FACULTÉ DES SCIENCES ET DE GÉNIE
FACULTY OF SCIENCE AND ENGINEERING

ON THE USE OF NEGACYCLIC CODES FOR SOURCE-ENCODING

S.G.S. Shiva, J.P. Dion
Dept. of Electrical Engineering
University of Ottawa,
Ottawa, Ontario. K1N 6N5.

UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

ON THE USE OF NEGACYCLIC CODES FOR SOURCE-ENCODING

S.G.S. Shiva, J.P. Dion
Dept. of Electrical Engineering
University of Ottawa,
Ottawa, Ontario. K1N 6N5.

Final report submitted to the Department of Supply and Services, for the Department of Communications (CRC) under Contract Serial No: OSU4-0162.

February 1977

PREFACE

This final report is in two parts. The first part entitled "A Joint Source and Channel Encoding Scheme" describes a technique, using negacyclic codes, which not only reduces the redundancy of the source, but also has error-detection capability. The second part entitled "Some Results on Negacyclic Codes" gives results which are parallel to certain known results for cyclic codes and which are useful for data-compression .

PART 1 : A JOINT SOURCE AND CHANNEL ENCODING SCHEME

1. INTRODUCTION

In the conventional communications system model, redundancy extant in the source output is reduced by source encoding, the encoded data being further encoded in order to increase channel noise immunity, source-encoded data being more vulnerable to channel noise than the original source data.

From the noiseless coding theorem (Shannon 1948, Ash 1965), it is known that for any source coding scheme, the average length of an encoded sequence can be no less than the source entropy divided by the capacity of the alphabet. Source coding schemes which approached this lower bound were suggested by Shannon (1948) and Fano (1961) and an optimal encoding procedure developed by Huffman (1952).

With the development of rate distortion theory (Berger 1971), it became clear that, while entropy describes the rate at which the source produces distortionless ("noiseless") information, the rate distortion function $R(\mathbb{D})$ describes the rate of information production by the source, subject to the condition that all distortion be constrained by a predefined value, D. Clearly, as D increases, R(D) decreases.

In this context, Shannon (1959) first suggested the use of group codes for source encoding, and Goblick (1962) proved the existence of group codes that perform arbitrarily close to the R(D) curve. Jelinek (1969) established a similar existence theorem for tree codes. Subsequently, many algorithms for sequential source encoding were developed (Jelinek and Anderson 1971, Anderson and Jelinek 1973, Anderson 1974).

Another approach has been to encode the source run lengths (Molinder 1974, Meyr, Rosdolsky and Huang 1974); this has been attempted in many different ways. Other methods include enumerative source encoding (Cover 1973), time-encoding schemes (Lynch 1974) and sliding-block source coding, to name only a few. A number of source encoding techniques are

mentioned by Wilkins and Wintz (1971), in Section I of their paper.

Hellman (1975) has suggested jointly encoding both the source and the channel using a convolutional encoder and a sequential decoder. In this paper, another type of joint source and channel encoding is presented, based on a variation of syndrome encoding. Using negacyclic codes (Berlekamp 1968), a type of block code, it results in data compression as well as an error detection capability. While it is demonstrated that this can be achieved without distortion, assuming that the source has the proper probability distribution, it is also shown that if distortion is allowed, both data compression and error detection can be achieved for other source probability distributions. Also, expressions for the average length of the encoded words, as well as the average distortion per digit, for the cases where distortion is allowed, are derived.

Before developing the main ideas, an introduction to negacyclic codes is in order; this is accomplished in the next section.

2. NEGACYCLIC CODES

Here we give a brief treatment of the material necessary for the development of the coding technique under consideration.

Let us consider a source the output of which is over the set $S = \{0, \pm 1, \pm 2, \ldots \pm t = \frac{p-1}{2}\}$ where p is prime and > 2. Under modulo -p operations, S is the same as GF(p).

With reference to a sequence $a_o a_1 a_2 \ldots a_{n-1}$ over S, the Lee weight of a_i is given by $|a_i|$. The Lee weight of the sequence is given by $\sum_{i=o}^{n-1} |a_i|$.

A negacyclic code V over S with block length $n = \frac{p^r-1}{2}$ is the set of all multiples of a generator polynomial which divides $1 + X^n$ over S. Here n is a nonmultiple of p and r is a positive integer. The code can be designed to correct any one error-pattern of Lee weight t or less. From this point on, it will be understood that Lee weight is meant whenever

weight is mentioned. Hereafter we mean by an error-pattern, one of weight t.

Corresponding to each distinct error-pattern there is a distinct (rt) - tuple, over S, which is called an error-syndrome. We also note that each error-pattern can be treated as an n-tuple, over S, with a Lee weight of t or less.

The ideas of the last paragraph form the basis of the encoding scheme to be discussed next.

3. ENCODING SCHEME WITHOUT DISTORTION

With reference to V, let us partition the source output into blocks of n digits. Let B be such a block with Lee weight w.

If $w = 0$, then B can be replaced with the (rt) - tuple of all zeros.

If $w \neq 0$, then let us write

$$w = mt + \beta, \quad 0 < \beta \leq t.$$

Now B can be treated as the sum of m+1 n-tuples $B_1, B_2, \ldots B_{m+1}$ such that the (m+1) -by-n matrix $\mathcal{B}$, in which the rows are $B_1, B_2, \ldots B_{m+1}$, satisfies the following properties : (i) every row of $\mathcal{B}$ has a Lee weight of t except the (m+1)st row which may have a Lee weight $< t$. (ii). Every column of $\mathcal{B}$ has at most 2 nonzero entries. (iii) The Lee weight of every column of $\mathcal{B}$ is t or less. (iv) If a column has two nonzero entries, then they are adjacent and are of the same sign.

For example let us consider the case of n = 12 and $S = \{0, \pm 1, \pm 2\}$. Suppose B = 0 1 -2 2 1 1 0 0 -2 1 1 1 .

Then we have

$$\mathcal{B} = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

Let σ_i represent the syndrome of B_i. Then the block B can be

replaced with the encoded sequence $\sigma = \sigma_1 \sigma_2 \cdots \sigma_{m+1}$ which, we note, is $(m+1)\,rt$ digits long.

We transmit σ over the channel. At the other end we receive, say, $\sigma_1' \sigma_2' \cdots \sigma_{m+1}'$. Corresponding to each syndrome σ_i, we get the error-pattern B_i' and form the matrix $\mathcal{B}'$ in which the ith row is B_i'.

If $\mathcal{B}'$ satisfies all of the properties satisfied by $\mathcal{B}$, then we add the rows of $\mathcal{B}'$ to obtain B' which, we decide, is the sequence put out by the source. If not, either we ask for a retransmission or just abandon the block.

3.1 AVERAGE LENGTH OF THE ENCODED SEQUENCE, DISTORTIONLESS CASE:

Let P_w represent the probability that B has Lee weight w. Then, the average length $\bar{L}$ of σ is given by

$$\bar{L} = Q_1 \cdot rt + Q_2 \cdot 2rt + Q_3 \cdot 3rt + \ldots + Q_n \cdot nrt$$

$$= rt\,(Q_1 + 2Q_2 + 3Q_3 + \ldots + nQ_n),$$

where

$$\left.\begin{aligned} Q_1 &= P_0 + P_1 + \ldots + P_t, \\ Q_2 &= P_{t+1} + P_{t+2} + \ldots + P_{2t}, \\ Q_3 &= P_{2t+1} + P_{2t+2} + \ldots + P_{3t}, \\ &\cdots\cdots \\ Q_n &= P_{(n-1)t+1} + \ldots + P_{nt}. \end{aligned}\right\} \qquad (3.1.1)$$

We note that $Q_1 = \text{Prob}\,(0 \le w \le t)$

$Q_j = \text{Prob}\,((j-1)t+1 \le w \le jt),\ j \ge 2$

Noting that

$$n = \frac{p^r - 1}{2} = \frac{(p-1)}{2}(1 + p + p^2 + \ldots + p^{r-1})$$

$$= t\,(1 + p + p^2 + \ldots + p^{r-1}),$$

we get

$$\frac{\bar{L}}{n} = \frac{r(Q_1 + 2Q_2 + 3Q_3 + \ldots + nQ_n)}{1 + p + p^2 + \ldots + p^{r-1}}.$$

Given the distribution $Q = \{Q_1, Q_2, \ldots Q_n\}$, the ratio $\frac{\bar{L}}{n}$ can be computed.

