A LONG PLAY DIGITAL AUDIO DISC SYSTEM

preprint no. 1442 (G-4)

by

Dr. TOSHI T. DOI, TAKASHI ITOH, HIROSHI OGAWA,
Sony Audio Technology Center, Tokyo, Japan

**Presented at
the 62nd Convention
March 13 through 16, 1979
Brussels, Belgium**

AN AUDIO ENGINEERING SOCIETY PREPRINT

A LONG PLAY DIGITAL AUDIO DISC SYSTEM

By Toshi Tada Doi
Takashi Itoh
Hiroshi Ogawa
(Sony Audio Technology Center, Tokyo, Japan)

Presented at the 62nd Convention
1979 March 13-16
Sheraton Brussels, Brussels Belgium

ABSTRACT

The development of video disc systems has turned out the possibility of a long play digital (PCM) audio disc systems. The band width required for two channel of digital audio signals is less than that of video signals and reduction of the revolution makes the longer playing time be possible. For the further improvement of packing density, a kind of run length limited code is adopted. The code is called as 3PM (three-position modulation), and the packing density can be 150 percent of MFM coding for the same minimum wave length to be recorded. This is achieved at the expence of decreasing jitter margine, which is relatively easy to solve in optical disc systems.

The playing time of two and a half hours is achieved on one side of optical disc with diameter of 30cm. The sampling rate is 44.056kHz and each of two channels is coded by 16 bit linear quantization. The revolution is 450 rpm.

Code errors are analyzed for each revolution of 1800, 900 and 450 rpm on the plane of bit error rate and bit error correlation coefficient. An effective error correcting scheme named "Cross Interleave" is developed which is possible to decode by various levels of decoder from simple erasure type to the complex "Corss Word" type with keeping full compatibility.

1. INTRODUCTION

A long play digital audio disc system has developed after three year studies of increasing packing density. Table 1 shows our steps of development[1],[2],[3], which indicates the playing time on the same size of a disc has increased by five times without decreasing the minimum wave length to be recorded.
The advance is mainly achieved by changing coding method from NRZ-FM (video format) into MFM and then into 3PM[4] accordingly revolution could be reduced as 1800 r.p.m., 900 r.p.m. and 450 r.p.m..

Table 1

Steps for Development of A Long Play
Digital Audio Disc System in Sony

Step	1	2	3
Completed Date	1976 Sep.	1977 Sep.	1978 Sep.
Coding	NRZ-FM (Analogue Video Format)	MFM	3PM
Playing Time (minutes)	30 minutes	60 minutes	150 minutes
Revolution (rpm)	1800	900	450
Track Pitch (μm)	2.5	1.7	1.3
Minimum Wave Length (μm)	1.1	3.4	2.4
Diameter of a Disc (mm)	303φ	303φ	303φ
Sampling Rate (kHz)	44.1	44.1	44.056
No of Channels	2	2	2
Quantization (bits)	13 Non Linear	13 Non Linear	16 Linear
Error Correcting Schemes	Conbination of Random and Burst Error Correction		Cross Interleave System
References		(1), (2)	(3)

The quantization is changed from 13 bit nonlinear into 16 bit linear, so that the quality of the master recording system will be exactly obtained until the final pressed disc to be played at homes.
Another great advance in these developments is that of error correcting shemes. The characteristics of code errors are different by the revolution of the discs, which are analyzed and an effective correcting method is described in this paper.

2. DIGITAL AUDIO DISC SYSTEMS

Fig 1 shows the total system of digital audio discs.
The mastering recorder is PCM-1600 , VTR based, two channel, 16 bit system, and all the signal processing is carried out in digital until the final D/A converters in the player.
For that reason, the sampling rate of the system is chosen as 44.056 kHz in line with PCM-1600.
The process from cutting to the production of discs is not different from that of video disc systems. The player consists of signal pick-up part (disc drive, tracking and focus servo, and he-ne laser tube and its optics) and signal processing part (PLL, frame and bit syncronization, demodulation of 3PM,[4] time base correction, error correction, and D/A). The former is almost the same as video disc players,[5] but as the revolution is 450 r.p.m. the higher stability is required. The latter is described in the following sections.

3. ANALYSIS OF CODE ERRORS

The causes of code errors of optical digital audio disc systems are expected as follows.

