MODULATION SYSTEMS FOR DIGITAL AUDIO DISCS WITH OPTICAL READOUT

by Kees A. Immink, Member IEEE

Philips Research Laboratories
The Netherlands

Abstract

This paper describes a new recording code format, Eight to Fourteen Modulation (EFM), designed for digital audio discs with optical readout. Attention is focused primarily on trade offs between conflicting parameters such as information density and d.c. content that led to the choice of the adopted format EFM.

Introduction

This paper describes the design of a modulation system of a digital audio disc with optical readout. The digital audio disc contains a spiral shaped track of successive shallow depressions, also called pits, in a reflective layer. The encoded audio information is stored in the lengths of these pits and in the distances between them. The read out is contactless; two servosystems follow the track in focus and radial direction within the desired accuracy. For details of optical recording the reader is referred to (9) and the special issue of applied optics July 1978.

The main specifications of the audio disc are: 2 audiochannels with a playing time larger than 1 hour, sample frequency 44.1 kc/s, 16 bits linear quantization.
Disc diameter: 120 mm.
The data rate after the anologue-digital converters is 32x44.1 kHz = 1.41 Mb/s. After the error control encoder (10) we yield a data rate of 4/3x1.41 = 1.88 Mb/s.
For the specified playing time we may calculate that this audiodisc is a storage medium with the phenomenal capacity of more than $6.5\ 10^9$ bits or 10^8 bits per cm^2.

A modulation system for an optical audio disc has to fulfil the following requirements:
- self clocking ability
- high information density
- small error propagation
- low power at low frequencies
- immunity against tolerances in the lightpath

This last requirement is not imposed because of limitations of optical recording to reproduce low frequencies as is mostly the case in magnetic recording. D.c. content causes interference with the servosystems and should be avoided.

In this paper we will discuss a code, called Eight to Fourteen Modulation (EFM), which meets a sound compromise between the aforementioned conflicting criteria.

Maximum data density, bandwidth restrictions

The present optical recording technology of Video Long Play and Compact Audio Disc restricts the input signal levels to only two: pit or land. So the information is just contained in the length of the pits and the distance between them. In a digital system these lengths and distances only take discrete values.

Considerations such as low power at low frequencies, selfclocking ability and limitations on intersymbol interference make it necessary that the input data be mapped (modulated or coded) into a sequence of binary data with some special properties. The self timing property e.g. sets an upper limit on the maximum length of the pits and lands. On the other hand intersymbolinterference imposes a lower limit upon the minimum length. The theory of binary sequences with restrictions on minimum and maximum feature size goes back to Kautz (1) and Tang et al. (2).

We adopt Tangs definitions:
A dk-limited sequence satisfies simultaneously the following conditions:

a. d-constraint - two logical ones are separated b a run of consecutive zeros of length at least d
b. k-constraint - any run of consecutive zeros is of length at most k.

When we integrate modulo 2 a dk sequence then the length of any run is at most k+1 and at least d+1. The number N of distinct dk sequences of a block of n bits can be calculated by a recursive relation. If we restrict here for convenience to the d-constraint then the number of distinct binary sequences of length n is given by (2):

$$N(n) = \begin{cases} 0 & n < 0 \\ n+1 & 0 \leq n \leq d+1 \\ N(n-1)+N(n-d-1) & n > d+1 \end{cases} \qquad (1)$$

The asymtotic information rate R of these sequence is determined by the specified constraints and is given by:

$$R = \lim_{n \to \infty} {}^2\log(N(n))/n \qquad (2)$$

We now come to the point to discuss the influence on intersymbol interference and consequently information density. Suppose we have an information source that transmits 1 bit per second in the format of an NRZ-code; i.e. a logical one is e.g. pit and a logical zero is land. In this case the minimum distance between transients is 1 second. Using integrated modulo 2 d-sequences for recording we are sure, that the transients are separated at least d, giving a physical distance of (d+1)-time units of the recorder clock. However the loss of rate, due to the redundancy of the d-sequence is R. So the actual minimum distance between transients in source time units is:

$$T_{min} = (d+1)R(d) \text{ (seconds)} \qquad (3)$$

In table I we tabulated for some values of d the minimum distance between level crossings that can be obtained.

d	R	(d+1).R
1	0.69	1.39
2	0.55	1.65
3	0.46	1.86
4	0.41	2.03
5	0.36	2.17

Table I. Rate and T_{min} versus d.

For instance, with d=4 we are able to guarantee a minimum distance between transients of about two data clock periods.

