INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

J.55
(ex CMTT.718)
(06/90)

TELEVISION AND SOUND TRANSMISSION

DIGITAL TRANSMISSION OF HIGH-QUALITY SOUND-PROGRAMME SIGNALS ON DISTRIBUTION CIRCUITS USING 480 kbit/s (496 kbit/s) PER AUDIO CHANNEL

ITU-T Recommendation J.55

(Formerly Recommendation ITU-R CMTT.718)

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.55 (formerly Recommendation ITU-R CMTT.718) 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 Similarly, in this reform process, the CCIR and the IFRB have been replaced by the Radiocommunication Sector (ITU-R).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.55[1)]

DIGITAL TRANSMISSION OF HIGH-QUALITY SOUND-PROGRAMME SIGNALS ON DISTRIBUTION CIRCUITS USING 480 kbit/s (496 kbit/s) PER AUDIO CHANNEL

(1990)

The CCIR,

CONSIDERING

a) that the distribution of high-quality sound-programme signals from the studio to transmitters and users does not require further audio downstream processing;

b) that in general more than one high-quality sound-programme signal has to be conveyed over distribution circuits as in the case of DBS;

c) that a sampling frequency of 32 kHz is recommended for the digital transmission of high-quality sound-programme signals (Recommendation 606);

d) that some digital broadcasting applications may require system performances that go beyond those offered by equipment complying with CCIR Recommendation 660;

e) that the high-quality sound-programme signals should interface the ISDN at the H1 level as specified in CCITT Recommendation I.412,

UNANIMOUSLY RECOMMENDS

1. that for distribution applications where a sampling frequency of 32 kHz is used and where a dynamic range corresponding to more than 14 bits is required the coding method described in § 1 of Annex I should be used on links providing a BER of less than 10^{-5};

2. that for transmission at H12 level, two stereophonic programmes or four monophonic programmes should be multiplexed according to the format described in § 2 of Annex I;

3. that for cases where a higher ancillary data capacity is required and dedicated 2048 kbit/s links are available, the coding method and multiplexing format described in Annexe II should be used.

Note – Exchange of international digital sound-programme signals on networks with other hierarchical rates (1544 kbit/s in North America) shall be made using Recommendation 660.

---

1) Formerly Recommendation ITU-R CMTT.718.

ANNEX I

DIGITAL TRANSMISSION OF HIGH-QUALITY SOUND-PROGRAMME SIGNALS ON DISTRIBUTION CIRCUITS USING 480 kbit/s AUDIO CHANNEL

1. Coding characteristics

1.1 *Sampling frequency*

The sampling frequency shall be 32 kHz. The sampling frequency tolerance shall be $\pm 5 \times 10^{-5}$, as specified in CCITT Recommendations G.732 and G.733 for primary PCM multiplex equipment. This sampling frequency is consistent with that indicated in CCIR Recommendation 606.

1.2 *Coding method*

16/14-bit floating point companding with a 2 ms coding block length (i.e. 64 consecutive samples per block) and a 3-bit scale factor (transmitted by signalling in parity).

The 16-bit samples of the sound signal are represented in a 2 s complement format. The first bit of each word is the MSB (sign bit, $0 \sim +$), and the last the LSB. Using a floating point system, the 16-bit samples are converted into 14-bit code words for transmission.

A 3-bit scale factor applying to a block of 64 samples indicates how many of the bits (0 . . . 7) following the sign bit (y_1) in all sampled words have the same value as the sign bit (Fig. 1a)). The redundancy indicated by the scale factor does not need to be transmitted. Instead, the samples and their relevant information must be shifted towards the sign bits (floating-point system). This allows the 15th and 16th bits of the source code words to be transmitted in the case of low signal amplitudes. The bits marked Z1 to Z5 have not yet been assigned (Fig. 1b)).

At the receiving end the scale factor is used to shift the bits of the samples back to their original value. This yields 16-bit samples and limits the effects of unrecognized bit errors to the amplitude range indicated by the scale factor.

1.3 *Sample error-protection*

After having applied the floating-point technique for reducing the bits per sample from 16 to 14, a parity bit is calculated on the seven most significant bits of each sample, such that the number of “1”s in the group of the seven protected bits and the parity bit is odd. Thus, the best protection against clicks due to bit errors is guaranteed provided concealment in the form of the arithmetic mean value of the samples adjacent to the faulty sample is performed at the receiving end. If the concealment is done after the 14/16 bit reconversion, errors in least significant bits can also be concealed in an optimum way.

