The $\mathbb{Z}_4$-Linearity of Kerdock, Preparata, Goethals and Related Codes*

*A. Roger Hammons, Jr.***
Hughes Aircraft Company
Network Systems Division, Germantown, MD 20876 U.S.A.

*P. Vijay Kumar***
Communication Science Institute, EE-Systems
University of Southern California, Los Angeles, CA 90089 U.S.A.

A. R. Calderbank and *N. J. A. Sloane*
Mathematical Sciences Research Center
AT&T Bell Laboratories, Murray Hill, NJ 07974 U.S.A.

Patrick Solé§
CNRS – I3S, 250 rue A. Einstein, bâtiment 4
Sophia – Antipolis, 06560 Valbonne, France

February 8, 2001

ABSTRACT

Certain notorious nonlinear binary codes contain more codewords than any known linear code. These include the codes constructed by Nordstrom-Robinson , Kerdock, Preparata, Goethals, and Delsarte-Goethals . It is shown here that all these codes can be very simply constructed as binary images under the Gray map of linear codes over $\mathbb{Z}_4$, the integers mod 4 (although this requires a slight modification of the Preparata and Goethals codes). The construction implies that all these binary codes are distance invariant. Duality in the $\mathbb{Z}_4$ domain implies that the binary images have dual weight distributions. The Kerdock and 'Preparata' codes are duals over $\mathbb{Z}_4$ — and the Nordstrom-Robinson code is self-dual — which explains why their weight distributions are dual to each other. The Kerdock and 'Preparata' codes are $\mathbb{Z}_4$-analogues of first-order Reed-Muller and extended Hamming codes, respectively. All these codes are extended cyclic codes over $\mathbb{Z}_4$, which greatly simplifies encoding and decoding. An algebraic hard-decision decoding algorithm is given for the 'Preparata' code and a Hadamard-transform soft-decision decoding algorithm for the Kerdock code. Binary first- and second-order Reed-Muller codes are also linear over $\mathbb{Z}_4$, but extended Hamming codes of length $n \geq 32$ and the Golay code are not. Using $\mathbb{Z}_4$-linearity, a new family of distance regular graphs are constructed on the cosets of the 'Preparata' code.

---

*A different version of this paper appeared in: *IEEE Trans. Inform. Theory*, **40** (1994), 301–319.

**The work of A. R. Hammons, Jr. and P. V. Kumar was supported in part by the National Science Foundation under Grant NCR-9016077 and by Hughes Aircraft Company under its Ph.D. fellowship program.

§P. Solé thanks the DIMACS Center and the IEEE for travel support.

The $\mathbb{Z}_4$-Linearity of Kerdock, Preparata, Goethals and Related Codes

A. Roger Hammons, Jr.[**]
Hughes Aircraft Company
Network Systems Division, Germantown, MD 20876 U.S.A.

P. Vijay Kumar[**]
Communication Science Institute, EE-Systems
University of Southern California, Los Angeles, CA 90089 U.S.A.

A. R. Calderbank and *N. J. A. Sloane*
Mathematical Sciences Research Center
AT&T Bell Laboratories, Murray Hill, NJ 07974 U.S.A.

Patrick Solé[§]
CNRS – I3S, 250 rue A. Einstein, bâtiment 4
Sophia – Antipolis, 06560 Valbonne, France

I. Introduction

Several notorious families of nonlinear codes have more codewords than any comparable linear code presently known. These are the Nordstrom-Robinson , Kerdock, Preparata , Goethals and Delsarte-Goethals codes [10], [28], [31], [32], [46], [56], [58], [61]. Besides their excellent error-correcting capabilities these codes are remarkable because the Kerdock and Preparata codes are 'formal duals', in the sense that although these codes are nonlinear, the weight distribution of one is the MacWilliams transform of the weight distribution of the other [56, Chap. 15]. The main unsolved question concerning these codes has always been whether they are duals in some more algebraic sense. Many authors have investigated these codes, and have found that (except for the Nordstrom-Robinson code) they are not unique, and indeed that large numbers of codes exist with the same weight distributions [2], [13], [43], [44], [45], [54]. Kantor [45] declares that the "apparent relationship between these [families of codes] is merely a coincidence."

Although this may be true for many versions of these codes, we will show that, when properly defined, Kerdock and Preparata codes are *linear* over $\mathbb{Z}_4$ (the integers mod 4), and that as $\mathbb{Z}_4$-codes they *are* duals. They are in fact just extended cyclic codes over $\mathbb{Z}_4$.

[**]The work of A. R. Hammons, Jr. and P. V. Kumar was supported in part by the National Science Foundation under Grant NCR-9016077 and by Hughes Aircraft Company under its Ph.D. fellowship program.

[§]P. Solé thanks the DIMACS Center and the IEEE for travel support.

The version of the Kerdock code that we use is the standard one, while our version of the Preparata code differs from the standard one in that it is not a subcode of the extended Hamming code but of a nonlinear code with the same weight distribution as the extended Hamming code. Our 'Preparata' code has the same weight distribution as Preparata's version, and has a similar construction in terms of finite field transforms. In our version, the Kerdock and 'Preparata' codes are $\mathbb{Z}_4$-analogues of first-order Reed-Muller and extended Hamming codes, respectively. Since the new construction is so simple, we propose that this is the 'correct' way to define these codes.

The situation may be compared with that for Hamming codes. It is known that there are many binary codes with the same weight distribution as the Hamming code — all are perfect single-error correcting codes, but one is distinguished by being linear (see [73], [59], [60] and also §5.4). Similarly, there are many binary codes with the same weight distributions as the Kerdock and Preparata codes; one pair is distinguished by being the images of a dual pair of linear extended-cyclic codes over $\mathbb{Z}_4$. It happens that Kerdock picked out the distinguished code, although Preparata did not.

Kerdock and Preparata codes exist for all lengths $n = 4^m \geq 16$. At length 16 they coincide, giving the Nordstrom-Robinson code [58], [66], [33]. The $\mathbb{Z}_4$ version of the Nordstrom-Robinson code turns out to be the 'octacode' [22], [23], a self-dual code of length 8 over $\mathbb{Z}_4$ that is used when the Leech lattice is constructed from eight copies of the face-centered cubic lattice.

The very good nonlinear binary codes of minimal distance 8 discovered by Goethals [31], [32], and the high minimal distance codes of Delsarte and Goethals [28], also have a simple description as extended cyclic codes over $\mathbb{Z}_4$, although our 'Goethals' code differs slightly from Goethals' original construction.

The decoding of all these codes is greatly simplified by working in the $\mathbb{Z}_4$-domain, where they are linear and it is meaningful to speak of syndromes. Decoding the Nordstrom-Robinson and 'Preparata' codes is especially simple.

These discoveries came about in the following way. Recently, a family of nearly optimal four-phase sequences of period $2^{2r+1} - 1$, with alphabet $\{1, i, -1, -i\}$, $i = \sqrt{-1}$, was discovered by Solé [67] and later independently by Boztaş, Hammons and Kumar [6], [7]. By replacing each element i^a by its exponent $a \in \{0, 1, 2, 3\}$, this family may be viewed as a linear code over $\mathbb{Z}_4$. Since the family has low correlation values, it also possesses a large minimal Euclidean distance and thus has the potential for excellent error-correcting capability.

When studying these four-phase sequences, Hammons and Kumar and later independently Calderbank, Sloane and Solé noticed the striking resemblance between the 2-adic (i.e. base 2) expansions of the quaternary codewords and the standard construction of the Kerdock codes. The reader can see this for himself by comparing the formulae on page 1107 of [7] (the common starting point for the two independent discoveries) and page 458 of [56].

Both teams then realized that the Kerdock code is simply the image of the quaternary code (when extended by an zero-sum check symbol) under the Gray map defined below (see (15)). This was a logical step to pursue since the Gray map translates a quaternary code with high minimal Lee or Euclidean distance into a binary code of twice the length with high minimal Hamming distance.

The discovery that the quaternary dual gives a code which is the 'correct' definition of the 'Preparata' code followed almost immediately.

The two teams worked independently until the middle of November 1992, when, discovering the considerable overlap between their work, they decided to join forces. The discoveries about the Kerdock and Preparata codes are in a paper [38] presented by Hammons and Kumar at the International Symposium on Information Theory (San Antonio, January 1993, but submitted in June 1992), in Hammons' dissertation [34], and in a manuscript [39] (now replaced by the present paper) submitted in early November 1992 to these *Transactions*. Hammons and Kumar realized in June 1992 that the $\mathbb{Z}_4$ Kerdock and 'Preparata' codes could be generalized to give the quaternary Reed-Muller codes $QRM(r, m)$ of Section 5.4.

In late October 1992, Calderbank, Sloane and Solé submitted a research announcement (now replaced by [11]) to the *Bulletin of the American Mathematical Society*, also containing the discoveries about the Kerdock and Preparata codes, as well as results (Sections 2.6 to 2.8) about the existence of quaternary versions of Reed-Muller, Golay and Hamming codes. They discovered the quaternary versions of the Goethals and Delsarte-Goethals codes in early November.

The present paper is a compositum of all our results.

The discovery that the Nordstrom-Robinson code is a quaternary version of the octacode was made by Forney, Sloane and Trott in early October 1992, and is described in [30]. (It was already known to Hammons and Kumar in June 1992 that the Nordstrom-Robinson code was linear over $\mathbb{Z}_4$, but they had not made the identification with the octacode.)

It can be shown that the binary nonlinear single-error-correcting codes found by Best [4], Julin [42], Sloane and Whitehead [65] and others can also be more simply described as codes over $\mathbb{Z}_4$ (although here the corresponding $\mathbb{Z}_4$-codes are nonlinear). This will be described elsewhere [24]. Large sequence families for code-division multiple-access (CDMA) that are supersets of the near optimum four-phase sequence families described above and which are related to the Delsarte-Goethals codes are investigated in [49].

The paper is arranged as follows. Section II discusses linear codes over $\mathbb{Z}_4$, their duals, and their images as binary codes under the Gray map. Necessary and sufficient conditions are given for a binary code to be the image of a linear code over $\mathbb{Z}_4$. Reed-Muller codes of length 2^m and orders $0, 1, 2, m-1, m$ satisfy these conditions, but extended Hamming codes and the Golay code do not. Cyclic codes over $\mathbb{Z}_4$ are studied by means of Galois rings $GR(4^m)$ rather than the Galois fields $GF(2^m)$ used to analyze binary cyclic codes, and Section III is devoted to these rings.

In Section IV we show that Kerdock codes are extended cyclic codes over $\mathbb{Z}_4$, and in fact are simply $\mathbb{Z}_4$-analogues of first-order Reed-Muller codes (see the generator matrix (49) and also §5.4). The Nordstrom-Robinson code is discussed in §4.5. Subsequent subsections give the weight distribution of the Kerdock codes and a soft-decision decoding algorithm for them.

