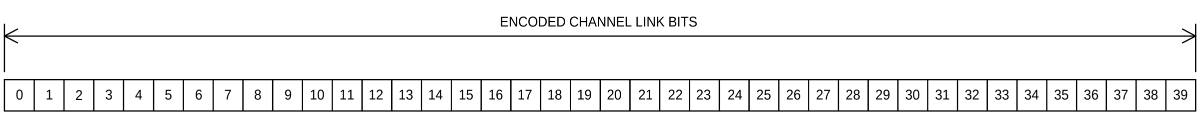

Fig. 4. Channel link format.

3.3.2 Synchronization Symbol

A synchronization symbol, 11000 10001, shall be inserted into the data stream at least once per frame period to ensure transmitter and receiver synchronization. Sufficient synchronization symbols shall be inserted by interleaving with the encoded data words to fill the total link capacity. The synchronization symbol is transmitted from the left. The synchronization symbol may only be inserted at 40-bit channel boundaries, but may be repeated between channels or during the idle period or both after the last channel has been transmitted in each frame capacity. The order placement of synchronization signals is not specified. Some examples of permissible positions of the syncronization symbol are shown in Fig. 5.

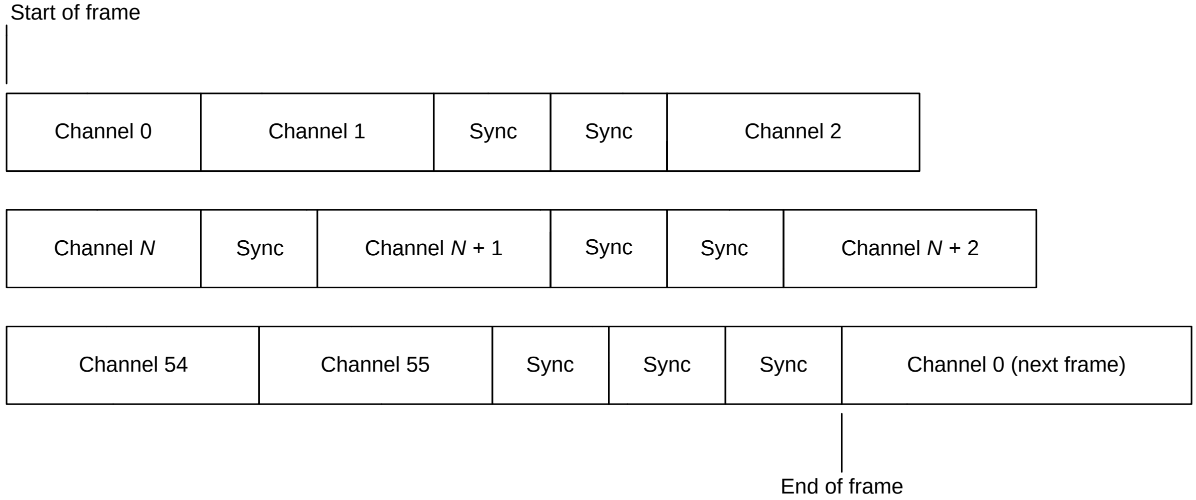

Fig. 5. Some permissible synchronization symbol positions.

3.3.3 Sequence of Transmission

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

3.3.4 NRZI Encoding

The link channel data are transmitted using an NRZI polarity-free encoding. Each high bit of the NRZ data is converted to a transition from the last bit; each low bit is represented by no transition. Therefore a 1 results in a high-to-low or a low-to-high transition and a 0 results in a static high or a static low.[2]

4 Sample Frequency and Rates

4.1 Sample Frequency

The nominal sample frequency at which the link operates is within the range of 32.0kHz to 48.0kHz. In addition, the facility to operate at ±12.5% of these frequencies is provided. Higher sample frequencies may be accommodated (e.g., 96kHz) by using two or more channels per audio sample on the link.

4.2 Link Transmission Rate

The link transmission rate shall be 125 megabits per second irrespective of the sample rate or number of active channels. (See Appendix A.)

4.3 Data Transfer Rate

The data transfer rate shall be 100 megabits per second. The difference between the data transfer rate and the link transmission data rate is due to the use of an encoding scheme. (See paragraph 3.3.1.)

4.4 Maximum Used Data Rate

The provision of 56 channels at 48kHz +12.5% gives a maximum used data rate of 96.768 megabits per second.

