A MULTI-TRACK DIGITAL AUDIO RECORDER FOR CONSUMER APPLICATIONS

1832(I-4)

W. J. van Gestel, L.M.H.E. Driessen,
and J.C.F. Moeskops
Philips Research Laboratories
Eindhoven, the Neterlands

**Presented at
the 70th Convention
1981 October 30-November 2
New York**

AN AUDIO ENGINEERING SOCIETY PREPRINT

A MULTI-TRACK DIGITAL AUDIO RECORDER FOR CONSUMER APPLICATIONS

BY

W.J. van Gestel, L.M.H.E. Driessen
and J.C.F. Moeskops

Philips' Research Laboratories
Eindhoven, the Netherlands

ABSTRACT

A prototype multi-track digital audio recorder is described. Two audio channels are sampled and coded by 14-bit linear quantization. The sampling rate is 44.1 kHz, the tape speed is 9.5 cm/sec. and the tape width is $\frac{1}{4}$". The channel code (a dc free block code) makes use of the statistics of the audio signal. Precautions are taken to limit the crosstalk in the 16-track ferrite head stack. Error correction and concealment techniques are applied to compensate for random and burst errors.

INTRODUCTION

The limitations of conventional analog recorders can be overcome by the use of digital techniques. Multi-track digital recorders for professional use are well known [1-2-3] . For consumer purposes, digital recorders based on video recorders are already standardized [4] but investigations on multi-track systems are still in progress. This paper deals with a prototype multi-track recorder which is meant for consumer use. The specifications of the audio quality and the digital parameters are shown in table 1.

Table I : Specifications

Number of audio channels	2
Frequency response	0-20 kHz
Dynamic range	85 dB
Wow and flutter	With compensation equal to crystal acc.
Tape speed	9.5 cm/sec
Tape	6.35 mm wide high density tape
Source coding	14 bit linear PCM
Sampling rate	44.1 kHz
Drop-out compensation	Error correction with double MDS code Error concealment with linear interpolation-muting
Number of tracks	16
Redundancy	28% (Error correction + frame sync.)
Channel code	7-8 (almost) dc free block code
Bit length on the tape	0.77 μm

THE SAMPLING RATE

A 44.1 kHz sampling rate is adopted because this sampling rate is also used in the digital Compact Disc system and in the domestic digital audio adaptors using PAL or SECAM video format. With these sources therefore recordings can be made without using the analog in and output circuitry or without using rate converters. The difference with the 44.056 kHz sampling frequency of the NTSC format adaptors is very small and can be handled by varying the tape speed. The 32 kHz sampling rate proposed by different Post Offices, is recorded and played back with a lower tape speed and crystal frequency. In the future there will be rate converters transforming 50.4 kHz sampling frequency (professional use) into 44.1 kHz.

The input filter and A/D converter are only needed with analog input signals (radio, disc, etc.). Taking into account the power density spectrum from these signals it can be said that the requirements on the input filter will not be so severe as is often thought. At the 44.1 kHz sampling frequency, a 20 kHz bandwidth is possible with a very low ripple (determined by input and output filters). If there is no pre- and de-emphasis of the audio signals, the peak value of the input signal is the same for all frequencies (power bandwidth 0-20 kHz).

THE DYNAMIC RANGE

A dynamic range of 85 dB seems to be sufficient at home. This is reached with 14-bit linear PCM coding of the samples. There is no need for headroom, the maximum output signal of all sources is well-known. With linear PCM the S/N ratio is equivalent to the dynamic range. Distortion , intermodulation, etc. can be kept very low. They are determined only by the input and output filter and by the A/D and D/A converters.

THE NUMBER OF TRACKS

Digitizing of the analog signals results in a very high bitrate ($\approx$ 2 Mbit/sec.). From the tape speed (set to 9.5 cm/sec) and the minimum allowable wavelength, we can calculate the number of tracks. Wavelengths longer than 2-3 μm result in low bit error rates. There is then no good compromise obtained between the tape consumption and what is possible with error correction. Furthermore with the tape speed fixed to 9.5 cm/sec this would result in a lot of tracks (> 20). Heads with a large number of tracks are difficult to realize, are more sensitive to mistracking and require a lot of electronics.

Wavelengths $< 1.5 \mu m$ (in situations with a low tape speed and a large contact area between head and tape) result in a rapid increase in random bit errors due to a low S/N ratio and burst errors. With a 16-track head, the minimum wavelength is in the desired region. These ferrite head stacks can be made with conventional techniques.

