THE DEVELOPMENT OF THE DIGITAL COMPACT CASSETTE SYSTEM

1861 (G-7)

Isao Owaki, Susumu Saito,
Shinji Nakamura
Victor Company of Japan, Ltd.
Yamato, Japan

Presented at the 71st Convention 1982 March 2-5 Montreux

AN AUDIO ENGINEERING SOCIETY PREPRINT

THE DEVELOPMENT
OF
THE DIGITAL COMPACT CASSETTE SYSTEM

Isao Owaki

Susumu Saito, Shinji Nakamura

Victor Company of Japan, Ltd.

Yamato, Japan

Abstract

This paper describes a Digital Compact Cassette System utilizing a regular Compact Cassette. Having two audio channels and operating at a tape speed of 7.14 cm/sec, this system enables recording and playback in both directions of tape travel. It is a compact PCM magnetic recording system equipped with a user's bit of 4.8k Baud in its frame format. The audio signal is quantized at a sampling rate of 33.6 kHz and recorded onto 4 tracks per audio channel with a powerful error correcting code.
This system is designed to operate at a linear density of 46.29k bpi (1.82k bits/mm) and a track density of 158 tpi.

Introduction

Recent applications of digital technology into audio and video equipment are remarkable. Among digital audio recorders which quantize the audio signal and record it onto magnetic tape are an open-reel PCM recorder with stationary heads, a PCM processor to be used together with a VTR, and other systems, on which many articles have already been published [1], [2], [3]. The majority of digital audio recorders reported there, are intended for professional use with one exception being a PCM processor for use with a home VTR.
As professional digital audio equipment becomes more and more popular, the quality of recorded master tapes has also been improved. From this, a demand has arisen for higher-quality consumer digital audio equipment, in particular for a digital audio tape recorder which is compact for portability and operates with stationary heads.
With this in mind, we have developed a Digital Compact Cassette System to be used with a regular Compact Cassette; our system is small in size and easy in operation with its tape transport mechanism greatly simplified.

1.Basic concept in development

When we started this project, we were convinced that digital recordings in the form of compact cassettes would give music lovers higher sound quality and more advanced functions than ever. Also both hardware and software would be able to be supplied at a reasonable price in the future.
In particular, our basic considerations were:

(1) The recording time should not be less than 1 hour.
As a music medium, the recording time of 1 hour is a minimum requirement.

(2) Recording should be possible on both sides of a cassette.
The feature of Compact Cassettes is not only their compactness, but their operability; i.e. they can be used in both directions.

(3) A tape transport mechanism should be as simple as possible.

(4) Most emphasis should be placed on reliability of recording and playback operation.
To realize this, sufficient error correcting capability must be provided against errors caused by drop-outs, etc.

(5) Mass-production of prerecorded tapes should be possible using a high-speed duplication system.

When we develop equipment which records digital signals onto a magnetic medium and plays them back, we meet a "circle" as shown in Fig. 1. The Digital Compact Cassette System needs high packing density recording. Considering consumer use, it should be designed within the limitations: lower speed, narrower track width and therefore lower S/N ratio. Keeping these factors in mind, first we investigated thoroughly the characteristics of a magnetic recording and playback system and then determined the parameters such as specifications of analog channels, tape speed, linear density, track density, redundancy rate, and error correcting capability so that they would be well balanced as a total system.
Table 1 shows the performance and structural specifications of the Digital Compact Cassette System.

2. Tape speed

Generally, consumer products are likely to be used under more severe environments than professional products. Also an ultra-high precision tape transport mechanism cannot be expected for consumer audio recorders. Therefore, reduction of the tape thickness would lead to errors and drop-outs because a thin tape is apt to be deformed (wavy edge or cupping) and is subject to residual elongation. At present, three types of tapes, C-60, C-90 and C-120, are available for Compact Cassettes. The typical thickness of each tape is 18, 12 and 9 μm respectively. However, in practical use, considering durability of tape the thickness of the C-90 cassette is the lower limit.
Because of its determined dimensions, the Compact Cassette sets a limit to the length of tape. With C-90 tape, about 135 m can be accommodated. To allow a recording of 30 minutes in one direction, our system uses a tape having the thickness of C-90 tape as a standard. From this, the upper limit of tape speed was determined to be 7.14 cm/sec, 1.5 times that of the analog Compact Cassette recorder.

3.Audio channel frequency response and sampling rate

As digital audio technology progresses, many studies have been made on the upper band limit frequency of the audio signal in connection with selection of sampling rates and several reports are available. It is natural that an increase of the sampling rate leads to an increased transmission rate, which requires a large amount of packing density. Table 2 shows variations of total transmission rates and transmission rates per track relative to variations of sampling rates with a redundancy rate of 0.25.