The relation between T_{min} and bandwidth properties of the code is not so clear. Some authors (3) claim, that the minimum bandwidth required to transmit the bitstream is simply related to T_{min}:

$$B = 1/(2\ T_{min}) \qquad (4)$$

[illegible] et al. (7) have shown that T_{min} is not a reliable indication of the bandwidth requirements. For a realistic channel (magnetic recording) bandwidths restrictions are derived by Jacoby (4). For optical recording the following results can be derived.

The most important part of the optical system is the objective lens. Its modulation transfer function MTF sets a limit to the maximum spatial frequency, that can be resolved. The MTF is determined by the numerical aperture (NA), the wavelength (λ) and the state of correction of the objective (5). The MTF is zero above the spatial cut off frequency (fc) given by:

$$f_c = 2NA/\lambda \qquad (5)$$

Noise in the optical read out system has been described by Heemskerk (6), who investigated several noise sources in a videodisk system. Most of his results apply to the case of a Compact Audio Disc. With this model of the read out mechanism and the noisesources we are now able to calculate the channel bit error rate (BER) of the theoretical channel vs. infodensity with T_{min} as parameter. We proceed in a way as presented by Tufts and Aaron (8). We assume, that the contribution to the interference is confined to m databits or equivalently to n = m/R channelbits (with R=R(d) or R(T_{min})). Corresponding to the N(m/R) possible d-limited choices of the truncated message sequence there are N(m/R) distinct "eye openings" at the sampling moments. The conditional errorprobabilities assuming a Gaussian probability function of the noise, are computed for each of the run length limited pulse sequences and then averaged with respect to the probability of occurence of these sequences.

In fig. 1 we plotted the locus of information density vs. T_{min} that can be attained at an arbitrarily high channel-bit error rate. We notice, that a maximum in information density can be attained if we choose a modulation system with d=2 or d=3. The difference with other choices is only slight. According to figure 1 the maximum info density equals for the specific conditions:

$$D/f_c \sim 2.2 \text{ at SNR} = 26 \text{ dB}$$

with D is information density (bits per meter).

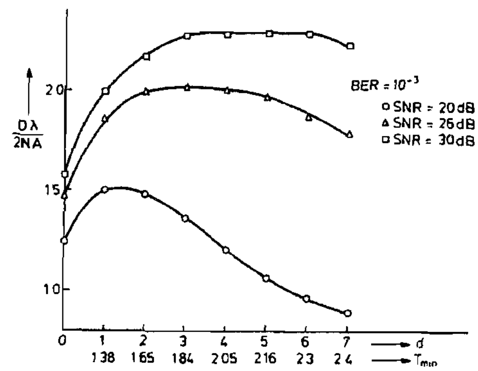

figure 1. Maximum achievable information density versus minimum pitsize (T_{min}) for a Gaussian channel. Substituting some practical values: SNR = 26 dB, NA = 0.4, λ = 780 nm yields a maximum density of 2.25 bits per micron.

In the calculations here presented we only took into account the "eye-opening" at the sampling moments. Another parameter, not yet discussed, is the self clocking ability of the modulation system. The desired accuracy of the timing (for a constant data rate) is proportional to R. However the clock content or average number of transients is related to the reciprocal of T_{min}. So from the point of view of timing we prefer a system with d=2 over d=3.

Another point in the favor of a d=2 system is that for practical implemented codes the rate is smaller (say 90%) than the theoretical maxima depicted in table 1. All these factors together lead to the choice of a modulation system with d=2.

Experiments with practical codes (in cooperation with engineers of SONY Corporation) led to the same conclusion.

EFM code

In the preceding chapter we studied the performance of some theoretical d-limited codes and noticed that a d=2 code is a quite optimal choice with respect to information density. A practical implementation of an encoder is most often based on a block code, that maps m consecutive data input bits into n channelbits. From the point of view of the error control system it is favourable to choose m=8, so that an 8-bitssymbol is encoded into an n bits channelsymbol. At the decoder we map the n bits channelsymbol again into an 8-bitssymbol. In this way we restrict errorpropagation to only one symbol. We assume that the receiver knows via some technical provision, the synchronization unit, where the beginning of each word is situated. From eq. (1) we derive that, for n=14, we have N=277 distinct 14-bits sequences satisfying the d=2 constraint. Deleting 21 patterns with the longest runlength yields an alphabet of 256 sequences, so that a unique one to one mapping by e.g. a look up table of 8 to 14 bits and vice versa is possible. The 14-bits blocks however cannot be concatenated without violating the d-constraint at the boundaries. The insertion of two merging bits between successive blocks, where normally no transient occurs, is sufficient. The two merging bits do not contain any information and are skippped by the decoder.

It appears, that the maximum runlength is k=10, if the merging bits are used (with preserving the d-constraint) to insert a single transient between the blocks if the runlength is larger than 10 by concatenating the blocks.
So the modulator looks one 8-bitssymbol ahead, looks up its 14-bits channelrepresentation and decides whether a transient is needed in the dummy bits to preserve the k-constraint.

