INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

J.57
(ex CMTT.724)
(06/90)

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

TELEVISION AND SOUND TRANSMISSION

TRANSMISSION OF DIGITAL STUDIO QUALITY SOUND SIGNALS OVER H1 CHANNELS

ITU-T Recommendation J.57

(Formerly Recommendation ITU-R CMTT.724)

FOREWORD

The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of the International Telecommunication Union. The ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Conference (WTSC), which meets every four years, established the topics for study by the ITU-T Study Groups which, in their turn, produce Recommendations on these topics.

ITU-T Recommendation J.57 (formerly Recommendation ITU-R CMTT.724) was elaborated by the former ITU-R Study Group CMTT. See Note 1 below.

NOTES

1 As a consequence of a reform process within the International Telecommunication Union (ITU), the CCITT ceased to exist as of 28 February 1993. In its place, the ITU Telecommunication Standardization Sector (ITU-T) was created as of 1 March 1993. Similarly, in this reform process, the CCIR and the IFRB have been replaced by the Radiocommunication Sector (ITU-R).

Conforming to a joint decision by the World Telecommunication Standardization Conference (Helsinki, March 1993) and the Radiocommunication Assembly (Geneva, November 1993), the ITU-R Study Group CMTT was transferred to ITU-T as Study Group 9, except for the satellite news gathering (SNG) study area which was transferred to ITU-R Study Group 4.

2 In this Recommendation, the expression “Administration” is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Recommendation J.57[1)]

TRANSMISSION OF DIGITAL STUDIO QUALITY SOUND SIGNALS OVER H1 CHANNELS [*)]

(1990)

The CCIR,

CONSIDERING

(a) that the source encoding for digital sound signals in broadcasting studios is given in Recommendation 646;

(b) that the 2-channel digital audio interface is specified in Recommendation 647;

(c) that a format for digital studio-quality sound signal connections should be based on these Recommendations;

(d) that the sound quality and the auxiliary information carried over the Recommendation 647 interface should be maintained as far as possible;

(e) that the hierarchical bit rates for digital networks are given in CCITT Recommendation G.702;

(f) that a digital hierarchy for interworking between networks using different transmission hierarchies is given in CCITT Recommandation G.802;

(g) that the access levels for the H1 channels in the ISDN are given in CCITT Recommendation I.412;

(h) that simple interworking between the different hierarchies is necessary;

(j) that account must be taken of network impairments such as signal bit errors, error bursts and controlled slips;

(k) that, for some applications, the introduction of excessive delay can cause operational problems,

RECOMMENDS

that for transmission of digital studio-quality sound signals the coding and multiplexing format given in Annex I should be used.

ANNEX I

CODING AND MULTIPLEXING FORMAT

1. Introduction

The transmission format is based on the 2-channel digital audio interface described in Recommendation 647, which should be read in conjunction with this text.

The 32 bits in each sub-frame of the Recommendation 647 interface are treated in the following way in the contribution format:

– Preamble (bit 0-3): not transmitted
– Bit 4-7: not transmitted
– Audio sample word (bit 8-27): companded
– Validity flag (bit 28): not transmitted
– User data (bit 29): transmitted transparently in H12 only
– Channel status (bit 30): subjected to data compression
– Parity bit (bit 31): not transmitted. Replaced by a parity bit for the audio sample word.

1) Formerly Recommendation ITU-R CMTT.724.

*) The United States of America withholds its approval of this Recommendation.

2. Compatibility between H11 and H12-based systems

The H12 level provides a total of 20 bits per sample, and H11 provides a total of 16 bits per sample. To simplify interworking between H11 and H12 channels, the companding of the audio signal is such that the samples are compressed for transmission in the H11 channel. In the H12 channel, extra bits may be conveyed, to improve resolution of the audio coding and provide a user data channel.

The essential data occupies the entire available capacity of the H11 channel, and the first 24 available octets of each frame of the H12 channel.

3. Coding law

Near instantaneous companding from 20 to 15 bits/sample is applied. A 1 ms companding block is used with 8 coding ranges.

The coding table is shown in Fig. 1a and the transmitted bits in Fig. 1b. Unused bits are set to 1 (one).

The 1 ms companding block introduces a fundamental delay of 2 ms per codec. In practice, the total delay will be slightly longer.

