SPECTRUM SHAPING CONSTRAINED CODES FOR RECORDING

B. Vasic[1], *S. Denic*[2], *R.Rathnakumar*[1]

[1] University of Arizona, USA, [2] University of Ottawa, Canada

ABSTRACT

The goal of this paper is to give a survey of spectrum shaping codes used for digital recording systems. This class of codes belongs to the broader class of modulation codes, which are widely used in recording systems for adjusting the source characteristics to the characteristics of the recording channel. Shannon noiseless capacities of recording channels will be considered, as well as the spectra of maxentropic sequences of *M*-ary recording constraints. In addition, some practical encoding and decoding schemes will be discussed.

1. INTRODUCTION

The spectrum shaping codes are widely used in recording systems [1]-[3]. Their role is to transform a source sequence into a channel sequence whose spectral characteristics correspond to the spectral characteristics of the channel. The simplest spectrum shaping codes are dc-free codes that suppress dc frequency component. In optical recording systems these codes are used to circumvent the interference between the recorded signal and the servo tracking system [4], [5]. There are class of spectrum shaping codes that supports the use of frequency multiplexing techniques for track following [1] and partial response techniques for high density data storage [6]. Both techniques require the spectral nulls of the recorded signal to be at frequencies other than f = 0 [7], [8]. The fourth characteristic group of spectrum shaping codes, that give rise to spectral lines, are used to provide reference information to head positioning servo systems, which positions and maintains the head accurately over a track in digital recorders [1]. Spectrum shaping codes are also expected to play a vital role in future high-density recording systems.

The goal of this paper is to give a survey of spectrum shaping codes for digital recording systems from the theoretical and the practical point of view. Section 2 describes briefly, the role of spectrum shaping codes in a recording system. Section 3 discusses dc-free codes and some theoretical bases for studying these codes. This discussion is also important when studying all other types of spectrum shaping codes. Section 4 discusses codes that have higher order spectral zeros at zero frequency, and codes that have spectral zeros at frequencies that are sub-multiples of the channel symbol frequency. Finally, in Section 5, compound constraint codes are briefly visited.

2. RECORDING SYSTEM AND SPECTRUM SHAPING CODES

The model of a recording system shown in Fig. 1 resembles that of a communication system. The spectrum shaping encoder receives a symbol stream $\mathbf{c}=\{c_k\}_{k=0}^{\infty}$ from an error correcting encoder and transforms it into a stream of channel symbols $\mathbf{x}=\{x_k\}_{k=0}^{\infty}$ that matches the spectral characteristics of the recording channel with impulse response h(t). Additive noise is denoted by n(t). The spectrum shaping decoder accepts data stream from the channel $\mathbf{y}=\{y_k\}_{k=0}^{\infty}$, and transforms it into the symbol stream $\hat{\mathbf{c}}=\{\hat{c}_k\}_{k=0}^{\infty}$ that represents the estimation of the symbol stream $\mathbf{c}$. From the point of view of an error correcting encoder and decoder, the spectrum shaping encoder and decoder are merely parts of the recording channel. In case of block codes, the input stream $\mathbf{c}$ is divided into sourcewords of length *k*, which are encoded into codewords of length *l* forming the output stream $\mathbf{x}$.

3. DC - FREE CONSTRAINED CODES

Dc-free codes belong to a class of spectrum shaping codes that transform the spectrum of an input sequence into an encoded

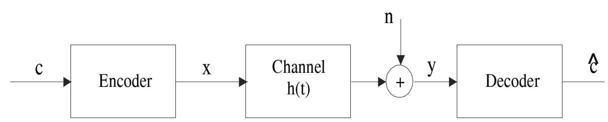

Fig. 1 Recording system

sequence whose spectrum has zero at the zero frequency. First important result for analysis and design of dc-free codes is by Pierobon [3]. In order to state the result of Pierobon, the notion of running digital sum has to be defined. The running digital sum (RDS) at moment *n*, z_n, of sequence of symbols $\mathbf{x}=\{x_k\}_{k=0}^{\infty}$ is defined as

$$z_n = \sum_{k=0}^{n} x_k \ . \tag{1}$$

Pierobon proved that the power spectral density (PSD) $S_x(f)$ of a sequence $\mathbf{x}$ vanishes at zero frequency, regardless of the sequence distribution, if and only if the RDS, $\mathbf{z}=\{z_k\}_{k=0}^{\infty}$ of the encoded sequence is bounded, i.e., $|z_n| \le K$, for each *n*, when -$K<\infty$. Any sequence that violates this condition on RDS will not be transmitted through the recording channel. The sequences that satisfy this constraint are called dc-free sequences. Thus, it follows that dc-free codes should map the source sequences into the sequences which have bounded RDS. The channel capacity *C* of a noiseless constrained channel is defined as

