New Binary Codes

NEIL J. A. SLOANE, MEMBER, IEEE, SUDHAKAR M. REDDY, MEMBER, IEEE, AND CHIN-LONG CHEN, MEMBER, IEEE

Abstract—**In this paper constructions are given for combining two, three, or four codes to obtain new codes. The Andryanov–Saskovets construction is generalized. It is shown that the Preparata double-error-correcting codes may be extended by about (block length)$^{1/2}$ symbols, of which only one is a check symbol, and that e-error-correcting BCH codes may sometimes be extended by (block length)$^{1/e}$ symbols, of which only one is a check symbol. Several new families of linear and nonlinear double-error-correcting codes are obtained. Finally, an infinite family of linear codes is given with $d/n = \frac{1}{3}$, the first three being the $(24,2^{12},8)$ Golay code, a $(48,2^{15},16)$ code, and a $(96,2^{18},32)$ code. Most of the codes given have more codewords than any comparable code previously known to us.**

INTRODUCTION

Definitions

An (n,M,d) code $\mathscr{C}$ is a set of M binary vectors of length n, any two of which differ in at least d places. The redundancy of this code is $r = n - \log_2 M$ and its rate is $(\log_2 M)/n$.

A *coset* of $\mathscr{C}$ is an arbitrary translation $a + \mathscr{C}$ of the codewords of $\mathscr{C}$ (where a is any binary vector of length n). If $\mathscr{C}$ is linear, then two cosets of $\mathscr{C}$ are either equal or disjoint, but this need not be true if $\mathscr{C}$ is nonlinear.

Two codes are said to be *equivalent* if they differ only by a permutation of the coordinates (Peterson [14, p. 33]).

Summary

Section I describes a construction for combining three codes to form a fourth. When applied to BCH codes it produces, among others, codes with the same parameters as those given by the Andryanov–Saskovets construction (Berlekamp [2, p. 333]). When applied to cyclic and Preparata codes, it yields several good new codes. Encoding and decoding methods are given for the new codes.

Section II describes a construction for combining four codes to form a fifth. When applied to e-error-correcting BCH codes, in the most favorable cases the codewords may be extended by about $n^{1/e}$ symbols, of which only one is a check symbol. Double-error-correcting codes are studied in detail in both Sections II and III, and several new infinite families, both linear and nonlinear, are obtained. The results are summarized in Table II, which gives for $6 \leq r \leq 35$ the length of the longest distance-5 code with redundancy r known to us. Again encoding and decoding methods are discussed.

Section III describes three constructions due to Goethals for combining a code and its dual. Applications are given to double- and triple-error-correcting BCH codes, to cyclic codes, and to quadratic-residue codes. It is shown in effect that double-error-correcting BCH codes may be extended by about $2\sqrt{n}$ symbols, of which only two are check symbols.

Finally in Section IV a construction for combining two different first-order Reed–Muller codes is used to obtain an infinite family of linear codes with $d/n = \frac{1}{3}$, the first three being the $(24,2^{12},8)$ Golay code, a $(48,2^{15},16)$ code, and a $(96,2^{18},32)$ code.

Most of the codes given as examples have more codewords than any comparable code previously known to us. However, apart from the Preparata codes, none of the codes mentioned is known to be optimal. (For extensive tables of upper and lower bounds on the sizes of codes see [10], [12], and [19].)

I. CONSTRUCTION X: COMBINING THREE CODES

The Construction

Suppose we are given an (n_1,M_1,d_1) code $\mathscr{C}_1$ and an $(n_1,M_2 = bM_1,d_2)$ code $\mathscr{C}_2$, with the property that $\mathscr{C}_2$ is the union of b disjoint cosets of $\mathscr{C}_1$,

$$\mathscr{C}_2 = (x_1 + \mathscr{C}_1) \cup (x_2 + \mathscr{C}_1) \cup \cdots \cup (x_b + \mathscr{C}_1)$$

for some set of vectors $S = \{x_1,x_2,\cdots,x_b\}$. Let $\mathscr{C}_3 = \{y_1,y_2,\cdots,y_b\}$ be any (n_3,b,Δ) code.

Let π be an arbitrary permutation of $\{1,2,\cdots,b\}$, so that $x_i \to y_{\pi(i)}$ defines a one–one mapping from S onto $\mathscr{C}_3$. Let (u,v) denote the vector formed by concatenating vectors u,v, and if S is a set of vectors, let (S,v) denote the set of all (u,v), $u \in S$.

The new code $\mathscr{C}_4$ is then defined to be

$$(x_1 + \mathscr{C}_1,y_{\pi(1)}) \cup (x_2 + \mathscr{C}_1,y_{\pi(2)}) \cup \cdots \cup (x_b + \mathscr{C}_1,y_{\pi(b)}).$$

Simply stated, $\mathscr{C}_2$ is divided into cosets of $\mathscr{C}_1$ and a different codeword of $\mathscr{C}_3$ is attached to each coset. See Fig. 1.

The parameters of the new code are given by the following theorem.

Theorem 1: $\mathscr{C}_4$ is an

$$(n_1 + n_3,M_2 = bM_1,d_4 = \min\{d_1,d_2 + \Delta\})$$

code.

Proof: Let $X = (x,y)$ and $X' = (x',y')$ be distinct codewords of $\mathscr{C}_4$. If x and x' belong to the same coset of $\mathscr{C}_1$, then $y = y'$ and $\text{dist}(X,X') = \text{dist}(x,x') \geq d_1$. If x

Manuscript received July 12, 1971. This work was supported in part by the National Science Foundation under Grants GK-10025 (SMR) and GK-24879 (CLC).

N. J. A. Sloane is with the Bell Telephone Laboratories, Inc., Murray Hill, N.J.

S. M. Reddy is with the Department of Electrical Engineering, University of Iowa, Iowa City, Iowa.

C.-L. Chen is with the Coordinated Science Laboratory, University of Illinois, Urbana, Ill.

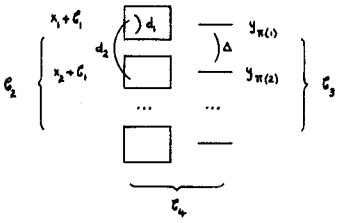

Fig. 1. Construction X.

and x' belong to different cosets, then $\text{dist}(x,x') \geq d_2$, $\text{dist}(y,y') \geq \Delta$, and $\text{dist}(X,X') \geq d_2 + \Delta$. Q.E.D.

A Simple Example

Let $\mathscr{C}_1$ be the (4,2,4) repetition code {0000,1111} and let $\mathscr{C}_2$ be the (4,8,2) even-weight code {0000,1111,0011,1100, 0101,1010,1001,0110};

$$\mathscr{C}_2 = \mathscr{C}_1 \cup (0011 + \mathscr{C}_1) \cup (0101 + \mathscr{C}_1) \cup (1001 + \mathscr{C}_1).$$

Then $b = 4$, so let $\mathscr{C}_3$ be the (4,3,2) code {000,011,101,110}. Attaching $\mathscr{C}_3$ to the tails of the cosets we obtain $\mathscr{C}_4$ = {0000000,1111000,0011011,1100011,0101101,1010101, 1001110,0110110}, a (7,8,4) linear code.

