BBC Engineering

September 1986

STOP PRESS
UK Government approved

This specification has now been
approved by the UK Government

Specification of a Standard for UK Stereo-with-Television Transmissions

DRAFT

British Broadcasting Corporation

Final

~~Draft~~ Specification of Standards for UK Stereo with Television Transmissions

Rev. 4

1. INTRODUCTION

This draft specification expresses the current BBC/IBA view as to the optimum characteristics of a system for providing two additional sound channels with System I television. It is offered as a basis for an agreed UK Standard.

Basic Starting Points

Certain factors were assumed from the start, namely:

1. In order to ensure a generous margin between the constraints of ruggedness in difficult reception areas and compatibility between the new and existing services, the development of the specification draws heavily on the results of tests carried out by the BBC involving transmissions from the Wenvoe and Crystal Palace transmitters, and associated laboratory investigations.
2. The sound coding system to be employed for UK DBS will be in accordance with the C-MAC/packet specification as published by the EBU (Doc. SPB 284), and the two-channel terrestrial service should employ a sound coding system which maximises the opportunity for commonality in DBS/terrestrial receiver circuits.
3. The specification embodies all of the features which may be provided at the start of the service. It allows some scope, however, for the addition of further features in a compatible fashion in the future, and these may be introduced as the need arises.

Services to be Provided

The data conveyed within the proposed signal comprises two high quality sound channels together with associated control information and a certain amount of additional data capacity.

One or both of the two sound channels could in future be used to carry data instead. Although it is not necessary to specify this data now, it is essential to give an indication of the use to which the channels are being put so that first generation receivers can be arranged to react appropriately when the data is broadcast. Thus the available options will be signalled by three control bits coded as defined in Section 2.2.2.2 and Table 1.

2. BASEBAND CHARACTERISTICS

2.1 Baseband Format

2.1.1 Frame Structure

The transmitted serial data stream is partitioned into 728-bit frames which are transmitted continuously without gaps. One frame is transmitted every millisecond; the overall bit-rate is thus 728 kbit/s made up as follows:

8-bit Frame Alignment Word	8 kbit/s	(see Section 2.2.1.)
5 bits for Control Information	5 kbit/s	(see Section 2.2.2.)
11 bits for Additional Data	11 kbit/s	(see Section 2.2.3.)
704 sound, parity or data bits	704 kbit/s	(see Section 2.2.4. and 2.2.5.)
Total:	728 kbit/s	

Diagrams of the frame structures for conveying stereo and mono sound signals are shown in Fig. 1. The 720 bits which follow the frame alignment word form a structure identical with that of the first-level protected, companded sound-signal blocks in the C-, D- and D2-MAC/packet DBS systems, so that decoding of the sound signals may be performed by the same type of decoder which is used in the above MAC systems.

Frame structures for data services use the same frame-alignment word, flag bit and additional data, with control bits as described in Section 2.2.2.2., but the audio samples are replaced by other data.

2.1.2. Bit Interleaving

Interleaving is applied to the block of 704 bits which follow the frame-alignment word (FAW) control bits and additional bits in order to minimise the effect of multiple-bit errors. The bits of each frame are transmitted in the following order:

FAW	5 Control Bits $C_0 \rightarrow C_4$	11 Additional Data Bits		704 bits of interleaved sound data 16-bits
1,2,3,4,5,6,7,8	9,10,11,12,13	14,15,16,17,18,19,20,21,22,23,24		25,69,113,157........................685
				26,70,114..............................686
				27,71,11.................................687
				28,72,116..............................688
			44 bits	— — — .... —
				— — — .... —
			4 x 11 bit companded samples	— — — .... —
				— — — .... —
				68,112,156............................728

The above interleaving pattern places data bits which are adjacent in the frame structure of Fig. 1 in positions at least 16 clock periods apart in the transmitted bit stream (i.e. at least 15 other bits occur between bits which are adjacent in Fig. 1).

2.1.3. Energy Dispersal Scrambling

The transmitted bit-stream is scrambled for spectrum-shaping purposes. The scrambling is done synchronously with the multiplex frame. The framing code is not scrambled, and is used to synchronise the pseudo-random sequence generator used for descrambling in the receiver. The other parameters are as follows:

(i) the bit which immediately follows the synchronisation word is the first scrambled bit;

(ii) the bit which immediately precedes the synchronisation word is the last scrambled bit;

(iii) scrambling takes place after interleaving (and descrambling is therefore prior to de-interleaving at the receiver);

(iv) the pseudo-random sequence is defined by the following generator polynomial and initialisation word.

Generator Polynomial: $x^9 + x^4 + 1$

Initialisation word: 1 1 1 1 1 1 1 1 1

The diagram for a possible generator for this sequence is given in Fig. 2. Thus the sequence starts: 0000 0111 1011 1110 0010.

2.2 Coding of Information

2.2.1. Frame-Alignment Word

The frame-alignment word is 01001110, the left-most bit being transmitted first.

2.2.2. Control Information

The control information is conveyed by a frame flag bit, C_0, three application control bits, C_1, C_2 and C_3 and, a reserve sound switching flag C_4, (see Fig.1).

2.2.2.1 The Frame Flag Bit

The frame flag bit, C_0, is set to '1' for 8 successive frames and to '0' for the next 8 frames; thus it defines a 16-frame sequence. The frames are numbered within the sequence as follows: the first frame (Frame no. 1) of the sequence is defined as the first of the 8 frames in which C_0 ='1'; hence the last frame (Frame no. 16) of the sequence is the last of the 8 frames in which C_0 = '0'. This frame sequence is used to synchronise changes in the type of information being carried in the channel.

2.2.2.2 The Application Control Bits

The last 704 bits in each frame may be used to convey sound samples, data or a mixture of both. The current application of these bits is defined by the three application control bits, C_1 C_2 and C_3, as indicated in Table 1 below.

These control bits can change to signal a new application on Frame 1 only of the 16-frame sequence. The 704-bit sound/data blocks then change to the new application on Frame 1 of the following 16-frame sequence.