In Section V we show that the binary images of the quaternary duals of the Kerdock codes are Preparata-like codes, having essentially the same properties as Preparata's original codes. Theorem 15, however, shows that the 'Preparata' codes are strictly different from the original construction. §5.2 provides a finite field transform characterization of the 'Preparata' codes and compares them with the original codes. The 'Preparata' codes have a very simple decoding algorithm (§5.3). (This is considerably simpler than any previous decoding algorithm — compare [5].) Section 5.4 defines a family of quaternary Reed-Muller codes $QRM(m, r)$ which generalizes the quaternary Kerdock and 'Preparata' codes. The final subsections are concerned with the automorphism groups of these codes (§5.5), and a new family of distance regular graphs defined on the cosets of the 'Preparata' code (§5.6).

In Section VI we show that the binary nonlinear Delsarte-Goethals codes [28] are also extended cyclic codes over $\mathbb{Z}_4$, and that their $\mathbb{Z}_4$-duals have essentially the same properties as the Goethals codes [31], [32] and the 'Goethals-Delsarte' codes of Hergert [40].

Postscript. After this paper was completed, V. I. Levenshtein drew our attention to an article by Nechaev [57]. In this article Nechaev considers the quaternary sequences $\{c_t\}$ given (in the notation of the present paper) by

$$c_t = (-1)^t \{T(\lambda \xi^t) + \delta\} ,$$

$0 \le t \le 2^{m+1} - 3$, $\lambda \in R$, $\delta \in \mathbb{Z}_4$, and their 2-adic expansions $c_t = a_t + 2b_t$, where a_t, $b_t \in \{0, 1\}$. The principal result of [57] shows that the set of $\{b_t\}$ is a nonlinear binary cyclic code which is equivalent to the binary Kerdock code punctured in two coordinates. However, [57] makes no mention of the fundamental isometry of Eq. (15), nor of Preparata codes and the sense in which they are duals of Kerdock codes.

II. Quaternary and related binary codes

2.1. Quaternary codes.

By a *quaternary code* $\mathcal{C}$ of length n we shall mean a linear block code over $\mathbb{Z}_4$, i.e. an additive subgroup of $\mathbb{Z}_4^n$. Such codes have been studied recently both in connection with the construction of sequences with low correlation ([6], [7], [67], [72]) and in a variety of other contexts (see [23] and the references contained therein).

We define an inner product on $\mathbb{Z}_4^n$ by $a \cdot b = a_1 b_1 + \cdots + a_n b_n \pmod 4$, and then the notions of *dual* code ($\mathcal{C}^\perp$), *self-orthogonal* code ($\mathcal{C} \subseteq \mathcal{C}^\perp$) and *self-dual* code ($\mathcal{C} = \mathcal{C}^\perp$) are defined in the standard way (cf. [47], [56]). For many applications there is no need to distinguish between $+1$ components of codewords and -1 components, and so we say that two codes are *equivalent* if one can be obtained from the other by permuting the coordinates and (if necessary) changing the signs of certain coordinates. Codes differing by only a permutation of coordinates are called *permutation-equivalent*. The automorphism group $\mathrm{Aut}(\mathcal{C})$ of $\mathcal{C}$ consists of all permutations and sign-changes of the coordinates that preserve the set of codewords.

Any code is permutation-equivalent to a code $\mathcal{C}$ with generator matrix of the form

$$G = \begin{bmatrix} I_{k_1} & A & B \\ 0 & 2I_{k_2} & 2C \end{bmatrix} , \tag{1}$$

where A and C are $\mathbb{Z}_2$-matrices and B is a $\mathbb{Z}_4$-matrix. The code is then an elementary abelian group of type $4^{k_1} 2^{k_2}$, containing $2^{2k_1 + k_2}$ codewords. We shall indicate this by saying that $\mathcal{C}$ has type $4^{k_1} 2^{k_2}$, or simply that $|\mathcal{C}| = 4^{k_1} 2^{k_2}$.

Eq. (1) illustrates a difference in point of view between ring theory and coding theory. Quaternary codes are $\mathbb{Z}_4$-modules. A ring theorist would point out, correctly, that a quaternary code is not in general a free module [41], and so need not have a basis. Although this is true,

(1) is a perfectly good generator matrix. Encoding is carried out by writing the information symbols in the form $u = u_1 \cdots u_{k_1} u_{k_1+1} \cdots u_{k_1+k_2}$, where $u_i \in \mathbb{Z}_4$ if $1 \le i \le k_1$, $u_i \in \mathbb{Z}_2$ if $k_1+1 \le i \le k_1+k_2$, and mapping u to the codeword uG. The code $\mathcal{C}$ is a free $\mathbb{Z}_4$-module if and only if $k_2 = 0$.

If $\mathcal{C}$ has generator matrix (1), the dual code $\mathcal{C}^\perp$ has generator matrix

$$\left[\begin{array}{ccc} -B^{tr} - C^{tr}A^{tr} & C^{tr} & I_{n-k_1-k_2} \\ 2A^{tr} & 2I_{k_2} & 0 \end{array} \right] \tag{2}$$

and type $4^{n-k_1-k_2}2^{k_2}$.

2.2. Weight enumerators.

Several weight enumerators are associated with a quaternary code $\mathcal{C}$. The *complete weight enumerator* (or c.w.e.) of $\mathcal{C}$ is

$$\mathrm{cwe}_{\mathcal{C}}(W,X,Y,Z) = \sum_{a\in\mathcal{C}} W^{n_0(a)} X^{n_1(a)} Y^{n_2(a)} Z^{n_3(a)} , \tag{3}$$

where $n_j(a)$ is the number of components of a that are congruent to $j \pmod 4$ (cf. [47], [56, p. 141]). Permutation-equivalent codes have the same c.w.e., but equivalent codes may have distinct c.w.e.'s. The appropriate weight enumerator for an equivalence class of codes is the *symmetrized weight enumerator* (or s.w.e.), obtained by identifying X and Z in (3):

$$\mathrm{swe}_{\mathcal{C}}(W,X,Y) = \mathrm{cwe}_{\mathcal{C}}(W,X,Y,X) . \tag{4}$$

The Lee weights of $0,1,2,3 \in \mathbb{Z}_4$ are $0,1,2,1$ respectively, and the Lee weight $wt_L(a)$ of $a \in \mathbb{Z}_4^n$ is the rational sum of the Lee weights of its components. This weight function defines a distance $d_L(\ ,\ )$ on $\mathbb{Z}_4^N$ called the *Lee metric*. The *Lee weight enumerator* of $\mathcal{C}$ is

$$\begin{aligned} \mathrm{Lee}_{\mathcal{C}}(W,X) &= \sum_{a\in\mathcal{C}} W^{2n-wt_L(a)} X^{wt_L(a)} \\ &= \mathrm{swe}_{\mathcal{C}}(W^2, WX, X^2) , \end{aligned} \tag{5}$$

a homogeneous polynomial of degree $2n$. Finally, the *Hamming weight enumerator* of $\mathcal{C}$, less useful than the others, is

$$\mathrm{Ham}_{\mathcal{C}}(W,X) = \mathrm{swe}_{\mathcal{C}}(W,X,X) . \tag{6}$$

We then have the following analogues of the MacWilliams identity, giving the weight enumerators for the dual code $\mathcal{C}^\perp$ ([47], [48], [23]):

$$\mathrm{cwe}_{\mathcal{C}^\perp}(W,X,Y,Z) = \frac{1}{|\mathcal{C}|}\mathrm{cwe}_{\mathcal{C}}(W+X+Y+Z, W+iX-Y-iZ, W-X+Y-Z, W-iX-Y+iZ) , \tag{7}$$

$$\mathrm{swe}_{\mathcal{C}^\perp}(W,X,Y) = \frac{1}{|\mathcal{C}|}\mathrm{swe}_{\mathcal{C}}(W+2X+Y, W-Y, W-2X+Y)\ , \tag{8}$$

$$\mathrm{Lee}_{\mathcal{C}^\perp}(W,X) = \frac{1}{|\mathcal{C}|}\mathrm{Lee}_{\mathcal{C}}(W+X, W-X)\ , \tag{9}$$

$$\mathrm{Ham}_{\mathcal{C}^\perp}(W,X) = \frac{1}{|\mathcal{C}|}\mathrm{Ham}_{\mathcal{C}}(W+3X, W-X)\ . \tag{10}$$

There are also several analogues of Gleason's theorem, giving bases for the weight enumerators of self-dual codes — see [47], [23].

2.3. Associated complex-valued sequences.

We may associate to every $\mathbb{Z}_4$-valued vector $a = (a_1, \ldots, a_n)$ an equivalent complex roots-of-unity sequence $s = i^a = (i^{a_1}, \ldots, i^{a_n})$, where $i = \sqrt{-1}$. Then, given a set $\mathcal{C}$ of quaternary vectors, we let

$$\Omega(\mathcal{C}) = \{i^a :\ a \in \mathcal{C}\}$$

denote the corresponding set of complex sequences. When $\mathcal{C}$ is regarded as a set of CDMA signature sequences, its effectiveness depends on the complex correlations (or Hermitian inner products) of the sequences in $\Omega(\mathcal{C})$. When $\mathcal{C}$ is regarded as a code, its error-correcting capability depends on the Euclidean distance properties of $\Omega(\mathcal{C})$. If a, b are quaternary vectors with associated vectors $s = i^a$, $t = i^b$, then

$$\begin{aligned} \|s-t\|^2 &= \|s\|^2 + \|t\|^2 - 2\ \mathrm{Re}\ \{s^H t\} \\ &= 2n - 2\ \mathrm{Re}\ \{\zeta(a-b)\}\ , \end{aligned} \tag{11}$$

where ^{H} denotes the Hermitian inner product, and

$$\zeta(a-b) = \sum_{r=1}^{n} i^{a_r - b_r} \tag{12}$$

is the *complex correlation* of a and b. Note that ζ depends only on the difference $a-b$. By (11), if the nontrivial correlations of $\Omega(\mathcal{C})$ are low in magnitude, then the set possesses large minimal Euclidean distance. We also see that

$$\|s-t\|^2 = 2d_L(a,b)\ . \tag{13}$$

2.4. Binary codes associated with quaternary codes; the Gray map.

In communication systems employing quadrature phase-shift keying (QPSK), the preferred assignment of two information bits to the four possible phases is the one shown in Fig. 1, in which adjacent phases differ by only one binary digit. This mapping is called *Gray encoding* and has the advantage that, when a quaternary codeword is transmitted across an additive white Gaussian noise channel, the errors most likely to occur are those causing a single erroneously decoded information bit.