There are cases where the merging bits are not uniquely determined by the concatenation rules, this freedom of choice is used for minimizing the power at low frequencies with e.g. the DSV (digital sum variation) as criterion. The DSV is defined as the integral of the modulation stream with a logical zero counted as minus one. Our experiments showed that this DC-control is not sufficient, so we increased the number of merging bits to 3, which means that in 65% of all mergings an extra transient can be set or omitted freely. This more effective DC-control costs however 1/17 or 6% of the information density. Depending on the exact strategy one may minimize the power at low frequencies. In fig. 2 a/b we depicted the power density spectrum of EFM with two different strategies. Curve a results if a one symbol look ahead strategy is implemented e.i. the decision of placing or omitting a transient (with preserving the d- and k constraint!) is only based on the knowledge of one future symbol.
The decision is based on the minimization of the DSV after the concatenation of the new block.
For comparison we plotted the power density spectrum of two modulation schemes known from the magnetic recording field M2 and 3PM (4,11,12).
Curve b is the result if a more sophisticated DC control is applied based on the knowledge of 4 future symbols. We note that a net 10 dB improvement is possible in the frequency range<10 kHz.

Experiments showed that no significant interference with the servosystems appeared if a one symbol look ahead strategy was applied.

Conclusions

A new digital coding format, EFM, has been described, which should be very useful in digital audio disk systems with optical read out. Background and design parameters which led to the actual choice of the adopted modulation system has been elucidated.

The choice of system with d=2, k=10 was based on many experiments and simulations, that showed a superior performance over other choices.

Small power at low frequencies is an important parameter that could be attained by using the degree of freedom in the merging bits that concatenate successive blocks. The blockstructure on an 8-data bits basis makes EFM very well suited for the adopted Reed-Solomon error correction system.

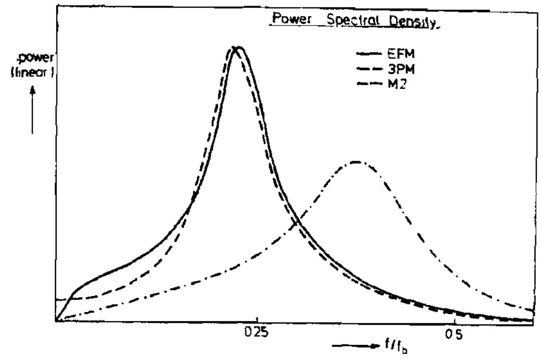

figure 2ª Power density function of EFM, one and four point look ahead strategy compared with 3PM and M2, f_b is the bitrate.

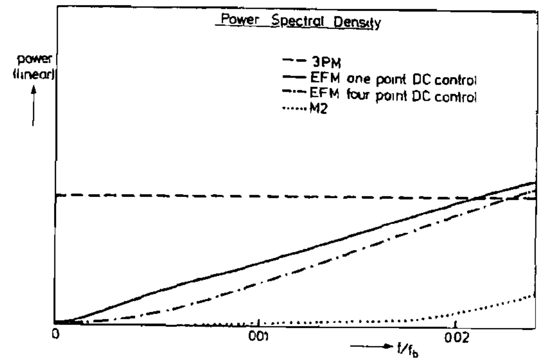

figure 2ᵇ Lower part of the power density functions.

References

(1) W.H. Kautz, IEEE Trans. Inform. Theory, vol. IT-11, p. 285 (1965).
(2) D.T. Tang and L.R. Bahl, Inform. Contr. vol. 17, p. 436, 1970.
(3) P.D. Shaft, IEEE Trans. Comm., vol. COM-16, p. 687, (1973).
(4) G.V. Jacoby, IEEE Trans. Magn., vol. MAG-13, no. 5, p. 1202 (1977).
(5) G. Bouwhuis and J. Braat, Applied Optics vol. 17, p. 1993, 1978.
(6) J. Heemskerk, Appl. Optics vol. 17, p. 2007, 1978).
(7) M.G. Pelchat and J.M. Geist, IEEE Trans. on Comm., vol. COM-23, no. 9, p. 878, 1975.
(8) M.R. Aaron and D.W. Tafts, IEEE Trans. Inform. Theory, IT-12, pp. 26 (1966).
(9) K. Compaan, Philips Tech. Rev., p. 1978 (1973).
(10) T. Doi, Conference Paper ASSP, Atlanta, 1981.
(11) J.C. Mallinson and J.W. Miller, Radio and Elec. Eng., 47, p. 172 (1977).
(12) D.A. Lindholm, IEEE Trans. on Magn., vol. MAG-14, p. 321, (1978).