In digital recording thin tapes (similar to video tapes) can be used, because there are no long wavelengths. In this way with a 9.5 cm/sec tape speed, a playing time of several hours of uninterrupted music is possible. Doubling the tape speed and halving the number of tracks in order to reduce the amount of electronics and the complexity of the heads, results in a playing time which is still long enough.

THE HEADS

Multi-track ferrite heads made with conventional methods (sawing-polishing-lapping) are used. The track positioning is given in fig. 1, different stages in the fabrication in fig. 2. The guard band between the cores makes it possible to use the tape in both directions. A proper tape guiding system is then required. The number of turns of each coil is limited by the available space.

A serious problem with multi-track heads is crosstalk. There are several kinds of crosstalk;

- Crosstalk during playback from a magnetized pattern on the tape to an adjacent head. This crosstalk may be neglected, because here wavelengths are short and a guard band is used.
- Crosstalk between the cores of adjacent heads. This crosstalk depends on the size of the opposing faces of the cores, the cross-section, the permability and the distance between the cores. It occurs during playback as well as during writing.

Crosstalk during playback

In the head design the distance between the cores is 225 μm, the width is 150 μm while the surface of the opposing yokes is large. Playback crosstalk can be as large as -12 dB. Each head except the outer ones does have two neighbours, so without crosstalk compensation the eye pattern will be almost closed. However, this kind of crosstalk is linear and in the frequency range used it is wavelength-independent (phase and amplitude). With the circuit of fig. 3 a remarkable reduction of this crosstalk is obtained.

Crosstalk during writing

During writing crosstalk from one core to another results in a magnetized pattern on the tape. The record current from one head acts as a bias field for the crosstalk field. The magnetization process is a non-linear process. It depends on the amplitude of the record current, the head saturation the coercivity of the tape etc. After playback this kind of crosstalk may even be higher than the playback crosstalk.

Crosstalk during writing is rejected in the following way. The record current of each head is duty-cycled (fig. 4). At least once per bit cell the record current is switched on during a short time, but heads close to each other never carry a current at the same time. In this way there is no bias field for the crosstalk fields while this crosstalk field alone is too weak to magnetize the tape. The spatial extension of the record field is large enough to magnetize the tape. With a record current optimized for short wavelengths, the record region is about 0.5 μm beyond the pole tips (see fig. 5). One record pulse every bit cell (0.77 μm) and every time on the same place of that bit interval is enough. This kind of duty-cycling is often used in combination with thin-film record heads, in order to limit the total record current and dissipation.

THE CHANNEL CODE

Channel coding is employed to match certain properties of the coded data to the characteristics of the recorder channel. The recorder channel is a bandpass, non-linear communication channel which suffers both amplitude and timing instabilities.

In systems with a low S/N ratio, amplitude fluctuations and a varying frequency response at high frequencies, equalizing by means of restoring the record current is a good choice. Restoring of the write current means integrating the differentiated output of the heads and possibly boosting the high frequencies. Due to the differentiator followed by an integrator the channel code should be dc free. In our case the S/N ratio is marginal and the electronics noise supersedes the tape noise.

The multi-track head suffers from track-to-track crosstalk. The compensation network will not eliminate all the crosstalk. Because of hum and 1/f noise of the pre-amplifier, a high pass filter with a cut-off frequency $f_c \approx 500$ Hz is needed in the equalizer. Although the code may be dc free, there will be symbol interference caused by low frequency components in the channel code. All these effects favour a channel code with an eye-opening as large as possible (large window ΔT and clocking at the maximum amplitude).

Block codes are more favourable than running codes because of the error propagation. In running codes the presence of a transition depends on the preceding data bits. There is no difference between one bit error and 14 erroneous bits in one sample if error concealment is used. Therefore if there are errors, they should be concentrated in one sample. This holds too for the error correction system.

To overcome all these difficulties an (almost) dc free block code (7-8) is developed with a clock window $\Delta T = 7/8$ T. The 14-bit samples are divided into two 7-bit symbols. A 14-16 block code is not practicable because of the large PROMs and the complicated error correction circuits that are needed.

The 7-8 block code cannot be made completely dc free, there are only 70 dc free code words. In the coding table we make use of the statistics of the audio signal. Fig. 8 gives the probability distribution of the signal amplitude levels for 10 minutes of orchestral music. Here we can see that small amplitude levels are much more likely to occur than high levels. The coding table is used in such a way that small amplitudesare coded in 8-bit symbols which are dc free in half the symbol length (4 bits). The high amplitudes are coded in non-favourable code words (Tmax = 5 ΔT); see also fig. 8. By using Tmax= 5 ΔT a practically unique sync symbol can be obtained.

