Digital Audio Technical Committee Report

MINUTES OF THE MEETING OF THE DIGITAL AUDIO TECHNICAL COMMITTEE

Date: *1979 November 4*

Time: *1300 hours*

Place: *Waldorf Astoria Hotel, New York, NY, USA*

Present: Chairman, Bart N. Locanthi (Pioneer North America); members, B. Blüthgen (Polygram); T. Buchler (Consultant); P. Burkowitz (Polygram); A. Conte (SMPTE); R. Ferrero (Analogic); D. Flygstad (Telex); J. Gibson (RCA Laboratories); W. Isom (RCA, retired); T. Kato (Pioneer); R. Koch (Ambichron); M. Kosaka (Matsushita); W. Korte (Hanover Technical University); R. Kragen (Recording for the Blind); R. McDonough (Harris Corporation); J. McKnight (Magnetic Reference Lab); P. Mitchell (Mystic Valley Audio); T. Muraoka (Audio Research Center of JVC); D. von Recklinghausen (Consultant); A. Sanchez (Analogic); R. Shinokawa (Toshiba); U. Swientek (Sonopress); N. Takahashi (Audio Research Center of JVC); K. Tanek (Mitsubishi); M. Willcocks (Audiomobile Division of Advent Corporation).

1. The Chairman opened the meeting, and the members introduced themselves.

2. John McKnight (Magnetic Reference Lab) discussed briefly the guidelines under which the committee would operate and that we would avoid the use of the term "Digital Audio Standards Committee" until such time as there exists an adequate number of digital audio equipment manufacturers in the marketplace (preferably more than three).

3. The Chairman discussed the proposed activities of the Digital Audio Technical Committee to get it moving in the direction started by John McKnight (Magnetic Reference Lab) in the fall of 1977:

3.1 To provide a forum for the free exchange of ideas concerning scientific and engineering problems and advanced topics of digital audio for the benefit of the AES members and other interested parties through the organization of informal workshops at the AES conventions.

3.2 To solicit and promote technical papers on digital audio in cooperation with the *Journal* editor.

3.3 To organize regular and special sessions at AES conventions and AES technical conferences for the purpose of discussing work to be done. For example:

3.3.1 To generate and publish a list of the pertinent parameters to be measured on digital audio equipment.

3.3.2 To generate and publish a list of critical listening tests for the evaluation of digital audio system performance.

3.3.3 To evaluate various coding, decoding, and transcoding schemes.

3.3.4 To generate and publish a list of unresolved problems concerning digital audio.

3.4 To collect and to publish data concerning the digital audio activities in the United States and in other countries with regard to standards. (Reports presented to the committee start on page 265.)

4. Mr. M. Kosaka (Matsushita), who is the secretariat of IEC SC60A WG 18 (Domestic Digital Audio Tape Recording), and who is also active in several digital audio committees in Japan, offered his services to keep the AES informed of activities in both areas. His help was gratefully accepted.

5. Mr. Blüthgen (Polygram) who is the secretariat of IEC SC60A WG 17 (Digital Audio Disk Recording) offered his services to keep us informed of the digital audio activity in his area, and his services were gratefully accepted.

6. The meeting was closed at 1400 hours.

7. The next meeting of the AES Digital Audio Technical Committee will be during the 1980 May AES Convention in Los Angeles. The location and time will be announced well in advance of the meeting.

BART LOCANTHI
Chairman

Consumer Use PCM Encoder-Decoder 1979-6

Electronic Industries Association of Japan Technical File of the Stereo Technical Committee and the Video Technical Committee STC-007

0 PREFACE

Among the many technical advancements in electronics in recent years, the development of digital technology and LSIs is outstanding. By applying these developments, new technology which heretofore had only been limited to theory has emerged in rapid succession.

In the field of audio engineering, digital technology has been utilized mostly in professional equipment design. However, a proliferation of consumer video cassette recorders has recently invited the digital technology into consumer audio equipment.

The PCM recording and playback system which was made possible by combining a VTR and digital technology has a strikingly broader dynamic range with practically no wow and flutter, compared to the current analog tape recorders. Its frequency and distortion characteristics are also much superior to that of the current system. With these various advantages the applications of digital technology in audio engineering are considered to be most effective.

However, unlike for current analog systems, establishing standards for the mechanical format and the electronic characteristics for a PCM system does not ensure compatibility between associated hardware or prerecorded tapes and hardware.

The PCM system requires an additional standard for signal transformation, that is, a unified signal format and signal processing.

Further, for VTRs there are differences in television systems with the 60-Hz/525-line type (NTSC) and the 50-Hz/625-line type (PAL, SECAM). There is no difference between the basic requirements for transferring audio signals to PCM signals of either TV format. However, when considering synchronizing signals, an appropriate signal processing for the VTR system to be used (which is different for NTSC or PAL and SECAM types) is desired.

This technical file deals mainly with VTRs in the NTSC system, which was adopted in Japan and the United States. However, studies will continue to establish suitable standards for the PAL and SECAM television systems.

For a PCM recording and playback system utilizing a VTR, there is a separate adapter unit which encases the encoder and the decoder, and a single unit in which the encoder–decoder is integrated inside the VTR unit.

