Upper Bound on the Efficiency of dc-Constrained Codes

By TA-MU CHIEN

(Manuscript received June 2, 1970)

We derive the limiting efficiencies of dc-constrained codes. Given bounds on the running digital sum (RDS), the best possible coding efficiency η, for a K-ary transmission alphabet, is $\eta = \log_2 \lambda_{max}/\log_2 K$, where λ_{max} is the largest eigenvalue of a matrix which represents the transitions of the allowable states of RDS. Numerical results are presented for the three special cases of binary, ternary and quaternary alphabets.

I. INTRODUCTION

In digital transmission systems, the transmission channel often does not pass dc. This causes the well-known problem of baseline wander. One way to overcome this difficulty is to restrict the dc content in the signal stream using suitably devised codes.[1–3] As a result many codes having a dc-constrained property have been studied.[4–9] The coding requirement is represented by the constraint put upon the running digital sum (RDS) of the coded signal stream. We expect that the efficiency of a dc-constrained code is related to the limits of RDS in some definite way. This is the subject to which we address ourselves in this paper. More specifically, we intend to answer the question: What is the best possible efficiency of any dc-constrained code satisfying a given limit on RDS?

Let $\{a_1, a_2, \cdots\}$ be the sequence of the transmitted symbols, the RDS of the signal stream at instant k is defined to be the sum $\sum_{i=1}^{k} a_i$. Taking the RDS at any instant as the state of the signal stream at that point, the limits on RDS define a set of allowable states, and each additional signal symbol may be considered as a transition from one state to another. This transition can be represented by a matrix-called naturally the transition matrix. For a K-ary signal alphabet, the best possible efficiency η of dc-constrained codes is found to be

$$\eta = \frac{\log_2 \lambda_{max}}{\log_2 K} \tag{1}$$

where λ_{max} is the largest eigenvalue of the transition matrix.

The efficiency of a code is defined to be the ratio of the average bits per symbol of the coded signal stream to that of the random (uncoded) signal stream.

McCullough[4] has derived the same result (1) for the special cases of $K = 2$ and 3. His approach is quite different from what will be presented in the sequel.

We first describe in detail the construction of a mathematical model for the case of a binary alphabet. Then we generalize the result of the binary case to include any alphabet set. Methods of effecting numerical calculation are discussed as well as approximation formulas. The numerical results for three important cases are presented and known codes are compared with the theoretical limits.

II. LIMITING EFFICIENCY OF THE BINARY CODES

In this section we confine our discussion to binary signals and direct our attenion to the intuitive reasoning which leads to the construction of a simple mathematical model and its interpretation.

Let M, a positive integer, be the desired bound on the RDS of the coded binary signal stream. This defines a subset $S_M(\infty)$ of the set $S(\infty)$ of all infinite binary sequences in the following way: An infinite sequence is in the subset $S_M(\infty)$ if the RDS of the sequence is nowhere larger than M or less than $-M$, i.e., $|\sum_{i=1}^{k} a_i| \leqq M$ for $k = 1, 2, \cdots$. A sequence in $S_M(\infty)$ is called an allowable sequence. Denoting by $N_M(\infty)$ and $N(\infty)$ the number of infinite sequences in $S_M(\infty)$ and $S(\infty)$ respectively, the average information per symbol for the sequences in $S_M(\infty)$ is given by

$$\eta = \frac{\log_2 N_M(\infty)}{\log_2 N(\infty)}, \tag{2}$$

assuming the ratio exists. If we interpret the set $S(\infty)$ as source data and $S_M(\infty)$ as the transmitted signal, then η defined in equation (2) is the efficiency of a dc-constrained code which maps one-to-one from $S(\infty)$ onto $S_M(\infty)$.*

Clearly for any code which satisfies the requirement that RDS be bounded by M, the coded signal stream must be a member of $S_M(\infty)$. Therefore, the set of allowable infinite sequences defined by any code satisfying the desired constraint on RDS must be a subset of $S_M(\infty)$.

* The puzzle of mapping a large set to a small set can be cleared mathematically by observing that the cardinality of both $S(\infty)$ and $S_M(\infty)$ are that of a continuum, and physically by demanding that the transmitter has a higher baud than the source.

Thus we conclude that the formal expression in (2) indeed gives the best possible efficiency for a given bound M. Our next step is to find a way to count the number of allowable sequences in $S_M(\infty)$.

Let us start by counting the allowable sequences of finite length L. Define an occupancy vector of the allowable states $\mathbf{u}_L$, $'$ denoting the transpose of a vector (or matrix),

$$\mathbf{u}_L = [u_{-M} \cdots u_0 \cdots u_{+M}]', \tag{3}$$

where u_k, $k = -M, \cdots, M$, is the number of allowable sequences of length L with their RDS at end equal to k, i.e., $\sum_{i=1}^{L} a_i = k$. The total number of allowable sequences of length L, $N_M(L)$ is simply

$$N_M(L) = \sum_{k=-M}^{M} u_k . \tag{4}$$

As $L \to \infty$, $N_M(L) \to N_M(\infty)$ and the total number of sequences of length L, $N(L) = 2^L \to N(\infty)$. Hence we can rewrite (2) as

$$\eta = \lim_{L\to\infty} \frac{\log_2 N_M(L)}{L}. \tag{5}$$

Now our job is to find a formula for the number of allowable sequences of finite length.

