INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

J.52
(07/96)

SERIES J: TRANSMISSION OF SOUND-PROGRAMME AND TELEVISION SIGNALS

Digital transmission of sound-programme signals

Digital transmission of high-quality sound-programme signals using one, two or three 64 kbit/s channels per mono signal (and up to six per stereo signal)

ITU-T Recommendation J.52

(Previously "CCITT Recommendation")

ITU-T J-SERIES RECOMMENDATIONS

TRANSMISSION OF SOUND-PROGRAMME AND TELEVISION SIGNALS

General Recommendations	J.1-J.9
General Recommendations concerning sound-programme transmissions	J.10-J.19
Performance characteristics of sound-programme circuits	J.20-J.29
Characteristics of equipment and lines used for setting up sound-programme circuits	J.30-J.39
Characteristics of equipment for coding analogue sound-programme signals	J.40-J.49
Digital transmission of sound-programme signals	**J.50-J.59**
Characteristics of circuits for television transmissions	J.60-J.69
Systems for television transmission over metallic lines and interconnection with radio-relay links	J.70-J.79
Digital transmission of television signals	J.80-J.89
Specific Recommendations for television transmission	J.90-J.99
Transmission of signals with multiplexing of video, sound and data, and signals of new systems	J.100-J.109
Interactive services	J.110-J.119

For further details, please refer to ITU-T List of Recommendations.

FOREWORD

The ITU-T (Telecommunication Standardization Sector) is a permanent organ of the International Telecommunication Union (ITU). 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, establishes the topics for study by the ITU-T Study Groups which, in their turn, produce Recommendations on these topics.

The approval of Recommendations by the Members of the ITU-T is covered by the procedure laid down in WTSC Resolution No. 1 (Helsinki, March 1-12, 1993).

ITU-T Recommendation J.52 was revised by ITU-T Study Group 9 (1993-1996) and was approved under the WTSC Resolution No. 1 procedure on the 11th of July 1996.

NOTE

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

CONTENTS

SUMMARY

The system described in this Recommendation makes possible the transmission of sound signals, coded by new methods recommended by ITU-R Study Group 10 and standardized by ISO/IEC 11172-3, on 1 to 6 B-channels of the N-ISDN or on permanent connections with a 2048 kbit/s (or 1544 kbit/s) frame. In the case of transmission via the N-ISDN, Recommendation H.221 is used to aggregrate up to six B-channels while maintaining the bit sequence integrity. The revised text contains clarifications that reflect experience gained in implementing the current version of this Recommendation, and a new chapter on transmission of program-associated data in the ancillary data field of the ISO frame.

INTRODUCTION

The ISO and IEC have approved the international standard 11172-3 for the reduction of the bit rate of high-quality digital sound-programme signals. With this system, it is possible to transmit high-quality mono sound-programme signals with bit rates in the range of 64 to 192 kbit/s. Stereo signals are encoded in single bit streams with bit rates in the range 128 to 384 kbit/s. If sound signals coded by ISO/IEC 11172-3 have to be transmitted in telecommunication networks, the properties of the networks are to be taken into account. The ITU-T has approved some Recommendations which describe the structure and function of 64 kbit/s channels for audiovisual applications (Recommendation H.200 contains a list of such Recommendations).

The N-ISDN allows the assembly of single channels with the bit rate of 64 kbit/s. If the bit rate is greater than 64 kbit/s (two times, three times etc.), measures are necessary to maintain the bit sequence. The equipment described in the following enables the transmission of source-coded sound signals on channels of N-ISDN or on connections with a 2048 kbit/s (or 1544 kbit/s) frame.

DIGITAL TRANSMISSION OF HIGH-QUALITY SOUND-PROGRAMME SIGNALS USING ONE, TWO OR THREE 64 kbit/s CHANNELS PER MONO SIGNAL (AND UP TO SIX PER STEREO SIGNAL)

(revised in 1996)

0 Scope

The ITU-R Recommendation BS.1115 recommends, based on extensive tests carried out by ITU-R Study Group 10 for different applications of the transmission in a broadcast chain (contribution, distribution, emission, commentary), the bit rate reduction system according to ISO/IEC 11172-3. The encoding layers and bit rates (per mono channel) recommended by ITU-R SG 10 depend on the application and are in the range of 120 - 180 kbit/s using Layer II, except for commentary where 60 kbit/s using Layer III may be used. For details, see ITU-R BS.1115.

For transmission of high-quality sound programmes and associated data it is desirable to use standardized 64 kbit/s channels or multiples thereof. This Recommendation defines formatter and reformatter for the transport of the bit rate reduced audio signals according to ISO/IEC 11172-3 and associated data on these standardized 64 kbit/s channels. The bit stream to be transmitted on 64 kbit/s channels could be protected optionally by means of error-correction measures.

The ITU-R has not (before the end of 1995) made any recommendation for the use of the ISO/IEC 13818-3 *lower sampling frequencies*. **If and only if** the extension to *lower sampling frequencies* is to be implemented, the provisions in Annex A shall apply. The implementation of the *lower sampling frequency* modes is in no way mandatory as they inherently limit the bandwidth, and would therefore not permit the high-quality criteria to be met.

1 General features

1.1 Purpose of the equipment

The purpose of the equipment is to process a bit rate reduced high-quality sound-programme signal for transmission on standardized 64 kbit/s channels. For this purpose, it is appropriate that the source encoder and the multiplexer form a unit with an output to the line interface (see Figure 1).

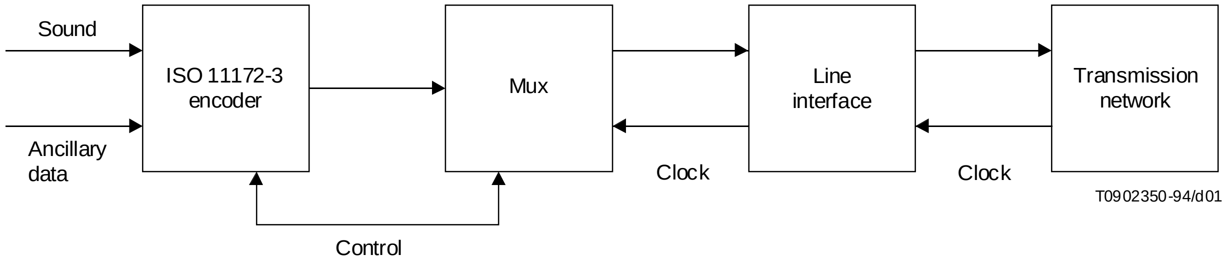

FIGURE 1/J.52

Connection between encoder, multiplexer and line interface

1.2 Incoming signals (PCM level)

The source encoder provides a bit stream according to ISO Standard 11172-3 (Layer I, Layer II or Layer III). Layer I is not recommended for broadcast applications, the block structure of Layers II and III with the block length 24 ms (in case of 48 kHz sampling frequency) is shown in Figure 2.

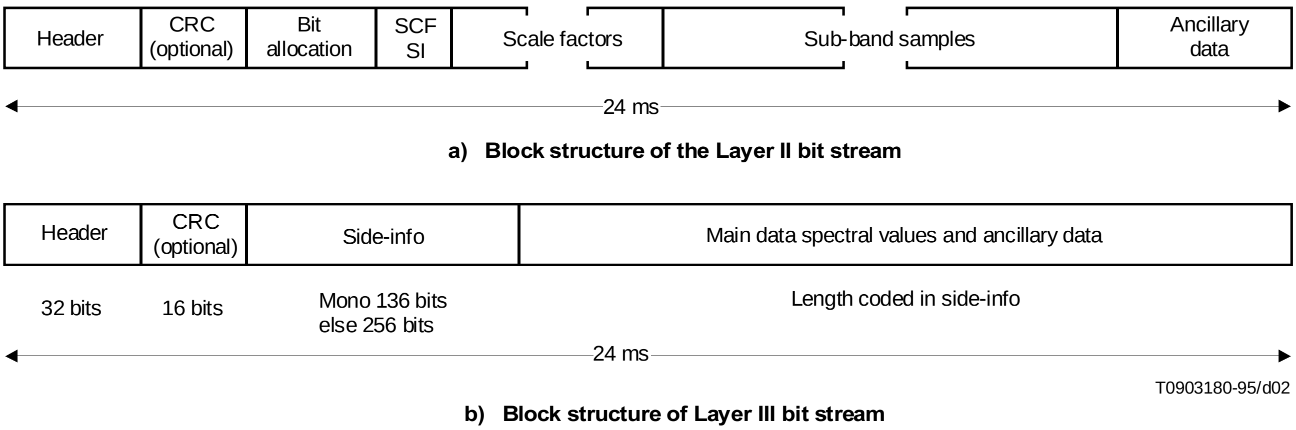

FIGURE 2/J.52

If possible, this bit stream has to be transmitted unchanged. In most cases, this is impossible for the following reasons:

– additional framing bits are necessary for the transmission on the ISDN B-channels;
– if it is necessary to include error control in order to correct the bit errors, another coding hierarchy level has to be used.

Table 1 gives all frame lengths defined in ISO 11172-3.

If the padding bit in the header is set, the frame contains an additional slot (1 byte for Layers II and III, 4 bytes for Layer I). The bit rate is encoded with 4 bits (bit rate index). The coding of the bit rate is different for the three layers (see ISO 11172-3). Layer I and sampling frequency 44.1 kHz are not recommended for broadcast applications.

2 Network interface formats

There is a distinction between dial-up connections (ISDN) and permanent connections. This distinction consists of the different routing of several B-channels constituting a virtual sound channel of more than 64 kbit/s, which requires a delay difference compensation in case of dial-up connections.