4.5 Minimum Used Data Rate

The provision of 56 channels at 32kHz –12.5% gives a minimum used data rate of 50.176 megabits per second.

[2] An example of the NRZ-to-NRZI conversion is given in the Appendix A.

5. Synchronization

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

5.1 Synchronizing Signal

Each transmitter and receiver shall be provided with an independently distributed master synchronizing signal. This signal shall be in accordance with AES3-1985.

5.2 Sample Timing

The link is not intended to carry sample timing information. The exact timing of connected equipment is controlled by the independently distributed master synchronizing signal, not by the multichannel audio digital interface.

5.3 Transmitted Frame Start Time

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.

5.4 Received Frame Start Time

A receiver shall be able to interpret correctly 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.

6. Electrical Characteristics

The transmission medium shall be either 75-ohm coaxial cable (see 6.1) or fibre optic (see 6.2). 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 that 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.

6.1 Coaxial Cable

6.1.1 Transmitter

6.1.1.1 Line Driver

The line driver shall have a single-ended output having an output impedance of 75 ohm ±2 ohm. The connection between the emitter-coupled logic (ECL) signal transmitter, for example, and the coaxial cable may be achieved by the circuits shown in Fig. 6.

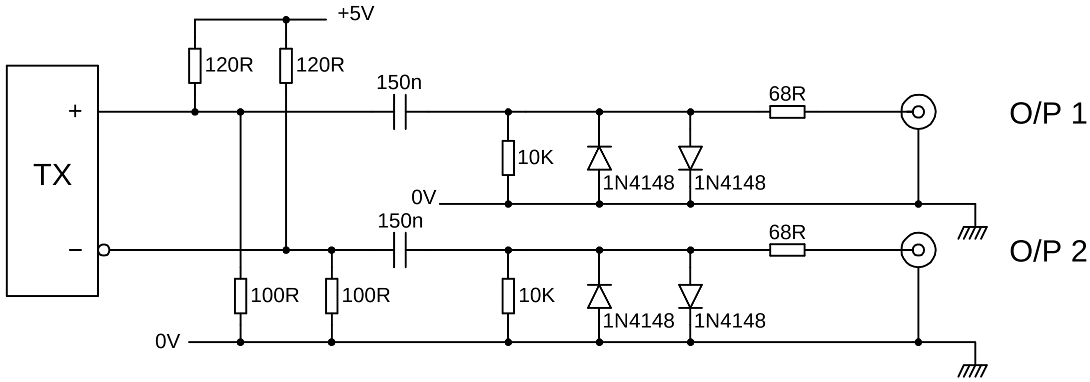

Fig. 6. MADI transmitter circuit buffer.

6.1.1.2 Mean Output

The average voltage of the output when terminated shall be 0V ±0.1V with reference to the signal ground terminal.

6.1.1.3 Peak Output

The peak-to-peak voltage of the output when terminated by a 75-ohm resistor shall lie between 0.3V and 0.6V.

6.1.1.4 Rise and Fall Times

When the output is terminated by a 75-ohm 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.

6.1.2 Receiver

6.1.2.1 Eye Pattern

The receiver shall decode the received signal correctly when the eye pattern is represented by the characteristics of Fig. 7 measured at the input terminal.

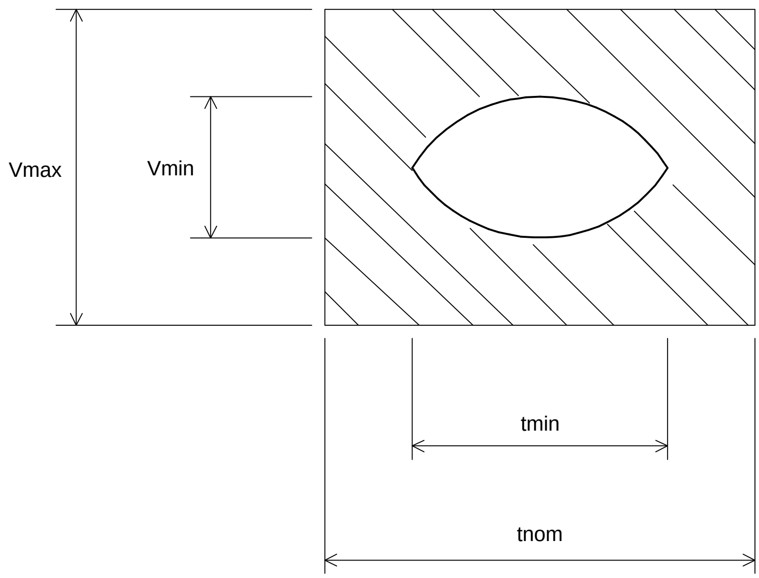

Fig. 7. Eye pattern diagram for maximum and minimum input signals: t_{nom} = 8 ns; t_{min} = 6 ns; V_{max} = 0.6 V; V_{min} = 0.15 V.