For recorded tape compatibility it must be taken into account that there will be arbitrary combinations of different VTRs and adapters or single-unit recorders in the field.

Tapes recorded by any combination or such a single-unit recorder must reproduce proper output signals without error or other malfunctions when played with another combination or single-unit recorder.

Both experiences and the repeated experiments conducted this time, revealed that in the case of a well-maintained recorder playing a good-quality tape, very few error signals were produced, and simple problems were easily correctable.

However, it must be noted that VTRs in general are primarily designed and manufactured for recording and reproducing TV signals up to their desired performance and tolerances that may not be appropriate for PCM audio recording. Thus chances are that the operation will not be troublefree when the adapters are out on the market.

Also, deterioration of PCM signal transmission is possible with VTRs or tapes that are misoperated or off standard due to prolonged use.

For this reason the standard prescribed here provided dual error correction to ensure and promote compatibility.

In the future, as the general quality of the machines on the market is upgraded and the users become efficient in operating the machines, part of the error-correcting words may be omitted, but for the time being this is not advisable.

1 OBJECT

This technical file is prescribed to establish the signal format and other necessary conditions for a consumer-use PCM encoder–decoder which records or reproduces audio signals in the form of pulse-code modulation (PCM), in conjunction with a consumer-use cassette video system or a part thereof, and to maintain optimum performance between the various equipments and the compatibility of recorded tapes[1] and of encoder–decoders.

2 SCOPE

This technical file applies to the consumer-use PCM encoder–decoder for transforming two sets of audio sig-

[1] This refers to compatibility within the same VTR format.

nals with a frequency range below 20 kHz, in the form of PCM which conforms to the 60-Hz/525-line television system.

3.2.2 Sampling Time

Sampling shall be done on A and B channels alternately for staggering the sampling time (see Table 1).

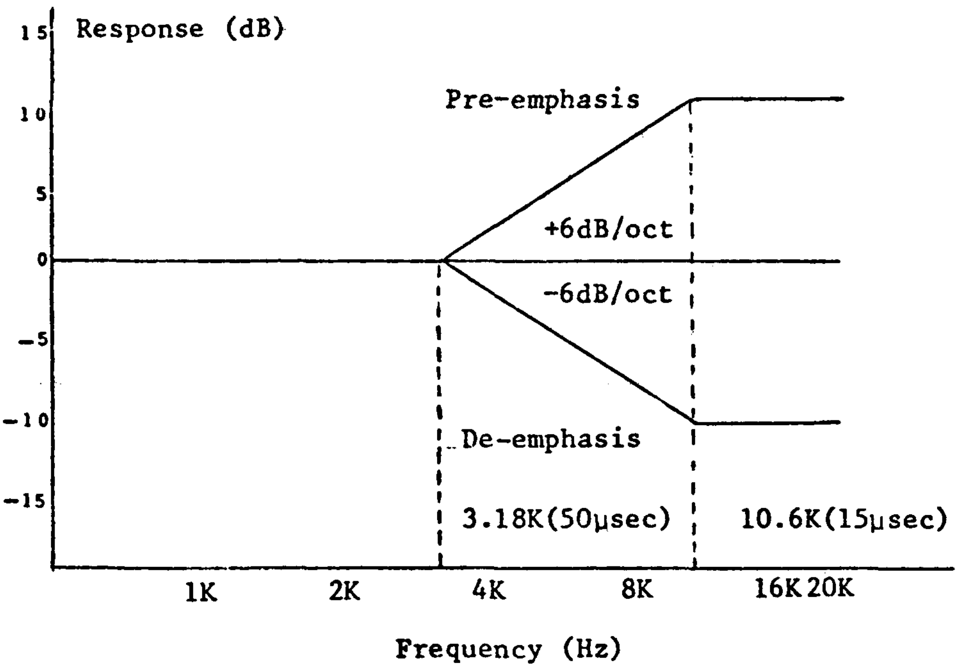

Fig. 1.

3 FORM OF SIGNALS

3.1 Audio Signals

3.1.1 Number of Audio Channels

The number of transmitted audio signals shall be two channels A and B. (Channels A and B correspond to L and R, respectively, for stereophonic use.)

3.1.2 Emphasis

Emphasis may be applied to audio signals. The time constants T_1 and T_2 for the emphasis shall be as follows:

$$T_1 = 50\ \mu s$$

$$T_2 = 15\ \mu s.$$

3.2 Signal Sampling and Quantization

3.2.1 Sampling Frequency

The sampling frequency shall be 44.056 ± 0.005 kHz.

3.2.3 Quantization

Recording shall be made in the form of a 14-bit linear slot.

3.2.4 Coding

A 2's complement binary code shall be used. A positive binary value represents the positive audio signal voltage.

3.2.5 Transmission Rate

2.643 Mbit/s.

3.3 PCM Signal Format

3.3.1 Signal Format

The PCM signal format shall be in accordance with the television standard, one field (262.5 H) contains a data block of 245 H and a control signal block of 1 H.

Table 1.

Sampling time	$2n/2\ T$	$2n+1/2\ T$	$2n+2/2\ T$	$2n+3/2\ T$ ...
A ch	An		An+1	...
B ch		Bn		Bn+1
T = 1/(44056 ± 5)(sec)				

3.3.2 Data Block

One data block shall consist of six sampled signal words, P and Q error-correcting words, and one error-detecting word (CRCC). The data block of nine words has a span of 1 H.

