## Alternant Codes\*

### H. J. HELGERT

Goddard Space Flight Center, NASA, Greenbelt, Maryland

### I. Introduction

It is an easily established fact that an (n, k) linear error correcting code with symbols from GF(q) can be completely described in terms of a so-called parity check matrix H which has n columns and t rows of elements from  $GF(q^m)$ , where t and m are any positive integers and such that  $tm \ge n - k$ . The usual representation in terms of elements from GF(q) is obtained by writing each element in  $GF(q^m)$  as an m-tuple of elements from GF(q) in column form. The row space of H is then just the set of all linear combinations of the rows of H over GF(q) and forms a vector space of dimension n - k. (Throughout this paper we assume k > 0.)

The code's minimum distance d is equal to the smallest number of columns in H which are linearly dependent over GF(q). Thus, if every  $t \times t$  submatrix of H in  $GF(q^m)$  is nonsingular, all combinations of t or fewer columns will be independent, even over GF(q) and the code must have minimum distance at least t+1.

Consider now a matrix of the form

$$H = \begin{bmatrix} y_1 g_1(x_1) & y_2 g_1(x_2) & \cdots & y_n g_1(x_n) \\ y_1 g_2(x_1) & y_2 g_2(x_2) & \cdots & y_n g_2(x_n) \\ \vdots & \vdots & & \vdots \\ y_1 g_t(x_1) & y_2 g_t(x_2) & \cdots & y_n g_t(x_n) \end{bmatrix}, \tag{1}$$

where the y's are any (not necessarily distinct) nonzero elements of  $GF(q^m)$ , the x's are distinct elements of  $GF(q^m)$ 

$$g_j(x) = c_{0j} + c_{1j}x + c_{2j}x^2 + \cdots + c_{t-1,j}x^{t-1}$$

is a polynomial of degree less than or equal to t-1 with coefficients from  $GF(q^m)$ , for j=1,2,...t.

\* This work was performed in part during the author's tenure as a National Research Council Senior Postdoctoral Resident Research Associate. Clearly

$$H = CXY$$

where

$$C = \begin{bmatrix} c_{01} & c_{11} & \cdots & c_{t-1,1} \\ c_{02} & c_{12} & \cdots & c_{t-1,2} \\ \vdots & \vdots & & \vdots \\ c_{0t} & c_{1t} & \cdots & c_{t-1,t} \end{bmatrix},$$

$$X = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & & x_n \\ x_1^2 & x_2^2 & & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_1^{t-1} & x_2^{t-1} & \cdots & x_n^{t-1} \end{bmatrix},$$

$$Y = \begin{bmatrix} y_1 & 0 & \cdots & 0 \\ 0 & y_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & y_n \end{bmatrix}.$$

If we choose any submatrix of order t from H and compute its determinant we get the product of |C|, a Vandermonde determinant with distinct elements and the determinant of a diagonal submatrix of Y.

Thus, if C is nonsingular, H is the parity check matrix of a linear code of minimum distance at least t+1. The number of parity check symbols is at most mt.

Determinants of the above type are known as alternants [Muir and Metzler, 1930] and for this reason we refer to the codes corresponding to (1) as alternant codes. Several prominent classes of codes such as the BCH, Srivastava and Goppa (1970, 1971) codes are readily obtained as special cases. Since the error correcting capability of these codes for short and moderate block lengths is optimum or close to it, one is tempted to conjecture that other good classes of codes within the general context of (1) can be found. Of course, whether or not any such codes are of more than theoretical interest depends on the complexity of the encoding and decoding algorithms. Little is known on this subject in general, but it appears that for certain subclasses of (1) the problem is only slightly more difficult than for cyclic codes of comparable performance.

Although, as is well known, the performance of primitive BCH codes deteriorates with increasing block length, the generalization inherent in (1) is sufficient to assure the existence of alternant codes whose performance is

asymptotically arbitrarily close to the Varsharmov-Gilbert bound. In fact, from Appendix A and Goppa's work one easily proves the existence of such codes for some Y and

$$C = I,$$
 
$$x_i = \alpha^{i-1} \qquad i = 1, 2, \dots n,$$

where  $\alpha$  is a primitive *n*-th root of unity. Note that the BCH codes are obtained as the special case  $y_i = \alpha^{m_0(i-1)} = x_i^{m_0}$ .