$C = \lim_{T\to\infty} \log_2 m(T)/T$ where m(T) is the number of constrained sequences of length T [9]. The importance of the channel capacity stems from the fact that it determines the maximal theoretical code rate R of a code [9]. In the case of dc-free codes, the channel capacity is determined by an upper bound on the RDS, K. In order to compute the channel capacity of dc-free constrained channels, the notion of digital sum variation (DSV) is introduced. If N_{max} is the maximum value of z, and N_{min} is its minimum value then the DSV is defined as $N= N_{max} - N_{min} + 1$. One way to compute the channel capacity of constrained channels is to use the connection matrix D of the finite state transition diagram (FSTD) that describes the channel constraint. An example of FSTD representing dc-free constraint is given in Fig. 2.

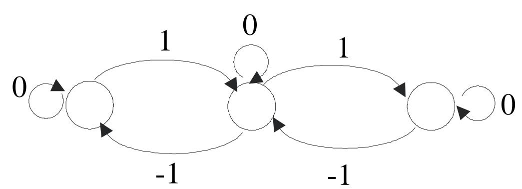

Fig. 2 FSTD for M=3, N=3 dc-free constraint

The state of the FSTD represents the value of RDS, z_k, at time instant k. A label above the edge denotes the channel symbol, which belongs to the set X={-1,0,1} generated at the time of transition between states. The number of states is equal to N. M denotes the cardinality of the set of channel symbols X. The connection matrix D of FSTD is a square matrix of dimension N, where the entry d_{ij} of matrix D represents number of edges emanating from state i and ending at state j. Each FSTD may be assigned a corresponding Markov chain $\mathbf{s} = \{s_k\}_{k=0}^{\infty}$. It was proved that the channel capacity is given by $C = \max H(\mathbf{s}) = \log_2 \lambda_{\max}$ [9], where λ_{max} is the maximum eigenvalue of the connection matrix D and $H(\mathbf{s})$ is the entropy of the Markov chain. Thus, the channel capacity is equal to the maximal entropy of the Markov chain **s** that corresponds to a particular FSTD. The sequences that are generated by the Markov chain that maximizes the entropy of corresponding FSTD are called maxentropic sequences.

3.1. Spectral characteristics of dc-free constraint

The performance of spectrum shaping codes is determined by observing the PSD (S_x) of the encoder's output. The parameter that is of great practical importance is the cut off frequency. The cut off frequency is defined as the value of the frequency at which $S_x(f_0)=0.5$. The cut off frequency determines the bandwidth, called the notch width (f=0 to $f=f_0$), within which the power spectral density of the recorded sequence is low. The performance is better if the cut-off frequency f_0 is larger. In [10], Justesen derived useful relationship between the sum variance (variance of RDS) $E[z_k^2]=\sigma_z^2(N)$, and the cut-off frequency ω_0. It was shown that

$$2\sigma_z^2(N)\omega_0 \approx 1 \tag{2}$$

This formula implies that larger the cut-off frequency ω_0, smaller the sum variance $\sigma_z^2(N)$. Another important measure of a code quality is the product of redundancy and the sum variance $(1-C(N))\ \sigma_z^2(N)$, which is tightly bounded from below and above for N>9 in the case of maxentropic dc-free sequences [1]. This shows that larger low frequency suppression requires larger redundancy and vice versa. Thus, the compromise has to be made in code design. Fig. 3 shows PSD of maxentropic sequences for M=3, and N=3, 4, 5 versus normalized frequency f where frequency is scaled by $f_s=1/T_s$, where T_s is the signaling interval.