3.3.3 Control Signal Block

The control signal block shall consist of five words which include a cueing signal word, a content identification signal word, an address signal word, a control signal word, and an error-detecting word (CRCC). The control signal block has a span of 1 H and is preceding the data block in each field.

3.3.4 Synchronizing Signal

Horizontal and vertical synchronizing signal formation shall be in accordance with the 60-Hz/525-line television system.

3.4 Data Block Structure

3.4.1 Sampled Signal Word

The sampled signal word shall consist of 14 bits, which are laid out in such a way that the most significant bit (MSB) is bit 1 and the least significant bit (LSB) is bit 14, as shown in Fig. 2.

3.4.3 Error-Detecting Word (CRCC)

The error-detecting word, expressed as a 16-bit word, shall be the remainder, which results when the sum of six sampling words, one of each of the P and Q words, and the 112-bit signal in total is divided by the following generating function:

$$G(x) = x^{16} + x^{12} + x^5 + 1 .$$

The generating operation is a forward system.

3.5 Data Block Assignment

3.5.1 Interleave

For the sampled signal words and the P and Q error-correcting words within each block, D = 16 block (16 H) interleave shall be applied. The interleave of D = 16 blocks is equivalent to the word interleave of $3D$ = 48 words.

3.5.2 Word Assignment

Each word after the interleave shall be arranged in the order of six sampled words, the P and Q error-correcting words, and the error-detecting word generated by application of the formula in Section 3.4.3.

The resulting data block is illustrated in Fig. 3.

The CRC shown in Fig. 3 may be expressed as follows:

MSB													LSB
1	2	3	4	5	6	7	8	9	10	11	12	13	14
14 bit													

Fig. 2.

3.4.2 Error-Correcting Words

Each of the error-correcting words P and Q shall consist of 14 bits and be generated as follows:

$$Pn = An \oplus Bn \oplus An+1 \oplus Bn+1 \oplus An+2 \oplus Bn+2$$

$$Qn = T^6An \oplus T^5Bn \oplus T^4An+1 \oplus T^3Bn+1 \oplus T^2An+2 \oplus T\,Bn+2$$

where n is an address expressed in 0 or multiples of 3; $\oplus$ is modulo 2 summation of the bits in each column; and T is the Q generating matrix as follows:

$$T = \begin{bmatrix} 0&0&0&0&0&0&0&0&0&0&0&0&0&1 \\ 1&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&1&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&1&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&1&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&1&0&0&0&0&0&1 \\ 0&0&0&0&0&0&0&0&1&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&1&0 \end{bmatrix} .$$

$$\mathrm{CRC} = (1 + b_1)x^{127} + (1 + b_2)x^{126} + \cdots + (1 + b_{16})x^{112} + b_{17}\cdot x^{111} + b_{18}\cdot x^{110} + \cdots + b_{112}\cdot x^{16} \bmod G(x).$$

Here b_1 is the MSB of An and so on in descending order, with b_{112} showing the LSB of $Qn^{-21}\,D$.

3.6 Control Signal Block Structure

3.6.1 Cueing Signal Word

The cueing word shall be a 56-bit configuration, having a content of

11001100..........1100.

3.6.2 Content-Identification Signal Word

The content-identification signal word shall be a 14-bit word. The configuration of the word is under consideration and is assumed tentatively as all 0.

3.6.3 Address Signal Word

The address signal word shall consist of 28 bits. The configuration is under consideration and is assumed tentatively as all 0.

3.6.4 Control Signal Word

The control signal word shall consist of 14 bits having the content as shown in Table 2.

Table 2.

Bit Number	Code Content	Application	Respective Bit
1–10	Not specified	—	0
11	Copy-prohibiting code	Not applied	0
12	*P* correction identification code	Applied	0
13	*Q* correction identification code	Applied	0
14	Preemphasis identification code	Applied	0

3.6.5 Error-Detecting Word (CRCC)

The error-detecting word shall be a 16-bit word which is derived, in accordance with the formula given in Section 3.4.3 from the 4-word 112 bit which includes the cueing signal word, the content identification signal word, the address signal word, and the control signal word.

3.7 Control Signal Block Assignment

3.7.1 Word Assignment

Each word shall be laid out in the order of the cueing signal word, the content-identification signal word, the address signal word, and the error-detecting word, and arranged into one data block as shown in Fig. 4.

3.8 PCM Signal Waveform

3.8.1 PCM Signal Waveform

The waveform shall be a nonreturn-to-zero (NRZ) signal waveform.

3.8.2 Assignment and Structure within 1 H

As shown in Fig. 5, the data synchronizing signal and the white reference signal shall be placed at the top and at the end, respectively, within one horizontal period.

The data synchronizing signal code shall be 1010.

One horizontal synchronizing segment is composed of 168 bits, which correspond to a time of 63.556 ± 0.007 μs.

3.8.3 Signal Waveform

Signal waveforms and levels are as shown in Fig. 6. The tolerance of each level is ± 10%.