Suppose we know the occupancy vector $\mathbf{u}_L$ and we want to calculate the occupancy vector $\mathbf{u}_{L+1}$. Clearly for any allowable sequence of length $L + 1$, its first L elements must be one of the allowable sequences of length L. We generate, therefore, the allowable sequences of length $L + 1$ from that of length L by adding one more binary symbol ($+1$ or -1). Therefore, the sequences of length $L + 1$ in the $-M$th state are generated by adding -1 to the sequences of length L in the $-M +$ 1st state; the sequences of length $L + 1$ in the $-M +$ 1st state are generated by adding $+1$ to the sequences of length L in the $-M$th state and by adding -1 to the sequences of length L in the $-M +$ 2nd state; etc. It is not difficult to see that the new state occupancy vector is

$$\mathbf{u}_{L+1} = \begin{bmatrix} u_{-M+1} \\ u_{-M} + u_{-M+2} \\ u_{-M+1} + u_{-M+3} \\ \vdots \quad\quad \vdots \\ u_{M-2} + u_M \\ u_{M-1} \end{bmatrix}. \tag{6}$$

Equivalently, $\mathbf{u}_{L+1}$ can be written as

$$\mathbf{u}_{L+1} = A_{2M+1}\mathbf{u}_L \tag{7}$$

where

$$A_{2M+1} = \begin{bmatrix} 0 & 1 & & & & 0 \\ 1 & 0 & 1 & & & \\ & 1 & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & 1 \\ & 0 & & & 1 & 0 \end{bmatrix} \tag{8}$$

is a square matrix of size $2M + 1$ with ones in the superdiagonal and the subdiagonal and zeros elsewhere. A_{2M+1} is the transition matrix of the allowable states. By the same reasoning, we have,

$$\begin{aligned} \mathbf{u}_L &= A_{2M+1}\mathbf{u}_{L-1} \\ \mathbf{u}_{L-1} &= A_{2M+1}\mathbf{u}_{L-2} \\ &\vdots \end{aligned} \tag{9}$$

and

$$\mathbf{u}_1 = A_{2M+1}\mathbf{u}_0$$

where $\mathbf{u}_0$ is the occupancy vector of the sequence of zero length. It is defined naturally with one at the zeroth state and zeros elsewhere,

$$\mathbf{u}_0 = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} . \tag{10}$$

From equation (9), we obtain, by successive substitution,

$$\mathbf{u}_L = A_{2M+1}^L \mathbf{u}_0 \ . \tag{11}$$

The total number of allowable sequences, from equation (4), is

$$N_M(L) = \mathbf{1}' A_{2M+1}^L u_0 \tag{12}$$

where

$$\mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}. \tag{13}$$

The class of matrices A_n defined in equation (8) has many interesting properties. Their investigation is relegated to the Appendix.

Using the result derived in the Appendix, and adapting the following ordering or eigenvalues of A_{2M+1}:

$$\lambda_{-M} < \lambda_{-M+1} < \cdots < \lambda_{M-1} < \lambda_M ,$$

we can rewrite equation (12),

$$N_M(L) = \mathbf{1}' P D_A^L P' \mathbf{u}_0 \tag{14}$$

where

$$D_A = \begin{bmatrix} \lambda_{-M} & & 0 \\ & \ddots & \\ 0 & & \lambda_M \end{bmatrix}, \tag{15}$$

$$P = \begin{bmatrix} \phi_0(\lambda_{-M})/\phi(\lambda_{-M}) & \cdots & \phi_0(\lambda_M)/\phi(\lambda_M) \\ \phi_1(\lambda_{-M})/\phi(\lambda_{-M}) & \cdots & \phi_1(\lambda_M)/\phi(\lambda_M) \\ \vdots & & \vdots \\ \phi_{2M}(\lambda_{-M})/\phi(\lambda_{-M}) & & \phi_{2M}(\lambda_M)/\phi(\lambda_M) \end{bmatrix} \tag{16}$$

from Lemma 7 of the Appendix and $\phi(\lambda)$ and $\phi_i(\lambda)$ are defined in equation (54) and (70). By straightforward multiplication, we can write, from equation (14),

$$N_M(L) = \sum_{i=-M}^{M} \lambda_i^L \phi_{M+1}(\lambda_i) \sum_{j=0}^{2M} \phi_j(\lambda_i), \tag{17}$$

where the normalization constants $\phi(\lambda_i)$ are omitted for simplicity. Denote by $\lambda_{\max}$ the largest absolute value of the eigenvalues of A_n, and from Lemmas 3 and 6 of the Appendix we know that

$$\lambda_{\max} = \lambda_M = -\lambda_{-M} . \tag{18}$$

Then from equation (17) and Lemma 2,

$$N_M(L) = \lambda_M^L\left\{\phi_{M+1}(\lambda_M) \sum_{j=0}^{2M} \phi_j(\lambda_M) + (-1)^L\phi_{M+1}(\lambda_{-M}) \sum_{j=0}^{2M} \phi_j(\lambda_{-M})\right\} + \sum_{i=-M+1}^{M-1} \lambda_i^L\phi_{M+1}(\lambda_i) \sum_{j=0}^{2M} \phi_j(\lambda_i)$$

$$= \lambda_M^L\phi_{M+1}(\lambda_M) \sum_{j=0}^{2M} [1 + (-1)^{L+M+j+1}]\phi_j(\lambda_M) + \sum_{i=-M+1}^{M-1} \lambda_i^L\phi_{M+1}(\lambda_i) \sum_{j=0}^{2M} \phi_j(\lambda_i). \qquad (19)$$

Since $\phi_j(\lambda_M) > 0$ for all j (see the proof of Lemma 4), the coefficient of the λ_M^L term in equation (19),

$$\phi_{M+1}(\lambda_M) \sum_{j=0}^{2M} [1 + (-1)^{L+M+j+1}]\phi_j(\lambda_M) > 0 \qquad (20)$$

independent of L.