1.4 *Ancillary data*

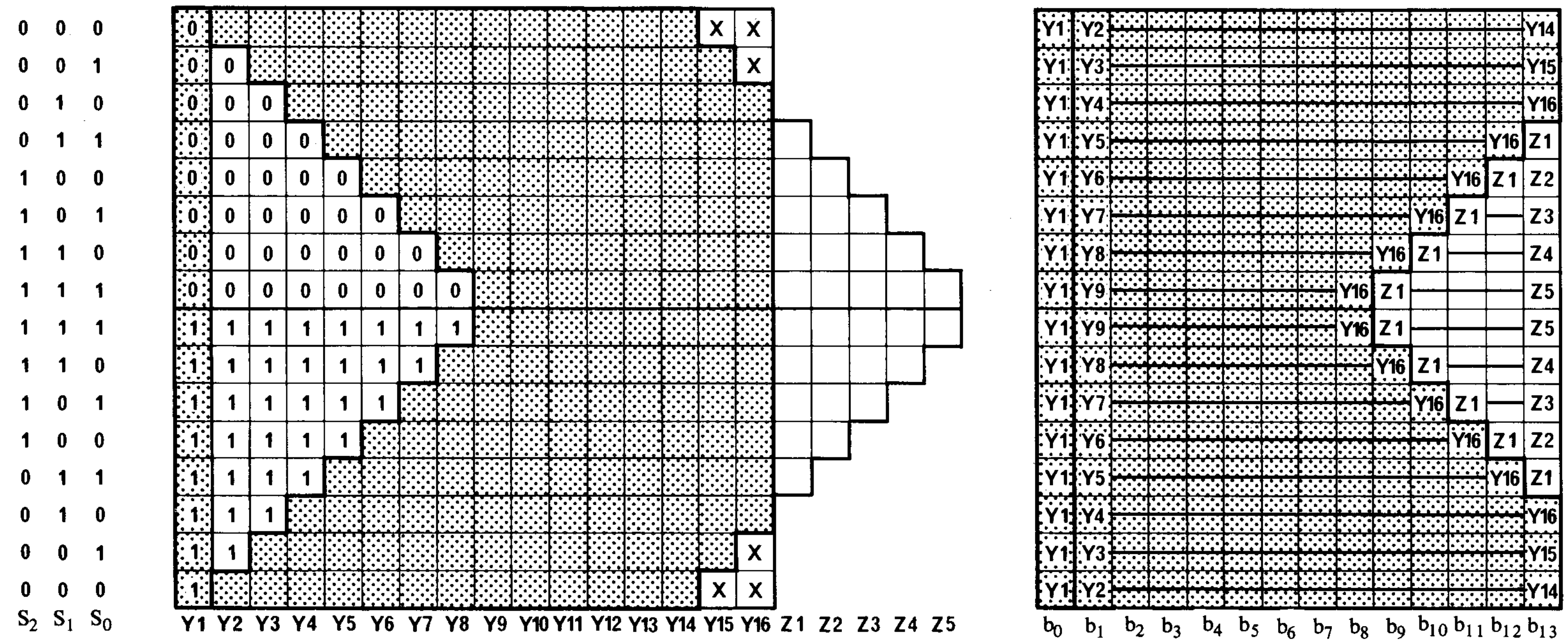
A data capacity of 4 kbit/s per channel is transmitted by signalling in parity.

1.5 *Signalling in parity* [Chambers, 1985]

Signalling in parity is achieved by transmitting the parity bits of an odd number of successive samples without inversion or after inversion, depending on the bit to be signalled. Inversion has to be done, if the bit to be transmitted is 1. The signalling technique used for the scale factor, its related parity bit, and the ancillary data is based on majority-decision logic at the receiver: for each channel, it processes twelve groups of five consecutive samples (three for the scale factor, one for its parity bit and eight for the ancillary data) to recognize simultaneously the odd or even parity of the sample and the data signalled in parity (see Fig. 2). A similar process is used to recognize the synchronization of the 2 ms frames through one group of four consecutive samples.