4. Sample error detection

One parity bit is applied to the 7 most significant bits of each transmitted sound sample such that the parity group is odd.

5. Sample interleaving

The companding block contains 96 sound samples (48 from each sound signal). The sound samples within the companding block are organized in 8 successive frames complying with CCITT Recommendation G.704. Each frame contains 6 samples from each sound signal, with the associated parity bits. Adjacent sound samples within the companding block are separated by four Recommendation G.704 frames. This is shown diagrammatically in Fig. 2. The first four frames carry all the odd-numbered samples from both sound signals; the second four frames carry all the even numbered samples.

In the event of an error burst corrupting consecutive bits for a period equivalent to up to four frames, the erroneous samples should be concealed by interpolation between adjacent samples (from part of the block unaffected by the error burst).

6. **Signalling in parity** [Chambers, 1985]

There are 96 parity bits per 1 ms companding block. Some extra data bits are carried by modifying the parity bits as follows:

6.1 *Scale factor transmission*

Each bit of each 3-bit scale factor word is carried in the parity of 8-sound samples defined in § 6.4, according to the following rule: a scale factor bit which is "0" causes the parity of the 8 samples to be unchanged; a scale factor bit which is "1" modifies the parity. At the decoder, a majority decision process is used to determine the scale factor bits and restore the original parity bit. The samples are then checked for errors in the normal way.

6.2 *Channel status*

Compressed channel status (see § 9) is carried in exactly the same manner as the scale factor bits.

6.3 *Framing signals*

Multiframe alignment signals (MFA) (see § 7) and frame slip detection signals (FSD) (see § 8) are carried by modifying the parity of single samples. These signals do not have the benefit of majority decision decoding, but are inherently predictable and may be decoded reliably.

6.4 *Signalling within the companding block*

The sample-interleaved companding block is represented in Fig. 3 by the rectangle at the bottom of the diagram. Each row within the rectangle represents a Recommendation G.704 frame, and is subdivided into 12 squares representing the 12 samples. The diagram further shows the modification of the parity bits associated with the 12 samples in each frame by the signalling in parity mentioned above.

Sign bit … LSB	Scale factor $S_2\ S_1\ S_0$
0 1 X X X X X X X X X X X X X X X X X X	0 0 0
0 0 1 X X X X X X X X X X X X X X X X X	0 0 1
0 0 0 1 X X X X X X X X X X X X X X X X	0 1 0
0 0 0 0 1 X X X X X X X X X X X X X X X	0 1 1
0 0 0 0 0 1 X X X X X X X X X X X X X X	1 0 0
0 0 0 0 0 0 1 X X X X X X X X X X X X X	1 0 1
0 0 0 0 0 0 0 1 X X X X X X X X X X X X	1 1 0
0 0 0 0 0 0 0 0 X X X X X X X X X X X X	1 1 1
1 1 1 1 1 1 1 1 X X X X X X X X X X X X	1 1 1
1 1 1 1 1 1 1 0 X X X X X X X X X X X X	1 1 0
1 1 1 1 1 1 0 X X X X X X X X X X X X X	1 0 1
1 1 1 1 1 0 X X X X X X X X X X X X X X	1 0 0
1 1 1 1 0 X X X X X X X X X X X X X X X	0 1 1
1 1 1 0 X X X X X X X X X X X X X X X X	0 1 0
1 1 0 X X X X X X X X X X X X X X X X X	0 0 1
1 0 X X X X X X X X X X X X X X X X X X	0 0 0

Bits truncated in H11 channels

Bits truncated in H11 and H12 channels

X = 1 or 0

FIGURE 1a – *Coding table*

d01-sc

Sign bit … LSB	Scale factor $S_2\ S_1\ S_0$
0 1 X X X X X X X X X X X X X X X X	0 0 0
0 1 X X X X X X X X X X X X X X X X	0 0 1
0 1 X X X X X X X X X X X X X X X X	0 1 0
0 1 X X X X X X X X X X X X X X X 1	0 1 1
0 1 X X X X X X X X X X X X X X 1 1	1 0 0
0 1 X X X X X X X X X X X X X 1 1 1	1 0 1
0 1 X X X X X X X X X X X X 1 1 1 1	1 1 0
0 X X X X X X X X X X X X 1 1 1 1 1	1 1 1
1 X X X X X X X X X X X X 1 1 1 1 1	1 1 1
1 0 X X X X X X X X X X X X 1 1 1 1	1 1 0
1 0 X X X X X X X X X X X X X 1 1 1	1 0 1
1 0 X X X X X X X X X X X X X X 1 1	1 0 0
1 0 X X X X X X X X X X X X X X X 1	0 1 1
1 0 X X X X X X X X X X X X X X X X	0 1 0
1 0 X X X X X X X X X X X X X X X X	0 0 1
1 0 X X X X X X X X X X X X X X X X	0 0 0

