Michael Robin, a 20-year veteran of the Canadian
Broadcasting Corporation Engineering Headquarters, is
an independent broadcast consultant located in
Montréal, Canada. He is the co-author of Digital
Television Fundamentals, published by McGraw-Hill
and a contributing writer to Broadcast Engineering Magazine.
mrobin@miranda.com

Technical Notes - 1

The AES/EBU Digital Audio Signal Distribution Standard

THE BACKGROUND

Early digital audio equipment, such as DASH (Digital Audio Stationary Head) tape recorders were used as drop-ins in an otherwise analog audio studio with analog in\out ports. It offered superior, near transparent, throughput performance, well above that of similar analog audio equipment. All these machines operated as a digital "black box" with analog in/out ports. Some proprietary bit-serial distribution formats were developed to allow for digital interconnection between various types of equipment. As the state of the art progressed, the need was felt for the development of a universal digital interconnect format. This led to the development of the AES/EBU digital interconnect format.

THE SOLUTION

The Audio Engineering Society (AES) together with the European Broadcasting Union (EBU) developed a digital audio transmission standard known as the AES/EBU standard as well as AES-1992, ANSI S.40-1992 or IEC-958. The transmission medium is copperwire, which has a wide bandwidth capability and allows for the bit-serial transmission of the digital audio data. The interface is primarily designed to carry monophonic or stereophonic signals in a studio environment at a 48 kHz sampling frequency and with a resolution of 16 to 20 or, optionally, 24 bits per sample.

The bit-parallel data words are serialized by sending the least significant bits (LSB) first. Word clock data is added to the bit stream to identify the start of each sample in the decoding process. The bit-serial data stream uses the non-return-to-zero (NRZ) coding. This means that a low voltage indicates binary zero (0) and a high voltage indicates binary one (1). NRZ results in the signal voltage remaining constant and not returning to zero between each data bit. As a consequence, information about signal polarity needs to be transmitted to correctly interpret the message. Since a single NRZ serial data stream contains no information about the signal polarity another coding format is required. The format chosen is the Bi-phase Mark Code (BPM).

THE STRUCTURE OF THE AES/EBU SIGNAL

Figure 1 shows the structure of the AES/EBU signal. The signal is transmitted as a succession audio blocks:

- Each block is made up of 192 frames numbered 0 to 191.
- Each frame is made up of two subframes, subframe A and subframe B.
- Each subframe is divided into 32 time slots numbered 0 to 31 and combines data from one audio source or channel, auxiliary data, sync data and associated data.