(1) defects of pits on discs
(2) defects of metalic coating on discs
(3) bubbles, irregular reflaction or other defects of discs
(4) dusts, scratches, finger-prints on discs
(5) de-focusing
(6) miss-tracking
(7) noise
(8) jitter
(9) fluctuation of RF levels
(10) inter-symbol interference

If the revolution of discs is reduced and the packing density is increased, the length of the code errors caused by (1)~(4) will be increased, but the bit error rate will remain in the same value. But the reduction of the revolution might make the flutter of the disc worse and the code errors by (5),(6) might increase. Besides that, the more the packing density increase, the severer be the code errors caused by (7)~(8), generally.

Design of error correcting shemes deeply depends on the characteristics of code errors and thus the selection of the revolution should be very careful.
In this paper, the analysis of code errors is carried out based on a simple statistical model called "Gilbert Model[7],[8]" which is shown in Fig 3. Where B and G express the states of error and no error, respectively, and the letters T and t are the transition probabilities.
The characteristics of error are expressed by two parameters, the bit error rate p, and the bit error correlation coefficient ξ.

$$p = \frac{T}{T + t} \qquad (1)$$

$$\xi = 1 - T - t \qquad (2)$$

Fig 4 shows an example of the difference of the error characteristics due to the revolution of the discs on the "$p - \xi$ plane". The calculation is carried out by comparing the block error rate before and after error correction. Error correction is chosen as (1) 1 bit error, (2) 2 bit random error, and (3) 15 bit burst error, while the length of the block is between 60 and 75 bits.
The error rate of n-bit-block is expressed as follows.

$$E_{BLK} = 1 - \frac{T}{T + t}(1 - T)^{n-1} \qquad (3)$$

The probability of 1 bit error E_1, 2 bit random error E_2 and that of burst error whthin 15 bits E_{15B} in n-bit-block are expressed as follows.

$$E_1 = \frac{Tt}{T + t}(1 - T)^{n-3}\left\{2(1 - T) + (n - 2)t\right\} \qquad (4)$$

$$\begin{aligned} E_2 &= \frac{Tt}{T + t}(1 - T)^{n-5}\Big[(1 - T)^2\left\{2(1 - t) + T\right\} \\ &\quad + (n - 3)(1 - T)t\left\{2T + (1 - t)\right\} \\ &\quad + \frac{1}{2}t^2T(n - 3)(n - 4)\Big] \end{aligned} \qquad (5)$$

$$E_{15B} = \sum_{i=1}^{15} \frac{Tt}{T + t}(1 - t)^{i-1}(1 - T)^{n-i-1}\left\{2 + (n - i - 1)t\right\} \qquad (6)$$

The data shown in Fig 4 is not meaningful enough, because the discs are produced in various times and in the normal room condition. Usually the quality of discs much depends on production lots rather than the revolution or the error correcting shemes.
If precise data are necessary, a disc including various revolution in various part should be prepared.

Nevertheless, a trend of errors is obviously found in Fig 4, that the value of bit error correlation coefficient decreases as the revolution reduced, and the bit error rate does not significantly depend on the revolution.

4. ERROR CORRECTING SCHEMES

In this system, an effective error correcting sheme named "Cross Interleave Correction" is adopted, which consists of delayed interleave and at least two sets of delayed parity word[9] Decoding is based on pointer-erasure method, where the pointer of the erroneous word is obtained by CRCC. For the further correcting ability, "Cross Word Decoding" is possible with keeping full compatibility to the pointer-erasure decoding.

4-1. ENCODER

Fig 5 shows the format and the encoder. One frame consists of 18 bit synchronization word and three sub-blocks. Each sub-block includes two imformation words (Li, Ri) and two parity words (Pi, Qi) and CRCC.
The encoder consists of simple delay memories, exclusive-or gate, and CRCC encoder. The parities Pi, Qi are composed by the following equation.

$$Pi = Li \oplus Ri \qquad (7)$$

$$Qi = Li \oplus R_{i-3} \oplus P_{i-8} \qquad (8)$$

Where, $\oplus$ means modulo-two summation (exclusive-or).