The 8-7 decoder makes it possible to detect some errors. As soon as an unused codeword appears at the intput of the PROM an erasure bit appears at the output; see fig. 6. The eye pattern of one track is shown in fig. 7.

THE ERROR CORRECTION SYSTEM

Errors in magnetic recording can be either random bit errors due to a low S/N ratio or burst errors mainly caused by dust particles, scratches etc. In general dust particles do not affect all tracks simultaneously but rather selectively degrade one or a few neighbouring tracks. Random bit errors should rarely occur; the system is designed in such a way that the S/N ratio is high enough. Then the error rate is determined by the burst errors, which cannot be prevented.

To choose the right drop-out compensation systems, statistical data of drop-outs must be known. However, the variation in drop-outs between tapes and between new and used tapes is more than 10 times.

From the perceptibility of errors we know that
- with error concealment (linear interpolation) a surprisingly good result is obtained;
- one click due to miscorrection or not detected errors may be more annoying then many interpolations.

The error detection should therefore be very strong.

A format has been created (fig. 9) in which parity check symbols are added in the longitudinal direction of the tape (horizontal) and in the cross-track direction (vertical). The detection is carried out with the P_h Q_h symbols. Often a CRC code is used for detection. We have opted for a shortened Reed-Solomon code. The detection proability is equivalent to the 14-bit CRC code, but the Reed-Solomon code gives the opportunity to correct random errors (one not identified symbol out of 16 symbols). This can be used in the error correction strategy.

The detection probability with 14 bits seems to be too low, but this detection is supplemented by the detection with 8-7 decoder. The detection probability during decoding of one (random) erroneous symbol is ≈ 0.5 (the unused code word). The detection probability of burst errors is much higher if a small hysteresis in the limiter is taken equal to peak noise level. During a drop-out the signal level decreases and the output of the limiter will remain in one position. These non dc free code words are detected. The Reed-Solomon code detects all single and double errors. Burst errors (> 2 symbols) will be detected by the Reed-Solomon code, the 8-7 decoder or by both of them.

The vertical parity symbols Pv and Qv are used to correct the detected erroneous tracks. A maximum of two tracks out of 16 tracks can be corrected. For horizontal as well as for vertical direction the same generator polynomial is taken. Thus with a square matrix (16x16 symbols) the electronic circuits for both situations are the same.

Fig. 10 shows some examples of correctable and non-correctable errors. As soon as correction fails, concealment is used. The concealment properties for large drop-outs with several tracks close to each other damaged, are increased by connecting the heads as shown in fig. 11.

No interleaving in the tracks is used; the multiplexing over the tracks makes correction of continuous drop-outs possible. Interleaving would only be useful with very large drop-outs like finger prints. However, this leads to very large memories. A method of avoiding these memories is given in fig. 12.
Here spatial interleaving is obtained due to the different positions of the gaps of the heads.

THE BLOCK DIAGRAM

The block diagram of the whole system is shown in fig. 13. Clocking signals are omitted. Small memories are needed to bridge over the time in which the parity symbols are fitted in the bit stream. The 8-bit symbols from the channel coder are multiplexed over 16 tracks.

The playback signal from each head is amplified, cross talk is compensated, equalized and clocked with its own clocking signal. Times base differences between tracks due to skewing of the tape or gap scatter between record and playback head are eliminated in the time base corrector. With this memory a crystal clock and a tape speed servo, wow and flutter can be compensated too.

The error correction circuit first detects erroneous tracks, and corrects all single symbol errors. Then it corrects in the vertical direction the erroneous tracks (two at the most). The uncorrectable errors are treated with a concealment method. The RAM capacity of the corrector is 4Kbit and of the time base corrector 2Kbit. Both memories can be combined.

CONCLUSIONS

A multi-track PCM recorder with a low tape consumption has been described. The format on the tape is designed in such a way that it can correct both random errors and burst errors. Even with large drop-outs in more tracks similtaneously, a good result is obtained with the concealment method.

The total memory capacity is about 6 kbit. To make a cheap recorder, which is needed for consumer applications, the electronic circuits should be integrated (of course after standardization). Then special attention should be paid to the analog circuitry of each track. It is expected that there will be progress in realizing cheap A/D and D/A converters.