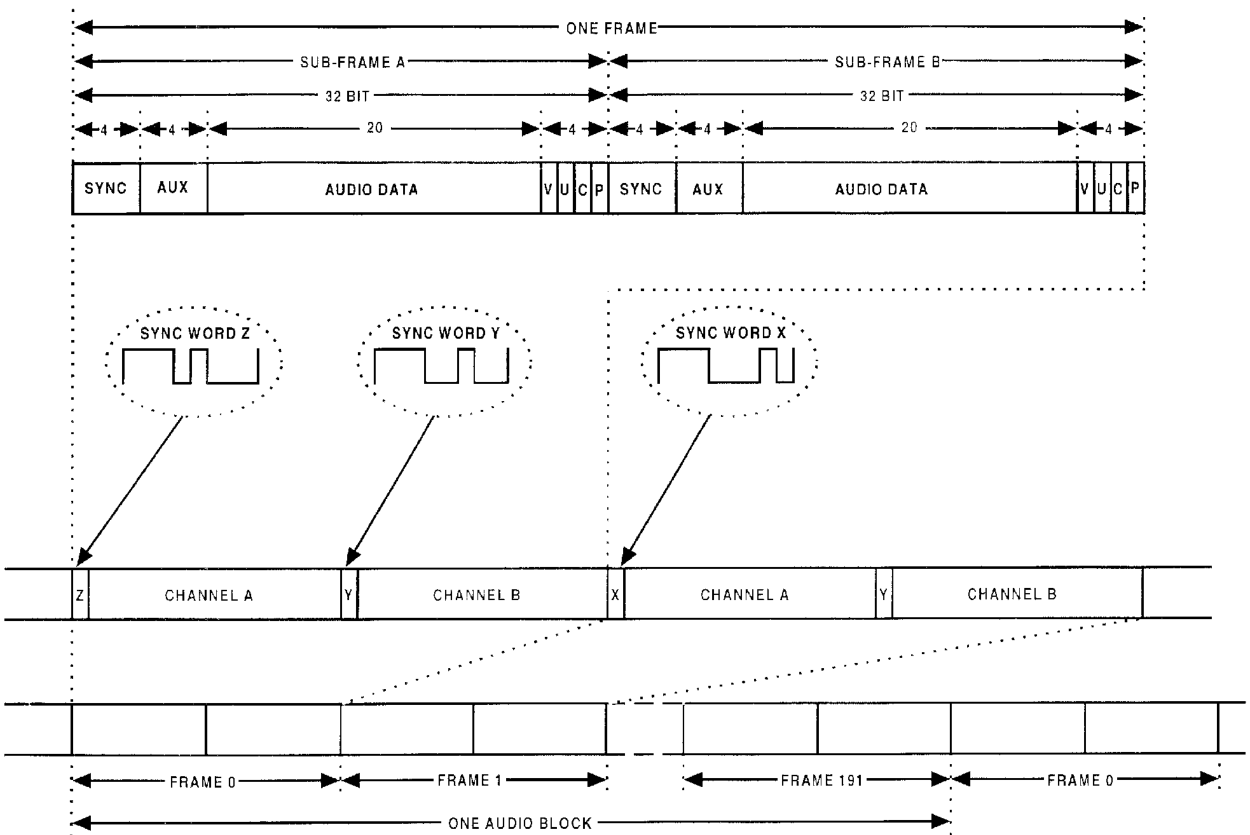

FIGURE 1 THE AES/EBU DIGITAL AUDIO DATA STRUCTURE

Data carried in time slots 0 to 3

These time slots carry one of the sync words denoted as X, Y or Z.

•Sync word Z: This bit sequence indicates the start of the first frame of an audio block.

•Sync word X: This bit sequence indicates the start of all remaining frames.

•Sync word Y: This bit sequence indicates the start of every B subframe.

The sync words are not BPM encoded. Their structure minimizes the DC component on the transmission line and facilitate clock recovery and subframe identification as they are unique in the data stream.

Data carried in time slots 4 to 7

These time slots can carry auxiliary information such as a low-quality auxiliary audio channel for producer talk-back or studio-to-studio communication. Alternately they can be used to augment the audio word length to 24 bits.

Data carried in time slots 8 to 27

These time slots carry 20 bits of audio information starting with LSB and ending with MSB. If the source provides fewer than 20 bits the unused LSB's will be set to the logical "0".

Data carried in time slots 28 to 31

These time slots carry associated bits as follows:

•Validity bit (V): The V bit is set to logical "0" if the audio sample word data are correct and suitable for D/A conversion, In the presence of defective audio samples the V bit is set to logical "1" instructing the receiving equipment to mute its output during the presence of defective samples. This capability has not been implemented by all manufacturers and some equipment may not generate or verify the sample word validity.

•User bit (U): The U bit in each subframe is sent to a memory array. The AES18-1992 recommended practice specifies the format of the user data channel of the interface.

•Channel status bit (C): The C bit carries, in a fixed format, information associated with each audio channel which is decodable by any interface user. Examples of information to be carried are length of audio sample words, pre-emphasis, sampling frequency and time codes.

•Parity bit (P): A parity bit is provided to permit the detection of an odd number of errors resulting from malfunctions of the interface. The P bit is always set to indicate an even parity.