4-2. THE SIMPLEST BASIC DECODER

Fig 6 shows the simplest decoder, which only utilize the parity words Pi (the parity words Qi are ignored). Each sub-block is checked by CRCC decoder and the error information of 1 bit is fed to the parity P-decoder after giving the same delay as the main imformation words.
In P-decoder, a syndrome word

$$Spi = L'i \oplus R'i \oplus P'i \qquad (9)$$

is calculated. Where the prime mark indicates the received word which might include errors. If Spi = 0, or all the error imformation bits indicate no error, the words L'i and R'i are considered not to include any code errors. If Spi $\neq$ 0, and the error, information bit of L'i indicates error(R'i and P'i are no error), for instance, L'i is supposed to be erroneous, and since the error pattern appears in Spi, the error correction is carried out by the following formula.

$$L_i = L_i' + Sp_i \qquad (10)$$

Error Correction is possible, as far as two of the related sub-blocks for one syndrome are not erroneous. Each related sub-block is separated 16 sub-blocks eachother, therefore a burst error within 16 sub-blocks (1290~1308 bits) can be corrected by this simplest decoder. While a guard space of 32 sub-blocks is necessary for that correction.

4-3. MULTIPLE DECODERS

Fig 7 shows a double decoder, which computes two syndromes Spi, and

$$Sq_i = L_i \oplus R_{i-3} \oplus P_{i-8} \oplus Q_i \qquad (11)$$

Using each syndromes Spi and Sqi, one error word can be corrected by the method described above. Supposing L_1' and R_1' are error, it can not be corrected by the basic decoder shown in Fig 6. But using double decoder (Fig 7), both words can be corrected by using Sqi.

$$Sq_1 = L_1^{x} \oplus R_{-2}' \oplus P_{-7}' \oplus Q_1' \qquad (12)$$

$$Sq_4 = L_4' \oplus R_1^{x} \oplus P_{-4}' \oplus Q_4' \qquad (13)$$

Where the uper script x indicates erroneous word. Supposing L_1, R_1, and R_{-2} are erroneous, two word errors are included both in Sp_1 and Sq_1, nevertheless, error correction is possible by the double decoder as the procedure shown below.

<Step 1: Q-Decoder>

$$Sq_1 = L_1^{x} \oplus R_{-2}^{x} \oplus P_{-7} \oplus Q_1' \qquad (14)$$

$$Sq_4 = L_4' \oplus R_1^{x} \oplus P_{-4} \oplus Q_4' \qquad (15)$$

R_1^{x} is corrected by Sq_4, but L_1^{x} and R_{-2}^{x} are still erroneous.

<Step 2: P-Decoder>

$$Sp_1 = L_1^{x} \oplus R_1' \oplus P_1' \qquad (16)$$

$$Sp_{-2} = L_{-2}' \oplus R_{-2}^{x} \oplus P_{-2}'$$

Since R_1^{x} is corrected by the preceding step, L_1^{x} is corrected by Sp_1, and R_{-2}^{x}, by Sp_{-2}.
Fig 8 shows the four times decoder. Supposing L_1', R_1', L_{-2}', R_{-2}', and Q_{-2}' are erroneous, four times correction is necessary.

<Step 1: Q-Decoder>

$$Sq_{-2} = L_{-2}^{x} \oplus R_{-5}' \oplus P_{-10}' \oplus Q_{-2}^{x} \qquad (18)$$

$$Sq_1 = L_1^{x} \oplus R_{-2}^{x} \oplus P_{-7}' \oplus Q_1' \qquad (19)$$

$$Sq_4 = L_4' \oplus R_1^{x} \oplus P_{-4}' \oplus Q_4' \qquad (20)$$

$R_1^{\times}$ is corrected by Sq_4, but others can not be corrected.

<Step 2: P-Decoder>

$$Sp_1 = L_1^{\times} \oplus R_1' \oplus P_1' \qquad (21)$$

$$Sp_{-2} = L_{-2}^{\times} \oplus R_{-2}^{\times} \oplus P_{-2}' \qquad (22)$$

Since R_1' is corrected at the preceding step, $L_1^{\times}$ can be corrected by Sp_1.

<Step 3: Q-Decoder>

$$Sq_1 = L_1' \oplus R_{-2}^{\times} \oplus P_{-7}' \oplus Q_1 \qquad (23)$$

$$Sq_{-2} = L_{-2}^{\times} \oplus R_{-5}' \oplus P_{-10}' \oplus Q_{-2}^{\times} \qquad (24)$$

$R_{-2}^{\times}$ is corrected by Sq_1.