Linear Codes

As in the last example, if $\mathscr{C}_1$, $\mathscr{C}_2$, and $\mathscr{C}_3$ are all linear, we can always choose π so as to make $\mathscr{C}_4$ a linear code. (For then $\mathscr{C}_2/\mathscr{C}_1$ and $\mathscr{C}_3$ are both Abelian groups of type $(1,1,\cdots,1)$, and so there exists an isomorphism π between them. See, for example, Carmichael [4, pp. 98–100].)

We give three principal applications of Construction X, using BCH, cyclic, and Preparata codes.

Example i): Using BCH Codes

When $d_1 > d_2$, the BCH code of designed distance d_1 is contained in that of designed distance d_2, so the construction may be applied to any pair of BCH codes.

If $d_1 \geq d_2 + 2$ and $b = 2^k$, then $\mathscr{C}_3$ may for example be taken to be the $(k + 1,2^k,2)$ even-weight code. In this case we are combining $(n_1,M_1,d_1 \geq d_2 + 2)$ and $(n_1,2^kM_1, d_2)$ BCH codes to obtain an $(n_1 + k + 1,2^kM_1,d_2 + 2)$ code. Thus we can construct linear codes having the same parameters as any of the codes given by the Andryanov–Saskovets construction (see Berlekamp [2, p. 333]). Examples follow:

$\mathscr{C}_1$	$\mathscr{C}_2$	$\mathscr{C}_4$
$(31,2^6,15)$	$(31,2^{11},11)$	$(37,2^{11},13)$
$(63,2^{45},7)$	$(63,2^{51},5)$	$(70,2^{51},7)$
$(63,2^{36},11)$	$(63,2^{39},9)$	$(67,2^{39},11)$
$(63,2^{10},27)$	$(63,2^{16},23)$	$(70,2^{16},25)$.

On the other hand, if d_1 is greater than $d_2 + 2$, good codes may sometimes be obtained by choosing $\mathscr{C}_3$ to be an $(n_3,b,\Delta = d_1 - d_2)$ code, where n_3 is as small as possible. The first and fourth of the preceding examples may be used to illustrate this:

$\mathscr{C}_1$	$\mathscr{C}_2$	$\mathscr{C}_3$	$\mathscr{C}_4$
$(31,2^6,15)$	$(31,2^{11},11)$	$(10,2^5,4)$	$(41,2^{11},15)$
$(63,2^{10},27)$	$(63,2^{16},23)$	$(11,2^6,4)$	$(74,2^{16},27)$.

Two other examples are

$(127,2^{71},19)$	$(127,2^{78},15)$	$(12,2^7,4)$	$(139,2^{78},19)$
$(127,2^{50},27)$	$(127,2^{57},23)$	$(12,2^7,4)$	$(139,2^{57},27)$.

Example ii): Using Cyclic Codes

Chen [5], [6] has found the minimum distance of a large number of binary cyclic codes of length ≤ 65. Using these data and Construction X, the codes shown in Table I are obtained. Here $\mathscr{C}_1$ and $\mathscr{C}_2$ are $(n_1,2^{k_1},d_1)$ and $(n_1, 2^{k_2},d_2)$ codes, respectively, the new code $\mathscr{C}_4$ is an $(n_4, 2^{k_2},d_4)$ code, and $\mathscr{C}_3$ can be deduced from the others. The table is arranged in order of increasing d_4.

Example iii): Using the Preparata Codes

We show that certain Hamming codes are expressible as a union of disjoint cosets of a Preparata code. This fact is then combined with Construction X to give an infinite family of nonlinear double-error-correcting codes, and will also be used in Section II to construct other families of codes.

For every even integer $m \geq 4$, Preparata [17] has constructed an optimal nonlinear

$$(2^m - 1,2^{2^m-2m},5)$$

code $\mathscr{K}_m$. The codewords are specified in terms of polynomials in the algebra $\mathscr{A}_{m-1}$ of polynomials modulo $x^{2^{m-1}-1} + 1$. Let α be a primitive element of $GF(2^{m-1})$ and let $M^{(i)}(x)$ be the minimal polynomial of α^i.

We first define three fixed polynomials. They are

$$u(x) = (x^{2^{m-1}-1} + 1)/(x + 1)$$

$$\phi(x) = (x^{2^{m-1}-1} + 1)/M^{(1)}(x)$$

and

$$f(x) = x^t\phi(x),$$

where t is chosen so that $f(x)^2 = f(x)$ in $\mathscr{A}_{m-1}$.

Then the codewords of $\mathscr{K}_m$ are all vectors of the form

$$(c(x) + q(x),i,c(x) + q(x)f(x) + (c(1) + i)u(x) + s(x)),$$

where $c(x)$, $q(x)$, i, and $s(x)$ are variables: $c(x)$ is any codeword in the Hamming code $\mathscr{H}_{m-1}$ of length $2^{m-1} - 1$; $q(x) = ax^j$ for $a = 0$ or 1 and $j = 0$ or 1 or $\cdots$ or $2^{m-1} - 2$; i is 0 or 1; and $s(x)$ is any codeword in the $d \geq 6$ BCH code $\mathscr{B}_{m-1}$ generated by $(x + 1)M^{(1)}(x)M^{(3)}(x)$. For the proof that $\mathscr{K}_m$ has minimum distance 5, see Preparata [17].

We shall show that the Hamming code $\mathscr{H}_m$ is a union of disjoint cosets of the Preparata code $\mathscr{K}_m$. In fact since the linear Vasil'ev code is equivalent to the Hamming code (see what follows), we shall show that the linear Vasil'ev code is a union of disjoint cosets of $\mathscr{K}_m$. (This generalizes a remark of Preparata [16], that $\mathscr{K}_4$ is a subcode of a linear Vasil'ev code.)