Two transmission formats are required:

– transmission of *one* mono/stereo signal using one to six 64 kbit/s channels or a single H_0 channel in a switched ISDN network;
– transmission of *one or more* mono/stereo signals using, for example, H_0 or H_1 channels for permanent connections.[1)]

[1)] For transmission of *one* mono/stereo signal at a bit rate of 1 to 6 × 64 kbit/s on permanent connections, the X.21 leased circuit option can be used. In this case, H.221 framing is optional.

TABLE 1/J.52

Possible frame lengths in bytes (without padding)

Sampling Frequency (Hz)	32 000			44 100			48 000		
Layer	I	II	III	I	II	III	I	II	III
Bit rate (bit/s)									
32 000	48	144	144	32	104	104	32	96	96
40 000	–	–	180	–	–	130	–	–	120
48 000	–	216	216	–	156	156	–	144	144
56 000	–	252	252	–	182	182	–	168	168
64 000	96	288	288	68	208	208	64	192	192
80 000	–	360	360	–	261	261	–	240	240
96 000	144	432	432	104	313	313	96	288	288
112 000	–	504	504	–	365	365	–	336	336
128 000	192	576	576	136	417	417	128	384	384
160 000	240	720	720	172	522	522	160	480	480
192 000	288	864	864	208	626	626	192	576	576
224 000	336	1008	1008	240	731	731	224	672	672
256 000	384	1152	1152	276	835	835	256	768	768
288 000	432	–	–	312	–	–	288	–	–
320 000	480	1440	1440	348	1044	1044	320	960	960
352 000	528	–	–	380	–	–	352	–	–
384 000	576	1728	–	416	1253	–	384	1152	–
416 000	624	–	–	452	–	–	416	–	–
448 000	672	–	–	484	–	–	448	–	–
NOTE – A dash "–" in the table means that the combination of sampling frequency, bit rate and layer is not specified in ISO/IEC 11172-3.									

2.1 Dial-up connections (switched ISDN network)

According to Recommendation I.412 the following possibilities exist:

a) *B-channel interface structures*

i) basic interface structure: 2 B-channels + 1 D-channel (16 kbit/s);

ii) primary rate B-channel interface structures:

– 1544 kbit/s primary rate: 23 B + 1 D (64 kbit/s);

– 2048 kbit/s primary rate: 30 B + 1 D (64 kbit/s).

b) *H-channel interface structures*

i) primary rate interface H_0-channel structures:

– 1544 kbit/s primary rate: 4 H_0;
3 H_0 + D.

– 2048 kbit/s primary rate: 5 H_0 + D.

For the following parameters, appropriate means have to be provided:

– clock synchronization between source coder and network;

– time synchronization (delay difference compensation) between several B-channels.

2.2 Permanent connections

2048 kbit/s or 1544 kbit/s signals are used. Framing is according to Recommendation G.704.

The following parameters need consideration:

- clock synchronization between source coder and network;
- octet sequence integrity within the same frame;
- time slot allocation.

3 Synchronization and frame alignment

3.1 Dial-up connections

For all 64 kbit/s channels of the basic interface structure in dial-up connections, the framing according to Recommendation H.221, in conjunction with Recommendation H.242, should be used. With the use of the H.221 framing there is a possibility to achieve both clock synchronization and time synchronization between up to six B-channels. This is permanently provided, also in case of a change of the routing during the transmission.

Furthermore, the addition of H.221 framing to dial-up connections (H_0 or 1 to 6 B-channels) provides a control channel harmonized with other N-ISDN audiovisual equipment.

The procedure for **initialization** of the communication at the start of a call shall be according to Recommendation H.242, except that, if there is no requirement for intercommunication with 3 kHz speech terminals, it is not necessary to transmit G.711 audio during the period until Sequence A is completed. The capabilaities to be transmitted during Sequence A include appropriate values from Table A.2/H.221, and a single command from that table introduces a mode switch to the required sound transmission mode.

Particular attention is drawn to clause 13/H.242 for the cases of transmission within a restricted network (multiple 56 kbit/s channels) and of interworking between unrestricted and restricted networks. Terminals on both types of network shall transmit within their capability sets the relevant values from attribute (110) of Table A.1/H.221.

3.1.1 Basic interface structure

To achieve the synchronization according to Recommendation H.221, an overhead [consisting of a Frame Alignment Signal (FAS) and Bit Rate Allocation Signal (BAS)] of 1.6 kbit/s is required per 64 kbit/s channel or in the first time slot of an H_0 channel.

For the transmission of bit rate reduced high-quality sound signals, Recommendation H.221 has been extended.

3.1.2 Realization of bit rates which are not explicitly listed in the bit rate table of the ISO/IEC 11172-3 Standard

Figure 3 describes the different sections of the transmission system. The network provides data channels each with 64 kbit/s. If H.221 framing and an additional error protection is applied, the full channel capacity cannot be used by the ISO 11172-3 encoder for audio data.

The ISO/IEC 11172-3 bit stream is sub-divided in audio frames corresponding to a sequence of 384 PCM audio samples for Layer I and 1152 PCM audio samples for Layers II and III. The audio frame length (see Table 1) is dependent on the bit rate index, the sampling frequency and the status of the padding bit, an information that is given in the audio frame header. The ISO/IEC 11172-3 Standard allows for 14 different, explicitly listed bit rates, indicated by the so-called bit rate index value. An additional format, the "free format", i.e. a user-defined audio frame length, can be chosen by using the bit rate index "0000".

Two methods may be used in order to realize those bit rates that are not explicitly listed in the bit rate tables of the ISO/IEC 11172-3 Standard. These methods are:

- using the ancillary data field (recommended for Layers I and II);
- dynamic bit rate switching (recommended for Layer III).

The use of free format is not recommended.

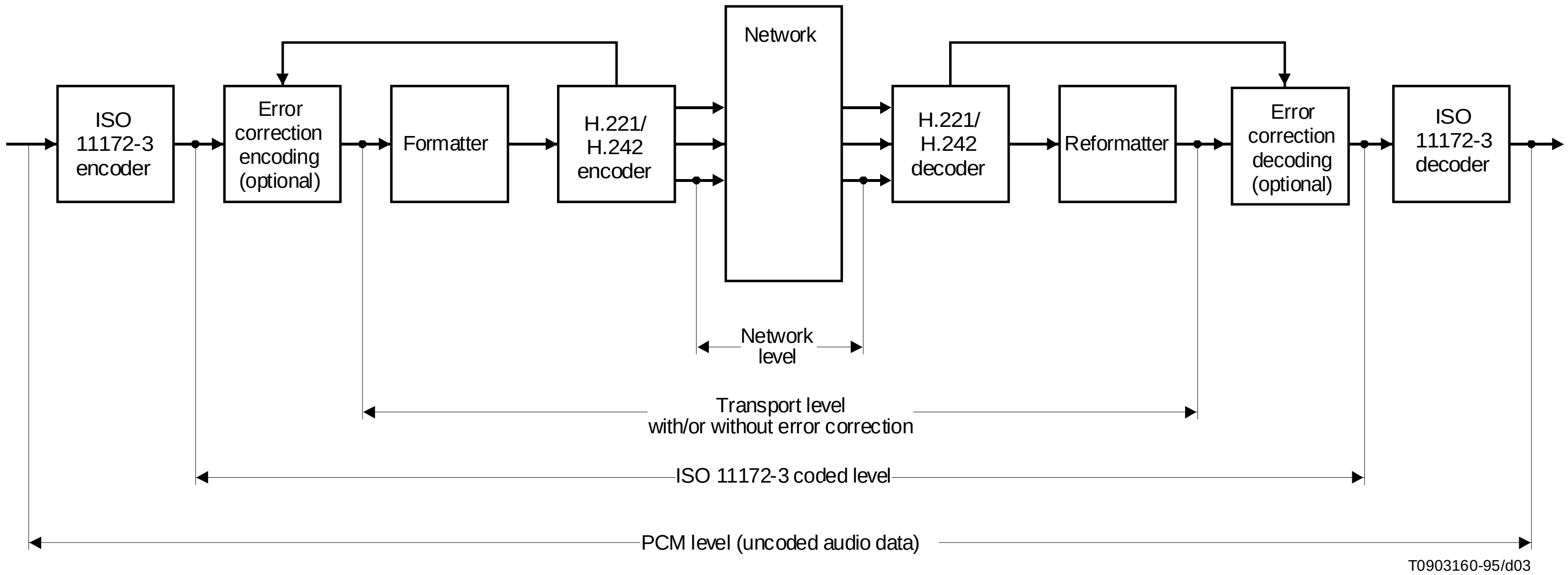

FIGURE 3/J.52

Transmission system with transport format and optional error control

3.1.2.1 Using the ancillary data field (recommended for Layers I and II)

The last field in a ISO/IEC 11172-3 frame is the "ancillary data" field, which consists of all bits that are not used for audio. A minimum length of this field can be guaranteed by constraining the bit allocation in the encoder. The desired bit rate can be obtained by selecting a bit rate higher than the desired bit rate, taking into account the constraining of the ISO/IEC Standard, and stripping off a number of bytes from each frame. The resulting frames are hereafter called "shortened frames". After transmission, the correct number of zero-valued bytes is inserted again, in order to reconstitute compliant MPEG audio frames, hereafter called "long frames".

In the following, an exact method is described to obtain the length of the short frame for each ISO/IEC 11172-3 frame, given a certain desired bit rate and sampling frequency. An ISO/IEC 11172-3 frame always consists of an integer number of slots, being one byte in Layer II and four bytes in Layer I. The number of slots in a short frame has to be calculated using the following formulae:

$$\text{Layer II: } N = 144 \frac{\text{desired bit rate}}{\text{sampling frequency}}$$

$$\text{Layer I: } N = 12 \frac{\text{desired bit rate}}{\text{sampling frequeny}}$$

If the calculation above yields a non-integer number, the result is truncated, and padding is required. In this case, the number of slots in a frame will vary between N and N + 1. The padding bit is set to "0" if the number of slots equals N, and otherwise to "1".