Consider now the class of alternant codes obtained from (1) as follows:

Let t=rs, where r and s are positive integers, each  $z_i$  a nonzero element of  $GF(q^m)$ , and each  $w_i$  a distinct element of  $GF(q^m)$  different from all the  $x_i$ , and set

$$\begin{split} y_i &= \frac{z_i}{\prod_{j=1}^s (x_i - w_j)^r}, \\ g_{(l-1)r+k}(x_i) &= \frac{z_i}{y_i (x_i - w_l)^k} \end{split}$$

for i = 1, 2, ..., n; l = 1, 2, ..., s; k = 1, 2, ..., r. Substituting into (1) we obtain the matrix

$$H = \begin{bmatrix} H_1 \\ \cdots \\ H_2 \\ \cdots \\ \vdots \\ \cdots \\ H_s \end{bmatrix}, \tag{2}$$

where for l=1, 2,... s

$$H_l = egin{bmatrix} rac{z_1}{(x_1 - w_l)^1} & rac{z_2}{(x_2 - w_l)^1} & \cdots & rac{z_n}{(x_n - w_l)^1} \ rac{z_1}{(x_1 - w_l)^2} & rac{z_2}{(x_2 - w_l)^2} & \cdots & rac{z_n}{(x_n - w_l)^2} \ dots & dots & dots & dots \ rac{z_1}{(x_1 - w_l)^r} & rac{z_2}{(x_2 - w_l)^r} & \cdots & rac{z_n}{(x_n - w_l)^r} \ \end{bmatrix}.$$

It can easily be shown that every determinant of (2) of order rs is nonzero (Helgert, 1972) and hence these codes have minimum distance at least rs + 1.

The number of parity check symbols is at most mrs and since the x's in (2) must be distinct and different from the distinct w's the code length n cannot exceed  $q^m - s$ .

For r = 1 the codes defined by (2) are generalized versions of Srivastava's codes, the latter being obtained as the special case  $z_i = x_i^k$ , for some integer k.

For s=1 the codes are generalizations of the BCH codes. Here the correspondence is  $w_1=0$ ,  $z_i=(\alpha^{m_0-a})^{i-1}$  and  $x_i=\alpha^{-a(i-1)}$ , in the usual notation.

In the general case the codes are modifications of Goppa's separable codes (Goppa, 1971).

In this paper we investigate a subclass of the codes defined by (2) that is obtained by restricting the  $x_i$  to subfields of  $GF(q^m)$ . For these codes we establish minimum distance and redundancy bounds and derive a number of interesting equivalence and invariance properties.

### II. Nonprimitive Alternant Codes

Let  $m = \lambda \mu$ , where  $\mu$  and  $\lambda$  are integers greater than 1. Then  $GF(q^{\lambda})$  is a proper subfield of  $GF(q^m)$ . In (2) set  $s = \mu$ ,  $n \leq q^{\lambda}$ ,  $x_i \in GF(q^{\lambda})(i = 1, 2, ..., n)$ ,  $z_i \in GF(q^{\lambda}) = 0$  and  $w_j = w^{q^{\lambda(j-1)}}(j = 1, 2, ..., \mu)$ , where w is an arbitrary element of  $GF(q^m)$  that is not contained in any proper subfield of  $GF(q^m)$ .

The condition on w assures that the  $w_i$  are distinct and different from the  $x_i$ . This choice of parameters is therefore legitimate.