<Step 4: P-Decoder>

$$Sp_{-2} = L_{-2}^{\times} + R_{-2}' + P_{-2}' \qquad (25)$$

$L_{-2}^{\times}$ is corrected by Sp_{-2}, and error correction is completed.

While if the double decoder (Fig 7) is used for the above example, L_{-2} and R_{-2} are remained as un correctable. This is one of the features of the Cross Interleave System that the more the steps of decoding increase, the better correctability is obtained.
Ofcause there is some probability of un correctable errors, however steps of decoding are prepared. One of the examples is shown below.

$$Sq_1 = L_1^{\times} \oplus R_{-2}' \oplus P_{-7}' \oplus Q_1^{\times} \qquad (26)$$

$$Sq_4 = L_4' \oplus R_1^{\times} \oplus P_{-4}' \oplus Q_4^{\times} \qquad (27)$$

$$Sp_1 = L_1^{\times} \oplus R_1^{\times} \oplus P_1 \qquad (28)$$

In this case, $L_1^{\times}$, $R_1^{\times}$, $Q_1^{\times}$, and $Q_4^{\times}$ are erroneous. All related syndromes, eqs. (26)~(28) include two erroneous words.

4-4. CROSS WORD DECODING (10),(11),(12),(13)

If an information word relates to plural syndromes, cross word decoding is possible by comparing those syndrome patterns. In above format, three types of syndrome Spi, Sqi and residual of CRCC can be utilized.
In the case shown above by eqs. (26)~(28), related syndromes by CRCC, Sci are as follows.

$$Sc_1 = Res(L_1', R_{-15}', P_{-31}', Q_{-34}', C_1') \qquad (29)$$

$$Sc_{17} = Res(L_{17}', R_1', P_{-15}', Q_{-18}', C_{17}') \qquad (30)$$

$$Sc_{36} = Res(L'_{36},\ R'_{20},\ P'_{4}\ ,\ Q'_{1}\ ,\ C'_{36}) \qquad (31)$$

$$Sc_{39} = Res(L'_{39},\ R'_{23},\ P'_{7}\ ,\ Q'_{4},\ C'_{39}) \qquad (32)$$

Where Res( ) means residual of CRCC, and $C'i$ is received CRCC word.
In the case described above as uncorrectable by erasure method, all the equations (26)~(32) are not zero. If the words L^x_1, R^x_1, Q^x_1, and Q^x_4 are really erroneous, they can not be corrected by all means. But the error position in eqs. (29)~(32) can not be indicated and the following error patterns are also considered as uncorrectable.

(i) $L^x_1,\ R^x_1,\ R^x_{20},\ C^x_{39}$

(ii) $L^x_1,\ L^x_{17},\ P^x_4,\ Q^x_4$

If the word error is considered as random, real uncorrectable errors are 1% of uncorrectable errors by above decoders, and 99% will be corrected by the Cross Word Decoder.
The principle of Cross Word decoding is to find the error location by comparing syndromes; for instance, in case (i) the equation

$$Sp_1 = Sq_1 + Sq_4 \qquad (33)$$

will be satisfied, and in case (ii),

$$Sq_1 = Sp_1\ . \qquad (34)$$

There is another method of Cross Word decoding, that is to put the parity words into CRCC decoder, and after appropriate shifting, compare with eqs. (29)~(32).[13] If the error is within 11 bit burst, the error location in eqs. (29)~(32) can be found by this method.

5. 3PM CODE

Among the various class of the run length limited code, 3PM (three-position modulation)[4] is selected because of its relatively simple hardware and fairly good efficiencies. The principle of 3PM, shown in Table 2, is to convert three bit of original data word into six-bit word, in which "1" means transition, and any 1's are always separated by two 0's. In other words, the minimum duration between transitions is 1.5 L, that is 50% larger than that of MFM, where L is the length of original data bit cell.

Table 2 3PM

No	Original Data Word			Transition Positions					
1	0	0	0	0	0	0	0	1	0
2	0	0	1	0	0	0	1	0	0
3	0	1	0	0	1	0	0	0	0
4	0	1	1	0	1	0	0	1	0
5	1	0	0	0	0	1	0	0	0
6	1	0	1	1	0	0	0	0	0
7	1	1	0	1	0	0	0	1	0
8	1	1	1	1	0	0	1	0	0
	|← L →|			|← L →|					

At the junction of the words, another consideration is necessary. If the word No.8 comes after No.1, the transition pattern is

0 0 0 0 1 0   1 0 0 1 0 0

No.1   No.8

, and there is a "1 0 1" pattern which violates above law. In this case, "1 0 1" pattern is changed into "0 1 0" pattern as follows.

