Recommendation G.722

7 kHz AUDIO-CODING WITHIN 64 KBIT/S

(Melbourne, 1988)

1 General

1.1 *Scope and outline description*

This Recommendation describes the characteristics of an audio (50 to 7 000 Hz) coding system which may be used for a variety of higher quality speech applications. The coding system uses sub-band adaptive differential pulse code modulation (SB-ADPCM) within a bit rate of 64 kbitB/Fs. The system is henceforth referred to as 64 kbit/s (7 kHz) audio coding. In the SB-ADPCM technique used, the frequency band is split into two sub-bands (higher and lower) and the signals in each sub-band are encoded using ADPCM.

The system has three basic modes of operation corresponding to the bit rates used for 7 kHz audio coding : 64, 56 and 48 kbit/s. The latter two modes allow an auxiliary data channel of 8 and 16 kbit/s respectively to be provided within the 64 kbit/s by making use of bits from the lower sub-band.

Figure 1/G.722 identifies the main functional parts of the 64 kbit/s (7 kHz) audio codec as follows:

i) 64 kbit/s (7 kHz) audio encoder comprising:

— a transmit audio part which converts an audio signal to

a uniform digital signal which is coded using 14 bits with

16 kHz sampling;

— a SB-ADPCM encoder which reduces the bit rate to 64 kbit/s.

ii) 64 kbitB/Fs (7 kHz) audio decoder comprising:

— a SB-ADPCM decoder which performs the reverse operation to the encoder, noting that the effective audio coding bit rate at the input of the decoder can be 64, 56 or 48 kbit/s depending on the mode of operation;

— a receive audio part which reconstructs the audio signal from the uniform digital signal which is encoded using 14 bits with 16 kHz sampling.

The following two parts, identified in Figure 1/G.722 for clarification, will be needed for applications requiring an auxiliary data channel within the 64 kbit/s:

— a data insertion device at the transmit end which makes use of, when needed, 1 or 2 audio bits per octet depending on the mode of operation and substitutes data bits to provide an auxiliary data channel of 8 or 16 kbit/s respectively;

— a data extraction device at the receive end which determines the mode of operation according to a mode control strategy and extracts the data bits as appropriate.

Paragraph 1.2 contains a functional description of the transmit and receive audio parts, § 1.3 describes the modes of operation and the implication of inserting data bits on the algorithms, whilst §§ 1.4 and 1.5 provide the functional descriptions of the SB-ADPCM encoding and decoding algorithms respectively. Paragraph 1.6 deals with the timing requirements. Paragraph 2 specifies the transmission characteristics of the 64 kbit/s (7 kHz) audio codec and of the transmit and receive audio parts, §§ 3 and 4 give the principles of the SB-ADPCM encoder respectively whilst §§ 5 and 6 specify the computational details of the Quadrature Mirror Filters (QMF) and of the ADPCM encoders and decoders respectively.

Networking aspects and test sequences are addressed in Appendices I and II respectively to this Recommendation.

Recommendation G.725 contains specifications for in-channel handshaking procedures for terminal identification and for mode control strategy, including interworking with existing 64 kbit/s PCM terminals.

Figure 1/G.722, p.

1.2 *Functional description of the audio parts*

Figure 2/G.722 shows a possible arrangement of audio parts in a 64 kbit/s (7 kHz) audio coding terminal shown simply to identify the audio parts and are not considered further in this Recommendation.

In order to facilitate the measurement of the transmission characteristics as specified in § 2, test points A and B need to be provided as shown. These test points may either be for test purposes only or, where the audio parts are located in different units from the microphone, loudspeaker, etc., correspond to physical interfaces.

The transmit and receive audio parts comprise either the following functional units or any equivalent items satisfying the specifications of § 2:

i) transmit:

— an input level adjustment device,

— an input anti-aliasing filter ,

— a sampling device operating at 16 kHz,

— an analogue-to-uniform digital converter with 14 bits and with 16 kHz sampling;

ii) receive:

— a uniform digital-to-analogue converter with 14 bits and with 16 kHz sampling,

— a reconstructing filter which includes x/sin x correction,

— an output level adjustment device.

Figure 2/G.722, p.

1.3 *Possible modes of operation and implications of inserting data*

The three basic possible modes of operation which correspond to the bit rates available for audio coding at the input of the decoder are defined in Table 1/G.722.

H.T. [T1.722]

TABLE 1/G.722

Basic possible modes of operation

Mode Auxiliary data channel bit rate }	7 kHz audio coding bit rate	{
1 2 3	64 kbit/s 56 kbit/s 48 kbit/s	0 kbit/s 8 kbit/s 16 kbit/s

Table 1/G.722 [T1.722], p.

See Appendix I for examples of applications using one or several of these modes and for their corresponding subjective quality.

The 64 kbit/s (7 kHz) audio encoder uses 64 kbit/s for audio coding at all times irrespective of the mode of operation. The audio coding algorithm has been chosen such that, without sending any indication to the encoder, the least significant bit or two least significant bits of the lower sub-band may

be used downstream from the 64 kbit/s (7 kHz) audio encoder in order to substitute the auxiliary data channel bits. However, to maximize the audio performance for a given mode of operation, the 64 kbit/s (7 kHz) audio decoder must be optimized to the bit rate available for audio coding. Thus, this Recommendation describes three variants of the SB-ADPCM decoder and, for applications requiring an auxiliary data channel, an indication must be forwarded to select in the decoder the variant appropriate to the mode of operation. Figure 1/G.722 illustrates the arrangement. It should be noted that the bit rate at the input of the 64 kbit/s (7 kHz) audio decoder is always 64 kbit/s but comprising 64, 56 or 48 kbit/s for audio coding depending on the mode of operation. From an algorithm viewpoint, the variant used in the SB-ADPCM decoder can be changed in any octet during the transmission. When no indication about the mode of operation is forwarded to the decoder, the variant corresponding to Mode 1 should be used.