Consider now the element

$$\frac{z_i}{(x_i - w_1)^l} = \frac{z_i}{(x_i - w)^l}$$

in the *l*-th row and *i*-th column of  $H_1$  . Raising this to the  $q^{\lambda(j-1)}$ -th power gives

$$\left\{ \frac{z_i}{(x_i - w)^i} \right\}^{q^{\lambda(j-1)}} = z_i \left\{ \frac{1}{(x_i - w)^{q^{\lambda(j-1)}}} \right\}^i \\
= \frac{z_i}{(x_i - w_j)^i} \quad j = 2, 3, ..., \mu.$$

Thus, the *l*-th row in  $H_j$  is the  $q^{\lambda(j-1)}$ -th power of the *l*-th row in  $H_1$  and is consequently redundant, for l=1,2,... r and j=2,3,...  $\mu$ . We now make the following

DEFINITION. Let  $m = \mu \lambda$ , where  $\mu$  and  $\lambda$  are integers greater than 1. The nonprimitive alternant codes are defined by the parity check matrix

$$H = \begin{bmatrix} \frac{z_1}{x_1 - w} & \frac{z_2}{x_2 - w} & \frac{z_3}{x_3 - w} & \cdots & \frac{z_n}{x_n - w} \\ \frac{z_1}{(x_1 - w)^2} & \frac{z_2}{(x_2 - w)^2} & \frac{z_3}{(x_3 - w)^2} & \cdots & \frac{z_n}{(x_n - w)^2} \\ \vdots & & & \vdots \\ \frac{z_1}{(x_1 - w)^r} & \frac{z_2}{(x_2 - w)^r} & \frac{z_3}{(x_3 - w)^r} & \cdots & \frac{z_n}{(x_n - w)^r} \end{bmatrix}, (3)$$

where  $n \leqslant q^{\lambda}$ , each  $x_i$  is a different element of  $GF(q^{\lambda})$ ,  $z_i \in GF(q^{\lambda}) - 0$  and w is any element of  $GF(q^m)$  not in a proper subfield of  $GF(q^m)$ .

As an immediate consequence of our discussion above we then have

Theorem 1. The nonprimitive alternant codes of length  $n \leq q^{\lambda}$  have minimum distance  $d \geqslant \mu r + 1$  and at most mr check symbols.

EXAMPLE. Let q=2, m=6,  $\mu=2$ ,  $\lambda=3$ ; r=2, n=8. Choosing  $z_i=1$  (i=1,2,...8) and w equal to a primitive element of  $GF(2^6)$  which is a root of  $x^6+x+1$  we get

$$H = egin{bmatrix} rac{1}{0-w} & rac{1}{1-w} & rac{1}{w^9-w} & rac{1}{w^{18}-w} \ & rac{1}{w^{27}-w} & rac{1}{w^{36}-w} & rac{1}{w^{45}-w} & rac{1}{w^{54}-w} \ & rac{1}{(0-w)^2} & rac{1}{(1-w)^2} & rac{1}{(w^9-w)^2} & rac{1}{(w^{18}-w)^2} \ & rac{1}{(w^{27}-w)^2} & rac{1}{(w^{36}-w)^2} & rac{1}{(w^{45}-w)^2} & rac{1}{(w^{54}-w)^2} \ \end{bmatrix}.$$

The second row is the square of the first row and is consequently redundant. Expanding the first row in GF(2) leads to

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

Thus H is the parity check matrix of an  $(n = 2^3, k \ge 2^3 - 6)$  code with minimum distance  $d \ge 2 \cdot 2 + 1 = 5$ . The code words are 00000000, 11110001, 01011110, 101011111. Note that this code is not cyclic, nor can it be made cyclic by appending an overall parity check. In terms of our first formulation of alternant codes we have for this code

Since  $|C| = (w + w^8)^4$ , which is nonzero, it follows from Appendix A that CXY and XY are parity check matrices for the same code. But X is the parity check matrix of the extended primitive (8, 1) BCH code with  $m_0 = 0$ , designed distance 8 and actual distance 8. The effect of Y is therefore to lower the minimum distance to 5 and increase the number of information symbols to 2. Note also that  $y_i = h^{-1}(x_i)$ , where  $h(z) = (z + w)^{18}$  is the Goppa polynomial of the code.

We next study a subclass of the codes defined above and derive a number of equivalence and invariance relations.