0 0 0 0 $\underline{0}$ $\underline{1}$ $\underline{0}$ 0 0 1 0 0

No.1   No.8

In order to allow this nonlinear junction, the last transition in Table 2 is always "0", and the coding is always proceeded watching the preceding and the following words.
Table 3 shows a comparison between 3PM and MFM. The packing density of 3PM can be 50% higher than that of MFM, but jitter margin is 50% worse and the maximum transition is 100% longer. Therefore, bit synchronization and the mechanisms of the player should be designed carefully.

Table 3 Comparison between 3PM and MFM

		MFM	3PM
Duration Between Transition	Minimum	L(1/2λmin)	1.5L(1/2λmin)
	Maximum	2L(λmin)	6L(2λmin)
Jitter Margin		0.5L(1/4λmin)	0.5L(1/6λmin)

L: length of data bit cell

λmin: minimum wave length to be recorded

6. CONCLUSION

A digital audio disc system realizing the playing time of two and a half hours on one side of a 30øcm-disc is described.
The redundancy of the error correcting schemes is 64.2% including synchronizing bits, and transmitting bit rate is 3.568536 Mb/s for two channels of 16 bit signals. But on account of 3PM coding, the maximum recording frequency is 1.189512 MHz, and the minimum wave length at the inside of discs(120ømm)is 2.4μm for 450 r.p.m. revolution.
These values are decided tentatively for experimental systems. For the systems to be comercialized, the redundancy and the minimum wave length might be able to be reduced on account of highly controled processes, and the more playing time or smaller size of the disc will be possible.
Another great problem for comercialization is the compatibility of the players between digital audio and video disc systems. The system described in this paper has that compatibility, but cousequently extremely long playing time is achieved, which might be unnecessary from software point of view. The final standard should be established by mutual studies of various software and hardware makers.

ACKNOWLEDGEMENT
The authors wish to express their thanks to Dr. H.Nakajima, managing director of Sony Corp. and Mr. S.Miyaoka, Mr. T.Ihashi and other members of disc group for their advice and great cooperation in developing this system.

REFERENCES

(1) T.T. Doi; "PCM Audio Disc Systems" preprint for press conference (1977 Sep.)

(2) H. Ogawa, T. Itoh, and T.T. Doi; "A System of Optical Digital Audio Disc", Tech. Rpt. of IECE No.EA78-27 (1978 July)

(3) T. Itoh, H. Ogawa, T.T. Doi, and K. Sekiguchi; "Advanced System of Digital Audio Disc", ASJ Conf. 3-P-13 (1978 Oct.)

(4) G.V. Jacoby, "A New Look-Ahead Code for Increased Data Density", IEEE Trans., Vol.MAG-13, No.5 (1977 Sep.)

(5) Y. Tsuchiya, T.T. Doi, T. Ohtsuki, and S. Kazami; "Professional PCM Audio Unit Using Error Correcting Code", ASJ Conf. No.3-2-10 (1977 Oct.)

(6) H. Nakajima, T.T. Doi, Y. Tsuchiya, and A. Iga; "A New PCM Audio System as an Adapter of Video Tape Recorders", 60th AES Conv. No.1352 (E-11), (1978 May)

(7) G. Fukuda, and T.T. Doi; "On Dropout Conpenration of PCM Systems; Computer Simulation Method and a New Error-Correcting Code", 60th AES Conv. No.1354 (E-7) (1978 May)

(8) T.T. Doi, Y. Tsuchiya, and G. Fukuda; "Statistical Analysis of Error Correcting Schemes of Stationary Head Digital Audio Recorders" Joint Meeting of ASA and ASJ, No.MM13, (1978 Nov.)

(9) T.T. Doi, G. Fukuda, and K. Odaka "Error Correction by Cross Interleave Code", IECE Conf. No.72 (1979 Mar.)

(10) T.T. Doi; "A Class of Error Correcting Code for a Transmission System with High Error Rate", ASJ Conf. No.3-2-14 (1977 Oct.)