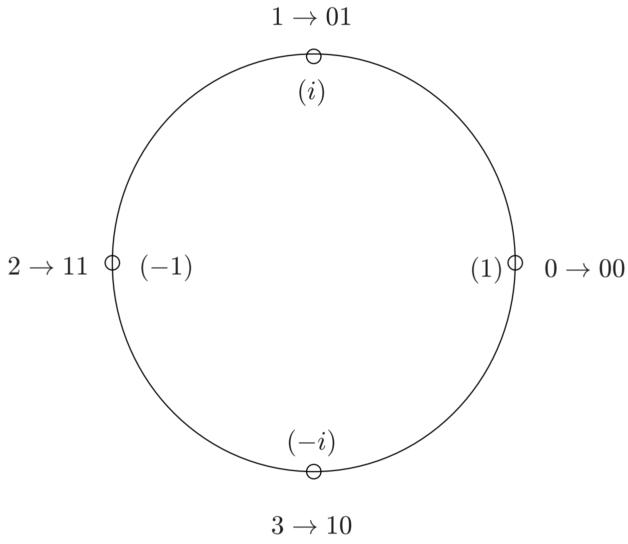


Figure 1: Gray encoding of quaternary symbols and QPSK phases.

Formally, we define three maps from $\mathbb{Z}_4$ to $\mathbb{Z}_2$ by

c	$\alpha(c)$	$\beta(c)$	$\gamma(c)$
0	0	0	0
1	1	0	1
2	0	1	1
3	1	1	0

and extend them in the obvious way to maps from $\mathbb{Z}_4^n$ to $\mathbb{Z}_2^n$. The 2-adic expansion of $c \in \mathbb{Z}_4$ is

$$c = \alpha(c) + 2\beta(c) \ . \tag{14}$$

Note that $\alpha(c) + \beta(c) + \gamma(c) = 0$ for all $c \in \mathbb{Z}_4$. We construct binary codes from quaternary codes using the *Gray map* $\phi : \mathbb{Z}_4^n \to \mathbb{Z}_2^{2n}$ given by

$$\phi(c) = (\beta(c), \gamma(c)), \quad c \in \mathbb{Z}_4^n \ . \tag{15}$$

When we speak of the binary image of a quaternary code $\mathcal{C}$, we will always mean its image $C = \phi(\mathcal{C})$ under the Gray map. We use script letters for quaternary codes, with the corresponding Latin letters for their binary images.

C is in general a nonlinear binary code of length $2n$. If $\mathcal{C}$ *is* linear, and $\mathcal{C}$ is defined by (1), then C has generator matrix

$$\begin{bmatrix} I_{k_1} & A & \alpha(B) & I_{k_1} & A & \alpha(B) \\ 0 & I_{k_2} & C & 0 & I_{k_2} & C \\ 0 & 0 & \beta(B) & I_{k_1} & A & \gamma(B) \end{bmatrix} . \tag{16}$$

We say that a binary code C is $\mathbb{Z}_4$-*linear* if its coordinates can be arranged so that it is the image under the Gray map ϕ of a quaternary code $\mathcal{C}$.

The crucial property of the Gray map is that it preserves distances.

Theorem 1. *ϕ is a distance-preserving map or* isometry *from*

$$(\mathbb{Z}_4^n, \textit{ Lee distance}) \quad \textit{to} \quad (\mathbb{Z}_2^{2n}, \textit{ Hamming distance}) .$$

Proof. It is easy to see from the definitions (and Fig. 1) that

$$wt(\phi(a)) = wt_L(a), \quad a \in \mathbb{Z}_4^n , \tag{17}$$

$$d(\phi(a), \phi(b)) = d_L(a, b), \quad a, b \in \mathbb{Z}_4^n , \tag{18}$$

where $wt(\ )$ and $d(\ ,\ )$ are the usual Hamming weight and distance functions for binary vectors. ■

From (13), (18), the Hamming distance between the binary images $\phi(a)$ and $\phi(b)$ is proportional to the squared Euclidean distance between the complex sequences i^a and i^b.

Two other binary codes $C^{(1)}$, $C^{(2)}$ are canonically associated with a quaternary code $\mathcal{C}$. These are the linear codes defined by

$$C^{(1)} = \{\alpha(c) : \ c \in \mathcal{C}\} , \tag{19}$$

$$C^{(2)} = \{\beta(c) : \ c \in \mathcal{C}, \ \alpha(c) = 0\} . \tag{20}$$

If $\mathcal{C}$ has generator matrix (1), then $C^{(1)}$ is an $[n, k_1]$ code with generator matrix

$$[I_{k_1} \ \ A \ \ \alpha(B)] , \tag{21}$$

while $C^{(2)} \supseteq C^{(1)}$ is an $[n, k_1 + k_2]$ code with generator matrix

$$\begin{bmatrix} I_{k_1} & A & \alpha(B) \\ 0 & I_{k_2} & C \end{bmatrix} \tag{22}$$

— compare (16). It is shown in [23] that given any binary codes C', C'' of length n with $C'' \supseteq C'$, there is a quaternary code $\mathcal{C}$ with $C^{(1)} = C'$, $C^{(2)} = C''$.

2.5. Weight and distance properties.

Since in general $C = \phi(\mathcal{C})$ is not linear, it need not have a dual. We define its $\mathbb{Z}_4$-*dual* to be $C_\perp = \phi(\mathcal{C}^\perp)$, as in the diagram

$$\begin{array}{ccl} \mathcal{C} & \xrightarrow{\phi} & C \;=\; \phi(\mathcal{C}) \\ {\scriptstyle \text{dual}}\Big\downarrow & & \\ \mathcal{C}^\perp & \xrightarrow{\phi} & C_\perp = \phi(\mathcal{C}^\perp)\ . \end{array} \tag{23}$$

Note that one cannot add an arrow marked 'dual' on the right side to produce a commuting diagram.

In this section we discuss the weight and distance properties of C and $C_\perp$. The principal results to be derived here are the following:

(1) C and $C_\perp$ are distance invariant.

(2) The weight distributions of C and $C_\perp$ are MacWilliams transforms of one another.

A binary code C is said to be *distance invariant* [56, p. 40] if the Hamming weight distributions of its translates $u + C$ are the same for all $u \in C$.

Theorem 2. *If $\mathcal{C}$ is a (linear) quaternary code, then its binary Gray representation $C = \phi(\mathcal{C})$ is distance invariant.*

Proof. $\mathcal{C}$ is distance invariant (with respect to Lee distance) because it is linear, and the result then follows from Theorem 1. ■

For a distance invariant code C of length n, the (Hamming) weight enumerator

$$\mathrm{Ham}_C(W, Z) = \sum_{c' \in C} W^{n-d(c',c)} X^{d(c',c)}$$

is independent of $c \in C$. If $C = \phi(\mathcal{C})$, it follows from Theorem 1 and (5) that

$$\mathrm{Ham}_C(W, X) = \mathrm{Lee}_{\mathcal{C}}(W, X) = \mathrm{swe}_{\mathcal{C}}(W^2, WX, X^2)\ . \tag{24}$$

Theorem 3. *If $\mathcal{C}$ and $\mathcal{C}^\perp$ are dual quaternary codes, then the weight distributions of the binary codes $C = \phi(\mathcal{C})$ and $C_\perp = \phi(\mathcal{C}^\perp)$ are related by the binary MacWilliams transform.*

Proof. From (24), (9) we have

$$\begin{aligned}\mathrm{Ham}_{C_\perp}(W,X) &= \mathrm{Lee}_{\mathcal{C}^\perp}(W,X) \\ &= \frac{1}{|\mathcal{C}|}\mathrm{Lee}_{\mathcal{C}}(W+X,\ W-X) \\ &= \frac{1}{|\mathcal{C}|}\mathrm{Ham}_{C}(W+X,\ W-X)\ .\end{aligned}$$

as required. ■

2.6. Existence and linearity conditions.

We now give necessary and sufficient conditions for a binary code to be $\mathbb{Z}_4$-linear, and for the binary image of a quaternary code to be a linear code. The reader who is primarily interested in Kerdock and Preparata codes should skip to Section III.

Since $\phi(-c) = (\gamma(c), \beta(c))$, it follows that if C is $\mathbb{Z}_4$-linear then C is fixed under the 'swap' map σ that interchanges the left and right halves of each codeword:

$$\sigma : (u_1\ u_2 \cdots u_n\ u_{n+1} \cdots u_{2n}) \mapsto (u_{n+1} \cdots u_{2n}\ u_1\ u_2 \cdots u_n)\ . \tag{25}$$

In other words σ applies the permutation

$$(1, n+1)(2, n+2)\cdots(n, 2n) \tag{26}$$

to the coordinates. This is a fixed-point-free involution in the automorphism group of C.

Theorem 4. *A binary, not necessarily linear, code C of even length is $\mathbb{Z}_4$-linear if and only if its coordinates can be arranged so that*

$$u, v \in C \Rightarrow u + v + (u + \sigma(u)) * (v + \sigma(v)) \in C\ , \tag{27}$$

where σ is the swap map that interchanges the left and right halves of a vector, and $*$ denotes the componentwise product of two vectors.

Proof. This is an immediate consequence of the easily-verified identity

$$\phi(a+b) = \phi(a) + \phi(b) + (\phi(a) + \sigma(\phi(a))) * (\phi(b) + \sigma(\phi(b)))\ , \tag{28}$$

for all $a, b \in \mathbb{Z}_4^n$. ■

Theorem 5. *The binary image $\phi(\mathcal{C})$ of a quaternary linear code $\mathcal{C}$ is linear if and only if*

$$a, b \in \mathcal{C} \Rightarrow 2\alpha(a) * \alpha(b) \in \mathcal{C} \tag{29}$$

Proof. This is an immediate consequence of the identity

$$\phi(a) + \phi(b) + \phi(a + b) = \phi(2\alpha(a) * \alpha(b)) \tag{30}$$

for all $a, b \in \mathbb{Z}_4^n$. ■

Theorem 6. *A binary linear code C of even length is $\mathbb{Z}_4$-linear if and only if its coordinates can be permuted so that*

$$u, v \in C \Rightarrow (u + \sigma(u)) * (v + \sigma(v)) \in C \ , \tag{31}$$

where σ is as in Theorem 4.

Proof. This is also a consequence of (28). ■

Conditions (29), (27) and (31) are very restrictive, and (we are now speaking informally) imply that most binary codes are not $\mathbb{Z}_4$-linear.