1.6 *Total bit rate*

With the parameters mentioned above the total bit rate required for one monophonic channel is 480 kbit/s [$32 \text{ kHz} \times (14 + 1)$ bit].



a) Coding scheme

b) Transmission format

FIGURE 1 – *16/14 bit floating-point method*

relevant sound signal code word range of the 16-bit source signal words

X X non-transmittable bits of the 16-bit source signal words

d01-sc

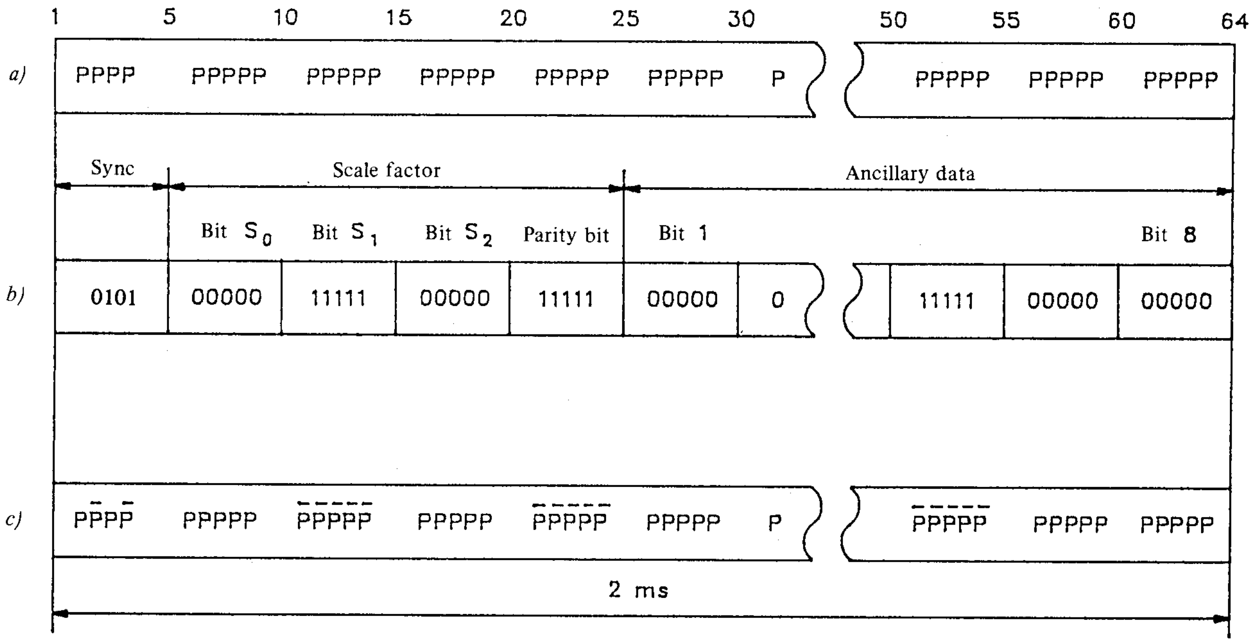


FIGURE 2 – *Signalling in parity coding for one audio channel*

a) Parity bits (P denotes the parity bit associated to each individual audio sample).

b) Sync pattern, scale factor, parity bit for the scale factor and ancillary data bits. The parity bit for the scale factor should be even as shown in the example.

c) Modified parity bits.

d02-sc

2. Frame structure for transmission

Within the ISDN, a usable bit rate of 1920 kbit/s for one multiplex frame is available according to CCITT Recommendation G.704.

In order to ensure the compatibility between the signals transmitted at 480 kbit/s according to this Recommendation with those transmitted at 384 kbit/s according to CCITT Recommendation G.737, the bits of each audio channel should be allocated according to the present proposal depicted in Fig. 3. The bits of each audio channel should be transmitted as a group of 30 bits within a half-frame.

Moreover, the bits of each sample are interleaved such that the most significant bit (MSB) is followed by the least significant bit (LSB), etc. (see Fig. 3). This arrangement of bits has proven to be a good protection against double errors, which otherwise cannot be recognized by a single parity bit and thus cannot be concealed.