(11) T.T. Doi; "A New Error Correcting/Detecting Code for Dropouts of Magnetic Tapes" Nikkei Electronics pp94-113 (1978 3-20)

(12) T.T. Doi; "Cross Word Code; A New Error Correcting Code for Digital Audio Systems" IECE Tech. Rpt. EA78-25 (1978 July)

(13) S. Kazami, T. Ohtsuki, I. Ajimine, T.T. Doi, O. Hamada; "Improvement of Cross Word Code", ASJ Conf. No.3-P-1 (1978 Oct.)

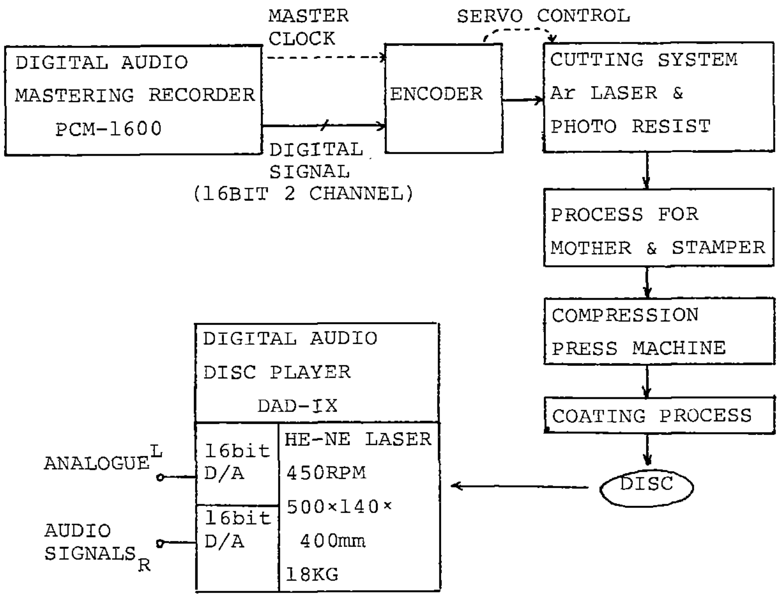
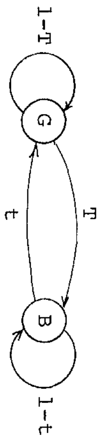


Fig 1. Digital Audio Disc Systems

Fig 2 Digital Audio Disc and the Player DAD-1X

Specifications

General	
Playing time:	Two and a half hours (maximum)
Revolution:	450 rpm
Number of channels:	2 channels, digital
Sampling frequency:	44.056kHz
Quantization:	16-bit linear/channel
Dropout compensation:	Newly developed error correcting code
Dynamic range:	Better than 95dB
Total harmonic distortion:	Better than 0.03%
Frequency response:	2Hz ~ 20kHz±0.25dB
Player system	
Signal detection:	Optical reflection type with a helium-neon laser
Wow and flutter:	Undetectable
Dimensions:	520(W) × 140(H) × 400mm(D)
Weight:	18 kg
Power consumption:	75W
Disc	
Dimensions:	303mm (dia.)
Thickness:	1.1mm
Material:	Polyvinyl chloride, coated with reflection material



