INFORMATION AND CONTROL **13**, 378–400 (1968)

A Class of Optimum Nonlinear Double-Error-Correcting Codes*

FRANCO P. PREPARATA

Coordinated Science Laboratory and Department of Electrical Engineering, University of Illinois at Urbana, Illinois 61801

Nonlinear double-error-correcting block codes of length $(2^n - 1)$ (n even) are presented in this paper. They have the largest possible number of code-words for their length and minimum distance and are formed by adjoining to a certain linear code (referred to as the "kernel") a specific subset of its cosets. The kernel results from the juxtaposition and superposition of Bose–Chaudhuri–Hocquenghem codes of length $(2^{n-1} - 1)$. The presented codes are systematic and are comparable to the corresponding linear codes with regard to the complexity of the encoding and decoding operations.

1. INTRODUCTION

Some examples of nonlinear binary codes have been reported in the literature over the past years (Vasil'ev, 1962; Nadler, 1962; Green, 1966). Particularly interesting for its structure and generality was the class discovered by Vasil'ev (1962), i.e., a class of perfect single-error-correcting group and nongroup codes containing the Hamming codes.

Recently some interest in nonlinear codes has been revived by the discovery made by Nordstrom and Robinson (1967) of a (15, 8) nonlinear double-error-correcting code, of which previously reported (12, 5) (Nadler, 1962) and (13, 6) (Green, 1966) nongroup codes were shortened versions. The (15, 8) code had the interesting features of being systematic and of meeting the Johnson's upper bound (1962) on the number of code words in a code of length 15 and distance 5. Subsequently the (15, 8) code has been described in terms of polynomial (i.e., linear) codes over GF(2) (Preparata, 1968a): This description proved to be a useful framework, since it led to the formal demonstration (Preparata,

* This work was supported in part by the Joint Services Electronics Program under contract DDAB-07-67-C-0199 and in part by NSF Grant GK-2339.

1968b) of the distance properties of the code, previously heuristically assessed.

A question which was first asked by Nordstrom and Robinson (1967) was whether the (15, 8) code was a member of a class of codes. The purpose of this paper is to answer this question in the affirmative. Nongroup double-error-correcting $(2^n - 1, 2^n - 2n)$ codes exist for each even $n \geqq 4$, and contain the (15, 8) code as a special case. Here again the polynomial description has been the essential device in the construction of these codes.

The interesting features of these codes can be summarized as follows: 1) They contain twice as many code words as the double-error-correcting BCH codes of the same length, which is the largest number of code words possible for given length and distance, i.e., they are optimal; 2) their decoding can be based on the calculation of syndrome-like quantites and its complexity is comparable to the corresponding BCH codes; 3) the codes are systematic and encoding can be accomplished very simply by shift-registers in as many time units as are required by the serial transmission of the information digits.

The following sections are devoted to the description of the codes and to the demonstration of the properties stated above.

2. DESCRIPTION OF THE CODES[1]

In the sequel all polynomials considered belong to the algebra A_{n-1} of polynomials over GF(2) modulo $(x^{2^{n-1}-1} + 1)(n \geqq 4)$. Given $a(x) \in A_{n-1}$, $W[a(x)]$ denotes the number of nonzero coefficients of $a(x)$; given $b(x) \in A_{n-1}$, $d[a(x), b(x)] = W[a(x) + b(x)]$ is the Hamming distance between $a(x)$ and $b(x)$. By the symbol $a(x)$ we shall also denote the row vector $[a_{2^{n-1}-2}, a_{2^{n-1}-3}, \cdots, a_0]$ where $a(x) = \sum a_j x^j$.

Let $\{m(x)\}$ be a single-error-correcting BCH code of length $2^{n-1} - 1$, generated by $g_1(x)$, a primitive polynomial of degree $(n - 1)$; that is, if by α we denote a primitive element of $\mathrm{GF}(2^{n-1})$, $g_1(\alpha) = 0$. Consider now the code $\{s(x)\}$, whose generator polynomial has roots α, α^3 and 1: clearly $\{s(x)\}$ is a BCH code of minimum weight 6 (see Peterson (1961), p. 167) and $\{s(x)\} \subset \{m(x)\}$. Clearly $\{s(x)\}$ exists only for $2^{n-1} - 1 \geqq 2(n - 1) + 1$, i.e., for $n \geqq 4$; when $n = 4$, $s(x)$ is identically 0. Finally by $u(x)$ we denote the polynomial $(x^{2^{n-1}-1}+1)/(x + 1)$.

[1] There is some overlap between this section and (Preparata, 1968b) since this work is a conceptual and chronological generalization of the latter.

Given two polynomials $a(x)$ and $b(x)$, $(a(x) \in A_{n-1}, b(x) \in A_{n-1})$ and a binary parameter i, we construct $(2^n - 1)$-component vectors over GF(2) of the form

$$[a(x), i, b(x)].$$

Given $m(x) \in \{m(x)\}$, $s(x) \in \{s(x)\}$ and arbitrary i, we now set $a(x) = m(x)$ and $b(x) = m(x) + (m(1) + i)u(x) + s(x)$. We obtain

$$\mathbf{v} = [m(x), i, m(x) + (m(1) + i)u(x) + s(x)] \qquad (1)$$

We claim that

LEMMA 1. *The vectors* $\mathbf{v}$ *given by* (1) *form a linear code* $\mathcal{C}_n$.

Proof. The statement follows immediately from the verification that $\mathcal{C}_n$ is a group with respect to addition over GF(2). In fact: i) $\mathcal{C}_n$ contains the additive unity [0, 0, 0], obtained by setting in (1) $m(x) = 0$, $s(x) = 0$, $i = 0$; ii) $\mathcal{C}_n$ is closed with respect to addition, since both $\{m(x)\}$ and $\{s(x)\}$ are group codes.

Q.E.D.

LEMMA 2. *The minimum distance between any two code words of* $\mathcal{C}_n$ *is at least* 6.

Proof. Since $\mathcal{C}_n$ is a linear code its minimum distance coincides with the minimum weight W of its nonzero code words, which we now determine. Assume first that $m(x) = 0$. If also $i = 0$, then $W = W[s(x)] \geqq 6$. If $i = 1$, then $W = 1 + W[u(x) + s(x)] \geqq 1 + W[u(x)] - \max W[s(x)]$ We know that $W[u(x)] = 2^{n-1} - 1$ and that $\max W[s(x)]$ is $2^{n-1} - 6$ for $n > 4$ or is 0 for $n = 4$ (since $\max W[s(x)]$ is the maximum *even* weight of code words of the double-error-correcting BCH code); hence $W \geqq 1 + 2^{n-1} - 1 - 2^{n-1} + 6 = 6$.

Assume now that $m(x) \neq 0$. If $m(x) \notin \{s(x)\}$, then $m^*(x) = m^*(x) + (m(1) + i)u(x) + s(x) \neq 0$ and $m^*(x) \in \{m(x)\}$. It follows that $W \geqq W[m(x)] + W[m^*(x)] \geqq 3 + 3 = 6$, since both $m(x)$ and $m^*(x)$ are nonzero and $\{m(x)\}$ has minimum weight 3. If, alternatively, $m(x) \in \{s(x)\}$, then $W[m(x)] \geqq 6$ and $W \geqq W[m(x)] \geqq 6$. Q.E.D.

The number of information bits of $\mathcal{C}_n$ is readily obtained when one considers that the independently selectable $m(x)$, $s(x)$ and i contribute $(2^{n-1} - n)$, $(2^{n-1} - 2n)$ and 1 information bits, respectively. Therefore $\mathcal{C}_n$ is a $(2^n - 1, 2^n - 3n + 1)$ linear code of minimum distance 6.

Consider now the polynomial $\varphi(x) = (x^{2^{n-1}-1} + 1)/g_1(x)$, i.e., a minimum degree maximum length sequence of length $(2^{n-1} - 1)$. We first show that

LEMMA 3. *There exists an* $s(0 \leqq s \leqq 2^{n-1} - 2)$ *such that* $(x^s\varphi(x))^2 = x^s\varphi(x)$.

Proof. We compute the product $\varphi(x)\varphi(x)$. Since $\varphi(x)$ is not divided by $g_1(x)$, $\varphi^2(x)$ is not zero; moreover, $\varphi^2(x)$ belongs to the code generated by $\varphi(x)$, i.e.

$$\varphi^2(x) = x^r\varphi(x) \tag{2}$$

for some r, $0 \leqq r \leqq 2^{n-1} - 2$. If we multiply (2) by x^{2s} we have $x^{2s}\varphi^2(x) = x^{r+2s}\varphi(x)$, i.e. $(x^s\varphi(x))^2 = x^s\varphi(x)\cdot x^{r+s}$. The lemma follows if $x^{r+s} = 1$, i.e., if $r + s = 0 \pmod{2^{n-1} - 1}$, or, equivalently, $s = 2^{n-1} - 1 - r \bmod (2^{n-1} - 1)$. Q.E.D.

We define $f(x) \triangleq x^s\varphi(x)$.

A polynomial $q(x) = ax^j(a = 0, 1; j = 0, 1, \cdots, 2^{n-1} - 2)$ is clearly a minimum weight coset leader of $\{m(x)\}$ for $a = 1$. We now construct vectors $\mathbf{u}$ of the form

$$\mathbf{u} = [q(x), 0, q(x)f(x)]. \tag{3}$$

We have the following lemmas:

LEMMA 4. *The polynomial* $q(x) + q(x)f(x)$ *belongs to* $\{m(x)\}$.