On this basis, channels using 480 kbit/s and 384 kbit/s can be combined according to Table I:

TABLE I

	Number of channels	
	480 kbit/s	384 kbit/s
I	4	0
II	3	1
III	2	2
IV	1	3
V	0	5

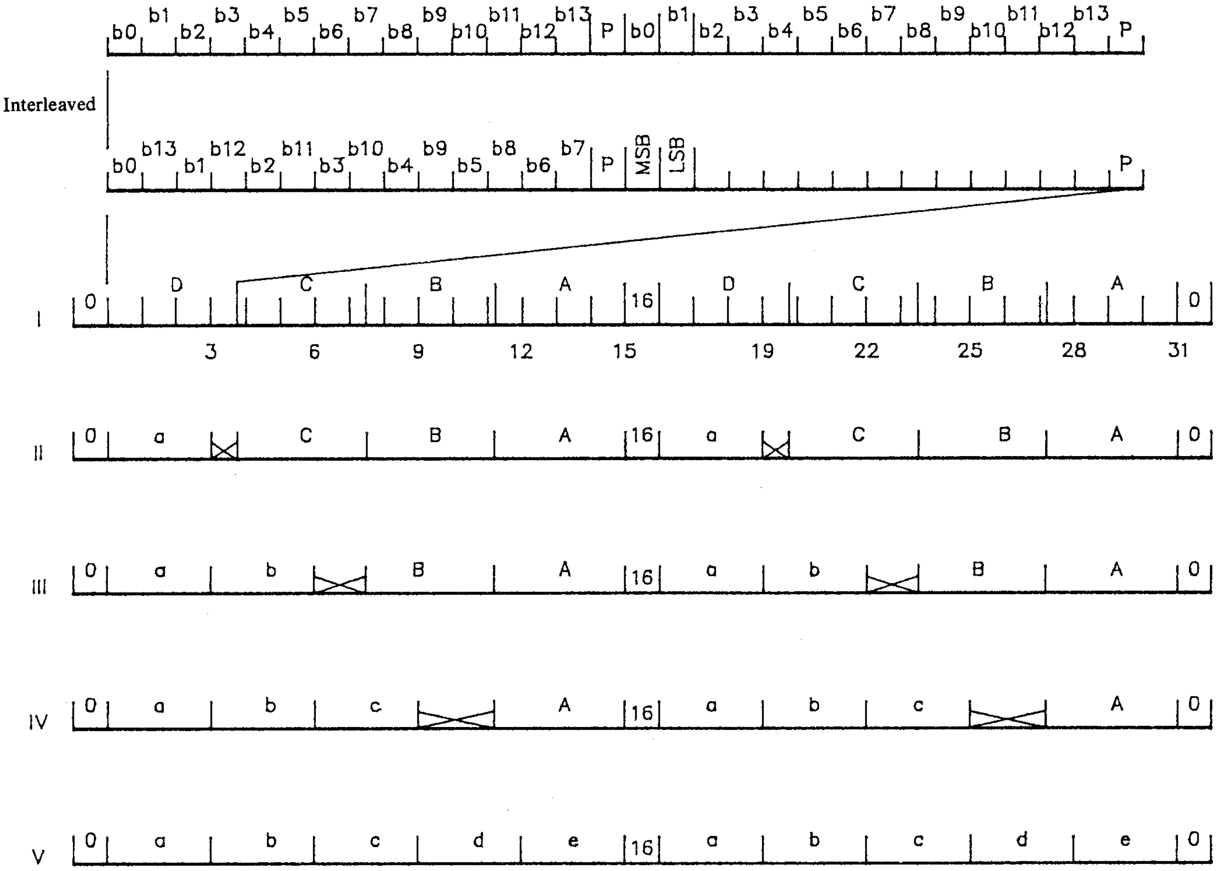


FIGURE 3

A, B, C, D: monophonic channel 480 kbit/s
a, b, c, d, e: monophonic channel 384 kbit/s
MSB: most significant bit
LSB: least significant bit
⋈ : not usable for sound transmission

d03-sc

Two monophonic channels of the same coding scheme can be combined to form a stereophonic channel. In some realized systems for 384 kbit/s channels only the combinations a, b and c, d are possible to form stereophonic channels.

REFERENCES

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

ANNEX II

DIGITAL TRANSMISSION OF HIGH-QUALITY SOUND-PROGRAMME SIGNAL ON DISTRIBUTION CIRCUITS USING 496 kbit/s PER AUDIO CHANNEL

1. Coding characteristics

1.1 *Sampling frequency*

See Annex I.