2.7. Reed-Muller and Hamming codes

Theorem 7. *The rth order binary Reed-Muller code $RM(r, m)$ of length $n = 2^m$, $m \geq 1$, is $\mathbb{Z}_4$-linear for $r = 0, 1, 2, m - 1$ and m.*

Proof. We leave to the reader the straightforward verification that $RM(r, m)$ is the image under ϕ of the quaternary code $ZRM(r, m - 1)$ (say) of length 2^{m-1} generated by $RM(r - 1, m - 1)$ and $2\ RM(r, m - 1)$, for $r = 0, 1, 2, m - 1, m$ (with the convention that $RM(-1, m - 1) = RM(m, m - 1) = 0$). ■

Let $(v_1, \ldots, v_{m-1})$ range over $\mathbb{Z}_2^{m-1}$, so that $RM(r, m-1)$ is generated (in the usual way, as a binary code) by the vectors corresponding to monomials in the Boolean functions v_i of degree $\leq r$ [56, Chap. 13]. Then $RM(1, m)$ is the binary image of the quaternary code $ZRM(1, m-1)$ generated by the vectors corresponding to $1, 2v_1, \ldots, 2v_{m-1}$, and $RM(2, m)$ is the image of the quaternary code $ZRM(2, m - 1)$ generated by $1, v_1, \ldots, v_{m-1}$, $2v_1v_2$, $2v_1v_3, \ldots, 2v_{m-2}v_{m-1}$.

For example, the $[16, 5, 8]$ code $RM(1, 4)$ and the $[16, 11, 4]$ code $RM(2, 4)$ are the binary images of the quaternary codes with generator matrices

$$\begin{array}{cc} ZRM(1,3) & \\ \begin{bmatrix} 11111111 \\ 00002222 \\ 00220022 \\ 02020202 \end{bmatrix} & \begin{array}{l} 1 \\ 2v_1 \\ 2v_2 \\ 2v_3 \end{array} \end{array} \ , \quad \begin{array}{cc} ZRM(2,3) & \\ \begin{bmatrix} 11111111 \\ 00001111 \\ 00110011 \\ 01010101 \\ 00000022 \\ 00000202 \\ 00020002 \end{bmatrix} & \begin{array}{l} 1 \\ v_1 \\ v_2 \\ v_3 \\ 2v_1v_2 \\ 2v_1v_3 \\ 2v_2v_3 \end{array} \end{array} \ . \tag{32}$$

In Eq. (31), if u, v are represented by Boolean functions of degree r, and $(u+\sigma(u))*(v+\sigma(v)) \neq 0$, then $(u+\sigma(u))*(v+\sigma(v))$ is a Boolean function of degree $2r-2$. So an rth order Reed-Muller code with $r \leq m/2$ satisfies (31) provided $r \leq 2$ (which gives an alternative proof of part of Theorem 7), but we conjecture that it does not satisfy (31) if $3 \leq r \leq m-2$. In other words we conjecture that if C is a binary Reed-Muller code $RM(r,m)$ with $3 \leq r \leq m-2$, then there is no permutation of the coordinates of C such that the permuted code is equal to $\phi(\mathcal{C})$ for some quaternary code $\mathcal{C}$. However, we have found a proof of this only for $(m-2)$nd order RM codes.

Theorem 8. *The binary code $RM(m-2,m)$, i.e. the extended Hamming code of length $n = 2^m$, is not $\mathbb{Z}_4$-linear for $m \geq 5$.*

Proof. Suppose H is a $[2^m, 2^m - m - 1, 4]$ extended Hamming code with its coordinates arranged so that $H = \phi(\mathcal{H})$ for some quaternary code $\mathcal{H}$. We will obtain a contradiction for $m \geq 5$. The codewords of weight 4 in H form a Steiner system $S(3, 4, 2^m)$ [56, p. 63]. From this it follows without difficulty that

$$\begin{array}{l}\text{for } m \geq 4,\ H \text{ contains codewords of} \\ \text{weight 4 that meet in just one coordinate.}\end{array} \tag{33}$$

Let F be the subcode of H fixed under the swap map σ of (25), and let ψ be the homomorphism $H \to F$ given by $\psi(x) = x + \sigma(x)$. Then $\operatorname{im}\psi \subseteq \ker\psi = F$. Since $\dim\ker\psi \leq 2^{m-1} - 1$, $\dim \operatorname{im} \psi \geq 2^{m-1} - m$. Let E consist of the right-hand halves of the codewords in $\operatorname{im} \psi$. Then E is a $[2^{m-1}, \geq 2^{m-1} - m, 2]$ code, containing say A_i words of weight i. We know from Theorem 6 that E is closed under componentwise multiplication.

Therefore the $A_2 + A_3$ words of weights 2 and 3 in E must be disjoint, or else E would contain a word of weight 1. Omitting these words from E, we are left with a code of length $2^{m-1} - 2A_2 - 3A_3$, dimension $\geq 2^{m-1} - m - A_2 - A_3$, and minimal distance 4. This violates the optimality of shortened Hamming codes unless $A_2 = A_3 = 0$ and E is itself an extended Hamming code of length 2^{m-1}. For $m \geq 5$ we now use (33) to deduce that E contains a word of weight 1, a contradiction. ∎

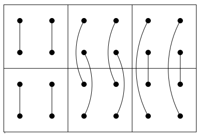
Theorem 8 demonstrates that a binary code can be $\mathbb{Z}_4$-linear, even though its dual is not. For $RM(1,m)$ is $\mathbb{Z}_4$-linear, while in general its dual, $RM(m-2,m)$, is not.

2.8. The Golay code.

Since the Nordstrom-Robinson code is $\mathbb{Z}_4$-linear (as we shall see in Theorem 12) and is closely connected with the Golay code ([56, p. 73], it is natural to ask if the Golay code itself is $\mathbb{Z}_4$-linear.

Theorem 9. *The* $[24, 12, 8]$ *Golay code* G *is not* $\mathbb{Z}_4$*-linear.*

Proof. Suppose on the contrary that G is the binary image of a quaternary linear code $\mathcal{G}$. The swap map σ (see (26)) is a fixed-point-free involution in $\mathrm{Aut}(G)$, the Mathieu group M_{24}. It is known ([19], [22]) that M_{24} contains a single conjugacy class of such involutions. Therefore, without loss of generality, we may suppose that this involution is the map defined by addition of the hexacodeword $11\,\omega\,\omega\,\overline{\omega}\,\overline{\omega}$ in the MOG description of G (see [22], Chap. 11, §9). In the MOG diagram this is the permutation



The diagram specifies the division of the 24 coordinates into twelve pairs, although we do not yet know which coordinate of each pair is on the left (in (25)) and which is on the right. Consider the Golay codewords

$$u = \begin{array}{|c|c|c|}\hline 11 & 11 & 00 \\ 11 & 11 & 00 \\ \hline 00 & 00 & 00 \\ 00 & 00 & 00 \\ \hline \end{array}\,, \qquad v = \begin{array}{|c|c|c|}\hline 01 & 11 & 11 \\ 10 & 00 & 00 \\ \hline 10 & 00 & 00 \\ 10 & 00 & 00 \\ \hline \end{array}\,.$$

Then

$$u + \sigma(u) = \begin{array}{|c|c|c|}\hline 00 & 11 & 00 \\ 00 & 11 & 00 \\ \hline 00 & 11 & 00 \\ 00 & 11 & 00 \\ \hline \end{array}\,, \qquad v + \sigma(v) = \begin{array}{|c|c|c|}\hline 11 & 11 & 11 \\ 11 & 00 & 00 \\ \hline 00 & 11 & 00 \\ 00 & 00 & 11 \\ \hline \end{array}$$

and $(u+\sigma(u))*(v+\sigma(v))$ (which by Theorem 6 must be in G) has weight 4, a contradiction. ■

III. Cyclic codes over $\mathbb{Z}_4$ and Galois rings

3.1. Galois rings.

To study BCH and other cyclic codes of length n over an alphabet of size q, it is customary to work in a Galois field $GF(q^m)$, an extension of degree m of a ground field $GF(q)$ [56]. The ground field $GF(q)$ is identified with the alphabet, and the extension field is chosen so that it contains an nth root of unity.

A similar approach is used for cyclic codes of length n over $\mathbb{Z}_4$, only now one constructs a Galois ring $GR(4^m)$ (not a field), that is an extension of $\mathbb{Z}_4$ of degree m containing an nth root of unity.

Galois rings have been studied by MacDonald [55], Liebler and Mena [52], Shankar [64], Solé [67], Yamada [71], Boztaş, Hammons and Kumar [7], among others, and of course the general machinery of commutative algebra, as described for example in Zariski and Samuel [76], is applicable to these rings. We list here some of the basic facts we shall need; proofs may be found in the above references.

Let $h_2(X) \in \mathbb{Z}_2[X]$ be a primitive irreducible polynomial of degree m. There is a unique monic polynomial $h(X) \in \mathbb{Z}_4[X]$ of degree m such that $h(X) \equiv h_2(X) \pmod 2$ and $h(X)$ divides $X^n - 1 \pmod 4$, where $n = 2^m - 1$ (see for example Yamada [71]). The polynomial $h(X)$ is a *primitive basic irreducible* polynomial, and may be found as follows.

Let $h_2(X) = e(X) - d(X)$, where $e(X)$ contains only even powers and $d(X)$ only odd powers. Then $h(X)$ is given by $h(X^2) = \pm(e^2(X) - d^2(X))$. This is Graeffe's method [70], [67] for finding a polynomial whose roots are the squares of the roots of $h_2(X)$. For example, when $m = 3$, $n = 7$ we may take $h_2 = X^3 + X + 1$. Then $e = 1$, $d = -X^3 - X$, $e^2 - d^2 = -X^6 - 2X^4 - X^2 + 1$, so

$$h(X) = X^3 + 2X^2 + X - 1 \ . \tag{34}$$

Table I in [7] gives all primitive basic irreducible polynomials of degree $m \leq 10$.

Let ξ be a root of $h(X)$, so that $\xi^n = 1$. Then the *Galois ring* $GR(4^m)$ is defined to be $R = \mathbb{Z}_4[\xi]$. There are two canonical ways to represent the 4^m elements of R (just as there are two canonical ways, multiplicative and additive, to represent elements of $GF(q^m)$).

In the first representation, every element $c \in R$ has a unique 'multiplicative' or 2-adic representation

$$c = a + 2b \ , \tag{35}$$

where a and b belong to the set

$$\mathcal{T} = \{0, 1, \xi, \xi^2, \ldots, \xi^{n-1}\} \ . \tag{36}$$

The map $\tau : c \mapsto a$ is given by

$$\tau(c) = c^{2^m}, \quad c \in R \ , \tag{37}$$

and satisfies

$$\tau(cd) = \tau(c)\tau(d) , \tag{38}$$
$$\tau(c+d) = \tau(c)+\tau(d)+2(cd)^{2^{m-1}} \tag{39}$$

(see [71]). Given c, one determines a from (37) and then b from (35).