Proof. The assertion follows immediately from Lemma 3, since

$$f(x)\{q(x) + q(x)f(x)\} = f(x)q(x) + f^2(x)q(x) = 0$$

i.e., $q(x) + q(x)f(x)$, being orthogonal to $f(x)$, is divided by $g_1(x)$. Q.E.D.

LEMMA 5. *The sum of two vectors* $\mathbf{u}_1$ *and* $\mathbf{u}_2$ *of the form* (3) *admits of the representation* $(n \geqq 4)$

$$\mathbf{u}_1 + \mathbf{u}_2 = \mathbf{v} + \mathbf{q} + \mathbf{p} \tag{4}$$

with

$$\mathbf{v} = [m'(x), 0, m'(x) + m'(1)u(x)], m'(x) \in \{m(x)\} \quad \text{i.e. } \mathbf{v} \in \mathcal{C}_n,$$

$$\mathbf{q} = [q(x), 0, q(x)], \tag{5}$$

$$\mathbf{p} = [0 \quad, 0, m''(x) \;], m''(x) \in \{m(x)\}. \tag{6}$$

If $q(x) = 0$, *then* $m'(x) = 0$; *if* $q(x) \neq 0$ *then either* $m'(x) = 0$ *or* $m'(x)$ *is a trinomial.*

Proof. Let $\mathbf{u}_1 = [q_1(x), 0, q_1(x)f(x)]$ and $\mathbf{u}_2 = [q_2(x), 0, q_2(x)f(x)]$. We have

$$\mathbf{u}_1 + \mathbf{u}_2 = [q_1(x) + q_2(x), 0, (q_1(x) + q_2(x))f(x)]. \tag{7}$$

Let $q(x)f(x) = (q_1(x) + q_2(x))f(x)$ and $m'(x) \triangleq q_1(x) + q_2(x) + q(x)$. Clearly, since $(q(x) + q_1(x) + q_2(x))f(x) = 0$, $m'(x) \in \{m(x)\}$. If $q_1(x) = q_2(x)$, it follows that $q(x) = 0$ and $m'(x) = 0$. If $q_1(x) \neq q_2(x)$, either $q(x) = q_i(x) (i = 1, 2)$ or $q_1(x)$, $q_2(x)$, $q(x)$ are nonzero and distinct: in the former case $m'(x) = 0$, in the latter $m'(x)$ is a trinomial.

This given we can write

$$q_1(x) + q_2(x) = m'(x) + q(x)$$

and rewrite (7) as

$$\begin{aligned}\mathbf{u}_1 + \mathbf{u}_2 = [m'(x), 0, m'(x) + m'(1)u(x)] + [q(x), 0, q(x)] \\ + [0, 0, q(x) + q(x)f(x) + m'(x) + m'(1)u(x)].\end{aligned}$$

It is now evident that $m''(x) \triangleq (q(x) + q(x)f(x)) + m'(x) + m'(1)u(x) \in \{m(x)\}$ since it is the sum of polynomials belonging to $\{m(x)\}$. Q.E.D.

LEMMA 6. *For any trinomial* $m(x) \in \{m(x)\}$

$$m(\alpha^3) = \alpha^{3s}(\alpha^h + \alpha^{2h})$$

where s and h are integers modulo $2^{n-1} - 1$, $h \neq 0$.

Proof. Let $m(x) = x^s + x^i + x^j$, with distinct s, i, j. Then $m(\alpha^3) = \alpha^{3s} + \alpha^{3i} + \alpha^{3j}$. Recalling that $m(\alpha) = \alpha^s + \alpha^i + \alpha^j = 0$ we have

$$\alpha^{3s} = (\alpha^i + \alpha^j)^3 = \alpha^{3i} + \alpha^{3j} + \alpha^i\alpha^j(\alpha^i + \alpha^j)$$

or equivalently $m(\alpha^3) = \alpha^i\alpha^j\alpha^s = \alpha^{3s}\alpha^{i-s}\alpha^{j-s}$. But $\alpha^{j-s} = 1 + \alpha^{i-s}$; hence, letting $i - s = h \neq 0$, the assertion is proved. Q.E.D.

Consider now the matrices:

$$\begin{aligned} H_1 &= [\alpha^{2^{n-1}-2}, \cdots, \alpha, 1], \\ H_3 &= [(\alpha^3)^{2^{n-1}-2}, \cdots, \alpha^3, 1], \\ U &= [1, \cdots, 1, 1]. \end{aligned}$$

The matrix $H = [H_1^T, H_3^T, U^T]^T$ is the parity check matrix of $\{s(x)\}$ (the superscript T denotes "transpose"). Given a polynomial $h(x)$, we define $h(x)H^T \triangleq [\beta_1, \beta_3, c]$ as the *characteristics* of $h(x)$. For arbitrary $s(x) \in \{s(x)\}$

$$W[h(x) + s(x)] \geqq W[k(x)]$$

where $k(x)$ is a minimum weight member of the coset of $\{s(x)\}$ to which $h(x)$ belongs. We now calculate the characteristics of some polynomials which we shall frequently use in the sequel:

$$\begin{cases} q(x) = \alpha^s & q(x)H^T = [\alpha^s, \alpha^{3s}, 1] \\ m(x) \in \{m(x)\} & m(x)H^T = [0, m(\alpha^3), m(1)] \\ m''(x) \text{ (see (6))} & m''(x)H^T = [0, m'(\alpha^3) + \alpha^{3s}, 1]. \end{cases} \tag{8}$$

The first relation is straightforward. The second follows from $m(\alpha) = 0$, since $m(x) \in \{m(x)\}$. To prove the last relation, recall that $m''(x) = q(x) + q(x)f(x) + m'(1)u(x) + m'(x)$, and that: $q(x)f(x)H^T = [\alpha^s, 0, 0]$, since $f(x)$ is divided by the minimum function $g_3(x)$ of α^3 and by $(x + 1)$; $m'(1)u(x)H^T = [0, 0, m'(1)]$, since $u(x)$ is divided by $g_3(x)$ and the minimum function $g_1(x)$ of α.

LEMMA 7. *For* $m''(x) = q(x) + q(x)f(x) + m'(1)u(x) + m'(x)$, *and arbitrary* $s(x) \in \{s(x)\}$,

$$W[m''(x) + q(x) + s(x)] \geqq \begin{cases} 4 \textit{ for even } n \\ 2 \textit{ for odd } n. \end{cases}$$

Proof. The characteristics of $(m''(x) + q(x))$ is $[\alpha^3, m'(\alpha^3), 0]$ (see (8)). $W[m''(x) + q(x) + s(x)]$ is the minimum number of columns of H which add to $[\alpha^3, m'(\alpha^3), 0]^T$. Since $c = 0$, this number is even. Let us assume that there are two elements x_1 and x_2 of GF(2^{n-1}) which satisfy the equations

$$\begin{cases} x_1 + x_2 = \alpha^s \\ x_1^3 + x_2^3 = m'(\alpha^3). \end{cases}$$

Since $\alpha^s \neq 0$ we make the substitution $y_1 = (x_1/\alpha^s)$, $y_2 = (x_2/\alpha^s)$. After easy manipulations we recognize that y_1 and y_2 are the solutions of the single equation

$$y^2 + y + 1 + m'(\alpha^3)/\alpha^{3s} = 0.$$

Since either $m'(x) = 0$ or $m'(x)$ is a trinomial (Lemma 5), Lemma 6 yields $m'(\alpha^3)/\alpha^{3s} = \alpha^h + \alpha^{2h}$ $(0 \leqq h \leqq 2^{n-1}-2)$ and the previous equation becomes

$$(y + \alpha^h)^2 + (y + \alpha^h) + 1 = 0$$

or, equivalently, letting $y + \alpha^h = z$

$$z^2 + z + 1 = 0. \tag{9}$$

But solutions of (9) are primitive cube roots of unity, whence (9) has solutions in $\mathrm{GF}(2^{n-1})$ only for odd n. Q.E.D.

We now construct $(2^n - 1)$-components vectors of the form

$$\mathbf{w} = [m(x) + q(x), i, m(x) + q(x)f(x) \\ + (m(1) + i)u(x) + s(x)] \tag{10}$$

where $m(x)$, $q(x)$, i, $s(x)$ are independently chosen and contribute $(2^{n-1} - n)$, $(n - 1)$, 1, $(2^{n-1} - 2n)$ information bits respectively, for a total of $(2^n - 2n)$ information bits. The vectors $\mathbf{w}$ form a $(2^n - 1, 2^n - 2n)$ code $\mathcal{K}_n$: the generic vector $\mathbf{w}$ can be decomposed as

$$\mathbf{w} = \mathbf{v} + \mathbf{u} \tag{11}$$

where $\mathbf{v}$ and $\mathbf{u}$ are defined by relations (1) and (3), respectively. Let $\mathbf{w}_1 = \mathbf{v}_1 + \mathbf{u}_1$ and $\mathbf{w}_2 = \mathbf{v}_2 + \mathbf{u}_2$ be two distinct code words of $\mathcal{K}_n$. Using relations (4) (Lemma 5) we have

$$\mathbf{w}_1 + \mathbf{w}_2 = (\mathbf{v}_1 + \mathbf{v}_2) + (\mathbf{u}_1 + \mathbf{u}_2) = \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v} + \mathbf{q} + \mathbf{p}$$

or

$$\mathbf{w}_1 + \mathbf{w}_2 = \mathbf{v}' + \mathbf{q} + \mathbf{p} \tag{12}$$

where $\mathbf{v}' \triangleq \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}$. Clearly $\mathbf{v}'$ is an arbitrary member of $\mathcal{C}_n$, but $\mathbf{q} + \mathbf{p}$ can be decomposed as

$$\begin{aligned} \mathbf{q} + \mathbf{p} &= [q(x), 0, q(x)f(x) + m'(x) + m'(1)u(x)] \\ &= [q(x), 0, q(x)f(x)] + [0, 0, m'(x) + m'(1)u(x)] \\ &= \mathbf{u}' + [0, 0, m'(x) + m'(1)u(x)]. \end{aligned}$$

When $m'(x) \neq 0$, we recall that $m'(x) \notin \{s(x)\}$ (Lemma 5), that is $[0, 0, m'(x) + m'(1)u(x)] \notin \mathcal{C}_n$: hence $\mathcal{K}_n$ is a nonlinear code. Furthermore, in (11) each nonzero $\mathbf{u}$ identifies a coset of $\mathcal{C}_n$, since $q(x) \neq 0$ identifies a coset of $\{m(x)\}$. Hence $\mathcal{K}_n$ can be seen as the set union of $\mathcal{C}_n$ and of a subset of its cosets, whose cardinality is $2^{n-1} - 1$.