G: State of No Error

B: State of Error

$\left.\begin{matrix} T \\ t \end{matrix}\right\}$: Transition Probability

$\gamma = \dfrac{t}{T + t}$: Bit Error Rate

$\varsigma = 1 - T - t$: Bit Error Correlation Coefficient

Fig 3 Gilbert Model

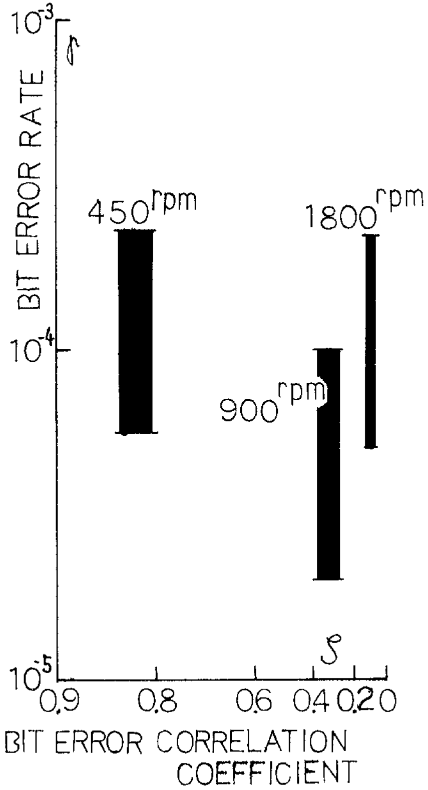
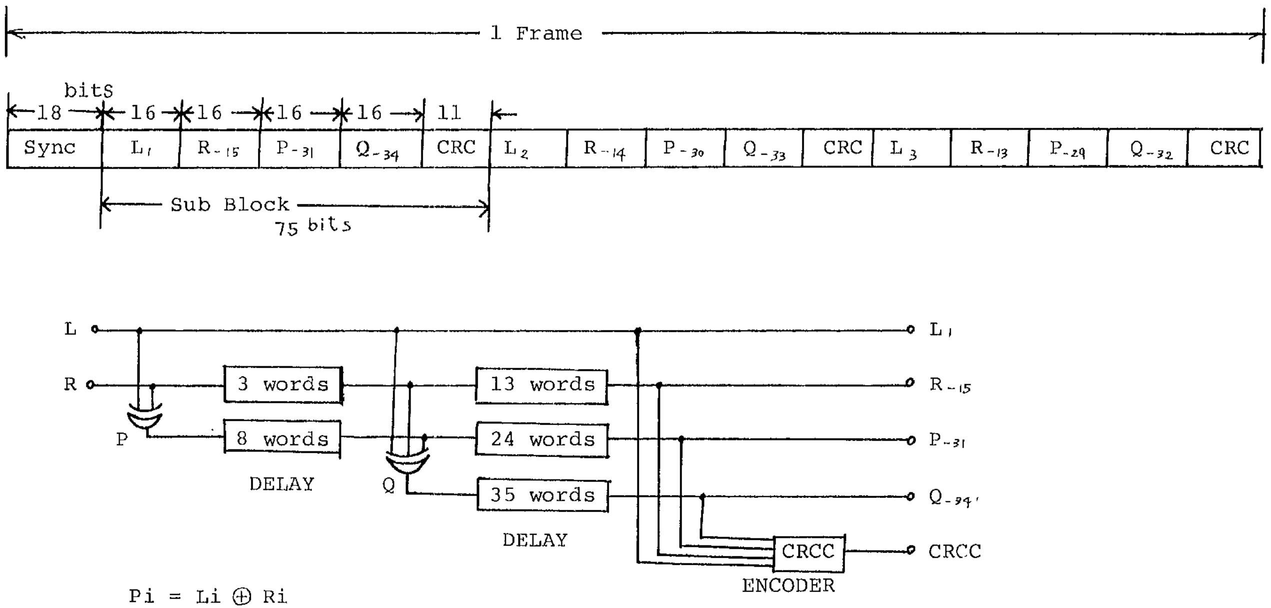


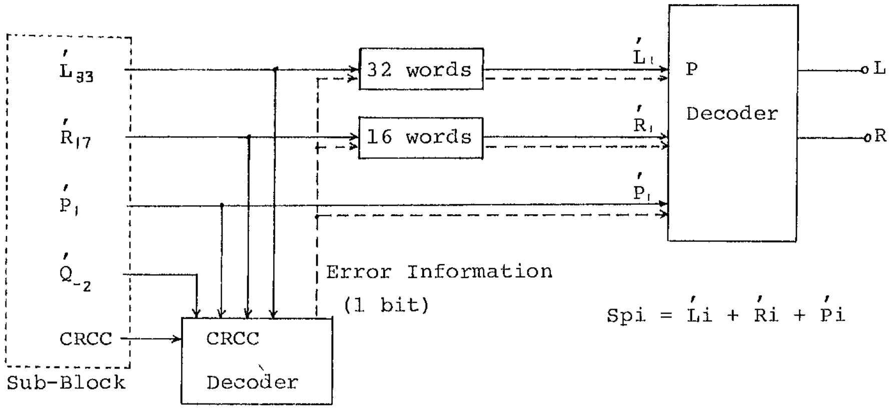
Fig 4 Code Error Analysis on γ-ς plane



$$P_i = L_i \oplus R_i$$

$$Q_i = L_i \oplus R_{i-3} \oplus P_{i-8}$$

Fig 5 Format and Encoder



$$Spi = L'i + R'i + P'i$$

Fig 6 Basic Decoder (Ignoring Parity Q)