For example if $Q_i = Q_1 \lambda^{i-1}$, we get

$$\frac{\bar{L}}{n} = \frac{r}{1+p+p^2+\ldots+p^{r-1}} \cdot \frac{Q_1}{1-\lambda} \left(\frac{1-\lambda^n}{1-\lambda} - n\lambda^n \right)$$

Since

$$Q_1+Q_2+\ldots+Q_n = Q_1 \cdot \frac{1-\lambda^n}{1-\lambda} = 1,$$

we further get

$$\frac{\bar{L}}{n} = \frac{r}{1+p+p^2+\ldots+p^{r-1}} \cdot \frac{1}{1-\lambda^n} \left(\frac{1-\lambda^n}{1-\lambda} - n\lambda^n \right)$$

or

$$\frac{\bar{L}}{n} = \frac{rt}{n} \left(\frac{1}{1-\lambda} - \frac{n\lambda^n}{1-\lambda^n} \right) = \frac{rt}{n(1-\lambda)} + \frac{\lambda^n rt}{\lambda^n - 1}$$

from which we get

$$\frac{\bar{L}}{n} \simeq \frac{rt}{n(1-\lambda)}$$

Here we note that $\frac{\bar{L}}{n}$ is very near the right hand side quantity, since λ^n is usually $<< 1$, so that we can write

$$\frac{\bar{L}}{n} = \frac{rt}{n(1-\lambda)} \qquad (3.1.2)$$

If the distribution Q is appropriate then $\frac{\bar{L}}{n}$ will be ≤ 1.

For instance, in the case of the distribution $Q_i = Q_1 \lambda^{i-1}$, if $\lambda \leq 1 - \frac{rt}{n}$, then $\frac{\bar{L}}{n} \leq 1$.

Figure 1 is a graph of equation (3.1.2) for the (n = 12, k = 8) 2-error-correcting negacyclic code over S, for the source probability distribution defined by equation (3.1.1). For this code, r = t = 2, and hence from the previous discussion, $\frac{\bar{L}}{n}$ should be unity or less for $\lambda \leq 0.666$; this is borne out in the figure.

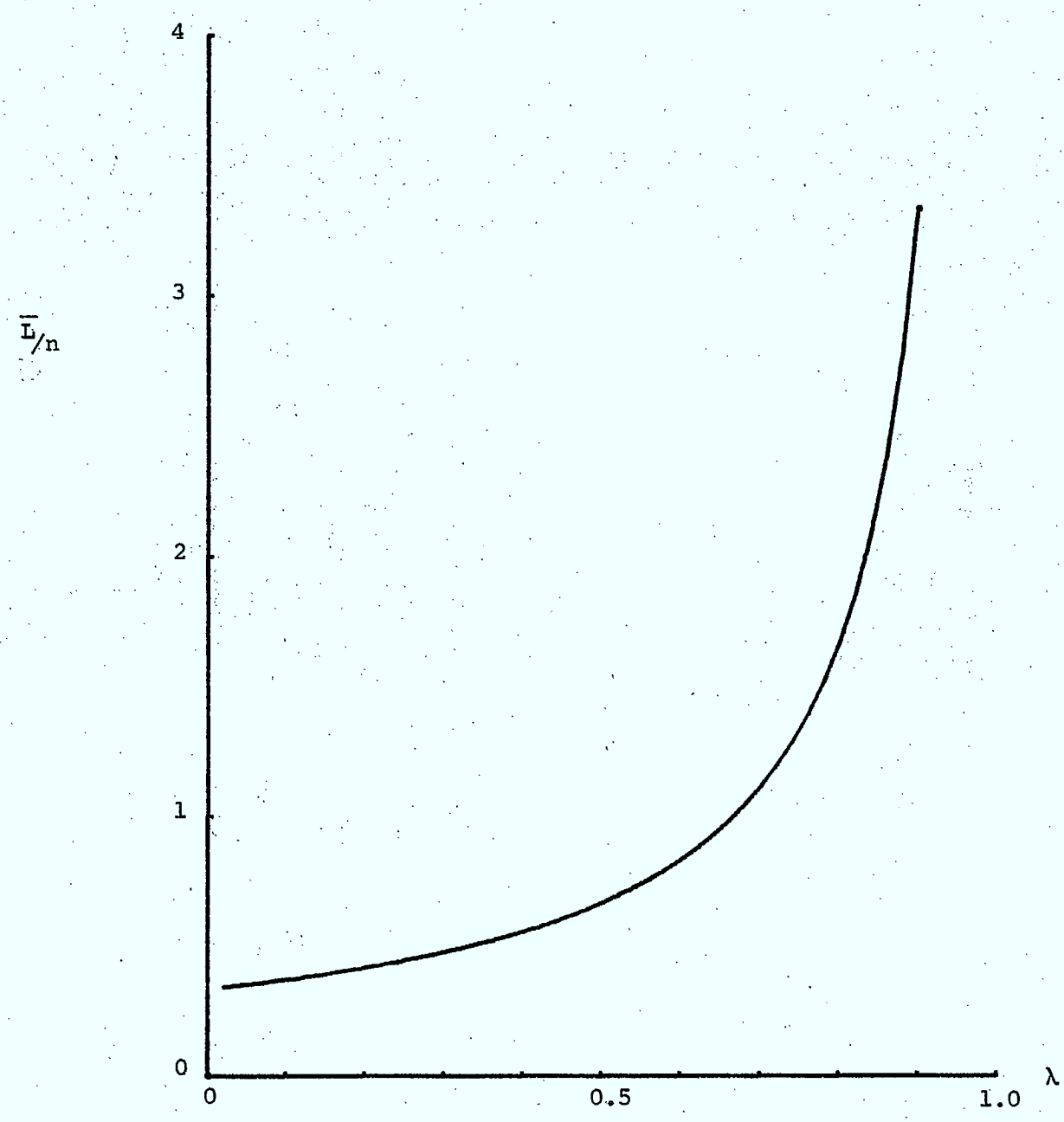

Fig. 1. Graph of the normalized average length of an encoded sequence for the distortionless case.

Henceforth, for purposes of comparison, all examples and graphs will make use of the aforementioned negacyclic code.

The case of $\frac{\bar{L}}{n} = 1$ is of particular interest. As already seen, every time $\mathcal{B}'$ does not satisfy all of those properties attributed to $\mathcal{B}$, we have detected an error situation. Thus the coding scheme has an error-detecting capability. In view of this, the condition $\frac{\bar{L}}{n} = 1$ implies that there has been some data compression, since otherwise $\bar{L}$ can not possibly be equal to n.

Thus we have a coding scheme which, for appropriate distributions Q, has both error-detection and data-compression capabilities. It is worth repeating that the scheme does not introduce distortion.

On the other hand, if the distribution Q is such that $\bar{L}$ works out to be $> n$, we can introduce distortion to lower the value of $\bar{L}$ down to n.

This last aspect will be discussed next.

4. ENCODING WITH DISTORTION

4.1 SCHEMES FOR THE INTRODUCTION OF DISTORTION

Distortion can be introduced in many ways.

(a) For instance, the source output can be clipped so that the resulting sequence is over the set $S_o = \{0, \pm 1, \pm 2, \ldots \pm t_o\}$ where $t_o < t$. If $t_o = t-1$, the average distortion per digit can be easily proved to be π_t, where π_t is the probability that the source puts out a $+t$ or $-t$. This scheme does not appreciably lower $\bar{L}$ unless there are frequently t or more $\pm$ t's per block.

For example, for $n = 12$, $S = \{0, \pm 1, \pm 2\}$, $(t = 2)$, if B = 0 1 -2 2 1 1 0 0 -2 1 1 1 then clipping to $t_o = t-1 = 1$ produces the sequence B = 0 1 -1 1 1 1 0 0 -1 1 1 1, with a distortion of 3, distributed over the sequence.