Substituting equation (19) in (5) and using (18), we have

$$\eta = \lim_{L\to\infty} \frac{1}{L}\left\{\log_2 \lambda_{\max}^L\left[z_{\max} + \sum_{i=-M+1}^{M-1} \left(\frac{\lambda_i}{\lambda_{\max}}\right)^L z_i\right]\right\}$$

$$= \log_2 \lambda_{\max} + \lim_{L\to\infty} \frac{1}{L} \log_2 \left[z_{\max} + \sum_{i=-M+1}^{M-1} \left(\frac{\lambda_i}{\lambda_{\max}}\right)^L z_i\right] \qquad (21)$$

where $z_{\max}$ and z_i are the coefficients of λ_M and λ_i, $i \neq \pm M$ in equation (19). The second term in equation (21) approaches zero as a limit since $z_{\max} > 0$. Thus we have the desired result for the binary case

$$\eta = \log_2 \lambda_{\max}. \qquad (22)$$

Actually, we have proved a result more general than (22). Observe that, in passing to the limit, the crucial point is that $z_{\max}$ in (21) be nonzero. From equation (20) and the fact that $\phi_i(\lambda_M) \neq 0$ for $i \leqq 2M$, we conclude that the particular $\mathbf{u}_0$ we use, though natural, is immaterial, and any vector with non-negative coordinates will serve the purpose. Observe also the actual values of the allowable RDS state nowhere enter into our discussion, hence, it is immaterial whether the bound on RDS be symmetric or not. We can consolidate our discussion by stating the following theorem.

Theorem 1: For a binary alphabet, if the RDS of a coded signal stream is required to be within some bound M^+ *and* M^-, *where* M^+ *and*

M^- are integers, then the best possible coding efficiency is given by

$$\eta = \log_2 \lambda_{\max}$$

where $\lambda_{\max}$ is the largest positive eigenvalue of the transition matrix A_n of size $n = M^+ - M^- + 1$ as defined in equation (8).

III. GENERALIZATION TO K-ary CODES

We now wish to extend the result derived in the previous section to an arbitrary K-ary alphabet set, $\{\alpha_1, \cdots, \alpha_K\}$. We shall restrict ourselves to symmetric alphabets. Namely, if K is even, α_i takes on the values $-(K-1), -(K-3), \cdots, -1, +1, \cdots, (K-1)$; if K is odd, α_i takes on the values $-(K-1)/2, -(K-2)/2, \cdots, -1, 0, 1, \cdots, (K-1)/2$. The transition matrix of allowable states is then given by

$$A_n = \sum_{\alpha_i \geq 0} H_n^{\alpha_i} + \sum_{\alpha_i < 0} F_n^{-\alpha_i} \tag{23}$$

where the size of the matrices is

$$n = M^+ - M^- + 1, \tag{24}$$

and M^+ and M^- are the desired upper and lower bound on RDS. If $\alpha_i = 0$ is a member of the alphabet, we follow the usual convention that $H_n^0 = 1_n$. The matrices

$$H_n = \begin{bmatrix} 0 & 1 & & & 0 \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ 0 & & & & 0 \end{bmatrix} \tag{25}$$

and

$$F_n = \begin{bmatrix} 0 & & & & 0 \\ 1 & 0 & & & \\ & 1 & \ddots & & \\ & & \ddots & \ddots & \\ 0 & & & 1 & 0 \end{bmatrix} \tag{26}$$

are known as superdiagonal and subdiagonal matrices respectively. To see that A_n given in equation (23) is indeed the transition matrix, we observe that each symbol α_i will generate a sequence in any allowable state to a state α_i unit away. Each term in (23) represents the transition of states due to a particular alphabet. As an example, taking the quaternary alphabet set $\{-3, -1, +1, +3\}$, the transition matrix is

$$A_n = \begin{bmatrix} 0 & 1 & 0 & 1 & & & 0 \\ 1 & 0 & 1 & 0 & 1 & & \\ 0 & 1 & 0 & 1 & \cdot & \cdot & \\ 1 & 0 & 1 & \cdot & \cdot & \cdot & 1 \\ & 1 & \cdot & \cdot & \cdot & \cdot & 0 \\ & & \cdot & \cdot & \cdot & \cdot & 1 \\ 0 & & & 1 & 0 & 1 & 0 \end{bmatrix}.$$

With these preliminaries out of the way, we now state a general result on the limiting efficiency of dc-constrained codes:

Theorem 2: If the RDS of the coded K-ary signal stream is required to be within some bound M^+ and M^-, then the best possible coding efficiency is given by

$$\eta = \frac{\log_2 \lambda_{\max}}{\log_2 K}, \tag{27}$$

where $\lambda_{\max}$ is the largest eigenvalue of the transition matrix A_n defined by equations (23) and (24).

Before we embark on the proof of Theorem 2, we need to establish an important auxiliary result.

Let $N_M(L)$ denote again the number of allowable K-ary sequences, then the limiting efficiency η, corresponding to equation (5), is

$$\eta = \lim_{L \to \infty} \frac{\log_2 N_M(L)}{L \log_2 K}. \tag{28}$$

In the set of allowable sequences $S_M(L)$, we can define a subset $S_{M|\alpha_i}(L)$ by restricting the first symbol to be α_i. Similarly we define a subset $S_{M|\alpha_i}, -\alpha_i(L)$ by restricting the first two symbols to be α_i and $-\alpha_i$. Clearly

$$S(L) \supset S_M(L) \supset S_{M|L_i}(L) \supset S_{M|\alpha_i,-\alpha_i}(L), \tag{29}$$

and it follows that

$$\eta \geqq \eta_{\alpha_i} \geqq \eta_{\alpha_i,-\alpha_i} \tag{30}$$

where η_{α_i} and $\eta_{\alpha_i,-\alpha_i}$ are the limiting efficiencies given by equation (28) with the additional restriction on the leading elements.

Considering now all the sequences in $S_{M|\alpha_i,-\alpha_i}(L+2)$, it is not difficult to see that the number of sequences in $S_{M|\alpha_i,-\alpha_i}(L+2)$ is equal to that in $S_M(L)$. Hence the efficiency,

$$\begin{aligned}\eta_{\alpha_i,-\alpha_i} &= \lim_{L\to\infty} \frac{\log_2 N_M(L)}{\log_2 K^{L+2}}, \\ &= \lim_{L\to\infty} \frac{\log_2 N_M(L)}{(L+2)\log_2 K}, \\ &= \eta.\end{aligned} \tag{31}$$

Coupled with equation (30), we have shown that

$$\eta = \eta_{\alpha_i} = \eta_{\alpha_i,-\alpha_i}. \tag{32}$$

A little reflection should convince us that any finite pattern at the beginning of the sequences does not affect the limiting efficiency η. In other words, the limiting efficiency is independent of starting point —a fact we observed in the previous section after the detailed study of the transition matrix. This fact enables us to prove Theorem 2 without going through a tedious mathematical analysis.