The padding should be applied to the bit stream such that the accumulated length of the coded frames, after a certain number of audio frames, does not deviate by more than (+0, –1 slot) from the following computed value, where *bit rate* is the desired bit rate:

$$\text{Accumulated frame length} = \sum_{\text{first frame}}^{\text{current frame}} \frac{(\text{frame size})(\text{bit rate})}{\text{sampling frequency}}$$

where frame size equals 384 for Layer I, and 1152 for Layer II.

The following method, described in C-syntax, can be used to determine whether or not to use padding:

for 1st audio frame:

```
rest = 0
padding = no;
```

for each subsequent audio frame:

```
if (Layer = 1) dif = (12 * bit rate) % sampling frequency;
else dif = (144 * bit rate) % sampling frequency;
rest = rest – dif;
if (rest < 0) {
    padding = yes;
    rest = rest + sampling frequency;
}
else padding = no;
```

If the padding bit is modified (P1→P2 in Figure 4), the ISO-CRC has to be recalculated.

Note that the resulting bit stream of short frames is very similar to an ISO/IEC 11172-3 free format bit stream. Exceptions are the value of the bit rate index (which is "0000" in free format mode) and the bit allocation tables for some combinations of sampling frequency and bit rate. It should be noted that in an integrated implementation of an MPEG audio encoder, formatter and H.221 encoder, it is possible to apply a special MPEG audio encoder that directly generates the short frames.

At the receiving end, the output of the H.221 decoder will be the short frames. In order to regenerate long frames, the following procedure has to be applied. The number of slots and the padding sequence of the long frames have to be calculated in the same way as previously described for the short frames, but now with the bit rate which corresponds to the bit rate index in the ISO/IEC 11172-3 header. The padding bit in the short frames has to be overwritten with the calculated padding bit value for the long frame (this value will always be zero for 32 and 48 kHz sampling frequency), and the short frame has to be filled up with zero-valued slots to the full length of the long frame.

If the padding bit has to be changed (P2→P3 in Figure 4), the following procedure has to be applied:

- if the ISO-CRC verification shows no errors, the ISO-CRC has to be recalculated, based upon the padding bit of the regenerated long frame;
- otherwise the ISO-CRC has to be transmitted as received.

It should be noted that in an integrated implementation of H.221 decoder, reformatter and ISO decoder, it might not be necessary to rebuild the long frames, but to apply a special MPEG audio decoder that works directly on the short frames. If ancillary data, such as Programme Associated Data (PAD), error control data or scale factor CRCs, have to be transmitted along with the audio data, the bit allocation of the encoder has to be further constrained as to allow insertion of this data in the short frames.

Figure 4 shows how not explicitly listed bit rates can be derived out of a compliant MPEG frame.

3.1.2.2 Dynamic bit rate switching (recommended for Layer III)

For Layer III the audio frame length can be changed dynamically from frame to frame. Using this method, on average additional bit rates that are not listed in the bit rate table of the ISO/IEC 11172-3 Standard are supported. The sequence of frames with different bit rates called switching period (P) which is necessary to realize the required bit rate has to be determined in the encoder. Hereby, the following formula has to be considered:

$$\text{Bit rate}_{/\text{kbit/s}} = \text{Average frame length}_{/\text{bits}} \times \text{Sampling frequency}_{/\text{kHz}}/1152$$

For example, for a bit rate of 62.4 kbit/s and a sampling frequency of 48 kHz, a sequence of one frame containing 168 bytes = 1344 bits (corresponding to a bit rate of 56 kbit/s) and four frames containing 192 bytes (corresponding to a bit rate of 64 kbit/s) has to be used. This sequence results in an average frame length of 187.2 bytes = 1497.6 bits. The switching period (P), the length of a short frame (L1) and a long frame (L2) is shown in Table 3. I1 and I2 are the number of frames with the length L1 or L2 respectively within a switching period P. In case of error control modes 2 and 3, the sum of the bits used for the H.221 coding and the error control has to be taken into account for the average bit rate.

If a long frame is used, some bits can be stored in the internal buffer and used for the next frame if necessary, to achieve a constant average bit rate for the encoding.

Example: Average bit rate 62.4 kbit/s, 48 kHz sampling frequency.

Bit rate index	Bits used for encoding	Bits put into/taken from the buffer
"64 000"	1497	39
"64 000"	1497	39
"64 000"	1498	38
"64 000"	1498	38
"56 000"	1498	–154

A nearly arbitrary switching sequence has to be taken into account, especially in asynchronous operating mode.

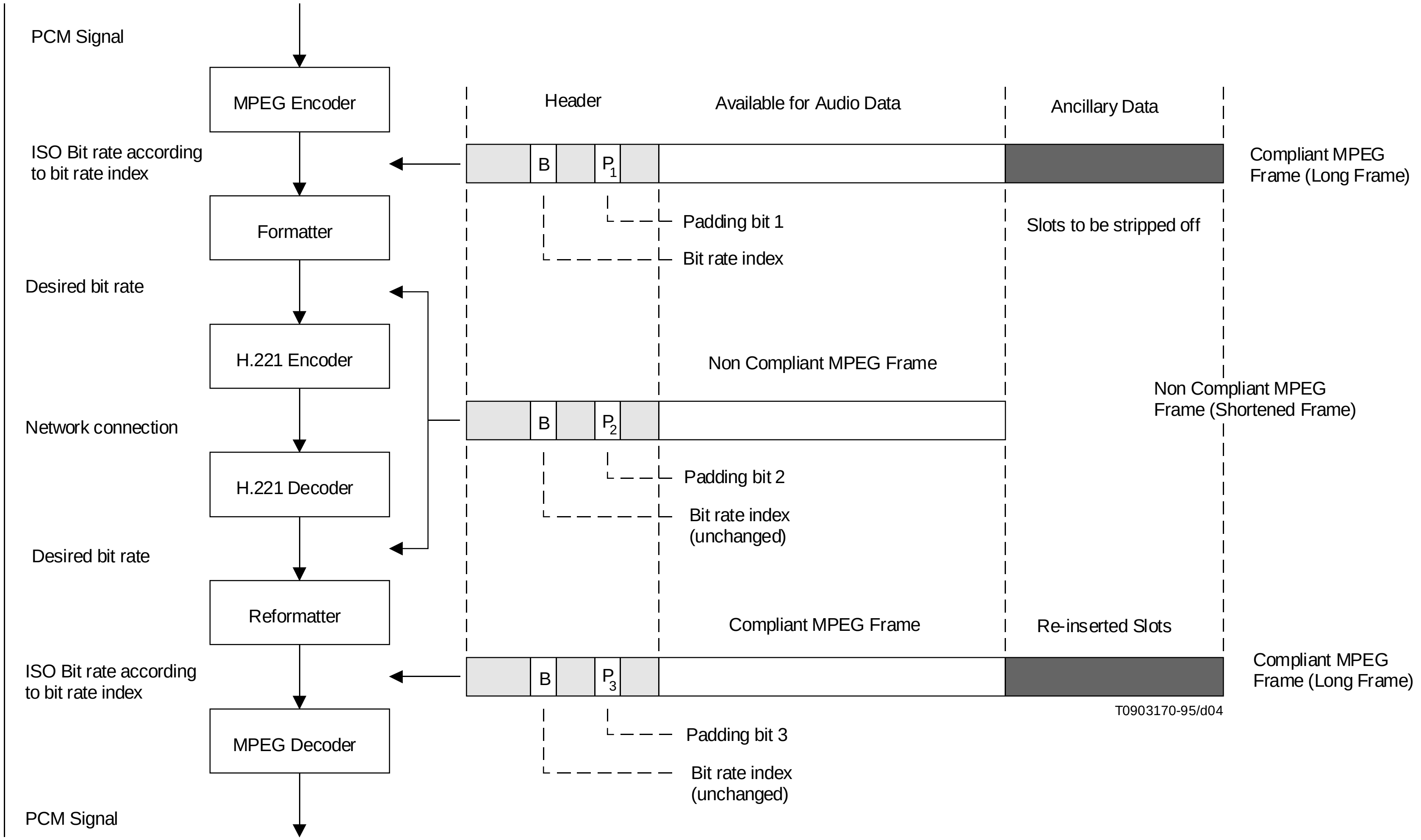

FIGURE 4/J.52

Ancillary data field method for realizing not explicitly listed bit rate (Mandatory for J.52)

3.1.2.3 Example

Tables 2 and 3 list details for an adaption of the explicitly listed bit rates according to ISO/IEC 11172-3 to the available channel bit rate for the methods "Using the ancillary data field" and "Dynamic bit rate switching". The contents of these tables are:

- *Channels*: Number of used 64 kbit/s channels.
- *Mode*: Inband — H.221 framing is used within each sub-channel.
  Inband 56k — H.221 framing is used within each 56 kbit/s channel.
- *Desired bit rate*: Data rate available for the ISO 11172-3 audio data.

The following values are given in Table 2:

L Length of long frame (as defined in ISO/IEC 11172-3);

S Length of a shortened frame.

TABLE 2/J.52

Adaption of the explicitly listed bit rates according to ISO/IEC 11172-3 to the desired bit rate (examples for Layer II)

Channels	Mode	Desired data rate	Sampling frequency					
			32 kHz		44.1 kHz		48 kHz	
			L	S	L	S	L	S
1	inband	62 400	192	187	208	203	288	280
2	inband	124 800	384	374	417	407	576	561
3	inband	187 200	576	561	626	611	864	842
4	inband	249 600	768	748	835	815	1152	1123
5	inband	312 000	960	936	1044	1018	1440	1404
6	inband	374 400	1152	1123	1253	1222	1728	1684
2	inband 56k	108 800	336	326	365	355	504	489

The following values are given in Table 3:

L1 Nominal frame length for a short frame (as defined in ISO/IEC 11172-3).

L2 Nominal frame length for a long frame (as defined in ISO/IEC 11172-3).