Table 1

Applications of 704-bit sound/data blocks

Application Control Information C_1	C_2	C_3*	Contents of 704-bit sound/data block
0	0	0	Stereo signal comprising alternate A-channel and B-channel samples.
0	1	0	Two independent mono sound signals transmitted in alternate frames (designated M1 and M2).
1	0	0	One mono signal and one 352 kbit/s transparent data channel transmitted in alternate frames.
1	1	0	One 704 kbit/s transparent data channel.

2.2.2.3 The Reserve Sound Switching Flag

Digital sound decoding equipment may be arranged so that it can switch the output of the conventional f.m. sound demodulator to replace the sound decoded from the digital signal in the event of the failure of the latter. Switching to the output of the f.m. demodulator is, of course, acceptable only if the f.m. carrier is modulated with the same sound programme as the failing digital signal; the means to inhibit such switching is incorporated in the control information.

A fifth control bit C_4, is set to "1" when the f.m. signal is carrying the same sound programme as the digital stereo signal, or the digital mono signal (only the "M1" mono signal in the case where two digital mono signals are being transmitted). When the f.m.signal is not carrying the same programme as the digital sound signal, the switching flag is set to "0". In this state it can be used to prevent switching to the f.m. sound.

2.2.3. Additional Data

Eleven additional data bits AD_0 to AD_{10} (see Fig. 1) are reserved for future applications yet to be defined.

2.2.4. The Sound/Data Block

The last 704 bits in any frame form a block of either sound or data information. The two types of information are not mixed within one frame. 64 sound samples (D_1 to D_{64})are transmitted. Fig. 1(a) shows the structure of a stereo sound frame, and Fig. 1(b) shows the mono sound frame.

If a stereo pair of sound signals is being transmitted ($C_1 = C_2 = C_3 =$ '0'), the odd-numbered samples (D_1, D_3, ..., D_{63}) are used to convey the A-channel, and the even-numbered samples (D_2, D_4, ..., D_{64}) the B-channel. Thus 32 samples of each channel are transmitted in every frame.

* $C_3 = 1$ provides for signalling additional sound or data coding options

(C_1 = '0', C_2 = '1', C_3 = '0'), M_1 is transmitted in odd-numbered frames, and M_2 in even-numbered frames (Section 2.2.2.1.). If one mono sound signal is being transmitted (C_1 = '1', C_2 = '0', C_3 = '0'), it is transmitted in odd-numbered frames and data is transmitted in even-numbered frames. Thus, for mono sound signals, each frame with sound information in it will contain 64 consecutive sound samples, which will span 2 complete companding blocks, shown as blocks n and (n + 1) in Fig. 1(b).

No format has currently been defined for data information.

2.2.5 Sound Signals

2.2.5.1. Near-Instantaneous Companding

Sound signals are sampled at 32 kHz and coded initially with a resolution of 14 bits per sample. For transmission, the number of bits per sample is reduced to 10, using near-instantaneous companing, and one parity bit is added to each 10-bit sample word for error detection and scale-factor signalling purposes.

The near-instantaneous compression process forms the 14-bit digital samples corresponding to each of the sound signals into blocks of 32. All of the samples in each 1 ms block are then coded, using a 10-bit two's complement code, to an accuracy determined by the magnitude of the largest sample in the block, and a scale-factor code is formed to convey the degree of compression to the receiver. Fig. 3 illustrates the coding of companded sound signals.

Pre-emphasis to CCITT Recommendation J.17 is applied to the sound signals prior to compression, either by using analogue pre-emphasis networks prior to digitisation or by using digital filters with the digital signals.

For stereophonic transmission the right and left signals are sampled simultaneously, the A samples convey the sound signal to be reproduced by the left-hand loudspeaker and the B samples convey the sound signal to be reproduced by the right-hand loudspeaker.

Table 2

Summary of Sound Coding Characteristics

Sampling frequency:	32 kHz
Initial Resolution:	14 bits/sample
Companding characteristics:	near-instantaneous, with compression to 10 bits/sample in 32-sample (1 ms) blocks
Coding for compressed samples:	2's complement (see Fig. 3)
Pre-emphasis:	CCITT Recommendation J.17 (6.5 dB attenuation at 800 Hz)
Audio overload level:	+ 12 dBmO

2.2.5.2. Error Protection for Sound Signals

One parity bit is added to each 10-bit sound sample, to check the six most significant bits for the presence of errors. The parity group thus formed is even (i.e. the modulo-2 sum of the six protected sample bits and the parity bit is zero). Subsequently, the parity bits are modified to signal the 3-bit scale-factor word associated with each sound signal block (see Section 2.2.5.3.).

Table 3 shows the coding ranges and protection ranges associated with each 3-bit scale-factor word. The five coding ranges indicate the degree of compression to which the block of samples has been subjected, for the near-instantaneous companding process.

In addition to signalling the coding range, the scale factor signals seven protection ranges. This information may be used in the receiver to provide extra protection for the most significant bits of the samples.

Table 3

Coding ranges	Protection ranges	Scale factor value R_2 R_1 R_0
1st range	1st range	1 1 1
2nd range	2nd range	1 1 0
3rd range	3rd range	1 0 1
4th range	4th range	0 1 1
5th range	5th range	1 0 0
5th range	6th range	0 1 0
5th range	7th range	0 0 1
5th range	7th range*	0 0 0

2.2.5.3. Scale-Factor Signalling in Parity for Sound Signals

The three-bit scale-factor R_2, R_1, R_0 (see Table 3) associated with each sound signal is conveyed by modification of the parity bits, in the samples used to convey that sound signal.