Let W denote the weight of $(\mathrm{w}_1 + \mathrm{w}_2)$. We can now prove the central result of this paper.

THEOREM 1. *For even* $n \geqq 4$, $\mathcal{K}_n$ *is a nonlinear* $(2^n - 1, 2^n - 2n)$ *code of minimum distance* 5.

Proof. If $q(x) = 0$, $\mathrm{w}_1 + \mathrm{w}_2 \in \mathcal{C}_n$ and, by Lemma 2, $W \geqq 6$. Assume now that $q(x) = x^s$ $(0 \leqq s \leqq 2^{n-1} - 2)$. In general, W is given by

$$\begin{aligned} W = {} & i + W[m(x) + q(x)] \\ & + W[m(x) + (m(1) + i)u(x) + s(x) + m''(x) + q(x)]. \end{aligned} \tag{13}$$

Depending upon the values of $m(1)$ and i we distinguish three cases:

1) $m(1) = 0, i = 1$. Relation (13) becomes

$$\begin{aligned} W & \geqq 1 + W[m(x) + q(x)] + W[u(x) + s(x) + m''(x)] \\ & \qquad - W[m(x) + q(x)] \\ & = 1 + W[u(x) + s(x) + m''(x)]. \end{aligned}$$

From relations (8) and $u(x)H^T = [0, 0, 1]$, we obtain $(u(x) + m''(x))H^T = [\beta_1, \beta_3, c] = [0, m'(\alpha^3) + \alpha^{3s}, 0]$. But by Lemma 5 and Lemma 6 $m'(\alpha^3) = \alpha^{3s}(\alpha^k + \alpha^{2k})$, $0 \leqq k \leqq 2^{n-1} - 2$. Hence, $m'(\alpha^3) + \alpha^{3s} = \alpha^{3s}(1 + \alpha^k + \alpha^{2k}) \neq 0$ in $\mathrm{GF}(2^{n-1})$, n even: it follows that $W[u(x) + s(x) + m''(x)] \neq 0$. Furthermore, $W[u(x) + s(x) + m''(x)]$ is even and >3, since $c = 0$ and $\beta_1 = 0$ (H_1 is the parity check matrix of a single-error-correcting code). We conclude that $W \geqq 1 + 4 = 5$.

2) $m(1) = 0, i = 0$. If $m(x) \neq 0$, relation (13) yields

$$W \geqq W[m(x)] + W[m(x) + m''(x) + s(x)] - 2W[q(x)].$$

Relations (8) give $(m(x) + m''(x))H^T = [\beta_1, \beta_3, c] = [0, m(\alpha^3) + m'(\alpha^3) + \alpha^{3s}, 1]$, that is, $W[m(x) + m''(x) + s(x)]$ is odd $(c = 1)$ and $\geqq 3$ $(\beta_1 = 0)$. Furthermore, $m(x) \neq 0$ and $m(1) = 0$ imply $W[m(x)] \geqq 4$, whence $W \geqq 4 + 3 - 2 = 5$. If $m(x) = 0$ relation (13) becomes

$$W \geqq W[q(x)] + W[m''(x) + q(x) + s(x)].$$

From Lemma 7, we have that $W[m''(x) + q(x) + s(x)] \geqq 4$ for n even, whence $W \geqq 1 + 4 = 5$.

3) $m(1) = 1$. In this case $W[m(x) + q(x)]$ is even and $\geqq 2$. Assume at first that $W[m(x) + q(x)] = 2$: this implies that $m(x) = x^s + x^i + x^j$ $(s, i, j$ distinct), whence by Lemma 6 $m(\alpha^3) = \alpha^{3s}(\alpha^h + \alpha^{2h})$ for some h, $1 \leqq h \leqq 2^{n-1} - 2$. Relation (13) yields

$$W = i + 2 + W[m(x) + (1 + i)\, u(x) + m''(x) + q(x) + s(x)].$$

For simplicity we let $k(x) \triangleq m(x) + (1 + i)u(x) + m''(x) + q(x) + s(x)$. With the help of relations (8) the characteristics of $k(x)$ is readily obtained as

$$k(x)H^T = [\beta_1, \beta_3, c] = [\alpha^s, m(\alpha^3) + m'(\alpha^3), i].$$

Since $m'(\alpha^3) = \alpha^{3s}(\alpha^k + \alpha^{2k})(0 \leqq k \leqq 2^{n-1} - 2)$, it follows that $m(\alpha^3) + m'(\alpha^3) = \alpha^{3s}(\alpha^r + \alpha^{2r})$ with $r = h + k$. Therefore, from Lemma 7, $W[k(x)] > 2$ for even n, that is, $W[k(x)] \geqq 4 - i$ (since i is the parity of $W[k(x)]$). We conclude that

$$W \geqq i + 2 + 4 - i = 6.$$

Finally assume that $W[m(x) + q(x)] \geqq 4$. Since $\beta_1 = \alpha^s$, we obtain $W[k(x)] \geqq 1$, whence $W \geqq i + 4 + 1 = i + 5$.

Q.E.D.

Note. It is interesting to consider the problem of extending the method employed for the construction of $\mathcal{K}_n$ to other values of the number of correctable errors, namely to $t = 1$ or to $t > 2$.

Two distinct schemes appear to be candidates for successful generalizations. Consider again relation (10) which describes the double-error-correcting $\mathcal{K}_n$, i.e.,

$$\mathrm{w} = [m(x) + q(x), i, m(x) + (m(1) + i)u(x) + s(x) + q(x)f(x)].$$

Here $\mathcal{K}_n$ is constructed in terms of two codes, i.e., $\{m(x)\}$ and $\{s(x)\}$, with $\{s(x)\} \subset \{m(x)\}$. Specifically, if α is primitive in GF(2^{n-1}), then $\{m(x)\}$ is characterized by the root α, and $\{s(x)\}$ by the roots $1, \alpha, \alpha^3$. Therefore two potential generalizations for t-error-correction are:

A. $\{m(x)\}$ has root α, and $\{s(x)\}$ has roots $1, \alpha, \alpha^3, \cdots, \alpha^{2t-1}$.

B. $\{m(x)\}$ has roots $\alpha, \alpha^3, \cdots, \alpha^{2t-3}$, and $\{s(x)\}$ has roots $1, \alpha, \alpha^3, \cdots, \alpha^{2t-1}$.

For $t = 1$, both schemes are successful and generate the same codes, as can be easily shown. Specifically with scheme B, $m(x)$ is the generic member of A_{n-1}, and $\{s(x)\}$ has $(x + 1)g_1(x)$ as its generator, which gives the code $\mathcal{K}_n^{(1)}$

$$\mathrm{w}_B = [m(x), i, m(x) + (m(1) + i)u(x) + s(x)]. \tag{14}$$

Surprisingly, $\mathcal{K}_n^{(1)}$ is a group code, as is apparent from (14). Moreover, it can be shown that it coincides with a Vasil'ev code (1962). In fact (14) can be expressed as

$$\mathrm{w}_B = [m(x), i, m(x) + p(x)]$$

where

$$p(x) = (m(1) + i)u(x) + s(x).$$

If we now impose the condition that $p(x)$ belong to the code generated by $g_1(x)$, this relation becomes an equation in the unknowns $s(x)$ and i, which can always be solved if

$$i = \text{parity } W[m(x)] + \text{parity } W[p(x)]$$

thereby yielding a linear Vasil'ev code (equivalent to a Hamming code).

For $t > 2$, the question whether either of the two outlined schemes produces a viable generalization remains entirely open.

3. THE FORM OF THE REDUNDANCY FUNCTIONS

Consider the expression (10) of the generic vector of $\mathcal{K}_n$, that is

$$\mathbf{w} = [m(x) + q(x), i, m(x) + (m(1) + i)u(x) + s(x) + q(x)f(x)].$$

It is easily seen that $\mathcal{K}_n$ can be encoded as a systematic code, i.e., $(2^n - 2n)$ binary information digits can be arbitrarily assigned in fixed positions and the remaining $(2n - 1)$ redundant digits can be computed as functions of the information digits. In this section we investigate the nature of these functions. For convenience, we now represent w as

$$\mathbf{w} = [i^{(0)}_{2^{n-1}-2}, \cdots, i^{(0)}_1, i^{((0)}_0, i, i^{(1)}_{2^{n-1}-2}, \cdots, i^{(1)}_{2n-1}, p_{2n-2}, \cdots, p_1, p_0]$$

where i's and p's denote information and redundancy digits, respectively.