A mode mismatch situation, where the variant used in the 64 kbit/s (7 kHz) audio decoder for a given octet does not correspond to the mode of operation, will not cause misoperation of the decoder. However, to maximize the audio performance, it is recommended that the mode control strategy adopted in

the data extraction device should be such as to minimize the duration of the mode mismatch. Appendix I gives further information on the effects of a mode mismatch. To ensure compatibility between various types of 64 kbit/s (7 kHz) audio coding terminals, it is recommended that, as a minimum, the variant corresponding to Mode 1 operation is always implemented in the decoder.

The mode control strategy could be derived from the auxiliary data channel protocol (see Recommendation G.725).

1.4 *Functional description of the SB-ADPCM encoder*

Figure 3/G.722 is a block diagram of the SB-ADPCM encoder. A functional description of each block is given below in §§ 1.4.1 to 1.4.4.

Figure 3/G.722, p.

1.4.1 *Transmit quadrature mirror filters* (QMFs)

The transmit QMFs comprise two linear-phase non-recursive digital filters which split the frequency band 0 to 8000 Hz into two sub-bands: the lower sub-band (0 to 4000 Hz) and the higher sub-band (4000 to 8000 Hz). The

input to the transmit QMFs, $x_{i\backslash dn}$, is the output from the transmit audio part and is sampled at 16 kHz. The outputs, x_L and x_H, for the lower and higher sub-bands respectively, are sampled at 8 kHz.

1.4.2 *Lower sub-band ADPCM encoder*

Figure 4/G.722 is a block diagram of the lower sub-band ADPCM encoder. The lower sub-band input signal, x_L after subtraction of an estimate, s_L, of the input signal produces the difference signal, e_L. An adaptive 60-level non linear quantizer is used to assign six binary digits to the value of the difference signal to produce a 48 kbit/s signal, I_L.

In the feedback loop, the two least significant bits of I_L are deleted to produce a 4-bit signal $I_{L\backslash \mathrm{dt}}$, which is used for the quantizer adaptation and applied to a 15-level inverse adaptive quantizer to produce a quantized difference signal, $d_{L\backslash \mathrm{dt}}$. The signal estimate, s_L is added to this quantized difference signal to produce a reconstructed version, $r_{L\backslash \mathrm{dt}}$, of the lower sub-band input signal. Both the reconstructed signal and the quantized difference signal are operated upon by an adaptive predictor which produce the estimate, s_L, of the input signal, thereby completing the feedback loop.

4-bit operation, instead of 6-bit operation, in the feedback loops of both the lower sub-band ADPCM encoder, and the lower sub-band ADPCM decoder allows the possible insertion of data in the two least significant bits as described in § 1.3 without causing misoperation in the decoder. Use of a 60-level quantizer (instead of 64-level) ensures that the pulse density requirements as described in Recommendation G.802 are met under all conditions and in all modes of operation.

Figure 4/G.722, p.

1.4.3 *Higher sub-band ADPCM encoder*

Figure 5/G.722 is a block diagram of the higher sub-band ADPCM encoder. The higher sub-band input signal, x_H after subtraction of an estimate, s_H, of the input signal, produces the difference signal, e_H. An adaptive 4-level non linear quantizer is used to assign two binary digits to the value of the difference signal to produce a 16 kbit/s signal, I_H.

An inverse adaptive quantizer produces a quantized difference signal, d_H, from these same two binary digits. The signal estimate, s_H, is added to this quantized difference signal to produce a reconstructed version, r_H, of the higher sub-band input signal. Both the reconstructed signal and the quantized difference signal are operated upon by an adaptive predictor which produces the estimate, s_H, of the input signal, thereby completing the feedback loop.

Figure 5/G.722, p.

1.4.4 *Multiplexer*

The multiplexer (MUX) shown in Figure 3/G.722 is used to combine the signals, I_L and I_H, from the lower and higher sub-band ADPCM encoders respectively into a composite 64 kbit/s signal, I, with an octet format for transmission.

The output octet format, after multiplexing, is as follows:

$$I_{H1} I_{H2} I_{L1} I_{L2} I_{L3} I_{L4} I_{L5} I_{L6}$$

where I_{H1} is the first bit transmitted, and where I_{H1} and I_{L1} are the most significant bits of I_H and I_L respectively, whilst I_{H2} and I_{L6} are the least significant bits of I_H and I_L respectively.

1.5 *Functional description of the SB-ADPCM decoder*

Figure 6/G.722 is a block diagram of the SB-ADPCM decoder. A functional description of each block is given below in §§ 1.5.1 to 1.5.4.

Figure 6/G.722, p.

1.5.1 *Demultiplexer*

The demultiplexer (DMUX) decomposes the received 64 kbit/s octet-formatted signal, I_r, into two signals, I_{Lr} and I_H, which form the codeword inputs to the lower and higher sub-band ADPCM decoders respectively.

1.5.2 *Lower sub-band ADPCM decoder*

Figure 7/G.722 is a block diagram of the lower sub-band ADPCM decoder. This decoder can operate in any of three possible variants depending on the received indication of the mode of operation.

Figure 7/G.722, p.

The path which produces the estimate, s_L, of the input signal including the quantizer adaptation , is identical to the feedback portion of the lower sub- band ADPCM encoder described in § 1.4.2. The reconstructed signal, r_L, is produced by adding to the signal estimate one of three possible quantized difference signals, $d_{L,6}$, $d_{L,5}$ or $d_{L,4}$ (= d_{Lt} - see note), selected according to the received indication of the mode of operation. For each indication, Table 2/G.722 shows the quantized difference signal selected, the inverse adaptive quantizer used and the number of least significant bits deleted from the input codeword.