The heads remain the most serious problem. Up to now no cheap multi-track ferrite heads nor multi-track thin film heads are available.

REFERENCES

1. J.A. McCracken. "A high performance digital audio recorder". AES 58th Conv. 1977.

2. H. Matsushima, K. Kanai, T. Mura and T. Kogure. "A new digital audio recorder for professional applications". AES 62nd Conv. 1979.

3. K. Tanaka et al., "On PCM Multi-channel tape recorder using powerful code format". AES 67th Conv. 1980.

4. EIAJ Technical File STC-007.

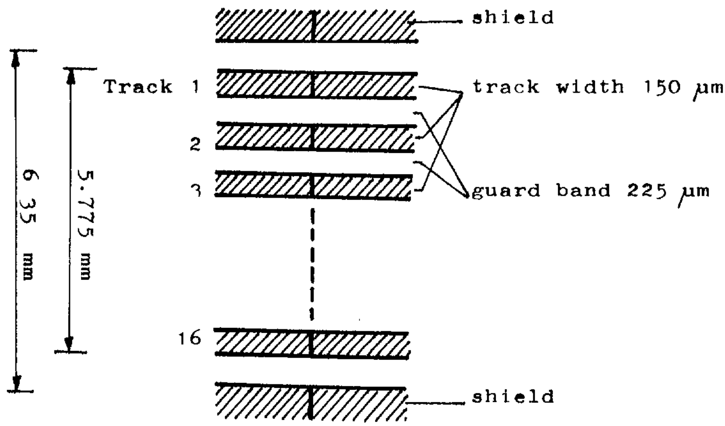


FIG 1: TRACK POSITIONING OF THE HEADS.

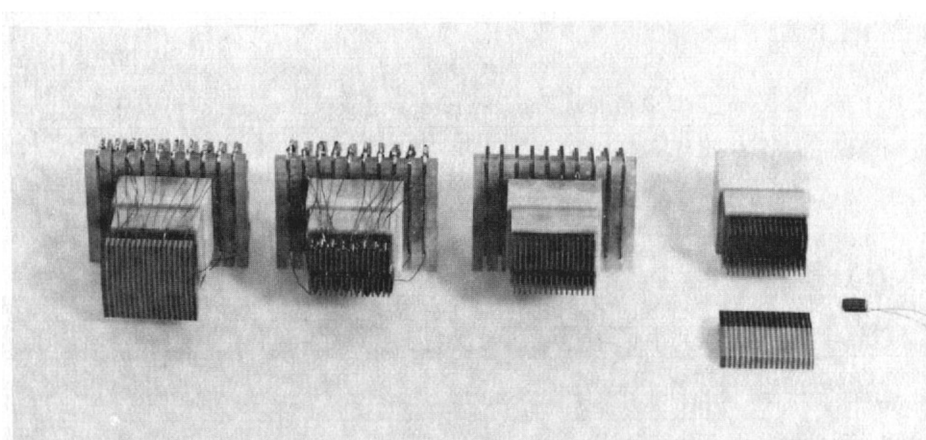


FIG 2: DIFFERENT STAGES IN THE FABRICATION OF THE HEADS.

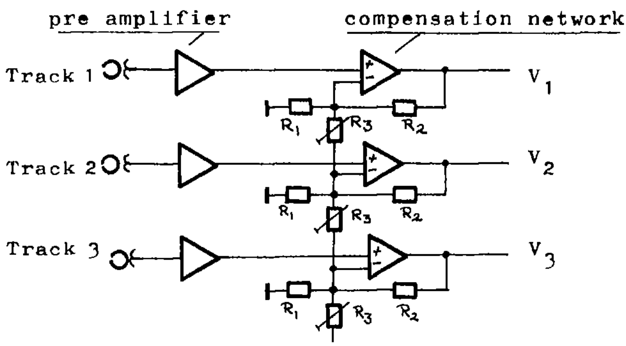


FIG 3: CROSSTALK COMPENSATION DURING PLAYBACK.

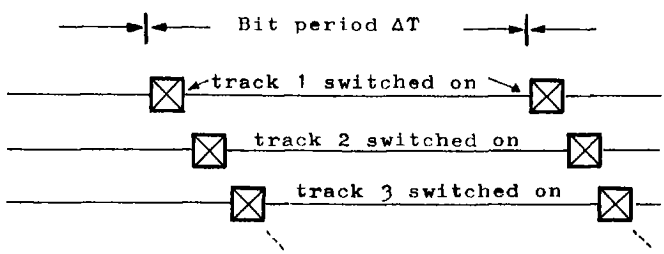


FIG 4: DUTY-CYCLING OF THE RECORD CURRENT IN ORDER TO REDUCE CROSSTALK DURING WRITING.