Proof of Theorem 2: The matrix A_n defined in equation (23) is real and symmetric. It can be diagonalized by an orthogonal transformation, i.e.,

$$A_n = PD_AP',\ PP' = 1 \tag{33}$$

where D_A is a diagonal matrix of real elements $\lambda_1, \cdots, \lambda_n$.

Using any $\mathbf{u}_0$, a constant vector with non-negative elements, we can generate a sequence of vectors $\mathbf{u}_L$,

$$\mathbf{u}_L = A_n^L\mathbf{u}_0,\ \text{for } L = 1, 2, \cdots. \tag{34}$$

Since A_n is a matrix with non-negative elements, it is easy to see that all $\mathbf{u}_L$'s are vectors with non-negative elements. Write

$$P = [\mathbf{p}_1\mathbf{p}_2 \cdots \mathbf{p}_n] \tag{35}$$

where $\mathbf{p}_i$ is a column vector. From equation (34) we have, using $\mathbf{u}_0$ with only a 1 in jth position,

$$\begin{aligned}\mathbf{u}_L &= PD_A^LP'\mathbf{u}_0 \\ &= \sum_{i=1}^{n} \lambda_i^L p_{ji}\mathbf{p}_i\end{aligned} \tag{36}$$

where p_{ji} is the jth element in $\mathbf{p}_i$.

Let $\lambda_{\max}$ denote the absolute value of the largest eigenvalue of A_n and assume that, in general,*

$$\lambda_1 = \lambda_2 = \cdots = \lambda_r = \lambda_{\max}, \tag{37}$$

$$\lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_{r+s} = -\lambda_{\max}.$$

We can rewrite (36)

$$\mathbf{u}_L = \lambda_{\max}^L \left\{ \sum_{i=1}^{r} p_{ji}\mathbf{p}_i + (-1)^L \sum_{i=r+1}^{r+s} p_{ji}\mathbf{p}_i + \sum_{i=r+s+1}^{n} \left(\frac{\lambda_i}{\lambda_{\max}}\right)^L p_{ji}\mathbf{p}_i \right\}. \tag{38}$$

Denote by $\mathbf{z}$ the first two sums in equation (38),

$$\mathbf{z} = \sum_{i=1}^{r} p_{ji}\mathbf{p}_i + (-1)^L \sum_{i=r+1}^{r+s} p_{ji}\mathbf{p}_i. \tag{39}$$

$\mathbf{z}$ must be non-negative for any j and L. If not, then for some large enough L, $\mathbf{u}_L$ will have negative elements, which is a contradiction. Since $\mathbf{z}$ is a linear combination of $\mathbf{p}_1, \cdots, \mathbf{p}_{r+s}$, a set of linearly independent vectors, $\mathbf{z} = \mathbf{0}$ only if $p_{ji} = 0$, $i = 1, \cdots, r + s$. Furthermore, if $p_{ji} = 0$ for all $j = 1, \cdots, n$, then the transformation matrix P has a row of zeros, which is again a contradiction. Thus we conclude that, for some choice of $\mathbf{u}_0$, i.e., for some j,

$$\mathbf{u}_L = \lambda_{\max}^L \left\{ \mathbf{z} + \sum_{i=r+s+1}^{n} \left(\frac{\lambda_i}{\lambda_{\max}}\right)^L p_{ji}\mathbf{p}_i \right\} \tag{40}$$

with $\mathbf{z}$ non-negative independent of L. The total number of allowable sequences is

$$N_M(L) = \lambda_{\max}^L \left\{ \mathbf{1}'\mathbf{z} + \sum_{i=r+s+1}^{n} \left(\frac{\lambda_i}{\lambda_{\max}}\right)^L p_{ji}\mathbf{1}'\mathbf{p}_i \right\}. \tag{41}$$

Substituting $N_M(\mathrm{L})$ in equation (28), and passing to limit, we get equation (27). The proof is now complete.

IV. NUMERICAL RESULTS AND DISCUSSIONS

4.1 *Numerical Calculation*

Using the digital computer, the calculation of $\lambda_{\max}$ of any transition matrix A_n is not difficult except maybe for large n. In the following, we discuss several alternative approaches to evaluating $\lambda_{\max}$, and we present results for three important cases.

---

* It can be shown that $r = 1$ and $s = 0$ or 1. But the proof that follows does not require this fact.

(*i*) Find λ_{max} by direct diagonalization of the matrix A_n. There are computer programs developed for this purpose. This is done for the binary case and the quaternary case, and the limiting efficiency η is plotted (solid curve) as a function of allowable states n in Figs. 1 and 2 respectively.