3.8.4 Assignment and Structure within One Field

Each field shall be headed by a vertical sync signal with equalizing pulses. The control signal block of one horizontal span shall occupy the 10th horizontal period for the odd field, and 10.5th period for the even field, respectively.

The control signal block is followed by a series of data blocks having 245 H. The rest, a series of horizontal periods in the field, is left blank (see Figs. 7 and 8.)

4 INPUT/OUTPUT LEVELS

4.1 Audio Signal Input–Output Levels

4.1.1 Reference Frequency[2]

419.58 Hz.

4.1.2 Reference Recording Level[3]

At sine wave, the positive peak value is 00010110110001 (05B1) and the negative peak value is 11101001001111 (3A4F).

4.1.3 Rated Input Level

Below 142 mV.

[2] 1/105 of sampling frequency and seven times the vertical synchronizing signal frequency.
[3] 15 dB below the maximum recording level.

A_n	B_{n-3D}	A_{n+1-6D}	B_{n+1-9D}	$A_{n+2-12D}$	$B_{n+2-15D}$	P_{n-18D}	Q_{n-21D}	CRCC

14x8+16 = 128 bit

Fig. 3.

Cueing Signal - Content-Identification - Address - Control - CRCC

56	14	28	14	16

128

Fig. 4.

4.1.4 Rated Output Level

142 mV.

4.2 PCM Signal Input– Output Levels

4.2.1 Encoder Output Level

1 volt peak to peak.

4.2.2 Decoder Input Level

1 volt peak to peak.

ADDENDUM

For the time being, both the P and the Q error-correcting words shall be used.

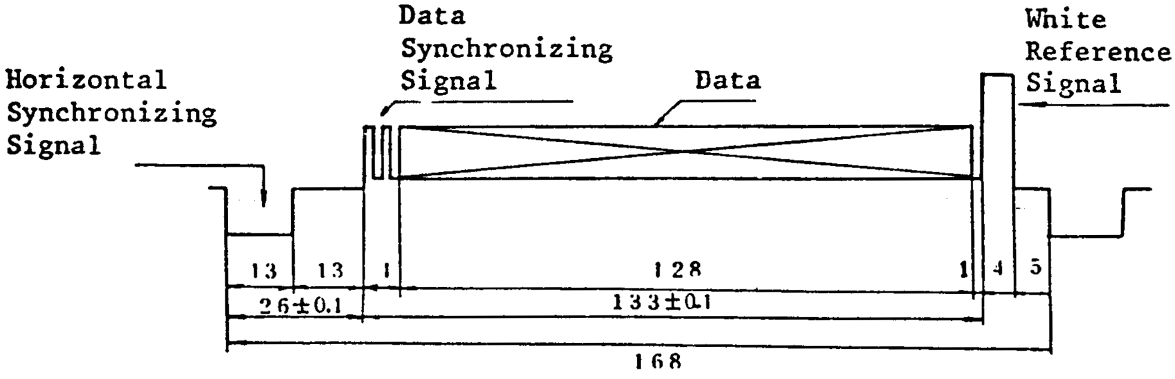

Fig. 5. Units are in bits.

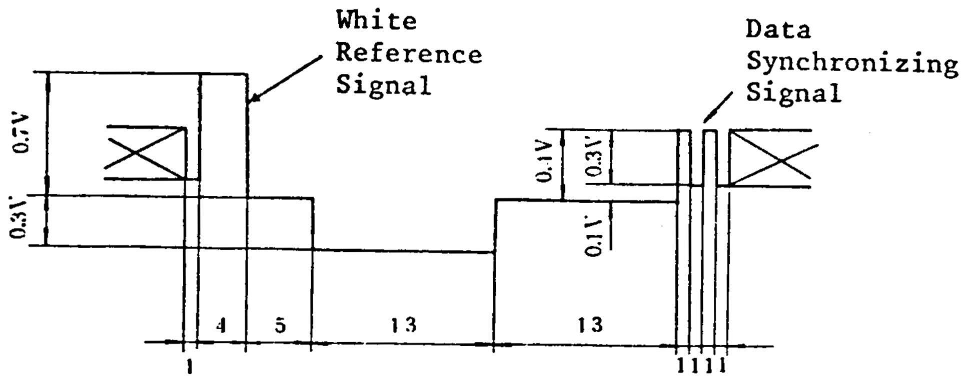

Fig. 6. Units are in bits.