III. The Case 
$$n = q^{\lambda}$$
 and  $z_i = z(i = 1, 2, ... n)$ 

If  $z_i = z$  for all *i* we can multiply each row of (3) by  $z^{-1}$  and the row space of *H* will remain invariant (see Appendix A). Thus without loss of generality we may and henceforth will assume z = 1. If we also set r = tq for some

positive integer t, then the (lq)-th row of H equals the q-th power of the l-th row (l = 1, 2, ..., t) and is therefore redundant. We thus have

Theorem 2. For  $z_i = 1 (i = 1, 2,... n)$  the nonprimitive alternant codes of length  $q^{\lambda}$  have minimum distance  $d \geqslant \mu tq + 1$  and at most (q - 1) tm check symbols, for any positive integer t = r/q.

For fixed values of the integers q, m, t and  $\lambda$  the codes are completely described by the parameter w. Since  $GF(q^{\lambda})$  is invariant under addition, multiplication by any nonzero element and raising all elements to the  $q^{l}$ -th power  $(l = 1, 2, ..., \lambda)$ , we have from Appendix A

THEOREM 3. For fixed q, m, t and  $\lambda$  the codes with w,  $w + \beta$ ,  $w\beta$  and  $w^{q^l}$  are equivalent for  $l = 1, 2, ... \lambda$  and any  $\beta \in GF(q^{\lambda}) - 0$ .

Next, consider the polynomial,

$$f(x) = x^n - \sigma x^{n-1} + \sigma,$$

where  $\sigma = (w^n - w)^{-1}$ . By simple substitution we can show that the roots of f(x) are the elements of the first row of H. Let us divide the l-th row of H by  $(-\sigma)^l$  (l=1,2,...r). By Appendix A this leaves the code invariant. Since for  $\mu=2$  we have  $\sigma^n=-\sigma$ , it follows easily that the elements of the first row of the new parity check matrix are the roots of  $x^n+x^{n-1}+1$ , independently of  $\sigma$ . Consequently we have:

THEOREM 4. For fixed q, m, t and  $\mu = 2$  all codes are equivalent.

We now form the extended codes by adding an overall parity check to each code word. The new parity check matrix is then

$$H_{0} = \begin{bmatrix} \frac{1}{1} & \frac{1}{x_{1} - w} & \frac{1}{x_{2} - w} & \frac{1}{x_{n} - w} & 0\\ \frac{1}{(x_{1} - w)^{2}} & \frac{1}{(x_{2} - w)^{2}} & \frac{1}{(x_{n} - w)^{2}} & 0\\ \vdots & \vdots & \vdots & \vdots\\ (x_{1} - w)^{r} & (x_{2} - w)^{r} & (x_{n} - w)^{r} & 0 \end{bmatrix}.$$
(4)

Clearly, the elements in the second row of  $H_e$  are the roots of the polynomial g(x) = xf(x).

Applying the transformation,

$$\left(\frac{1}{x_i - w}\right)^l \rightarrow \left(\frac{\alpha}{x_i + w} + b\right)^l, \qquad l = 0, 1, 2, \dots, r; \qquad i = 1, 2, \dots, n, \quad (5)$$

$$0 \rightarrow b,$$

where b is any root of g(x) and  $a = (\sigma - b)/\sigma$ , we note first that this leaves the row space of  $H_e$  invariant (see Appendix A). Second, we can show by direct substitution that for  $\sigma^n = -\sigma$ , i.e.,  $\mu = 2$ 

$$f\left(\frac{a}{x_i - w} + b\right) = 0$$
  $i = 1, 2, ..., n$ .

This transformation is therefore nothing more than a permutation of the columns of  $H_e$ . There are n+1 such transformations corresponding to the n+1 roots of g(x) and they form an Abelian group. It is easily shown that for i, j = 1, 2, ... n the respective choices,

$$b = \frac{1}{x_i - w}, \quad b = \frac{-1}{(x_i - w^n)}, \quad b = \frac{x_i - x_j}{(x_i - w^n)(x_j - w)},$$

permute the last column of  $H_e$  into the *i*-th column, the *i*-th column into the last column and the *i*-th column into the *j*-th column. Thus the group is also transitive.

We summarize these conclusions in:

THEOREM 5. For fixed q, m, t and  $\mu = 2$  all extended codes are invariant under a transitive group of permutations.