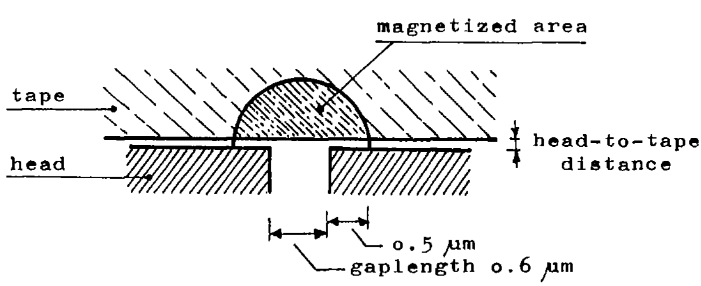


FIG 5: MAGNETIZED AREA BY A RECORD CURRENT PULSE

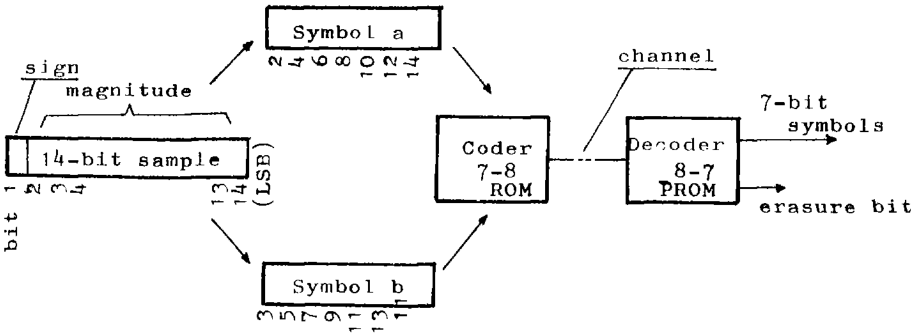


FIG 6: BLOCK DIAGRAM OF THE CHANNEL CODER-DECODER.

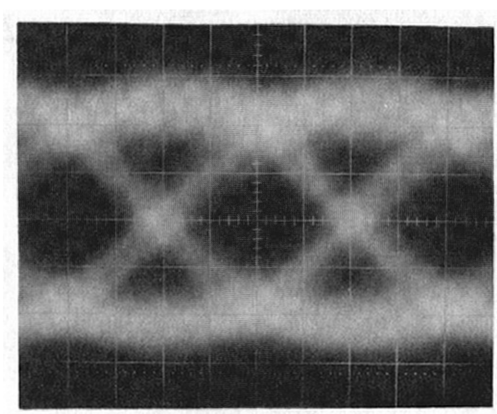


FIG 7: EYE PATTERN OF THE EQUALIZED SIGNAL.

Tape speed 9.5 cm/sec.
CrO_2 video tape.
With crosstalk compensation.