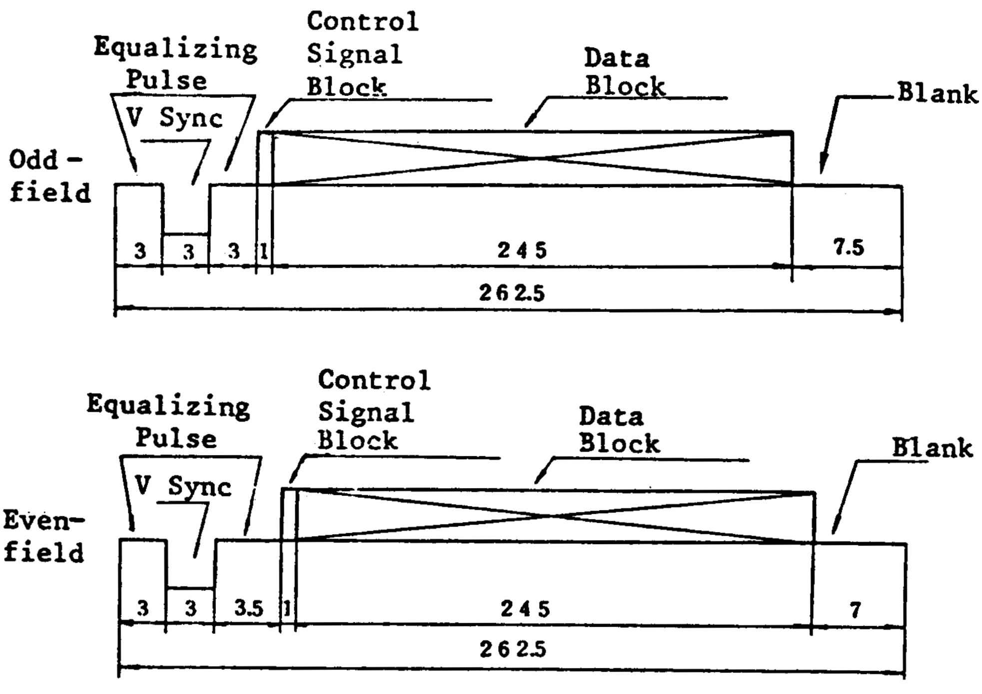

Fig. 7. Units are in H.

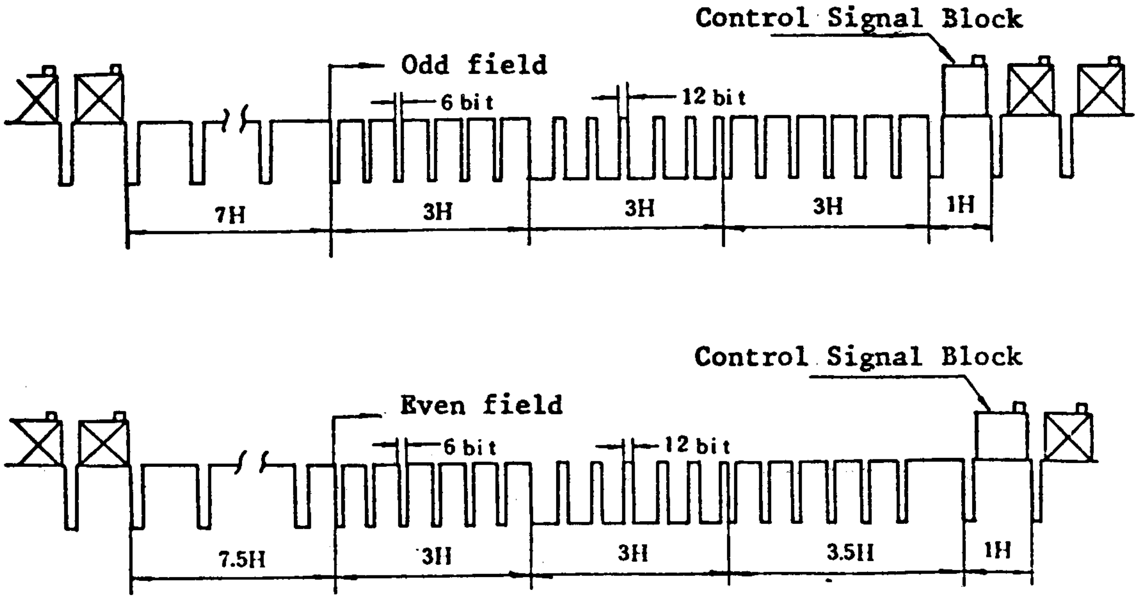

Fig. 8.

International Electrotechnical Commission Technical Committee 60: Recording Subcommittee 60A: Sound Recording

REPORT OF SUBCOMMITTEE 60A, WORKING GROUP 15, DIGITAL AUDIO RECORDING

Date: *1979 October*

Place: *Holiday Inn, Palo Alto, CA, USA*

Present: B. Blüthgen (West Germany); D. Browning (United Kingdom); A. Heaslett (USA); M. Kano (Japan); J. Kerstens (Netherlands); M. Kosaka (Japan); F. Larsen (Norway); L. Märtin (West Germany); G. Nathan (United Kingdom); J. Potgiesser (Netherlands); R. Siefken (USA); M. Stübbe (West Germany); A. Weisser (France); R. Youngquist (Secretary) (USA).

A series of meetings were held 1979 October 16 and 17, covering digital audio recording. The first was on professional digital audio tape recording, WG16, with Mr. Browning as the secretariat. Following this was a digital audio disk recording, WG17, meeting with Mr. Blüthgen as the secretariat, and then a domestic digital audio tape recording, WG18, meeting was held with Mr. Kosaka as the secretariat. Each of the meetings discussed the tasks of the groups and a short report on each follows.

WG16, Professional Digital Audio Tape Recording

An ad-hoc meeting of WG16 was held on the above date. The discussion was general and far reaching in an effort to establish a program for future work.

It was generally agreed that a vocabulary of the terminology is required as an aid to further discussion. The secretary agreed to collate and circulate a basic vocabulary and asked for input from members to aid this task.