b_0 … b_{14} … b_{17}

Bits transmitted in H11 channels

Bits transmitted in H12 channels

X = 1 or 0

FIGURE 1b – *Transmitted bits*

d02-sc

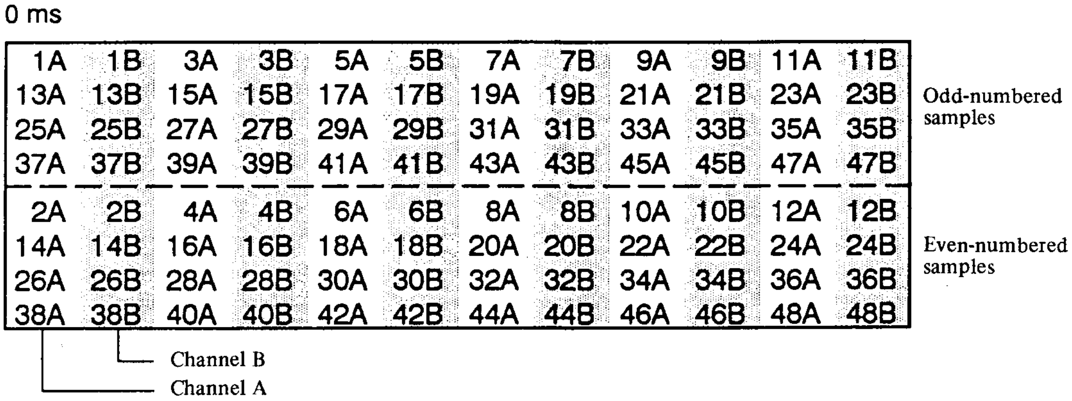

FIGURE 2 – *Sample interleaving within the companding block*

1 ms = 1 companding block

d03-sc

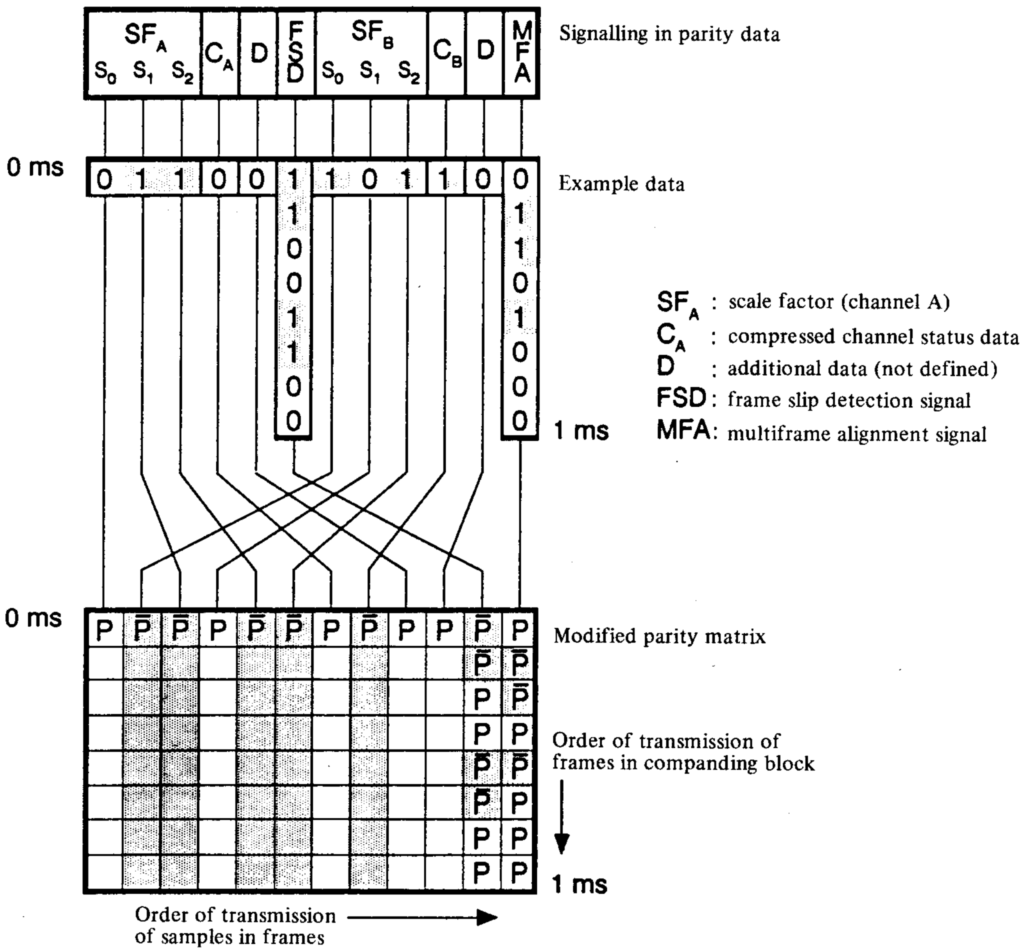

FIGURE 3 – *Signalling in parity matrix and data*

d04-sc

7. Synchronization and frame alignment

The sampling frequency of the sound signal must be synchronous with the bit clock of the transmission system.

Frame alignment of the companding block and multiframe (see § 9) is accomplished by application of a 1536-bit chain code, MFA. The generator is shown in Fig. 4a and the corresponding circuit for locking in Fig. 4b [Chambers, 1985].

The start word of the MFA generator is 10000110100 and it corresponds to the first frame of the multiframe. The generator will produce a sync pulse after 1536 bits (192 ms) and automatically reset to the start word.

The start of the multiframe may be locked to the interface Z-preamble as shown in Fig. 6.

The 1 ms companding block sync is generated by decoding the corresponding register contents.

The multiframe alignment signal (MFA) is signalled in parity (see § 6).

