On the Preparata and Goethals codes

of the codes, it is only the B_j to which Theorem 2 applies directly.

References

[1] C. W. Curtis and I. Reiner, *Representation Theory of Finite Groups and Associative Algebras*. New York: Wiley, 1962.
[2] P. Delsarte and R. J. McEliece, "Zeros of functions in finite Abelian group algebras," *Amer. J. Math.*, vol. 26, pp. 145–153, 1971.
[3] L. Dornhoff, *Group Representation Theory*. New York: Dekker, 1971.
[4] B. Huppert, *Endliche Gruppen I*. Berlin: Springer-Verlag, 1967.
[5] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam: North-Holland, 1977.
[6] J. Milnor and D. Husemoller, *Symmetric Bilinear Forms*. New York: Springer-Verlag, 1973.
[7] N. J. A. Sloane, "Self-dual codes and lattices," in *Relations Between Combinatorics and Other Parts of Mathematics* (Proc. Symp. Pure Math. 34). Providence, RI: Amer. Math. Soc., 1979, pp. 273–308.
[8] H. N. Ward, "Combinatorial polarization," *Discrete Math.*, vol. 26, pp. 185–197, 1979.
[9] ——, "Divisible codes," *Arch. Math.* (Basel), vol. 36, pp. 485–494, 1981.
[10] ——, "Multilinear forms and divisors of codeword weights," *Quart. J. Math. Oxford Ser.* (2), vol. 34, no. 133, 1983.

On the Preparata and Goethals Codes

RONALD D. BAKER, JACOBUS H. VAN LINT, AND RICHARD M. WILSON

DEDICATED TO JESSIE MACWILLIAMS ON THE OCCASION OF HER RETIREMENT FROM BELL LABORATORIES

Abstract—**Simple descriptions of Preparata and Goethals codes are provided.**

I. Introduction

In a paper on the partitioning of affine planes [1] the first author pointed out that some of his methods also lead to a simple description of the *Preparata codes* (cf. [2], [6], [7]). Since the known descriptions of these codes involve rather messy calculations it seems worthwhile to give this simple description. We shall show that the same ideas can be used to treat the *Goethals codes* (cf. [3], [6]). Several authors have observed that a Hamming code can be partitioned into extended Preparata codes (cf. [1], [8]). The methods of this paper allow us to also show this fact in a simple way.

II. Preparata Codes

In the following m is odd ($m \geqslant 3$), $n = 2^m - 1$. Let $\mathbb{F}$ be the field $GF(2^m)$ and let $x \mapsto x^\sigma$ be an automorphism of $\mathbb{F}$, i.e., σ is a power of 2. We require that both $x \mapsto x^{\sigma+1}$ and $x \mapsto x^{\sigma-1}$ are one-to-one mappings, i.e., $(\sigma \pm 1, 2^m - 1) = 1$. (This is true, for example, for $\sigma = 2$.)

For the admissible values of σ we shall define a code $\overline{\mathcal{P}}(\sigma)$ of length $2n + 2 = 2^{m+1}$. The codewords will be described by pairs (X, Y), where $X \subset \mathbb{F}$, $Y \subset \mathbb{F}$. As usual we interpret the pair (X, Y) as the corresponding pair of characteristic functions, i.e., as a (0,1)-vector of length 2^{m+1}. We shall let the zero element of $\mathbb{F}$ correspond to the first position in the X-part.

Definition 1: The extended *Preparata code* $\overline{\mathcal{P}}(\sigma)$ of length 2^{m+1} consists of the codewords described by all pairs (X, Y) satisfying

a) $|X|$ is even, $|Y|$ is even,

b) $\sum_{x \in X} x = \sum_{y \in Y} y$,

c) $\sum_{x \in X} x^{\sigma+1} + \left(\sum_{x \in X} x\right)^{\sigma+1} = \sum_{y \in Y} y^{\sigma+1}$.

The code $\mathcal{P}(\sigma)$ is obtained by deleting the first coordinate.

Remark: It is not difficult to check that the usual complicated definition of the Preparata codes (cf. [2]) actually coincides with Definition 1 for $\sigma = 2$.