When a stereo sound signal is being sent, let FE1** be the scale-factor word R_{2A}, R_{1A}, R_{0A}, associated with the A samples, and FE2 the scale-factor word R_{2B}, R_{1B}, R_{0B}, associated with the B samples. Now if P_i is the parity bit of the i^{th} sample, then this is modified to P'_i, by modulo-two addition of one bit of one of the scale-factor words according to the following relationship:

* It would be possible to add a further protection range, however the last scale factor code indicates "7th protection range" (not 8th) in order to maintain the maximum commonality with the EBU C-MAC/packet system.

** The initial letters "FE" (facteur d'echelle) for scale-factor have been used to conform with the EBU Specification SPB 284.

$P'_i = P_i \oplus R_{2A}$ for i = 1, 7, 13, 19, 25, 31, 37, 43, 49
$P'_i = P_i \oplus R_{1A}$ for i = 3, 9, 15, 21, 27, 33, 39, 45, 51
$P'_i = P_i \oplus R_{0A}$ for i = 5, 11, 17, 23, 29, 35, 41, 47, 53

$P'_i = P_i \oplus R_{2B}$ for i = 2, 8, 14, 20, 26, 32, 38, 44, 50
$P'_i = P_i \oplus R_{1B}$ for i = 4, 10, 16, 22, 28, 34, 40, 46, 52
$P'_i = P_i \oplus R_{0B}$ for i = 6, 12, 18, 24, 30, 36, 42, 48, 54

When a mono sound signal is being sent, FE1 is the scale-factor word R_{2n}, R_{1n}, R_{0n} associated with the first block of 32 samples in the frame, and FE2 is the scale-factor word R_{2n+1}, R_{1n+1}, R_{0n+1} associated with the second block of 32 samples in the frame. As in the case of stereo sound, the parity bit of the i^{th} sample (P_i) is modified (to P_i') by modulo-two addition of one bit of one of the scale-factor words. However, the modification of the parity bits in the mono case relates to the block structure of the mono signal, as follows:

$P'_i = P_i \oplus R_{2n}$ for i = 1, 4, 7, 10, 13, 16, 19, 22, 25
$P'_i = P_i \oplus R_{1n}$ for i = 2, 5, 8, 11, 14, 17, 20, 23, 26
$P'_i = P_i \oplus R_{0n}$ for i = 3, 6, 9, 12, 15, 18, 21, 24, 27

e

$P'_i = P_i \oplus R_{2n+1}$ for i = 28, 31, 34, 37, 40, 43, 46, 49, 52
$P'_i = P_i \oplus R_{1n+1}$ for i = 29, 32, 35, 38, 41, 44, 47, 50, 53
$P'_i = P_i \oplus R_{0n+1}$ for i = 30, 33, 36, 39, 42, 45, 48, 51, 54

(Note: some of the scale-factor information for the second block of samples is conveyed in the parity coding of samples 28 to 32, which are in the first block. This conforms with the coding in the EBU Specifications for DBS.)

Scale-factor coding range and protection range information are extracted at the decoder by majority decision logic. Subsequently the original parity is restored for the purposes of error concealment.

The 10 parity bits associated with samples 55 to 64 are not used to signal scale-factor (or any other) information.

3. RADIO FREQUENCY CHARACTERISTICS

3.1 Characteristics of the vision and f.m. sound components

These are defined in 'Specification of Television Standards for 625-line System-I Transmission' as published jointly by the BBC and the IBA, with the exception that the ratio between peak vision carrier power and f.m. sound carrier power shall be approximately 10:1.

3.2 Specification of the digitally-modulated carrier

3.2.1. Type of modulation

Differentially encoded quadrature phase shift keying (QPSK),* i.e. four-state phase modulation in which each change of state conveys two data bits.

3.2.2. Differential coding

The input data stream at the modulator is differentially encoded by the following processes (see Fig. 4):

i) Serial to two-bit parallel conversion

The input data stream is formed into bit-pairs by a serial to two-bit parallel convertor

ii) Coding of transmitted phase changes

The amounts of the changes of carrier phase which correspond to the four possible values of the input bit-pairs (A_n, B_n) are:

Input bit-pair		Amount by which the carrier changes phase
A_n	B_n	
0	0	0° (i.e. no change)
0	1	-90°
1	0	-270^0
1	1	-180°

* This type of modulation is also known as 4-phase differentially encoded phase shift keying (DPSK)

where, as indicated in Fig. 4, A_n is the input bit at some arbitrary time, and B_n is the input bit one bit-rate clock period later.

Thus the carrier-phase can dwell in one of four rest-states 90° apart, as illustrated in Fig. 5(a). An input bit-pair will shift the carrier phase into a different rest state by the amount of phase-change assigned to that particular value of bit-pair. The transmitted phase-changes and subsequent carrier rest-states for the input bit-pair sequence 00, 10, 11, and 01 are illustrated in Fig. 5(b).

In the receiver the transmitted data-stream may be unambiguously recovered by determining the phase-changes between one bit-pair and the next.

3.2.3. Bit-rate

(a) 728 kbit/s

(b) the long-term stability is $\pm$ 1 part per million

3.2.4. Carrier Frequency

(a) 6.552 MHz above the frequency of the transmitted vision carrier

(b) the frequency 6.552 MHz may be obtained by multiplying the transmitted bit-rate by 9

(c) the long-term stability of the intercarrier frequency (6.552 MHz) is therefore $\pm$ 1 part per million

3.2.5. Carrier level

The ratio between the peak vision carrier power level and the power level of the modulated digital sound signal is approximately 100:1.

3.2.6. Spectrum of the transmitted digital sound signal

Figure 6 shows the frequency band occupied by the transmitted digital sound signal in relation to that ofthe vision and f.m. sound signal components.

Data-signal spectrum-shaping may be implemented as shown schematically in Fig. 4. The two baseband data-streams (shown as I_n' and Q_n' in Fig. 4, and comprising impulses at the symbol rate of 364 kHz) at the inputs to the modulators are each shaped by a low-pass filter with amplitude frequency response $H_T(f)$, where:

$$H_T(f) = \begin{cases} \cos\frac{\pi f t_s}{2} & \text{if } 0 \leqslant f \leqslant \frac{1}{t_s} \\ 0 & \text{if } f > \frac{1}{t_s} \end{cases} \qquad \ldots\ldots\ldots\ldots\ldots(1)$$

and here:

$$t_s = \frac{1}{364000}\ \text{s}$$

and the filter has a constant group delay for all frequencies $f \leqslant \frac{1}{t_s}$.