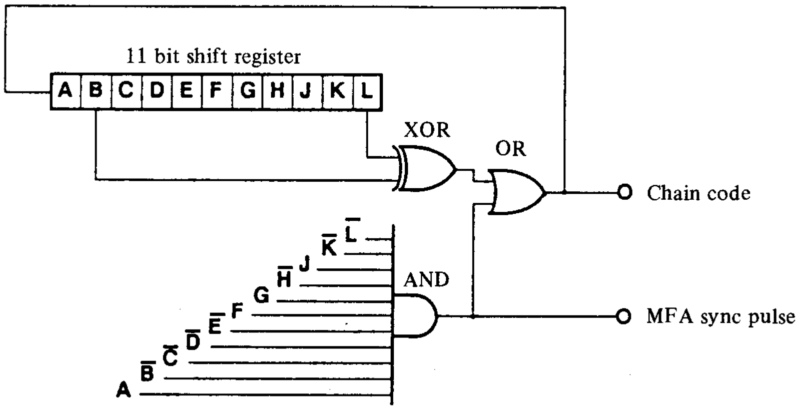

FIGURE 4a – *Chain code generator*

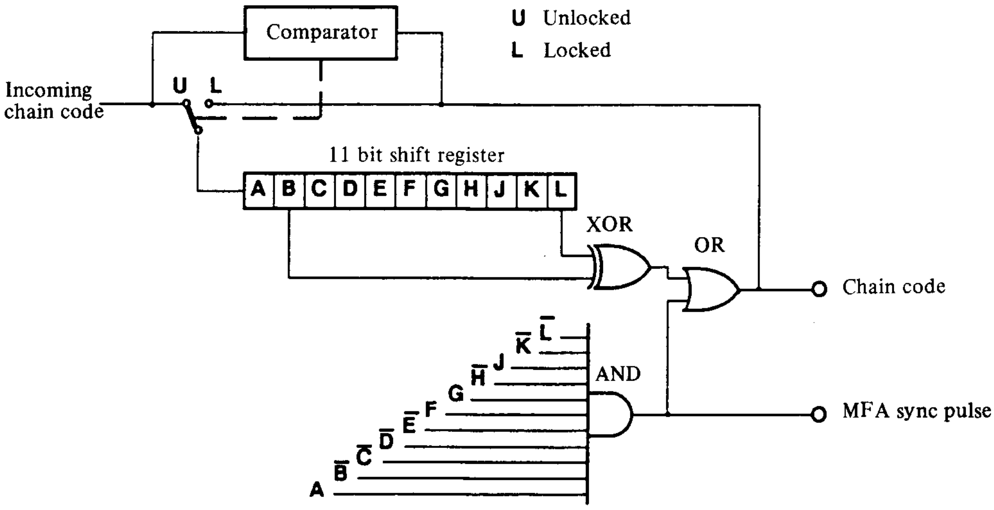

FIGURE 4b – *Chain code synchronization circuit*

d06-sc

8. Frame slip detection and management

A controlled slip is defined as suppression or repetition of a Recommendation G.704 frame.

A fixed of bits, FSD, (. . . 110011001100 . . .) is signalled in parity to assist decoders to detect controlled slips during transmission. The FSD is transmitted as shown in Fig. 3 (i.e. the FSD is frame-aligned with the companding block).

With a companding block of 1 ms and traditional framing methods, it would normally take a number of such frames (7-8 ms) to detect a slip and re-align the companding frame.

With the FSD, it will only take a few Recommendation G.704 frames to detect that a slip has occurred as the phase of the sequence will shift + or –90° (depending on whether a frame has been suppressed or repeated). Through a modulo-2 operation on the received and the expected FSD it is possible to detect, within two frames or more precisely, in which "pair" (1 and 2, 3 and 4, 5 and 6, or 7 and 8) of frames within the companding block the slip has occurred.

A suggested strategy for interpolation after a detected slip is shown in Fig. 5a and 5b, where the decoder produces a decoded block of the same length as the received block. Only channel A is shown; channel B is handled identically (Fig. 2 shows the transmitted sample sequence).

Positive slip – one frame is repeated

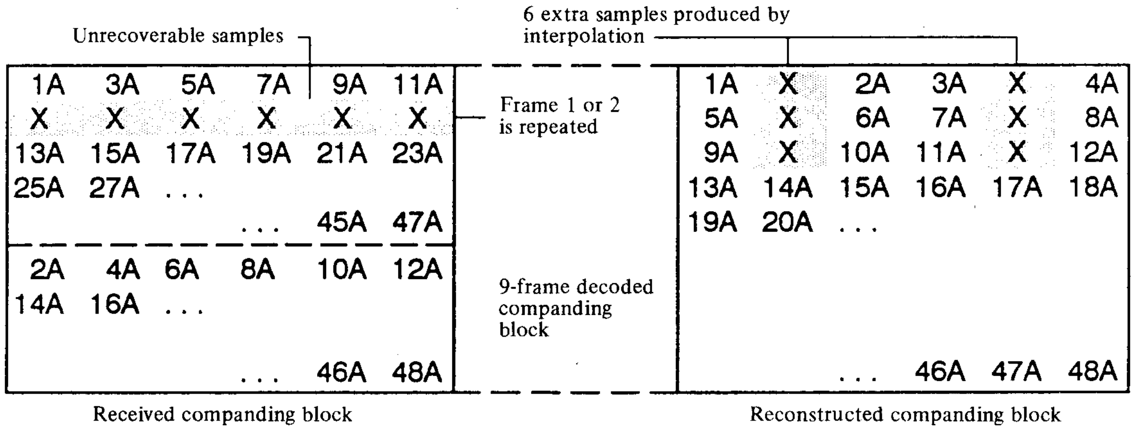

FIGURE 5a – *Strategy for interpolation after a detected positive slip*

d07-sc

By suitable decoder arrangements the scale factor can still be correctly decoded.

As all sound samples are stored for 1 ms (the length of the companding block) in the decoder, it is possible to move the companding frame boundary pointer such that the correct scale factor is applied to the correct number of samples.