(b) From the point of view of reducing $\bar{L}$, an improvement over the previous scheme (a) would be to "clip" the weight of the block put out by the source to some value, say, W. For example let us consider the case of n = 12 and t = 2. Maximally the block weight is 24. Suppose we set W = 8. Suppose B = 0 1 -2 2 1 1 0 0 -2 1 1 1 ; it has a weight of 12. Clipping this B to a weight of 8 produces B = 0 1 -2 2 1 1 0 0 -1 0 0 0, the sequence which would be encoded and transmitted. The distortion in this case would be 4, concentrated in the last 4 positions.

(c) From the point of view of distributing the distortion over the block, it is desirable to combine (a) and (b). For the example considered in (a) and (b), this would transform B = 0 1 -2 2 1 1 0 0 -2 1 1 1 into B = 0 1 -1 1 1 1 0 0 -1 1 1 0, for t_o = 1 and W = 8. The distortion in this case would be 4, distributed as follows : 1 unit of distortion in the third, fourth, ninth and twelfth digits of B.

We may note that, for the examples considered, both (b) and (c) give the same total distortion of 4, while (a) gives a distortion of 3.

4.2 AVERAGE LENGTH $\bar{L}_a$ OF THE ENCODED SEQUENCE:

The average length $\bar{L}_a$ of σ in the case of (a) is given by

$$\bar{L}_a = \left\langle \frac{n\left[\pi_1 + 2\pi_2 + 3\pi_3 + \ldots + (t_o - 1)\,\pi_{t_o - 1} + t_o(\pi_{t_o+1} + \ldots + \pi_t)\right]}{t} \right\rangle rt$$

Where $\langle x \rangle$ means the "higher" integer if x is not an integer and $\langle x \rangle = x$ if x is an integer.

4.3 AVERAGE LENGTH $\bar{L}_c$ OF THE ENCODED SEQUENCE :

The average length $\bar{L}_c$ of σ in the case of (b) and (c) is given by

$$\bar{L}_c = rt\,(Q_1 + 2Q_2 + 3Q_3 + \ldots + \nu Q_\nu) ,$$

Where ν is given by

$$\nu = \left\langle \frac{W}{t} \right\rangle$$

For the distribution $Q_i = Q_1 \lambda^{i-1}$, we get

$$\frac{\bar{L}_c}{n} = \frac{r}{1+p+p^2+\ldots+p^{r-1}} \quad \frac{1}{1-\lambda^n} \left(\frac{1-\lambda^\nu}{1-\lambda} - \nu \lambda^\nu \right) \qquad (4.3.1)$$

Figure 2 is a graph of equation (4.3.1) for the source probability distribution defined by equation (3.1.1), with W as parameter.

It can be readily observed that as W decreases, thus increasing the potential distortion, $\frac{L_c}{n}$ decreases.

The decrease in the average length is given by

$$\bar{L} - \bar{L}_c = rt[Q_{\nu+1}(\nu+1) + Q_{\nu+2}(\nu+2) + \ldots + Q_n n] \qquad (4.3.2)$$

Figure 3 is a graph of equation (4.3.2) (normalized by division by n) with parameter W, for the source probability distribution defined by equation (3.1.1).

As expected, as W decreases, the difference in lengths increases.

4.4 TOTAL AVERAGE DISTORTION FOR CASES (a), (b) AND (c) :

Let π_i represent the probability that the source puts out $+i$ or $-i$, where i belongs to S.

Then in scheme (a) the total average distortion per block is given by $n\pi_t(t-t_o) + n\pi_{t-1}(t-t_o-1) + n\pi_{t-2}(t-t_o-2) + \ldots + n\pi_{t_o+1}(1)$ so that the average distortion D_a per digit is given by

$$D_a = \pi_t(t-t_o) + \pi_{t-1}(t-1-t_o) + \pi_{t-2}(t-2-t_o) + \ldots + \pi_{t_o+1}(t_o+1-t_o)$$

$$= \pi_t(\delta) + \pi_{t-1}(\delta-1) + \pi_{t-2}(\delta-2) + \ldots + \pi_{t_o+1}(1)$$

or

$$D_a = \pi_{t_o+1}(1) + \pi_{t_o+2}(2) + \pi_{t_o+3}(3) + \ldots + \pi_t(\delta),$$

where $\delta = t-t_o$.

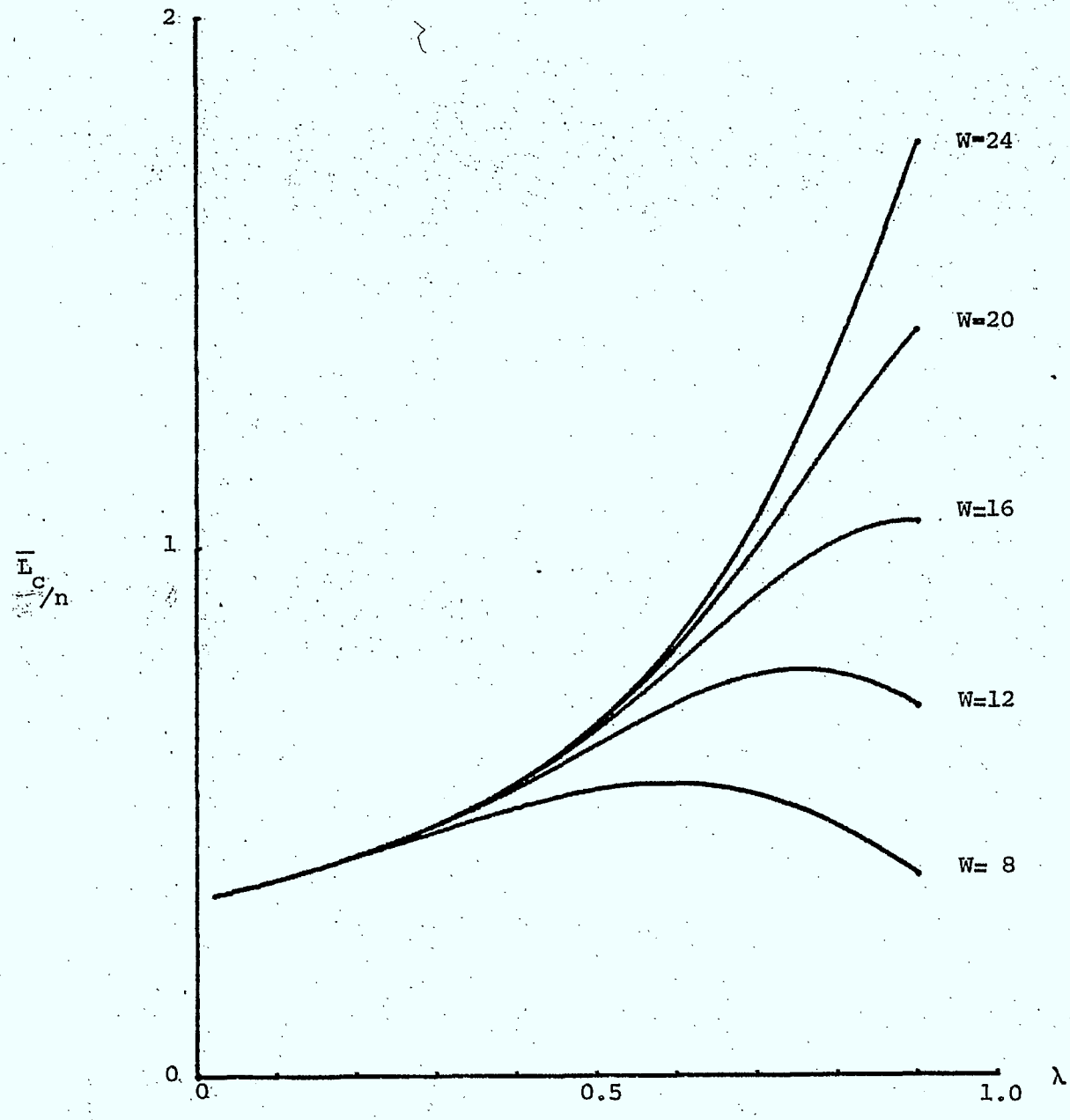

Fig. 2. Graph of the normalized average length of an encoded sequence for the case with distortion, with W as parameter.