It is reported [4], [5] that, as far as consumer audio products are concerned, 15 kHz is sufficient as the upper limit frequency of the audio signal. Based on this theory, we decided to employ 15 kHz as the upper limit frequency of an audio channel. As the sampling rate, we considered two points: production of low-pass filters should be simple and the sampling rate of this system should be proportional to that of professional digital audio recorders at a simple integer ratio. So we established 33.6 kHz, corresponding to 16/21 of 44.1 kHz, the sampling rate of the professional system.

4. Track format

As shown in Table 2, increasing the number of tracks per audio channel reduces the transmission rate per track, resulting in reduction of linear density. However, an increased number of tracks means a reduced track width; this would require a highly accurate, complicated tape transport mechanism and more dimensional accuracy of tape width in production. As a result, the system would be subject to more degradation of S/N ratio and errors due to even the minutest dust particles. In our system, it is difficult to increase the number of tracks per audio channel because both sides A and B should be able to be used and so high an S/N ratio cannot be expected with such a low tape speed.
On the other hand, if several problems are considered which would occur in the production of a coil-wound type multi-track head of single crystal ferrite, the possible range of the track width may be 150 to 200 μm. Consequently, we adopted a track format having 8 tracks for 2 audio channels, with the track width (Tw) of 160 μm and the track spacing (Ts) of 240 μm, plus an auxiliary track for information so that new features possible only with digital systems such as address for random access, are available. As a result, the head has a 9-track configuration for each recording direction of the tape. Fig. 2 shows the track format.
Because a single head is used for both recording and playback in this system, considering the amount of recording current and core saturation, the gap length (ℓ) was determined to be 0.4 to 0.5 μm.

5. Transmission rate

It is reported that the limit of packing density in digital magnetic recording is 10^8 to 10^9 bits/inch^2 [6]. However, this is a theoretical value completely free from any mechanical defect. In practice, it is impossible to be free from limitations resulting from interaction among the medium, head, mechanism and circuitry [7]. The most typical losses involved in magnetic recording are gap loss, spacing loss and thickness loss.

Particularly, in regards to the gap loss (Lg) during playback, the equation is given as follows:

$$L_g = \frac{\sin(\pi \cdot \ell/\lambda)}{\pi \cdot \ell/\lambda} \cdot \frac{5-4(\lambda/\ell)^2}{4-4(\lambda/\ell)^2}$$

where λ is the wavelength to be recorded.
Assuming that ℓ = 0.4 to 0.5 μm, the above equation means that the signal with a wavelength λ = 0.45 to 0.56 μm cannot be played back. At a tape speed of 7.14 cm/sec (v), fg (=v/λ)is 158.7 to 127.5 kHz. It will be assumed that fg is the upper limit frequency of this recording and playback system.
Fig. 3 shows the recording/playback frequency response before equalization of this system.
With everything considered, we determined the transmission rate per track to be 130.2k bits/s, with a rolloff factor of 1 taken into account.

6. Modulation/Demodulation [8], [9], [10]

A digital audio recorder using a stationary multi-track head requires a number of modulation and demodulation circuits corresponding to the number of tracks. Therefore, if the recorder is designed for consumer use, its modulation and demodulation system and circuitry should be as simple as possible.
The process of reading out signals using a coil-wound type head offers a differentiation characteristic. In order to make effective use of the limited bandwidth, we adopted a partial response system (1,-1) taking full advantage of the differentiation characteristics and intersymbol interference. In this system, a stream of data is transmitted to the precoder 1/(1-D), then recordings are made on a tape medium. It is subject to three-level detection in playback. This system does not require any special demodulator. And also it is characteristic in that there is no DC component in the frequency spectrum. Therefore, in the vicinity of rolloff factor 1, the timing margin of the eye pattern is less subject to variations.
Fig. 4 shows computer-simulated eye patterns with rolloff factors 1 and 0.7.
Fig. 5 shows the unequalized isolated readback pulse during playback, isolated readback pulse after equalization and eye pattern before detection.

7. Frame configuration

Fig. 6 shows the frame configuration of our system.
Each frame consists of a sync word (SYNC) of 8 bits, a user's bit, 16 data words (W) or error correction words (P,

Q) of 12 bits and a CRC word of 16 bits; in total 217 bits. The CRC word represents a residual polynomial obtained by dividing the information polynomial from the user's bit to the data word W or error correction word P,Q by the generator polynomial $G(X) = X^{16} + X^{12} + X^{5} + 1$ (CRC-CCITT). Here X is an indeterminant. The user's bit can be used for literal information on the recordings such as a liner-note of music.
The frame frequency is 600 Hz and one frame corresponds to the tape length of 119 μm.