Negative slip – one frame is deleted

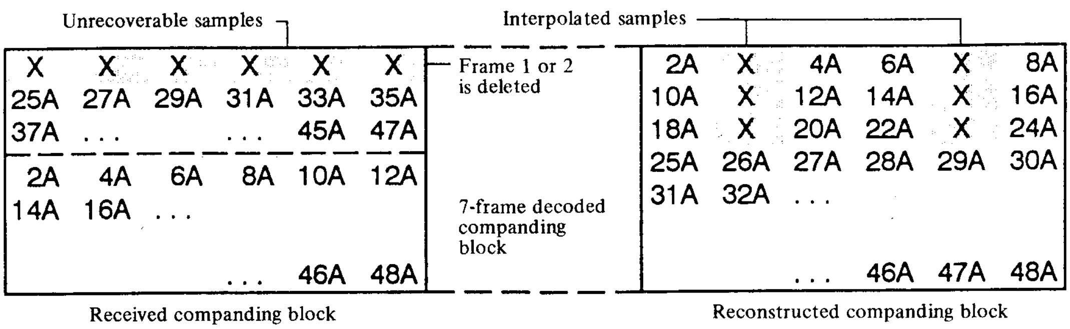

FIGURE 5b – *Strategy for interpolation after a detected negative slip*

Note that it is not possible to tell which frame has been deleted. Consequently, of the 12 transmitted samples, 6 are not received and 6 cannot be identified. In the reconstructed companding block, 6 samples are simply omitted (to adjust the length of the reconstructed block) and 6 are replaced by interpolation.

d08-sc

9. Channel status

Channel status data in the digital audio interface consists of a 192 bit (24 byte) cycle, which is repeated in 4 ms (one block of the interface).

The channel status is signalled in parity according to the description in § 6.2 and 6.4. This method of signalling provides one channel status bit for each audio signal per 1 ms companding block, enabling the system to carry one channel status block every 192 ms. This is shown in Fig. 6.

Because only one data block out of 48 is transmitted, the two counters (local sample and time of day address code) must be incremented in the decoder by the appropriate amount.

The start of the multiframe is signalled by the multiframe alignment signal defined in § 7. The time codes carried in the compressed channel status refer to the timing of the first sample in the multiframe.