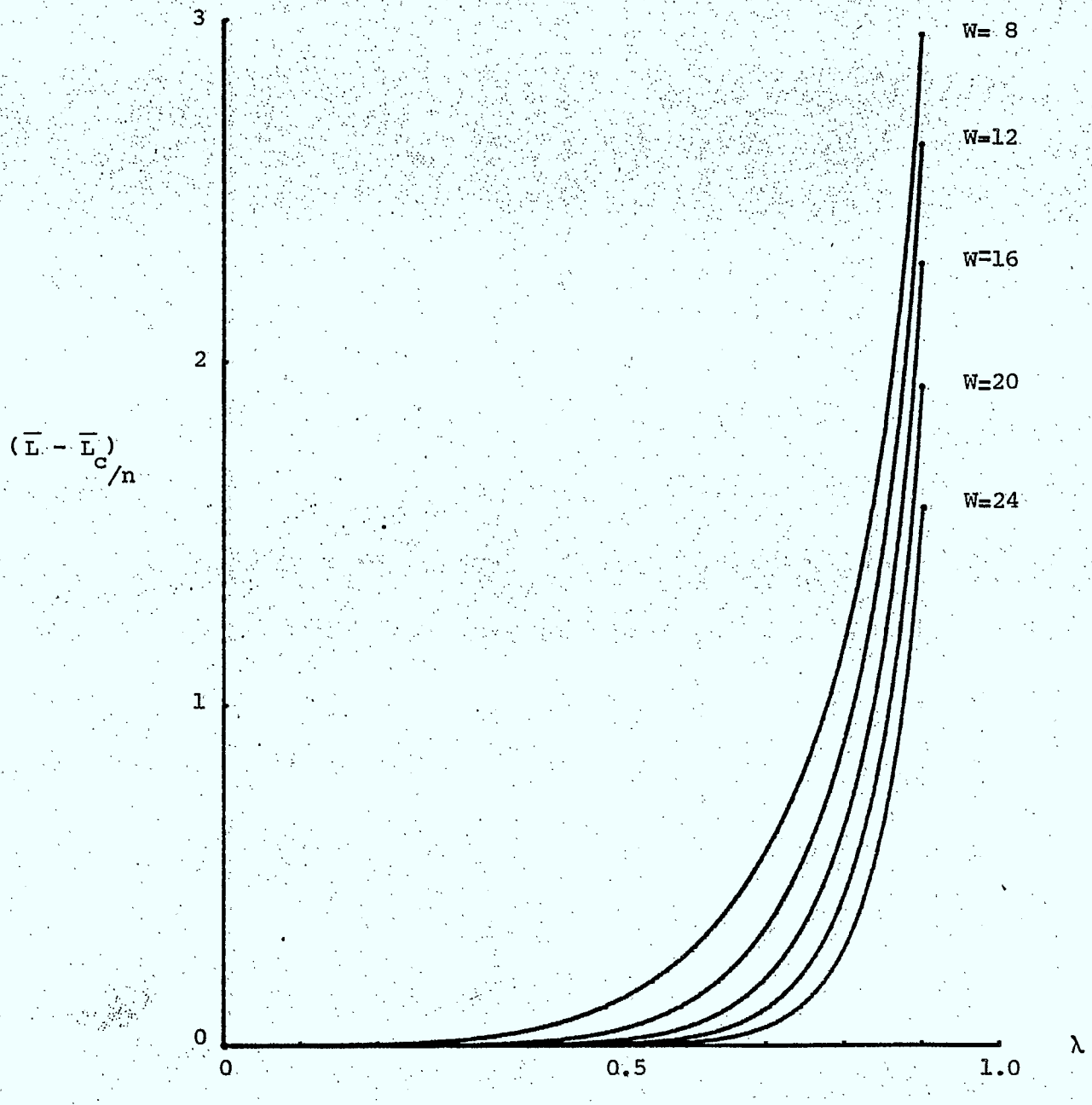

Fig. 3. Graph of the differences in normalized average lengths between the cases with and without distortion, with W as parameter.

The average weight $\bar{w}$ per block is given by

$$\bar{w} = \pi_o n(o) + \pi_1 n(1) + \pi_2 n(2) + \ldots + \pi_t n(t)$$

$$= n(\pi_1 + 2\pi_2 + 3\pi_3 + \ldots + t\pi_t) \qquad (4.4.1)$$

Therefore the average distortion D per digit in schemes (b) and (c) is given by

$$D_c = \pi_1 + 2\pi_2 + 3\pi_3 + \ldots + t\pi_t - \frac{W}{n} .$$

In the next section, again for the distortionless case, an expression for the average length of an encoded sequence is derived for another type of source probability distribution.

5. AVERAGE LENGTH $\bar{L}$ of σ :

If π_i is as defined in section 4.4, then the average weight $\bar{w}$ per block B is given by (4.4.1), i.e.

$$\bar{w} = n(\pi_1 + 2\pi_2 + 3\pi_3 + \ldots + t\pi_t).$$

Expressing

$$\bar{w} = Nt + \rho \ , \ o < \rho \leq t,$$

we see that $\mathcal{B}$ has N+1 rows so that the average length $\bar{L}$ of σ is given by

$$\bar{L} = (N+1)\, rt$$

$$= < \frac{n(\pi_1 + 2\pi_2 + 3\pi_3 + \ldots + t\pi_t)}{t} > rt$$

Since

$$n = \frac{p^r - 1}{2} = \frac{(p-1)}{2}(1 + p + p^2 + \ldots + p^{r-1}) = t(1 + p + p^2 + \ldots + p^{r-1}),$$

we can also write

$$\bar{L} = <(1 + p + p^2 + \ldots + p^{r-1})\ (\pi_1 + 2\pi_2 + \ldots + t\pi_t) > rt$$

Given a distribution $\pi = \{\pi_o, \pi_1, \ldots \pi_t\}$,

we can compute $\bar{L}$.

For instance, if $\pi_i = \pi_o \lambda^i$, then

$$\bar{w} = \frac{n\lambda}{1-\lambda^{t+1}} \left(\frac{1 - \lambda^t}{1 - \lambda} - t\lambda^t\right)$$

so that

$$\bar{L} = \left\langle (1+p+p^2+\ldots+p^{r-1}) \frac{\lambda}{1-\lambda^{t+1}} \left(\frac{1-\lambda^t}{1-\lambda} - t\lambda^t \right) \right\rangle rt$$

or

$$\bar{L} = \left\langle \frac{n\lambda}{t(1-\lambda^{t+1})} \left(\frac{1-\lambda^t}{1-\lambda} - t\lambda^t \right) \right\rangle > rt \qquad (5.1)$$

In deriving this we note that

$$\pi_o = \frac{1-\lambda}{1-\lambda^{t+1}}$$

As an example let us consider the case of $n = 12$, $r = 2$, $t = 2$, $\lambda = \frac{1}{2}$.

Then $\bar{L} = 16$ as against $n = 12$.

Figure 4 is a graph of equation (5.1) (normalized by division by n) for the new source probability distribution. This should be contrasted with figure 1.