Assume for a moment that $s(x) = 0$. Then the leftmost 2^{n-1} digits $[i^{(0)}_{2^{n-1}-2}, \cdots, i]$ completely determine the $2^{n-1} - 1$ rightmost ones; we denote the latter ones by $[\varphi_{2^{n-1}-2}, \cdots, \varphi_0]$, and analyze their dependence upon the former set. Let

$$i(x) \triangleq \sum i^{(0)}_j x^j, \qquad q(x)f(x) \triangleq c(x) = \sum c_j x^j, \qquad f(x) = \sum f_j x^j,$$
$$m(x) = \sum m_j x^j$$

where all summations run for $j = 0, 1, \cdots, 2^{n-1} - 2$. If $q(x) = 0$, then $c_j = 0$ for every j. If $q(x) = x^s$, then due to the unique property of the maximum length sequence (see Peterson (1961), p. 148), $c_{s+b}c_{s+b-1} \cdots c_{s+b-n+2} = 1$ and $c_{j+b} \cdots c_{j+b-n+2} = 0$ for $j \neq s$, where b is such that $f_b f_{b-1} \cdots f_{b-n+2} = 1$. We readily have

$$q(x) = \sum c_{j+b} \cdots c_{j+b-n+2} x^j, \qquad m_j = i^{(0)}_j + c_{j+b} \cdots c_{j+b-n+2}$$

and

$$\varphi_j = i^{(0)}_j + c_{j+b} \cdots c_{j+b-n+2} + i + \sum_k (i^{(0)}_k + c_{k+b} \cdots c_{k+b-n+2}) + c_j$$

or, after regrouping the terms

$$\varphi_j = \Big\{\sum_{k\neq j} i_k^{(0)} + i + c_j\Big\} + \Big\{\sum_{h\neq j+b} c_h \cdots c_{h-n+2}\Big\}. \tag{15}$$

We now recall that, since $c(x) = q(x)f(x) = i(x)f(x)$, $c_j = \sum f_{j-k} i_k^{(0)}$ is a linear function of the variables $i_0^{(0)}, i_1^{(0)}, \cdots, i_{2^{n-1}-2}^{(0)}$. Specifically, since $f(x)$ is a maximum length sequence, for distinct r and s there is a t such that $c_r + c_s = c_t$. If $g_1(x)$ is a trinomial, for $s = (r + n - 1)$, t satisfies the relations $r < t < r + n - 1$. Hence

$$c_h c_{h-1} \cdots c_{h-n+2} + c_{h-1} c_h \cdots c_{h-n+1}$$
$$= c_{h-1} \cdots c_{h-n+2}(c_h + c_{h-n+1}) = c_{h-1} \cdots c_{h-n+2}\,.$$

It follows that in the last term of (15), which is the sum of $2^{n-1} - 2$ products, each pair of consecutive products of $(n - 1)$ factors is contracted into a single product of $(n - 2)$ factors, for a total of $2^{n-2} - 1$ products. In conclusion we obtain

$$\varphi_j = \left\{i + i_j^{(0)} + \sum_{k=0}^{2^{n-1}-2} (1 + f_{j-k}) i_k^{(0)}\right\} + \sum_{h=0}^{2^{n-2}-2} \prod_{s=0}^{n-3} c_{j+b+1+2h-s} \tag{16}$$

which shows that $\varphi_j(i_{2^{n-1}-2}^{(0)}, \cdots, i_0^{(0)}, i)$ is the sum of a strictly linear function and of a nonlinear function of degree at most $(n - 2)$. As a check, for the (15, 8) code, $n = 4$, the latter function is quadratic.

This given, let h_{ij} be the generic entry of the parity check matrix H^* of $\{s(x)\}$ in systematic form, i.e., the $(2n - 1)$ rightmost columns of H^* form the unity matrix and the index j runs from right to left. Then the relations

$$p_i = \varphi_i + \sum_{j=2n-1}^{2^{n-1}-2} h_{ij}(i_j^{(1)} + \varphi_j) \qquad (i = 0, 1, \cdots, 2n - 2) \tag{17}$$

give the sought redundancy functions.

Expressions (15) and (17) are suggestive of a very simple implementation of encoding. In fact, φ_j is a cyclic function of its arguments. Hence it can be realized by a recirculating nonlinear convolutional encoder consisting of a cyclic shift register and of a combinational circuit realizing $\varphi = \varphi_{2^{n-1}-2}$ (see Figure 1). The complete encoder consists of three shift-registers SR1, SR2, SR3 with 1, $(2^{n-1} - 1)$ and $(2n - 1)$ stages, respectively. The operation is organized in four phases G_1, G_2, G_3, G_4, whose durations are 1, $(2^{n-1} - 1)$, $(2^{n-1} - 2n)$ and $(2n - 1)$ time

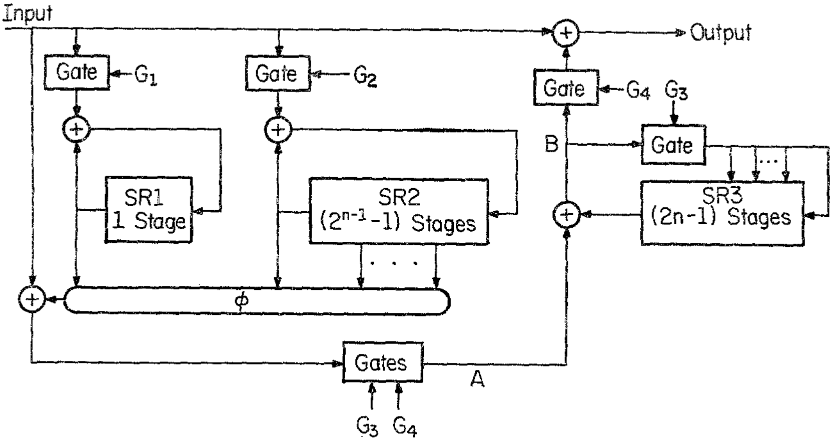

FIG. 1. Encoder for the $\mathcal{K}_n$ code.

units, respectively. The indicated gates are permissive when the applied signals are active. All registers are initially set to 0. The information digits are fed in the sequence i, $\dot{i}^{(0)}_{2^{n-1}-2}$, $\cdots$, $\dot{i}^{(0)}_0$, $\dot{i}^{(1)}_{2^{n-1}-2}$, $\cdots$, $\dot{i}^{(1)}_{2n-1}$, one per time unit. Then during G_1 the digit i is fed to SR1 and during G_2 $\dot{i}^{(0)}_{2^{n-1}-2}$, $\cdots$, $\dot{i}^{(0)}_0$ are fed to SR2 (while they are concurrently sent to the output): both SR1 and SR2 are recirculating, as shown. During phase G_3, $\varphi_j + \dot{i}^{(1)}_j$ appears at point A to be fed to SR3, which is a feedback shift register performing the division of a polynomial by $(x + 1)g_1(x) \cdot g_3(x)$ (see Peterson (1961), p. 149); then at the end of G_3 SR3 contains the parity checks $\sum_j h_{ij}(\dot{i}^{(1)}_j + \varphi_j)$. During G4 the input is 0 and at point B the functions p_i are formed and fed to the output. Therefore the calculation of the redundant digits takes no longer than the serial transmission of the information digits.

4. OPTIMALITY OF THE $\mathcal{K}_n$ CODES

A code $\mathcal{K}_n$ is a $(2^n - 1, 2^n - 2n)$ double-error-correcting code. It contains one information digit more than the corresponding linear code, i.e., the BCH double-error-correcting code of the same length (which is a $(2^n - 1, 2^n - 1 - 2n)$ code).

In this section we prove a stronger statement, namely, that a $\mathcal{K}_n$ code has the largest number of code words for its length and minimum distance, since it meets the Johnson's bound $A(N, d)$ (Johnson, 1962) for

$N = 2^n - 1$ (n even) and $d = 5$.[2] In fact the Johnson's bound for $d = 2t + 1$ is given by

$$A(N, d) \leqq \frac{2^N}{\sum_{j=0}^{t} \binom{N}{j} + \dfrac{\binom{N}{t+1} - \binom{d}{t} R(N, d, t)}{\left[\dfrac{N}{t+1}\right]}} \tag{18}$$

where $[a]$ is the integral part of a, and $R(N, d, t)$ satisfies the upper bound

$$R(N, d, t) \leqq \left[\frac{N}{d}\left[\frac{N-1}{d-1}\left[\cdots\left[\frac{N-t}{d-t}\right]\cdots\right]\right]\right]. \tag{19}$$

When $N = 2^n - 1$ and $t = 2$, relations (18) and (19) specialize as

$$A(2^n - 1, 5) \leqq \frac{2^{2^n-1}}{1 + \binom{2^n-1}{1} + \binom{2^n-1}{2} + \dfrac{\binom{2^n-1}{3} - \binom{5}{2} R(2^n - 1, 5, 2)}{\left[\dfrac{2^n-1}{3}\right]}} \tag{20}$$

$$R(2^n - 1, 5, 2) \leqq \left[\frac{2^n-1}{5}\left[\frac{2^n-2}{4}\left[\frac{2^n-3}{3}\right]\right]\right]. \tag{21}$$