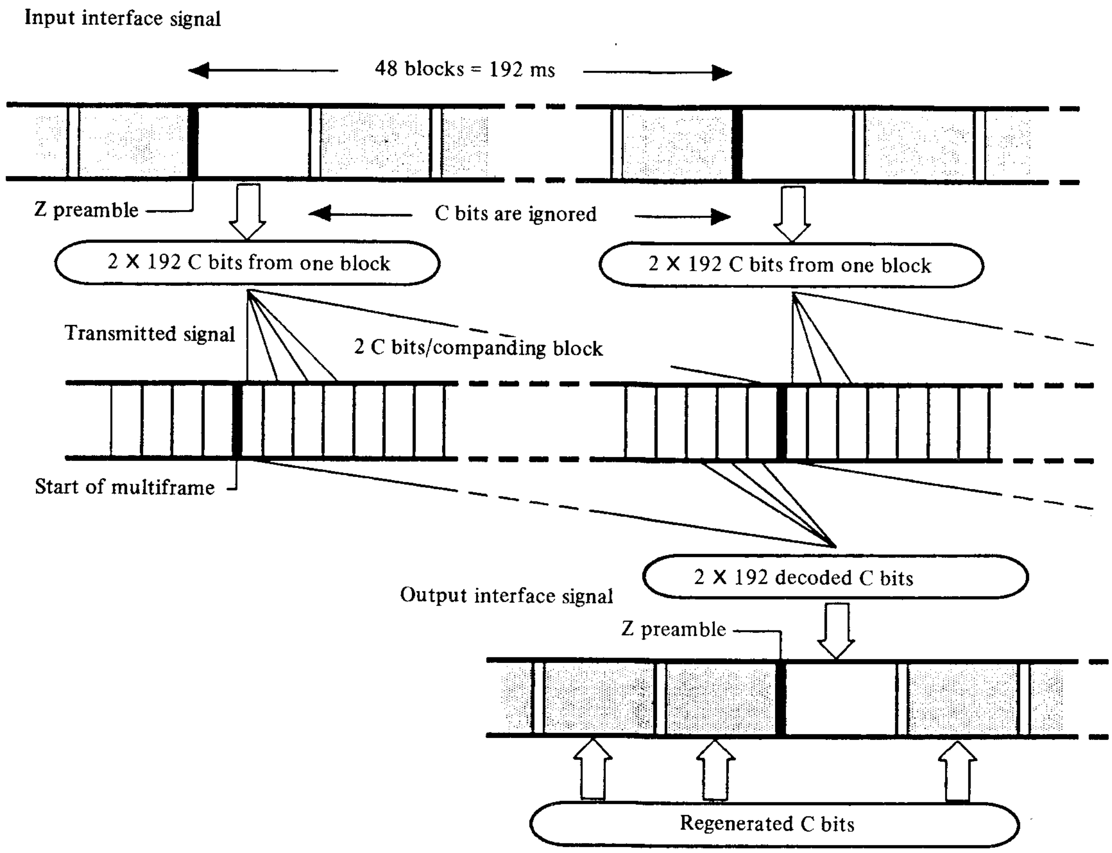

FIGURE 6 – *Transmission of channel status data*

d09-sc

10. User data

The user data bit in each sub-frame of the interface is transmitted transparently in the H12 channel only.

11. Frame structure and bit interleaving

The Recommendation G.704 frame for the H11 channel contains 192 usable bits, and that for the H12 channel contains 240 usable bits. Twelve 15-bit sound samples, each accompanied by its parity bit, may be carried in either type of frame. Additionally, the H12 frame may carry sufficient extra audio bits to increase the length of the compressed audio samples to 18 bits, and one user data bit per sample.

The organization of the 24 octets of data which are common to both H11 and H12 channels are identical in both types of