I1 Number of frames using L1.

I2 Number of frames using L2.

P Switching period of frame length.

dL Average frame length.

Table 3 is valid for a sampling frequency of 48 kHz.

3.2 Transmission on permanent connections

The structure of a 2048 kbit/s frame consists of 32 Time Slots (TS) with a capacity of 64 kbit/s each. The structure of a 1544 kbit/s frame consists of 24 TS, also with the capacity of 64 kbit/s per TS.

TABLE 3/J.52

Adaptation of the explicity listed bit rates according to ISO/IEC 11172-3 to the desired bit rate (examples for Layer III)

Channels	Mode	Desired data rate	L1 (byte)	I1	L2 (byte)	I2	P	dL
1	inband	62 400	168	1	192	4	5	187.2
2	inband	124 800	336	1	384	4	5	374.4
3	inband	187 200	480	3	576	17	20	561.6
4	inband	249 600	672	1	768	4	5	748.8
5	inband	312 000	768	1	960	7	8	936
2	inband 56	108 800	288	1	336	4	5	326.4

TS 0 is used for frame alignment. For 2 Mbit/s frames, TS 16 is reserved for signalling and other network purposes. In case of channel-associated signalling, transmission of low speed data channels (associated with sound channels) in TS 16 is possible.

In case of error control modes 1, 2 and 3, capacity for redundancy is necessary. This capacity either:

- is taken from the ancillary data field (refer to 3.1.2.1); or
- is provided by dynamic bit rate switching (refer to 3.1.2.2).

In 2 Mbit/s (or 1.5 Mbit/s) connections there are two operation modes, a flexible multiplex mode and a fixed multiplex mode. In the flexible multiplex mode only time slot TS 1 has the H.221 framing, all other TS without are framing and can transmit the full capacity of 64 kbit/s. In the fixed multiplex mode all TS can transmit the full capacity of 64 kbit/s without H.221 framing. The channels are allocated in a fixed order.

3.2.1 Flexible multiplex

In this flexible multiplex mode the Multiple Byte Extension (MBE) is used (see Recommendation H.221). TS 1 has a frame structure according to Recommendation H.221. It is not to be used for the transmission of sound signals, but contains FAS and BAS and can be used for the transmission of one G.722 channel and/or Low Speed Data (LSD). The audio signals are transmitted in the TS 2 to 15 and 17 to 31 for 32 TS systems, and in the TS 2 to 23 for 24 TS systems.

Each TS except TS 1 is an unframed 64 kbit/s channel with full capacity, which means it is possible to transmit:

- 64 kbit/s in one TS;
- 128 kbit/s in two TS;
- 192 kbit/s in three TS, etc.

In case of transmission of more than 64 kbit/s, two or more TS constitute a virtual channel.

In Recommendation H.221, the details for control and signalling purposes (capability exchange, command exchange) are indicated.

3.2.2 Fixed multiplex

The fixed multiplex mode is established via 2048 kbit/s frames according to Recommendation G.704 and in principle the time slots could be allocated individually.

In the case of transmission of more than 64 kbit/s, two or more time slots constitute a virtual channel.

If the frame is in accordance with Recommendation G.735 or Recommendation G.737, the following time slot allocations are recommended (see Tables 4-1 and 4-2):

TABLE 4-1/J.52

	A	B	C	D	E
1	1 – 17	4 – 20	7 – 23	10 – 26	13 – 29
2	2 – 18	5 – 21	8 – 24	11 – 27	14 – 30
3	3 – 19	6 – 22	9 – 25	12 – 28	15 – 31
NOTE – The fifteen possible 128 kbit/s channels in a 2048 kbit/s stream are numbered A1 to E3. Preferably the channel pairs A1-A2, A2-A3, A3-B1, ... E2-E3 should be used for stereophonic transmission, or as a 256 kbit/s channel.					

TABLE 4-2/J.52

	A	B	C	D	E
a	1 – 2 – 3	4 – 5 – 6	7 – 8 – 9	10 – 11 – 12	13 – 14 – 15
b	17 – 18 – 19	20 – 21 – 22	23 – 24 – 25	26 – 27 – 28	29 – 30 – 31
NOTE – The ten possible 192 kbit/s channels in a 2048 kbit/s stream are numbered Aa to Eb. Preferably the channel pairs Aa-Ab and Ba-Bb and Ca-Cb and Da-Db and Ea-Eb should be used for stereophonic transmission, or as a 384 kbit/s channel.					

If the frame is in accordance with Recommendation G.738, the following time slot allocations are recommended for establishing 320 kbit/s channels (see Table 4-3):

TABLE 4-3/J.52

A	B	C	D	E	F
1 – 2 – 3 – 4 – 5	6 – 7 – 8 – 9 – 10	11 – 12 – 13 – 14 – 15	17 – 18 – 19 – 20 – 21	22 – 23 – 24 – 25 – 26	27 – 28 – 29 – 30 – 31

3.3 Synchronization of the sampling frequencies to the clock frequency

In real-time transmission systems, proper synchronization of encoder and decoder is a very important topic. A synchronous or an asynchronous operating mode can be used.

3.3.1 Synchronous operating mode

Usually, the transmission system provides the master clock at both sides, i.e. the sampling clock at the input and output of the system is synchronized by the transmission clock (see Figure 5). Therefore, with digital audio inputs, either the audio source has to be synchronized to the transmission clock or a sample rate converter has to be used.

3.3.2 Asynchronous operating mode

The sampling clock of the audio input is asynchronous to the network clock. The codec has to adapt the ratio between sampling frequency and bit rate to the network clock.

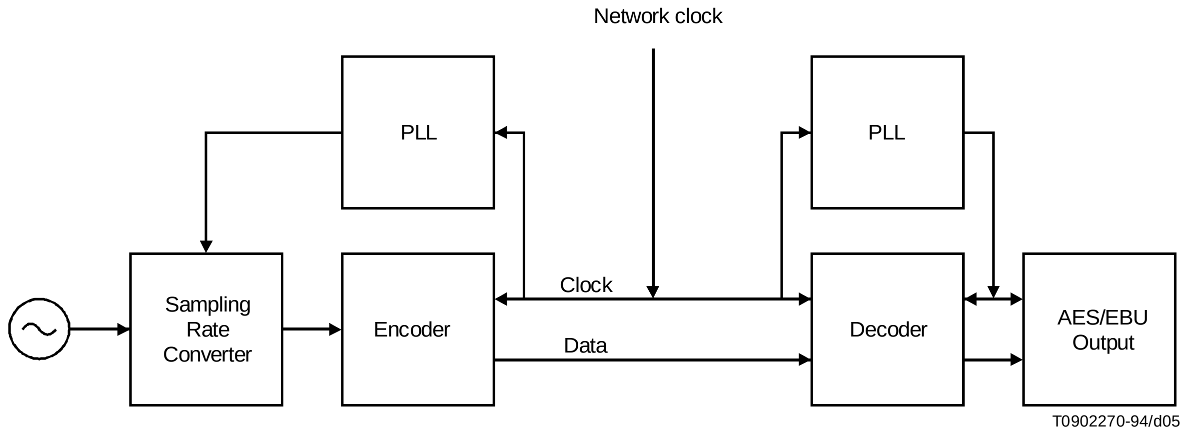

FIGURE 5/J.52

Synchronous operating mode

With Layer III, the use of any data rate is possible. Therefore, a completely asynchronous operation is feasible. The encoder counts the number of data bits that have been transmitted within a certain time interval and compares it with the target number (the nominal data rate, multiplied by the time interval). If the actual number of data bits is too small, the encoder has to use a frame with a smaller data rate for the next frame and vice versa. Using a similar approach, the decoder controls its sampling frequency to set it to the same value as in the encoder (see Figure 6).

If H.221 framing is applied, asynchronous operating mode is also possible for Layer II using padding (except the data rate 320 kbit/s).

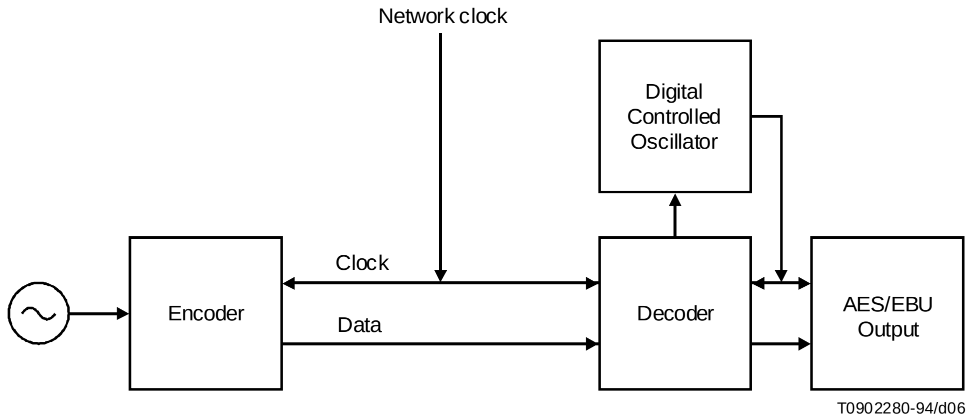

FIGURE 6/J.52

Asynchronous operating mode

4 Error control

The use of the ISO CRC in the source coded audio signal is mandatory for the transmission application. (The CRC bits allow error detection in the header and in the most important parts of the side information.)

If Recommendation H.221 is applied, the CRC-4 procedure has to be used. Both ISO CRC and CRC-4 can be utilized to detect errors, but not to correct them. If necessary, error correction can be applied in addition.

The following four error control modes are provided:

- Mode 0: only ISO CRC
- Mode 1: unequal error control typical redundancy $r \approx 1.0\%$
- Mode 2: low equal error control typical redundancy $r \approx 2.5\%$
- Mode 3: high equal error control typical redundancy $r \approx 10\%$

The error control is performed by:

- error correction by a Reed-Solomon code;
- error concealment with the CRC-16.