Vasil'ev [22] obtained a class of perfect $(2^m - 1,2^{2^m-m-1}$,

TABLE I
NEW $(n_4, 2^{k_2}, d_4)$ CODES FROM CONSTRUCTION X

n	k_1	d_1	Type	r_1	d_2	Const.	n′	M	d′
63	46	7	Cyclic	17	22	Y1	41	2^{25}	7
31	10	12	BCH	21	5	Y2	25	320	9
48	24	12	QR	24	12	Y2	35	$3\cdot2^{15}$	9
63	36	11	BCH	27	14	Y2	49	$7\cdot2^{24}$	9
48	24	12	QR	24	12	Y1	35	2^{13}	11
63	28	15	Cyclic	35	8	Y2	55	2^{24}	13
90	45	18	QR	45	18	Y2	71	$9\cdot2^{29}$	15
63	18	21	BCH	45	8	Y3	55	$29\cdot2^{11}$	17
63	21	18	Cyclic	42	6	Y1	56	2^{16}	17
90	45	18	QR	45	18	Y1	71	2^{28}	17
63	18	21	BCH	45	8	Y2	55	2^{14}	19
63	16	23	BCH	47	6	Y3	57	2^{15}	19
104	52	20	QR	52	20	Y1	83	2^{33}	19
63	16	23	BCH	47	6	Y2	57	$6\cdot2^{11}$	21
63	16	23	BCH	47	6	Y1	57	2^{11}	23

3) codes, both linear and nonlinear, by the following construction. Let λ be any mapping that assigns the values 0 and 1 to the codewords of $\mathcal{H}_{m-1}$ and satisfies $\lambda[0] = 0$. Then the Vasil'ev code $\mathcal{V}_m^{\lambda}$ has as codewords all vectors

$$(p(x),p(1) + \lambda[b(x)],p(x) + b(x)),$$

where $p(x)$ and $b(x)$ are variables: $p(x) \in \mathcal{A}_{m-1}$ and $b(x) \in \mathcal{H}_{m-1}$. In particular, if λ is the mapping $\lambda[b(x)] = b(1)$, the resulting code is linear and is denoted simply by $\mathcal{V}_m$. Since the Hamming code $\mathcal{H}_m$ is the unique perfect linear code with minimum distance 3, $\mathcal{H}_m$ and $\mathcal{V}_m$ are equivalent.

Theorem 2: The linear Vasil'ev code $\mathcal{V}_m$, or equivalently the Hamming code $\mathcal{H}_m$, is a union of disjoint cosets of the Preparata code $\mathcal{K}_m$, where m is any even integer ≥ 4.

Proof: We shall establish the theorem by constructing vectors $x_0 = 0,x_1,\cdots,x_{2^{m-1}-1}$ in $\mathcal{V}_m$ such that any $v \in \mathcal{V}_m$ has a unique representation of the form $v = x_i + w$, where $w \in \mathcal{K}_m$.

Let $v \in \mathcal{V}_m$, so

$$v = (p(x),p(1) + b(1),p(x) + b(x)).$$

Since $\mathcal{H}_{m-1}$ is a perfect single-error-correcting code, $p(x)$ has a unique representation as $p(x) = c(x) + q(x)$, where $c(x) \in \mathcal{H}_{m-1}$ and $q(x) = 0$ or x^j for $j = 0$ or 1 or $\cdots$ or $2^{m-1} - 2$. Let $i = c(1) + q(1) + b(1)$. Then

$$\begin{aligned} v &= (c(x) + q(x),i,c(x) + q(x) + b(x)) \\ &= (c(x) + q(x),i,c(x) + q(x)f(x) \\ &\quad + (q(1) + b(1))u(x) + \hat{s}(x)) \end{aligned}$$

say, where $\hat{s}(x) = q(x)(1 + f(x)) + b(x) + (q(1) + b(1))u(x)$. By Preparata [17, lemma 4], $q(x)(1 + f(x)) \in \mathcal{H}_{m-1}$. Since $f(1) = 0$, $u(1) = 1$, it follows that $\hat{s}(1) = 0$. Therefore $\hat{s}(x) \in \mathcal{H}_{m-1}^{+}$, the even-weight subcode of $\mathcal{H}_{m-1}$.

Since $\mathcal{B}_{m-1}$ is a subcode of $\mathcal{H}_{m-1}^{+}$, we may choose a set of vectors $y_0,y_1,\cdots,y_{2^{m-1}-1} \in \mathcal{H}_{m-1}^{+}$ so that any $\hat{s}(x) \in \mathcal{H}_{m-1}^{+}$ has a unique representation of the form $\hat{s}(x) = y_i + s(x)$, where $s(x) \in \mathcal{B}_{m-1}$. Therefore

$$\begin{aligned} v &= (c(x) + q(x),i,c(x) + q(x)f(x) \\ &\quad + (i + c(1))u(x) + s(x)) + (0,0,y_i) \\ &= w + (0,0,y_i), \end{aligned}$$

where $w \in \mathcal{K}_m$, and this representation is unique. Taking $x_i = (0,0,y_i)$ proves the theorem. Q.E.D.

New Codes

By Theorem 2, we may apply Construction X with $\mathcal{C}_1$ equal to the Preparata code $\mathcal{K}_m$, $\mathcal{C}_2$ equal to the Hamming code $\mathcal{H}_m$, and $\mathcal{C}_3$ equal to the $(m,2^{m-1},2)$ even-weight code. We obtain an infinite family of nonlinear

$$(2^m + m - 1,2^{2m-m-1},5), \qquad m = 4,6,8,\cdots$$

codes. When $m = 4$, this is a $(19,2^{11},5)$ code. For $m \geq 6$, Theorem 2 will be used in Section II to construct codes with more information symbols than this family.

As a last example we consider the following.

Example iv): Adding an Overall Parity Check

Suppose $\mathcal{C}$ is an $(n,2^k,d)$ linear code, with d odd. The set of all codewords of even weight forms an $(n,2^{k-1},d_{\text{even}})$ subcode $\mathcal{C}^{+}$, with $d_{\text{even}} \geq d + 1$. By applying Construction X with $\mathcal{C}_1 = \mathcal{C}^{+}$, $\mathcal{C}_2 = \mathcal{C}$, and $\mathcal{C}_3 = \{0,1\}$, we obtain an $(n + 1,2^k,d + 1)$ code; we have just added an overall parity check to $\mathcal{C}$ (Berlekamp [2, p. 333]).

Remark

Construction X works even if the cosets of $\mathcal{C}_1$ in $\mathcal{C}_2$ are not of equal size. A trivial application is the addition of an overall parity check to a nonlinear code in which more than half of the codewords have odd weight (e.g., the (8,20,3) code described in [20]).

Encoding and Decoding for Construction X

a) Encoding: Let $\mathcal{C}_4$ be obtained from codes $\mathcal{C}_1,\mathcal{C}_2,\mathcal{C}_3$ by Construction X. For simplicity we assume that all the codes are linear and let $M_1 = 2^{k_1}$, $M_2 = 2^{k_2}$, $M_3 = 2^{k_2-k_1}$. Then $\mathcal{C}_4$ contains 2^{k_2} codewords. We assume that encoders and decoders for $\mathcal{C}_1,\mathcal{C}_2,\mathcal{C}_3$ are available and correspond to generator matrices

$$G_1 = \begin{array}{|c|c|}\hline A_1 & I_{k_1} \\ \hline \end{array}_{k_1 \times n_1}$$

$$G_2 = \begin{array}{|c|c|c|}\hline \multicolumn{2}{|c|}{A_1} & I_{k_1} \\ \hline A_2 & I_{k_2-k_1} & 0 \\ \hline \end{array}_{k_2 \times n_1}$$

$$G_3 = \begin{array}{|c|c|}\hline A_3 & I_{k_2-k_1} \\ \hline \end{array}_{(k_2-k_1) \times n_3}$$

for $\mathcal{C}_1$, $\mathcal{C}_2$, and $\mathcal{C}_3$, respectively. Here I_k denotes the $k \times k$ identity matrix and A_1,A_2,A_3 are nonzero matrices. We use Encoder_1 to denote the encoder for $\mathcal{C}_1$, and so forth.