The basic physical characteristics of the tape and transport interface were discussed. Present knowledge of existing and proposed systems shows that a range of physical formats exists. A questionnaire will be circulated to working group members to provide basic data and a starting point for action in this area.

Examination of existing source encoding characteristics showed that there is some common practice and that two recorders already in use and two in the development stage have the following characteristics:

1) Sampling frequency: 50 kHz,
2) Quantization: 16-bit linear slot,
3) Two's complement binary code;
4) No source encoding emphasis.

The sampling frequency choice was established primarily by a 20-kHz bandwidth requirement, with secondary considerations being a variable-speed facility and the possibility of practical audio-to-video format transform.

There is also a wide range of other values under discussion for these and other parameters.

From the discussions the scope of the work for the group can be proposed as follows:

1) The interface characteristics of the recording medium.

2) The digital input–output signal format of recorders.

3) Signal processing, which can be divided into source encoding, formatting, and channel encoding, and the equivalent inverse functions.
4) Interface characteristics for editing.
5) Methods of measurement.

WG17, Digital Audio Disk Recording

The plenary meeting of WG15 discussed four submitted documents:
1) Document SC60A, WG15 (Mr. Kosaka) 1, "Report of Digital Audio Standardization Activities in Japan" (1979 October).
2) Document SC60A, WG15 (Mr. Kosaka) 2, "The Status of the DAD (Digital Audio Disk) Study Group Meeting" (1979 January 6).
3) Document SC60A, WG15 (Mr. Kosaka) 3, "The Target Specification of the DAD System" (1979 October).
4) Document SC60A, WG15 (Mr. Kosaka) 4, "Report on Digital Audio Disk Study Meeting" (1979 June).

It was stated that besides the group of Japanese manufacturers participating at the local (Japan) DAD meetings and the working group activities, four foreign manufacturers are also acting as members.

This DAD group hopes that the now established WG17 becomes the body for further evaluation of the DAD standardization activity.

Another briefly discussed item was the DAD application for broadcasters' use.

WG17 Scope (Draft): The digital audio disk standard is applied to all disks and their recording as well as reproducing equipment interchange characteristics in general domestic and professional use.

Classification of Standards Applied:
1) General standard,
2) Professional including broadcasting applications,
3) Domestic application,
4) Signal carrier properties:
   a) Mechanical, b) Physical,
5) Signal channel format encoding,
6) Recording equipment interface characteristics,
7) Reproducing equipment interface characteristics,
8) Testing and measuring methods.

Present Task: Discussion of "general standard" related items, like sampling rate(s), number of quantization bits, number of transmission channels, play time, and disk size, by submitting proposals to WG15 derived from respective questionnaires.

In order to speed up its activity, the working group meetings are intended to be held also before the forthcoming official IEC meeting.

WG18, Domestic Digital Audio Tape Recording

In addition to the attendees listed, Mr. Herman Lia of Norway was present at this meeting.

1) Discussions were focused mainly on the following fields:
1.1) Area of discussion,
1.2) Reports and/or comments generated from delegates of each national committee,
1.3) Discussions and exchange of opinions among members on the "Domestic Digital Audio Tape Recording."

2) Summary of discussions on above-mentioned fields.
2.1) Area of discussion. Domestic digital audio tape recording would be mainly put into practice by utilizing consumer video tape recorders, and these video tape recorders are adapted to rotary magnetic heads. Consequently, the area of discussion within WG18 will mainly be limited to PCM encoder/decoders, either as processing adaptors, or units built into PCM decks.

Domestic digital audio recording using a longitudinal type video tape recorder will be in the area of WG18, but general understanding will be the future work.

The stationary-head type digital audio recorder for domestic use and the rotary-head type digital audio recorder which will come in the future may probably be treated.

2.2) Reports and/or comments of attendees of WG18. Two reports were submitted to WG18 from the Japan National Committee: SC60A WG15 (M. Kosaka)1, "Report of Digital Audio Standardization Activities in Japan," and SC60A WG15 (M. Kosaka)5, "Electronic Industries Association of Japan—Consumer Use PCM Encoder–Decoder."

Mr. M. Kosaka reported on the standardization activities taking place in Japan based on and using the above two reports.

There were questions and comments on these reports. They were mainly related to clarifying the understanding of these reports.

Mr. M. Kosaka also reported that the status of the Japanese standard covered by the above reports is tentative, but that it may become a formal standard in the future.

2.3) Discussion and exchange of opinions. The main discussions were considered to have been as follows:

a) A stationary-head recorder should be discussed in the future for domestic digital audio recording and to set up a standard format which may be independent of the TV signal format.

b) An adaptive[4] pre-/deemphasis can be incorporated in a domestic digital audio recording.

Following these three meetings, WG15, Digital Audio Recording, met. This group is composed of R. Youngquist, secretariat, and D. Browning, B. Blüthgen, and M. Kosaka. They discussed its activities and defined the following tasks:

1) Coordination of all internal activities of WG16, 17, and 18. Specifically, three areas of concern arose:
a) Terminology. This will be common for all groups and overseen by WG15.
b) Measurement methods. Coordination will occur so that, when required, common methods will be used in the groups.
c) Sampling frequencies. It is evident that there are already several sampling frequencies in existence. It will be

[4] Varying emphasis characteristics depending on a distribution of the recording signal spectrum.

necessary for WG15 to coordinate and make sure that if sampling frequency relationships are required between WG16, 17, and 18, these will occur, and that the individual groups will not go in separate directions.