In the binary case we also have:

THEOREM 6. For fixed m, t, and  $\mu = q = 2$  all codes have odd minimum distance. For t = 1 the minimum distance equals 5 and the number of check symbols is exactly  $2\lambda$ , the smallest possible.

The first part of this result follows immediately from Appendix B and Theorem 5 and the second part is a consequence of Hamming's bound (Peterson and Weldon, 1972, p. 83).

In Table I we list the parameters of all nontrivial binary codes of this type for  $8 \le m \le 12$  and selected others for m = 14 and m = 16. Here  $d^*$  is the minimum distance of the dual code and the parameters n', k' and d' refer to shortened codes obtained by eliminating an appropriate set of  $d^*$  columns

from H. Many of these codes and their duals are optimum or as good as the best known, while the three shortened codes for m=14 and the last three shortened codes for m=16 represent an improvement over previously known values (Sloane, 1972).

TABLE I										
Para	ameters of	Some No	onprimit	ive Alterr	ant Code	sα				
4	47	7.	J							

m	t	n	k	d	$d^*$	n'	k'	d'
8	1	16	8	5	5	11	4	5
10	1	32	22	5	11	21	12	5
10	2	32	12	9	≪3	29	10	9
	1	64	52	5	25	39	28	5
10	2	64	40	9	≤15	49	26	9
12	3	64	28	13	≤5	59	24	13
	5	64	16	21	<b>≤</b> 5	59	12	21
	1	128	114	5	54	74	61	5
14	2	128	100	9	≤37	91	64	9
	3	128	86	13	≤30	98	57	13
	1	256	240	5	113	143	128	5
16	2	256	224	9	≪89	167	136	9
16	3	256	208	≤15	≪81	175	128	13
	4	256	192	€23	€75	181	118	17

 $<sup>^{</sup>a}$   $(q = \mu = 2; z_{i} = 1).$ 

For  $\mu > 2$  the codes defined in this section are generally not equivalent. For example, when m=15 and  $\mu=3$  there exist binary (32, 17) codes whose weight spectra for two values of w are given below. Here  $A_i$  is the number of codewords of weight i and  $\alpha$  is a primitive element of  $GF(2^{15})$ . Note, however, that if we extend these codes by appending an overall parity check the weight spectra become identical.

In general we have:

Theorem 7. For fixed q, m, t, and  $\mu = 3$  all extended codes are equivalent.

The proof is simple and consists of showing that under the transformation (5), with

$$a = \frac{1}{\sigma + \sigma^{n^2}} \qquad b = \frac{-\sigma}{\sigma + \sigma^{n^2}},$$

the elements of the second row of (4) are the roots of the polynomial  $x^{n+1} + x + 1$ , which is independent of  $\sigma$ .

For q=2 and t=1 it follows readily from the Griesmer or Hamming bounds that the minimum distance of these extended codes is 8. The number of check symbols is, of course, upper bounded by  $3\lambda+1$ , which is well within the Varsharmov-Gilbert bound. Hamming's bound can also be used to produce a lower bound on the number of check symbols and yields  $3\lambda-1$ .

TABLE II

	$A_7$	$A_8$	$A_9$	$A_{10}$	$A_{11}$	$A_{12}$	$A_{13}$	$A_{14}$	$A_{\mathbb{R}}$	15	$A_{16}$
$w = \alpha$	128	400	816	1887	4000	6948	10464	14316	174	40	18389
$w = \alpha^5$	128	400	826	1877	3944	7004	10592	14188	172	88	18541
	A	17	$A_{18}$	$A_{19}$	$A_{20}$	$A_{21}$	$A_{22}$	$A_{23}$	$A_{24}$	$A_{2!}$	<sub>5</sub> A <sub>26</sub>
$w = \alpha$	17	217	14450	10464	4 6756	5 4128	3 2076	736	274	144	39
$w = \alpha^5$	17	316	14350	10424	4 6796	5 4144	2060	728	282	146	37

#### Appendix A

Let  $\alpha$  be a primitive element of  $GF(q^m)$  which is a root of some primitive polynomial of degree exactly m and coefficients in GF(q). Then every element of  $GF(q^m)$  can be expressed as a polynomial in  $\alpha$  of degree at most m-1, with coefficients from GF(q).