If the Reed-Solomon code (RS code) is overloaded, then the error concealment based on the ISO CRC-16 should be applied. The error control modes 1, 2, 3 use the same coding scheme.

- Symbol length: $m = 8$ bits (1 byte)
- Codeword length: N byte (variable)
- Code dimension: $K = N - 4$ bytes

  Field generator polynomial: $f(x) = x^8 + x^4 + x^3 + x^2 + 1$
- Code generator polynomial:

$$g(x) = \prod_{i=1}^{4} (x + \alpha^{125+i})$$

$$= x^4 + \alpha^{201}x^3 + \alpha^{246}x^2 + \alpha^{201}x + 1$$

A data byte (d_7, d_6, ..., d_1, d_0) is identified with the element:

$$d_7\alpha^7 + d_6\alpha^6 + ... + d_1\alpha + d_0 \text{ in GF(256).}$$

For the calculation of RS-code and transmission of redundancy, the most significant byte and the most significant bit of a byte (d_7) are the first transmitted.

4.1 Unequal error correction (only Layer II)

In any frame, error correction applies to:

- bits 16...31 of the header;
- bits of CRC check;
- bits of bit allocation;
- the maximum number of bits in:
  - scalefactor select information;
  - scalefactors.

Other parts of the signal are not protected.

Single channel signals always use one Reed-Solomon codeword per frame. Stereo and dual channel signals at low bit rates use one and at high bit rates two Reed-Solomon codewords per frame (see Tables 5, 6 and 7).

If there are two codewords, they are interleaved byte-by-byte according to the following scheme: The first transmitted byte is the first byte of codeword one, the second transmitted byte is the first byte of codeword two, the third transmitted byte is the second byte of codeword one, and so on. This means that the information itself is not interleaved.

Redundancy of code $$r_{code} = \frac{4}{N} 100\ (\%)$$

Redundancy per frame $$r_{frame} = 4 \frac{\text{codewords per frame}}{\text{byte per frame}} 100\ (\%)$$

TABLE 5/J.52

Code parameters for a sampling frequency of 48 kHz
(unequal error-control mode – mode 1 – for Layer II only)

Bit rate (kbit/s)	Bytes per frame	Single channel				Stereo or dual channel			
		No. of codewords	K (bytes)	r_{code} (%)	r_{frame} (%)	No. of codewords	K (bytes)	r_{code} (%)	r_{frame} (%)
32	96	1	28	12.5	4.2	–	–	–	–
48	144	1	28	12.5	2.8	–	–	–	–
56	168	1	83	4.6	2.4	–	–	–	–
64	192	1	83	4.6	2.1	1	51	7.3	2.1
80	240	1	83	4.6	1.7	–	–	–	–
96	288	1	83	4.6	1.4	1	51	7.3	1.4
112	336	1	83	4.6	1.2	2	81	4.7	2.4
128	384	1	83	4.6	1.0	2	81	4.7	2.1
160	480	1	83	4.6	0.8	2	81	4.7	1.7
192	576	1	83	4.6	0.7	2	81	4.7	1.4
224	672	–	–	–	–	2	81	4.7	1.2
256	768	–	–	–	–	2	81	4.7	1.0
320	960	–	–	–	–	2	81	4.7	0.8
384	1152	–	–	–	–	2	81	4.7	0.7

TABLE 6/J.52

Code parameters for a sampling frequency of 44.1 kHz
(unequal error-control mode – mode 1 – for Layer II only)

Bit rate (kbit/s)	Bytes per frame	Single channel				Stereo or dual channel			
		No. of codewords	K (bytes)	r_{code} (%)	r_{frame} (%)	No. of codewords	K (bytes)	r_{code} (%)	r_{frame} (%)
32	104.5	1	28	12.5	3.8	–	–	–	–
48	156.7	1	28	12.5	2.6	–	–	–	–
56	182.9	1	83	4.6	2.2	–	–	–	–
64	209.0	1	83	4.6	1.9	1	51	7.3	1.9
80	261.2	1	83	4.6	1.5	–	–	–	–
96	313.5	1	91	4.2	1.3	1	51	7.3	1.3
112	365.7	1	91	4.2	1.1	2	81	4.7	2.2
128	418.0	1	91	4.2	1.0	2	81	4.7	1.9
160	522.4	1	91	4.2	0.8	2	81	4.7	1.5
192	626.9	1	91	4.2	0.6	2	89	4.3	1.3
224	731.4	–	–	–	–	2	89	4.3	1.1
256	835.9	–	–	–	–	2	89	4.3	1.0
320	1044.9	–	–	–	–	2	89	4.3	0.8
384	1253.9	–	–	–	–	2	89	4.3	0.6

TABLE 7/J.52

Code parameters for a sampling frequency of 32 kHz
(unequal error-control mode – mode 1 – for Layer II only)

Bit rate (kbit/s)	Bytes per frame	Single channel				Stereo or dual channel			
		No. of codewords	K (bytes)	r_{code} (%)	r_{frame} (%)	No. of codewords	K (bytes)	r_{code} (%)	r_{frame} (%)
32	144	1	39	9.3	2.8	–	–	–	–
48	216	1	39	9.3	1.9	–	–	–	–
56	252	1	83	4.6	1.6	–	–	–	–
64	288	1	83	4.6	1.4	1	74	5.1	1.4
80	360	1	83	4.6	1.1	–	–	–	–
96	432	1	91	4.2	0.9	1	74	5.1	0.9
112	504	1	91	4.2	0.8	2	81	4.7	1.6
128	576	1	91	4.2	0.7	2	81	4.7	1.4
160	720	1	91	4.2	0.6	2	81	4.7	1.1
192	864	1	91	4.2	0.5	2	89	4.3	0.9
224	1008	–	–	–	–	2	89	4.3	0.8
256	1152	–	–	–	–	2	89	4.3	0.7
320	1440	–	–	–	–	2	89	4.3	0.6
384	1728	–	–	–	–	2	89	4.3	0.5

4.2 Equal error correction (only Layers II and III)

The RS code defined above is used for protecting a complete frame. The more powerful error control of mode 3 is achieved by using more and shorter codewords per frame. The L codewords per frame are interleaved, in order to increase the error correction capability in case of burst errors. By this simple coding scheme only one encoder implementation, one decoder implementation and a flexible interleaver is required for all applications considered in this subclause. See Figure 7 for examples.

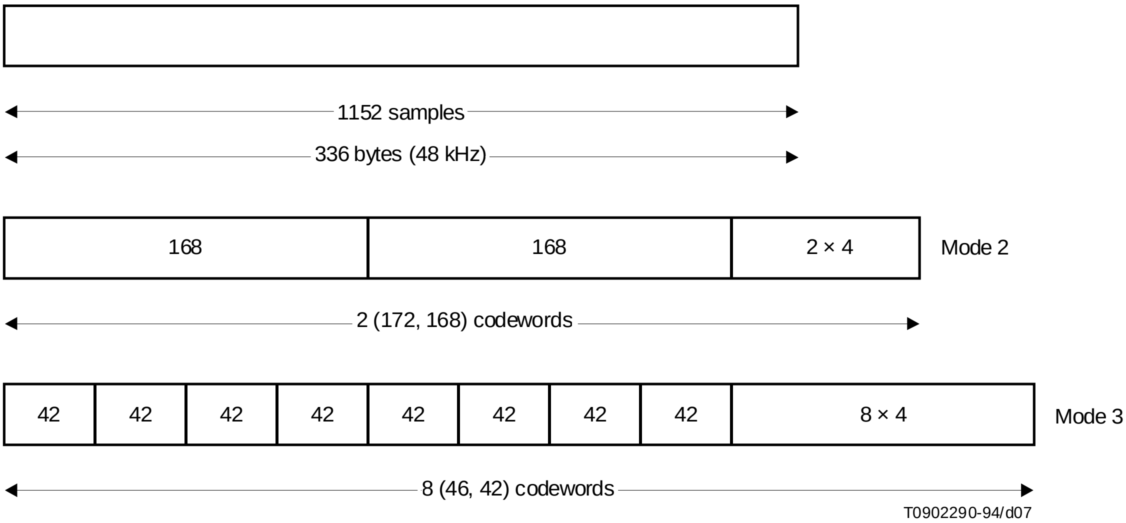

FIGURE 7/J.52

Examples for the coding scheme before interleaving

For the sampling frequencies 48 kHz, 44.1 kHz and 32 kHz and for different bit rates, the code length N and the number L of codewords per frame are given in Tables 8, 9 and 10.

In many cases two different code lengths N and N – 1 are necessary in each frame. The number of codewords of length N is denoted by L_N and the number of codewords of length N – 1 is denoted by L_{N-1}, where $L = L_N + L_{N-1}$. The redundancy is:

$$r = \frac{4L}{\text{Number of bytes per frame}}$$

For Layer III, the RS code is calculated based on the length of the ISO-frame taking into account the padding bit. If the padding bit is set, the following modifications occur:

for $L_{N-1} > 0$: $\quad L_N = L_N + 1 \quad L_{N-1} = L_{N-1} - 1$,

or for $L_{N-1} = 0$: $\quad N = N + 1 \quad L_N = 1 \quad L_{N-1} = L - 1$

For details of the performance, see Appendix I.