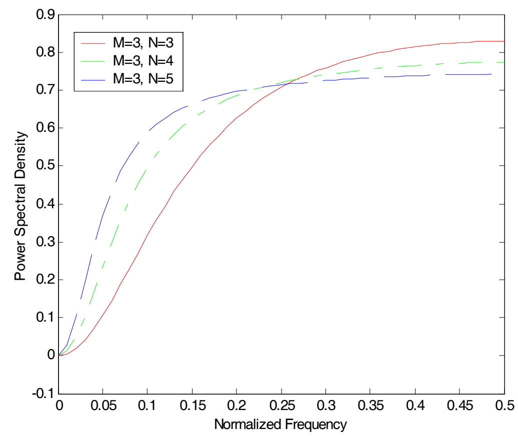

Fig. 3 PSD of M=3, N=3,4,5 maxentropic sequences

3.2. Encoding and decoding of dc-free constraints

The quality of a particular encoder for a constraint channel is determined by the encoder efficiency E, which is defined as [1]

$$E = \frac{(1-C(N))\sigma_z^2(N)}{(1-R)s_z^2} \tag{3}$$

The encoder efficiency compares the product of redundancy and the sum variance of the maxentropic sequences with that of the specific encoder for the same channel. The maxentropic sequences were taken as a reference because it achieves the channel capacity of a dc-free constraint channel.

In general, encoders, and decoders can be divided into two groups, namely state independent and dependent. State independent encoders are usually realized in the form of look-up tables, while state dependent encoders are designed as finite state machines [11]. The advantage of the state dependent encoding is that sometimes more efficient codes can be constructed (in terms of a code rate) with shorter codewords [1]. Similarly, the state independent and dependent decoders can be described [11]. The weakness of state dependent decoders is catastrophic error propagation and as a result, state independent decoders were invented, which are called sliding window decoders [11].

Having in mind that RDS has to be bounded, and that the recorded symbols could take both positive and negative values, the notion of codeword disparity is introduced. The disparity d of codeword $\mathbf{x}=(x_1\ x_2 \ldots x_n)$ of length n is defined as

$$d = \sum_{i=1}^{n} x_i \quad (4)$$

Based on the codeword disparity, there are three simple ideas that are employed in the design of state independent encoders. The first is the use of codewords that have zero disparity. This type of state independent encoding is inefficient because the number of possible codewords with zero disparity of length n is limited. Another way for achieving better efficiency is to use so-called low disparity codes [12]. The drawback of this technique is the increase in the power of low frequency components as compared to zero disparity coding. The polarity bit coding is yet another simple method that generates dc-free sequences [13]. One extra bit, called polarity bit, is added to n-1 source bits comprising codeword of length n. The polarity bit is set to 1. If the disparity of the codeword has the same sign as RDS at the moment of sending the codeword, the inverted codeword is recorded. Otherwise original codeword is recorded. It is shown that low disparity codes outperform zero disparity codes, and zero disparity codes outperform polarity bit codes [1]. Also, for small codeword lengths, low and zero disparity codes have unit efficiency, while the efficiency drops as the codeword length increases. For other encoding techniques that do not employ look-up tables (for example, those that are based on enumeration algorithms) see [1]. For an example of state dependent encoders see Section 5 of this paper.

4. CODES WITH HIGHER ORDER SPECTRAL ZEROS

Further progress of spectrum shaping codes was in two directions. One corresponded to better suppression of low frequency components, while the other corresponded to the introduction of spectral zeros at frequencies other than zero (related to use of partial response signaling techniques for high density recording systems [6]). The improvement was achieved by generalization of the basic concepts given in previous sections [4].

The k^{th} order RDS at frequency f = pf_s/q (K-RDS_f) of sequence $\mathbf{x} = \{x_k\}_{k=0}^{n}$ is defined as

$$\sigma_k^f(x) = \sum_{i_1=0}^{n} \sum_{i_2=0}^{i_1} \cdots \sum_{i_k=0}^{i_{k-1}} w^{i_k} x_{i_k} \quad (5)$$

where p and q are relatively prime and $w=e^{-j2\pi p/q}$. An FSTD is said to be a K-RDS_f FSTD if there is a mapping ψ from the set of states $\boldsymbol{\Sigma}$ onto a finite set of complex numbers such that $x(D)=wz(D)(1-w^{-1}D)^K$, where $\mathbf{x} = \{x_k\}_{k=0}^{n}$ is a channel sequence, $z_n=\psi(s_{n+1})$, $\mathbf{s} = \{s_k\}_{k=0}^{n}$ is a state Markov process ($s_k \in \Sigma$), and $a(D) = \sum_{k=0}^{\infty} a_k D^k$ defines D transform of a sequence $\mathbf{a}=\{a_k\}_{k=0}^{\infty}$.