This specified transmitter (and ideal receiver) low-pass filter response is illustrated in Fig. 7(a).

For best performance in the presence of random noise the amplitude/frequency response of data-spectrum-shaping filters at the receiver (or the low-pass equivalents of bandpass filtering) should be identical to that at the transmitter, i.e. as given above in Equation (1) and they should have constant group delay. The overall data-channel spectrum shaping would then be 100% cosine roll-off, as illustrated in Fig. 7(b).

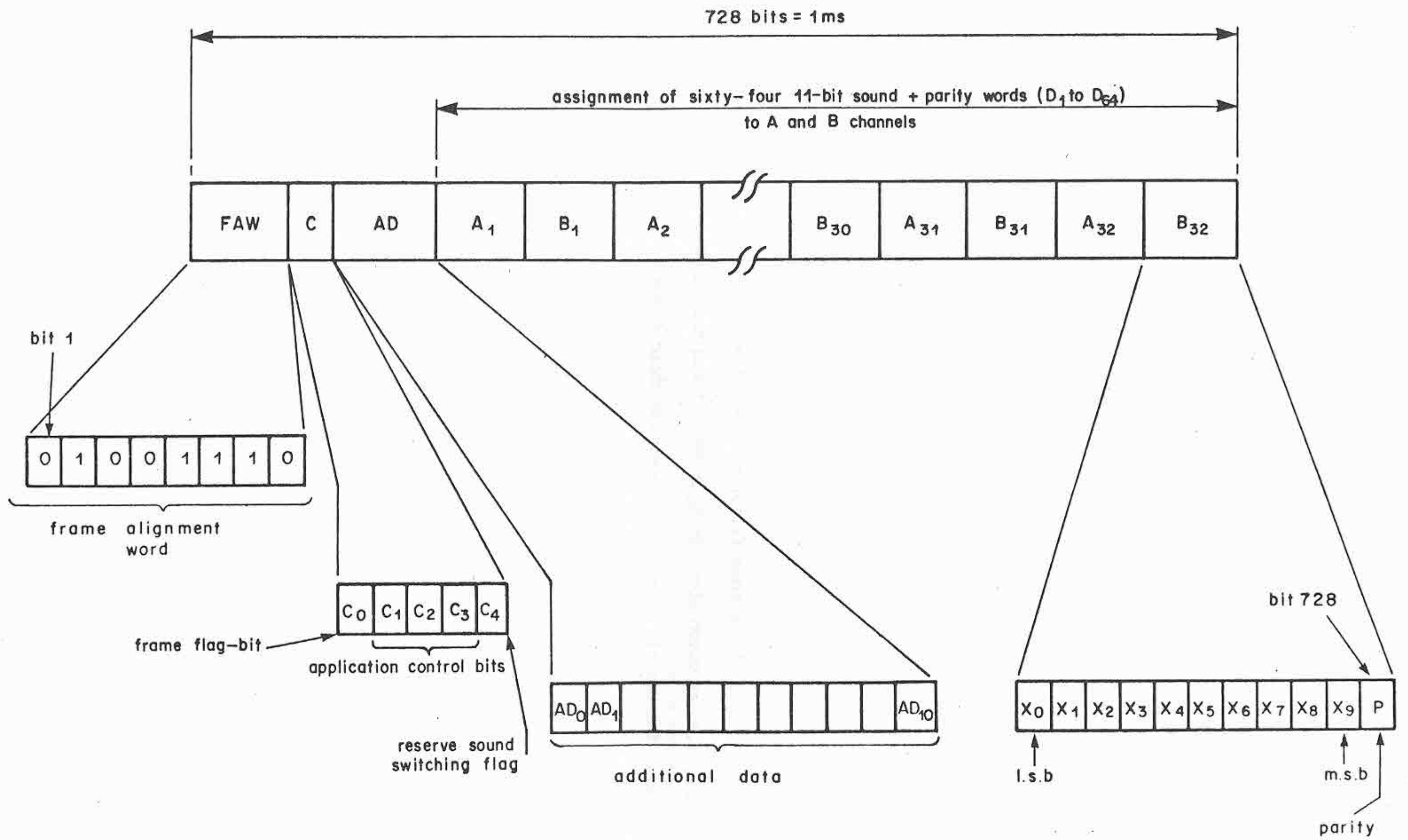

Fig.1 (a) Structure of a 728 - bit frame containing a stereo sound signal (before interleaving)