Let $I = (i_1,i_2,\cdots,i_{k_2})$ be the string of information symbols to be encoded by $\mathscr{C}_4$. The encoding is accomplished by feeding I into Encoder$_2$, producing an output u (say) and feeding $(i_{k_2-k_1+1},\cdots,i_{k_2})$ into Encoder$_3$, producing an output v (say). Then (u,v) is the corresponding codeword of $\mathscr{C}_4$. This corresponds to using the generator matrix

$$G_4 = \begin{array}{|cc|c|cc|} \multicolumn{2}{|c|}{A_1} & I_{k_1} & \multicolumn{2}{c|}{0} \\ \hline A_2 & I_{k_2-k_1} & 0 & A_3 & I_{k_2-k_1} \end{array}_{k_2 \times (n_1+n_3)}$$

for $\mathscr{C}_4$.

b) Decoding: This is a modification of the decoding algorithm for product codes given in [18]. The minimum distance of $\mathscr{C}_4$ is $d_4 = \min\{d_1,d_2 + d_3\}$. We suppose that $e \leq [\frac{1}{2}(d_4 - 1)]$ errors have occurred.

Let $R = (r_1,r_2,\cdots,r_{n_1+n_3})$ be received. We feed $(r_1,\cdots,r_{n_1})$ into Decoder$_2$ and $(r_{n_1+1},\cdots,r_{n_3})$ into Decoder$_3$. For Decoder$_i$, where $i = 2$ or 3, let e_i be the number of errors found and let $\mu_i = d_i - 2e_i$ be the associated "reliability" (see lemma following). If both Decoder$_2$ and Decoder$_3$ have made a decoding error, then at least

$$[\tfrac{1}{2}(d_2 - 1)] + [\tfrac{1}{2}(d_3 - 1)] + 2 > [\tfrac{1}{2}(d_4 - 1)]$$

errors have occurred, which contradicts our hypothesis. So at least one of the decoders has made a correct decision. Because of the following lemma, we decide that the correct decoder is that with the largest μ_i.

Lemma: If Decoder$_2$ is correct and Decoder$_3$ is incorrect, then $\mu_2 \geq \mu_3$, and vice versa.

Proof: Let a_i be the actual number of errors in Decoder$_i$. Since Decoder$_3$ is in error, by [18, Lemma 1], $a_3 \geq d_3 - e_3$. By hypothesis $a_2 + a_3 \leq \frac{1}{2}(d_2 + d_3 - 1)$. Then $\mu_2 = d_2 - 2e_2 = d_2 - 2a_2$, $\mu_3 \leq 2a_3 - d_3$, and $\mu_2 - \mu_3 \geq 0$. Q.E.D.

If Decoder$_2$ was correct, we now know the information symbols $(i_{k_2-k_1+1},\cdots,i_{k_2})$. By feeding these into Encoder$_1$ and substracting the result from the output of Decoder$_2$, the remaining information symbols are recovered correctly. A similar discussion applies if Decoder$_3$ was correct.

Thus we may decode up to $[\frac{1}{2}(d_4 - 1)]$ errors in $\mathscr{C}_4$.

II. Construction $X4$. Combining Four Codes

The Construction

Suppose we are given four codes: an (n_1,M_1,d_1) code $\mathscr{C}_1$, an $(n_1,M_2 = bM_1,d_2)$ code $\mathscr{C}_2$, an (n_3,M_3,d_1) code $\mathscr{C}_3$, and an $(n_3,M_4 = bM_3,d_4)$ code $\mathscr{C}_4$, with the properties that i) $\mathscr{C}_2$ is a union of b disjoint cosets of $\mathscr{C}_1$,

$$\mathscr{C}_2 = (x_1 + \mathscr{C}_1) \cup (x_2 + \mathscr{C}_1) \cup \cdots \cup (x_b + \mathscr{C}_1),$$

and ii) $\mathscr{C}_4$ is a union of b disjoint cosets of $\mathscr{C}_3$,

$$\mathscr{C}_4 = (y_1 + \mathscr{C}_3) \cup (y_2 + \mathscr{C}_3) \cup \cdots \cup (y_b + \mathscr{C}_3)$$

for some sets of vectors $S = \{x_1,x_2,\cdots,x_b\}$ and $T = \{y_1,y_2,\cdots,y_b\}$.

As in Construction X, let π be an arbitrary permutation of $\{1,2,\cdots,b\}$, so that $x_i \to y_{\pi(i)}$ defines a one–one mapping from S onto T.

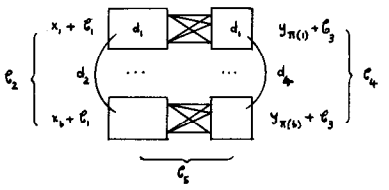

Fig. 2. Construction $X4$.

If S_1 and S_2 are arbitrary sets of vectors, we define $S_1 \times S_2$ to be the set of all possible concatenations (s_1,s_2), where $s_1 \in S_1$, $s_2 \in S_2$.

Finally, the new code $\mathscr{C}_5$ is defined to be

$$\mathscr{C}_5 = \bigcup_{i=1}^{b} (x_i + \mathscr{C}_1) \times (y_{\pi(i)} + \mathscr{C}_3).$$

Simply stated, the vectors of the ith coset of $\mathscr{C}_1$ are concatenated in every possible way with the vectors of the $\pi(i)$th coset of $\mathscr{C}_3$. See Fig. 2.

The parameters of the new code are given by the following theorem, whose proof is immediate.

Theorem 4: $\mathscr{C}_5$ is an

$$(n_1 + n_3, M_2M_3, d_5 = \min\{d_1,d_2 + d_4\})$$

code.

Linear Codes

As in Construction X, if $\mathscr{C}_1$–$\mathscr{C}_4$ are all linear, we can always choose π so as to make $\mathscr{C}_5$ a linear code.

We will apply Construction $X4$ to double-error-correcting codes and then to e-error-correcting BCH codes. First we define some nonprimitive BCH codes.

Nonprimitive BCH Codes of Length $2^m + 1$

As pointed out by several authors [8], [2, pp. 139–140], [13], [11] nonprimitive BCH codes sometimes contain more information symbols with the same redundancy as primitive BCH codes.