6.1.3 Cable

The coaxial cable shall have a 75 ohm ±2-ohm characteristic impedance and an attenuation of <0.1 dB/m within the range 1MHz to 100MHz.

6.1.4 Connectors

BNC connectors shall be used throughout. All equipment shall utilize chassis mounting connectors that have protruding bayonet lugs. Interconnecting cables shall utilize mating free connectors.

6.1.5 Cable Length

The maximum cable length shall be 50 m. Equalization is not permitted. The signal shall have an eye pattern no worse than that defined by Fig. 7.

6.1.6 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 Fig. 8.

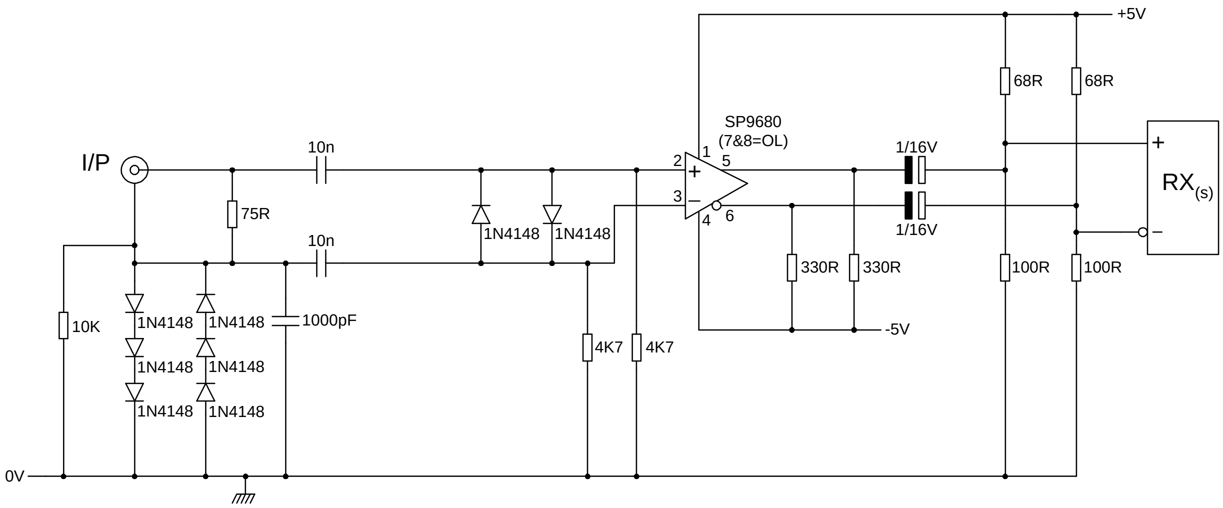

Fig. 8. MADI buffer circuits.

6.1.7 Grounding

The coaxial cable shield shall be grounded at the transmitter. 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. The coaxial cable shall be grounded to the receiver chassis at radio frequencies above 30 MHz. It is recommended that this connection 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.

6.2 Fiber Optic

Not yet defined.

[This appendix is not part of the AES Recommended Practice for Digital Audio Engineering – Serial Multichannel Audio Digital Interface (MADI), AES10-1991, but is included to facilitate its use.]

Appendix A

Example of Link Encoding

Suppose the channel data word 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

This translates 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
Link (NRZ) :	11010	10110	01011	11101	11110	11010	10101	11110
Link (NRZI):	01001	10010	00110	10100	10101	10110	01100	10101

⟵—— Direction of transmission

Note: the NRZI bit stream shown could also be inverted, the polarity depending upon the last bit of the preceeding data.

Note: The tolerance of the link transmission rate of 125 megabits per second as mentioned in 4.2 should be ± 100 ppm.