H.T. [T2.722]

TABLE 2/G.722

Lower sub-band ADPCM decoder variants

{
Received indication of mode of operation
}
Quantized difference signal selected
}
Inverse adaptive quantizer used
}
Number of least significant bits deleted from input codeword,
I
L
r
}
	{
	{
	{		

Mode 1	d L , 6	60-level	0
Mode 2	d L , 5	30-level	1
Mode 3	d L , 4	15-level	2

Note — For clarification purposes, all three inverse quantizers have been indicated in the upper portion of Figure 7/G.722. In an optimized implementation, the signal d L t, produced in the predictor loop, could be substituted for d L , 4.

Table 2/G.722 [T2.722], p.

1.5.3 *Higher sub-band ADPCM decoder*

Figure 8/G.722 is a block diagram of the higher sub-band ADPCM decoder. This decoder is identical to the feedback portion of the higher sub-band ADPCM encoder described in § 1.4.3, the output being the reconstructed signal, r_H.

Figure 8/G.722, p.

1.5.4 *Receive QMFs*

The receive QMFs shown in Figure 6/G.722 are two linear-phase non-recursive digital filters which interpolate the outputs, r_L and r_H, of the lower and higher sub-band ADPCM decoders from 8 kHz to 16 kHz and which then produce an output, $x_{o\du\dt}$, sampled at 16 kHz which forms the input to the receive audio parts.

Excluding the ADPCM coding processes, the combination of the transmit and the receive QMFs has an impulse response which closely approximates a simple delay whilst, at the same time, the aliasing effects associated with the 8 kHz sub-sampling are cancelled.

1.6 *Timing requirements*

64 kHz bit timing and 8 kHz octet timing should be provided by the network to the audio decoder.

For a correct operation of the audio coding system, the precision of the 16 kHz sampling frequencies of the A/D and D/A converters must be better than ± | 0 | (mu | 0^{D}lF261^{6}.

2 Transmission characteristics

2.1 *Characteristics of the audio ports and the test points*

Figure 2/G.722 indicates the audio input and output ports and the test points (A and B). It is for the designer to determine the characteristics of the audio ports and the test points (i.e. relative levels, impedances, whether balanced or unbalanced). The microphone, pre-amplifier, power amplifier and loudspeaker should be chosen with reference to the specifications of the audio parts: in particular their nominal bandwidth, idle noise and distortion.

It is suggested that input and ouput impedances should be high and low, respectively, for an unbalanced termination whilst for a balanced termination these impedances should be 600 ohms. However, the audio parts should meet all audio parts specifications for their respective input and output impedance conditions.

2.2 *Overload point*

The overload point for the analogue-to-digital and digital-to-analogue converters should be + 9 dBm0 ± | .3 dB. This assumes the same nominal speech level (see Recommendation G.232) as for 64 kbit/s PCM, but with a wider margin for the maximum signal level which is likely to be necessary with conference arrangements. The measurement method of the overload point is under study.

2.3 *Nominal reference frequency*

Where a nominal reference frequency of 1000 Hz is indicated below, the actual frequency should be chosen equal to 1020 Hz. The frequency tolerance should be +2 to —7 Hz.

2.4 *Transmission characteristics of the 64 kbit/s (7 kHz) audio codec*

The values and limits specified below should be met with a 64 kbit/s (7 kHz) audio encoder and decoder connected back-to-back. For practical reasons, the measurements may be performed in a looped configuration as shown in Figure 9a)/G.722. However, such a looped configuration is only intended to simulate an actual situation where the encoder and decoder are located at the two ends of a connection.

These limits apply to operation in Mode 1.

2.4.1 *Nominal bandwidth*

The nominal 3 dB bandwidth is 50 to 7000 Hz.

2.4.2 *Attenuation/frequency distortion*

The variation with frequency of the attenuation should satisfy the limits shown in the mask of Figure 10/G.722. The nominal reference frequency is 1000 Hz and the test level is —10 dBm0.

2.4.3 *Absolute group delay*

The absolute group delay, defined as the minimum group delay for a sine wave signal between 50 and 7000 Hz, should not exceed 4 ms. The test level is —10 dBm0.

2.4.4 *Idle noise*

The unweighted noise power measured in the frequency range 50 to 7000 Hz with no signal at the input port (test point A) should not exceed —66 dBm0. When measured in the frequency range 50 Hz to 20 kHz the unweighted noise power should not exceed —60 dBm0.

Figure 9/G.722, p.

Figure 10/G.722, p.

2.4.5 *Single frequency noise*

The level of any single frequency (in particular 8000 Hz, the sampling frequency and its multiples), measured selectively with no signal at the input port (test point A) should not exceed —70 dBm0.

2.4.6 *Signal-to-total distortion ratio*

Under study.

2.5 *Transmission characteristics of the audio parts*

When the measurements indicated below for the audio parts are from audio-to-audio, a looped configuration as shown in Figure 9b)/G.722 should be used. The audio parts should also meet the specifications of § 2.4 with the measurement configuration of Figure 9b)/G.722.

2.5.1 *AttenuationB/Ffrequency response of the input anti-aliasing filter*

The in-band and out-of-band attenuation/frequency response of the input anti-aliasing filter should satisfy the limits of the mask shown in Figure 11/G.722. The nominal reference frequency is 1000 Hz and the test level for the in-band characteristic is —10 dBm0. Appropriate measurements should be made to check the out-of-band characteristic taking into account the aliasing due to the 16 kHz sampling.

2.5.2 *AttenuationB/Ffrequency response of the output reconstructing filter*

The in-band and out-of-band attenuation/frequency response of the output reconstructing filter should satisfy the limits of the mask shown in Figure 12/G.722. The nominal reference frequency is 1000 Hz and the test level for the in-band characteristic is —10 dBm0. Appropriate measurements should be made to check the out-of-band characteristic taking into account the aliasing due to the 16 kHz sampling. The mask of Figure 12/G.722 is valid for the whole of the receive audio part including any pulse amplitude modulation distortion and x/sin x correction.