Consider relation (21). For even n, $(2^n - 4)$ is divisible by 3, hence $[(2^n - 3)/3] = (2^n - 4)/3$. Moreover $(2^n - 4)$ is divisible by 4. We must now show that $(2^n - 1)(2^n - 2)(2^{n-2} - 1)$ is divisible by 5. This follows immediately from the observation that the residues modulo 5 of 2^n (n even) alternate as 1 and 4, i.e., the residue of $(2^n - 1)$ alternate as 0 and 3: since $(2^n - 1)(2^n - 2)(2^{n-2} - 1)$ contains two consecutive even powers of 2, we have

$$R(2^n - 1, 5, 2) \leqq \frac{(2^n - 1)(2^n - 2)(2^n - 4)}{60}$$

[2] The observation that $A(2^n - 1, 5)$ (n even) is a power of 2 is originally due to J. P. Robinson. Prior to this, the author formulated a conjecture, based on rather fuzzy geometric arguments, that nonlinear codes of length $(2^n - 1)$ and distance 5, analogous to the (15, 8) code, existed only for even n (private communications, Jan. and March 1968).

from which we readily obtain for even n

$$\frac{\binom{2^n-1}{3} - \binom{5}{2} R(2^n-1, 5, 2)}{\left[\frac{2^n-1}{3}\right]} \geqq 3 \frac{\frac{(2^n-1)(2^n-2)}{6}}{(2^n-1)} = 2^{n-1} - 1. \qquad (22)$$

We then conclude that

$$A(2^n-1, 5) \leqq \frac{2^{2^n-1}}{2^{2n-1} - 2^{n-1} + 1 + (2^{n-1} - 1)} = 2^{2^n-2n} \qquad (n \text{ even})$$

which is exactly the number of code words of $\mathcal{K}_n$. Clearly for odd n, ratio (22) is strictly larger than $2^{n-1} - 1$, since $[(2^n-3)/3] = (2^n-5)/3$: which also shows, from a different angle, the unrealizability of $\mathcal{K}_n$ codes for odd n.

5. DECODING OF A $\mathcal{K}_n$ CODE

In this section we show that decoding of a $\mathcal{K}_n$ code can be easily accomplished through the calculation and examination of syndrome-like quantities.

With the vector

$$\mathbf{e} = [e_0(x), e, e_1(x)]$$

we represent an error pattern, where $e_j(x) \in A_{n-1}$ and e is a binary parameter. The distance properties of $\mathcal{K}_n$ give the following condition on $\mathbf{e}$ for correctability

$$W[e_0(x)] + W[e_1(x)] + e \leqq 2. \qquad (23)$$

In general the received vector is $\mathbf{r} = [r_0(x), r, r_1(x)] = \mathbf{w} + \mathbf{e}$, with $\mathbf{w} \in \mathcal{K}_n$. We now compute the following functions:

$$\begin{cases} \sigma_0 \triangleq r_0(x) H_1^T, \\ \sigma_1 \triangleq r_1(x) H_1^T, \\ \sigma \triangleq (r_0(x) + r_1(x)) H_3^T, \\ d \triangleq r + r_1(x) U^T. \end{cases}$$

Since $r_0(x) = m(x) + q(x) + e_0(x)$, and $m(x)H_1^T = 0$, letting $q(x) = bx^s$, we have $\sigma_0 = b\alpha^s + e_0(\alpha)$. Similarly, from $r_1(x) = m(x) +$

$(m(1) + i)u(x) + q(x)f(x) + s(x) + e_1(x)$ and $u(x)H_1^T = 0$, $s(x)H_1^T = 0$, $q(x)f(x)H_1^T = b\alpha^s$ we obtain $\sigma_1 = b\alpha^s + e_1(\alpha)$. From $r_0(x) + r_1(x) = q(x) + q(x)f(x) + s(x) + e_0(x) + e_1(x) + (m(1) + i)u(x)$, recalling that $s(x)H_3^T = 0, f(x)H_3^T = 0, u(x)H_3^T = 0$, we obtain $\sigma = b\alpha^{3s} + e_1(\alpha^3) + e_0(\alpha^3)$. Finally since $W[q(x)f(x)]$, $W[s(x)]$, $W[m(x) + m(1)u(x)]$ are even, $d = r + i + e_1(1) = e + e_1(1)$. This is summarized as follows:

$$\begin{cases} \sigma_0 = b\alpha^s + e_0(\alpha), \\ \sigma_1 = b\alpha^s + e_1(\alpha), \\ \sigma \ \ = b\alpha^{3s} + e_0(\alpha^3) + e_1(\alpha^3), \\ d \ \ = e + e_1(1). \end{cases} \tag{24}$$

The quadruple $\Sigma \equiv (\sigma_0, \sigma_1, \sigma, d)$ is conventionally termed the *syndrome* of r.

We now give a lemma which is based on rather well-known results of the theory of finite fields.[3]

LEMMA 8. *The set Θ of all $\theta \in \mathrm{GF}(2^{n-1})$ for which $y^2 + y + \theta = 0$ has solutions over $\mathrm{GF}(2^{n-1})$ is a vector space of dimension $(n - 2)$, given by the even linear combinations of a normal basis $\beta, \beta^2, \beta^4, \cdots, \beta^{2^{n-2}}$ of $\mathrm{GF}(2^{n-1})$.*

Proof. It is well-known (see, e.g., Albert (1956) p. 121) that there are bases of $\mathrm{GF}(2^{n-1})$ consisting of complete sets of conjugates (normal basis): let $\beta, \beta^2, \cdots, \beta^{2^{n-2}}$ be one such set of linearly independent conjugates. Then every $\gamma \in \mathrm{GF}(2^{n-1})$ is uniquely expressible as

$$\gamma = c_0\beta + c_1\beta^2 + \cdots + c_{n-2}\beta^{2^{n-2}} \quad (c_j \in \mathrm{GF}(2)).$$

Since

$$\beta^{2^{n-1}} = \beta, \quad \text{then} \quad \gamma^2 = c_{n-2}\beta + c_0\beta^2 + \cdots + c_{n-3}\beta^{2^{n-2}}$$

and

$$\gamma^2 + \gamma = d_0\beta + d_1\beta^2 + \cdots + d_{n-2}\beta^{2^{n-2}} \quad (d_j \in \mathrm{GF}(2)) \tag{25}$$

with $d_j = c_j + c_{j-1}$ (the subscripts are modulo $n - 1$). But the right side of (25) is the generic element of Θ: assume then that $d_0, d_1, \cdots, d_{n-2}$ are given. We then have $c_{n-2} = d_0 + c_0 = d_0 + d_1 + c_1 = \cdots =$

[3] The argument given here is substantially borrowed from Albert (1956). A very similar theorem was proved by Berlekamp et al. (1962, Thm. 1). A particularly illuminating reference is Berlekamp (1968), which also contains a generalization of the lemma (p. 166). Since the statement given here is particularly geared to subsequent considerations, lemma and proof are given in full.

$d_0 + \cdots + d_{n-2} + c_{n-2}$, i.e.,

$$d_0 + d_1 + \cdots d_{n-2} = 0$$

i.e., the number of nonzero d_j's is even. Q.E.D.

This lemma provides a rule for testing whether $\gamma \in \mathrm{GF}(2^{n-1})$ is a member of Θ. In fact, we must first find a normal basis $\beta, \beta^2, \cdots, \beta^{2^{n-2}}$ of $\mathrm{GF}(2^{n-1})$, (see, e.g., Berlekamp (1968), pp. 253–254). Let $\boldsymbol{\gamma}$ denote the column vector representation over GF(2) of $\gamma \in \mathrm{GF}(2^{n-1})$ with respect to the basis $1, \alpha, \cdots, \alpha^{n-2}$, and let $M \triangleq [\boldsymbol{\beta}, \cdots, \boldsymbol{\beta}^{2^{n-2}}]$, a nonsingular $(n-1) \times (n-1)$ matrix. Then $\boldsymbol{\gamma}$ is related to the representation $[d_0, \cdots, d_{n-2}]$ of γ with respect to the basis $\beta, \beta^2, \cdots, \beta^{2^{n-2}}$ by

$$\boldsymbol{\gamma} = M\,[d_0, \cdots, d_{n-2}]^T$$

i.e., $M^{-1}\boldsymbol{\gamma} = [d_0, \cdots, d_{n-2}]^T$. Premultiplying both sides by the row vector $\mathbf{u} = [1, 1, \cdots, 1]$ we have the condition

$$\mathbf{u} \cdot [d_0, \cdots, d_{n-2}]^T = \begin{cases} 0 & \text{if} \quad \gamma \in \Theta \\ 1 & \text{if} \quad \gamma \notin \Theta \end{cases}$$

which, denoting by $\boldsymbol{\lambda}^T$ the row sum of M^{-1}, is translated into

$$\boldsymbol{\lambda}^T\boldsymbol{\gamma} = \begin{cases} 0 & \text{if} \quad \gamma \in \Theta \\ 1 & \text{if} \quad \gamma \notin \Theta. \end{cases} \tag{26}$$

The following lemma provides some insight into the distance relationship between the generic vector $\mathbf{r}$ and the members $\mathbf{w}$ of $\mathcal{K}_n$.

LEMMA 9. *Given any vector* $\mathbf{r} = [r_0(x), r, r_1(x)]$ *there exists a* $\mathbf{w} \in \mathcal{K}_n$ *such that* $\mathbf{r} + \mathbf{w} = [0, e, e(x)]$ with $W[e(x)] \leqq 3$.