For later use we define the following codes of length $n = 2^m + 1$. Let α be a primitive nth root of unity, and let $M^{(i)}(x)$ be the minimal polynomial of α^i. Let $\overline{\mathscr{B}}(m,\lambda)$ be the nonprimitive BCH code of length $n = 2^m + 1$ having generator polynomial $g(x) = M^{(0)}(x)M^{(1)}(x)M^{(3)}(x)\cdots M^{(\lambda)}(x)$, for λ odd. Since $2^m = -1$ (modulo n), if β is a root of $g(x)$, so is β^{-1}. Thus $g(x)$ has roots α^i for $i = 0, \pm 1, \pm 2, \cdots, \pm(\lambda + 1)$ and so by the BCH bound, $\overline{\mathscr{B}}(m,\lambda)$ has minimum distance at least $2\lambda + 4$. Let $\mathscr{B}(m,\lambda)$ denote the punctured code obtained by deleting any parity-check symbol from $\overline{\mathscr{B}}(m,\lambda)$.

For $\lambda = 1$ it is not difficult to show that $\overline{\mathscr{B}}(m,1)$ contains exactly 2^{2m-2m} codewords for $m \geq 4$, and so $\mathscr{B}(m,1)$ is a

$$(2^m,2^{2m-2m},5)$$

code. This contains one more information symbol than the

corresponding

$$(2^m - 1, 2^{2^m-1-2m}, 5)$$

primitive BCH code.

Double-Error-Correcting Codes

Let N_r denote the length of the longest possible code of minimum distance 5 and redundancy r. We already have the following bounds.

i) $N_{4a+3} \geq 2^{2a+2} - 1$ for all $a \geq 1$, from the Preparata code $\mathscr{K}_{2a+2}$.

ii) $N_{2m} \geq 2^m + m - 1$ for all even $m \geq 4$, from the nonlinear codes constructed in Section I.

iii) $N_{4a+2} \geq 2^{2a+1}$ for all $a \geq 2$, from the BCH code $\mathscr{B}(2a + 1,1)$.

Some linear codes to be given in Section III show that

iv) $N_{4a+1} \geq 2^{2a} + 2^a - 1$ for all $a \geq 1$.

Let us apply Construction $X4$ with $\mathscr{C}_1 = \mathscr{K}_m$, $\mathscr{C}_2 =$ the Hamming code $\mathscr{H}_m$ (which is justified by Theorem 2), $\mathscr{C}_3 =$ the longest distance-6 code with redundancy m, with length $N_{m-1} + 1$, and $\mathscr{C}_4 =$ the even-weight code of length $N_{m-1} + 1$. The new code $\mathscr{C}_5$ is a nonlinear distance-5 code of length $2^m + N_{m-1}$ and redundancy $2m$, so

v) $N_{2m} \geq 2^m + N_{m-1}$ for all even $m \geq 4$.

Thus the Preparata code $\mathscr{K}_m$ has been extended by $\sqrt{n+1}$ symbols, of which only one is a check symbol.

With the same $\mathscr{C}_1$ and $\mathscr{C}_2$, $\mathscr{C}_3 =$ the longest distance-6 code with redundancy $m + 1$, with length $N_m + 1$, and $\mathscr{C}_4$ an even-weight code, we find

vi) $N_{2m+1} \geq 2^m + N_m$ for all even $m \geq 4$.

From i), iv), v), vi) we obtain

vii) $N_{4a} \geq 2^{2a} + 2^{a-1} + 2^{(a-1)/2} - 1$ for all $a \geq 2$, so that iii) and vii) together improve on ii), and

viii) $N_{8a} \geq 2^{4a} + 2^{2a} - 1$ for all $a \geq 1$,

ix) $N_{16a+1} \geq 2^{8a} + 2^{4a} + 2^{2a} - 1$ for all $a \geq 1$, and so on.

Table II gives N_r^*, the length of the longest code of minimum distance 5 and redundancy r presently known to us, for $6 \leq r \leq 35$. Included here are the $(19,2^{11},5)$ code constructed in Section I, the $(23,2^{14},5)$ quasi-perfect code found by Wagner [23], and the $(73,2^{60},5)$, $(277,2^{260},5)$ Srivastava codes found by Helgert [9]. The remaining codes are obtained from the preceding discussion. Most are nonlinear.

The results of Table II show that the bounds iv) and vii) can always be improved on.

Two distance-5 codes that do not appear in Table II are worth mentioning. They are the (11,24,5) Hadamard code with redundancy $r = 6.415\cdots$; and the $(20,5 \cdot 2^9,5)$ code with $r = 8.678\cdots$ obtained from Construction $X4$ by taking $\mathscr{C}_1 = \mathscr{K}_4$, $\mathscr{C}_2 = \mathscr{K}_4$ plus any other four of the eight cosets of $\mathscr{K}_4$ in $\mathscr{H}_4$, $\mathscr{C}_3 = \{00000,11111\}$, and $\mathscr{C}_4 = \{00000,11111,11000,00111,10100,01011,10010,01101,10001,01110\}$.

e-Error-Correcting Codes

Let us apply Construction $X4$ to extend e-error-correcting BCH codes of length $2^m - 1$, for $e \geq 1$. Let $m = \alpha e - \beta$,

TABLE II

N_r^*, THE LONGEST DOUBLE-ERROR-CORRECTING CODE OF REDUNDANCY r KNOWN TO US

r	N_r^*	r	N_r^*	r	N_r^*
6	8	16	271	26	2^{13}
7	15	17	277	27	$2^{14}-1$
8	19	18	512	28	$2^{14}+73$
9	23	19	$2^{10}-1$	29	$2^{14}+128$
10	32	20	$2^{10}+23$	30	2^{15}
11	63	21	$2^{10}+32$	31	$2^{16}-1$
12	70	22	2^{11}	32	$2^{16}+255$
13	73	23	$2^{12}-1$	33	$2^{16}+271$
14	128	24	$2^{12}+63$	34	2^{17}
15	255	25	$2^{12}+70$	35	$2^{18}-1$

where $0 \leq \beta < e$, and take $\mathscr{C}_1$, $\mathscr{C}_2$, and $\mathscr{C}_3$ equal to $(2^m - 1, 2^{2^m-em-1}, 2e + 1)$, $(2^m - 1, 2^{2^m-(e-1)m-1}, 2e - 1)$ and $(2^\alpha, 2^{2^\alpha-e\alpha-1}, 2e + 2)$ BCH codes. Take $\mathscr{C}_4$ equal to $\mathscr{C}_3$ plus $2^m - 1$ other cosets of $\mathscr{C}_3$ in the $(2^\alpha, 2^{2^\alpha-1}, 2)$ even-weight code.

Then the new code $\mathscr{C}_5$ is a

$$(2^m + 2^\alpha - 1, 2^{2^m+2^\alpha-em-\beta-2}, 2e + 1)$$

linear code of redundancy $em + \beta + 1$. In the most favorable case, when m is divisible by e and $\beta = 0$, $\mathscr{C}_1$ has been extended by $2^{m/e} \approx n^{1/e}$ symbols, of which only one is a check symbol. For large m this is an improvement over the Andryanov–Saskovets construction, which gives an extension by only $m \approx \log n$ symbols. In Section III it is shown that in effect triple-error-correcting BCH codes can be extended by about $2\sqrt{n}$ symbols, of which only two are check symbols.