TABLE 8/J.52

Code parameters for a sampling frequency of 48 kHz

Bit rate (kbit/s)	Bytes per frame	N (Mode 2)	No. of codewords L	L_N	L_{N-1}	r (%)	N (Mode 3)	No. of codewords L	L_N	L_{N-1}	r (%)
32	96	100	1	1	0	4.17	52	2	2	0	8.33
40	120	124	1	1	0	3.33	44	3	3	0	10.00
48	144	148	1	1	0	2.78	52	3	3	0	8.33
56	168	172	1	1	0	2.38	46	4	4	0	9.52
64	192	196	1	1	0	2.08	43	5	2	3	10.42
80	240	244	1	1	0	1.67	44	6	6	0	10.00
96	288	148	2	2	0	2.78	46	7	1	6	9.72
112	336	172	2	2	0	2.38	46	8	8	0	9.52
128	384	196	2	2	0	2.08	47	9	6	3	9.38
160	480	164	3	3	0	2.50	44	12	12	0	10.00
192	576	196	3	3	0	2.08	46	14	2	12	9.72
224	672	172	4	4	0	2.38	44	17	9	8	10.12
256	768	158	5	3	2	2.60	45	19	8	11	9.09
320	960	164	6	6	0	2.50	44	24	24	0	10.00
384	1152	169	7	4	3	2.43	44	29	21	8	10.07
		Mode 2					Mode 3				

TABLE 9/J.52

Code parameters for a sampling frequency of 44.1 kHz

Bit rate (kbit/s)	Bytes per frame	N (Mode 2)	No. of codewords L	L_N	L_{N-1}	r (%)	N (Mode 3)	No. of codewords L	L_N	L_{N-1}	r (%)
32	104.4	108	1	1	0	3.83	56	2	2	0	7.66
40	130.6	134	1	1	0	3.06	48	3	1	2	9.19
48	156.7	160	1	1	0	2.55	43	4	4	0	10.21
56	182.8	186	1	1	0	2.19	50	4	2	2	8.75
64	208.9	212	1	1	0	1.91	46	5	3	2	9.57
80	261.2	135	2	1	1	3.06	48	6	3	3	9.19
96	313.4	161	2	1	1	2.55	44	8	1	7	10.21
112	365.7	187	2	1	1	2.19	45	9	5	4	9.84
128	417.9	213	2	1	1	1.91	46	10	7	3	9.57
160	522.4	178	3	1	0	2.30	45	13	2	11	9.95
192	626.9	161	4	2	2	2.55	46	15	11	4	9.57
224	731.4	187	4	3	1	2.19	45	18	11	7	9.84
256	835.9	171	5	5	0	2.39	44	21	16	5	10.05
320	1044.8	178	6	6	0	2.30	45	26	4	22	9.95
384	1253.8	161	8	5	3	2.55	45	31	13	18	9.89
		Mode 2					Mode 3				

TABLE 10/J.52

Code parameters for a sampling frequency of 32 kHz

Bit rate (kbit/s)	Bytes per frame	N	No. of codewords			r (%)	N	No. of codewords			r (%)
			L	L_N	L_{N-1}			L	L_N	L_{N-1}	
32	144	148	1	1	0	2.78	52	3	3	0	8.33
40	180	184	1	1	0	2.22	49	4	4	0	8.89
48	216	220	1	1	0	1.85	48	5	1	4	9.26
56	252	130	2	2	0	3.17	46	6	6	0	9.52
64	288	148	2	2	0	2.78	46	7	1	6	9.72
80	360	184	2	2	0	2.22	44	9	9	0	10.00
96	432	148	3	3	0	2.78	44	11	3	8	10.19
112	504	172	3	3	0	2.38	46	12	12	0	9.52
128	576	196	3	3	0	2.08	46	14	2	12	9.72
160	720	184	4	4	0	2.22	44	18	18	0	10.00
192	864	177	5	4	1	2.31	46	21	3	18	9.72
224	1008	172	6	6	0	2.38	45	25	8	17	9.92
256	1152	169	7	4	3	2.43	44	29	21	8	10.07
320	1440	164	9	9	0	2.50	44	36	36	0	10.00
384	1728	162	11	1	10	2.55	45	43	8	35	9.95
		Mode 2					Mode 3				

4.2.1 Computation of RS code for Layer II, by using the ancillary data field method

In this case some bit slots need to be reserved within the ISO frames capacity according to 3.1.2.1. See Figure 8.

R_h is the integer number of bytes to be stripped off for H.221 purposes and R_{rs} is the integer number of bytes to be reserved for RS computation according to Tables 5, 6, 7, 8, 9 and 10. If H.221 framing is not used, $R_h = 0$.

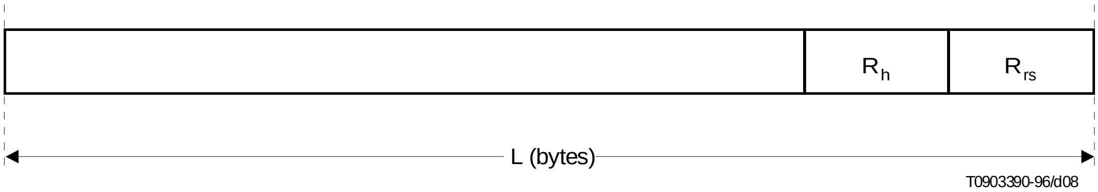

FIGURE 8/J.52

Audio frame with reserved slots for J.52 and FEC

RS code is computed according to the following procedure:

– R_h and R_{rs} fields have to be set to zero;

– Redundancy code is computed according to clause 4 for the words defined in Tables 5, 6, 7, 8, 9 and 10 on the whole frames including L bytes as defined in Table 2. L is the Length of the ISO-frame according to Table 2.

Due to the differences between audio and H.221 frames' lengths, the padding bit is set from time to time according to the description made in clause 3.1.2.1.

In case of setting the padding bit in the frame, in order to keep the parameters of the RS FEC equal to the parameters of frames not using the padding bit, the $(R_{rs} + R_h) - 1$ slot (1 slot = 1 byte for Layer II and 4 bytes for Layer I) is set to zero at the end of the frame and the FEC on the whole frame (length = L bytes according to Table 2) is computed.

In this case the extra slot belonging to the audio information is also protected.

4.2.2 Interleaving

Block interleaving is proposed, where the first transmitted byte is the first byte of the first codeword, the second transmitted byte is the first byte of the second codeword, etc. The interleaving scheme is shown in Figure 9.

Interleaving is accomplished in a way that the $(i + k \cdot L)^{th}$ byte of each frame is the $(k + 1)^{th}$ unaltered information byte of the i^{th} codeword with the following conventions:

- i: i^{th} codeword ($i = 1, 2, ... L$).
- k: k^{th} bit of the codeword ($k = 0, 1, ... N - 5$).
- L: Number of codewords.
- N: Length of the codeword.

The encoding of each frame starts with encoding of the L_N codewords of length N byte.

The effect of this block interleaving scheme and the systematic encoding is that the order and the values of the transmitted information bytes are not affected by the interleaving process and the encoding process, respectively. The additional 4 L redundant bytes, derived in the RS-encoder, are transmitted at the end of each frame.

Examples for this interleaving scheme are given in Figure 9.

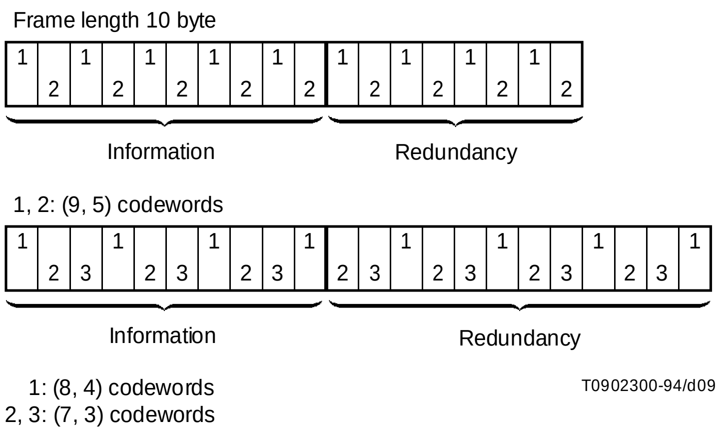

FIGURE 9/J.52

Interleaving scheme

4.3 Position of redundancy

In all cases, redundancy for error correction is inserted just before the header and corresponds to the following frame. Using the dynamic bit rate switching method, the length of the redundancy, located before the corresponding ISO-frame may vary from frame to frame. Only two different lengths are possible. To establish a proper synchronization both possible locations of the ISO syncword have to be tested. The decision is then made according to which length is proper for the following frame.

5 Transmission of data

The ancillary data field is defined by International Standard ISO/IEC 11172-3 to be the part of the frame, which is neither header nor audio data. For layers I and II the data field is at the end of the ISO frame, while for layer III the position is determined by pointers in the ISO-frame Side Information part. The International Standard does **not** define any structure or recommended use of the ancillary data field, but mutually incompatible application specific or proprietary formats exists.

This Recommendation defines the structure for the ancillary data field. The use of this structure is optional. The decoder can detect whether this structure is in use by checking the pattern of the identification bit at the end of the ancillary data field (for further details see 5.2).

The data format defined here is intended both for use with or without H.221 framing and/or FEC. The format supports a byte oriented transparent data channel, transmission of scalefactor CRCs, and other data.

5.1 Types of data

5.1.1 Byte oriented transparent channel

The transparent data channel allows for a byte oriented transfer of 8 bit characters, MSB first. The data bytes are put in reverse order in the bit stream, i.e. the last data byte in a frame corresponds to the first byte on the data interface. Applications should be tolerant to data loss.

5.1.2 Scalefactor CRC

For detection of errors within the three MSB's of the scalefactors, CRC check words can optionally be inserted in the J.52 Layer I or Layer II frame. The CRC check words cover the scalefactors of the following subbands:

– scfcrc0: subbands 0 ... 3 (subband group 0);

– scfcrc1: subbands 4 ... 7 (subband group 1);

– scfcrc2: subbands 8 ... 15 (subband group 2);

– scfcrc3: subbands 16 and above (subband group 3).