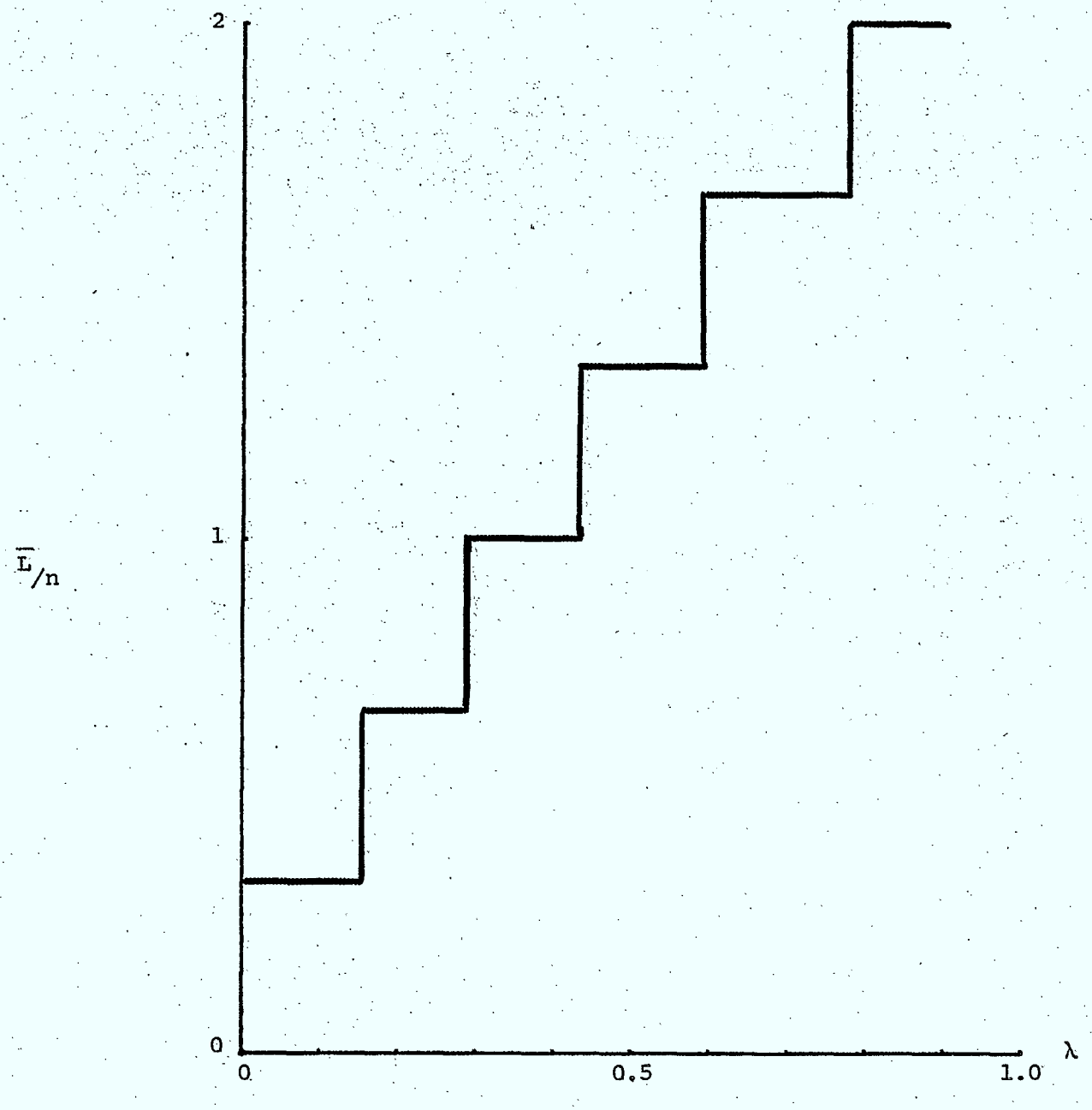

Fig. 4. Graph of the normalized average length of an encoded sequence for the distortionless case.

6. CONCLUDING REMARKS

After a short description of negacyclic codes, two variations of a scheme for joint source and channel encoding were presented ; the first was distortionless and always reduced source redundancy, providing the source statistics were appropriate. In the second method, distortion was permitted and this allowed the constraints on the source statistics to be loosened. Both cases resulted in an error-detection capability.

It is clear from the discussion that the scheme could easily be modified such that the decomposition of a source n-tuple produces an orthogonal matrix $\mathcal{B}$, orthogonal in the sense that no column or row of $\mathcal{B}$ would contain more than one non-zero entry. Consequently, on the average, the number of rows of $\mathcal{B}$, and hence, the length of the encoded sequence, would increase, but so would the error-detection capability, since more constraints are being placed on $\mathcal{B}$, and hence, on $\mathcal{B}'$.

For example, if we consider the usual negacyclic code of length 12, then for B = 0 1 -2 2 1 1 0 0 -2 1 1 1 , we have

$$\mathcal{B} = \begin{bmatrix} 0&1&0&0&0&0&0&0&0&0&0&0 \\ 0&0&-2&0&0&0&0&0&0&0&0&0 \\ 0&0&0&2&0&0&0&0&0&0&0&0 \\ 0&0&0&0&1&0&0&0&0&0&0&0 \\ 0&0&0&0&0&1&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&-2&0&0&0 \\ 0&0&0&0&0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&0&0&0&0&1 \end{bmatrix}$$

if no distortion is permitted. Thus the number of rows of $\mathcal{B}$ is equal to the number of nonzero elements in B. We further note that each row contains exactly one nonzero entry, and each column, at most, one nonzero entry.

REFERENCES

Anderson, J.B., 1974, IEEE Trans. Inform. Theory, vol. 20, 211.

Anderson, J.B., and Jelinek, F., 1973, IEEE Trans. Inform. Theory, IT-19, 77.

Ash, R., 1965, Information Theory (New York: Interscience).

Berger, T., 1971, Rate Distortion Theory (Englewood Cliffs, N.J.: Prentice-Hall).

Berlekamp, E.R., 1968, Algebraic Coding Theory (New York : McGraw-Hill) .

Cover, T.M., 1973, IEEE Trans. Inform. Theory, IT-19, 73.

Fano, R.M., 1961, Transmission of Information (New York: MIT Press/ Wiley).

Goblick, T.J., Jr., 1962, Coding for a Discrete Information Source with a Distortion Measure, Ph.D. dissertation, M.I.T., Cambridge, Mass.

Gray, R.M., 1975, IEEE Trans. Inform. Theory, IT-21, 357.

Hellman, M.E., 1975, IEEE Trans. Inform. Theory, IT-21, 651.

Huffman, D.A., 1952, Proc. I.R.E., Vol. 40, 1098.

Jelinek, F., 1969, IEEE Trans. Inform. Theory, IT-15, 584.

Jelinek, F., and Anderson, J.B., 1971, IEEE Trans. Inform. Theory, IT-17, 118.

Lynch, T.J., 1974, IEEE Trans. on Comm., COM 22, 151.

Meyr, H., Rosdolsky, H.G., and Huang, T.S., 1974, IEEE Trans. on Comm., COM-22, 826.

Molinder, J.I., 1974, IEEE Trans. Inform. Theory, IT-20, 336 .

Shannon, C.E., 1948, Bell System Technical Journal, vol. 27, 379.

Shannon, C.E., 1959, IRE Nat'l Conv. Rec., part 4, 142.

Wilkins, L.C. and Wintz, P.A., 1971, IEEE Trans. Inform. Theory, IT-17, 180.

PART 2: SOME RESULTS ON NEGACYCLIC CODES

1. INTRODUCTORY MATERIAL

In this part we present a few results concerning negacyclic codes (Berlekamp 1968). Some of these results have a bearing on the problem of data-compression.

Since the metric used in negacyclic codes is the Lee metric, (Berlekamp 1968) we begin the discussion with this metric.

The LEE WEIGHT of a sequence $a_0\, a_1\, a_2 \ldots a_{n-1}$ over GF(p), where p is an odd prime, is the sum of the Lee weights of a_i's. The Lee weight of a_i is a_i if $0 \le a_i \le \frac{p-1}{2}$ and $p - a_i$ if $\frac{p+1}{2} \le a_i \le p-1$.

From the definition of Lee weight it is clear that we can treat the negacyclic code as being over $\mathcal{S} = \{ -\frac{p-1}{2}, -\frac{p-3}{2}, \ldots -1, 0, 1, 2, \ldots, \frac{p-1}{2} \}$ instead of being over GF(p).

By the LEE DISTANCE between two n-tuples $a_0\, a_1 \ldots a_{n-1}$ and $b_0\, b_1 \ldots b_{n-1}$, both over GF(p), is meant the Lee weight of the n-tuple $a_0 - b_0\; a_1 - b_1\; a_2 - b_2 \ldots a_{n-1} - b_{n-1}$.

To indicate the situations where the Lee metric is more relevant than the Hamming metric, we mention the following known points (Berlekamp 1968): while the Hamming metric is well-suited to orthogonal modulation schemes, the Lee metric is well-suited to phase-modulation schemes. For channels using amplitude modulation, both the Hamming and Lee metrics are not flawless. However, if the alphabet size is large, the Lee metric provides a better approximation than the Hamming metric. Generally speaking, wherever the magnitude of the error is to be taken into account, the Lee metric is a better choice than the Hamming metric.

The codes which are best suited to the Lee metric are the negacyclic codes invented by Berlekamp.

A NEGACYCLIC CODE of block length n and over GF(p), where p is an odd prime and n is a nonmultiple of p, is the set of all of the

multiples of a generator polynomial g(x) which divides $1 + x^n$ over GF(p).