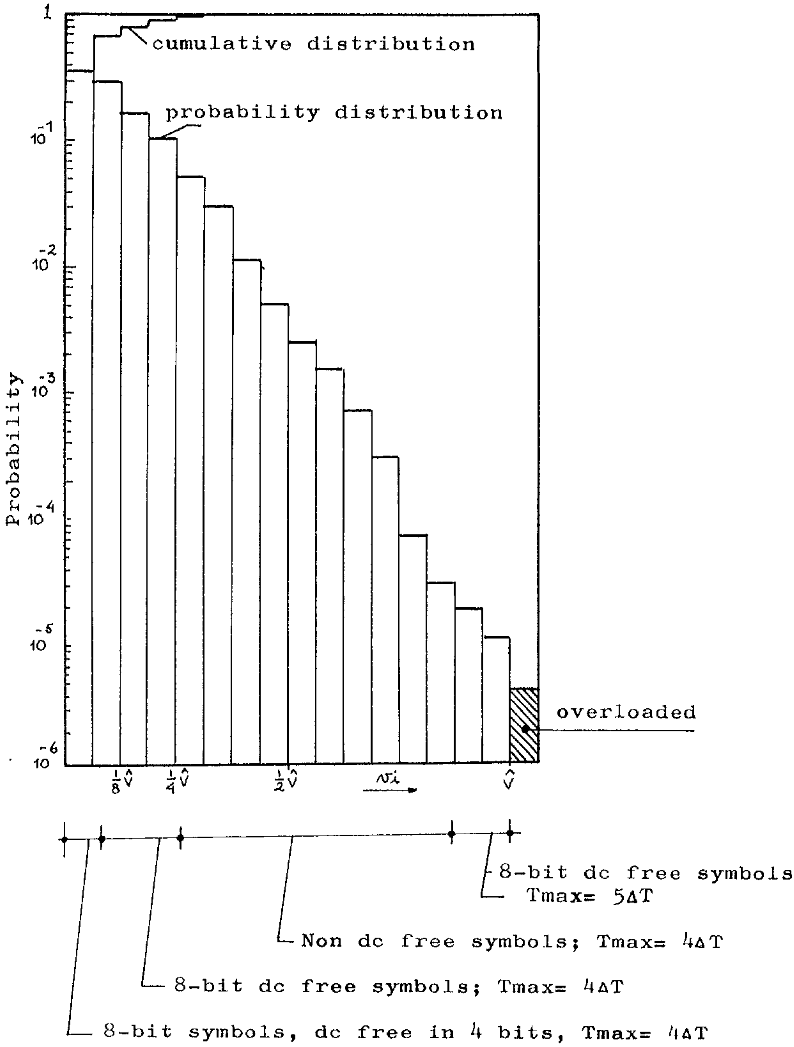
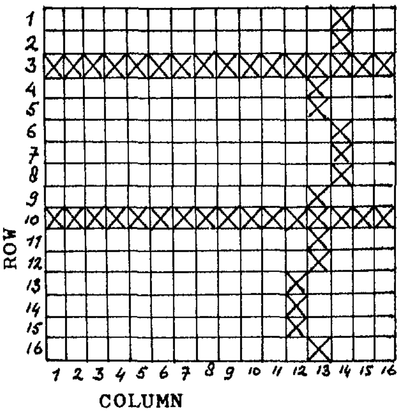
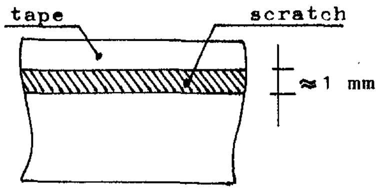


FIG 8: PROBABILITY DISTRIBUTION OF RECORDED MUSIC.
(10 minutes of orchestral music from disc)