In the second representation, each element $c \in R$ has a unique 'additive' representation

$$c = \sum_{r=0}^{m-1} b_r \xi^r, \quad b_r \in \mathbb{Z}_4 . \tag{40}$$

For example, if $m = 3$ and h is given by (34), the additive representations for the elements of $\mathcal{T}$ and $2\mathcal{T}$ are

element	b_0	b_1	b_2	$2b_0$	$2b_1$	$2b_2$
0	0	0	0	0	0	0
1	1	0	0	2	0	0
ξ	0	1	0	0	2	0
ξ^2	0	0	1	0	0	2
ξ^3	1	3	2	2	2	0
ξ^4	2	3	3	0	2	2
ξ^5	3	3	1	2	2	2
ξ^6	1	2	1	2	0	2

(41)

This table may be produced (just as for Galois fields) by a (modulo 4) shift register whose feedback polynomial is $h(X)$. By using (35), the table gives the additive representation of every element of R.

One essential difference between $R = GR(4^m)$ and a Galois field is that R contains zero divisors: these are the elements of the radical $2R$, the unique maximal ideal in R (R is a local ring). Let μ denote the map $R \to R/2R$. Then $\theta = \mu(\xi)$ is a root of $h_2(X)$, and we can identify $R/2R$ with $GF(2^m)$, taking the elements of $GF(2^m)$ to be

$$\mu(\mathcal{T}) = \{0, 1, \theta, \theta^2, \ldots, \theta^{n-1}\} . \tag{42}$$

We denote the set of *regular* or *invertible* elements of R by $R^* = R \setminus 2R$. Every element of R^* has a unique representation in the form $\xi^r(1+2t)$, $0 \le r \le n-1$, $t \in \mathcal{T}$. R^* is a multiplicative group of order $(2^m-1)2^m$ which is a direct product $H \times \mathcal{E}$, where H is the cyclic group of order $2^m - 1$ generated by ξ, and $\mathcal{E}$ is the group of *principal units* of R, that is, elements of the form $1+2t$, $t \in \mathcal{T}$. $\mathcal{E}$ has the structure of an elementary abelian group of order 2^m and is isomorphic to the additive group of $GF(2^m)$.

3.2. Frobenius and trace maps.

The *Frobenius map* f from R to R is the ring automorphism that takes any element $c = a+2b \in R$ to

$$c^f = a^2 + 2b^2 \ . \tag{43}$$

f generates the Galois group of R over $\mathbb{Z}_4$, and f^m is the identity map. The *relative trace* from R to $\mathbb{Z}_4$ is defined by

$$T(c) = c + c^f + c^{f^2} + \cdots + c^{f^{m-1}}, \quad c \in R \ . \tag{44}$$

For comparison, the usual trace from $GF(2^m)$ to $\mathbb{Z}_2$ is given by

$$tr(c) = c + c^2 + c^{2^2} + \cdots + c^{2^{m-1}}, \quad c \in GF(2^m) \ , \tag{45}$$

and the Frobenius map is simply the squaring map

$$f_2(c) = c^2, \quad c \in GF(2^m) \ . \tag{46}$$

The following commutativity relationships between these maps are easily verified:

$$\mu \circ f = f_2 \circ \mu \ , \tag{47}$$

$$\mu \circ T = tr \circ \mu \ . \tag{48}$$

In particular, since tr is not identically zero, it follows that the Galois ring trace is nontrivial. In fact, T is an onto mapping from R to $\mathbb{Z}_4$. The set of elements of R invariant under f is identical with $\mathbb{Z}_4$.

3.3. Dependencies among ξ^j.

For later use we record some results about dependencies among the powers ξ^j.

(P1) $\pm\xi^j \pm \xi^k$ is invertible for $0 \le j < k < 2^m - 1$, for $m \ge 2$. Proof. If on the contrary we had $\pm\xi^j \pm \xi^k = 2\lambda$, $\lambda \in R$, then applying μ we obtain $\theta^j + \theta^k = 0$, which contradicts the fact that θ is primitive in $GF(2^m)$. ■

(P2) $\xi^j - \xi^k \ne \pm\xi^l$ for distinct j, k, l in the range $[0, 2^m - 2]$, for $m \ge 2$. Proof. Otherwise, after rearranging, we have $1 + \xi^a = \xi^b$ for $a \ne b$. Squaring gives $1 + 2\xi^a + \xi^{2a} = \xi^{2b}$, but applying the Frobenius map gives $1 + \xi^{2a} = \xi^{2b}$, so $2\xi^a = 0$, a contradiction. ■

(P3) Suppose i, j, k, l are in the range $[0, 2^m - 2]$ and $i \neq j$, $k \neq l$, $m \geq 3$. Then

$$\xi^i - \xi^j = \xi^k - \xi^l \Leftrightarrow i = k \quad \text{and} \quad j = l \ ,$$

Proof. Suppose $1 + \xi^a = \xi^b + \xi^c$. Squaring and subtracting the result of applying the Frobenius map gives $2\xi^a = 2\xi^{b+c}$. Therefore $\xi^a \equiv \xi^{b+c} \pmod 2$, so if we write $x = \theta^a$, $y = \theta^b$, $z = \theta^c$ we have $x = yz$. But also $1 + x = y + z$, so $(y+1)(z+1) = 0$, which since θ is primitive in $GF(2^m)$ implies y or $z = 1$. ■

(P4) For odd $m \geq 3$,

$$\xi^i + \xi^j + \xi^k + \xi^l = 0 \Rightarrow i = j = k = l \ .$$

Proof. Suppose $\xi^a + \xi^b + \xi^c = -1$. Arguing as in the previous proof we obtain $x^2 + y^2 = (x+1)(y+1)$, hence $u^2 + v^2 = uv$, with $x = u + 1$, $y = v + 1$. Substituting $v = tu$ we find $u^2(t^2 + t + 1) = 0$. But $t^2 + t + 1 \neq 0$ in $GF(2^m)$, m odd, since $tr(t^2 + t + 1) = m \neq 0$, so $u = 0$, $x = 1$, therefore $a = b = c = 0$. ■

Properties P2, P3 and P4 are also consequences of the fact that errors of weight ≤ 2 in the 'Preparata' code can be decoded uniquely, as shown in §5.3.

3.4. The ring $\mathcal{R}$.

As usual when studying cyclic codes of length n it is convenient to represent codewords by polynomials modulo $X^n - 1$. We identify $v = (v_0, v_1, \ldots, v_{n-1})$ with the polynomial $v(X) = \sum_{r=0}^{n-1} v_r X^r$ in the ring $\mathcal{R} = \mathbb{Z}_4[X]/(X^n - 1)$. We must be careful when working with $\mathcal{R}$: it is not a unique factorization domain — for example $X^4 - 1$ has two distinct factorizations into irreducible polynomials in $\mathcal{R}$:

$$\begin{aligned} X^4 - 1 &= (X-1)(X+1)(X^2+1) \\ &= (X+1)^2(X^2 + 2X - 1) \ . \end{aligned}$$

Note also that every element $1 + 2\lambda$, $\lambda \in R$, is a root of $X^2 - 1$. On the other hand $\mathcal{R}$ *is* a principal ideal domain: just as in the binary case, cyclic codes have a single generator (the proof is given in Calderbank and Sloane, Modular and p-adic cyclic codes, *Designs, Codes and Cryptography*, to appear).

IV. Kerdock codes

The main result of this section is a very simple quaternary construction for Kerdock codes.

4.1. The Kerdock code is an extended cyclic code over $\mathbb{Z}_4$.

Let $h(X)$ be a primitive basic irreducible polynomial of degree m, as above, and let $g(X)$ be the reciprocal polynomial to $(X^n - 1)/((X-1)h(X))$, where $n = 2^m - 1$.

Theorem 10. *let $\mathcal{K}^-$ be the cyclic code of length n over $\mathbb{Z}_4$ with generator polynomial $g(X)$, and let $\mathcal{K}$ be obtained from $\mathcal{K}^-$ by adjoining a zero-sum check symbol. Then for odd $m \geq 3$ the binary image $K = \phi(\mathcal{K})$ of $\mathcal{K}$ under the Gray map (15) is a nonlinear code of length 2^{m+1}, with 4^{m+1} words and minimal distance $2^m - 2^{(m-1)/2}$ that is equivalent to the Kerdock code. This code is distance invariant.*

Note that $\mathcal{K}^-$ has parity check polynomial $(X-1)h(X)$. There are two equivalent generator matrices for $\mathcal{K}$. The first is

$$\begin{bmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & \xi & \xi^2 & \cdots & \xi^{n-1} \end{bmatrix}, \tag{49}$$

where the entries in the second row are to be replaced by the corresponding m-tuples $(b_0 b_1 \cdots b_{m-1})'$ (the prime indicating transposition) obtained from (40). Alternatively, let $g(X) = \sum_{j=0}^{\delta} g_j X^j$, $\delta = 2^m - m - 2$, $g_j \in \mathbb{Z}_4$, and let $g_\infty = -\sum_{j=0}^{\delta} g_j$. Then the second form for the generator matrix for $\mathcal{K}$ is

$$\begin{bmatrix} g_\infty & g_0 & g_1 & \cdots & g_\delta & 0 & \cdots & 0 \\ g_\infty & 0 & g_0 & \cdots & g_{\delta-1} & g_\delta & \cdots & 0 \\ \cdot & \cdot & \cdot & & \cdot & \cdot & & \\ g_\infty & 0 & 0 & \cdots & g_0 & g_1 & \cdots & g_\delta \end{bmatrix}. \tag{50}$$

$\mathcal{K}$ is a code of type 4^{m+1}. The binary code $K^{(1)}$ associated with $\mathcal{K}$ (see (19)) is $RM(1,m)$.

For example, with $m = 3$ and h given by (34), we find $g = x^3 + 2x^2 + x - 1$, so the two equivalent generator matrices are

$$\begin{bmatrix} 1 & 3 & 1 & 2 & 1 & 0 & 0 & 0 \\ 1 & 0 & 3 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 3 & 1 & 2 & 1 & 0 \\ 1 & 0 & 0 & 0 & 3 & 1 & 2 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 2 & 3 & 1 \\ 0 & 0 & 1 & 0 & 3 & 3 & 3 & 2 \\ 0 & 0 & 0 & 1 & 2 & 3 & 1 & 1 \end{bmatrix} \tag{51}$$

(the second one being read from (41)).

For $m = 5$, we may take $h(X) = \sum_{i=0}^{5} h_i X^i$, $g(X) = \sum_{i=0}^{25} g_i X^i$, where $h_0 \ldots$ and $g_0 \ldots$ are 323001 and 11120122010303133013212213.