1.2 *Coding method*

See Annex I.

1.3 *Sample error-protection*

See Annex I.

1.4 *Ancillary data*

12 kbit/s per channel are available for transmission of ancillary data.

1.5 *Signalling in parity* [Chambers, 1985]

Signalling in parity is achieved by transmitting the parity bits of an odd number of successive samples without inversion or after inversion, depending on the bit to be signalled. Inversion has to be done, if the bit to be transmitted is “1”. The signalling technique used for the scale factor is based on majority-decision logic at the receiver: for each channel, it processes three groups of twenty-one consecutive samples to recognize simultaneously the odd or even parity of the sample and the data signalled in parity (see Fig. 4).

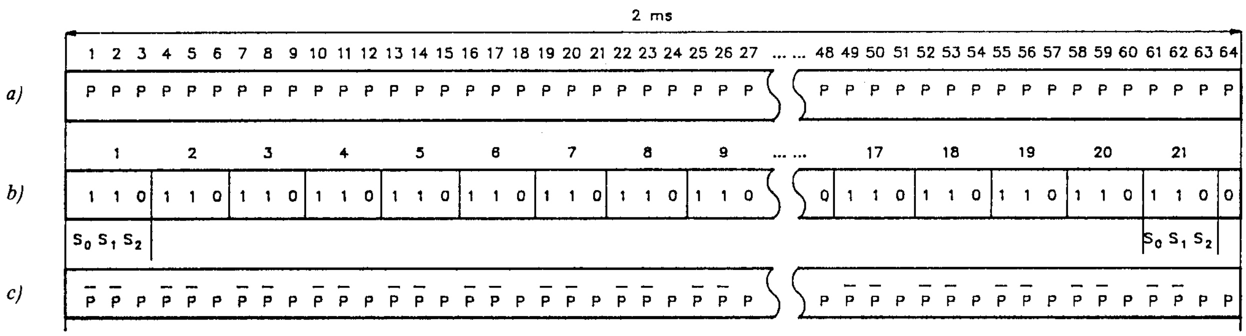


FIGURE 4 – *Signalling in parity coding for one audio channel*

a) **Parity bits (P denotes the parity bit associated to each individual audio sample)**
b) **Scale factor bits**
c) **Modified parity bits**

d04-sc

1.6 *Synchronization*

4 kbit/s are used for synchronization of the floating point companding block.

1.7 *Total bit rate*

With the parameters mentioned above the total bit rate required for one monophonic channel is 496 kbit/s [32 kHz × (14 +1) bits + 12 kbit/s + 4 kbit/s].

2. Frame structure for transmission

The frame structure for transmission is based on an interface operating at 1024 kbit/s.

Two of the 496 kbit/s channels are combined with a frame alignment signal to form a multiplexed signal of 1024 kbit/s, which constitutes the bit rate of the interface. The structure of the frame is almost the same as that specified in CCITT Recommendation G.704 for the primary hierarchical level of 2048 kbit/s. It should be noted that the frame repetition rate is 4 kHz instead of 8 kHz. The frame format is shown in Fig. 5.