Figure 11/G.722, p.

Figure 12/G.722, p.

2.5.3 *Group-delay distortion with frequency*

The group-delay distortion, taking the minimum value of group delay as a reference, should satisfy the limits of the mask shown in Figure 13/G.722.

Figure 13/G.722, p.

2.5.4 *Idle noise for the receive audio part*

The unweighted noise power of the receive audio part measured in the frequency range 50 to 7000 Hz with a 14-bit all-zero signal at its input should not exceed —75 dBm0.

2.5.5 *Signal-to-total distortion ratio as a function of input level*

With a sine wave signal at a frequency excluding simple harmonic relationships with the 16 kHz sampling frequency, applied to test point A, the ratio of signal-to-total distortion power as a function of input level measured unweighted in the frequency range 50 to 7000 Hz at test point B, should satisfy the limits of the mask shown in Figure 14/G.722. Two measurements should be performed, one at a frequency of about 1 kHz and the other at a frequency of about 6 kHz.

Figure 14/G.722, p.

2.5.6 *Signal-to-total distortion ratio as a function of frequency*

With a sine wave signal at a level of —10 dBm0 applied to test point A, the ratio of signal-to-total distortion power as a function of frequency measured unweighted in the frequency range 50 to 7000 Hz at test point B should satisfy the limits of the mask shown in Figure 15/G.722.

Figure 15/G.722, p.

2.5.7 *Variation of gain with input level*

With a sine wave signal at the nominal reference frequency of 1000 Hz, but excluding the sub-multiple of the 16 kHz sampling frequency, applied to test point A, the gain variation as a function of input level relative to the gain at an input level of —10 dBm0 measured selectively at test point B, should satisfy the limits of the mask shown in Figure 16/G.722.

Figure 16/G.722, p.

2.5.8 *Intermodulation*

Under study.

2.5.9 *Go/return crosstalk*

The crosstalk from the transmit direction to the receive direction should be such that, with a sine wave signal at any frequency in the range 50 to 7000 Hz and at a level of +6 dBm0 applied to test point A, the crosstalk level measured selectively at test point B should not exceed —64 dBm0. The measurement should be made with a 14-bit all-zero digital signal at the input to the receive audio part.

The crosstalk from the receive direction to the transmit direction should be such that, with a digitally simulated sine wave signal at any frequency in the range of 50 to 7000 Hz and a level of +6 dBm0 applied to the input of the receive audio part, the crosstalk level measured selectively and with the measurement made digitally at the output of the transmit audio part should not exceed —64 dBm0. The measurement should be made with no signal at test point A, but with the test point correctly terminated.

2.6 *Transcoding to and from 64 kbit/s PCM*

For compatibility reasons with 64 kbit/s PCM, transcoding between 64 kbit/s (7 kHz) audio coding and 64 kbit/s PCM should take account of the relevant specifications of Recommendations G.712, G.713 and G.714. When the audio signal is to be heard through a loudspeaker, more stringent specifications may be necessary. Further information may be found in Appendix I.

3 SB-ADPCM encoder principles

A block diagram of the SB-ADPCM encoder is given in Figure 3/G.722. Block diagrams of the lower and higher sub-band ADPCM encoders are given respectively in Figures 4/G.722 and 5/G.722.

Main variables used for the descriptions in §§ 3 and 4 are summarized in Table 3/G.722. In these descriptions, index (j) indicates a value corresponding to the current 16 kHz sampling interval, index (j —l) indicates a value corresponding to the previous 16 kHz sampling interval, index (n) indicates a value corresponding to the current 8 kHz sampling interval, and index (n —1) indicates a value corresponding to the previous 8 kHz sampling interval. Indices are not used for internal variables, i.e. those employed only within individual computational blocks.

3.1 *Transmit QMF*

A 24-coefficient QMF is used to compute the lower and higher sub-band signal components. The QMF coefficient values, h_i, are given in Table 4/G.722.

The output variables, $x_L(n)$ and $x_H(n)$, are computed in the following way:

3.2 *Difference signal computation*

The difference signals, $e_L(n)$ and $e_H(n)$, are computed by subtracting predicted values, $s_L(n)$ and $s_H(n)$, from the lower and higher sub-band input values, $x_L(n)$ and $x_H(n)$:

H.T. [T3.722]

TABLE 3/G.722

Variables used in the SB-ADPCM encoder and decoder descriptions

Variable	Description
x i n Input value (uniform representation) }	{
x L, x H	QMF output signals
S L p, S H p Pole-predictor output signals }	{
a L , i, a H , i Pole-predictor coefficients }	{
r L, r L t, r H Reconstructed signals (non truncated and truncated) }	{
b L , i, b H , i Zero-predictor coefficients }	{
d L, d L t, d H Quantized difference signals (non truncated and truncated) }	{
S L z, S H z Zero-predictor output signals }	{
S L, S H	Predictor output signals
e L, e H Difference signals to be quantized }	{
∇ L, ∇ H Logarithmic quantizer scale factors }	{
?63 L, ?63 H Quantizer scale factor (linear) }	{
I L, I L t, I H Codewords (non truncated and truncated) }	{
P L t, P H Partially reconstructed signals }	{
I L r Received lower sub-band codeword }	{
X o u t	Output value (uniform)

Note — Variables used exclusively within one section are not listed. Subscripts L and H refer to lower sub-band and higher sub-band values. Subscript Lt denotes values generated from the truncated 4-bit codeword as opposed to the nontruncated 6-bit (encoder) or 6-, 5- or 4-bit (decoder) codewords.