Encoding and Decoding for Construction X4

a) Encoding: Let $\mathscr{C}_5$ be obtained from $\mathscr{C}_1$–$\mathscr{C}_4$ by Construction $X4$, where $\mathscr{C}_i$ is an $(n_i,2^{k_i},d_i)$ linear code for $i = 1$–5. As in the case of Construction X, we assume encoders and decoders are available, corresponding to generator matrices $G_1 = [A_1,I_{k_1}]$, $G_3 = [A_3,I_{k_3}]$,

$$G_2 = \begin{array}{|c|c|c|} \multicolumn{2}{c}{A_1} & I_{k_1} \\ \hline A_2 & I_{k_2-k_1} & 0 \end{array}$$

and

$$G_4 = \begin{array}{|c|c|c|} A_4 & I_{k_4-k_3} & 0 \\ \hline \multicolumn{2}{c}{A_3} & I_{k_3} \end{array}$$

for $\mathscr{C}_1$, $\mathscr{C}_3$, $\mathscr{C}_2$, and $\mathscr{C}_4$, respectively.

Let $(i_1,i_2,\cdots,i_{k_2+k_3})$ be the string of information symbols to be encoded by $\mathscr{C}_5$. Encoding is accomplished by feeding $(i_1,\cdots,i_{k_1+1},\cdots,i_{k_2})$ into Encoder$_2$ and $(i_{k_1+1},\cdots,i_{k_2+1},\cdots,i_{k_2+k_3})$ into Encoder$_4$, and concatenating the outputs. This corresponds to using the generator matrix

$$G_5 = \begin{array}{|c|c|c|c|c|c|} \hline A_1 & & I_{k_1} & 0 & & \\ \hline A_2 & I_{k_2-k_1} & 0 & A_4 & I_{k_4-k_3} & 0 \\ \hline 0 & & & A_3 & & I_{k_3} \\ \hline \end{array}$$

for $\mathscr{C}_5$. (Notice that $k_2 - k_1 = k_4 - k_3$ by hypothesis.)

b) Decoding: Decoding is then analogous to that for Construction X, and corrects up to $[\frac{1}{2}(d_5 - 1)]$ errors.

III. Construction $Y1$, $Y2$, $Y3$: Combining a Code and Its Dual

In this section we state three general constructions for linear and nonlinear codes due to Goethals [7]. These constructions are then applied to extended double- and triple-error-correcting BCH codes, to quadratic-residue (QR) codes, and to cyclic codes. First we state the constructions.

Construction Y1

Let $\mathscr{C}_1$ be an $(n,2^{k_1},d_1)$ linear code and let $\mathscr{C}_2$ be its $(n,2^{r_1},d_2)$ dual code, with coordinates chosen so that there is a minimum weight codeword $1\cdots10\cdots0$ in $\mathscr{C}_2$. Let S be the subgroup of $\mathscr{C}_1$ in which the first $d_2 - 1$ coordinates are zero. Then the d_2th coordinates of S are also zero. If the initial d_2 zeros are deleted from S we are left with an

$$(n - d_2,\ 2^{k_1-d_2+1},\ d_1)$$

linear code.

Construction Y2

Let T be the union of S and all of the $d_2 - 1$ cosets of S in $\mathscr{C}_1$ with coset leaders $110^{n-2},1010^{n-3},\cdots,10^{d_2-2}10^{n-d_2}$. By deleting the first d_2 coordinates of T we obtain an

$$(n - d_2,\ d_2 2^{k_1-d_2+1},\ d_1 - 2)$$

nonlinear code.

Construction Y3

Let U be the union of S and all of the $\binom{d_2}{2}$ cosets of S in $\mathscr{C}_1$ with coset leaders of weight 2. By deleting the first d_2 coordinates of U we obtain an

$$\left(n - d_2,\ \left(1 + \binom{d_2}{2}\right) 2^{k_1-d_2+1},\ d_1 - 4\right)$$

nonlinear code.

Applications

i) To Double-Error-Correcting BCH Codes[1]*:* Let $\mathscr{C}_1$ be a $(2^{m+1},2^{2^{m+1}-2m-3},6)$ extended double-error-correcting BCH code. From Berlekamp [2, table 16.5], the dual code $\mathscr{C}_2$ has minimum distance d_2 equal to

$$2^m - 2^{(m+1)/2}, \qquad m \text{ is odd}$$

or

$$2^m - 2^{m/2}, \qquad m \text{ is even.}$$

From Construction $Y1$, we obtain linear codes with minimum distance 5, redundancy $2m + 1$, and length $2^m + 2^{(m+1)/2} - 1$ if m is odd, or length $2^m + 2^{m/2} - 1$ if m is even, for all $m \geq 3$. See Section II for a comparison of these with other distance-5 codes.

ii) To Triple-Error-Correcting BCH Codes[1]*:* Now choose $\mathscr{C}_1$ to be a $(2^{m+1},2^{2^{m+1}-3m-4},8)$ BCH code. Again using [2, table 16.5], we obtain linear codes with minimum distance 7, redundancy $3m + 2$, and length $2^m + 2^{(m+2)/2} - 1$ if m is even, or length $2^m + 2^{(m+3)/2} - 1$ if m is odd, for all $m \geq 3$.

Thus we have in effect extended these codes by about $2\sqrt{n}$ symbols, of which only two are check symbols. Unfortunately the method cannot be applied to BCH codes of distance greater than 8, because for such codes the exact minimum weight of the dual is not known.

iii) To QR and Cyclic Codes: We apply Constructions $Y1$, $Y2$, and $Y3$ to the table[2] of QR codes on [2, p. 360], and to Chen's table of cyclic codes [5], [6]. The new codes obtained are shown in Table III. In the table an $(n,2^{k_1},d_1)$ code of the given type is combined with its $(n,2^{r_1},d_2)$ dual code, to produce a new (n',M,d') code. The table is arranged in order of increasing d'.

IV. A Generalization of the Golay Code: Combining Two Codes

Turyn [1] showed in 1967 that the length 24 Golay code can be obtained by combining two first-order Reed–Muller codes. This technique is used here to give a simple construction of an infinite family of linear codes with $d/n = \frac{1}{3}$. The first three codes of the family are the $(24,2^{12},8)$ Golay code, a $(48,2^{15},16)$ code, and a $(96,2^{18},32)$ code. For large block lengths the rate approaches zero.

Let $n = 2^m - 1$, and let α be a primitive element of the field $F = GF(2^m)$. Let $M^{(i)}(x)$ be the minimal polynomial of α^i.

The roots of $M^{(1)}(x)$ are $\alpha,\alpha^2,\alpha^4,\cdots,\alpha^{2^{m-1}}$, and so $M^{(1)}(\alpha^{-1}) \neq 0$. Thus $M^{(1)}(x)$ and $M^{(-1)}(x)$ are distinct. Let $\mathscr{C}_1$ and $\mathscr{C}_{-1}$ be the $(2^m - 1,2^m,2^{m-1})$ codes having check polynomials $M^{(1)}(x)$ and $M^{(-1)}(x)$, respectively. (Since the codewords are the coordinate vectors of a simplex inscribed in the n-cube, these are called simplex codes.)