(*ii*) In the binary case, the characteristic polynomial $\phi_n(\lambda)$ of A_n satisfies a simple recursive relation (56). Treating (56) as a difference equation of $\phi_n(\lambda)$'s, one can express $\phi_n(\lambda)$ in an alternate form:[4]

$$\phi_n(\lambda) = \frac{\sin [(m+1)\cos^{-1}(\lambda/2)]}{\sin [\cos^{-1}(\lambda/2)]}. \tag{42}$$

The roots of $\phi_n(\lambda)$ are, as easily seen in equation (42),

$$\lambda_k = 2\cos\frac{k\pi}{n+1}, \qquad k = 1, \cdots, n. \tag{43}$$

(*iii*) The ternary case can be reduced to the binary case by replac-

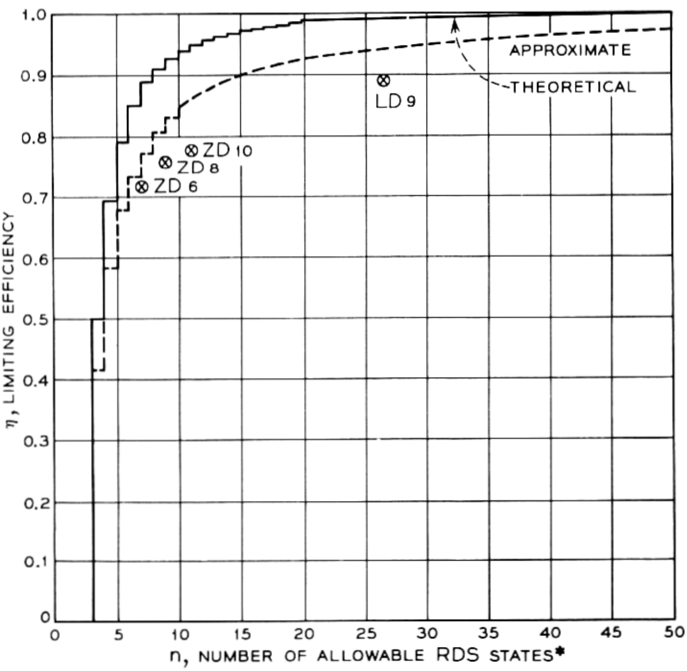

Fig. 1—Limiting efficiency vs allowable states binary alphabet (+1, −1). *$n = M^+ - M^- + 1$, where M^+ and M^- are the upper and lower bound of the RDS.

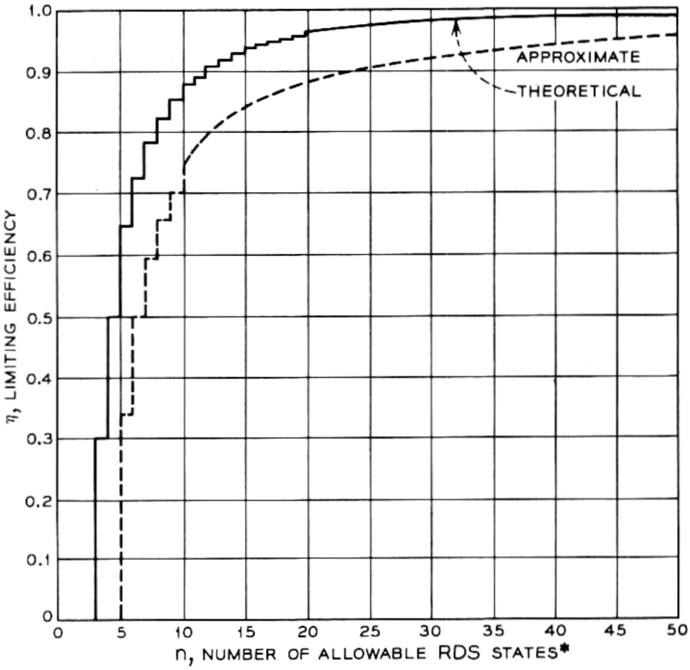

Fig. 2—Limiting efficiency vs allowable states quarternary alphabet (+3, +1, −1, −3). $*n = M^+ - M^- + 1$, where M^+ and M^- are the upper and lower bound of RDS.

ing λ in $\phi_n(\lambda)$ by $\lambda + 1$. Therefore, one gets $\lambda_{\max}$ of the ternary case by adding 1 to the corresponding (same n) $\lambda_{\max}$ of the binary case. The top curve in Fig. 3 is plotted in this way.

(*iv*) From the well-known formula[10]

$$\lambda_{\max} = \max_{\|\mathbf{x}\|=1} \mathbf{x}' A_n \mathbf{x} \tag{44}$$

where

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \tag{45}$$

and the norm of a vector $||x||$, is

$$||x|| = \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}} \tag{46}$$

we have

$$\lambda_{\max} = \max_{\|\mathbf{x}\|=1} \left\{ \sum_{\alpha_i \geq 0} \sum_{i=1}^{n-\alpha_i} x_i x_{i+\alpha_i} + \sum_{\alpha_i < 0} \sum_{i=1}^{n+\alpha_i} x_i x_{i-\alpha_i} \right\} \tag{47}$$

where the α_i's are members in the alphabet set. For example,

$$\lambda_{\max} = \max_{\|\mathbf{x}\|=1} 2 \sum_{i=1}^{n-1} x_i x_{i+1}, \tag{48}$$

in the binary case;

$$\lambda_{\max} = \max_{\|\mathbf{x}\|=1} \left\{ \sum_{i=1}^{n} x_i^2 + 2 \sum_{i=1}^{n} x_i x_{i+1} \right\} \tag{49}$$

in the ternary case; and

$$\lambda_{\max} = \max_{\|\mathbf{x}\|=1} 2 \left\{ \sum_{i=1}^{n-1} x_i x_{i+1} + \sum_{i=1}^{n-3} x_i x_{i+3} \right\} \tag{50}$$

in the quaternary case.

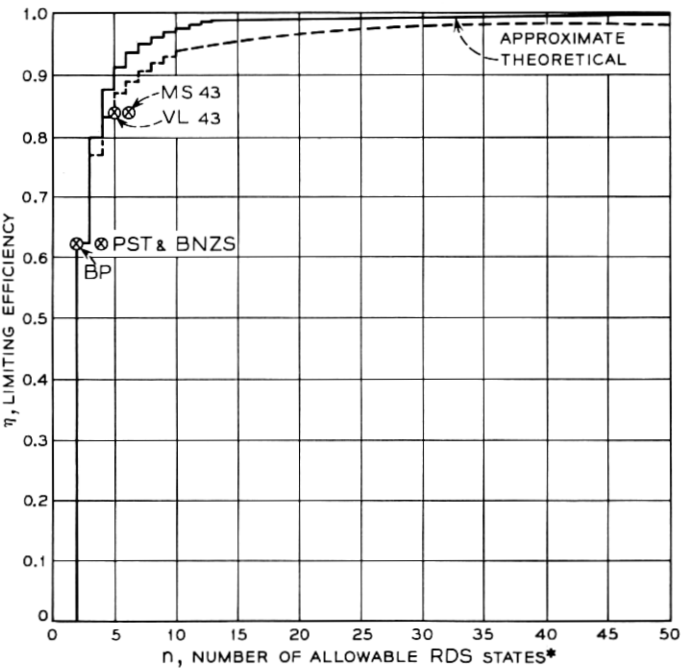

Fig. 3—Limiting efficiency vs allowable states ternary alphabet (+1, 0, −1). *$n = M^+ - M^- + 1$, where M^+ and M^- are the upper and lower bound of the RDS.