For a discussion of the properties of these codes we make the following conventions concerning notation. The symmetric difference of two sets X_1, X_2 is denoted by $X_1 \triangle X_2$ (this corresponds to addition of codewords). The

Manuscript received March 4, 1982; revised May 17, 1982. This research was supported in part by NSF Grant MCS 7821599. This work was presented at the Oberwolfach Meeting on Information Theory, April 4–10, 1982.

R. D. Baker is with the Department of Mathematics, North Carolina State University, Box 5548, Raleigh, NC 27650.

J. H. van Lint and R. M. Wilson are with the Department of Mathematics, California Institute of Technology, Pasadena, CA 91109.

0018-9448/83/0500-0342$01.00

set $\{x+\alpha | x \in X\}$ is denoted by $X+\alpha$. Many of the calculations depend on the following equality:

$$(a+b)^{\sigma+1} = a^{\sigma+1} + a^{\sigma}b + ab^{\sigma} + b^{\sigma+1}. \qquad (1)$$

We shall show that the Preparata codes are nearly perfect and hence completely regular. The weaker assertion that $\overline{\mathcal{P}}(\sigma)$ is distance invariant can be proved directly.

Theorem 1: The code $\overline{\mathcal{P}}(\sigma)$ is distance invariant.

Proof: We compare a codeword (X_0, Y_0) with $(\varnothing, \varnothing) = \mathbf{0}$. Let $\alpha = \sum_{x \in X_0} x$. The mapping $(X, Y) \mapsto (U, V)$, where $U = (X \triangle X_0) + \alpha$, $V = Y \triangle Y_0$ is clearly one-to-one. We show that if (X, Y) is a codeword then so is (U, V) and vice versa. For Definition 1 a) and b) this is trivial. We check Definition 1 c). Using (1) we find

$$\sum_{x \in U} x^{\sigma+1} + \Big(\sum_{x \in U} x\Big)^{\sigma+1}$$
$$= \sum_{x \in X} (x+\alpha)^{\sigma+1} + \sum_{x \in X_0} (x+\alpha)^{\sigma+1} + \Big(\sum_{x \in X} x + \alpha\Big)^{\sigma+1}$$
$$= \sum_{x \in X} x^{\sigma+1} + \sum_{x \in X_0} x^{\sigma+1} + \Big(\sum_{x \in X} x\Big)^{\sigma+1} + \alpha^{\sigma+1}$$
$$= \sum_{y \in Y} y^{\sigma+1} + \sum_{y \in Y_0} y^{\sigma+1} = \sum_{y \in V} y^{\sigma+1}$$

□

The proofs of the main properties of these codes become simpler if we first find some automorphisms of the codes.

Theorem 2: The group Aut $\overline{\mathcal{P}}(\sigma)$ *contains the permutations*

a) $(X, Y) \mapsto (X+c, Y+c), \quad c \in \mathbb{F}$,
b) $(X, Y) \mapsto (Y, X)$,
c) $(X, Y) \mapsto (\alpha X, \alpha Y), \quad \alpha \in \mathbb{F}^*$,
d) $(X, Y) \mapsto (X^{\varphi}, Y^{\varphi}), \quad \varphi \in \operatorname{Aut}\mathbb{F}$.

Proof: In the case of a) one checks Definition 1 c) using (1). All the other properties are trivially true. □

We remark that the permutations a) and b) generate all the translations of the $(m+1)$-dimensional vector space $V = \mathbb{F} \oplus \mathrm{GF}(2)$. The complete group Aut $\overline{\mathcal{P}}(\sigma)$ was determined by W. M. Kantor [4].

Theorem 3: $\overline{\mathcal{P}}(\sigma)$ has minimum distance 6.

Proof: By Theorem 1 it is sufficient to show that the minimum weight is 6. There are obviously no words of weight 2. So we must show that weight 4 can not occur. There are two cases:

1) If $(\{x_1, x_2\}, \{y_1, y_2\})$ is a codeword we may assume that $x_1 = 0$ (by Theorem 2). Then Definition 1 c) yields
$$y_1^{\sigma+1} + y_2^{\sigma+1} = 0,$$
and then the condition on σ implies that $y_1 = y_2$, a contradiction.