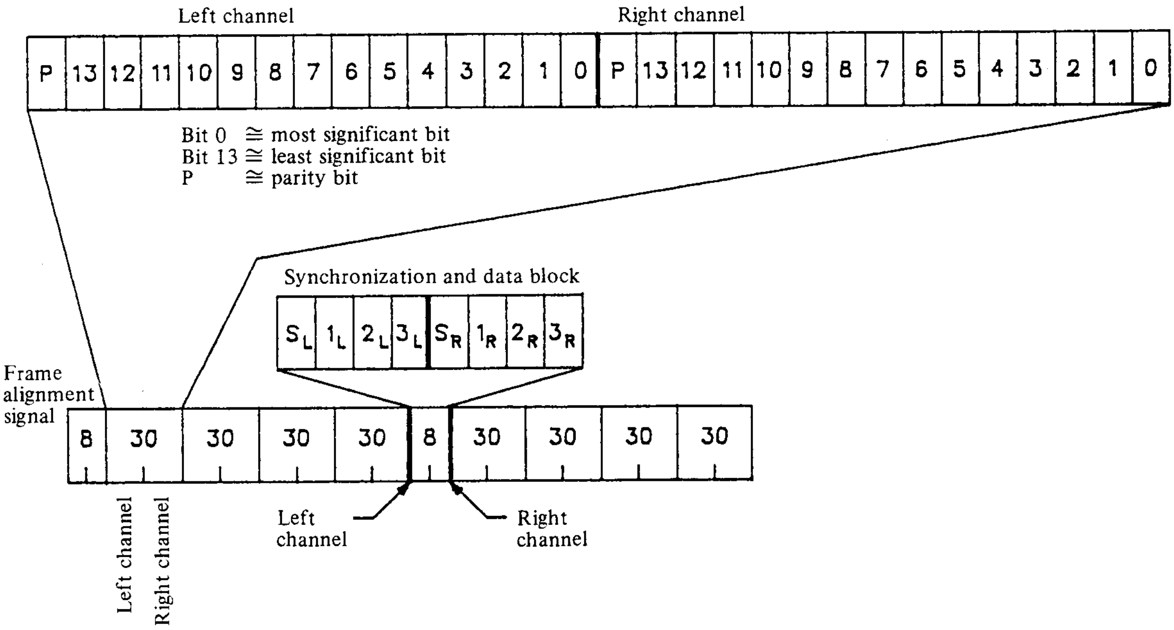


FIGURE 5 – *Frame for the 1024 kbit/s signal with a length of 256 bits and a 4 kHz repetition rate*

S_L, S_R: synchronization signal for floating point companding

1_L, 2_L, 3_L / 1_R, 2_R, 3_R } ancillary data signal

d05-sc

For each channel, the synchronization of the floating point companding block (64 samples carried in 8 frames) is achieved by means of an 8-bit synchronization word.

Two synchronization words are defined:

$$S_y = 00011011$$

$$S_1 \ldots\ldots S_8$$

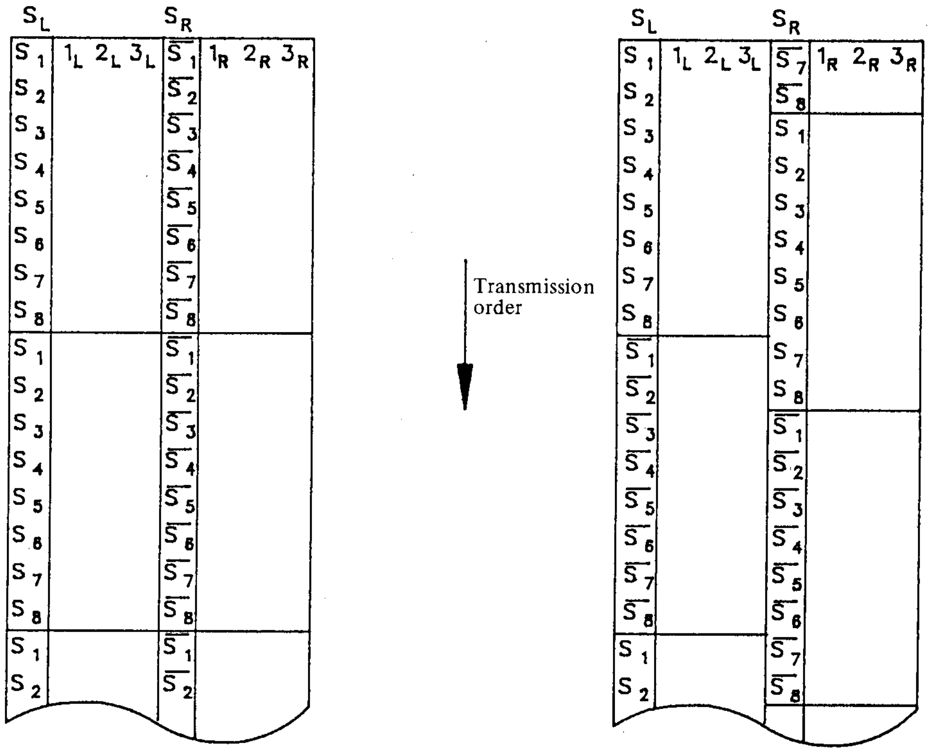
and the inverted form

$$\overline{S_y} = 11100100$$

$$\overline{S_1} \ldots\ldots \overline{S_8}$$

For a stereophonic pair the synchronization word used is S_y for the left channel and $\overline{S_y}$ for the right channel.