As the maximum number of subbands that can be allocated depends on the layer, sampling frequency and the bit rate per channel, the subband groups 2 and 3 can in some cases not have an allocation. Consequently, for those subband groups no scalefactor CRCs are contained in the bit stream:

Layer I:

for all combinations of sampling frequency and bit rate: 4 scfcrcs;

Layer II:

sblimit = 8: 2 scfcrcs (fs = 44.1 and 48 kHz, bit rate 32 and 48 kbit/s per channel);

sblimit = 12: 3 scfcrcs (fs = 32 kHz, bit rate 32 and 48 kbit/s per channel);

sblimit = 27 or 29 4 scfcrcs (fs = 32, 44.1 and 48 kHz, bit rate 56 kbit/s per channel or higher and free format).

The error detection method used is "CRC-8", whose generator polynomial is:

$$G_2(x) = x^8 + x^4 + x^3 + x^2 + 1$$

The bits included in the CRC-check are the three MSBs of all scalefactors of the corresponding subband group, in the order of their occurrence in the bit stream.

The initial state of the scalefactor crc-register is binary "0000 0000", i.e. all zeroes (note that the initial state for the crc-register defined by ISO/MPEG is binary "1111 1111 1111 1111" or hex $FF FF. The ISO/MPEG CRC protects the bit allocation and scalefactor select information).

The method for the calculation of the CRC is the same as for the ISO CRC, and can be found in ISO/IEC 11172-3.

The scfcrcs are put in reverse order in the bit stream to keep the position of scfcrc0 and scfcrc1 independent from the sampling frequency and bit rate.

The scfcrcs apply to the scalefactors in the following ISO frame. See Table 11.

TABLE 11/J.52

Number of scfcrcs	Order in bit stream
2	scfcrc1, scfcrc0
3	scfcrc2, scfcrc1, scfcrc0
4	scfcrc3, scfcrc2, scfcrc1, scfcrc0

5.1.3 Presentation time stamp

Transmission of Presentation Time Stamps (PTSs) in each audio frame is possible by using a 5 byte data block in the ancillary data field. The 33 least significant bits represent the PTS, and the seven most significant bits are set to zero. The MSB is transmitted first. The PTS has a resolution of 90 kHz, and represents the time of the first sample in the following audio frame. For further details see 2.4.3.3 and 2.4.4.3 of ISO/IEC 11172-1.

The transmission of PTS is optional on a frame-to-frame basis. The PTS can be recalculated by the decoder to other time code formats.

5.2 J.52 data format

The data format is self-identifying to allow easy coexistence with other pre-existing ancillary data formats, such as e.g. the PAD formats defined by Eureka 147 DAB or ADR. See Figure 10.

If equal Error Protection is in use, the data channel is protected together with the rest of the ISO frame (or shortened frame in case H.221 framing is being used).

Available for audio data	Extension data	Extension header	D_{L-1} … Data … D_0	Data header	X
		1 byte (optional)	L bytes	1 ... 2 bytes	

For Layer I and II, X denotes the header of the next ISO frame, or, in case FEC is used, the beginning of the FEC redundancy field. For Layer III, X denotes the beginning of the main data of the next ISO frame. D_0 is the first byte and D_{L-1} the last byte received by the data interface.

FIGURE 10/J.52

J.52 ancillary data format

The data header is defined as follows for all layers as shown in Figure 11.

15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
L								L5	L4	L3	L2	L1	L0	ext	ID
Optional (exists if L5...L0 = "11 1111"								Length indication							

ID Identification, a serial bit channel. In subsequent frames, the pattern "00 0001" is repeated.

ext extension, indicating whether more data than the transparent data channel is contained in the bit stream. Equals "0" if no extension data is present, "1" if extension data is present.

L5 ... L0 Length indication, MSB first (number of data bytes). If L5 ... L0 = "11 1111", then there is an additional length byte (MSB first) which can indicate a length from 0 ... 255.

If the "ext" bit in the data header equals "1", an extension header with the format shown in Figure 12 precedes the transparent data bytes.

FIGURE 11/J.52

J.52 Data header

	7	6	5	4	3	2	1	0
	Pr	Fu5	Fu4	Fu3	Fu2	Fu1	PT	SC

SC To indicate whether Scalefactor CRCs are contained in the bit stream. Equals "0" if no Scalefactor CRCs are present, "1" if Scalefactor CRCs are present.

PT To indicate whether presentation time stamp data are contained in the bit stream. Equals "0" if no PTS data are present, "1" if PTS data are present.

Fu1 ... Fu5 Bits reserved for future use, not defined yet. Zero at present.

Pr Bit for private use.

FIGURE 12/J.52

J.52 Data extension header

The extension data field has the format shown in Figure 13.

Presentation Time Stamp	Scalefactor CRCs	extension header
5 bytes (optional)	2 … 4 bytes (optional)	

FIGURE 13/J.52

J.52 Extensión data

Annex A

Digital transmission of medium-quality sound-programme signals using a 64 kbit/s channel (or a part thereof)

(This annex forms an integral part of this Recommendation)

A.1 Introduction

The International Standard ISO/IEC 13818-3 contains an extension of ISO/IEC 11172-3 to lower sampling frequencies. This extension provides an improved audio quality at the lowest bit rates, e.g. below 64 kbit/s per channel, at sampling frequencies which are half the sampling frequencies of ISO/IEC 11172-3 (16, 22.05 and 24 kHz).

Although encoding to this part of ISO/IEC 13818-3 is not yet (end 1995) recommended by ITU-R Study Group 10, it is recognized that there is a demand from the market to transmit such signals over the transmission channels addressed in this Recommendation, and that this only requires minor modifications to this Recommendation. These adaptations are given in this annex. The implementation of the lower sample frequencies as such is not mandatory, but, when decided, must conform to this specification.

A.2 Possible bit rates

See Table A.1.

TABLE A.1/J.52

Bit rate index	Bit rate specified for fs = 16, 22.05, 24 kHz (kbit/s)	
	Layer I	Layer II, Layer III
"0000"	free	free
"0001"	32	8
"0010"	48	16
"0011"	56	24
"0100"	64	32
"0101"	80	40
"0110"	96	48
"0111"	112	56
"1000"	128	64
"1001"	144	80
"1010"	160	96
"1011"	176	112
"1100"	192	128
"1101"	224	144
"1110"	256	160
"1111"	forbidden	forbidden

A.3 Framelength

For Layers I and II the framelength and padding sequence for the lower sampling frequencies (16, 22.05, 24 kHz) are to be calculated as in ISO/IEC 11172-3:

$$\text{Layer I: framelength} = 12\ \frac{\text{Bit rate}}{\text{Sampling frequency}}\ \text{slots} \qquad \text{(one slot equals 4 bytes)}$$

$$\text{Layer II: framelength} = 144\ \frac{\text{Bit rate}}{\text{Sampling frequency}}\ \text{slots} \qquad \text{(one slot equals one byte)}$$

For Layer III, the framelength is to be calculated by:

$$\text{Layer III: framelength} = 144\ \frac{\text{Bit rate}}{2\ (\text{Sampling frequency})}\ \text{slots} \qquad \text{(one slot equals one byte)}$$

A.4 Realization of bit rates which are not explicitly listed in the bit rate table of the ISO/IEC 13818-3 Standard

Adaptation of the explicitly listed bit rates accordingly to ISO/IEC 13818-3 Layer II to the available channel bit rate is done by using the ancillary data field method, as with the higher sampling frequencies. L and S are the length in bytes of a long frame and a short frame respectively, both without padding slot. See Table A.2.

TABLE A.2/J.52

Adaption of the explicity listed bit rates according to ISO/IEC 11172-3 to the desired bit rate (examples for Layer II)

Channnels	Mode	Useful data rate	Sampling frequency					
			16 kHz		22,05 kHz		24 kHz	
			L	S	L	S	L	S
1	in-band	62 400	576	561	417	407	384	374
2	in-band	124 800	1152	1123	835	815	768	748
2	in-band 56k	108 800	1008	979	731	710	672	652

Adaptation of the explicitly listed bit rates according to ISO/IEC 13818-3 Layer III to the available channel bit rate is done by using the dynamic bit rate switching method, as with the higher sampling frequencies. The tables for the sampling frequencies 32, 44.1 and 48 kHz in this Recommendation can also be used for the corresponding half sampling frequencies.

A.5 Error control

For the lower sampling frequencies, Mode 1 (unequal error control) is currently not defined. Modes 2 and 3 (equal error control) are only defined for Layers II and III. The method is the same as for the higher sampling frequencies. Tables A.3, A.4 and A.5 apply for Layer II.