Let $\mathscr{C}_2$ be the $(2^m,2^{m+1},2^{m-1})$ first-order Reed–Muller code obtained from $\mathscr{C}_1$ by including the complements of all codewords and adding an overall parity check. Similarly, let $\mathscr{C}_{-2}$ be obtained from $\mathscr{C}_{-1}$.

Theorem 4: The code $\mathscr{C}$ consisting of all codewords $(a + x,b + x,a + b + x)$, where $a \in \mathscr{C}_2$, $b \in \mathscr{C}_2$, $x \in \mathscr{C}_{-2}$, is a

$$(3 \cdot 2^m,2^{3m+3},2^m)$$

linear code, for $m = 3,4,5,\cdots$.

[1] Suggested by E. R. Berlekamp.

[2] The following corrections should be made to this table: for $n = 73$, $d \leq 13$; delete the line $n = 97$; for $n = 103$, $d = 19$; for $n = 113$, $d = 15$; for $n = 151$, $d = 19$; for $n = 191$, $d \leq 27$; for $n = 223$, $d \leq 31$.

TABLE III
NEW (n', M, d') CODES FROM CONSTRUCTIONS *Y1*, *Y2*, *Y3*

n_1	k_1	d_1	k_2	d_2	n_4	d_4
63	46	7	49	5	67	7
63	36	11	42	7	74	11
31	6	15	11	11	41	15
63	28	15	30	13	66	15
63	21	18	24	14	66	17
63	17	20	20	18	66	19
63	19	19	21	16	68	19
63	19	19	22	15	70	19
63	19	19	25	15	74	19
127	71	19	78	15	139	19
63	18	21	24	15	77	21
63	11	26	17	22	73	25
63	10	27	16	23	74	27
127	50	27	57	23	139	27
63	7	31	13	24	78	29

Proof: It is easy to see that the code is linear and contains 2^{3m+3} codewords. Let w be the weight of $(a + x, b + x, a + b + x)$, and for convenience let us write u for 2^{m-1}. Each of a, b, and x has weight O, u, or 2^m. We consider three cases. i) If at most one of a, b, and x has weight u, then $w = 0$ or $w \geq 2^m$. ii) If two of a, b, and x have weight u, then we find that $w \geq \min\{2^m, u + 2d_1\}$, where $d_1 = \min \text{ wt } \{a + x \mid a \in \mathscr{C}_2, x \in \mathscr{C}_{-2}, a \neq 0, x \neq 0\}$. By Corollary 6 following, $d_1 \geq 2^{m-1} - 2^{m/2} - \frac{1}{2}$. Direct calculation shows that for $m = 3$, when $\mathscr{C}_2$ and $\mathscr{C}_{-2}$ are extended Hamming codes, $d_1 = 2$. Hence $w \geq 2^m$ holds for $m \geq 3$. iii) Suppose wt (a) = wt (b) = wt $(x) = u$ and $a = b$. Then $w \geq u + 2d_1 \geq 2^m$ for $m \geq 3$. For the remaining case when $a \neq b$ we use the following lemma.

Lemma—[1]: For any binary vectors a,b,x,

$$\text{wt } |a + x| + \text{wt } |b + x| + \text{wt } |a + b + x| \geq 2 \text{ wt } |a + b + ab| - \text{wt } |x|.$$

The proof is straightforward and is omitted.

Returning to the proof of the theorem, in the last case we have wt $|a + b + ab| = 3 \cdot 2^{m-2}$ and wt $|x| = u$, so by the lemma $w \geq 2^m$. Q.E.D.

For example, when $m = 3$, $\mathscr{C}$ is a $(24,2^{12},8)$ linear code, which must be the Golay code since Pless [15] has shown that code to be unique.

It remains to prove Theorem 5 and its corollary.

Theorem 5: Let $d_{\mu\nu}$ be the Hamming distance between the μth codeword of $\mathscr{C}_1$ and the νth codeword of $\mathscr{C}_{-1}$. Then either both codewords are zero or

$$\tfrac{1}{2}n - 2^{m/2} \leq d_{\mu\nu} \leq \tfrac{1}{2}n + 2^{m/2}.$$

Proof: The proof uses the Carlitz–Uchiyama [3] bound for Kloosterman sums. Let

$$T(x) = \sum_{r=0}^{m-1} x^{2^r}.$$

Then it is known (see, for example, Solomon [21, p. 252]) that the polynomial $f_1(x) = T(\mu x)$, for μ any element of F, has the property that $f_1(\alpha^{-l})$ is the lth bit, $c_1(\mu,l)$, say, of the μth codeword of $\mathscr{C}_1$. Similarly if $f_{-1}(x) = T(\nu x)$, for ν any element of F, $f_{-1}(\alpha^l)$ is the lth bit, $c_{-1}(\nu,l)$ say, of the νth codeword of $\mathscr{C}_{-1}$. (This sets up a one–one correspondence between F and the codewords of $\mathscr{C}_1$ and of $\mathscr{C}_{-1}$. The polynomials f_1 and f_{-1} are the Mattson–Solomon polynomials.)

Then

$$S_{\mu\nu} \triangleq \sum_{l=0}^{2^m-2} (-1)^{c_1(\mu,l)+c_{-1}(\nu,l)} = n - 2d_{\mu\nu},$$

since the sum contributes $+1$ when the codewords agree and -1 when they disagree.

$$S_{\mu\nu} = \sum_{l=0}^{2^m-2} (-1)^{T(\mu\alpha^{-l}+\nu\alpha^l)} = \sum_{\beta \in F, \beta \neq 0} (-1)^{T(\beta+\gamma/\beta)},$$

where $\beta = \mu\alpha^{-l}$, $\gamma = \nu\mu$. The last expression is a Kloosterman sum and it is shown in [3] that $|S_{\mu\nu}| \leq 2^{\frac{1}{2}m+1}$. The theorem then follows. An immediate consequence of Theorem 5 is the following.

Corollary 6: Let

$$d_1 = \min \text{ wt } \{a + x \mid a \in \mathscr{C}_2, x \in \mathscr{C}_{-2}, a \neq 0, x \neq 0\}.$$

Then $d_1 \geq 2^{m-1} - 2^{m/2} - \frac{1}{2}$.

Remark

It is perhaps worth pointing out that the code $\mathscr{C}$ constructed in Theorem 4 is a union of cosets of the direct-product code $\mathscr{C}_2 \times \mathscr{C}_3$, where $\mathscr{C}_3$ is the code $\{000,111\}$. (For the definition of a direct-product code, see [14, p. 81] or [2, p. 339].)