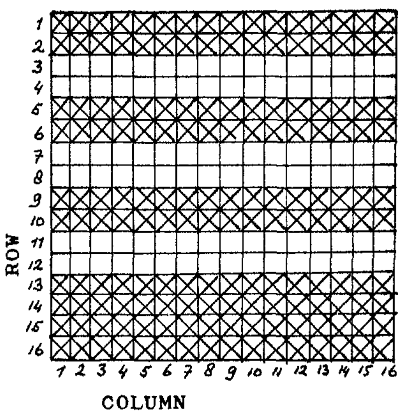
sync	L_{1a}	L_{1b}	L_{8a}	L_{8b}	L_{15a}	L_{15b}	L_{22a}	L_{22b}	L_{29a}	L_{29b}	L_{36a}	L_{36b}	L_{43a}	L_{43b}	P_{h1}	Q_{h1}	sync	ROW 1
sync	R_{1a}	R_{1b}	R_{8a}	R_{8b}	R_{15a}	R_{15b}	R_{22a}	R_{22b}	R_{29a}	R_{29b}	R_{36a}	R_{36b}	R_{43a}	R_{43b}	P_{h2}	Q_{h2}	sync	2
sync	L_{2a}	L_{2b}	L_{9a}	L_{9b}	L_{16a}	L_{16b}	L_{23a}	L_{23b}	L_{30a}	L_{30b}	L_{37a}	L_{37b}	L_{44a}	L_{44b}	P_{h3}	Q_{h3}	sync	3
sync	R_{2a}	R_{2b}	R_{9a}	R_{9b}	R_{16a}	R_{16b}	R_{23a}	R_{23b}	R_{30a}	R_{30b}	R_{37a}	R_{37b}	R_{44a}	R_{44b}	P_{h4}	Q_{h4}	sync	4
sync	L_{3a}	L_{3b}	L_{10a}	L_{10b}	L_{17a}	L_{17b}	L_{24a}	L_{24b}	L_{31a}	L_{31b}	L_{38a}	L_{38b}	L_{45a}	L_{45b}	P_{h5}	Q_{h5}	sync	5
sync	R_{3a}	R_{3b}	R_{10a}	R_{10b}	R_{17a}	R_{17b}	R_{24a}	R_{24b}	R_{31a}	R_{31b}	R_{38a}	R_{38b}	R_{45a}	R_{45b}	P_{h6}	Q_{h6}	sync	6
sync	L_{4a}	L_{4b}	L_{11a}	L_{11b}	L_{18a}	L_{18b}	L_{25a}	L_{25b}	L_{32a}	L_{32b}	L_{39a}	L_{39b}	L_{46a}	L_{46b}	P_{h7}	Q_{h7}	sync	7
sync	R_{4a}	R_{4b}	R_{11a}	R_{11b}	R_{18a}	R_{18b}	R_{25a}	R_{25b}	R_{32a}	R_{32b}	R_{39a}	R_{39b}	R_{46a}	R_{46b}	P_{h8}	Q_{h8}	sync	8
sync	L_{5a}	L_{5b}	L_{12a}	L_{12b}	L_{19a}	L_{19b}	L_{26a}	L_{26b}	L_{33a}	L_{33b}	L_{40a}	L_{40b}	L_{47a}	L_{47b}	P_{h9}	Q_{h9}	sync	9
sync	R_{5a}	R_{5b}	R_{12a}	R_{12b}	R_{19a}	R_{19b}	R_{26a}	R_{26b}	R_{33a}	R_{33b}	R_{40a}	R_{40b}	R_{47a}	R_{47b}	P_{h10}	Q_{h10}	sync	10
sync	L_{6a}	L_{6b}	L_{13a}	L_{13b}	L_{20a}	L_{20b}	L_{27a}	L_{27b}	L_{34a}	L_{34b}	L_{41a}	L_{41b}	L_{48a}	L_{48b}	P_{h11}	Q_{h11}	sync	11
sync	R_{6a}	R_{6b}	R_{13a}	R_{13b}	R_{20a}	R_{20b}	R_{27a}	R_{27b}	R_{34a}	R_{34b}	R_{41a}	R_{41b}	R_{48a}	R_{48b}	P_{h12}	Q_{h12}	sync	12
sync	L_{7a}	L_{7b}	L_{14a}	L_{14b}	L_{21a}	L_{21b}	L_{28a}	L_{28b}	L_{35a}	L_{35b}	L_{42a}	L_{42b}	L_{49a}	L_{49b}	P_{h13}	Q_{h13}	sync	13
sync	R_{7a}	R_{7b}	R_{14a}	R_{14b}	R_{21a}	R_{21b}	R_{28a}	R_{28b}	R_{35a}	R_{35b}	R_{42a}	R_{42b}	R_{49a}	R_{49b}	P_{h14}	Q_{h14}	sync	14
sync	P_{V1}	P_{V2}	P_{V3}	P_{V4}	P_{V5}	P_{V6}	P_{V7}	P_{V8}	P_{V9}	P_{V10}	P_{V11}	P_{V12}	P_{V13}	P_{V14}	P_{h15}	Q_{h15}	sync	15
sync	Q_{V1}	Q_{V2}	Q_{V3}	Q_{V4}	Q_{V5}	Q_{V6}	Q_{V7}	Q_{V8}	Q_{V9}	Q_{V10}	Q_{V11}	Q_{V12}	Q_{V13}	Q_{V14}	P_{h16}	Q_{h16}	sync	16
COLUMN	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16		

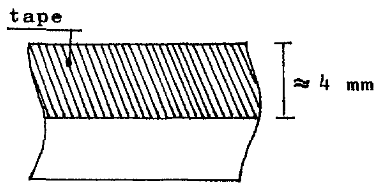
FIG 9 : FORMAT ON THE TAPE. (each symbol is 7 data bits - 8 channel bits).





Full error correction is possible with two damaged tracks and one erroneous symbol in the other tracks.





Error correction is not possible.
Error concealment with linear interpolation is still possible

FIG 10: EXAMPLES OF ERROR CORRECTION AND CONCEALMENT.