8. Error correction [11]

In recording, the error correction words P and Q are generated from the data word W through encoding algorithm and recorded onto tape according to the frame configuration shown in Fig. 6.
In playback, the resulting words W, P, Q and CRC are subjected to error detection operation so that an error pointer is labeled with W, P and Q.
Word errors in the magnetic recording and playback system are corrected by the decoding circuit using error pointers and encoding algorithm.

(1) Encoding

Fig. 7 shows the encoding diagram according to the format of this system. The generator algorithm for the error correction words P and Q is as follows:

$$P_m = W_m \oplus W_{m+1} \oplus W_{m+2} \oplus W_{m+3} \oplus W_{m+4} \oplus W_{m+5} \oplus W_{m+6} \oplus W_{m+7} \oplus W_{m+8} \oplus W_{m+9} \oplus W_{m+10} \oplus W_{m+11} \oplus W_{m+12} \oplus W_{m+13} \quad \text{----} (1)$$

$$Q_m = D^{0\cdot i} W_m \oplus D^{1\cdot i} W_{m+1} \oplus D^{2\cdot i} W_{m+2} \oplus D^{3\cdot i} W_{m+3} \oplus D^{4\cdot i} W_{m+4} \oplus D^{5\cdot i} W_{m+5} \oplus D^{6\cdot i} W_{m+6} \oplus D^{7\cdot i} P_m \oplus D^{8\cdot i} W_{m+7} \oplus D^{9\cdot i} W_{m+8} \oplus D^{10\cdot i} W_{m+9} \oplus D^{11\cdot i} W_{m+10} \oplus D^{12\cdot i} W_{m+11} \oplus D^{13\cdot i} W_{m+12} \oplus D^{14\cdot i} W_{m+13} \quad \text{----------} (2)$$

Here D is a delay operator; D^i represents the delay of word i. Therefore, we obtain $D^{\ell\cdot m} W_{m+k} = W_{m+k-\ell\cdot m}$.
$\oplus$ means vector addition on GF (2) and each word (W, P, Q) corresponds to its relevant vector.
The word W is recorded on tracks 1 through 7 and the words P and Q are recorded on an exclusive track (track 8) according to the frame configuration shown in Fig.6.

(2) Decoding

Among several alternative ways of decoding, considering the practical error correction capability, this system adopts a method as shown in Fig. 8. In playback, the

words W, P and Q coming from each track are subjected to error detecting operation as expressed by the *generator polynomial G(X) and CRC word, then written* into memory with labeled error pointers. Correction is performed after the arrangement at the time of generation of P and Q is recovered. The principle of correction is expressed by the encoding equations (1) and (2). Correction according to equation (1) is P-correction and one according to equation (2) is Q-correction.
The following are examples of the P-correction and Q-correction when the W_{n+3} is error-labeled.

P-correction

$$
\begin{aligned}
W_{n+3} = {} & W_n \oplus W_{n+1} \oplus W_{n+2} \oplus W_{n+4} \oplus W_{n+5} \oplus W_{n+6} \oplus W_{n+7} \\
& \oplus W_{n+8} \oplus W_{n+9} \oplus W_{n+10} \oplus W_{n+11} \oplus W_{n+12} \oplus W_{n+13} \oplus P_n
\end{aligned}
$$

Q-correction

$$
\begin{aligned}
W_{n+3} = {} & D^{-3\cdot i} \big(Q_n \oplus D^{0\cdot i} \cdot W_n \oplus D^{1\cdot i} \cdot W_{n+1} \oplus D^{2\cdot i} \cdot W_{n+2} \oplus D^{4\cdot i} \cdot W_{n+4} \oplus D^{5\cdot i} \cdot W_{n+5} \\
& \oplus D^{6\cdot i} \cdot W_{n+6} \oplus D^{7\cdot i} \cdot P_n \oplus D^{8\cdot i} \cdot W_{n+7} \oplus D^{9\cdot i} \cdot W_{n+8} \oplus D^{10\cdot i} \cdot W_{n+9} \\
& \oplus D^{11\cdot i} \cdot W_{n+10} \oplus D^{12\cdot i} \cdot W_{n+11} \oplus D^{13\cdot i} \cdot W_{n+12} \oplus D^{14\cdot i} \cdot W_{n+13} \big) \\
= {} & D^{-3\cdot i} \cdot Q_n \oplus D^{-3\cdot i} \cdot W_n \oplus D^{-2\cdot i} \cdot W_{n+1} \oplus D^{-1\cdot i} \cdot W_{n+2} \oplus D^{1\cdot i} \cdot W_{n+4} \oplus D^{2\cdot i} \cdot W_{n+5} \\
& \oplus D^{3\cdot i} \cdot W_{n+6} \oplus D^{4\cdot i} \cdot P_n \oplus D^{5\cdot i} \cdot W_{n+7} \oplus D^{6\cdot i} \cdot W_{n+8} \oplus D^{7\cdot i} \cdot W_{n+9} \\
& \oplus D^{8\cdot i} \cdot W_{n+10} \oplus D^{9\cdot i} \cdot W_{n+11} \oplus D^{10\cdot i} \cdot W_{n+12} \oplus D^{11\cdot i} \cdot W_{n+13}
\end{aligned}
$$

In this correction method, first Q-correction (1) is applied and all possible corrections are made. Then after writing into memory, P-correction (1) is applied. Q-correction (2) and P-correction (2) follow in the same way. In Q-correction (1), failures in CRC error detection are also checked by syndrome check.