2) By Theorem 1 and Theorem 2 it remains to check the possibility $|X| = 4$, $Y = \varnothing$, where $X = \{0, a, b, c\}$. From Definition 1 b) and c) we find
$$a+b+c = 0,$$
$$a^{\sigma+1} + b^{\sigma+1} + c^{\sigma+1} = 0.$$
Substituting the first equation into the second, using (1), yields $ab(a^{\sigma-1} + b^{\sigma-1}) = 0$, i.e., $a = b$, a contradiction; (here we use the fact that $x \mapsto x^{\sigma-1}$ is one-to-one). Finally we show that there are indeed codewords of weight 6. Given a, b, c (distinct), define y by $y^{\sigma+1} = a^{\sigma+1} + b^{\sigma+1} + c^{\sigma+1}$ and define x by $x = a+b+c+y$. Then $(\{0, x\}, \{a, b, c, y\})$ is a codeword. □

From the original definition of the Preparata codes one immediately finds the number of codewords. In the case of Definition 1 this is more difficult.

Theorem 4: $|\overline{\mathcal{P}}(\sigma)| = 2^k$, where $k = 2^{m+1} - 2m - 2$.

Proof: In Definition 1 we can choose the set X in 2^n ways, satisfying a). We count how many sets $Y \subset \mathbb{F}^*$ satisfy b) and c) and to each such set add the element 0 if necessary to satisfy a). Let ω be a primitive element of $\mathbb{F}$ and $m_i(x)$ the minimal polynomial of ω^i. The two equations, Definition 1 b) and c), for the elements y are equations over $\mathbb{F}$. Considering $\mathbb{F}$ as m-dimensional space over GF(2) these become $2m$ linear equations over GF(2). We claim that these equations are independent. This is so because $(\sigma+1, n) = 1$ and hence $m_{\sigma+1}(x)$ has degree m, i.e., the cyclic code over GF(2) with length n and generator $m_1(x)m_{\sigma+1}(x)$ has dimension $n - 2m$. It follows that for each choice of X the equations, Definition 1 b) and c), have 2^{n-2m} solutions Y with $Y \subset \mathbb{F}^*$. This proves our assertion. □

From Theorem 3 and Theorem 4 it follows that the codes $\mathcal{P}(\sigma)$ are nearly perfect (cf. [2]). In the next section we shall show that the extended Hamming code is a union of translates of $\overline{\mathcal{P}}(\sigma)$.

The arguments above do not show that different values of σ produce different codes. However it was shown by W. M. Kantor that $\overline{\mathcal{P}}(\sigma)$ and $\overline{\mathcal{P}}(\tau)$ are equivalent if and only if $\sigma = \tau$ or $\sigma\tau = 2^m$ (cf. [4]).

III. A Partition of the Hamming Code into Translates of $\overline{\mathcal{P}}(\sigma)$.

We define a number of translates of $\overline{\mathcal{P}}(\sigma)$ as follows. Let $\mathcal{C}_0 = \overline{\mathcal{P}}(\sigma)$ and if $\alpha \in \mathbb{F}^*$ then let $\mathcal{C}_\alpha$ be the code obtained by adding the word corresponding to $(\{0, \alpha\}, \{0, \alpha\})$ to the codewords of $\overline{\mathcal{P}}(\sigma)$.

Lemma 1: The code $\mathcal{C}_\alpha$ has minimum weight 4, ($\alpha \in \mathbb{F}^*$).

Proof: By Theorem 1 and Theorem 3 we only have to show that no word has weight 2. Weight 2 is possible only if $\overline{\mathcal{P}}(\sigma)$ contains a word of the form $(\{0, \alpha\}, \{0, \alpha, \beta, \gamma\})$. By Definition 1 b) this is not so. □

By Theorem 3 the codes $\mathcal{C}_\alpha$, where $\alpha \in \mathbb{F}$, are pairwise disjoint. We define

$$\mathcal{H} = \bigcup_{\alpha \in \mathbb{F}} \mathcal{C}_\alpha. \tag{2}$$

From Theorem 4 we find $|\mathcal{H}| = 2^m|\overline{\mathcal{P}}(\sigma)| = 2^{2n-m}$ which is the cardinality of the extended Hamming code of length $2n+2$.