ROW	written in TRACK
1	2
2	3
3	6
4	7
5	10
6	11
7	14
8	15
9	4
10	5
11	8
12	9
13	12
14	13
15	1
16	16

FIG 11: INTERCONNECTIONS BETWEEN ROWS AND TRACKS.
(In order to improve concealment properties)

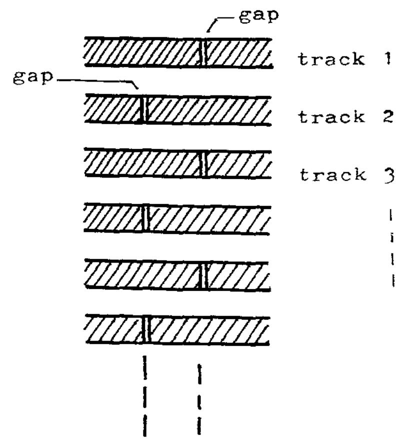


FIG 12: SPATIAL INTERLEAVING WITH THE HEADS.

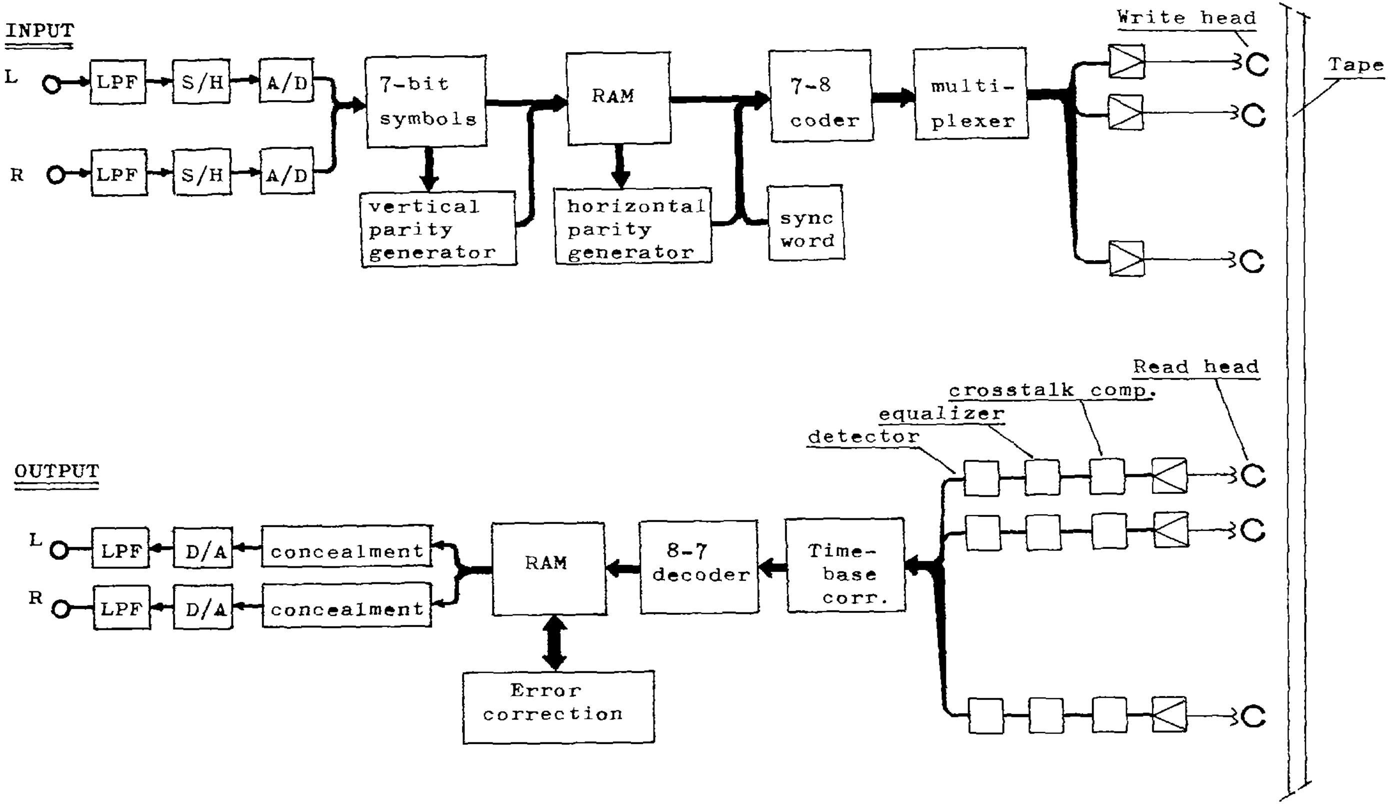


FIG 13: BLOCK DIAGRAM OF THE 2-CHANNEL, 16-TRACK PCM TAPE RECORDER.