TABLE A.3/J.52

Code parameters for a sampling frequency of 16 kHz (for Layer II)

Bit rate (kbit/s)	Bytes per frame	N	No. of codewords			r (%)	N	No. of codewords			r (%)
			L	L_N	L_{N-1}			L	L_N	L_{N-1}	
32	288	148	2	2	0	2.78	46	7	1	6	9.72
40	360	184	2	2	0	2.22	44	9	9	0	10.00
48	432	148	3	3	0	2.78	48	10	2	8	9.26
56	504	172	3	3	0	2.38	46	12	12	0	9.52
64	576	148	4	4	0	2.78	46	14	2	12	9.72
80	720	148	5	5	0	2.78	44	18	18	0	10.00
96	864	148	6	6	0	2.78	46	21	3	18	9.72
112	1008	148	7	7	0	2.78	45	25	8	17	9.92
128	1152	148	8	8	0	2.78	46	28		28	9.72

TABLE A.4/J.52

Code parameters for a sampling frequency of 22.05 kHz (for Layer II)

Bit rate (kbit/s)	Bytes per frame	N	No. of codewords			r (%)	N	No. of codewords			r (%)
			L	L_N	L_{N-1}			L	L_N	L_{N-1}	
32	208	212	1	1	0	1.92	46	5	3	2	9.62
40	261	135	2	1	1	3.07	48	6	3	3	9.20
48	313	161	2	1	1	2.56	49	7	5	2	8.95
56	365	126	3	2	1	3.29	45	9	5	4	9.86
64	417	143	3	3	0	2.88	46	10	7	3	9.59
80	522	178	3	3	0	2.30	45	13	2	11	9.96
96	626	161	4	2	2	2.56	46	15	11	4	9.58
112	731	151	5	1	4	2.74	45	18	11	7	9.85
128	835	144	6	1	5	2.87	46	20	15	5	9.58

TABLE A.5/J.52

Code parameters for a sampling frequency of 24 kHz (for Layer II)

Bit rate (kbit/s)	Bytes per frame	N	No. of codewords			r (%)	N	No. of codewords			r (%)
			L	L_N	L_{N-1}			L	L_N	L_{N-1}	
32	192	196	1	1	0	2.08	52	4	4	0	8.33
40	240	124	2	2	0	3.33	44	6	6	0	10.00
48	288	148	2	2	0	2.78	46	7	1	6	9.72
56	336	172	2	2	0	2.38	46	8	8	0	9.52
64	384	196	2	2	0	2.08	47	9	6	3	9.38
80	480	164	3	3	0	2.50	44	12	12	0	10.00
96	576	148	4	4	0	2.78	46	14	2	12	9.72
112	672	139	5	2	3	2.98	46	16	16	0	9.52
128	768	158	5	3	2	2.60	45	19	8	11	9.90

For Layer III, Tables 8, 9 and 10 for the sampling frequencies 32, 44.1 and 48 kHz in this Recommendation shall also be used for the corresponding half sampling frequencies.

A.6 Transmission of data

The transmission of data is the same as for the higher sampling frequencies, except for the number of scalefactor CRCs (only Layers I and II). In the case of the lower sampling frequencies, there are 4 scalefactor CRCs, for all combinations of sampling frequency and bit rate.

Appendix I

Performance of the forward error correction

(This appendix does not form an integral part of this Recommendation)

Within each codeword, up to two ($t < 2$) arbitrarily distributed byte errors can be corrected. The burst error correction capability b within each frame depends on the number L of codewords and the amount t of error correction:

$$b \leq [(Lt - 1)\ 8 + 1] \text{ bits} \qquad \text{for } t > 0$$

In Figure I.1, the bit-error probability P_b after decoding is shown as a function of the number t of corrected symbol errors, the code length N, and the bit-error probability BER of the transmission channel, where statistically independent bit errors are assumed. For burst errors with the same BER within a frame, the performance is much better than shown in Figure I.1.

The probability P_f of a falsely decoded word is shown in Figure I.2. The probability P_F of a falsely decoded frame (containing L codewords) is given by:

$$P_F = 1 - (1 - P_f)^L$$

It must be mentioned that even for $t = 2$, a high amount of uncorrectable error patterns can be detected by the RS-decoder. The reliability of error detection can further be increased by reducing t.

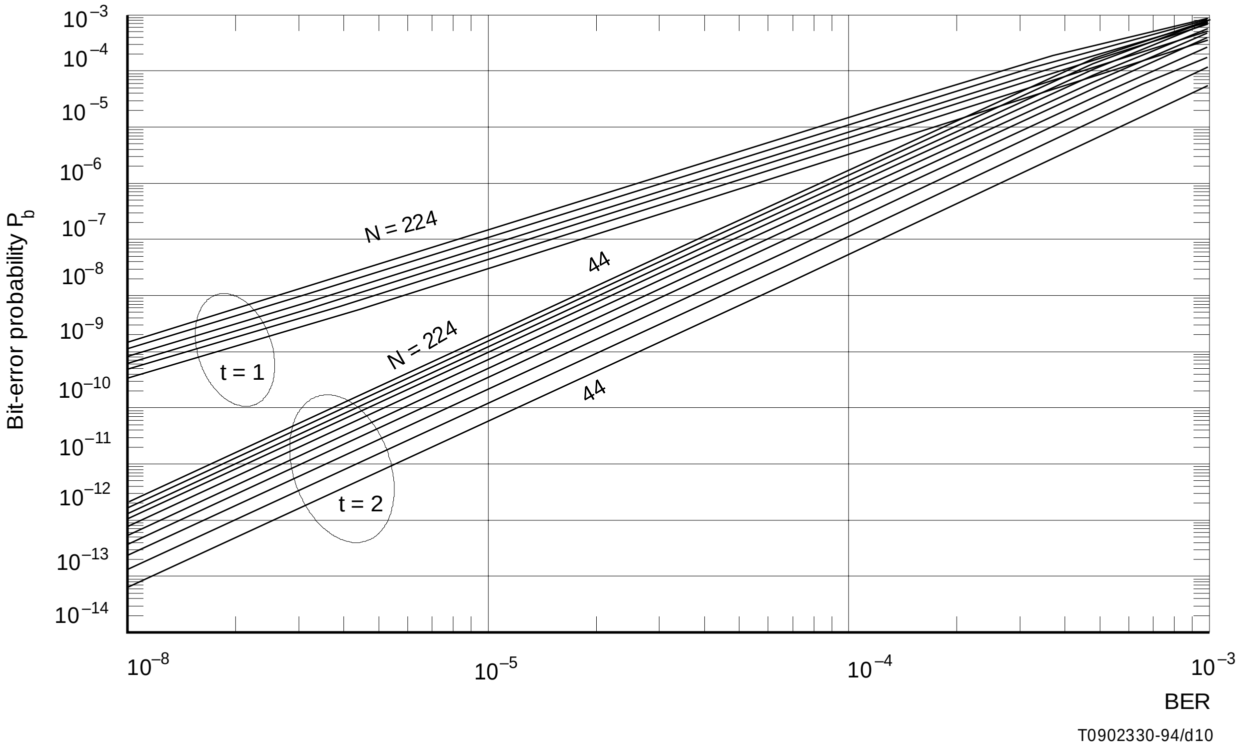

FIGURE I.1/J.52

Bit-error probability P_b after decoding N = 44, 64, ..., 224

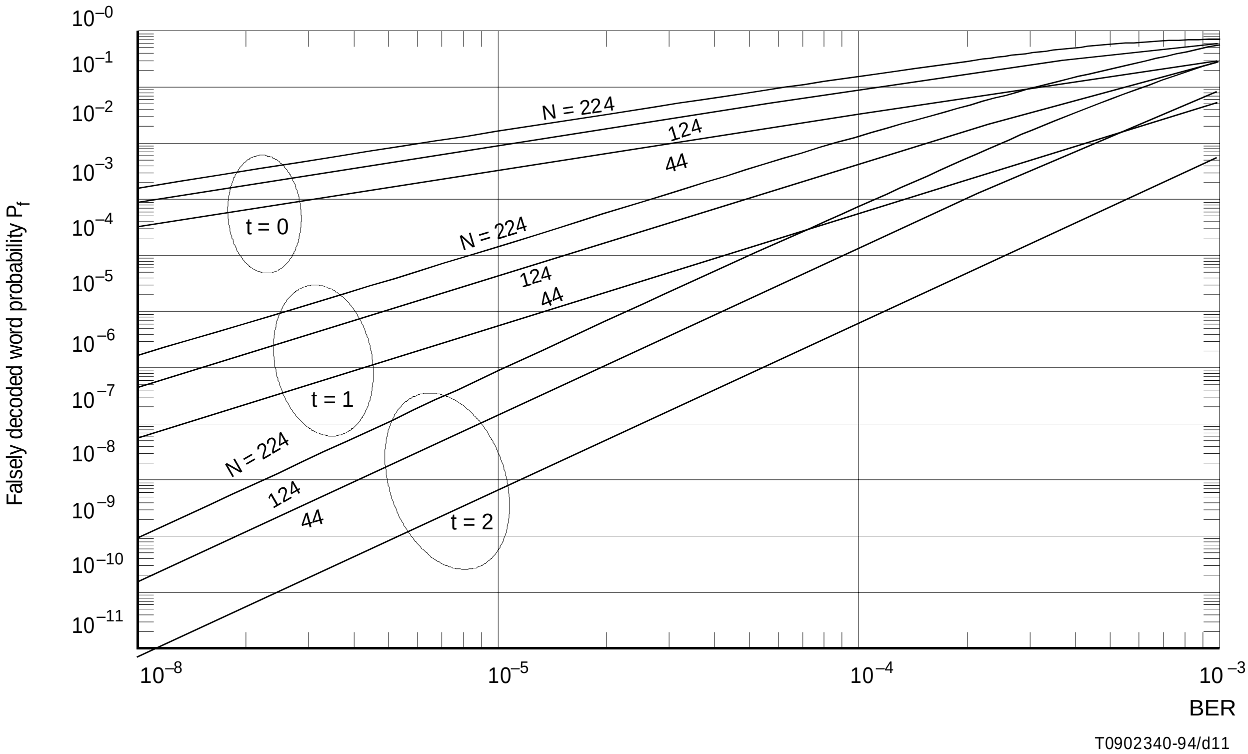

FIGURE I.2/J.52

Probability P_f of a falsely decoded codeword

ITU-T RECOMMENDATIONS SERIES

Series A Organization of the work of the ITU-T

Series B Means of expression

Series C General telecommunication statistics

Series D General tariff principles

Series E Telephone network and ISDN

Series F Non-telephone telecommunication services

Series G Transmission systems and media

Series H Transmission of non-telephone signals

Series I Integrated services digital network

Series J Transmission of sound-programme and television signals

Series K Protection against interference

Series L Construction, installation and protection of cables and other elements of outside plant

Series M Maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits

Series N Maintenance: international sound-programme and television transmission circuits

Series O Specifications of measuring equipment

Series P Telephone transmission quality

Series Q Switching and signalling

Series R Telegraph transmission

Series S Telegraph services terminal equipment

Series T Terminal equipments and protocols for telematic services

Series U Telegraph switching

Series V Data communication over the telephone network

Series X Data networks and open system communication

Series Z Programming languages