For monophonic applications, the synchronization word used for each channel is alternatively S_y and $\overline{S_y}$ (see Fig. 6).



a) Stereophonic pair

Note – The relation between S_L and S_R should be as shown.

b) Two monophonic channels

Note – For monophonic application there is no need for a fixed relation between S_L and S_R.

FIGURE 6 – *Transmission of the synchronization words*

d06-sc

By way of synchronous insertion into a 2048 kbit/s frame structure, it is not necessary to transmit the frame alignment signal (FA) which is contained in the frame structure shown in Fig. 5.

Thus, for one channel the net bit rate for transmission is still 496 kbit/s.

In cases where dedicated 2048 kbit/s links are available, up to four 496 kbit/s channels can be combined in one multiplex frame according to CCITT Recommendation G.704.

In order to ensure compatibility between the signals transmitted at 496 kbit/s according to this Annex II with those transmitted at 384 kbit/s according to CCITT Recommendation G.737, the available transmission capacity should be allocated as shown in Fig. 7 [CCIR, 1982-86a].

In Fig. 7 each block L I, R I, L II, R II comprises 60 bits corresponding to a capacity equivalent to 4 samples. Each block results from bit- and sample-interleaving over two consecutive Recommendation G.704 frames.

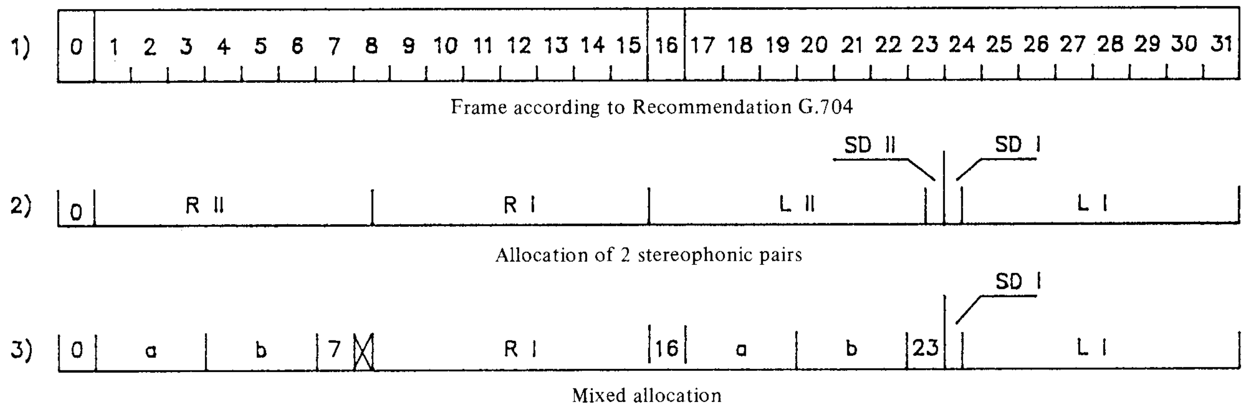


FIGURE 7 – *Allocation of available transmission capacity*

L I: 7.5 × 64 kbit/s = 480 kbit/s for left channel I
R I: 480 kbit/s for right channel I
L II: 480 kbit/s for left channel II
R II: 480 kbit/s for right channel II

SD I: 0.5 × 64 kbit/s = 32 kbit/s, to be divided into 2 × 16 kbit/s.
In the frames containing the frame alignment word the data capacity is allocated to the left channel I, in the other frames to right channel I. In each case 4 kbit/s are used for synchronization of the floating point companding block and 12 kbit/s are used for the transmission of ancillary data

SD II: as SD I but for channels L II and R II

a, b: each 2 × 3 × 64 kbit/s = 384 kbit/s monophonic channel

⊠ : capacity not usable (32 kbit/s) in this combination

Note – In line 3) above time slots 7 and 23 are available for telephone; the signalling channel (time slot 16) is not occupied by a sound channel.

d07-sc

REFERENCES

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

CCIR Documents

[1982-86]: **a.** CMTT/214 (Germany (Federal Republic of)).