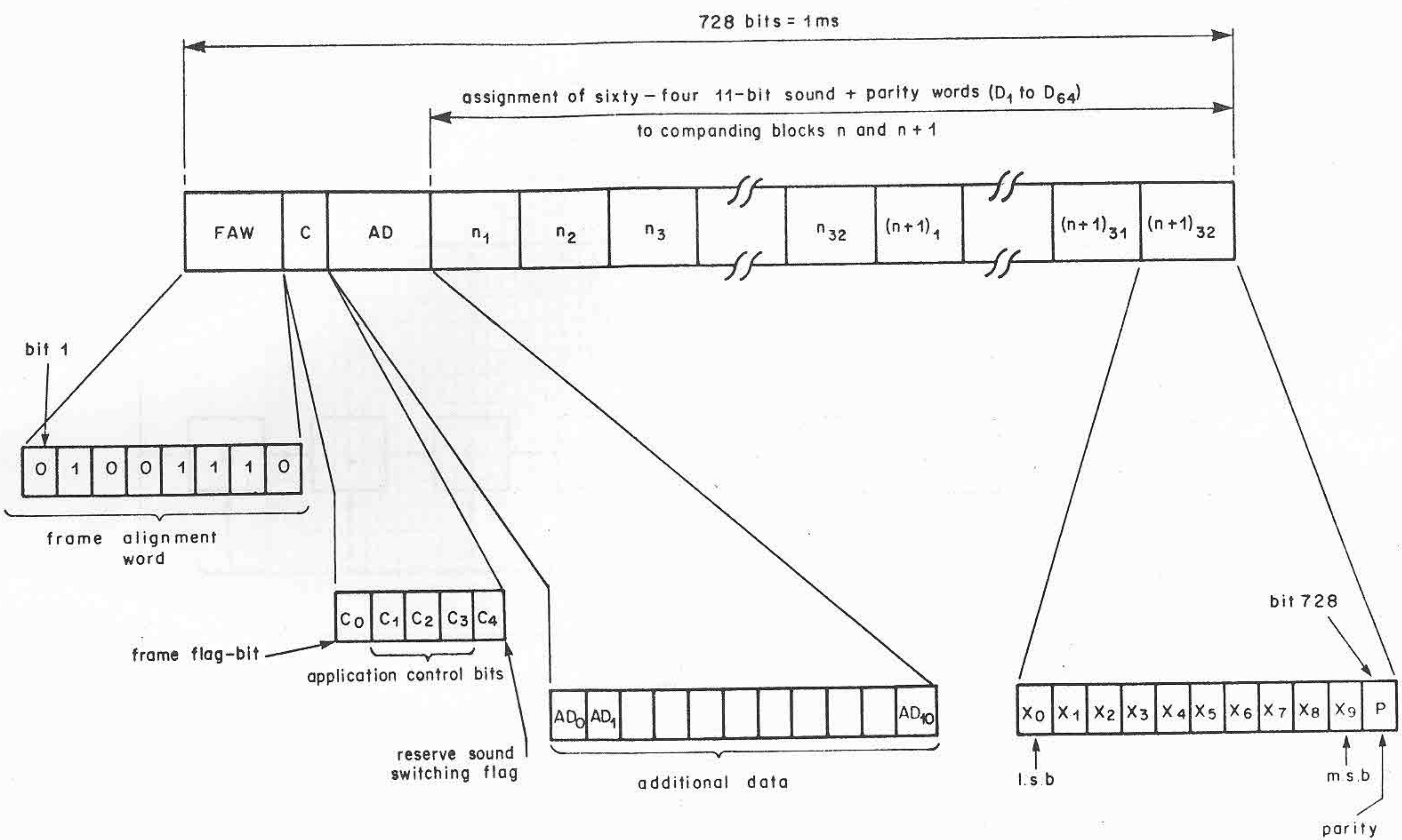

Fig.1(b) Structure of a 728-bit frame containing a mono sound signal (before interleaving)