Since the metric used for a negacyclic code is the Lee metric, we will treat the code as being over $\mathcal{S}$ rather than as being over GF(p) whenever it is convenient to do so.

Defining the NEGACYCLIC SHIFT of an n-tuple $a_0\ a_1\ a_2 \cdots a_{n-1}$ over $\mathcal{S}$ as the sequence $-a_{n-1}\ a_0\ a_1\ a_2 \cdots a_{n-2}$, where we note, for the sake of clarity, that the minus sign goes with only a_{n-1}, we see that the negacyclic shift of a word, of a negacyclic code, also belongs to the code since g(x) divides $1+x^n$. This explains why the word negacyclic was coined by Berlekamp.

We remark that the definition of a negacyclic shift is parallel to that of a cyclic shift, which is relevant to cyclic codes.

Following the usage in the case of cyclic codes, a negacyclic code of block length $n = \frac{p^r-1}{2}$, where r is any positive integer, is said to be a PRIMITIVE NEGACYCLIC CODE (Van Chinh 1974).

On the other hand if the block length n is a divisor of $\frac{p^r-1}{2}$, then the code is said to be a NONPRIMITIVE NEGACYCLIC CODE.

Like cyclic codes, negacyclic codes can also be used for DATA COMPRESSION in many ways. One such method follows (Berger 1971). With reference to the source-output which is over GF(p), a negacyclic code of block length n is chosen and the standard array formed. We note that the standard array is nothing but an arrangement of the code and its cosets in a certain manner. The source-output is partitioned into blocks of n digits. For a given block the relevant code word is determined from the standard array. This assures us that the Lee distance between the block and the code word determined is the least possible. Then the k information digits corresponding to this code word are determined and the n-digit block replaced with this k-tuple. It is this k-tuple that is finally transmitted. The resulting compression-ratio C is given by

$$C = \frac{n}{k} \qquad (1)$$

and the average distortion D per digit is given by

$$D = \frac{\sum_{\omega=1}^{n} K_\omega \omega}{p^{n-k} n}, \qquad (2)$$

where K_ω is the number of cosets, with leaders of Lee weight ω, in the standard array. With respect to (1) and (2) it is assumed that the source puts out all of the p symbols with the same probability.

From the previous discussion it is clear that the code has to be so chosen that C is as high as possible for a given D or D is as low as possible for a given C.

Now we come to a discussion of the error-correcting capability of negacyclic codes (Berlekamp 1968). The most important result in this regard is contained in the following known

THEOREM 1: With α as a primitive root of $x^{2n}-1$, if the roots of the generator polynomial of a negacyclic code over $\mathcal{S}$ include $\alpha, \alpha^3, \alpha^5, \ldots \alpha^{2t-1}$, where $t \leq \frac{p-1}{2}$, then the code can correct any single error-pattern of Lee weight t or less.

Recalling the rule of construction for BCH codes, we see a parallel between that rule and Theorem 1. However, there is a major difference in that the condition of $t \leq \frac{p-1}{2}$ does not exist in the case of BCH codes.

The Theorem 1 can clearly be used to obtain the generator polynomials g(x) of negacyclic codes. In this connection we note that α is a root of $x^n + 1$ since it has order 2n. Thus, to obtain g(x), we have to examine the irreducible factors of x^n+1, whereas in the case of cyclic codes, g(x) divides x^n-1.

By now it should have become clear that there is a great deal of similarity between cyclic codes and negacyclic codes. In fact many of the techniques obtained for cyclic codes can easily be used, with appropriate modification, for negacyclic codes. As an example of this situation we may

mention the Bose-Caldwell technique for synchronization in the case of cyclic codes (Van Chinh 1974).

An open problem in the case of negacyclic codes is how to obtain majority logic-decodable codes which are negacyclic or derived from negacyclic codes (Van Chinh 1974). We will not discuss this aspect except to give, for the sake of interest, the following example :
let us consider the negacyclic code which has $p = 5$ and $r = 2$ so that $n = \frac{p^r - 1}{2} = 12$ and $t = \frac{p-1}{2} = 2$. One possible $g(x)$ for this code is

$$g(x) = (x^2 - 2x+3)(x^2+ 2) = x^4 - 2x^3 + x+ 1.$$

Now let us consider the (11,4) shortened subset, of the negacyclic code, generated by

$$g_1(x) = g(x)(1-x+ 2x^3) = 1 -x^2 - x^5 + x^6 + 2x^7.$$

Every word $V(x)$ of this subset code can be expressed in the form

$$V(x) = g_1(x)(C_o + C_1 x + C_2 x^2 + C_3 x^3)$$

$$= v_o + v_1 x + x_2 x^2 + \ldots + v_{10} x^{10},$$

where $C(x) = C_0 + C_1 x + C_2 x^2 + C_3 x^3$ is the information polynomial.
Comparing the coefficients of the product $g_1(x)\, C(x)$ with those of $v_0 + v_1 x + v_2 x^2 + \ldots + v_{10} x^{10}$, and after suitable manipulation, we get

$$2C_0 = 2v_0,$$
$$2C_0 = -2v_2 - 2v_4,$$
$$2C_0 = 2v_1 + 2v_6,$$
$$2C_0 = -2v_5 - v_{10},$$
$$2C_0 = v_7 + 2v_8 + 2v_9.$$

Since these relations are orthogonal, any error-pattern of Lee weight 2 or less will affect at most 2 out of 5 relations. Hence we conclude that

the subset shortened code is 1-step majority-logic-decodable.

At this point we derive a result which will be required in later discussion, regarding the number of n-tuples, over $\mathcal{S}$, having any Lee weight. With N_ω representing the number of n-tuples with Lee weight ω, we readily see that $N_o = 0$ and that $N_1 = 2n$, since we can have a + 1 or -1 in any one of the n positions. As regards $\omega = 2$, we have $N_2 = 2n + \binom{n}{2} 2^2$, where 2n is the number of single errors of Lee weight 2 and $\binom{n}{2} 2^2$ is the number of 2-error-patterns, each error having a Lee weight 1. From the preceding discussion it is clear that the number N_ω depends on the number of ways in which ω can be expressed as the sum of numbers 1,2,3,... p-1. Since ω can be expressed as the sum of j numbers in $\binom{\omega-1}{j}$ ways, we conclude the following

THEOREM 2: The number N_ω of n-tuples, over $\mathcal{S}$, which have Lee weight ω is given by

$$N_\omega = 2^1 \binom{n}{1} \binom{\omega-1}{0} + 2^2 \binom{n}{2} \binom{\omega-1}{1} + 2^3 \binom{n}{3} \binom{\omega-1}{2} + \ldots + 2^\omega \binom{n}{\omega} \binom{\omega-1}{\omega-1}.$$

2. PARALLEL RESULTS

Next we present a few results, for negacyclic codes, which are parallel to those known for cyclic codes.

Parallel to the concept of cyclic classes in the case of cyclic codes, we define a NEGACYCLIC CLASS as the set of a code word and all of the distinct code words obtained by negacyclically shifting this code word. The number of code words in a negacyclic class is 2n or a divisor of 2n. This number is also called the PERIOD of the negacyclic class. Here we recall that a cyclic class has period n or a divisor of n.

Since the maximal-length cyclic code of block length n is a constant weight code in the sense that all of the nonzero words have the same Hamming weight of $\frac{n+1}{2}$, it seems reasonable to expect a similar result in the case of negacyclic codes. It is known that this is indeed the case.

Here we state, without proof, the following known (Berlekamp 1968)

THEOREM 3: The maximal - length negacyclic code of block length $n = \frac{p^r-1}{2}$ has a constant Lee weight of $\frac{p^{r-1}(p^2-1)}{8}$.

For the sake of completeness we mention that a MAXIMAL - LENGTH NEGACYCLIC CODE of block length $n = \frac{p^r-1}{2}$ is the code generated by $g(x) = \frac{x^n+1}{h(x)}$, where h(x) is an irreducible polynomial, over GF(p), which has degree r and has exponent 2n in the sense that h(x) divides $x^{2n}-1$ and not $x^{n'}-1$ for any $n' < 2n$.