Kerdock codes contain more codewords than any known linear code with the same minimal distance (although we are not aware of any theorem to guarantee this, except at length 16).

4.2. Family $\mathcal{A}$.

If we omit the factor $X-1$ from the parity check polynomial for $\mathcal{K}^-$, we obtain a cyclic code containing 4^m codewords. Let $\mathcal{A}$ denote the family of cyclically distinct vectors obtained from this code by deleting the zero vector and failing to distinguish between a vector and any of its cyclic shifts. The corresponding collection $\Omega(\mathcal{A})$ of complex-valued sequences has been studied in [6], [7], [67], [72] as a family of asymptotically optimal CDMA signature sequences (referred to as Family $\mathcal{A}$ in [7]). Since the sequences of $\Omega(\mathcal{A})$ have low values of auto- and cross-correlation, the set $\Omega(\mathcal{A})$ also has large minimal Euclidean distance.

4.3. Trace description of Kerdock code and proof of Theorem 10.

Theorem 11. *The codes $\mathcal{K}^-$ and $\mathcal{K}$ have the following trace descriptions over the ring R.*

(a) $c = (c_0, c_1, \ldots, c_{n-1})$ is a codeword in $\mathcal{K}^-$ if and only if, for some $\lambda \in R$ and $\epsilon \in \mathbb{Z}_4$,

$$c_t = T(\lambda \xi^t) + \epsilon, \quad t \in \{0, 1, \ldots, n-1\} \ . \tag{52}$$

Thus

$$\mathcal{K}^- = \{\epsilon\, \mathbf{1} + v^{(\lambda)} : \ \epsilon \in \mathbb{Z}_4, \quad \lambda \in R\} \ , \tag{53}$$

where

$$v^{(\lambda)} = (T(\lambda), T(\lambda\xi), T(\lambda\xi^2), \ldots, T(\lambda\xi^{n-1})) \ .$$

(b) $c = (c_\infty, c_0, c_1, \ldots, c_{n-1})$ is a codeword in $\mathcal{K}$ if and only if, for some $\lambda \in R$ and $\epsilon \in \mathbb{Z}_4$,

$$c_t = T(\lambda \xi^t) + \epsilon, \quad t \in \{\infty, 0, 1, \ldots, n-1\} \ , \tag{54}$$

with the convention that $\xi^\infty = 0$.

This theorem is essentially equivalent to Theorem 3 of [7].

Proof. (a) Let C be the code defined by (53). If $c(X)$ is the polynomial form of a codeword in C, then $c(X)(X-1)h(X) = 0$ [the all 1's vector is annihilated by $X-1$ and the $v^{(\lambda)}$ by $h(X)$]. Therefore $C \subseteq \mathcal{K}^-$. Since C and $\mathcal{K}^-$ contain the same number of codewords, $C = \mathcal{K}^-$. (b) follows because the zero-sum check for $\epsilon\,\mathbf{1}$ is ϵ and for $v^{(\lambda)}$ it is 0. ■

Proof of Theorem 10. We consider an arbitrary codeword $c \in \mathcal{K}$ in the form (54). We will show that c_t has 2-adic expansion

$$c_t = a_t + 2b_t, \quad t \in \{\infty, 0, 1, \ldots, n-1\} \ , \tag{55}$$

given by

$$a_t = tr(\pi\theta^t) + A \,, \tag{56}$$

$$b_t = tr(\eta\theta^t) + Q(\pi\theta^t) + B \,, \tag{57}$$

where the elements π, $\eta \in GF(2^m)$ and $A, B \in \mathbb{Z}_2$ are arbitrary,

$$Q(x) = \sum_{j=1}^{(m-1)/2} tr(x^{1+2^j}) \,, \quad x \in GF(2^m) \,,$$

and we adopt the convention that $\theta^\infty = 0$.

Let $\lambda = \xi^r + 2\xi^s$, r, $s \in \{\infty, 0, \ldots, n-1\}$, so that

$$c_t = \epsilon + T(\xi^{r+t}) + 2T(\xi^{s+t}) = a_t + 2b_t \,.$$

Projecting modulo 2, we obtain

$$a_t = \alpha(\epsilon) + tr(\pi\theta^t) \,,$$

where $\pi = \mu(\xi^r)$, $\theta = \mu(\xi)$. To find b_t, we compute $c_t - c_t^2 = 2b_t$ (since $a_t = 0$ or 1) and obtain

$$\begin{aligned} 2b_t &= (\epsilon - \epsilon^2) + (T(\xi^{r+t}) - T^2(\xi^{r+t})) + 2\epsilon T(\xi^{r+t}) + 2T(\xi^{s+t}) \\ &= 2\beta(\epsilon) + 2 \sum_{0 \le j < k \le m-1} (\xi^{r+t})^{2^j + 2^k} + 2T((\epsilon\xi^r + \xi^s)\xi^t) \,. \end{aligned}$$

Thus

$$b_t = \beta(\epsilon) + Q(\pi\theta^t) + tr(\eta\theta^t) \,,$$

where $\eta = \mu(\epsilon\xi^r + \xi^s)$.

The next step is to observe that the vectors (b_t) and $(a_t + b_t)$ defined by (56), (57) are the left and right halves of the codewords in Kerdock's original definition ([46]; [56, p. 458]). But the Gray map ϕ sends c to $(\beta(c), \gamma(c)) = ((b_t), (a_t + b_t))$.

The fact that $\phi(\mathcal{K})$ is distance invariant follows from Theorem 2. ■

It is shown in [8] that when m is odd, the family of binary sequences $\{Q(\pi\theta^t) + tr(\eta\theta^t) : \eta, \pi$ in $GF(2^m)$, not both zero$\}$ has Gold-like correlation properties, but a larger linear span.

4.4. The first-order Reed-Muller subcode.

The vectors for which $\pi = 0$ in (56), (57) form a linear subcode of $\mathcal{K}$, with generator matrix

$$\begin{bmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & 2 & 2\xi & 2\xi^2 & \cdots & 2\xi^{n-1} \end{bmatrix} ,$$

whose binary image is the first-order Reed-Muller code contained in the Kerdock code.

4.5. The Nordstrom-Robinson code.

The case $m = 3$ is particularly interesting. The Kerdock and Preparata codes of length 16 coincide, giving the Nordstrom-Robinson code ([58]; see also [62]). This is the unique binary code of length 16, minimal distance 6, containing 256 words [66], [33]. In this case $\mathcal{K}$ is the 'octacode', whose generator matrix is given in (51). The octacode may also be characterized as the unique self-dual quaternary code of length 8 and minimal Lee weight 6 [23], or as the 'glue code' required to construct the 24-dimensional Leech lattice from eight copies of the face-centered cubic lattice [22, Chap. 24]. Thus the following theorem is a special case of Theorem 10.

Theorem 12. *The Nordstrom-Robinson code is the binary image of the octacode under the Gray map.*

The symmetrized weight enumerator of the octacode is ([23])

$$W^8 + 16X^8 + Y^8 + 14W^4Y^4 + 112WX^4Y(W^2 + Y^2) \ ,$$

and the weight distribution of the Nordstrom-Robinson code is then given by (24).

4.6. Weight distribution.

The weight distribution of any Kerdock code is also easily determined from the new quaternary description.

Theorem 13. *The binary Kerdock code $K = \phi(\mathcal{K})$ of length 2^{m+1} (m odd ≥ 3) has the following weight distribution:*

$$\begin{array}{cc} i & A_i \\ 0 & 1 \\ 2^m - 2^{(m-1)/2} & 2^{m+1}(2^m - 1) \\ 2^m & 2^{m+2} - 2 \\ 2^m + 2^{(m-1)/2} & 2^{m+1}(2^m - 1) \\ 2^{m+1} & 1 \end{array} \qquad (58)$$

(cf. [56], Fig. 15.7).

Proof. This is a slight modification of the argument used in [7] to obtain the correlation distribution of the associated complex sequences. We assume the codewords $c \in \mathcal{K}$ are defined as in Theorem 11. As mentioned in §4.5, the words for which $\pi = 0$ (and $\lambda \notin R^*$ in (53)) form a first-order Reed-Muller code, and account for the words of weights 0, 2^m and 2^{m+1}.

We now consider a word $v^{(\lambda)} \in \mathcal{K}^-$ for $\lambda \in R^*$. Let $n_j = n_j(v^{(\lambda)})$ (see (3)). We claim that there exist $\delta_1, \delta_2 = \pm 1$ so that

$$\begin{aligned} n_0 &= 2^{m-2} - 1 + \delta_1 2^{(m-3)/2}, \quad & n_1 &= 2^{m-2} + \delta_2 2^{(m-3)/2} , \\ n_2 &= 2^{m-2} - \delta_1 2^{(m-3)/2}, & n_3 &= 2^{m-2} - \delta_2 2^{(m-3)/2} . \end{aligned} \tag{59}$$

Let

$$S = \sum_{j=0}^{2^m-2} i^{T(\lambda\xi^j)} = n_0 - n_2 + i(n_1 - n_3) .$$

Then

$$|S|^2 = 2^m - 1 + \sum_{j \neq k} i^{T(\lambda(\xi^j - \xi^k))} .$$

We use properties (P1), (P2), (P3) to rewrite this as

$$|S|^2 = 2^m - 1 + \sum_{\nu \in R^*} i^{T(\nu)} - S - \overline{S} .$$

But it is easily verified that

$$\sum_{\nu \in R^*} i^{T(\nu)} = 0$$

(see [7, p. 1104]), hence

$$(S+1)(\overline{S}+1) = 2^m ,$$

$$(n_0 - n_2 + 1)^2 + (n_1 - n_3)^2 = 2^m .$$

The diophantine equation $X^2 + Y^2 = 2^m$ has a unique solution, so

$$n_0 - n_2 = -1 \pm 2^{(m-1)/2} , \tag{60}$$

$$n_1 - n_3 = \pm 2^{(m-1)/2} . \tag{61}$$

We also know that $\mu(v^{(\lambda)})$ is in the simplex code, so

$$n_1 + n_3 = 2^{m-1} , \tag{62}$$

$$n_0 + n_2 = 2^{m-1} - 1 . \tag{63}$$

(59) follows from (60)–(63).