In this formulation, λ_{max} becomes the extreme value of a quadratic form with an equality constraint. There are a number of ways to effect a numerical solution.

4.2 *Approximation Formula*

To search for λ_{max} using equation (47) is not an easy alternative, but it leads to an estimation of λ_{max}. If we let $x_1 = x_2 = \cdots = x_n$, then, from equation (47), we have

$$\lambda_{max} \geqq \sum_{\alpha_i \geqq 0} \frac{n - \alpha_i}{n} + \sum_{\alpha_i < 0} \frac{n + \alpha_i}{n}. \tag{51}$$

Any other choice of the x_i's will lead to a different estimate of λ_{max} which may be better or worse than that of equation (51). We justify the present choice by noting the simplicity of equation (51). Using equation (51), we obtain an approximation formula for the limiting efficiency η,

$$\eta = \frac{\log_2 \left(\sum_{\alpha_i \geqq 0} \frac{n - \alpha_i}{n} + \sum_{\alpha_i < 0} \frac{n + \alpha_i}{n} \right)}{\log_2 K} \tag{52}$$

where α_i's are members of the K-ary alphabet set. The approximate η are also plotted in Figs. 1 to 3 (dashed curve).

As expected, the approximation is reasonably good for large n and it's for large n that we may have to rely on the approximation formula.

4.3 *Discussion*

(*i*) It is of some interest to see how the efficiencies of various codes with the dc-constrained property compare with the limiting curves. We have located the following known codes:

ZDN (Zero-Disparity Binary Code of Block Length N)[5] and
LDN (Low-Disparity Binary Code of Block Length N)[6] in Fig. 1; and
BP (Bipolar Code)[1], $n = 2$, $\eta = 0.63$[†],
PST (Paired Selected Ternary Code)[7], $n = 4$, $\eta = 0.63$,
BNZS (Bipolar with N Zero Substitution Codes)[8], $n = 4$, $\eta = 0.63$,
VL43 (Variable Length Ternary Code)[9], $n = 5$, $\eta = .84$, and
MS43 (Fixed Length Ternary Code)[9], $n = 6$, $\eta = 0.84$, in Fig. 3.

* $N = 6$. The allowable states for BP are -1 and 0 or 0 and $+1$ depending upon whether the first pulse transmitted be -1 or $+1$. A similar situation exists for PST, BNZS.

The BP and BNZ* codes are used in T-1 and T-2 systems[1,2] respectively and PST is used in the experimental T-4 system.[3] In comparison with their limiting efficiencies, one gets the impression that, barring the fact that these codes have other properties in addition to the dc-constraint, there is some room for improving the coding efficiency. VL43, and MS43 are examples along this direction.

It should be pointed out* that the real engineering problem is to control baseline wander. A true comparison of different coding schemes should therefore be done on this basis. The relation between RDS and baseline wander is an elusive one. It depends upon the detailed structure of the code in question and the channel to which the signal is applied. In terms of RDS, it depends not only on the bounds and the distribution of the allowable RDS but also on its dynamics. By dynamics we mean the "speed" of moving from one state to another, the dynamic behavior is of importance because the channel has "failing memory", so to speak. For example, it can be shown[11] under certain conditions that a quick jump from one extreme state to the other results in a larger amount of baseline wander than would occur from staying in an extreme state for a long time.

(*ii*) As a general observation, the limiting curves saturate rapidly. This implies that, to make a high efficiency code possible, a physical system should be designed to operate beyond the fast rising portion of the curves. It would be reasonable then to expect that a simple ternary block code could be found with 90 percent efficiency or better for, say, $n = 15$.

(*iii*) An interesting question which arises naturally in connection with the limiting efficiency is its realizability. If one accepts infinite delay, the answer is affirmative. If one thinks in terms of block codes of finite length, then the limiting efficiency cannot be realized.

V. CONCLUSION

We have shown that, for a dc-constrained code, the limiting efficiency is related to the number of allowable RDS states in a very simple way. The result is effective in the sense that it lends itself easily to numerical evaluation.

The underlying mathematical fact in our proof is the property of non-negative matrices and vectors. Using the theorem of Frobenius on non-negative matrices,[10] our result can be proved in a few steps.

---

* The discussion here is heuristic in nature. A thorough treatment of the subject is beyond the scope of this paper.

We retain our approach for the reasons that it only requires elementary knowledge of matrix theory and it gives more insight to the problem.

The technique developed in this paper can be used to investigate bounds for some other classes of codes, say timing codes. This will be done elsewhere.

VI. ACKNOWLEDGMENTS

The author would like to express his thanks to V. I. Johannes and P. E. Rubin for their challenging comments and useful suggestions and to R. W. Kolor for his effort and interest in calculating the numerical results.

APPENDIX

The class of matrices which we want to investigate has the following general form*

$$A_n = \begin{bmatrix} 0 & 1 & & & 0 \\ 1 & 0 & 1 & & \\ & 1 & \cdot & \cdot & \\ & & \cdot & \cdot & 1 \\ 0 & & & 1 & 0 \end{bmatrix}. \tag{53}$$

Each matrix A_n is a square matrix of size n and it has ones in the super- and sub-diagonal and zeros elsewhere.