Proof. Let $\{t(x)\}$ be the double-error-correcting BCH code generated by $g_1(x)g_3(x)$. We decompose $r_0(x)$ as $r_0(x) = m_0(x) + q_0(x)$ and form $r_1^*(x) = r_1(x) + m_0(x) + (m_0(1) + r)u(x) + q_0(x)f(x)$. Next $r_1^*(x)$ is decomposed as $r_1^*(x) = t(x) + e(x)$, where $t(x) \in \{t(x)\}$ and $e(x)$ is a minimum weight coset leader of $\{t(x)\}$: it is known (Gorenstein, et al. (1960)) that $W[e(x)] \leqq 3$. It is also of immediate verification that $(t(x) + t(1)u(x)) \in \{s(x)\}$. We then form the code word

$$\begin{aligned} \mathbf{w} = [m_0(x) + q_0(x), r + t(1), m_0(x) + (m_0(1) \\ + r + t(1))u(x) + q_0(x)f(x) + t(x) + t(1)u(x)] \\ = [r_0(x), r + t(1), r_1(x) + e(x)] \end{aligned}$$

Letting $t(1) = e$, $\mathbf{r} + \mathbf{w} = [0, e, e(x)]$. Q.E.D.

Hereafter the subscript j of σ_j or $e_j(x)$ is to be considered modulo 2. We define $\rho \triangleq \sigma + (\sigma_0 + \sigma_1)^3$ and prove the following basic Lemma.

LEMMA 10. *The conditions* $\rho + \sigma_j^3 = 0$ ($j = 0$ *and* 1), $d = 0$ *hold if and only if* $\mathbf{r} \in \mathcal{K}_n$, i.e., *they characterize the code* $\mathcal{K}_n$.

Proof. From Lemma 9, we can assume without loss of generality that the discrepancy between $\mathbf{r}$ and some $\mathbf{w} \in \mathcal{K}_n$ be of the form $[0, e, e(x)]$, $W[e(x)] \leqq 3$. Then relations (24) become

$$\begin{cases} \sigma_0 = b\alpha^s, \\ \sigma_1 = b\alpha^s + e(\alpha), \\ \sigma = b\alpha^{3s} + e(\alpha^3), \\ d = e + e(1). \end{cases} \tag{27}$$

The direct statement follows immediately by setting $e(x) = 0$, $e = 0$ in (27). To prove the converse, assume that $\rho + \sigma_j^3 = 0$ ($j = 0, 1$). This implies $\sigma_0^3 = \sigma_1^3$, and, due to the uniqueness of the cubic root in $\mathrm{GF}(2^{n-1})$, n even, $\sigma_0 = \sigma_1$. From (27) it follows that $e(\alpha) = 0$. We then have: $\rho + \sigma_j^3 = \sigma + \sigma_j^3 = e(\alpha^3) = 0$. Since $[H_1^T, H_3^T]$ is the parity check matrix of a double-error-correcting code, and $W[e(x)] \leqq 3$ (Lemma 9), from $e(\alpha) = e(\alpha^3) = 0$ we conclude that $e(x) = 0$. Finally $d = 0$ yields $e = e(1) = 0$. Q.E.D.

We readily recognize that $\rho + \sigma_j^3 = 0$ ($j = 0, 1$), $d = 0$ are equivalent to

$$\sigma_0 = \sigma_1, \qquad \sigma = \sigma_0^3 = \sigma_1^3, \qquad d = 0 \tag{28}$$

which characterize the code.

Following is a sequence of three theorems (2.1, 2.2 and 2.3) which establish a correspondence between sets of syndromes and sets of correctable error configurations. The statements and the relative proofs follow an almost identical pattern. The necessary condition ("if") is demonstrated by showing through relations (24) that an error configuration of the prescribed type produces a syndrome of the prescribed type. The converse ("only if") is demonstrated as follows: we form a "correction" vector $\mathbf{c} = [c_0(x), c, c_1(x)]$ which is a function of the syndrome Σ alone and such that $c + W[c_0(x)] + W[c_1(x)] \leqq 2$; then we show that $\mathbf{r} + \mathbf{c} \in \mathcal{K}_n$, since the syndrome $\Sigma^* \equiv (\sigma_0^*, \sigma_1^*, \sigma^*, d^*)$ calculated for $(\mathbf{r} + \mathbf{c})$, that is

$$\begin{cases} \sigma_j^* = \sigma_j + c_j(\alpha) & (j = 0, 1), \\ \sigma^* = \sigma + c_0(\alpha^3) + c_1(\alpha^3), \\ d^* = d + c + c_1(1) \end{cases} \tag{29}$$

satisfies the conditions of Lemma 10 (or, equivalently, (28)); finally, due to the distance properties of $\mathcal{K}_n$, we conclude that $\mathbf{e} = \mathbf{c}$ is the only correctable error configuration which could have produced $\mathbf{r}$. Clearly each of these theorems also yields a decoding rule, embodied by the calculation of the vector $\mathbf{c}$ from Σ. After this introduction the proof of each theorem will be simply sketched.

THEOREM 2.1. *For correctable* $\mathbf{e}$ *the condition* $\rho + \sigma_j^3 = 0$ *is verified for exactly one value of* $j = 0, 1$ *if and only if* $W[e_0(x)] + W[e_1(x)] = 1$.

Proof. "*If*": $e_j(x) = x^k$, $e_{j+1}(x) = 0$ give $\sigma_j = b\alpha^s + \alpha^k$, $\sigma_{j+1} = b\alpha^s$, $\sigma = b\alpha^{3s} + \alpha^{3k}$, whence $\rho = b\alpha^{3s}$. Then $\rho + \sigma_{j+1}^3 = 0$ and

$$\rho + \sigma_j^3 = \alpha^{3k}[(b\alpha^{s-k})^2 + b\alpha^{s-k} + 1] \neq 0$$

since $z^2 + z + 1 \neq 0$ for any value of $z \in \mathrm{GF}(2^{n-1})$, n even.

"*Only if*": If $\rho + \sigma_j^3 \neq 0$, $\rho + \sigma_{j+1}^3 = 0$, we calculate $\sigma_0 + \sigma_1 = \alpha^h$; then we set $c_j(x) = x^h$, $c_{j+1}(x) = 0$, $c = d + c_1(1)$ and compute Σ^*, i.e.,

$$\begin{aligned} \sigma_j^* &= \sigma_j + \alpha^h = \sigma_{j+1} = \sigma_{j+1}^*, \qquad d^* = 0, \\ \sigma^* &= \sigma + \alpha^{3h} = \sigma + (\sigma_0 + \sigma_1)^3 = \rho = \sigma_{j+1}^3 = \sigma_{j+1}^{*3} \end{aligned}$$

i.e., (28) are satisfied with $c + W[c_0(x)] + W[c_1(x)] \leqq 2$. Q.E.D.

Theorem 2.1 yields the following decoding rule:

Rule 1. If $\sqrt[3]{\rho} = \sigma_j$, $\sqrt[3]{\rho} \neq \sigma_{j+1}$, then $c_{j+1}(x) = x^h$ and $c = d + c_1(1)$ where $\alpha^h = \sigma_0 + \sigma_1$.

THEOREM 2.2. *For correctable* $\mathbf{e}$ *the conditions* $\rho + \sigma_j^3 \neq 0$ $(j = 0, 1)$, $d = 1$ *hold if and only if* $e = 0$ *and* $W[e_j(x)] = 1$ $(j = 0, 1)$.

Proof. "*if*": $e_j(x) = x^{k_j}$, $e = 0$ give $\sigma_j = b\alpha^s + \alpha^{k_j}$, $\sigma = b\alpha^{3s} + \alpha^{3k_j} + \alpha^{3k_{j+1}}$, $d = 1$, whence $\rho = b\alpha^{3s} + \alpha^{k_j}\alpha^{k_{j+1}}(\alpha^{k_j} + \alpha^{k_{j+1}})$. We then have

$$\rho + \sigma_j^3 = \alpha^{3k_j}\left[\left(\frac{\sigma_{j+1}}{\alpha^{k_j}}\right)^2 + \frac{\sigma_{j+1}}{\alpha^{k_j}} + 1\right] \neq 0$$

in $GF(2^{n-1})$, n even.

"*Only if*". If $d = 1$, $\rho + \sigma_j^3 \neq 0$ $(j = 0, 1)$ we calculate $\rho' \triangleq \sigma + \sigma_0\sigma_1(\sigma_0 + \sigma_1)$ and obtain $\sigma_{j+1} + \sqrt[3]{\rho'} = \alpha^{k_j}$. Then we set $\mathbf{c} = [x^{k_0}, 0, x^{k_1}]$ and compute Σ^*, i.e.,

$$\begin{aligned} \sigma_j^* &= \sigma_j + \alpha^{k_j} = \sigma_j + \sigma_{j+1} + \sqrt[3]{\rho'} = \sigma_{j+1}^* \\ \sigma^* &= \sigma + \alpha^{3k_j} + \alpha^{3k_{j+1}} = \sigma + (\sigma_{j+1} + \sqrt[3]{\rho'})^3 + (\sigma_j + \sqrt[3]{\rho'})^3 \\ &= (\sigma + \sigma_j\sigma_{j+1}(\sigma_{j+1} + \sigma_j) + \rho') + (\sigma_j + \sigma_{j+1} + \sqrt[3]{\rho'})^3 = \sigma_j^{*3} \end{aligned}$$

since $\sigma + \sigma_j\sigma_{j+1}(\sigma_{j+1} + \sigma_j) = \rho'$. Finally $d^* = d + c_1(1) = 0$. Relations (28) are satisfied with $c + W[c_0(x)] + W[c_1(x)] = 2$. Q.E.D.