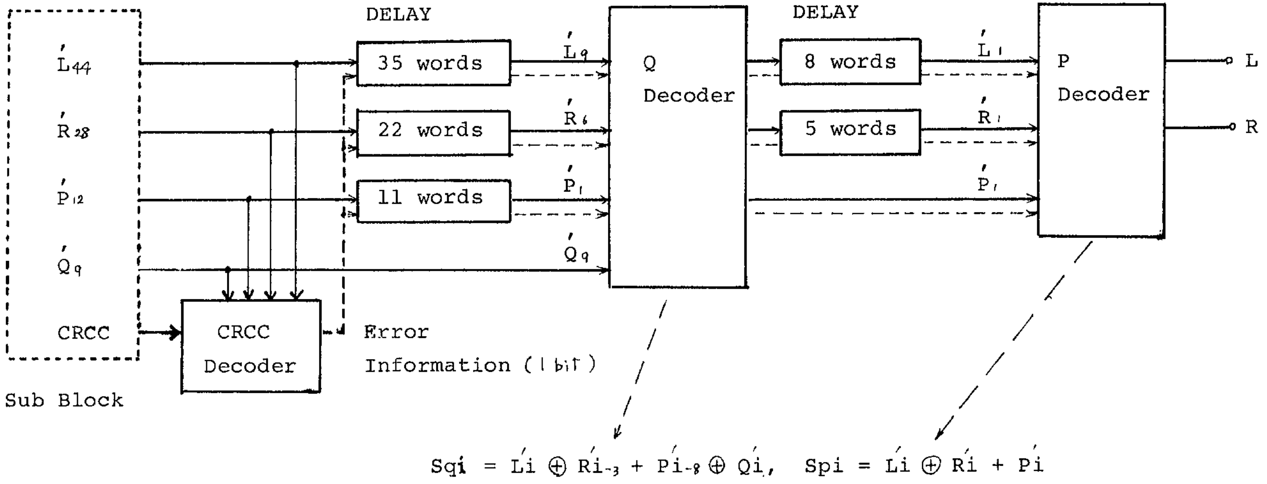


Fig 7 Double Decoder

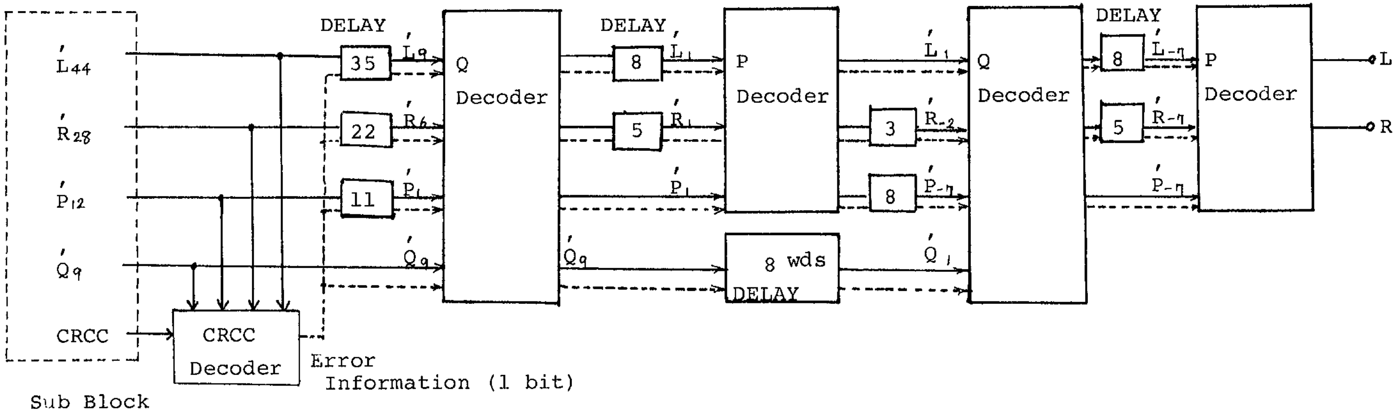


Fig 8 4-Times Decoder