Conclusion

Our prototype made in accordance with the format described in this paper was successful. The prototype system uses a Compact Cassette which contains a metal tape having a coercive force of 1350 Oe.

In order for prerecorded music tapes to be mass-produced with a high-speed duplication system, an appropriate track width and a long-life magnetic head are necessary. The format used in our system meets all these requirements. Our decision to use 4 tracks per audio channel contributes to obtaining an appropriate track width and simplifying the playback circuit arrangement.
We propose this system as a well-balanced system resulting from the consideration of many factors.
Here we wish to extend our gratitude to Mr. T. Inoue, M. Fujimoto, M.Yamazaki and N. Takahashi who have given us much assistance in this project and to all of our members of this project team.

References

[1] K. Tanaka, T. Yamaguchi, and Y. Sugiyama "Improved Two Channel PCM Tape Recorder For Professional Use" AES 64th Conv. Preprint, 1979
[2] H. Matsushima, K. Kanai, T. Mura, and T. Kogure "A New Digital Audio Recorder For Professional Applications" AES 62nd Conv. Preprint 1979
[3] EIAJ Technical File STC-007
[4] T. Muraoka, M. Iwahara, and Y. Yamada "Examination Of Audio-Bandwidth Requirements For Optimum Sound Signal Transmission" J. Audio Eng. Soc. Vol. 29, No.1/2, 1981, Jan./Feb.
[5] G. Plenge, H. Jakubowski, and P. Schöne "Which Bandwidth Is Necessary For Optimal Sound Transmission" AES 62nd Conv. Preprint 1979
[6] J. C. Mallinson "On Extremely High Density Digital Recording" IEEE, Trans. on Mag. Vol. MAG-10, No. 2, June 1974
[7] G. Bate, G. J. Hampton, and B. J. Latta "A 5 Megbyte Flexible Disk" IEEE, Trans. on Mag. Vol. MAG-17, No. 4, July 1981
[8] A. Lender "Correlative Level Coding For Binary-Data Transmission" IEEE Spectrum, Feb. 1966
[9] H. Kobayashi and D. T. Tang "Application Of Partial-Response Channel Coding To Magnetic Recording Systems" IBM Jour. Res. Develop, July 1970
[10] S. Nakagawa, K. Yokogawa, and H. Katayama "A Study On Detection Methods Of NRZ Recording" IEEE Trans. On Mag. Vol. MAG-16, No. 1, Jan. 1980
[11] Y. Yamada, Y. Fujii, M. Moriyama and S. Saito "Professional-Use PCM Audio Processor With A High Efficiency Error Correction System" AES 66th Conv. Preprint 1980