The spectrum of a sequence $\mathbf{x}$ generated by the state process $\mathbf{s}$ assigned to an FSTD has a spectral null of order K at f = pf_s/q, if and only if FSTD is a K-RDS_f FSTD [4]. It should be noted that spectral null of order K guarantees that first 2K-1 derivatives of PSD $S_x(f)$ vanishes at frequency f. It is shown in [4] that the dynamics of K-RDS_f FSTD with respect to time can be described by

$$\sigma_{n+1}^f = \mathbf{A}\sigma_n^f + w^n \mathbf{1} x_n \quad (6)$$

where $\sigma_n^f = w^n \psi_n$ is a K-dimensional column vector, $\boldsymbol{A}$ is a lower triangular all one matrix, and $\mathbf{1}$ is a K-dimensional all one vector. If w=1 then this formula is valid for the case of spectral null at f=0. If $\sigma_0^f = 0$, then the k^{th} element of vector σ_n^f represents K-RDS_f of sequence $\mathbf{x}$ at moment n. The equation (6) is related to the concept of canonical state transition diagram [7], denoted by D_K^f. These diagrams have a countable infinite number of states, and there exists a finite state sub-diagram of D_K^f that generates sequences with K^{th} order spectral null at frequency f. These finite state sub-diagrams enable the computation of capacities, PSD's of the constraints and encoder/decoder design. In Fig. 4, the PSD's of two maxentropic sequences are shown, i) with first order spectral null, and parameters M=2, N=3, and ii) with second order spectral null, and parameters M=2, V_1= V_2=3, at frequency f=0, where V_1 is the number of values that 1-RDS_0 can take, and V_2 is the number of values that 2-RDS_0 can take.

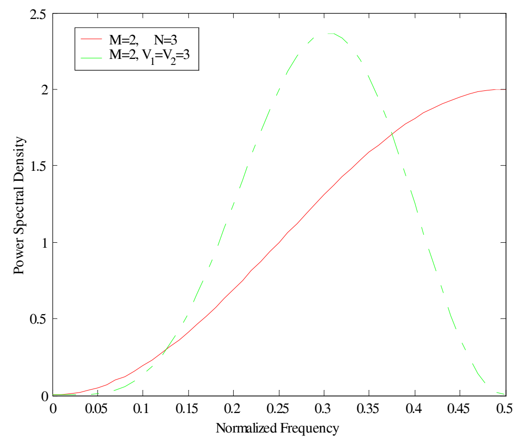

Fig. 4 PSD of higher order spectral zeros

The sequences with second order spectral null have better rejection of low frequency components although the notch width is wider in the case of sequences with first order spectral null. The relation between cut-off frequency, and sum variance (2) is no longer valid.

A number of constructions of codes with higher order spectral zeros, based on FSTD's can be found in [4]. Finally, it should be noted that codes that generate K-RDS_f constrained sequences improve the error correcting capabilities of recording systems [14].

5. COMPOSITE CONSTRAINED AND COMBINED ENCODING

Like any communication system, a recording system employs different encoding techniques such as source encoding, channel encoding and modulation encoding. It means that a channel sequence has to satisfy different kinds of constraints in order to be reliably recorded. Therefore, there is a need for codes that generate sequences, which satisfy composite constraints. One example of such codes are the ones that are both RLL (run length limited) and dc-free [15], [16]. RLL codes are widely used in digital recording systems. They confine minimal and maximal number of consecutive like symbols in a recording channel to prevent intersymbol interference and to enable clock recovery. In a wide sense, the combination of dc-free codes and error correcting codes that improves the performance of dc-free codes on noisy recording channels, also belong to this group. Discussion on these codes can be found in [17]-[20].

As an example, we present M=3, N=3 (d=1,k=2) RLL dc-free code in Fig. 5. The Shannon capacity of corresponding constraint is C=0.4650 bits/sym, and the chosen code rate is R=1/3 bits/sym. In order to get the encoder, the algorithm from [11] was applied. The sliding window decoder for this code can be found in [16].

At the end, in Fig. 6, the spectra of M-ary RLL dc-free maxentropic sequences are shown, for fixed DSV, N=7, (d=1,k=3) and different values of M, i) M=3, ii) M=4, iii) M=5, iv) M=7.