We now consider the four words of $\mathcal{K}$ obtained from $\epsilon\mathbf{1} + v^{(\lambda)}$ $(\epsilon = 0, 1, 2, 3)$ by appending the zero-sum check symbol ϵ. For $\mathbf{1} + v^{(\lambda)}$, for example, we have

$$n_1 = 2^{m-2} + \delta_1 2^{(m-3)/2}, \quad n_2 = 2^{m-2} + \delta_2 2^{(m-3)/2} \ ,$$

$$n_3 = 2^{m-2} - \delta_1 2^{(m-3)/2}, \quad n_0 = 2^{m-2} - \delta_2 2^{(m-3)/2} \ ,$$

which is a word of Lee weight

$$n_1 + n_3 + 2n_2 = 2^m + \delta_2 2^{(m-1)/2} \ .$$

Of these four words obtained from $v^{(\lambda)}$, two have Lee weight $2^m + 2^{(m-1)/2}$ and two have Lee weight $2^m - 2^{(m-1)/2}$. This holds for all $2^m(2^m - 1)$ words $v^{(\lambda)}$, $\lambda \in R^*$, and establishes (58). ∎

When m is even, $m \geq 2$, a similar argument shows that $\phi(\mathcal{K})$ is a nonlinear code of length 2^{m+1}, with 4^{m+1} codewords, minimal distance $2^m - 2^{m/2}$, and weight distribution

i	A_i
0	1
$2^m - 2^{m/2}$	$2^m(2^m - 1)$
2^m	$2^{m+1}(2^m + 1) - 2$
$2^m + 2^{m/2}$	$2^m(2^m - 1)$
2^{m+1}	1

This code is not as good as a double-error-correcting BCH code.

4.7. Soft-decision decoding of Kerdock codes.

Although in the theoretical development we make a distinction between the quaternary code $\mathcal{K}$ and the associated nonlinear binary code $K = \phi(\mathcal{K})$ (and similarly in Section V between $\mathcal{P} = \mathcal{K}^\perp$ and $P = \phi(\mathcal{P})$), they are really two different descriptions of the same code. For instance, a decoder for the quaternary code obviously provides a decoder for the binary code and conversely.

The following is a new soft-decision decoding algorithm for the Kerdock code. This is comparable in complexity to previously known techniques that were derived from the binary description of the code.

The idea is to extend the fast Hadamard transform (FHT) soft-decision decoding algorithm for the binary first-order Reed-Muller code to the Kerdock code. This provides substantial savings over brute-force correlation decoding. Define

$$\Delta = \{\infty, 0, 1, 2, \ldots, n-1\}, \quad n = 2^m - 1 \ .$$

Brute-force decoding of a received vector $\{v_t : t \in \Delta\}$ requires the computation of its correlation with all possible received signals. In particular, the decoder must compute the correlation

$$\zeta(\lambda, \delta) = \sum_{t \in \Delta} v_t i^{-[T(\lambda \xi^t) + \delta]}$$

for all $\lambda = \xi^r + 2\xi^s$, $r, s \in \Delta$ and all $\delta \in \mathbb{Z}_4$, and find that pair (λ, δ) for which Real$\{\zeta(\lambda, \delta)\}$ is a maximum. Computed directly, this technique requires $4^{m+1}2^m$ multiplications and $4^{m+1}(2^m - 1)$ additions.

An immediate reduction in complexity is obtained by writing

$$\zeta(\lambda, \delta) = i^{-\delta} \sum_{t \in \Delta} v_t i^{-T(\xi^{t+r})} (-1)^{tr(\theta^{t+s})} \ ,$$

where we adopt the convention that for l in Δ, $l + \infty = \infty$. The correlation sums $\zeta(\xi^r + 2\xi^s, \delta)$ may now be viewed (after some reordering of indices) as $i^{-\delta}$ times the Hadamard transform of the 2^m complex vectors $\{v_t i^{-T(\xi^{t+r})}\}$ of length 2^m. Using the FHT, each of these can be computed using $m2^m$ additions/subtractions. Thus the overall requirement is for about 4^m multiplications (one multiplicand is always a power of i) and $m4^m$ additions/subtractions.

This complex-data FHT decoding algorithm is of the same order of complexity as recently published real-data FHT decoders for the Kerdock codes [1], [29] based on the general super-code decoding method of Conway and Sloane [20]. These real-data algorithms perform 2^m FHTs of size 2^{m+1}. Finally, we note that the case $m = 3$ corresponds to decoding the Nordstrom-Robinson code.

V. Preparata codes

In this section we show that the binary image of the dual code $\mathcal{P} = \mathcal{K}^\perp$ is a Preparata-like code with essentially the same properties as Preparata's original code (yet is much simpler to construct).

5.1. The 'Preparata' code is an extended cyclic code over $\mathbb{Z}_4$.

Let $h(X)$ and $g(X)$ be defined as in §4.1.

Theorem 14. *Let $\mathcal{P}^-$ be the cyclic code of length $n = 2^m - 1$ with generator polynomial $h(X)$, and let $\mathcal{P}$ be obtained from $\mathcal{P}^-$ by adjoining a zero-sum check symbol, so that $\mathcal{P} = \mathcal{K}^\perp$. Then for odd $m \geq 3$ the binary image $P = \phi(\mathcal{P})$ of $\mathcal{P}$ under the Gray map (15) is a nonlinear code of length $l = 2^{m+1}$, with 2^{l-2m-2} codewords and minimal distance 6. This code is distance invariant and its weight distribution is the MacWilliams transform of the weight distribution of the Kerdock code of the same length.*

Note that $\mathcal{P}^-$ has parity check polynomial $g(X)$, and that (49), (50) are equivalent parity check matrices for $\mathcal{P}$. Also $\mathcal{P}$ is a code of type 4^{2^m-m-1}. The code $P = \phi(\mathcal{P})$ is the $\mathbb{Z}_4$-dual of K, and we refer to it as a 'Preparata' code, using the quotes to distinguish it from Preparata's original code. It is known that the Preparata code (and P) contains more codewords than any linear code with the same minimal distance [10]. The binary code $P^{(1)}$ associated with $\mathcal{P}$ (see (19)) is $RM(m-2, m)$.

Proof of Theorem 14. It follows from Theorems 2 and 3 that P is distance invariant and its weight distribution is the MacWilliams transform of that of K. By Theorem 24 of [56], Chapter 15, P has the same weight distribution as the original Preparata code. ■

Semakov, Zinoviev and Zaitsev [63] had already shown in 1971 that any code with the same parameters as the Preparata code must be distance invariant.

The decoding algorithm given below provides an alternative proof that $\mathcal{P}$ has minimal Lee weight 6, for odd $m \geq 3$. For even $m \geq 2$, $\mathcal{P}$ contains words of Lee weight 4. For ξ satisfies $\xi^{3t} = 1$, where $t = (2^m - 1)/3$, and since $\xi^t - 1 \in R^*$, by (P1), $\xi^{2t} + \xi^t + 1 = 0$, yielding a word of Lee weight 3 in $\mathcal{P}^-$.

There is one essential difference between P and the original Preparata code. It is known that the latter is contained in the extended Hamming code spanned by its codewords.

Theorem 15. *For odd $m \geq 5$, P is contained in a nonlinear code with the same weight distribution as the extended Hamming code of the same length, and the linear code spanned by the codewords of P has minimal weight 2.*

Proof. The first assertion follows by considering the binary images of the following sequence of codes:

$$ZRM(1, m) \subseteq \mathcal{K} \subseteq ZRM(2, m) \subseteq \cdots \subseteq ZRM(2, m)^\perp \subseteq \mathcal{P} \subseteq ZRM(1, m)^\perp \ . \qquad (64)$$

For the second assertion we use the fact that $\mathcal{P}$ is an extended cyclic code with generator polynomial $h(X) = \sum_{j=0}^m h_j X^j$ (say). Let $h_\infty = -h(1) = \pm 1$, since $h_2(1) = 1$. Then $\mathcal{P}$ has a generator matrix of the form

$$\begin{array}{c} a \\ \\ \\ b \\ \\ \end{array} \begin{bmatrix} h_\infty & h_0 & h_1 & \cdots & h_m & & & \\ h_\infty & 0 & h_0 & \cdots & & h_m & & \\ \cdot & \cdot & \cdot & \cdots & & \cdot & & \\ h_\infty & 0 & 0 & \cdots & 0 & h_0 & \cdots & h_m \\ \cdot & \cdot & \cdot & \cdots & & & & \end{bmatrix} . \tag{65}$$

It follows from (30) that the linear span of $P = \phi(\mathcal{P})$ contains all words of the form $\phi(2\alpha(a) * \alpha(b))$, for $a, b \in \mathcal{P}$. Taking a and b to be as indicated in (65) produces a word of weight 2 in the linear span. ■

Again there is a result of Zaitsev, Zinoviev and Semakov that is relevant: they showed in [75] that any code with the same parameters as the Preparata code is a subcode of a possibly nonlinear code with the same parameter as an extended Hamming code. Theorem 15 answers a question raised in that paper, by providing an example where the Hamming-type code is indeed nonlinear.

As a quaternary linear code, $ZRM(1,m)^\perp$ (see (64)) is the union of $2^m - 1$ translates of $\mathcal{P}$, each nonzero translate having minimal Lee weight 4. The codewords of weight 4 in the binary image of $ZRM(1,m)^\perp$ (a nonlinear code with the same parameters an extended Hamming code) form a Steiner system $S(3,4,2^{m+1})$. It is not difficult to show that this $S(3,4,2^{m+1})$ is identical to the Steiner system formed by the codewords of weight 4 in the classical extended Hamming code of length 2^{m+1}. The blocks of this design are divided equally among the binary images of the $2^m - 1$ nonzero cosets of $\mathcal{P}$. The blocks falling in the binary image of a fixed coset form a Steiner system $S(2,4,2^{m+1})$.

5.2. Transform-domain characterization of 'Preparata' codes.

In spite of the previous theorem, in this section we shall show that the 'Preparata' code $P = \phi(\mathcal{P})$ and Preparata's original code have similar characterizations by finite field transforms.

We define the *Galois ring transform* $\widehat{c} = (\widehat{c}(\lambda))$, $\lambda = 0, 1, \ldots, n-1, n = 2^m - 1$, of a quaternary sequence $c = (c_t)$, $t = 0, 1, \ldots, n-1$, by

$$\widehat{c}(\lambda) = \sum_{t=0}^{2^m-2} c_t \xi^{\lambda t} .$$

The inversion formula

$$c_t = -\sum_{\lambda=0}^{n-1} \widehat{c}(\lambda) \xi^{-\lambda t}$$

follows in the usual way from the fact that

$$\sum_{\lambda=0}^{n-1} \xi^{\lambda} = \frac{1-\xi^n}{1-\xi} = 0 \ .$$

We define the *finite field transform* $\tilde{a} = (\tilde{a}(\lambda))$, $\lambda = 0, 1, \ldots, n-1$, of a *binary* sequence $a = (a_t)$, $t = 0, 1, \ldots, n-1$, by

$$\tilde{a}(\lambda) = \sum_{t=0}^{n-1} a_t \theta^{\lambda t} \ ,$$

where $\theta \in GF(2^m)$ is the image of $\xi \in R$ after reduction modulo 2 (as in §3.1). We define the *half-convolution* $\mathcal{H}(\tilde{a}, \lambda) \in GF(2^m)$ of the sequence $\tilde{a}$ at lag λ by

$$\mathcal{H}(\tilde{a}, \lambda) = \sum_{\substack{\lambda_1 \le \lambda_2 \\ \lambda_1 + \lambda_2 = \lambda}} \tilde{a}(\lambda_1)\tilde{a}(\lambda_2) \ ,$$

where λ_1, $\lambda_2 = 0, 1, \ldots, n-1$. The summation is a half rather than full convolution because we exclude the cases $\lambda_1 > \lambda_2$.