We have therefore the following decoding rule:

Rule 2. If $\sqrt[3]{\rho} \neq \sigma_0$, $\sqrt[3]{\rho} \neq \sigma_1$, $d = 1$, then $c = 0$ and $c_j(x) = x^{k_j}$, where $\alpha^{k_j} = \sigma_{j+1} + \sqrt[3]{\sigma + \sigma_0\sigma_1(\sigma_0 + \sigma_1)}$.

Before giving Theorem 2.3, we notice that subject to $(\sigma_0 + \sigma_1) \neq 0$ the functions $\tau_j \triangleq (\rho + \sigma_j^3)/(\sigma_0 + \sigma_1)^3 (j = 0, 1)$ are related by

$$\tau_j + \tau_{j+1} + \frac{\sigma_0\sigma_1}{(\sigma_0 + \sigma_1)^2} + 1 = 0.$$

Expressing these elements of $\mathrm{GF}(2^{n-1})$ as column vectors with respect to the basis $1, \alpha, \cdots, \alpha^{n-2}$ and premultiplying by $\boldsymbol{\lambda}^T$ (see (26)) we have

$$\boldsymbol{\lambda}^T\boldsymbol{\tau}_j + \boldsymbol{\lambda}^T\boldsymbol{\tau}_{j+1} = 1$$

since $\boldsymbol{\lambda}^T \cdot \mathbf{1} = 1$ and $(\sigma_0\sigma_1)/(\sigma_0 + \sigma_1)^2 \in \Theta$ (in fact $\sigma_0/(\sigma_0 + \sigma_1)$ solves the equation $y^2 + y + [\sigma_0\sigma_1/(\sigma_0 + \sigma_1)^2] = 0$). This proves the following lemma.

LEMMA 11. *If* $(\sigma_0 + \sigma_1) \neq 0$, *then exactly one of the two functions* τ_j, τ_{j+1} *belongs to* Θ.

THEOREM 2.3. *For correctable* $\mathbf{e}$ *the conditions* $\rho + \sigma_j^3 \neq 0$ $(j = 0, 1)$, $d = 0$, $(\sigma_0 + \sigma_1) \neq 0$ *hold if and only if* $e_j(x) = 0$, $W[e_{j+1}(x)] = 2$, $e = 0$.

Proof. *"If"*: $e = 0$, $e_j(x) = 0$, $e_{j+1}(x) = x^{k_1} + x^{k_2}$ give $\sigma_j = b\alpha^s$, $\sigma_{j+1} = b\alpha^s + e_{j+1}(\alpha)$, $\sigma = b\alpha^{3s} + e_{j+1}(\alpha^3)$, whence $(e_{j+1}(\alpha) \neq 0)$

$$\rho + \sigma_j^3 = e_{j+1}^3(\alpha) + e_{j+1}(\alpha^3)$$
$$\rho + \sigma_{j+1}^3 = e_{j+1}^3(\alpha) + e_{j+1}(\alpha^3) + e_{j+1}^3(\alpha)\left[1 + \frac{\sigma_j}{e_{j+1}(\alpha)} + \left(\frac{\sigma_j}{e_{j+1}(\alpha)}\right)^2\right].$$

Recalling that $e_{j+1}^3(\alpha) + e_{j+1}(\alpha^3) = \alpha^{k_1}\alpha^{k_2}(\alpha^{k_1} + \alpha^{k_2})$ and letting $\gamma \triangleq \alpha^{k_1}/\alpha^{k_2} \neq 0$, $y \triangleq \sigma_j/e_{j+1}(\alpha)$ we have $\rho + \sigma_j^3 = \alpha^{3k_2}\gamma(1 + \gamma) \neq 0$ since $\gamma \neq 1 (k_1 \neq k_2)$, and

$$\rho + \sigma_{j+1}^3 = \alpha^{3k_2}(1 + \gamma)^3\left(y^2 + y + 1 + \frac{\gamma}{1 + \gamma^2}\right) \neq 0$$

since $1 \notin \Theta$ and $\gamma/(1 + \gamma^2) \in \Theta$ imply: $1 + \gamma/(1 + \gamma^2) \notin \Theta$. Moreover, $d = 0$ and $\sigma_0 + \sigma_1 = e_{j+1}(\alpha) \neq 0$.

"Only if": If $d = 0$, $\rho + \sigma_j^3 \neq 0$ $(j = 0, 1)$, $(\sigma_0 + \sigma_1) \neq 0$ we obtain

$\alpha^{k_1}/(\sigma_0 + \sigma_1)$ and $\alpha^{k_2}/(\sigma_0 + \sigma_1)$ as the solutions of

$$y^2 + y + \frac{\rho + \sigma_j^3}{(\sigma_0 + \sigma_1)^3} = 0 \tag{30}$$

(that (30) has solutions over $\mathrm{GF}(2^{n-1})$ for exactly one value of j is guaranteed by Lemma 11). Set $c_{j+1}(x) = x^{k_1} + x^{k_2}$, $c_j(x) = 0$, $c = 0$ and compute Σ^*, i.e.,

$$\begin{aligned}
d^* &= d + e_1(1) = 0,\\
\sigma_{j+1}^* &= \sigma_{j+1} + \frac{\alpha^{k_1} + \alpha^{k_2}}{\sigma_j + \sigma_{j+1}}(\sigma_j + \sigma_{j+1}) = \sigma_j = \sigma_j^*,\\
\sigma^* &= \sigma + \frac{\alpha^{3k_1} + \alpha^{3k_2}}{(\sigma_0 + \sigma_1)^3}(\sigma_0 + \sigma_1)^3\\
&= \sigma + (\sigma_0 + \sigma_1)^3\left(1 + \frac{\alpha^{k_1}\alpha^{k_2}}{(\sigma_0 + \sigma_1)^2}\,\frac{\alpha^{k_1} + \alpha^{k_2}}{\sigma_0 + \sigma_1}\right)\\
&= \sigma + (\sigma_0 + \sigma_1)^3\left(1 + \frac{\rho + \sigma_j^3}{(\sigma_0 + \sigma_1)^3}\right) = \sigma_j^3 = \sigma_j^{*3}
\end{aligned}$$

since $(\alpha^{k_1} + \alpha^{k_2})/(\sigma_0 + \sigma_1) = 1$ and $\alpha^{k_1}\alpha^{k_2}/(\sigma_0 + \sigma_1)^2 = (\rho + \sigma_j^3)/(\sigma_0 + \sigma_1)^3$, being the sum and the product of the solutions of (30), respectively. Relations (28) are satisfied with $c + W[c_0(x)] + W[c_1(x)] = 2$. Q.E.D.

This yields the following decoding rule:

Rule 3. If $\sqrt[3]{\rho} \neq \sigma_0$, $\sqrt[3]{\rho} \neq \sigma_1$, $d = 0$, $\sigma_0 + \sigma_1 \neq 0$, then set $c = 0$, $c_j(x) = 0$ and $c_{j+1}(x) = x^{k_1} + x^{k_2}$, where α^{k_1} and α^{k_2} are the solutions of $z^2 + (\sigma_0 + \sigma_1)z + (\rho + \sigma_j^3)/(\sigma_0 + \sigma_1) = 0$.

Rules 1, 2, 3 constitute an algorithm which encompasses the correction of all the correctable error patterns. What is the behavior of this algorithm when the received $\mathbf{r}$ is at distance ≥ 3 from any $\mathbf{w} \in \mathcal{K}_n$? The answer to this question is implicitly provided by the previous three theorems, which give necessary and sufficient conditions for the existence of a code word within distance 2 from the received word $\mathbf{r}$. Therefore $\mathbf{r}$ lies at distance ≥ 3 from any code word if and only if $\rho + \sigma_j^3 \neq 0$ $(j = 0, 1)$ (Theorem 2.1), $d = 0$ (Theorem 2.2) and $\sigma_0 + \sigma_1 = 0$ (Theorem 2.3). When Σ satisfies these conditions, clearly we can no longer perform the correction. In fact, while the distance properties of $\mathcal{K}_n$ guarantee that an existing correction vector $\mathbf{c}$ of weight ≤ 2 is also unique, more than one **c** of weight 3 can be constructed when Rules 1, 2 and 3 are inapplicable. This is shown by the following argument. Assume that the conditions $\sqrt[3]{\rho} \neq \sigma_j$ $(j = 0, 1)$, $d = 0$, $\sigma_0 = \sigma_1$ hold for **r**. We determine α^h such that $(1 + (\sigma + \sigma_0^3)/\alpha^{3h}) \in \Theta$: there are 2^{n-2} values of h which meet this requirement, since α^3 generates the multiplicative group of GF(2^{n-1}) and, for fixed $(\sigma + \sigma_0^3)$, $(1 + (\sigma + \sigma_0^3)/\alpha^{3h})$ spans the set $\{0, \alpha, \alpha^2, \cdots, \alpha^{2^n-2}\}$ which contains Θ for even n (Lemma 8). We then form a correction vector **c** as follows: $c = 0$, $c_0(x) = x^h$, $c_1(x) = x^{k_1} + x^{k_2}$, where $\alpha^{k_i} = \sigma_0 + z_i\alpha^h$ $(i = 1, 2)$ and z_1, z_2 are the solutions of

$$z^2 + z + 1 + \frac{\sigma + \sigma_0^3}{\alpha^{3h}} = 0.$$

We notice that $\alpha^{k_1} + \alpha^{k_2} = (z_1 + z_2)\alpha^h = \alpha^h$ since $z_1 + z_2 = 1$. Recalling that $\sigma_0 = \sigma_1$, the syndrome Σ^* yields (see (29))

$$\begin{aligned} \sigma_0^* &= \sigma_0 + \alpha^h = \sigma_1 + \alpha^{k_1} + \alpha^{k_2} = \sigma_1^*, \\ \sigma^* &= \sigma + \alpha^{3h} + \alpha^{3k_1} + \alpha^{3k_2} = \sigma + \alpha^{k_1}\alpha^{k_2}(\alpha^{k_1} + \alpha^{k_2}) \\ &= \sigma + \alpha^h(\sigma_0 + z_1\alpha^h)(\sigma_0 + z_2\alpha^h) = (\sigma_0 + \alpha^h)^3 = \sigma_0^{*3}, \\ d^* &= c + c_1(1) = 0 \end{aligned}$$

i.e., $\mathbf{r} + \mathbf{c} \in \mathcal{K}_n$. This discussion proves that there are several[4] code words at distance 3 from **r** (but none at distance <3) and yields the following error detection rule:

Rule 4. If $\sqrt[3]{\rho} \neq \sigma_0$, $\sqrt[3]{\rho} \neq \sigma_1$, $d = 0$, $\sigma_0 + \sigma_1 = 0$, then the received **r** is at distance $\geqq 3$ from any code word.