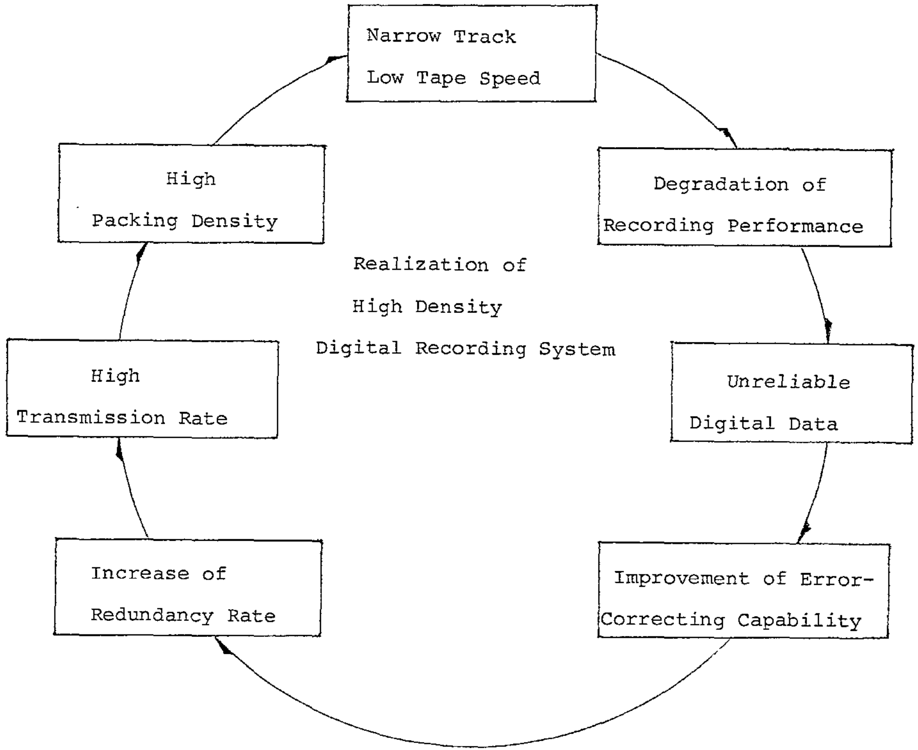

Narrow Track
Low Tape Speed
Degradation of
Recording Performance
Unreliable
Digital Data
Improvement of Error-
Correcting Capability
Increase of
Redundancy Rate
High
Transmission Rate
High
Packing Density
Realization of
High Density
Digital Recording System

Fig. 1, " Circle "

(1) Performance Specifications	
Number of Audio Channel	2 Audio Channels (Each Direction)
Frequency Response	DC - 15KHz $\pm$ 0.5dB
Dynamic Range	more than 84dB
Distortion	less than 0.03% T.H.D.
Wow and Flutter	Crystal Oscillator Accuracy
Tape Speed	7.14 cm/s
Tape	3.81mm wide,Metal Tape
(2) Structural Specifications	
Track Configuration	8 Tracks/2 Audio Channels 1 AUX. Track
Sampling Rate	33.6KHz
Data Word Bits	12 bits 7-segments,Compress/Expand
Total Transmission Rate	1.0416M b/s
Redundancy Rate	0.226
Linear Density	46.29K bpi (1.82K bit/mm)
Mod./Demodulation	Partial Response (1,-1)
Error Correcting Code	Triple Error Correction with Error Pointer

Table 1, Specifications

		Sampling Rate (Hz)			
		32K	33.6K	44.1K	48K
Data Word 12 Bits		1.024M B/S	1.075M B/S	1.411M B/S	1.536M B/S
Transmission	4 Tracks per Audio Channel	128K	134.375K	176.4K	192K
Rate	8 Tracks per Audio Channel	64K	67.188K	88.2K	96K
per Track	16 Tracks per Audio Channel	32K	33.594K	44.1K	48K
Data Word 14 Bits		1.195M B/S	1.254M B/S	1.646M B/S	1.792M B/S
Transmission	4 Tracks per Audio Channel	149.375K	156.8K	205.8K	224K
Rate	8 Tracks per Audio Channel	74.687K	78.4K	102.9K	112K
per Track	16 Tracks per Audio Channel	37.344K	39.2K	51.45K	56K
Data Word 16 Bits		1.365M B/S	1.434M B/S	1.882M B/S	2.048M B/S
Transmission	4 Tracks per Audio Channel	170.625K	179.2K	235.2K	256K
Rate	8 Tracks per Audio Channel	85.313K	89.6K	117.6K	128K
per Track	16 Tracks per Audio Channel	42.656K	44.8K	58.8K	64K

2 Audio Channels, Redundancy Rate : 0.25

Table 2, Total Transmission Rate and Transmission Rate per Track vs. Sampling Rate