From Theorem 3, we readily get the following

COROLLARY 1: The maximal-length negacyclic code over $\mathcal{S}$ corrects any single error-pattern of Lee weight $\left[\frac{\frac{p^{r-1}(p^2-1)}{8}-1}{2}\right]$ or less, where $[\nu]$ is the largest integer contained in the positive number ν.

With reference to Theorem 3 and Corollary 1, we see that

$$\frac{p^{r-1}(p^2-1)}{8} = p\,\frac{p^{r-2}(p^2-1)}{8} \geq p, \text{ for } r \geq 2,$$

indicating that the condition of $t \leq \frac{p-1}{2}$ in Theorem 1 is sufficient, but not necessary.

In the context of the fact that the binary Hamming codes, which have block length 2^m-1 and are single-error-correcting, are perfect, we state the following known

THEOREM 4: The negacyclic code V with $n = \frac{p^r-1}{2}$ and $t = 1$ is perfect.

We note that this V has $k = n-r$.

Here we also recall that a code with error-correcting capability t is said to be PERFECT if the maximum coset - leader - weight is t. The code is s and to be QUASI-PERFECT if the maximum coset-leader-weight is $t + 1$.

From Theorem 4 it follows that V has one coset with the leader

of all zeros and 1 coset with leader of weight 1. Using this fact and the fact that $k = n - r$ in (1) and (2), we obtain the following

COROLLARY 2: The negacyclic code V, with $n = \frac{p^r - 1}{2}$ and $t = 1$, has $C = \frac{n}{n - r}$ and $D = \frac{1}{n + \frac{1}{2}}$.

It is known that the binary Hamming code with length $2^m - 1$ has $\frac{(2^m - 1)(2^m - 2)}{6}$ words of Hamming weight 3. As a parallel to this result we now prove the following

THEOREM 5: The negacyclic code V, with $n = \frac{p^r - 1}{2} = \frac{3^r - 1}{2}$ and $t = 1$, has $\frac{2n(n-1)}{3}$ words of weight 3.

We begin the proof by noting that V is over $\mathcal{F} = \{-1, 0, 1\}$ so that every word $V_3(x)$, of weight 3, has the form $V_3(x) = a_u x^u + a_v x^v + a_w x^w$ where, a_u, a_v and a_w are from $\{-1, 1\}$. This means, in view of the definition of a negacyclic class, the following

LEMMA 1: In V of Theorem 5, every negacyclic class with words of weight 3, has exactly 3 words of the form $1 + a_v x^v + a_w x^w$.

From Theorem 4 we see that V has one coset with the leader of all zeros and 2n cosets with leaders of weight 1, thus accounting for all of the $p^{n-k} = 3^r$ cosets. Since V has minimum weight 3, all of the $2^2\binom{n}{2} = 2n(n-1)$ n-tuples of weight 2, as computed from Theorem 2, are contained in the cosets with leaders of weight 1. This means, since V is negacyclic and all of the n-tuples of weight 1 have been used as coset leaders, that the number of n-tuples of weight 2 per coset with leader of weight 1 is given by $\frac{2n(n-1)}{2n} = n - 1$. We repeat this fact in the following

LEMMA 2: In V of Theorem 5, every coset with leader of weight 1, has n-1 n-tuples of weight 2.

Now let us consider, without losing generality, the coset with leader x^o. This leader can produce an n-tuple of weight 2 only with a code word

of the form $1 + a_v x^v + a_w x^w$. This means, in view of Lemma 2, that V has n-1 words of the form $1 + a_v x^v + a_w x^w$. Therefore, using Lemma 1, we now get the following

LEMMA 3: V of Theorem 5 has $\frac{n-1}{3}$ negacyclic classes containing words of weight 3.

Since 3 does not divide $n = \frac{3^r-1}{2}$, it follows that a word of weight 3 produces a negacyclic class with 2n words. Combining this fact with Lemma 3 we get Theorem 5.

This concludes the proof of Theorem 5.

It is known that the shortened-by-1-bit Hamming code is quasi-perfect. Now we prove the following parallel

THEOREM 6: The (n-1,k-1) code V' obtained by shortening V of Theorem 5 by 1 bit is quasi-perfect.

Since p=3, every word of V with weight 3 has the form $a_u x^u + a_v x^v + a_w x^w$, where a_u, a_v and a_w are taken from $\{-1, 1\}$. As already seen a word like this produces a negacyclic class with 2n words. This means that every negacyclic class with words of weight 3, in V, has 6 words ending with a nonzero digit. Using this fact with Lemma 3, we see that V has $\frac{6(n-1)}{3} = 2(n-1)$ words of weight 3 ending in a nonzero digit. This leads, in view of Theorem 5, to

LEMMA 4: V' of Theorem 6 has $\frac{2n(n-1)}{3} - 2(n-1) = \frac{2(n-1)(n-3)}{3}$ words of weight 3.

Now let us consider the cosets of V'. Since V' has t=1, we have 1 coset with the leader of all zeros and 2(n-1) cosets with leaders of weight 1, accounting for 1+2(n-1) = 1+2n-2 = 2n-1 cosets. This means that we have yet to account for $3^r - (2n-1) = (2n+1) - (2n-1) = 2$ cosets. According to Lemma 4, V' has $\frac{2(n-1)(n-3)}{3}$ words of weight 3. Since all of the 2(n-1) coset leaders of weight 1 have been used, the number of words of weight 2, in the cosets of V', which have already occurred is

$\frac{2(n-1)(n-3)}{3} \cdot 3 = 2(n-1)(n-3)$. But the total number of $(n-1)$-tuples of weight 2 is $2^2\binom{n-1}{2} = 2(n-1)(n-2)$. Therefore we have the following

LEMMA 5: With respect to V' of Theorem 6, there are $2(n-1)(n-2) - 2(n-1)(n-3) = 2(n-1)$ $(n-1)$-tuples, of weight 2, available for the construction of cosets with leaders of weight 2.

A leader of weight 2 can produce a word of weight 2 only with those words of V' which have weight 3 or 4.

Without losing generality let us consider the $(n-1)$ - tuple $\lambda_2 = 1\ 1\ 0\ 0 \ldots 0$. A word, of V', which has a distance of 2 from λ_2 has to be of the form $1\ -1\ x\ x\ x \ldots x$ or $-1\ 1\ x\ x\ x \ldots x$, where only one x is nonzero. Since the distance between these two forms has to be at least 3, we conclude the following

LEMMA 6: With respect to V' of Theorem 6, a leader of weight 2 can produce, with codewords of weight 3, at most 2 words of weight 2.

Next we have to investigate the number of $(n-1)$ tuples, of weight 2, that can exist at most in a coset with leader of weight 2, because of code words of weight 4 in V'. This number is the same as the number ξ_ν of ν - tuples possible such that all of them have Lee weight 4, have the form of, say, $1\ 1\ x\ x\ x \ldots x$ with only 2 x's nonzero, the mutual distance being at least 3. For $\nu = 3$, ξ_3 is clearly zero. For $\nu = 4$, $\xi_4 = 1$, since we can choose $1\ 1\ 1\ 1$, $1\ 1\ -1\ 1$, $1\ 1\ 1\ -1$ or $1\ 1\ -1\ -1$. Suppose we choose $1\ 1\ 1\ 1$, without losing generality. For $\nu = 5$, $\xi_5 = 3$, since we can have $1\ 1\ 1\ 1\ 0$, $1\ 1\ -1\ 0\ 1$ and $1\ 1\ 0\ -1\ -1$, or $1\ 1\ 1\ 1\ 0$, $1\ 1\ -1\ 0\ -1$ and $1\ 1\ 0\ -1\ 1$. Suppose we choose the former three words, again without losing generality. For $\nu = 6$, $\xi_6 = 3$, since we cannot add one more word. For $\nu = 7$, $N_7 = 4$, since we can have $1\ 1\ 1\ 1\ 0\ 0\ 0$, $1\ 1\ -1\ 0\ 1\ 0\ 0$, $1\ 1\ 0\ -1\ -1\ 0\ 0$, $1\ 1\ 0\ 0\ 0\ 1\ 1$, or $1\ 1\ 1\ 1\ 0\ 0\ 0$, $1\ 1\ -1\ 0\ 1\ 0\ 0$,

1 1 0 - 1 - 1 0 0, 1 1 0 0 0 1 - 1, or 1 1 1 1 0 0 0, 1 1 - 1 0 1 0 0, 1 1 0 - 1 - 1 0 0, 1 1 0 0 0 - 1 1, or 1 1 1 1 0 0 0, 1 1 - 1 0 1 0 0, 1 1 0 - 1 - 1 0 0, 1 1 0 0 0 - 1 - 1. Continuing this line of reasoning we get the following

LEMMA 7: The number ξ_ν of ν - tuples, over GF(3), which are such that all of them have weight 4 and a mutual distance of at least 3, and match in two given positions, is given by $\xi_\nu = 3(\psi - 1)$ for $\nu = 3$ and $\nu = 3\psi - 1$, and $\xi_\nu = 3(\psi - 1) - 2$ for $\nu = 3\psi - 2$, where ψ is a positive integer.