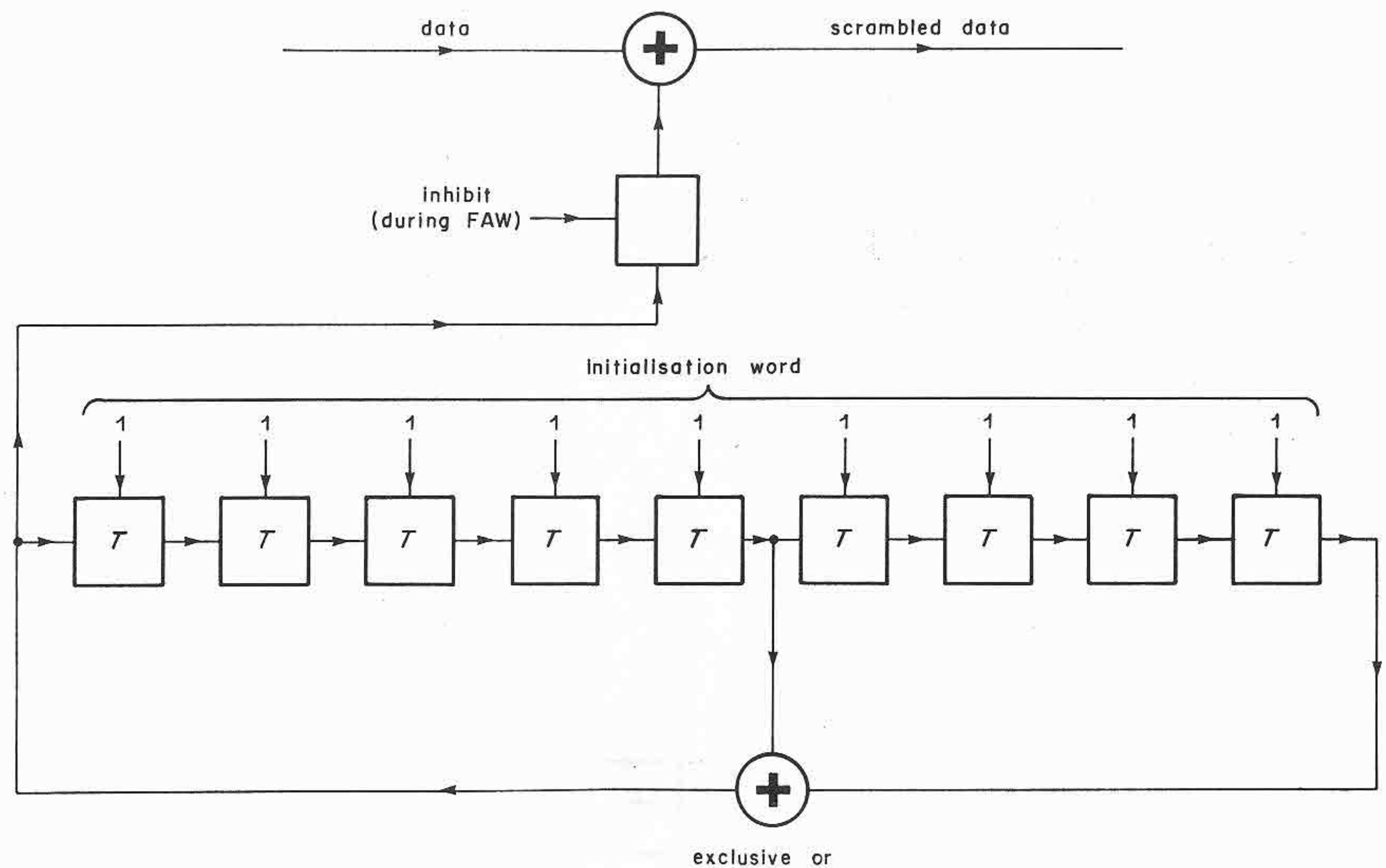

Fig. 2 : Pseudo – random sequence generator for energy dispersal scrambling

MSB													LSB	Scale factor			Coding range	Protection range
0	1	1	1	1	1	1	1	1	1	1	1	1	1					
:	:	:	:	:	:	:	:	:	:	:	:	:	:	1	1	1	1	1
0	1	0	0	0	0	0	0	0	0	0	0	0	0					
0	0	1	X	X	X	X	X	X	X	X	X	X	X	1	1	0	2	2
0	0	0	1	X	X	X	X	X	X	X	X	X	X	1	0	1	3	3
0	0	0	0	1	X	X	X	X	X	X	X	X	X	0	1	1	4	4
0	0	0	0	0	1	X	X	X	X	X	X	X	X	1	0	0		5
0	0	0	0	0	0	1	X	X	X	X	X	X	X	0	1	0		6
0	0	0	0	0	0	0	1	X	X	X	X	X	X					
0	0	0	0	0	0	0	0	1	1	1	1	1	1					
0	0	0	0	0	0	0	X	X	X	X	X	X	X	0	0	1		
0	0	0	0	0	0	0	0	0	0	0	0	0	0	or			5	7
1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0		
1	1	1	1	1	1	1	X	X	X	X	X	X	X					
1	1	1	1	1	1	1	1	0	0	0	0	0	0					
1	1	1	1	1	1	1	0	X	X	X	X	X	X					
1	1	1	1	1	1	0	X	X	X	X	X	X	X	0	1	0		6
1	1	1	1	1	0	X	X	X	X	X	X	X	X	1	0	0		5
1	1	1	1	0	X	X	X	X	X	X	X	X	X	0	1	1	4	4
1	1	1	0	X	X	X	X	X	X	X	X	X	X	1	0	1	3	3
1	1	0	X	X	X	X	X	X	X	X	X	X	X	1	1	0	2	2
1	0	1	1	1	1	1	1	1	1	1	1	1	1					
:	:	:	:	:	:	:	:	:	:	:	:	:	:	1	1	1	1	1
1	0	0	0	0	0	0	0	0	0	0	0	0	0					