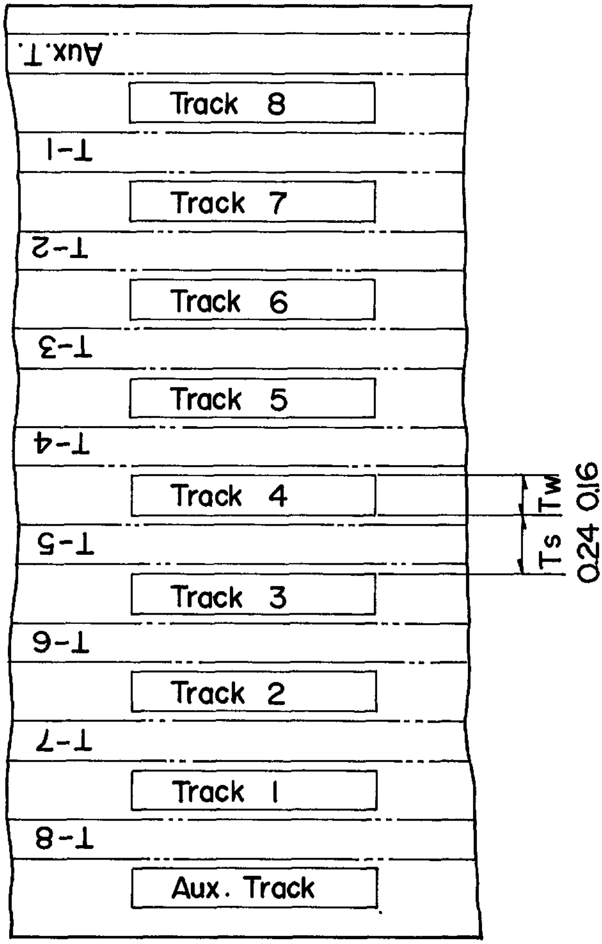


Fig. 2, Track Format

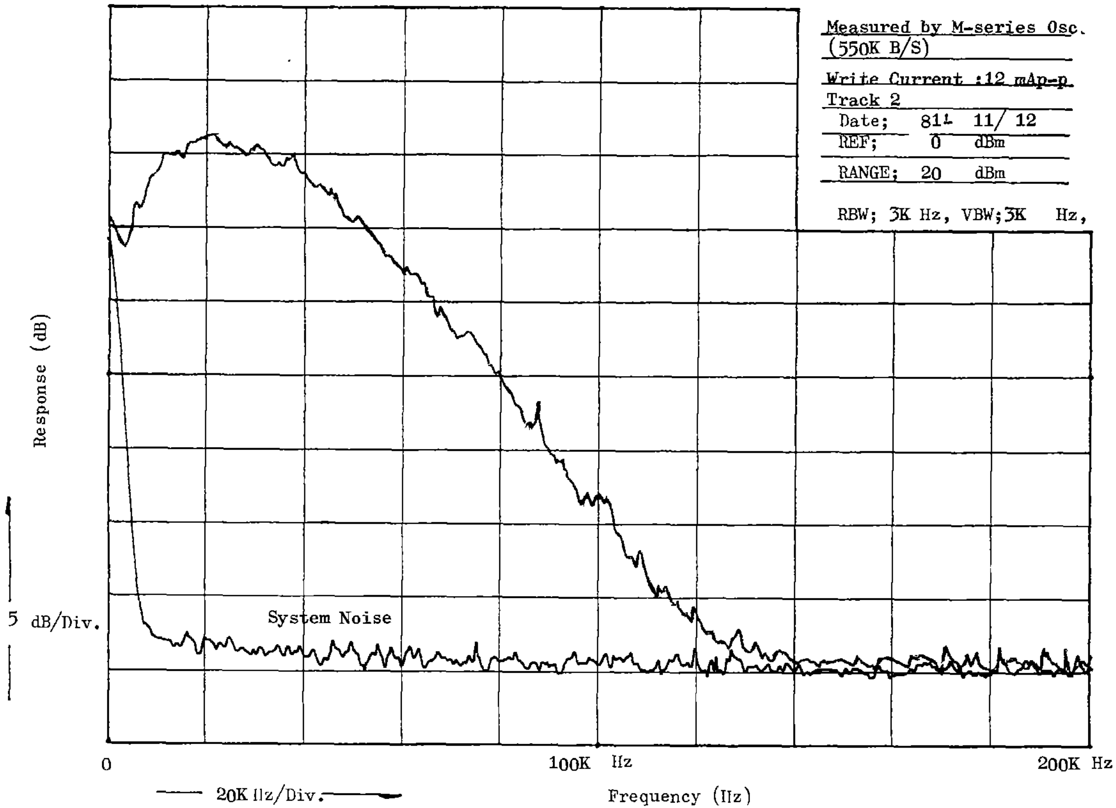
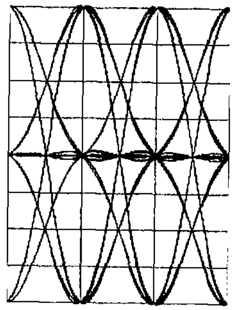
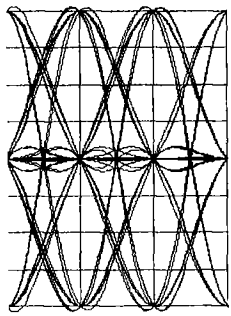