An "extra bonus" of the same discussion is that given any **r** there are code words at distance $\leqq 3$ from **r**: this property is analogous to the one found by Gorenstein et al. (1960) for BCH double-error-correcting codes.

In Figure 2 we sketch a possible organization of a decoder for a $\mathcal{K}_n$ code. The serially received message is stored in three recirculating registers SR1, SR2, SR3, corresponding to the homologous registers of Figure 1. The received message is also fed to the SYNDROME COMPUTER, which, once reception is completed, stores the functions d, σ, σ_0, σ_1, i.e., the syndrome Σ. These functions constitute the inputs of combinational networks, which we now describe (in the illustration heavy

[4] Another weight 3 correction vector is obtained through Rule 2, i.e., $\mathbf{c} = [\alpha^k, 1, \alpha^k]$ where $\alpha^k = \sigma_0 + \sqrt[3]{\sigma} = \sigma_1 + \sqrt[3]{\sigma}$.

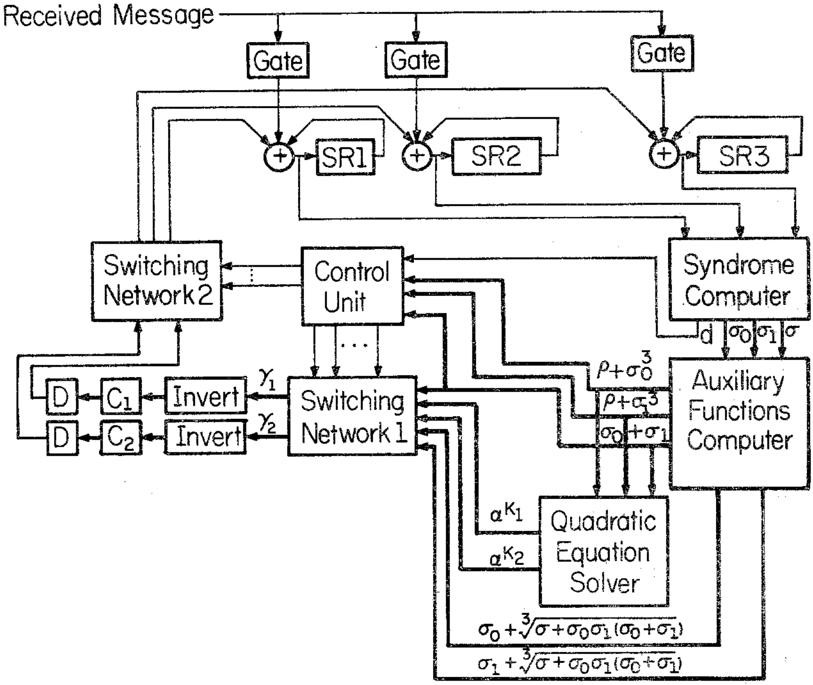

FIG. 2. Decoder for the $\mathcal{K}_n$ code.

lines denote bundles of $(n - 1)$ binary lines). The AUXILIARY FUNCTIONS COMPUTER produces $\rho + \sigma_0^3$, $\rho + \sigma_1^3$, $\sigma_0 + \sigma_1$, $\sigma_0 + \sqrt[3]{\sigma + \sigma_0\sigma_1(\sigma_0 + \sigma_1)}$ and $\sigma_1 + \sqrt[3]{\sigma + \sigma_0\sigma_1(\sigma_0 + \sigma_1)}$. Of these, $\rho + \sigma_0^3$, $\rho + \sigma_1^3$, $\sigma_0 + \sigma_1$, together with d, are fed to the CONTROL UNIT, which determines which decoding rule must be applied. In parallel, $\rho + \sigma_0^3$, $\rho + \sigma_1^3$, $\sigma_0 + \sigma_1$ are fed to the QUADRATIC EQUATION SOLVER, where $\tau_0 = (\rho + \sigma_0^3)/(\sigma_0 + \sigma_1)^3$ and $\tau_1 = (\rho + \sigma_1^3)/(\sigma_0 + \sigma_1)^3$ are computed and tested for membership in Θ. If $\tau_j \in \Theta$, $T\tau_j$ and $(1 + T\tau_j)$ are the solutions of the equation $y^2 + y + \tau_j = 0$, where T is an appropriate square matrix (see, e.g., Berlekamp et al., (1967)); $T\tau_j$ and $(1 + T\tau_j)$ are then multiplied in GF(2^{n-1}) by $(\sigma_0 + \sigma_1)$ in order to obtain the solutions α^{k_1} and α^{k_2} of $z^2 + (\sigma_0 + \sigma_1)z + (\rho + \sigma_j^3)/(\sigma_0 + \sigma_1) = 0$ (see Rule 3). Since $W[c_0(x)] + W[c_1(x)] \leqq 2$, at most two correction bits must be produced. This is accomplished as follows: 1) $\sigma_0 + \sigma_1$, $\sigma_0 + \sqrt[3]{\sigma + \sigma_0\sigma_1(\sigma_0 + \sigma_1)}$, $\sigma_1 + \sqrt[3]{\sigma + \sigma_0\sigma_1(\sigma_0 + \sigma_1)}$, α^{k_1}, α^{k_2} are fed to the SWITCHING NETWORK 1: here signals from the control unit govern the selection of two correction functions γ_1, γ_2 in the form of the vector representations of two elements of $GF(2^{n-1})$; 2) the combinational circuit INVERT computes the inverse of $\gamma_j = \alpha^{h_j}$ (if $\gamma_j = 0$, the output of INVERT is conventionally 0) and α^{-h_j} is loaded in the Galois field counter C_j. It must be noticed that, assuming no delay in the combinational elements, loading of C_j with $\alpha^{-h_j} (j = 1, 2)$ occurs simultaneously with the production of d, σ, σ_0, σ_1. At this point the contents of SR2 and SR3 are recirculated synchronously with the stepping of C_1 and C_2: once the condition $10 \cdots 0$ is detected in C_j a time unit duration signal is generated by D and routed through the SWITCHING NETWORK 2 to perform the required correction of the contents of the registers. The decoding operation therefore terminates $(2^{n-1} - 1)$ time units after the serial reception of the message is completed.

This completes the presentation of the decoding procedure.

6. ACKNOWLEDGMENT

The author gratefully acknowledges the very valuable suggestions of E. R. Berlekamp and J..P. Robinson.

RECEIVED: June 28, 1968; revised September 11, 1968

REFERENCES

ALBERT, A. A. (1956), "Fundamental Concepts of Higher Algebra." University of Chicago Press, Chicago.

VASIL'EV, YU. L. (1962), O negruppovykh plotno upakovannykh kodakh (On nongroup close packed codes). *Problemy Kibernetiki.* **8,** 337–339.

BERLEKAMP, E. R., RUMSEY, H. AND SOLOMON, G. (1967). On the solution of algebraic equations over finite fields. *Inform. Control.* **10,** 553–564.

BERLEKAMP, E. R. (1968), "Algebraic Coding Theory." McGraw-Hill, New York.

GORENSTEIN, D., PETERSON, W. W. AND ZIERLER, N. (1960). Two-error correcting Bose–Chaudhuri codes are quasi-perfect. *Inform. Control.* **3,** 291–294.

GREEN, M. V. (1966). Two heuristic techniques for block-code construction. *IEEE Trans. on Inform. Theory.* **IT-12,** 273.

JOHNSON, S. M. (1962). A new upper-bound for error-correcting codes. *IRE Trans. on Inform. Theory.* **IT-8,** 203–207.

NADLER, M. (1962). A 32-point $n = 12$, $d = 5$ Code. *IRE Trans. on Inform. Theory.* **IT-8,** 58.

NORDSTROM, A. W. AND ROBINSON, J. P. (1967). An optimum nonlinear code. *Inform. Control.* **11,** 613–616.

PETERSON, W. W. (1961). "Error Correcting Codes." M.I.T. Press and John Wiley & Sons, New York.

PREPARATA, F. P. (1968a). An alternate description and a new decoding procedure of Nordstrom–Robinson optimum code. Proc. 2nd Princeton Conf. on Information Sciences and Systems, 131–134.

PREPARATA, F. P. (1968b). Weight and distance structure of Nordstrom–Robinson Quadratic Code. *Inform. Control.* **12,** 466–473. (see also Erratum, **13,** no. 7).