From Lemma 7 we directly get the following

LEMMA 7': The number of (n-1) - tuples, of weight 2 that can exist in a coset of V' with a leader of weight 2, because of words of weight 4 in V', is at most n-4.

Combining Lemmas 6 and 7', we see that a leader of weight 2 can produce at most n-4+2 = n-2 (n-1)-tuples of weight 2 in a coset of V'. This means, including the leader itself, that there can be at most n-1 (n-1)-tuples of weight 2 in a coset of V' with a leader of weight 2. This in turn implies, in view of Lemma 5, that there are 2 cosets of V' with leaders of weight 2.

To summarize the discussion so far made, we have accounted for all of the $p^r = 3^r$ cosets of V'. There are, in addition to V' itself, 2(n-1) cosets with leaders of weight 1 and 2 cosets with leaders of weight 2, implying that V' of Theorem 6 is quasiperfect.

This concludes the proof of Theorem 6.

With reference to (1) and (2), Theorem 6 implies the following

COROLLARY 3: For the code V' of Theorem 6, $C = \frac{n-1}{n-1-r}$ and $D = \frac{n+1}{(n+\frac{1}{2})(n-1)}$.

It is known that the code obtained by interleaving two Hamming codes

is quasi perfect. Now we prove the following parallel

THEOREM 7: Let V_1 be the $(n_1 = \frac{p^{r_1}-1}{2}, k_1 = n_1 - r_1)$ negacyclic code with $t = 1$ and V_2 the $(n_2 = \frac{p^{r_2}-1}{2}, k_2 = n_2 - r_2)$ negacyclic code with $t = 1$. Then the $(n = n_1 + n_2, k = k_1 + k_2)$ code V, obtained by interleaving V_1 and V_2, is quasiperfect.

We begin the proof by noting that V can correct any error-pattern which, on deinterleaving, reduces to a single error, of weight 1, per code. This means that V can correct 1 error-pattern of weight 0, $2n = 2n_1 + 2n_2$ error-patterns of weight 1 and $2n_1\, 2n_2 = 4n_1 n_2$ error-patterns of weight 2. Since $1 + 2n + 4n_1 n_2 = 1 + 2(n_1 + n_2) + 4n_1 n_2 = 1 + 2(\frac{p^{r_1}-1}{2} + \frac{p^{r_2}-1}{2}) + 4(\frac{p^{r_1}-1}{2})(\frac{p^{r_2}-1}{2}) = p^{r_1+r_2} = p^{n-k}$ and since V has t=1, we conclude that V is quasiperfect.

This concludes the proof of Theorem 7.

With reference to (1) and (2), Theorem 7 implies the following

COROLLARY 4: For the code V of Theorem 7,

$$C = \frac{n_1+n_2}{n_1+n_2-r_1-r_2} \quad \text{and} \quad D = \frac{n_1}{n_1+n_2}\,\frac{1}{n_1+\frac{1}{2}} + \frac{n_2}{n_1+n_2}\,\frac{1}{n_2+\frac{1}{2}}.$$

From the expression for D in Corollary 4, we see, with reference to Corollary 2, that the average distortion per digit of the interleaved code is less than the sum of the distortions of the component codes.

Now we consider interleaving two codes of Theorem 6. Specifically we prove the following

THEOREM 8: If V_1 is the code obtained by shortening the $(n_1 = \frac{p^{r_1}-1}{2}, k_1 = n_1 - r_1)$ negacyclic code with $t = 1$ by 1 digit and V_2 is

the code obtained by shortening the ($n_2 = \frac{p^{r_2}-1}{2}$, $k_2 = n_2 - r_2$) negacyclic code with t = 1 by 1 digit, where p = 3, then for the ($n = n_1+n_2$, $k = k_1+k_2$) interleaved code $C = \frac{n_1+n_2-2}{n_1+n_2-2-r_1-r_2}$ and $D = \frac{8n_1n_2+16}{(2n_1+1)(2n_2+1)(n_1+n_2-2)}$) .

The proof of Theorem 8 is as follows:

that $C = \frac{n_1+n_2-2}{n_1+n_2-2-r_1-r_2}$ follows trivially from (1). To obtain the expression for D we proceed in the following way.

Defining N_{ij} to be the number of cosets of V_i with leaders of weight j, i = 1,2, we see from Theorem 6 that $N_{i0} = 1$, $N_{i1} = 2(n_1-1)$, $N_{i2} = 2$. Since V can correct any error-pattern which, on de-interleaving, reduces to coset leaders of V_1 and V_2, it follows that $N_0=1$, $N_1 = 2(n_1-1+n_2-1)$, $N_2 = 2(n_1-1)\,2(n_2-1) + 2 + 2$, $N_3 = 2(n_2-1) + 2(n_1-1)$, $N_4 = 2.2$, where N_ω is the number of cosets, of V, with leaders of weight ω. This means that V gives, with reference to (2),

$$D = \frac{(0)(1) + (1)\;2(n_1-1+n_2-1) + (2)\,[\,4 + 4(n_1-1)(n_2-1)\,] + (3)\,[\,2(n_2-1) + 2(n_1-1)\,] + (4)(4)}{3^{r_1+r_2}}$$

$$= \frac{2(n_1+n_2-2) + 8\,[\,1 + (n_1-1)(n_2-1)\,] + 6[\,n_2-1+n_1-1\,] + 1}{(2n_1+1)(2n_2+1)(n_1+n_2-2)}$$

or

$$D = \frac{2(n_1+n_2-2) + 8(-n_1-n_2+n_1n_2+2) + 6\,[\,n_1+n_2-2\,] + 1}{(2n_1+1)(2n_2+1)(n_1+n_2-2)}$$

$$= \frac{(n_1+n_2-2)\,[2-8+6] + 8n_1n_2 + 16}{(2n_1+1)(2n_2+1)(n_1+n_2-2)}$$

$$= \frac{8n_1n_2+16}{(2n_1+1)(2n_2+1)(n_1+n_2-2)} .$$

This completes the proof of Theorem 8.

3. CONCLUDING REMARKS

We conclude this part by noting that, given the component codes $V_1, V_2, \ldots, V_u$ the coset-enumerations of which are known, it is easy to obtain the coset-enumeration of the interleaved code V. By a coset-enumeration we mean, with reference to (2), the set of numbers K_ω's. If $K_{1\omega}, K_{2\omega}, \ldots K_{u\omega}$ refer to $V_1, V_2, \ldots V_u$ respectively and if K_ω refers to V, then K_ω is given by

$K_\omega = \sum_{\omega_1+\omega_2+\ldots+\omega_j=\omega} K_{1\omega_1} K_{2\omega_2} K_{2\omega_3} \ldots K_{u\omega_u}$. The validity of this can be easily established starting with the fact that, if V_i is (n_i, k_i), then $K_{i,o} + K_{i,1} + \ldots + K_{i,n_i} = p^{n_i - k_i}$. An implication of the preceding discussion is that the compression-ratio C and the average distortion D for V can be computed easily.

REFERENCES

Berger, T., 1971, Rate Distortion Theory (Englewood Cliffs, N.J. : Prentice-Hall), pp. 200-207.

Berlekamp, E.R., 1968, Algebraic Coding Theory (New York: McGraw-Hill), pp 200-218, pp 305-316.

Van Chinh, C., 1974, A Study of Negacyclic Codes, M.A.Sc. thesis Univ. of Ottawa, Ottawa Canada, Ch. 2.