Fig. 3, Unequalized Frequency Response of The Digital Compact Cassette System

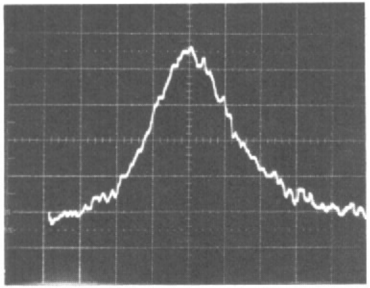


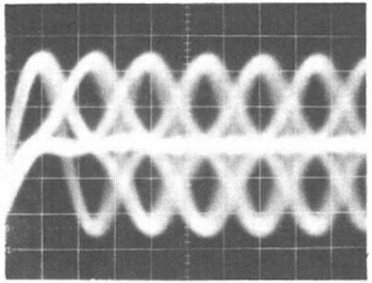
(1) Rolloff Factor : 1



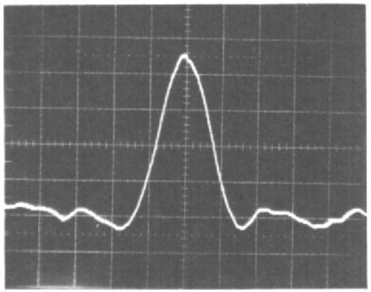
(2) Rolloff Factor : 0.7

Fig. 4, Simulated Eye Patterns



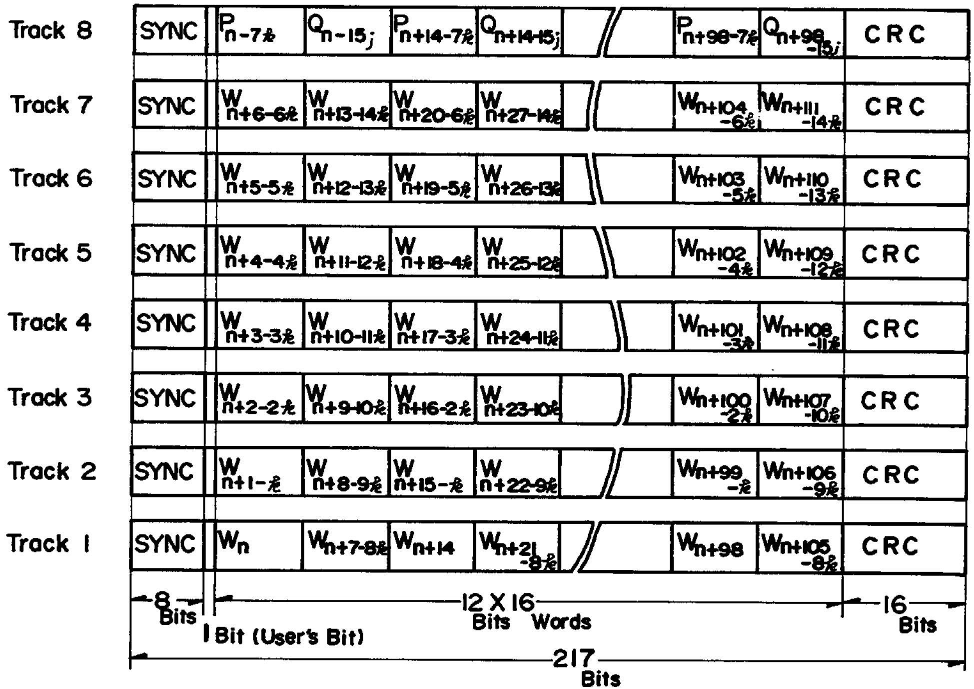


5µs/div.
Eye Pattern of The Digital Compact Cassette System



5µs/div.
Equalized Isolated Readback Pulse

Fig. 5, Isolated Readback Pulses and Eye Pattern of The Digital Compact Cassette System.