Tableau 3/G.722 [T3.722], p. 19

Blanc

H.T. [T4.722]

TABLE 4/G.722

Transmit and receive OMF coefficient values

{h0↓1, h23}	—0.366211E—03
{h1↓1, h22}	—0.134277E—02
{h2↓1, h21}	—0.134277E—02
{h3↓1, h20}	—0.646973E—02
{h4↓1, h19}	—0.146484E—02
{h5↓1, h18}	—0.190430E—01
{h6↓1, h17}	—0.390625E—02

{

h	
7↓**1**	
, h	
1	
6	
}	{
—	
0.441895E—01	
}	
{	
h	
8↓**1**	
, h	
1	
5	
}	—0.256348E—01
{	
h	
9↓**1**	
, h	
1	
4	
}	—0.982666E—01
h 1 0, h 1 3	— 0.116089E+00
h 1 1, h 1 2	— 0.473145E+00

Tableau 4/G.722 [T4.722], p. 20

3.3 *Adaptive quantizer*

The difference signals, $e_L(n)$ and $e_H(n)$, are quantized to 6 and 2 bits for the lower and higher sub-bands respectively. Tables 5/G.722 and 6/G.722 give the decision levels and the output codes for the 6- and 2-bit quantizers respectively. In these tables, only the positive decision levels are indicated, the negative levels can be determined by symmetry. m_L and m_H are indices for the quantizer intervals. The interval boundaries, *LL* 6, *LU* 6, *HL* and *HU* , are scaled by computed scale factors, ?63 $_L(n)$ and ?63 $_H(n)$ (see § 3.5). Indices, m_L and m_H, are then determined to satisfy the following:

for the lower and higher sub-bands respectively.

The output codes, *ILN* and *IHN* , represent negative intervals, whilst the output codes, *ILP* and *IHP* , represent positive intervals. The output codes, $I_L(n)$ and $I_H(n)$, are then given by:

for the lower and higher sub-bands respectively.

H.T. [T5.722]

TABLE 5/G.722

Decision levels and output codes for the 6-bit lower sub-band quantizer

m L	LL6	LU6	ILN	ILP
1 2	0.00000 0.06817	0.06817 0.14103	111111 111110	111101 111100
3 4 5 6	{			
0.14103				
0.21389				
0.29212				
0.37035				
}	{			
0.21389				
0.29212				
0.37035				
0.45482				
}	011111 011110 011101 011100	111011 111010 111001 111000		
7 8 9 10	{			
0.45482				
0.53929				
0.63107				
0.72286				
}	{			
0.53929				
0.63107				
0.72286				
0.82335				
}	011011 011010 011001 011000	110111 110110 110101 110100		
11 12 13 14	{			
0.82335				
0.92383				
1.03485				
1.14587				
}	{			
0.92383				
1.03485				
1.14587				
1.26989				
}	010111 010110 010101 010100	110011 110010 110001 110000		
15 16 17 18	{			
1.26989				
1.39391				
1.53439				
1.67486				
}	{			
1.39391				
1.53439				
1.67486				
1.83683				
}	010011 010010 010001 010000	101111 101110 101101 101100		
19 20 21 22	{			
1.83683				
1.99880				
2.19006				
2.38131				
}	{			
1.99880				
2.19006				
2.38131				
2.61482				
}	001111 001110 001101 001100	101011 101010 101001 101000		
23 24 25 26	{			
2.61482				
2.84833				
3.14822				

3.44811 } 2.84833 3.14822 3.44811 3.86796 }	{ 001011 001010 001001 001000	100111 100110 100101 100100		
27 28 29 30 3.86796 4.28782 4.99498 5.70214 } 100011 100010 100001 100000 }	{ 4.28782 4.99498 5.70214 ∞	000111 000110 000101 000100	{	

Note — If a transmitted codeword for the lower sub-band signal has been transformed, due to transmission errors to one of the four suppressed codewords ‘‘0000XX’’, the received code word is set at ‘‘111111’’.

Table 5/G.722 [T5.722], p.

H.T. [T6.722]

TABLE 6/G.722

Decision levels and output codes for the 2-bit higher sub-band quantizer

m H	HL	HH	IHN	IHP
1 2	0 1.10156	1.10156 ∞	01 00	11 10

Table 6/G.722 [T6.722], p.

3.4 *Inverse adaptive quantizers*

3.4.1 *Inverse adaptive quantizer in the lower sub-band ADPCM encoder*

The lower sub-band output code, $I_L(n)$, is truncated by two bits to produce $I_{L\backslash dt}(n)$. The 4-bit codeword, $I_{L\backslash dt}(n)$, is converted to the truncated quantized difference signal , $d_{L\backslash dt}(n)$, using the QL 4^{D}lF261^{1} output values of Table 7/G.722, and scaled by the scale factor , ?63 $_L(n)$:

where sgn $[I_{L\backslash dt}(n))$ is derived from the sign of $e_L(n)$ defined in Equation 3-9.

There is a unique mapping, shown in Table 7/G.722, between four adjacent 6-bit quantizer intervals and the QL 4^{D}lF261^{1} output values. QL 4^{D}lF261$^{1}[I_{L\backslash dt}(n)]$ is determined in two steps: first determination of the quantizer interval index , m_L, corresponding to $I_L(n)$ from Table 5/G.722, and then determination of Q_L | $^{\mathrm{D}}$lF261$^{1}(m_L)$ by reference to Table 7/G.722.