frame, to facilitate remultiplexing at the interface between H11 and H12 channels. The common data occupies the entire available capacity of the H11 frame, and the first 24 available octets in the H12 frame, as shown in Fig. 7. The remaining 6 octets in the H12 frame carry user bits and extra audio bits.

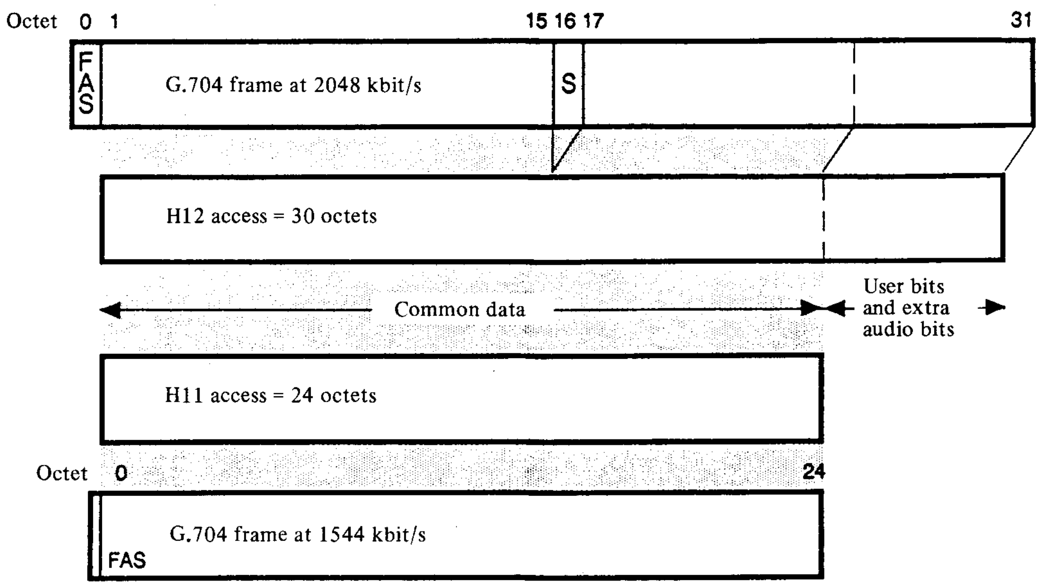

FIGURE 7 – *Organization of H11 and H12 frames*

FAS: frame alignment signal
S: signalling

d10-sc

11.1 *Organization of the 24 octets common to both H11 and H12*

The number of bits available for interleaving is 192. The bit-interleaved sample block can be described by a 8 × 24 matrix (Fig. 8). The numbers designate the samples within any frame of a companding block, interleaved as described in § 5 (see Fig. 2), in transmission order.