k=i+j

Fig. 6, Frame Configuration

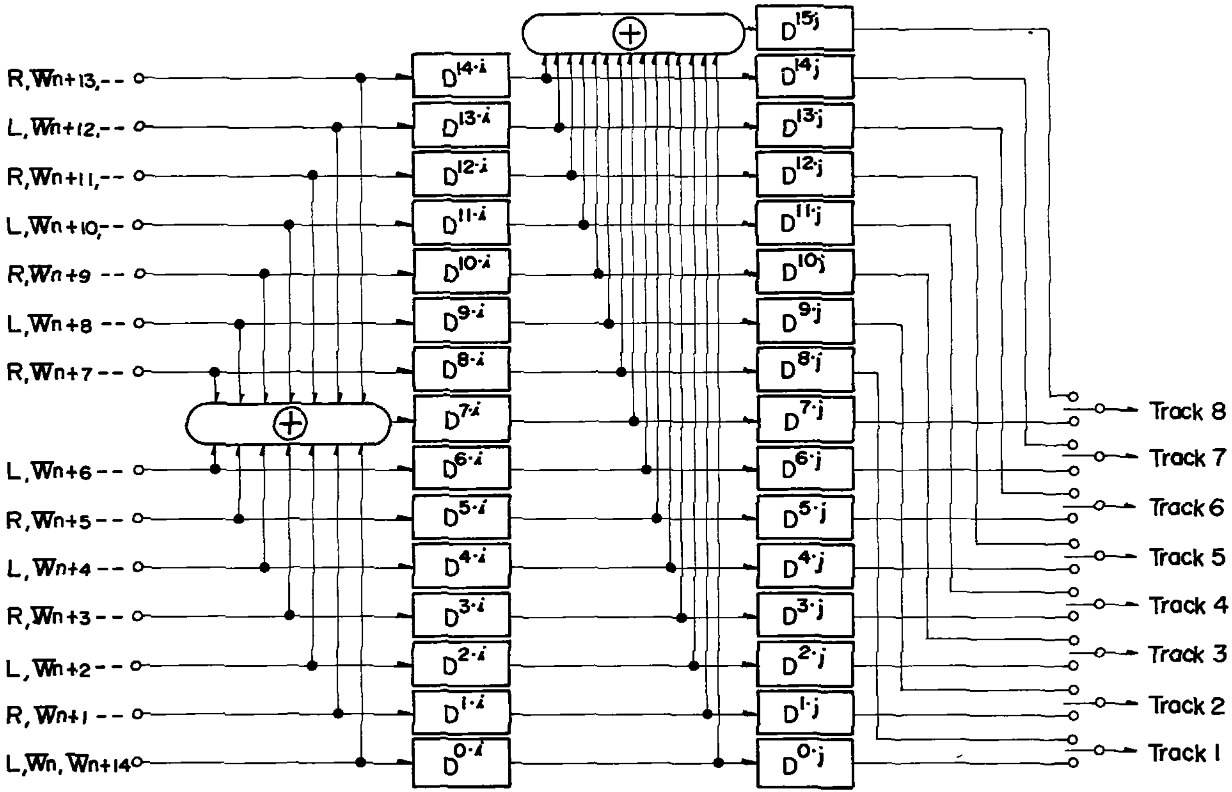


Fig. 7, Encoding Diagram

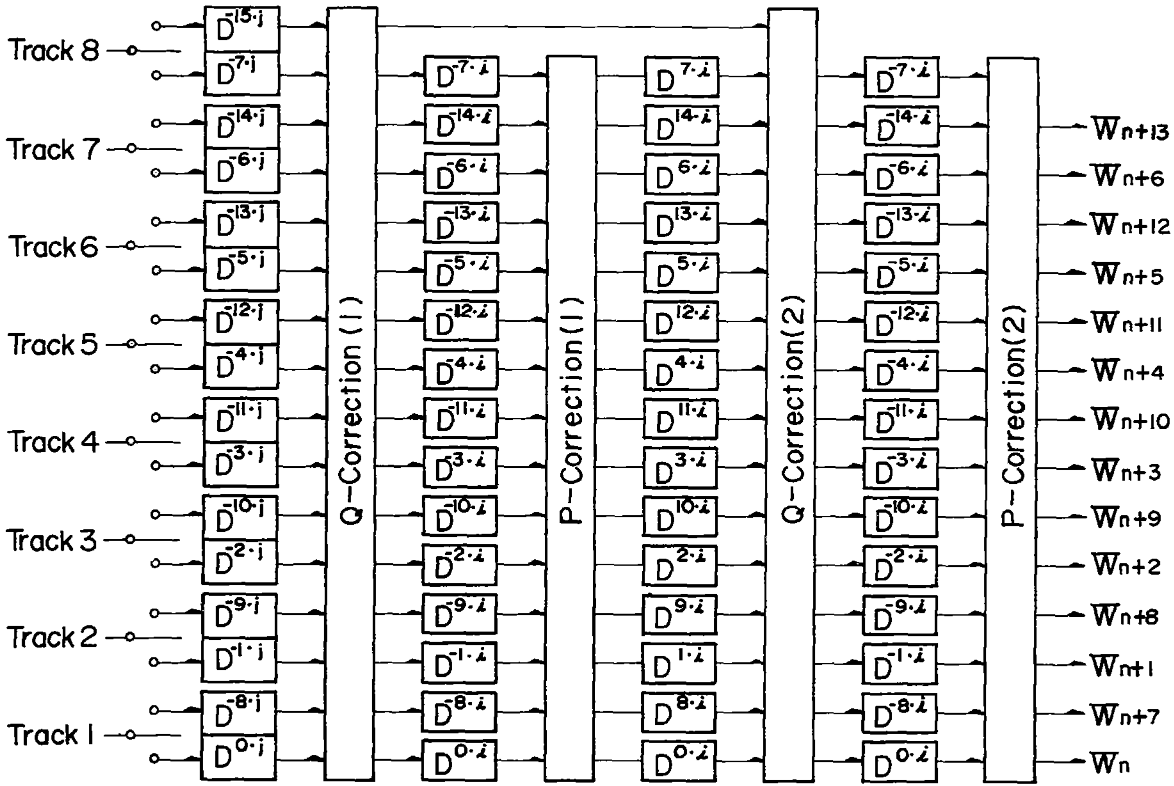


Fig. 8, Decoding Diagram