H.T. [T7.722]

TABLE 7/G.722

Output values and multipliers for 6, 5 and 4-bit lower sub-band inverse quantizers

m L	$QL6^{D}lF261^{1}$	$QL5^{D}lF261^{1}$	$QL4^{D}lF261^{1}$	W L
. 1 2	. 0.03409 0.10460	{		
.				
0.06817				
.				
}	{			
0.0000				
.				
.				
}	{			
—0.02930				
.				
.				
}				
3 4 5 6	{			
0.17746				
0.25300				
0.33124				
0.41259				
}	{			
0.21389				
.				
0.37035				
.				
}	{			
.				
0.29212				
.				
.				
}	{			
.				
—0.01465				
.				
.				
}				
7 8 9 10	{			
0.49706				
0.58518				
0.67697				
0.77310				
}	{			
0.53929				
.				
0.72286				
.				
}	{			
.				
0.63107				
.				
.				
}	{			
.				
0.02832				
.				
.				
}				
11 12 13 14	{			
0.87359				
0.97934				
1.09036				
1.20788				
}	{			

0.92383
.
1.14587
.
} {
.
1.03485
.
.
} {
.
0.08398
.
.
}

---

15 16 17 18 {
1.33191
1.46415
1.60462
1.75585
} {
1.39391
.
1.67486
.
} {
.
1.53439
.
.
} {
.
0.16309
.
.
}

---

19 20 21 22 {
1.91782
2.09443
2.28568
2.49806
} {
1.99880
.
2.38131
.
} {
.
2.19006
.
.
} {
.
0.26270
.
.
}

---

23 24 25 26 {
2.73157
2.99827
3.29816
3.65804

} 2.84833 . 3.44811 . } . 3.14822 . . } . 0.58496 . . }	{ { {			
27 28 29 30 4.07789 4.64140 5.34856 6.05572 } 4.28782 . 5.70214 . } . 4.99498 . . } . 1.48535 . . }	{ { { {			

Table 7/G.722 [T7.722], p.

3.4.2 *Inverse adaptive quantizer in the higher sub-band ADPCM encoder*

The higher sub-band output code, $I_H(n)$ is converted to the quantized difference signal, $d_H(n)$, using the Q 2^{D}lF261^{1} output values of Table 8/G.722 and scaled by the scale factor, ?63$_H(n)$:

where sgn[$I_H(n)$] is derived from the sign of $e_H(n)$ defined in Equation (3-10), and where Q_2^{D}lF261^{1}[$I_H(n)$] is determined in two steps: first determine the quantizer interval index, m_H, corresponding to $I_H(n)$ from Table 6/G.722 and then determine Q 2^{D}lF261^{1}(m_H) by reference to Table 8/G.722.

H.T. [T8.722]

TABLE 8/G.722

Output values and multipliers for the 2-bit higher sub-band quantizer

m H	Q2^{D}lF261^{1}	W H
1	0.39453	—0.10449
2	1.80859	0.38965

Tableau 8/G.722 [T8.722],

3.5 *Quantizer adaptation*

This block defines ?63 $_L(n)$ and ?63 $_H(n)$, the scaling factors for the lower and higher sub-band quantizers. The scaling factors are updated in the log domain and subsequently converted to a linear representation. For the lower sub-band, the input is $I_{L\backslash dt}(n)$, the codeword truncated to preserve the four most significant bits. For the higher sub-band, the 2-bit quantizer output, $I_H(n)$, is used directly.

Firstly the log scaling factors, ?63 $_L(n)$ and ?63 $_H(n)$, are updated as follows:

where W_L and W_H are logarithmic scaling factors multipliers given in Tables 7/G.722 and 8/G.722, and B is a leakage constant equal to 127/128.

Then the log scaling factors are limited, according to:

Finally, the linear scaling factors are computed from the log scaling factors, using an approximation of the inverse $\log_2$ function:

where ?63 $_{m\backslash di\backslash dn}$ is equal to half the quantizer step size of the 14 bit analogue-to-digital converter.

3.6 *Adaptive prediction*

3.6.1 *Predicted value computations*

The adaptive predictors compute predicted signal values, $s_L(n\ )$ and $s_H(n\ )$, for the lower and higher sub-bands respectively.

Each adaptive predictor comprises two sections: a second-order section that models poles, and a sixth-order section that models zeroes in the input signal.

The second order pole sections (coefficients $a_{L\backslash\mathrm{d},i}$ and $a_{H\backslash\mathrm{d},}\ i$) use the quantized reconstructed signals, $r\mathrm{L}_t(n\ )$ and $r_H(n\ )$, for prediction. The sixth order zero sections (coefficients $b_{L\backslash\mathrm{d},i}$) and $b_{H\backslash\mathrm{d},}\ i$) use the quantized difference signals, $d_{L\backslash\mathrm{d}t}(n\ )$ and $d_H(n\ )$. The zero-based predicted signals, $s_{L\backslash\mathrm{dz}}(n\ )$ and $s_{H\backslash\mathrm{dz}}(n\ )$, are also employed to compute partially reconstructed signals as described in § 3.6.2.

Firstly, the outputs of the pole sections are computed as follows:

Similarly, the outputs of the zero sections are computed as follows:

Then, the intermediate predicted values are summed to produce the predicted signal values:

3.6.2 *Reconstructed signal computation*

The quantized reconstructed signals, $r_{L\backslash\mathrm{d}t}(n\ )$ and $r_H(n\ )$, are computed as follows:

The partially reconstructed signals, $p_{L\backslash\mathrm{d}t}(n\ )$ and $p_H\ (n\ )$, used for the pole section adaptation, are then computed:

3.6.3 *Pole section adaptation*

The second order pole section is adapted by updating the coefficients, $a_{L\backslash d,\backslash d1}$, $a_{H\backslash d,\backslash d1}$, $a_{H\backslash d,\backslash d2}$, using a simplified gradient algorithm:

where

with

and

Then the following stability constraints are imposed:

$a_{H\backslash d,\backslash d1}(n)$ and $a_{H\backslash d,\backslash d2}(n)$ are similarly computed, replacing $a_{L\backslash d,\backslash d1}(n)$, $a_{L\backslash d,\backslash d2}(n)$ and $P_{L\backslash dt} \mid n)$ by $a_{H\backslash d,\backslash d1}(n)$, $a_{H\backslash d,\backslash d2}(n)$ and $P_H(n)$, respectively.