Figure 2 shows a conceptual block diagram of an AES/EBU encoder. Figure 3 shows the BPM encoded signal waveform as obtained form an NRZ data stream.The NRZ is characterized by logical "1's" having a determined high value and logical "0's" having a determined low value. This means that long strings of 0's and 1's have no transitions and result in difficult clock recovery in the receiver. BPM alleviates this condition by introducing transitions in the middle of each "one" bit interval. At a 48 kHz sampling rate the total data rate is:

32 bits x 2 subframes x 48,000 Hz = 3.072 Mbps

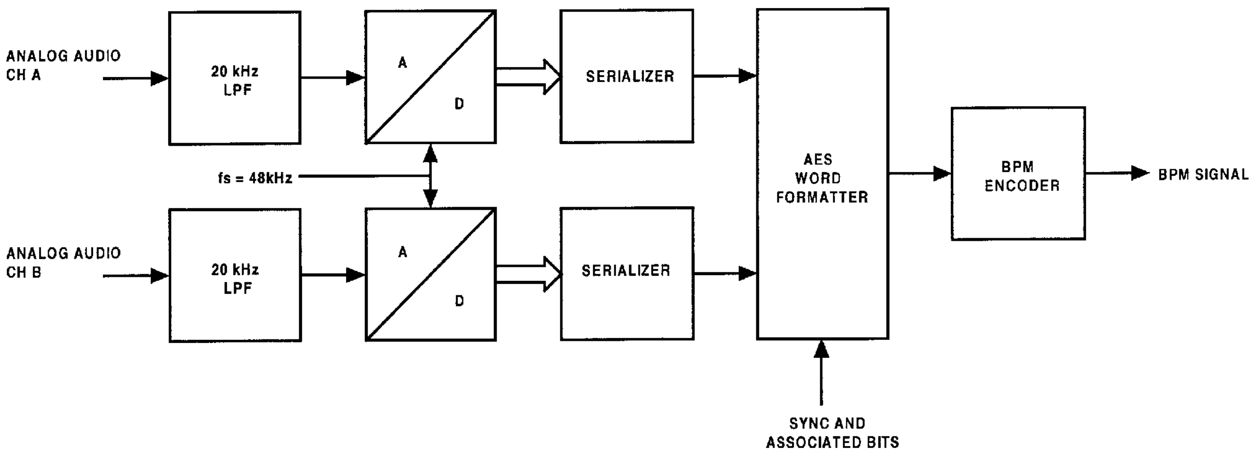

FIGURE 2 CONCEPTUAL BLOCK DIAGRAM OF AES/EBU CHANNEL ENCODER

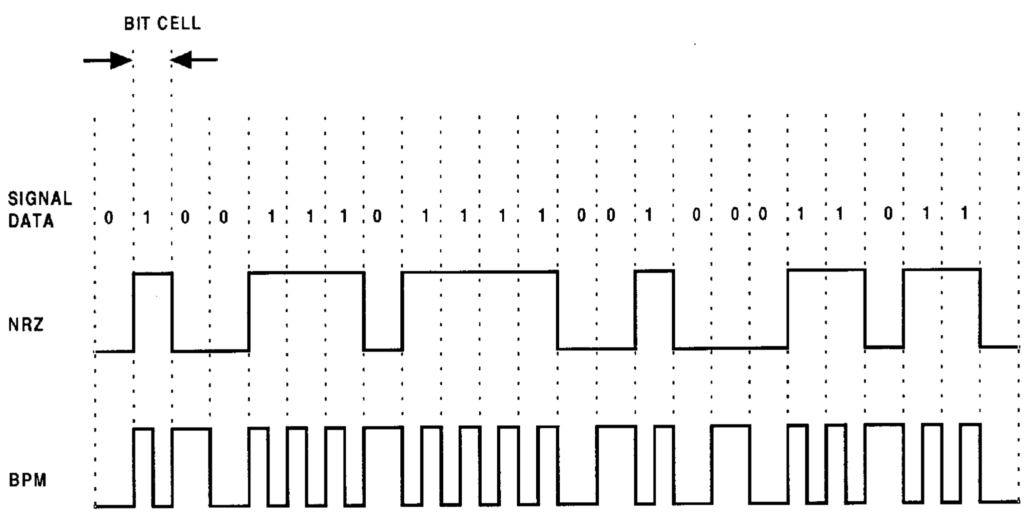

FIGURE 3 BI-PHASE MARK ENCODED SIGNAL WAVEFORM

After BPM encoding the data stream rate is doubled to 6.144 Mbps which yields a Nyquist frequency of 3.072 MHz. As shown in Figure 4 the resulting spectrum exhibits nulls at multiples of 6.144 MHz.

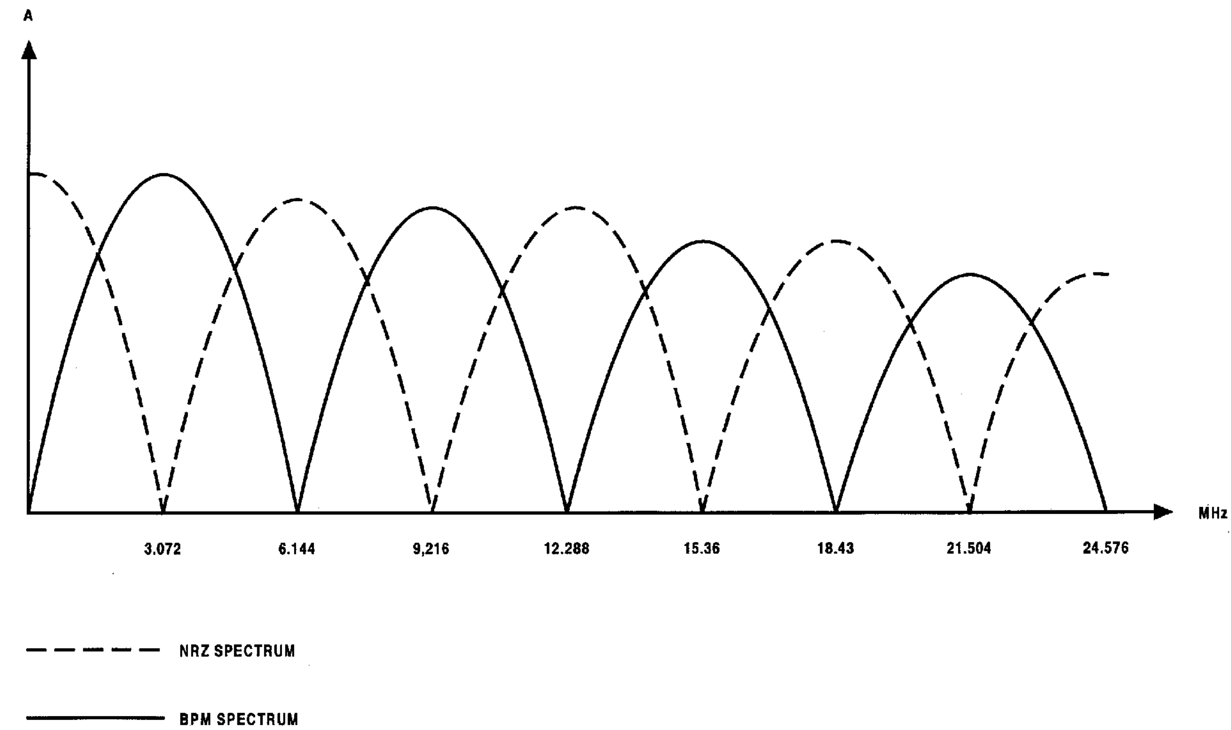

FIGURE 4 SPECTRUM OF AES/EBU DIGITAL AUDIO SIGNALS

Figure 5 shows a conceptual block diagram of an AES/EBU decoder.