Lemma 2: $\mathcal{H}$ is a linear code.

Proof: Let (X_1, Y_1) and (X_2, Y_2) be codewords in $\overline{\mathcal{P}}(\sigma)$ and let $\alpha \in \mathbb{F}$, $\beta \in \mathbb{F}$. We define $s_i = \sum_{x \in X_i} x$ $(i = 1,2)$. For $\gamma \in \mathbb{F}$ we define X and Y by

$$X \triangle \{0\} \triangle \{\gamma\} = X_1 \triangle X_2 \triangle \{\alpha\} \triangle \{\beta\},$$

$$Y \triangle \{0\} \triangle \{\gamma\} = Y_1 \triangle Y_2 \triangle \{\alpha\} \triangle \{\beta\}.$$

We must show that there is a choice for γ such that $(X, Y) \in \overline{\mathcal{P}}(\sigma)$. For each choice of γ the sets X and Y satisfy Definition 1 a) and b). Substitution in Definition 1 c) yields the equation

$$(s_1 + s_2 + \alpha + \beta + \gamma)^{\sigma+1} = s_1^{\sigma+1} + s_2^{\sigma+1},$$

which has a unique solution γ. □

Theorem 5: $\mathcal{H}$ is the extended Hamming code of length 2^{m+1}.

Proof: $\mathcal{H}$ is linear and it has the required minimum distance and cardinality. □

We remark that the fact that $\mathcal{P}(\sigma)$ is nearly perfect with minimum distance 5 and wordlength $\equiv 0 \pmod 3$ implies that we can obtain the Hamming code of the same length by taking the Preparata code and all words which have distance 3 from this code (cf. [2], [8]).

IV. Two Lemmas on Cyclic Codes

At the Oberwolfach Meeting on Information Theory in April 1982 a new bound for the minimum distance of cyclic codes was presented by C. Roos. It turned out that the following two lemmas are both applications of this bound. They have been included as examples in the paper by C. Roos ([9] elsewhere in this issue). As a consequence the proofs which we give below are too difficult for the present problem, but since they are of independent interest we have not changed them.

Let $m = 2t+1$ and let ω be a primitive element of $\mathbb{F}$. Let $\rho = 2^{t-1}$, $\sigma = 2^t$, $r = \rho + 1$, $s = \sigma + 1$.

Lemma 3: The cyclic code $\mathcal{D}$ of length n, generated by $m_r(x)m_s(x)$, has minimum distance at least 5.

Proof: Among the zeros of any codeword we find ω^j with j respectively r, s, $2r = s+1$, $s \cdot 2^{t+1} = 2s - 1, 2s$, $r \cdot 2^{t+2} = 2^{t+2} + 1$. The values s, $s+1$, and $2s-1, 2s$ show, using the Hartmann–Tzeng bound (cf. [5]), that $\mathcal{D}$ has minimum distance at least 4. We now follow an idea of Goethals [3]. Suppose $x^i + x^j + x^k + x^l$ is a codeword of weight 4 in $\mathcal{D}$. Let $S = \{\omega^i, \omega^j, \omega^k, \omega^l\}$ and let $\langle S \rangle$ be the linear space spanned by S. Define the linearized polynomial $\pi(y)$ by

$$\pi(y) = \prod_{\xi \in \langle S \rangle} (y - \xi) = \sum_{u=a}^{4} \pi_u \cdot y^{2^u}, \qquad a \geqslant 0, \pi_a \neq 0.$$

Define S_ν by

$$S_\nu = \sum_{\xi \in S} \xi^{1+2^\nu}.$$

We saw above that $S_\nu = 0$ for $\nu = t-1, t, t+1, t+2$. From the equation

$$0 = \sum_{\xi \in S} \xi (\pi(\xi))^{2^b} = \sum_{u=a}^{4} \pi_u S_{b+u},$$

it follows that if $S_\nu = 0$ for four consecutive values of ν then $S_\nu = 0$ for all values of ν. This would imply that $S_0 = S_1 = S_2 = 0$, i.e., $\omega^2, \omega^3, \omega^5$ are zeros of our codeword, contradicting the Bose–Chaudhuri–Hocquenghem bound. □

Lemma 4: The cyclic code $\mathcal{D}'$ of length n, generated by $m_1(x)m_r(x)m_s(x)$, has minimum distance at least 7.