3.6.4 *Zero section adaptation*

The sixth order zero predictor is adapted by updating the coefficients $b_{L\backslash d,i}$ and $b_{H\backslash d,i}$ using a simplified gradient algorithm:

for $i = 1, 2 . \mid \mid 6$

and with

where $b_{L\backslash d,i} \mid n)$ is implictly limited to $\pm \mid$.

$b_{H\backslash d,i}(n)$ are similarly updated, replacing $b_{L\backslash d,i}(n)$ and $d_{L\backslash dt}(n)$ by $b_{H\backslash d,i}(n)$ and $d_H(n)$ respectively.

4 SB-ADPCM decoder principles

A block diagram of the SB-ADPCM decoder is given in Figure 6/G.722 and block diagrams of the lower and higher sub-band ADPCM decoders are given respectively in Figures 7/G.722 and 8/G.722.

The input to the lower sub-band ADPCM decoder, $I_{L\backslash dr}$, may differ from I_L even in the absence of transmission errors, in that one or two least significant bits may have been replaced by data.

4.1 *Inverse adaptive quantizer*

4.1.1 *Inverse adaptive quantizer selection for the lower sub-band ADPCM decoder*

According to the received indication of the mode of operation the number of least significant bits which should be truncated from the input codeword $I_{L\backslash\mathrm{dr}}$, and the choice of the inverse adaptive quantizer are determined, as shown in Table 2/G.722.

For operation in mode 1, the 6-bit codeword, $I_{L\backslash\mathrm{dr}}(n)$, is converted to the quantized difference, $d_L(n)$, according to *QL* 6^{D}lF261^{1} output values of Table 7/G.722, and scaled by the scale factor, ?63 $_L(n)$:

where sgn[$I_{L\backslash\mathrm{dr}}(n)$] is derived from the sign of $I_L(n)$ defined in equation (3-9).

Similarly, for operations in mode 2 or mode 3, the truncated codeword (by one or two bits) is converted to the quantized difference signal, $d_L(n)$, according to *QL* 5^{D}lF261^{1} or *QL* 4^{D}lF261^{1} output values of Table 7/G.722 respectively.

There are unique mappings, shown in Table 7/G.722, between two or four adjacent 6-bit quantizer intervals and the *QL* 5^{D}lF261^{1} or *QL* 4^{D}lF261^{1} output values respectively.

In the computations above, the output values are determined in two steps: first determination of the quantizer interval index, m_L, corresponding to $I_{L\backslash\mathrm{dr}}(n)$ from Table 5/G.722, and then determination of the output values corresponding to m_L by reference to Table 7/G.722.

The inverse adaptive quantizer, used for the computation of the predicted value and for adaptation of the quantizer and predictor, is described in § 3.4.1, but with $I_L(n)$ replaced by $I_{L\backslash\mathrm{dr}}(n)$.

4.1.2 *Inverse adaptive quantizer for the higher sub-band ADPCM decoder*

See § 3.4.2.

4.2 *Quantizer adaptation*

See § 3.5.

4.3 *Adaptive prediction*

4.3.1 *Predicted value computation*

See § 3.6.1.

4.3.2 *Reconstructed signal computation*

See § 3.6.2.

The output reconstructed signal for the lower sub-band ADPCM decoder, $r_L(n)$, is computed from the quantized difference signal, $d_L(n)$, as follows:

4.3.3 *Pole section adaptation*

See § 3.6.3.

4.3.4 *Zero section adaptation*

See § 3.6.4.

4.4 *Receive QMF*

A 24-coefficient QMF is used to reconstruct the output signal, $x_{out}(j)$, from the reconstructed lower and higher sub-band signals, $r_L(n)$ and $r_H(n)$. The QMF coefficient values, h_i, are the same as those used in the transmit QMF and are given in Table 4/G.722.

The output signals, $x_{o\du\dt}(j)$ and $x_{o\du\dt}(j+1)$, are computed in the following way:

where

5 Computational details for QMF

5.1 *Input and output signals*

Table 9/G.722 defines the input and output signals for the transmit and receive QMF. All input and output signals have 16-bit word lengths, which are limited to a range of —16384 to 16383 in 2's complement notation. Note that the most significant magnitude bit of the A/D output and the D/A input appears at the third bit location in XIN and XOUT, respectively.

H.T. [T9.722]

TABLE 9/G.722

Representation of input and output signals

Transmit QMF			
	Name	Binary representation	Description
Input S, S, —2, —3, . | | , —14, —15 } Input value (uniformly quantized) }	XIN {	{	
Output S, S, —2, —3, . | | , —14, —15 } Output signal for lower sub-band encoder }	XL {	{	
Output S, S, —2, —3, . | | , —14, —15 } Output signal for higher sub-band encoder }	XH {	{	
Receive QMF			
	Name	Binary representation	Description
Input S, S, —2, —3, . | | , —14, —15 } Lower sub-band reconstructed signal }	RL {	{	
Input S, S, —2, —3, . | | , —14, —15 } Higher sub-band reconstructed signal }	RH {	{	
Output S, S, —2, —3, . | | , —14, —15 } Output value (uniformly quantized) }	XOUT {	{	

Note — XIN and XOUT are represented in a sign-extended 15-bit format, where the LSB is set to ''0'' for 14-bit converters.

Table 9/G.722 [T9.722], p.