X is 1 or 0

Bits used for companded code

Fig. 3 Coding of companded sound signals

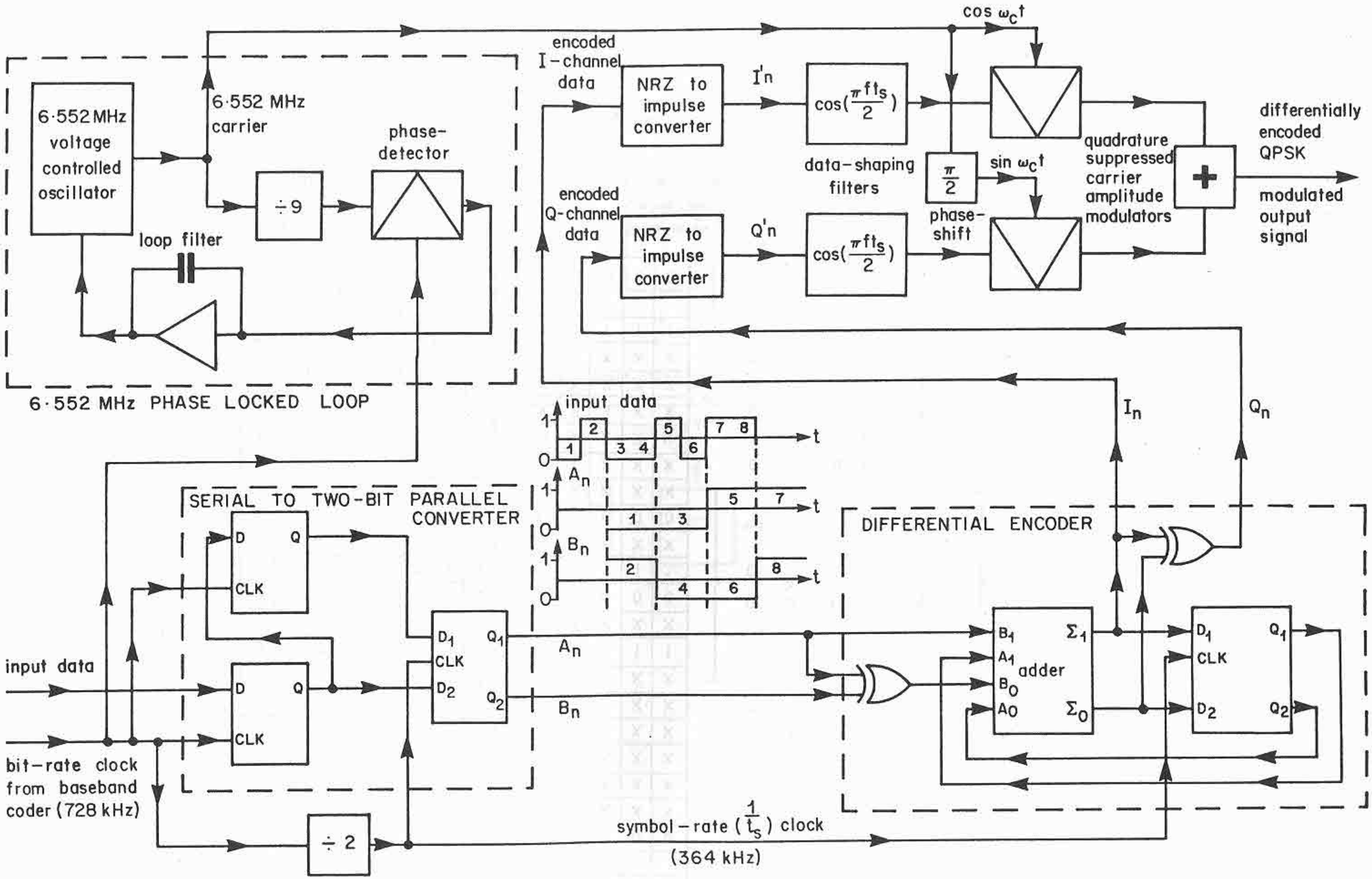

Fig. 4 Block diagram showing the processes of differential encoding, data-signal spectrum shaping and modulation at the transmitter.

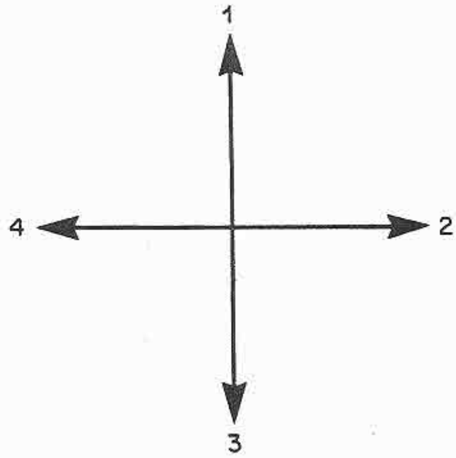

Fig. 5 a Rest states of carrier phase 90° apart

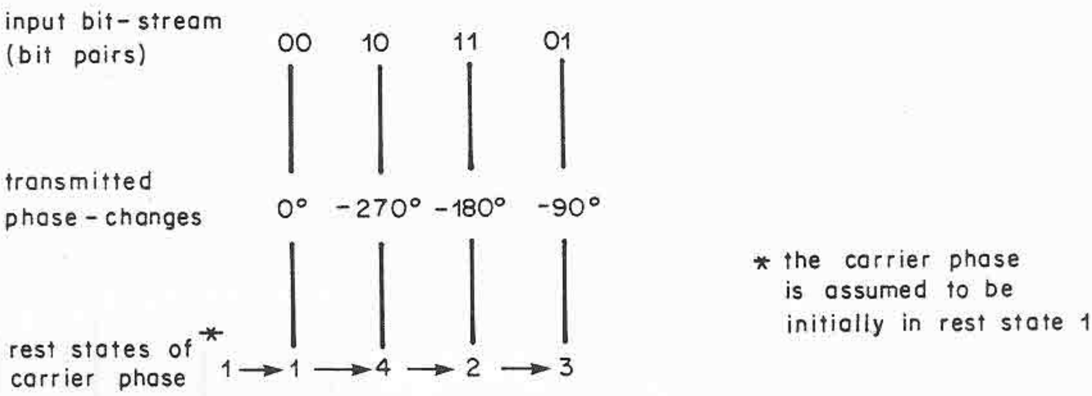

Fig. 5b The transmitted phase-changes and rest states of carrier phase for the input bit-pair sequence 00, 10, 11, 01, assuming the carrier to be initially in rest state 1