Let $\phi_n(\lambda)$ denote the characteristic polynomial of A_n. By definition,

$$\phi_n(\lambda) \triangleq \det [\lambda 1_n - A_n]$$

$$= \det \begin{bmatrix} \lambda & -1 & & & 0 \\ -1 & \lambda & -1 & & \\ & -1 & \cdot & \cdot & \\ & & \cdot & \cdot & -1 \\ 0 & & & -1 & \lambda \end{bmatrix}. \tag{54}$$

The first few polynomials $\phi_n(\lambda)$ are

* A special class of Jacobi matrix.[12]

$$\begin{aligned}\phi_0(\lambda) &= 1,\\ \phi_1(\lambda) &= \lambda,\\ \phi_2(\lambda) &= \lambda^2 - 1,\\ \phi_3(\lambda) &= \lambda^3 - 2\lambda,\\ \phi_4(\lambda) &= \lambda^4 - 3\lambda^2 + 1,\end{aligned} \tag{55}$$

where $\phi_0(\lambda)$ is defined to be 1. The polynomials $\phi_n(x)$ have some interesting properties. We state them as lemmas.

Lemma 1: $\phi_n(x)$ satisfies the following recursive relations:
For $n \geqq 2$

$$\phi_n(\lambda) = \lambda\phi_{n-1}(\lambda) - \phi_{n-2}(\lambda), \tag{56}$$

and for $n > m$,

$$\phi_n(\lambda) = \phi_m(\lambda)\phi_{n-m}(\lambda) - \phi_{n-1}(\lambda)\phi_{n-m-1}(\lambda). \tag{57}$$

Proof: To prove equation (56), we expand the determinant in (54) with respect to the first column. To prove (57), we expand the determinant with respect to the first m columns and observe that the only two nonzero products of minors are of size m and $n - m$.

Lemma 2: ϕ_n (λ) is an even (odd) polynomial if n is even (odd), i.e.,

$$\phi_n(\lambda) = (-1)^n\phi_n(-\lambda). \tag{58}$$

Proof: Assume that (58) is true for $\phi_{n-1}(\lambda)$ and $\phi_{n-2}(\lambda)$, then by (56),

$$\begin{aligned}\phi_n(-\lambda) &= -\lambda\phi_{n-1}(-\lambda) - \phi_{n-2}(-\lambda),\\ &= (-1)^n\lambda\phi_{n-1}(\lambda) - (-1)^{n-2}\phi_{n-2}(\lambda),\\ &= (-1)^n[\lambda\phi_{n-1}(\lambda) - \phi_{n-2}(\lambda)],\\ &= (-1)\phi_n(\lambda).\end{aligned}$$

Since (58) is true for $\phi_0(\lambda)$ and $\phi_1(\lambda)$ by inspection of (55), (58) is true for any n by induction.

Lemma 3: If λ_0 is a root of $\phi_n(\lambda)$ and $\lambda_0 \neq 0$, then —λ_0 is also a root of $\phi_n(\lambda)$.

Proof: Follows directly from the previous lemma.

By definition, the roots of $\phi_n(\lambda)$ are the eigenvalues of matrix A_n. Since A_n is real symmetric, it follows that all the roots of $\phi_n(\lambda)$ are real. Let $\lambda_{\max}^{(n)}$ denote the root of $\phi_n(\lambda)$ with largest absolute value. In view of the result in Lemma 3, $\lambda_{\max}^{(n)}$ can always be taken to be positive.

Lemma 4: For any finite n, $\lambda_{\max}^{(n)}$ of $\phi_n(\lambda)$ has the following ordering:

$$\lambda_{\max}^{(1)} < \lambda_{\max}^{(2)} < \cdots < \lambda_{\max}^{(n)} < 2. \tag{59}$$

Proof: We will prove this lemma again by induction. Clearly, $\phi_n(\lambda) \to \infty$ as $\lambda \to \infty$ for any $n = 1, 2, \cdots$. It follows that $\phi_n(\lambda) > 0$ for $\lambda > \lambda_{\max}^{(n)}$ the largest positive root of $\phi_n(\lambda)$. Now, assume that (59) holds for $\lambda_{\max}^{(n-1)}$ and $\lambda_{\max}^{(n-2)}$, then, from (56),

$$\begin{aligned} \phi_n(\lambda_{\max}^{(n-1)}) &= \lambda_{\max}^{(n-1)}\phi_{n-1}(\lambda_{\max}^{(n-1)}) - \phi_{n-2}(\lambda_{\max}^{(n-1)}) \\ &= -\phi_{n-2}(\lambda_{\max}^{(n-1)}) < 0. \end{aligned} \tag{60}$$

Hence $\phi_n(\lambda)$ changes sign at least once between $\lambda_{\max}^{(n-1)}$ and ∞. This implies that

$$\lambda_{\max}^{(n)} > \lambda_{\max}^{(n-1)}. \tag{61}$$

Since (59) holds for $n = 1$ and 2, it holds for any n.

To show that $\lambda_{\max}^{(n)} < 2$ for any finite n, we make the observation that for $\lambda > 2$, the matrix $\lambda 1_n - A_n$ is a dominant matrix,[13] and therefore nonsingular. For $\lambda = 2$,

$$\begin{aligned} \phi_n(2) &= n\phi_1(2) - (n-1)\phi_0(2), \\ &= n + 1 \neq 0, \end{aligned} \tag{62}$$

by repeated use of the recursive relation (56).

Lemma 5: For some number λ_0, if $\phi_n(\lambda_0) = 0$, then $\phi_{n-1}(\lambda_0) \neq 0$.

Proof: Assume the contrary, i.e.,

$$\phi_n(\lambda_0) = \phi_{n-1}(\lambda_0) = 0. \tag{63}$$

Then from the recursive formula (56),

$$\phi_{n-2}(\lambda_0) = 0. \tag{64}$$

Repeating the same argument, we conclude

$$\phi_n(\lambda_0) = \cdots = \phi_1(\lambda_0) = \phi_0(\lambda_0) = 0 \tag{65}$$

which is impossible.