2) WG15 will provide liaison as required with external groups.

3) Any documents to be submitted to SC60A will be submitted by WG15, even though initially generated in WG16, 17, or 18.

4) Copies of all correspondence of the individual secretariats will be supplied to the secretariat of WG15, who will provide any required information to the other secretariats.

Questionnaire Prior to Digital Audio Standardization

BJÖRN BLÜTHGEN

Polygram, Hannover, West Germany

General

The economic approach to digital audio system standardization will be found only by being aware of the users' limits of analog audio perceptibility, the state of the art technique, the circuitry/equipment costs, and the environmental influence, such as ambient temperature as well as electric energy interferences.

Therefore it becomes obvious that worst-case tolerance limits of already known system characteristics must be specified carefully, especially as one element in the system chain may strongly influence corresponding other ones.

The following text aims at contributing to the above by raising questions which have not yet been clearly answered concerning software-related aspects regarding elements involved, such as filter, analog-to-digital converter, error correction, and data reduction methods.

Filter (F)

The input low-pass filter must be designed for the attenuation of audio signal spectral components and for spectral components created by different sampling rate energy above the human audio bandwidth. The output low-pass filter attenuates aliasing components. An aperture correction filter is required for the sample and hold circuit passband equalization. Pre-/deemphasis filters may sometimes be advantageous for some applications.

Questions

F1) Who investigates and reports on the mandatory reasons for the upper bandwidth limit for the analog program production chain (20 kHz), compared to the fundamental requirements of forthcoming digital production techniques? It is claimed that digital techniques eliminate quality degradation per processing stage with respect to phase distortion, compared with today's analog circuitry chain elements. Such an investigation should include common mixing, equalization and editing procedures.

F2) Is it feasible for either economic, technical, or marketing reasons to consider a prerecorded program bandwidth not exceeding 15 kHz when the international communication network agrees upon 32 kHz as its audio signal sampling frequency? The industry might be blamed later on for wasting playing time, carrier material, and so on, when it will be found by the customer that there is no audible difference compared to the same program material sampled at higher rates.

The expense for necessary filter techniques should thus, as far as possible, remain on the professional side to obtain simple playback equipment circuitry.

F3) Which guard bandwidth and stopband attenuation is *practically* (environmental and audible perception related) required for the input–output filter design for 15 kHz and 20 kHz as Nyquist frequencies?

F3a) Do we expect other applications with diverging demands?

F3b) Can we accept suppressing particular environmental or application-depending interference frequencies by additional components in order to simplify standard equipment?

F3c) What are the tolerances for the input–output low-pass filter ripple (separate/total)?

F3d) Which phase distortion (separate/total) is acceptable for 15/20-kHz filter types?

F3e) What can/should be filtered in the digital domain and where?

F3f) Which digital filtering reduces cost or circuit complexity?

F3g) Do we expect new PCM decoding methods requiring less expense?

Sampling Frequency (SF)

The choice of sampling frequencies must be seen with respect to their application fields:

a) Broadcasting/communication networks with 32,

b) Video: PAL, SECAM, and NTSC compatibility (44.056/50.4 kHz),

c) Film: 24-Hz frame resolution,

d) Consumer: fidelity, material cost, and playing time.

It is not believed that only one standard frequency will be agreed upon. But more than two sampling frequencies will create complex problems for transcoding equipment and filter circuitry of interconnected recorder and mixer (aliasing frequency suppression and operational problems, such as black-box handling) for the studio operator with diverging sampling rates feeding a mixer.

Real-time sampling rate conversion *fx/fy* in the digital domain, such as interpolation by an integer ratio M and decimation by an integer ratio N, does not result in a significant hardware cost difference if the number of filter coefficients necessary is not too large. Specially designed VLSIS are already available and reported to become more adapted to filter techniques.

This gives some freedom in the choice of appropriate frequencies. Up to now three sampling frequencies are of major concern:

a) 32 kHz for international communication network,

b) 44.056/44.1 kHz for VTR equipment.

NOTE: 44.1 kHz could be accepted as the nominal standard, while NTSC applications require 44.056 kHz (difference 0.1%), which can easily be accomplished by external equipment synchronization without any audible effects.

c) The European broadcasting corporations as well as the German phonographic industry do not believe in mandatory relations between the choice of a professional sampling frequency and VTR equipment. Longitudinal magnetic recorder types, successfully demonstrated by the thin-film head (Matsushita) or in LVR methods (BASF), may succeed the VTR period. Appropriate professionally recorded audio program transcoding from, for example, 48 kHz into any professional video/film rate-related frequency is not believed to be a technical drawback for those applications where it is really necessary.

Thus 48 kHz might be seen as the wrong frequency, only if unnecessary data samples produced by this frequency waste tape, playing time, and so on.