5.2 *Description of variables and detailed specification of sub-blocks*

This section contains a detailed expansion of the transmit and receive QMF. The expansions are illustrated in Figures 17/G.722 and 18/G.722 with the internal variables given in Table 10/G.722, and the QMF coefficients given in Table 11/G.722. The word lengths of internal variables, XA, XB and WD must be equal to or greater than 24 bits (see Note). The other internal variables have a minimum of 16 bit word lengths. A brief functional description and the full specification is given for each sub-block.

The notations used in the block descriptions are as follows:

> > | denotes an *n* -bit arithmetic shift right operation (sign extension),

\+ denotes arithmetic addition with saturation control which forces the result to the minimum or maximum representable value in case of underflow or overflow, respectively,

— denotes arithmetic subtraction with saturation control which forces the result to the minimum or maximum representable value in case of underflow or overflow, respectively.

* denotes arithmetic multiplication which can be performed with either truncation or rounding,

< denotes the ‘‘less than’’ condition as *x* < | fIy ; *x* is less than *y* ,

> denotes the ‘‘greater than’’ condition, as *x* > | fIy ; *x* is greater than *y* ,

= denotes the substitution of the right-hand variable for the left-hand variable.

Note 1 — Some freedom is offered for the implementation of the accumulation process in the QMF: the word lengths of the internal variables can be equal to or greater than 24 bits, and the arithmetic multiplications can be performed with either truncation or rounding. It allows a simplified implementation on various types of processors. The counterpart is that it excludes the use of digital test sequence for the test of the QMF.

H.T. [T10.722]

TABLE 10/G.722

Representation of internal processing variables and QMF coefficients

Transmit QMF		
Name	Binary representation	Description
XA S, —1, —2, —3, .		, —y+1, —y } Output signal of sub-block, ACCUMA } XB S, —1, —2, —3, .		, —y+1, —y } Output signal of sub-block, ACCUMB } XIN1, XIN2, .		, XIN23 S, S, —2, —3, .		, —14, —15 } Input signal with delays 1 to 23 }	{ { { { { {	
Receive QMF		
Name	Binary representation	Description
XD, XD1, .		, XD11 S, —1, —2, —3, .		, —14, —15 } Input signal for sub-block, ACCUMC, with delays 0 to 11 } XOUT1 S, S, —2, —3, .		, —14, —15 } XOUT2 S, S, —2, —3, .		, —14, —15 } XS, XS1, .		, XS11 S, —1, —2, —3, .		, —14, —15 } Input signal for sub-block, ACCUMD, with delays 0 to 11 } WD S, —1, —2, —3, .		, —y+1, —y }	{ { { 8 kHz sampled output value { 8 kHz sampled output value { { { Partial sum	
QMF coefficients		
Name	Binary representation	Description
H0, H1, .		, H23 S, —2, —3, —4, .		, —12, —13 } Filter coefficient values }	{ {	

Note — y is equal to or greater than 23.

Table 10/G.722 [T10.722], p.

H.T. [T11.722]
TABLE 11/G.722
QMF coefficient

Coefficient	Scaled values (see Note)
H0 , H23	3
H1 , H22	—11
H2 , H21	—11
H3 , H20	53
H4 , H19	12
H5 , H18	—156
H6 , H17	32
H7 , H16	362
H8 , H15	—210
H9 , H14	—805
H10 , H13	951
H11 , H12	3876

Note — QMF coefficients are scaled by 2^{13} with respect to the representation specified in Table 10/G.722.

Table 11/G.722 [T11.722], p.

5.2.1 *Description of the transmit QMF*

Figure 17/G.722, p.

DELAYX

Input: x

Output: y

Note — Index (*j*) indicates the current 16-kHz sample period, while index (*j* — 1) indicates the previous one.

Function: Memory block. For any input x, the output is given by:

y (*j*) = *x* (*j* — 1)

ACCUMA

Inputs: XIN, XIN2, XIN4, . | | , XIN22

Output: XA

Note 1 — H0, H2, . | | , H22 are obtained from Table 11/G.722.

Note 2 — The values XIN, XIN2, . | | , XIN22 and H0, H2, shifted before multiplication, if so desired. The result XA must be rescaled accordingly, In performing these scaling operations the following rules must be obeyed:

1) the precision of XIN, XIN2, . | | , XIN22 and H0, H2, . | | , H22 as given in Table 9/G.722 and Table 10/G.722 must be retained,

2) the partial products and the ouptut signal XA must be retained to a significance of at least 2^{D}lF261^{23},

3) no saturation should occur in the calculation of the function XA.

Note 3 — No order of summation is specified in accumulating the partial products.

Function: Multiply the even order QMF coefficients by the appropriately delayed input signals, and accumulate these products.

XA = (XIN * H0) + (XIN2 * H2) + (XIN4 * H4) + . | | + (XIN22 * H22)

ACCUMB

Inputs: XIN1, XIN3, XIN5, . | | , XIN23

Output: XB

Note 1 — H1, H3, . | | , H23 are obtained from Table 11/G.722.

Note 2 — The values XIN1, XIN3, . | | , XIN23 and H1, H3, shifted before multiplication, if so desired. The result XB must be rescaled accordingly. In performing these scaling operations the following rules must be obeyed:

1) the precision of XIN1, XIN3, . | | , XIN23 and H1, H3, . | | , H23 as given in Table 9/G.722 and Table 10/G.722 must be retained,

2) the partial products and the output signal X3 must be retained to a significance of at least 2^{D}lF261^{23},

3) no saturation should occur in the calculation of the function XB.

Note 3 — No order of summation is specified in accumulating the partial products.

Function: Multiply the odd order QMF coefficients by the appropriately delayed input signals, and accumulate these products.

XB = (XIN1 * H1) + (XIN3 * H3) + (XIN5 * H5) + . | | + (XIN23 * H23)

LOWT

Inputs: XA, XB

Output: XL

Function: Compute the lower sub-band signal component.

XL = (XA + XB) > > (y — 15) [Formula Deleted]