Now if H is a matrix of l rows and n columns with elements from  $GF(q^m)$ , each element can be represented by the m coefficients of its corresponding polynomial arranged in a column. In this manner H expands into a matrix of ml rows and n columns in which the rows are n-tuples of elements from GF(q). The row space V of H is the set of all linear combinations of the rows of H over GF(q) and forms a vector subspace whose dimension equals the number of linearly independent rows in H.

We state here some simple properties of V.

Theorem A.1. V is invariant under a permutation of the rows of H in  $GF(q^m)$ .

THEOREM A.2. V is invariant under multiplication of the elements of any row of H in  $GF(q^m)$  by any nonzero element of  $GF(q^m)$ .

THEOREM A.3. V is invariant under the addition of any multiple of one row of H in  $GF(q^m)$  to any other row of H in  $GF(q^m)$ .

THEOREM A.4. V is invariant under the operation of raising the elements of any row of H in  $GF(q^m)$  to the q-th power.

Theorem A.5. V is invariant under premultiplication of H in  $GF(q^m)$  by any  $l \times l$  nonsingular matrix in  $GF(q^m)$ .

The proofs of these theorems require nothing more than elementary algebra and will be omitted.

# APPENDIX B

Theorem B.1. Let  $A_i$  be the number of codewords of weight i in a binary linear code C of length n and let  $B_i$  be the number of codewords of weight i in the dual code. If the extended code obtained from C by appending an overall parity check to each codeword is invariant under a transitive group of permutations of the code digits, then

$$A_i = \left(\frac{n+1-i}{i}\right) A_{i-1} \qquad \text{for even } i \tag{B1}$$

and

$$B_i = \left(\frac{n+1-i}{i}\right) B_{n+1-i}$$
 for  $i = 1, 2, ..., n$ . (B2)

*Proof.* Since equality (B1) is well known (Berlekamp, 1968, p. 228), we omit its proof.

To show (B2), let H be the parity check matrix of C. Then the parity check matrix of the extended code is

and if the extended version of C is invariant under a transitive group of permutations, so is the row space of  $H_e$ . Now this row space is the union of two sets of vectors A and  $A_c$ , where A is the row space of H with a zero appended to each vector and  $A_c$  is obtained from A by complementing all digits. Hence, if  $C_i$  is the total number of vectors of weight i in the row space of  $H_e$ , then

$$C_i = B_i + B_{n+1-i}$$
.

If we list these  $C_i$  vectors as the rows of an array, then because of the invariance property, every column must have the same total weight and from the last column of  $H_e$  we see that this total weight must equal  $B_{n+1-i}$ . Therefore,

$$(n+1) B_{n+1-i} = iC_i + B_{n+1-i}$$

and this implies (B2).

A simple consequence of Theorem B.1 is:

COROLLARY B.2. Let C be a binary linear code of length n whose extended code is invariant under a transitive group of permutations. If d,  $d^*$  and  $D^*$  are the minimum weight of C and the minimum and maximum weights of the dual of C, respectively, then d is odd and  $d^* + D^* = n + 1$ .

RECEIVED: August 27, 1973; REVISED: May 30, 1974

#### REFERENCES

Berlekamp, E. R. (1968), "Algebraic Coding Theory," McGraw-Hill, New York. Goppa, V. D. (1970), A New Class of Linear Error-Correcting Codes, *Probl. Peredaci Informacii* 6, 24.

GOPPA, C. D. (1971), Rational Representations of Codes and (L, g)-Codes, Probl. Peredaci Informacii 7, 41.

HELGERT, H. J. (1972), Noncyclic generalizations of BCH and Srivastava codes, Inform. Contr. 21, 280.

Muir, T. and Metzler, W. H. (1930), "A Treatise on the Theory of Determinants," privately published, Albany, NY.

Peterson, W. W. and Weldon, E. J. (1972), "Error-Correcting Codes," 2nd ed. MIT Press, Cambridge, MA.

SLOANE, N. J. A. (1972), A survey of constructive coding theory, and a table of binary codes of highest known rate, *Discrete Math.* 3, 265.