When economic filter design cannot be taken as the major reason for a 20/48-kHz digital audio system, the benefits of a 44-kHz choice become also evident with respect to the transcoder frequency ratio of 32/44 = 8/11.

Questions

SF1) Since serious reports on audible bandwidth investigations have found 15 kHz to be sufficient for audiophile experts, and since practical realizations on 44.056-kHz sampled material proved to be excellent, what are the drawbacks for an agreement upon 44.1/44.056 kHz as the professional standard frequency?

SF2) Do we expect any audible limitation, and on which application, when processing ~ 44-kHz recorded program material at speed/sampling rate variations (20 → 15 kHz) down to −27%?

SF3) Is everybody aware that effects (such as pitch adjustments) easily established in the analog domain, when performed digitally require sophisticated frequency conversion?

SF4) What are the detailed specific applications for transcoders? Or,

SF5) Where are the economic and/or practical transcoder insert locations, and how many will be required?

SF6) What are the hardware aspects involved for required filter type realization?

Quantization (Q)

Basic system specifications with respect to noise/distortion are determined by the choice of the element numbers per sample word. Additional noise caused by physical conversion limitations and tolerances involved during the analog-to-digital and digital-to-analog processing also contributes to the noise figure measured at the system end. Therefore a total state-of-the-art Codec circuitry must always be considered for all perception tests evaluating standard figures.

Questions

Q1) Which measuring method results in equal perceptible signal-to-noise ratio figures measured on digital equipment compared to common analog-related methods?

Q2) What are the practical measured (for method see Q1) differences in noise/distortion figures of existing common 14- and 16-bit Codec realizations in a specified ambient temperature range?

Q3) Are 14-bit element words sufficient also for reverberant program decay monitored at ~ 90 dB sound pressure level and equal system noise floor?

Q4) Which benefit is really achieved by a common 16-bit PCM system featuring relatively high practical dynamic inaccuracy, compared to 14-bit PCM systems with narrow tolerances?

Q5) What are the effects of analog-to-digital converter circuit peak overload?

Q6) Which present type of peak overload is perceptible, as well as that in the digital domain for professional and/or commercial applications?

Q7) How are adequate (see Q5 and Q6) peak-level indicator characteristics to be defined?

Q8) Which static and/or dynamic reference level definition can be proposed for a standard?

Q9) What are the (dis)advantages of pre-/deemphasis application, especially on 14-bit PCM systems?

Error Concealment (EC)

A 100% correction method available for first-order error probability will be mandatory for most applications. Excessive cumulative errors of less distribution probability must be concealed by other methods in order to reduce redundant data transmission if sample word interleaving by specific channel encoding methods is not sufficient.

Questions

EC1) How many random erroneous sample words can be replaced by last/first correct word value interpolation with respect to special selected program material and perception?

EC2) How many random erroneous sample words can be replaced by equal last correct sample word values with respect to special selected program material and perception?

EC3) Which fadeout characteristic/method provides minimum audible effects on excessive drop-out periods?

EC4) Is it feasible to cover excessive drop-out periods (see EC3) by smooth signal replacement (crossfade) of an available corresponding channel program?

EC5) How to specify carrier-related tolerance limits for carrier- and application-related probability categories:

a) First-order errors,
b) Error concealment (see EC1 and EC2),
c) Errors exceeding the concealment ability.

Data Reduction (DR)

Signal encoding with inaudible/perceptible data reduction methods applied ensures economic carrier modulation methods with respect to carrier capacity (playing time) or data packing density. Packing density has a major concern with drop-out rates and the tolerances involved in the mass production of prerecorded sound carriers.

Data reduction methods must cope with hardware-related requirements such as constant time slot size per sample word and its equal data block location for economic sample and error identification during data retrieval.

Several reduction methods have already been investigated extensively with PCM companding methods being the most appropriate ones. The BBC research department reports (BBC-RD-1978/26) that its NICAM-3 system proved best in perception tests. NICAM-3 based on program-depending blockwise coded PCM floating point encoding, is proposed for EBU application. Similar codes developed by others are discussed at the same institution.

In order to provide a minimum treatment to the original data content obtaining optimum sound quality, before standardization, a new promising data reduction proposal should be considered:

Binary PCM sample difference words of subsequent sample values can be seen as coded slew-rate values which, by the theory of music, only seldom exceed a certain value. Maximum slew-rate threshold values thus can be selected in 6-dB steps by PCM difference word sizes as references for selective sample encoding with respect to human perception.

Questions

DR1) How is the selected program material relevant to word size "overflow" distribution of sample values exceeding 8-, 9-, 10-, 11-, 12-, 13-, or 14-bit elements per difference word?

DR2) What is the probability distribution according to DR1 with respect to time and cumulative occurrence?

DR3) Which audible effects occur at selected program material when slew-rate peaks above chosen thresholds are replaced by: a) A companded PCM word of equal size (PCM difference word plus polarity bit)? b) The reference value?

NOTE: Absolutely no data deterioration is achieved when slew-rate peaks above system limits are encoded into two corresponding code words (no limitation), presuming a buffer memory is placed before the digital-to-analog converter circuit.

As references for the above PCM code word difference-channel encoding method, the first sample in each data block may represent a full-size PCM sample or, depending on the application, the difference to "zero" value.