Proof: First suppose that there is a codeword of weight 6, say with nonzero coordinates in the positions α, β, γ, δ, ϵ, μ, where $\mu = \alpha + \beta + \gamma + \delta + \epsilon$. Then we have $\alpha^r + \beta^r + \gamma^r + \delta^r + \epsilon^r = \mu^r$ and a similar equation for the exponent s. An easy calculation shows that this implies that the word with nonzero coordinates in the positions $\alpha + \epsilon$, $\beta + \epsilon$, $\gamma + \epsilon$, $\delta + \epsilon$, $\alpha + \beta + \gamma + \delta$ also belongs to the code. By Lemma 3 it is now sufficient to show that $\mathcal{D}'$ does not have minimum distance 5. In the same way as in the proof of Lemma 3 we assume that there is a codeword of weight 5 and look at the space spanned by the nonzero coordinate positions. Because $m_1(x)$ divides the generator, this space has dimension at most 4. The proof of Lemma 3 shows that this leads to a contradiction. □

V. The Goethals Codes

We use the notation of the previous section. The Goethals code is the intersection of $\overline{\mathcal{P}}(\rho)$ and $\overline{\mathcal{P}}(\sigma)$, without the restriction on ρ and σ which we made in Section II. A direct definition analogous to Definition 1 is the following.

Definition 2: The *Goethals code* $\mathcal{G}$ of length 2^{m+1} consists of the codewords described by all pairs (X, Y) satisfying

a) $|X|$ is even, $|Y|$ is even,

b) $\sum_{x \in X} x = \sum_{y \in Y} y,$

c) $\sum_{x \in X} x^r + \left(\sum_{x \in X} x\right)^r = \sum_{y \in Y} y^r,$

d) $\sum_{x \in X} x^s + \left(\sum_{x \in X} x\right)^s = \sum_{y \in Y} y^s.$

From the proof of Theorem 1 we see that $\mathcal{G}$ is also distance invariant. The automorphisms of Theorem 2 are clearly also in Aut $\mathcal{G}$.

Theorem 6: $\mathcal{G}$ has minimum distance 8.

Proof: Again it is sufficient to show that $\mathcal{G}$ has minimum weight 8. By Theorem 3 there are only two possibilities which we must consider. The first of these is $X = \varnothing$, $|Y| \geqslant 6$. In this case Y corresponds to a codeword in $\mathfrak{D}'$, so $|Y| \geqslant 8$ by Lemma 4. The second possibility is $|X| = 2$, $|Y| \geqslant 4$. The automorphisms a) and c) of Theorem 2 show that we may assume without loss of generality that $X = \{0, 1\}$. From Definition 2 c) and d) we find that Y corresponds to a codeword in $\mathfrak{D}$, i.e., $|Y| \geqslant 6$ by Lemma 3. Finally we observe that $|X| = |Y| = 4$ is possible by taking $X = Y = \{0, \alpha, \beta, \alpha + \beta\}$. □

To find the cardinality of $\mathcal{G}$ we can use exactly the same method as in the proof of Theorem 4. Since $(n, r) = (n, s) = 1$ the polynomials $m_r(x)$ and $m_s(x)$ have degree m. Hence $\mathfrak{D}'$ has dimension $n - 3m$. The argument of Theorem 4 now shows that $|\mathcal{G}| = 2^l$, where $l = 2^{m+1} - 3m - 2$.