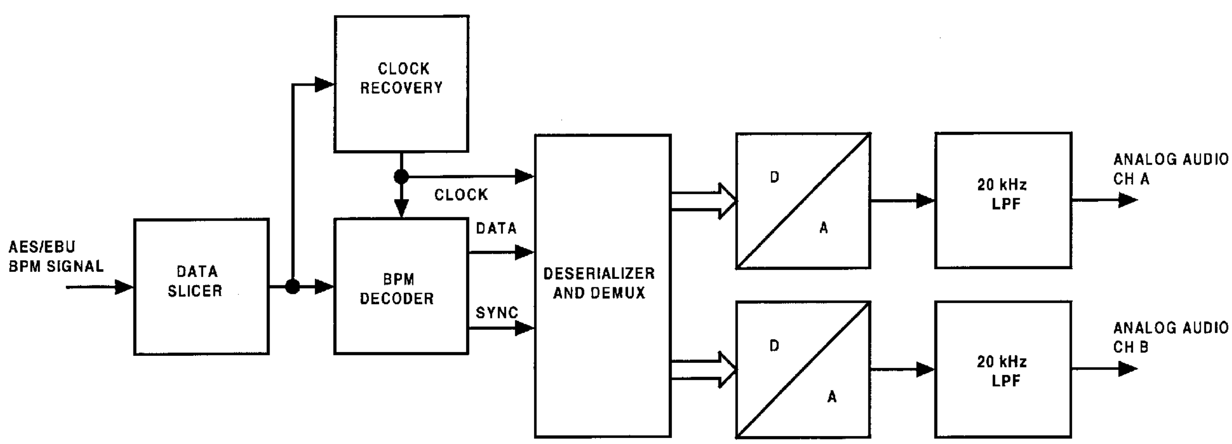

FIGURE 5 CONCEPTUAL BLOCK DIAGRAM OF AES/EBU DECODER

THE INTERFACE CHARACTERISTICS

The original AES3-1985 standard defined the distribution of AES/EBU signals through a twisted-pair shielded audio cable. It specified a transmitter source impedance of 110 Ω, a receiver input impedance of 250 Ω and stipulated that up to four receivers could be connected in parallel across the audio cable. It gave, however, no guidance on precautions needed to be taken by the user or systems integrator. This resulted in difficulties with reflections and standing waves as the performance of the distribution link was unpredictable and depended on the wide variety of installation conditions encountered in practice. The unpredictability was compounded by the loose specification of the output signal amplitude which puts an additional stress on the receiver.

The standard was revised and reissued as AES3-1992. Table 1 lists the characteristics of the 110 Ω balanced interface as defined by this standard. This second version specifies a receiver input impedance of 110 Ω and warns against the use of more than one receiver across the feeding cable. Other specifications remained unchanged. Figure 6 shows a typical interconnection as per the AES3-1992 specifications. Note that the screen is directly connected to ground at the sending end and through a 0.1 µF capacitor at the receiving end. This effectively grounds the shield at both ends at high frequencies thus reducing RF radiation and EMI problems and avoids low frequency ground loops which are the source of hum.