ACKNOWLEDGMENT

This paper owes a lot to our friends. E. R. Berlekamp contributed some of the constructions of Section II and the idea of using Kloosterman sums in Section IV; F. P. Preparata helped with the proof of Theorem 2; and conversations with F. J. MacWilliams and C. W. Hoffner II were also very helpful.

Note Added in Proof: A different method of extending Preparata codes from the ones given here has been described by H. Miyakawa, H. Imai, and I. Nakajima, "Modified Preparata codes—Optimum systematic nonlinear double-error-correcting codes," *Electron. Commun. Japan*, vol. 53A, pp. 25–32, 1970. However, their method does not appear to be as powerful as ours, because for example they obtain only a $(266,2^{250},5)$ code, whereas we construct a $(271,2^{255},5)$ code with the same redundancy.

REFERENCES

[1] E. F. Assmus, Jr., H. F. Mattson, Jr., and R. J. Turyn, "Research to develop the algebraic theory of codes," Air Force Cambridge Res. Lab., Bedford, Mass., Sci. Rep. AFCRL-67-0365, 1967.
[2] E. R. Berlekamp, *Algebraic Coding Theory*. New York: McGraw-Hill, 1968.
[3] L. Carlitz and S. Uchiyama, "Bounds for exponential sums," *Duke Math. J.*, vol. 24, pp. 37–41, 1957.

[4] R. D. Carmichael, *Introduction to the Theory of Groups of Finite Order*. New York: Dover, 1956.
[5] C. L. Chen, "Some results on algebraically structured error-correcting codes," Ph.D. dissertation, Dep. Elec. Eng., Univ. Hawaii, Honolulu, Jan. 1969.
[6] ——, "Computer results on the minimum distance of some binary cyclic codes," *IEEE Trans. Inform. Theory* (Corresp.), vol. IT-16, pp. 359–360, May 1970.
[7] J. M. Goethals, "On the Golay perfect binary code," *J. Comb. Theory*, vol. 11A, pp. 178–186, Sept. 1971.
[8] H. D. Goldman, M. Kliman, and H. Smola, "The weight structure of some Bose-Chaudhuri codes," *IEEE Trans. Inform. Theory* (Corresp.), vol. IT-14, pp. 167–169, Jan. 1968.
[9] H. J. Helgert, "Srivastava codes," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 292–297, Mar. 1972.
[10] H. J. Helgert and R. D. Stinaff, "Minimum distance bounds for binary linear codes," to be published.
[11] C. W. Hoffner II, "Equivalence class decoding," Ph.D. dissertation, Dep. Elec. Eng., Univ. Iowa, Iowa City, Jan. 1970.
[12] S. M. Johnson, "On upper bounds for unrestricted binary error-correcting codes," *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 466–478, July 1971.
[13] D. Knee and H. D. Goldman, "Quasi-self-reciprocal polynomials and potentially large minimum distance BCH codes," *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 118–121, Jan. 1969.
[14] W. W. Peterson, *Error-Correcting Codes*. Cambridge, Mass.: M.I.T. Press, 1961.
[15] V. Pless, "On the uniqueness of the Golay code," *J. Comb. Theory*, vol. 5, pp. 215–228, Nov. 1968.
[16] F. P. Preparata, "Weight and distance structure of Nordstrom–Robinson quadratic codes," *Inform. Contr.*, vol. 12, pp. 466–473, May–June 1968.
[17] F. P. Preparata, "A class of optimum nonlinear double-error-correcting codes," *Inform. Contr.*, vol. 13, pp. 378–400, Oct. 1968.
[18] S. M. Reddy, "On decoding iterated codes," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 624–627, Sept. 1970.
[19] N. J. A. Sloane, "A survey of constructive coding theory, and a table of binary codes of highest known rate," *Discrete Math.*, to be published.
[20] N. J. A. Sloane and D. S. Whitehead, "A new family of single-error correcting codes," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 717–719, Nov. 1970.
[21] G. Solomon, "Algebraic coding theory," in *Communication Theory*, A. V. Balakrishnan, Ed. New York: McGraw-Hill, 1968, ch. 6.
[22] Yu. L. Vasil'ev, "O negruppovykh plotno upakovannykh kodakh," *Probl. Kibern.*, vol. 8, pp. 337–339, 1962.
[23] T. J. Wagner, "A search technique for quasi-perfect codes," *Inform. Contr.*, vol. 9, pp. 94–99, 1966.

Correspondence

Coding Theorem and Its Converse for Continuous Incrementally Stationary Channels With Finite Incremental Memory

T. T. KADOTA

Abstract—**We define continuous incrementally stationary channels with finite incremental input- and output-memories and prove the coding theorem and its converse for such channels. These channels include, as special cases, stationary channels with finite input- and output-memories and incrementally stationary and memoryless channels. The former is defined here and the latter has been defined previously. It is emphasized that, with an elementary measure-theoretic formulation, the standard method of proving the coding theorem for discrete channels becomes directly applicable for continuous channels. Consequently, the tedious step of representing a continuous channel by an infinite series of discrete channels can be avoided entirely.**

I. Introduction

Mutual information has long been defined for continuous channels where the input and the output are functions [1]–[3]. Yet the coding theorem and its converse have been proved primarily for discrete channels where the input and the output are sequences [4], [5]. In a few specific cases they have been proved by first representing the continuous channel by an infinite series of discrete channels [5, pp. 355–441], [6]. In this correspondence we note that, once continuous channels are properly characterized, the standard method of proving the coding theorem and its converse for discrete channels becomes directly applicable. Thus, we can avoid the tedious representation by infinite series and an additional continuity condition for the convergence. Of course, some elementary use of measure theory is necessary both for characterizing channels and for proving the coding theorem.

We define the class of continuous incrementally stationary channels with finite incremental memory. This class of channels includes, as special cases, the stationary channel with finite memory and the incrementally stationary and memoryless channel. The former is a generalization of the discrete stationary channel with finite memory [4] and is defined for the first time here. The latter is not a generalization of the discrete stationary memoryless channel and has been defined previously [7]. Mathematical definitions of these continuous channels are given in Section II. The coding theorem and its converse are stated in Section III, where in the remarks we give a comprehensive interpretation of the theorem and the converse. The explicit proof is omitted due to space limitations. For the benefit of mathematically inclined readers, however, we have made copies of the proof available.[1]

Before we begin the mathematical presentation, it is instructive to give a heuristic characterization of continuous channels by using simple examples. First, consider the following additive-noise channel:

$$y(t) = x(t) + v(t), \qquad -\infty < t < \infty, \tag{1}$$

where x and y are the input and the output of the channel and v is the additive noise. Observe that the probability distribution of the output for each fixed input, i.e., the transitional measure of the channel, is the noise probability distribution with its mean shifted by the signal. Suppose v is a stationary zero-mean

Manuscript received December 28, 1971.
The author is with the Bell Telephone Laboratories, Inc., Murray Hill, N.J. 07974.

[1] Shortly after this correspondence was written, we proved the coding theorem and the converse for a more general channel [8]. This is another reason for omitting the proof here.