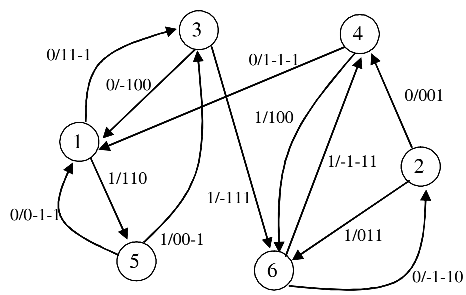

Fig. 5 Encoder of M=3, (d=1,k=2), N=3, RLL dc-free codes

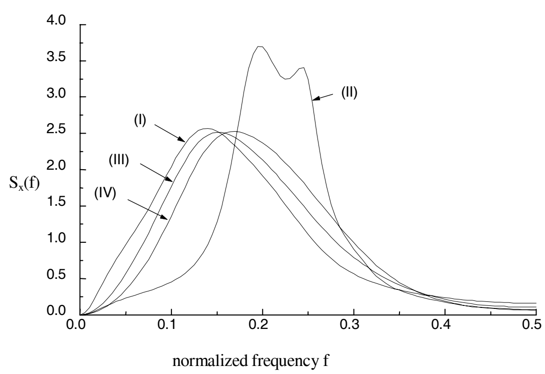

Figure 6 Spectrum of RLL dc-free maxentropic sequences

6. CONCLUSION

This paper provides an overview of basic concepts and ideas of spectrum shaping codes for digital recording systems. We considered the theoretical and the practical aspects of four groups of spectrum shaping codes, dc-free codes, codes with higher order spectral null at f=0, codes with spectral nulls at the submultiples of channel symbol frequency, and codes with composite constraints.

7. ACKNOWLEDGEMENT

This work is supported by the NSF Grant CCR- 0208597 and by Seagate Technology.

8. REFERENCES

[1] Immink, K.A.S., *Coding Techniques for Digital Recorders*, 1st ed., Prentice Hall International, New York, 1991.

[2] Immink, K.A.S., Siegel, P.H, and Wolf, J.K., Codes for digital recorders, *IEEE Trans. Inform. Theory,* 44, 2260, 1998.

[3] Pierobon, G.L., Codes for zero spectral density at zero frequency, *IEEE Trans. Inform. Theory,* 30, 435, 1984.

[4] Eleftheriou, E., Cideciyan, R.D., On codes satisfying Mth order running digital sum constraint, *IEEE Trans. Inform. Theory*, 37, 1294, 1991.

[5] Immink, K.A.S., Beenker, G.F.M, Binary transmission codes with higher order spectral zeros at zero frequency, *IEEE Trans. Inform. Theory*, 33, 452, 1987.

[6] Kobayashi, H., Correlative level coding and maximum likelihood decoding, *IEEE Trans. Inform. Theory*, 17, 586, 1971

[7] Marcus, B.H., Siegel, P.H., On codes with spectral nulls at rational submultiples of the symbol frequency, *IEEE Trans. Inform. Theory*, 33, 557, 1987.

[8] Karabed, R., Siegel, P.H., Matched spectral-null codes for partial-response channels, *IBM Res. Rep.* RJ 7092, Apr. 1990.

[9] Shannon, C.E., A mathematical theory of communications, *Bell Syst. Tech. J.,* 27, 379, 1948.

[10] Justesen, J., Information rate and power spectra of digital codes, *IEEE Trans. Inform. Theory,* 28, 457, 1982.

[11] Marcus, B.H., Siegel, P.H., Wolf, J.K., Finite state modulation codes for data storage, *IEEE Jour. Selec. Area Commun.,* 10, 5, 1992.

[12] Cattermole, K.W., Principles of digital line coding, *Int. Jour. Electron.*, 55, 3, 1983.

[13] Carter, R.O., Low disparity binary coding system, *Electronics Letters*, 1, 65, 1965.

[14] Immink, K.A.S., Beenker, G.F.M, Binary transmission codes with higher order spectral zeros at zero frequency, *IEEE Trans. Inform. Theory*, 33, 452, 1987.

[15] Norris, K., Bloomberg, D.S., Channel capacity of charge-constrained run-length limited codes, *IEEE Trans. Magn.*, 17, 3452, 1981.

[16] Denic, S.Z., Vasic, B., Stefanovic, M.C., M-ary RLL dc-free codes for optical recording channels, *Elect. Letters,* 36, 1214, 2000.

[17] Popplewell, A., O'Reilly, J., A simple strategy for constructing a class of DC-free error-correcting codes with minimum distance 4, *IEEE Trans. Inform. Theory*, 41, 1134, 1995.

[18] Deng, R.H., Li, Y.X., Herro, M.A., DC-free error-correcting convolutional codes, *Elect. Letters* , 29, 1910, 1993.

[19] Lee,S.I., DC- and Nyquist-free error correcting convolutional codes, *Elect. Letters* , 32, 2196, 1996.

[20] Etzion, T., Constructions of error-correcting DC-free block codes, *IEEE Trans. Inform. Theory*, 36, 899, 1990.