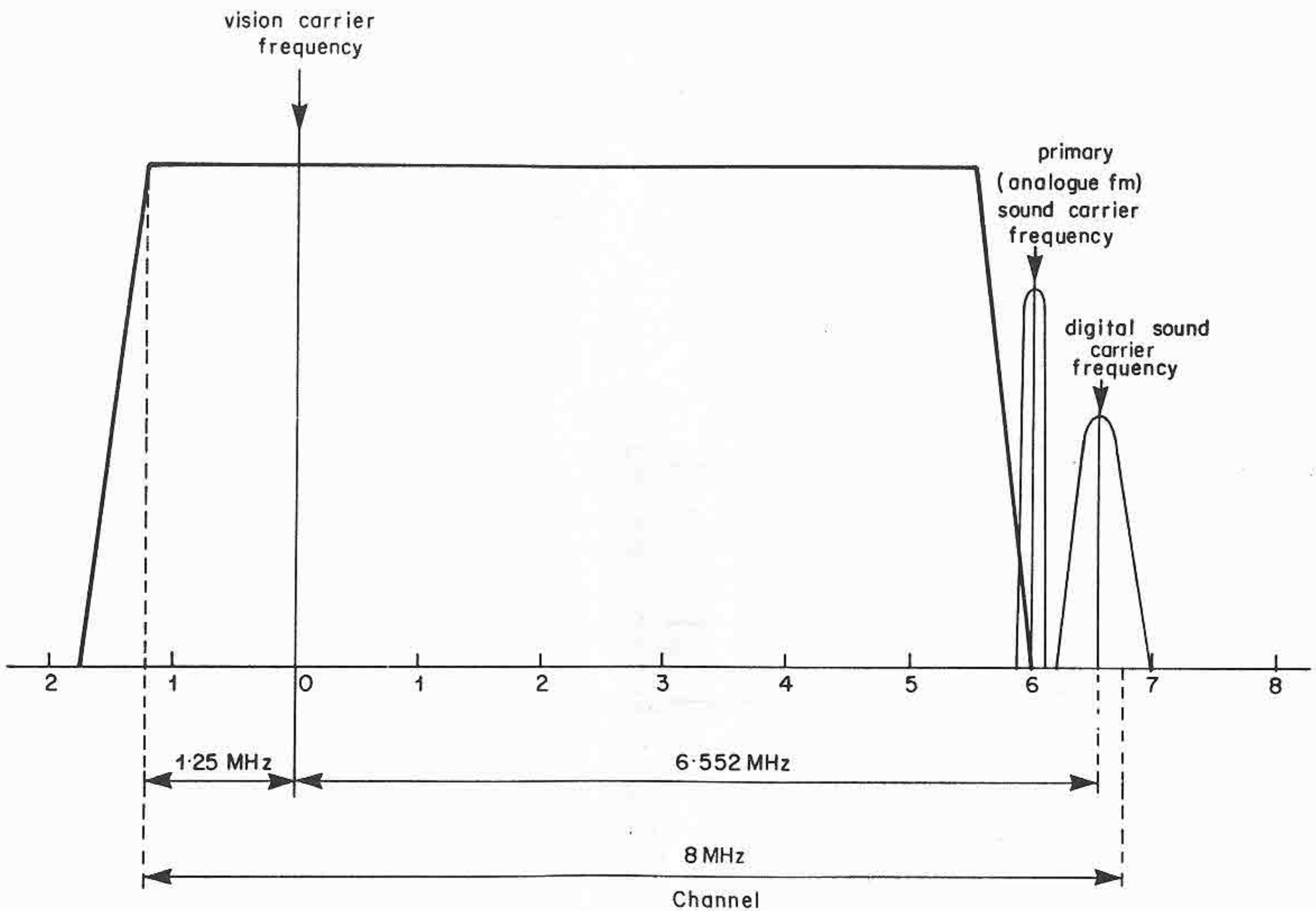

Fig. 6 The frequency bands occupied by the digital sound signal in relation to that of the picture and primary (analogue, fm) sound signal components of the transmitted signal

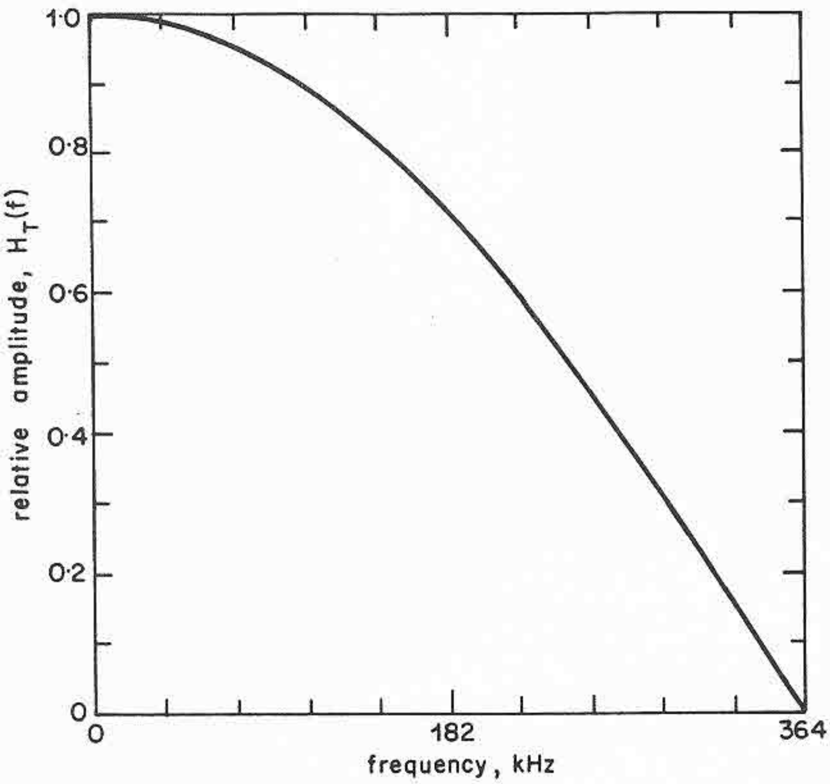

Fig. 7(a) Amplitude response of the specified transmitter (or ideal receiver) data-shaping filter

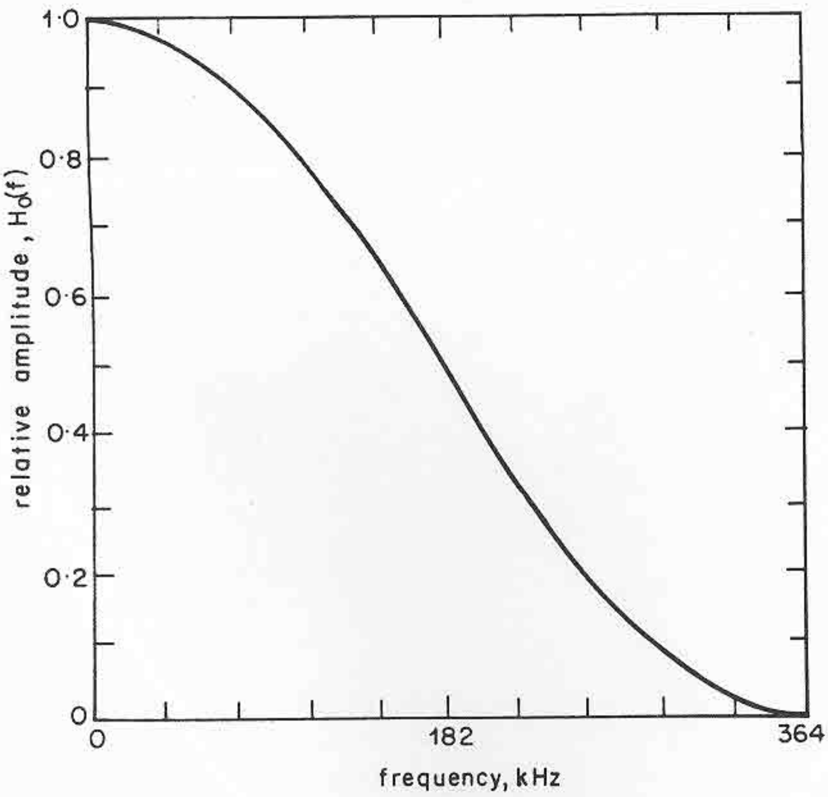

Fig. 7(b) Amplitude response of the combined transmitter and ideal receiver data-shaping filters