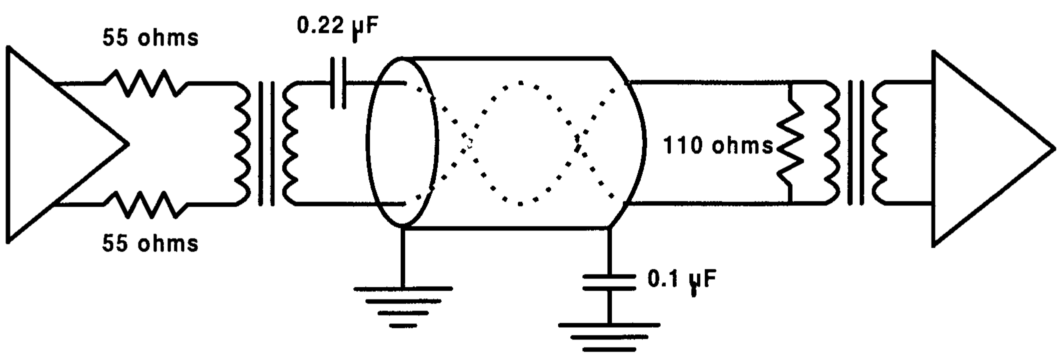

FIGURE 6 TYPICAL AES3-1992 DISTRIBUTION INTERFACE

TABLE 1 ELECTRICAL CHARACTERISTICS OF THE AES-1992 INTERFACE

Transmitter characteristics	Balanced output with XLR connector Source impedance: 110 Ω ± 20% Balance: <-30 dB (to 6 MHz) Launch signal amplitude: 2 - 7 Vp-p across 110Ω load Rise and fall time: 5 to 30 ns Jitter: <20 ns p-p
Receiver characteristics	Balanced input with XLR connector Input impedance: 110 Ω ±20% CMMR: Up to 7 V p-p to 20 MHz Maximum accepted signal level: 7V p-p Cable specification: Shielded twisted pair Cable length: 100 to 200 m maximum Cable equalization: Optional

TABLE 2 ELECTRICAL CHARACTERISTICS OF THE AES-3id-1996 INTERFACE

Transmitter characteristics	Balanced output with BNC connector Source impedance: 75 Ω nominal Return loss: >25 dB (0.1 to 6 MHz) Launch signal amplitude: 1 Vp-p across 75 Ω load D.C. offset: 0.0 V ±10% Rise and fall time:30 to 44 ns Jitter: <20 ns p-p
Receiver characteristics	Unbalanced input with BNC connector Input impedance: 75 Ω nominal Return loss: >25 dB (0.1 to 6 MHz) Minimum input level sensitivity: 100 mV Cable equalization: Optional

The AES3id-1996 standard defines the unbalanced 75 Ω impedance interface. Table 2 lists the characteristics of the interface. This version recognizes the need to narrowly specify impedance tolerances in terms of "return loss" as well as transmitter output signal levels and, if properly implemented, results in a more predictable performance as it is based on well-known SDTV analog video signal distribution concepts. However, most digital audio equipment is equipped with XLR connectors and conversion to BNC connectors including the use of 100 Ω to 75 Ω balun transformers and signal amplitude normalizers must be considered. Figure 7 shows a typical interconnection as per AES3id-1996 between a 110 Ω high level source and a 75 Ω cable and receiver. Note again that the cable shield is grounded through a 0.1 µF capacitor at the destination to avoid low frequency ground loops. There are many other possible interconnect configurations including active impedance converters.

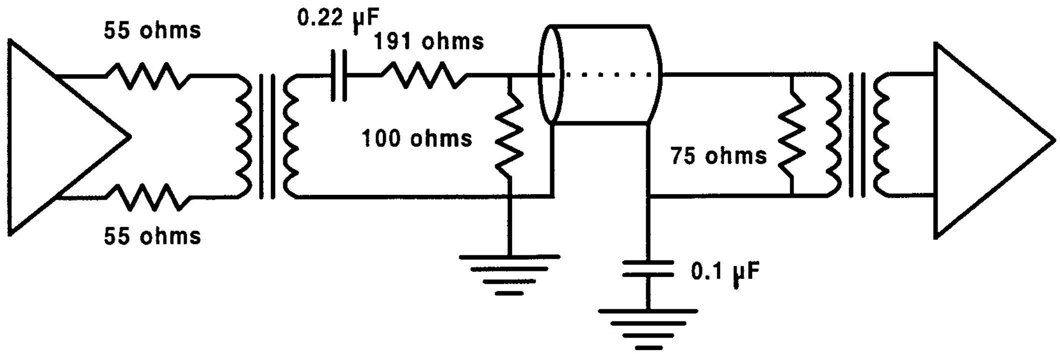

FIGURE 7 TYPICAL AES3id-1996 DISTRIBUTION INTERFACE