Lemma 6: $\phi_n(\lambda)$ has only simple roots.

Proof: Let λ_0 be a root of $\phi_n(\lambda)$, then the matrix $[\lambda_0 1_n - A_n]$ is singular. On the other hand, from the previous lemma, $[\lambda_0 1_{n-1} - A_{n-1}]$ is nonsingular. Hence the null space of the matrix $[\lambda_0 1_n - A_n]$

is one-dimensional. From the fact that A_n is diagonalizable, we conclude that λ_0 must be a simple root of $\phi_n(\lambda)$.

Lemma 7: Write

$$A_n = PD_A P' \tag{66}$$

where

$$D_A = \text{diag}\,[\lambda_1, \cdots, \lambda_n] \tag{67}$$

and

$$PP' = 1; \tag{68}$$

then in general, P can be expressed in terms of $\phi(\lambda)$'s

$$P = \begin{bmatrix} \frac{\phi_0(\lambda_1)}{\phi(\lambda_1)} & \frac{\phi_0(\lambda_2)}{\phi(\lambda_2)} & \cdots & \frac{\phi_0(\lambda_n)}{\phi(\lambda_n)} \\ \vdots & \vdots & & \vdots \\ \frac{\phi_{n-1}(\lambda_1)}{\phi(\lambda_1)} & \frac{\phi_{n-1}(\lambda_2)}{\phi(\lambda_2)} & & \frac{\phi_{n-1}(\lambda_n)}{\phi(\lambda_n)} \end{bmatrix} \tag{69}$$

where

$$\phi(\lambda) = \left[\sum_{i=0}^{n-1} \phi_i^2(\lambda)\right]^{\frac{1}{2}} \tag{70}$$

is a normalization constant.

Proof: Let $\mathbf{x}$ be an eigenvector corresponding to an eigenvalue λ of A_n,

$$A_n\mathbf{x} = \lambda\mathbf{x}. \tag{71}$$

Write

$$\mathbf{x} = [x_1, \cdots, x_n]'$$

and expand (71), we have

$$\begin{aligned} x_2 &= \lambda x_1, \\ x_3 &= \lambda x_2 - x_1, \\ x_4 &= \lambda x_3 - x_2 \\ &\vdots \\ x_n &= \lambda x_{n-1} - x_{n-2}, \end{aligned} \tag{72}$$

and

$$x_{n-1} = \lambda x_n .$$

Delete the last equation in (72) and then compare with equations (55) and (56). We can make the following identification:

$$\begin{aligned} x_1 &= \phi_0(\lambda), \\ x_2 &= \phi_1(\lambda) \\ &\vdots \\ x_n &= \phi_{n-1}(\lambda). \end{aligned} \tag{73}$$

To normalize $\mathbf{x}$, we divide (73) by the inner product of $\mathbf{x}$. Denoting the inner product by $\phi(\lambda)$,

$$\phi(\lambda) = \left[\sum_{i=0}^{n-1} \phi_i^2(\lambda) \right]^{\frac{1}{2}}, \tag{74}$$

the normalized eigenvector corresponding to the eigenvalue λ is given by

$$\left[\frac{\phi_0(\lambda)}{\phi(\lambda)} \cdots \frac{\phi_{n-1}(\lambda)}{\phi(\lambda)} \right]'. \tag{75}$$

It is well known that eigenvectors corresponding to different eigenvalues are orthogonal. Therefore, for P as defined in (69), $P'P = 1_n$.

Since each column of P is an eigenvector of A_n, it follows that

$$A_n P = P D_A$$

and equation (66) is immediate.

REFERENCES

1. Aaron, M. R., "PCM Transmission in the Exchange Plant," B.S.T.J., *41*, No. 1 (January 1962), pp. 99–141.
2. Davis, J. H., "T2: A 6.3 Mb/s Digital Repeatered Line," IEEE International Conference on Communications, Boulder, Colorado, June 1969.
3. Dorros, I., Sipress, J. M., and Waldhauer, F. D., "An Experimental 224 Mb/s Digital Repeatered Line," B.S.T.J., *45*, No. 7 (September 1966), pp. 993–1043.
4. McCullough, R. H., "Ternary Codes for Regenerative Digital Transmission Systems," Ph.D. Dissertation, Polytechnic Institute of Brooklyn, June 1967.
5. Cattermole, K. W., "Low-Disparity Codes and Coding for PCM," Proc. IEEE Conf. Transmission Aspects of Commun. Networks, London, February 1964, pp. 179–182.
6. Carter, R. O., "Low-Disparity Binary Coding System," Elec. Letters, *1*, No. 3 (May 1965), pp. 67–68.
7. Sipress, J. M., "A New Class of Selected Ternary Pulse Transmission Plans for Digital Transmission Lines," IEEE Trans. on Commun. Technology, *13*, No. 3 (September 1965), pp. 366–372.

8. Johannes, V. I., Kaim, A. G., and Walzman, T., "Bipolar Pulse Transmission with Zero Extraction," Trans. IEEE Commun. Technology, *17*, No. 2 (April 1969), pp. 303–310.
9. Franaszek, P. A., "Sequence-State Coding for Digital Transmission," B.S.T.J., *47,* No. 1 (January 1968), pp. 143–157.
10. Gantmacher, F. R., *The Theory of Matrices,* Vol. II, New York: Chelsea Publishing Co., 1959, pp. 50–60.
11. Johannes, V. I., unpublished work.
12. Marcus, M., and Ming, H., *A Survey of Matrix Theory and Matrix Inequalities,* Boston: Allyn and Bacon, 1964, pp. 166–167.
13. Bellman, R., *Introduction to Matrix Analysis,* New York: McGraw-Hill, 1960, p. 295.