References

[1] R. D. Baker, "Partitioning the planes $AG_{2m}(2)$ into 2-designs," *Discrete Math.*, vol. 15, pp. 205–211, 1976.
[2] P. J. Cameron and J. H. van Lint, "Graphs, codes and designs," *London Math. Soc. Lecture Note Series*, vol. 43, Cambridge Univ., 1980.
[3] J. M. Goethals, "Nonlinear codes defined by quadratic forms over GF(2)," *Inform. Contr.*, vol. 31, pp. 43–74, 1976.
[4] W. M. Kantor, "On the equivalence of generalized Preparata Codes," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 345–348, May 1983.
[5] J. H. van Lint, *Introduction to Coding Theory*. New York: Springer Verlag, 1982.
[6] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam: North-Holland, 1977.
[7] F. P. Preparata, "A class of optimum non-linear double-error-correcting codes," *Inform. Contr.*, vol. 13, pp. 378–400, 1968.
[8] G. V. Zaitsev, V. A. Zinovjev, and N. V. Semakov, "Interrelation of Preparata and Hamming codes and extension of Hamming codes to new double-error-correcting codes," B. N. Petrov and F. Csaki, Eds., in *Proc. 2nd Int. Symp. Inform. Theory*, Akadémiai Kiadó, Budapest, pp. 257–263, 1973.
[9] C. Roos, "A new lower bound for the minimum distance of a cyclic code," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 330–332, May 1983.

On the Inequivalence of Generalized Preparata Codes

WILLIAM M. KANTOR

DEDICATED TO JESSIE MACWILLIAMS ON THE OCCASION OF HER RETIREMENT FROM BELL LABORATORIES

***Abstract*—If m is odd and $\sigma \in \operatorname{Aut} GF(2^m)$ is such that $x \to x^{\sigma^2 - 1}$ is $1-1$, there is a $[2^{m+1} - 1, 2^{m+1} - 2m - 2]$ nonlinear binary code $P(\sigma)$ having minimum distance 5. All the codes $P(\sigma)$ have the same distance and weight enumerators as the usual Preparata codes (which rise as $P(\sigma)$ when $x^\sigma = x^2$). It is shown that $P(\sigma)$ and $P(\tau)$ are equivalent if and only if $\tau = \sigma^{\pm 1}$, and Aut $P(\sigma)$ is determined.**

I. Introduction

In [13], Preparata introduced a family of $[2^{m+1} - 1, 2^{m+1} - 2m - 2]$ nonlinear binary 2-error correcting codes, where m is odd and $m > 1$. These have remarkable combinatorial properties: they are nearly perfect codes (Goethals and Snover [7]; Cameron and van Lint [4, ch. 16]) and, in particular, they are uniformly packed (Semakov, Zinovjev, and Zaitsev [14]); they give rise to designs [14], [15], [7], [12, p. 473], [4, pp. 89–90]; and they produce parallelisms of the lines of $PG(m,2)$ [15]; [1]. The published descriptions of these codes [13], [15], [12, § 15.6], [4] are complicated and difficult to work with. Fortunately, Baker and Wilson [2] have found a relatively simple description which led to a generalization of Preparata's codes.

Let m be odd, $m > 1$, and let $\sigma \in \operatorname{Aut} GF(2^m)$, where $x \to x^{\sigma^2 - 1}$ is $1 - 1$. (Thus, if $x^\sigma = x^{2^i}$ for all x then i and m are relatively prime.) Baker and Wilson constructed a code $P(\sigma)$ having the same parameters as Preparata's codes (cf. (1)), and hence having the same combinatorial properties. Moreover, their description makes a group of $(2^m - 1)m$ automorphisms very visible. We will show that this group is precisely $\operatorname{Aut}(P(\sigma)$ when $m > 3$, and that two generalized Preparata codes $P(\sigma)$ and $P(\tau)$ are equivalent if and only if $\tau = \sigma^{\pm 1}$. Similar results are obtained for the extended codes $\bar{P}(\sigma)$ of length 2^{m+1}.

All the codes $P(\sigma)$ (for fixed m) have the same distance and weight enumerators (by Goethals and Snover [7, p. 85]). One of the many curious properties of the extended Preparata codes is that their weight enumerators are related to those of the Kerdock codes [11] in exactly the same manner as are the enumerators of a linear code and its dual [11], [7], [12, p. 468]. This naturally leads to speculations as

Manuscript received September 11, 1981; revised March 23, 1982.
The author is with the Department of Mathematics, University of Oregon, Eugene, OR 97403.

0018-9448/83/0500-0345$01.00