Row/column	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	1	1	1	2	1	7	1	8	1	3	1	4	1	9	1	10	2	5	2	6	2	11	2	12
2	2	1	2	2	2	7	2	8	7	3	7	4	7	9	7	10	7	5	7	6	7	11	7	12
3	8	1	8	2	8	7	8	8	8	3	8	4	8	9	8	10	3	5	3	6	3	11	3	12
4	3	1	3	2	3	7	3	8	4	3	4	4	4	9	4	10	4	5	4	6	4	11	4	12
5	9	1	9	2	9	7	9	8	9	3	9	4	9	9	9	10	10	5	10	6	10	11	10	12
6	10	1	10	2	10	7	10	8	5	3	5	4	5	9	5	10	5	5	5	6	5	11	5	12
7	6	1	6	2	6	7	6	8	6	3	6	4	6	9	6	10	11	5	11	6	11	11	11	12
8	11	P1	11	P2	11	P7	11	P8	12	P3	12	P4	12	P9	12	P10	12	P5	12	P6	12	P11	12	P12

Protected bits and parity bits
Unprotected bits

Order of transmission

Numbering convention

1 corresponds to bits from samples 1A, 13A, 25A, ... or 38A
2 corresponds to bits from samples 1B, 13B, 25B, ... or 38B
3 corresponds to bits from samples 3A, 15A, 27A, ... or 40A
etc.

See Fig. 2

FIGURE 8 – *Bit-interleaving matrix*

d11-sc

Each even-numbered column represents the 7 protected bits (b_0-b_6) and the parity bit of one sample. These bits are filled into the matrix column by column. The progression of samples is 1-2-7-8-3-4-9-10-5-6-11-12 in order to maximize the distance between any two neighbouring samples of one channel and at the same time keep co-timed samples (in stereo mode) together. The transmission sequence always starts with the LSB and the MSB always precedes the P-bit.

The 8 least significant bits of the 15-bit word (b_7-b_{14}) are filled into the odd columns of the matrix, but in this case row by row. These bits will be kept close after interleaving, in order to minimize the number of impaired samples when long error-bursts occur. The transmission sequence is: b_7-b_9-b_{11}-b_{13}-b_8-b_{10}-b_{12}-b_{14}. Bits read from columns 1, 5, 9, 13, 17 and 21 are inverted prior to transmission.

The first 24 available octets of the Recommandation G.704 frame are filled using the data represented in Fig. 8, by reading row by row from the matrix. The distance between protected sample-bits or parity bits from each sample is 24.

Octets 0 and 16 of the 2048 kbit/s Recommandation G.704 frame are not interleaved.

Note – When the connection is established over 1544 kbit/s circuits which are not bit-sequence independent, the minimum pulse density requirement may be guaranteed by forcing to digital “1” bits in columns 7, 15 or 23, where necessary. This operation is equivalent to the “z-operation” described in CCITT Recommendation G.802, § 2.1, and should be performed at the interface with such circuits.

11.2 *Organization of the last 6 octets in H12*

The last 3 LSB and the user bit from each sample are transmitted in the remaining 6 octets in the H12 frame by reading row by row from the matrix shown in Fig. 9. As before, the numbers designate samples, and the order of transmission is LSB first.

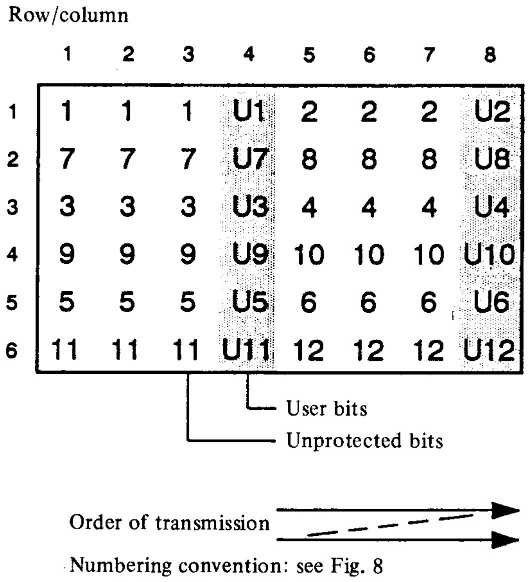

FIGURE 9 – ***Interleaving of supplementary audio bits and user bit in the H12 channel***

d12-sc

REFERENCES

CHAMBERS, J. P. [1985] Signalling in parity: a brief history. British Broadcasting Corporation, BBC RD 1985/15.