Theorem 16. *The quaternary 'Preparata' code* $\mathcal{P}$ *consists of all vectors* $c = (c_t) \in \mathbb{Z}_4^n$, $t \in \{\infty, 0, 1, \ldots, n-1\}$ *satisfying the Galois ring transform constraints*

$$\begin{aligned} c_\infty + \widehat{c}(0) &= 0 \ , \\ \widehat{c}(1) &= 0 \ . \end{aligned} \tag{66}$$

Proof. This follows from the definition of the Galois ring transform and the parity check matrix for $\mathcal{P}$ given in Eq. (49). ■

Theorem 17. *The binary 'Preparata' code* P *consists of all vectors* $(b, a+b)$ *for which* $a, b \in \mathbb{Z}_2^n$ *satisfy*

$$\begin{aligned} \tilde{a}(0) + a_\infty &= 0 \ , \\ \tilde{a}(1) &= 0 \ , \\ \tilde{b}(0) + b_\infty &= \mathcal{H}(\tilde{a}, 0) + a_\infty \ , \\ \tilde{b}(1) &= \mathcal{H}(\tilde{a}, 1) \ . \end{aligned} \tag{67}$$

Note that equations (66) are over R, whereas equations (67) are over $GF(2^m)$.

Proof. Consider a codeword $c = (c_t) \in \mathcal{P}$, where $c_t = a_t + 2b_t$, $t \in \{\infty, 0, \ldots, n-1\}$. It follows from the previous theorem that

$$\begin{aligned} a_\infty + 2b_\infty + \widehat{a}(0) + 2\widehat{b}(0) &= 0 \ , \\ \widehat{a}(1) + 2\widehat{b}(1) &= 0 \ . \end{aligned} \tag{68}$$

The next step is identify the constraints that (68) places on $\tilde{a}(\lambda)$, $\tilde{b}(\lambda)$. Given $\lambda \in \{0, 1, \ldots, n-1\}$, let

$$\widehat{a}(\lambda) = e_\lambda + 2f_\lambda, \quad \text{where} \quad e_\lambda,\ f_\lambda \in \mathcal{T} \ . \tag{69}$$

We find f_λ indirectly, starting from the inversion formula

$$a_t = -\sum_{\lambda=0}^{n-1} \widehat{a}(\lambda)\xi^{-\lambda t} \ .$$

After squaring and also applying the Frobenius map we obtain

$$a_t = a_t^2 = \sum_{\lambda=0}^{n-1} e_\lambda^2 \xi^{-2\lambda t} + 2 \sum_{\lambda_1 < \lambda_2} e_{\lambda_1} e_{\lambda_2} \xi^{-(\lambda_1+\lambda_2)t}$$

and

$$\begin{aligned} a_t &= -\sum_{\lambda=0}^{n-1} (e_\lambda^2 + 2f_\lambda^2)\xi^{-2\lambda t} \\ &= \sum_{\lambda=0}^{n-1} e_\lambda^2 \xi^{-2\lambda t} + 2\sum_{\lambda=0}^{n-1} (e_\lambda^2 + f_\lambda^2)\xi^{-2\lambda t} \end{aligned}$$

respectively. Comparing these two expressions, and using the uniqueness of the Galois ring transform coefficients, we find

$$2(e_\lambda^2 + f_\lambda^2) = 2 \sum_{\substack{0 \le \lambda_1 < \lambda_2 \le n-1 \\ \lambda_1+\lambda_2 = 2\lambda}} e_{\lambda_1} e_{\lambda_2} \ . \tag{70}$$

Now $\mu(e_\lambda) = \tilde{a}(\lambda)$, $\mu(e_{2\lambda}) = \tilde{a}(2\lambda) = \tilde{a}(\lambda)^2 = \mu(e_\lambda)^2$, so (70) implies

$$\mu(f_\lambda^2) = \sum_{\substack{0 \le \lambda_1 \le \lambda_2 \le n-1 \\ \lambda_1+\lambda_2 = 2\lambda}} \tilde{a}(\lambda_1)\tilde{a}(\lambda_2) \ ,$$

an equation in $GF(2^m)$. Taking the square root of both sides we obtain

$$\mu(f_\lambda) = \mathcal{H}(\tilde{a}, \lambda) \ . \tag{71}$$

From (68), (69), (71) we see that

$$a_\infty + 2b_\infty + e_0 + 2f_0 + 2\widehat{b}(0) = 0 \ ,$$

which implies

$$a_\infty + \mu(e_0) = a_\infty + \tilde{a}(0) = 0 \ ,$$

$$\mu\left(\sqrt{a_\infty e_0}\right) + b_\infty + \mathcal{H}(\tilde{a}, 0) + \tilde{b}(0) = 0$$

and the first and third equations of (67) now follow. The second and fourth equations follow easily from the second equation of (68). ■

For comparison with (67), a transform characterization of Preparata's original code (of the same length 2^{m+1}) can be readily derived from the description given by Baker, van Lint and Wilson [2]: a vector $(b, a+b)$ is in this code if and only if

$$\begin{aligned} \tilde{a}(0) + a_\infty &= 0 \\ \tilde{a}(1) &= 0 \\ \tilde{b}(0) + b_\infty &= 0 \\ \tilde{b}(1)^3 &= \tilde{a}(3) \ . \end{aligned} \tag{72}$$

The similarity between (67) and (72) is evident. At length 16 (the case $m = 3$) the two descriptions must coincide, since the Nordstrom-Robinson code is unique (see §4.5). This may be verified directly as follows.

Theorem 18. *When $m = 3$ the 'Preparata' code P coincides with Preparata's original code.*

Proof. It is enough to show that $\mathcal{H}(\tilde{a}, 0) = a_\infty \tilde{a}(0)$ and $\mathcal{H}(\tilde{a}, 1) = \tilde{a}(3)^{1/3} = \tilde{a}(3)^5$. The cyclotomic cosets mod 7 are $\{0\}$, $\{1, 2, 4\}$ and $\{3, 5, 6\}$. Hence

$$\tilde{a}(2) = \tilde{a}(1)^2, \ \tilde{a}(4) = \tilde{a}(1)^4, \ \tilde{a}(6) = \tilde{a}(3)^2, \ \tilde{a}(5) = \tilde{a}(3)^4 \ .$$

Since $\tilde{a}(1) = 0$ and $\tilde{a}(0) = a_\infty$ are given, we have

$$\begin{aligned} \mathcal{H}(\tilde{a}, 0) &= \tilde{a}(0)^2 + \tilde{a}(1)\tilde{a}(6) + \tilde{a}(2)\tilde{a}(5) + \tilde{a}(3)\tilde{a}(4) \\ &= \tilde{a}(0)^2 \ = \ \tilde{a}(0) a_\infty \ , \\ \mathcal{H}(\tilde{a}, 1) &= \tilde{a}(0)\tilde{a}(1) + \tilde{a}(2)\tilde{a}(6) + \tilde{a}(3)\tilde{a}(5) + \tilde{a}(4)^2 \\ &= \tilde{a}(3)\tilde{a}(5) \ = \ \tilde{a}(3)^5 \ , \end{aligned}$$

as required. ■

As we have already seen in §4.5, the appropriate quaternary code in the case $m = 3$ is the self-dual octacode.

5.3. Decoding the quaternary 'Preparata' code in the $\mathbb{Z}_4$ domain.

There is a very simple decoding algorithm for the 'Preparata' code $\mathcal{P}$, obtained by working in the $\mathbb{Z}_4$ domain. This is an optimal syndrome decoder: it corrects all error patterns of Lee weight at most 2, detects all errors of Lee weight 3, and detects some errors of Lee weight 4. A decision tree for the algorithm is shown in Fig. 2. We use the parity check matrix H given in (49), and assume m is odd and ≥ 3.

Let $v = (v_\infty, v_0, \ldots, v_{n-1}) \in \mathbb{Z}_4^{n+1}$ be the received vector. The syndrome Hv' has two components, which we write as

$$\begin{aligned} t &= \sum_{j=0}^{n-1} v_j + v_\infty \ , \\ A + 2B &= \sum_{j=0}^{n-1} v_j \xi^j \ , \end{aligned}$$

where $A, B \in \mathcal{T}$.

In Theorem 13 we saw that exactly four nonzero weights occur in the Lee weight distribution of the quaternary Kerdock code $\mathcal{K} = \mathcal{P}^\perp$, and hence also in the Hamming weight distribution of K. It follows that the covering radius of P is at most 4 ([26]; [56], Theorem 21 of Chap. 6), i.e. the Lee distance $d_L(v, \mathcal{P})$ from a vector $v \in \mathbb{Z}_4^{n+1}$ to $\mathcal{P}$ satisfies $d_L(v, \mathcal{P}) \leq 4$. Note that $t = \pm 1$ if and only if $d_L(v, \mathcal{P}) = 1$ or 3.

Single errors of Lee weight 1 or 2. If $t = 1$ and $B = 0$, or if $t = -1$ and $A = B$, we decide that there is a single error of Lee weight 1 in column $(1, A)'$. If $t = 1$ and $B \neq 0$, or if $t = -1$ and $A \neq B$, then $d_L(v, \mathcal{P}) = 3$. If $t = 2$ and $A = 0$, we decide that there is a single error of Lee weight 2 in column $(1, B)'$.

Double errors of Lee weight 2. We begin by supposing that $t = 0$ and

$$A + 2B = X - Y \ ,$$

where $X, Y \in \mathcal{T}$ and $X \neq Y$. Note that $A \neq 0$ since by (P1) $X - Y$ is invertible. We have

$$\begin{aligned} A &= X + Y + 2X^{2^{m-1}} Y^{2^{m-1}} \ , \\ B &\equiv Y + X^{2^{m-1}} Y^{2^{m-1}} \pmod 2 \ . \end{aligned}$$

Let x, y, a, b respectively be the images of X, Y, A, B in $GF(2^m)$ after reduction mod 2 using the map μ. Then

$$a = x + y \ , \quad b = y + x^{2^{m-1}} y^{2^{m-1}} \ ,$$

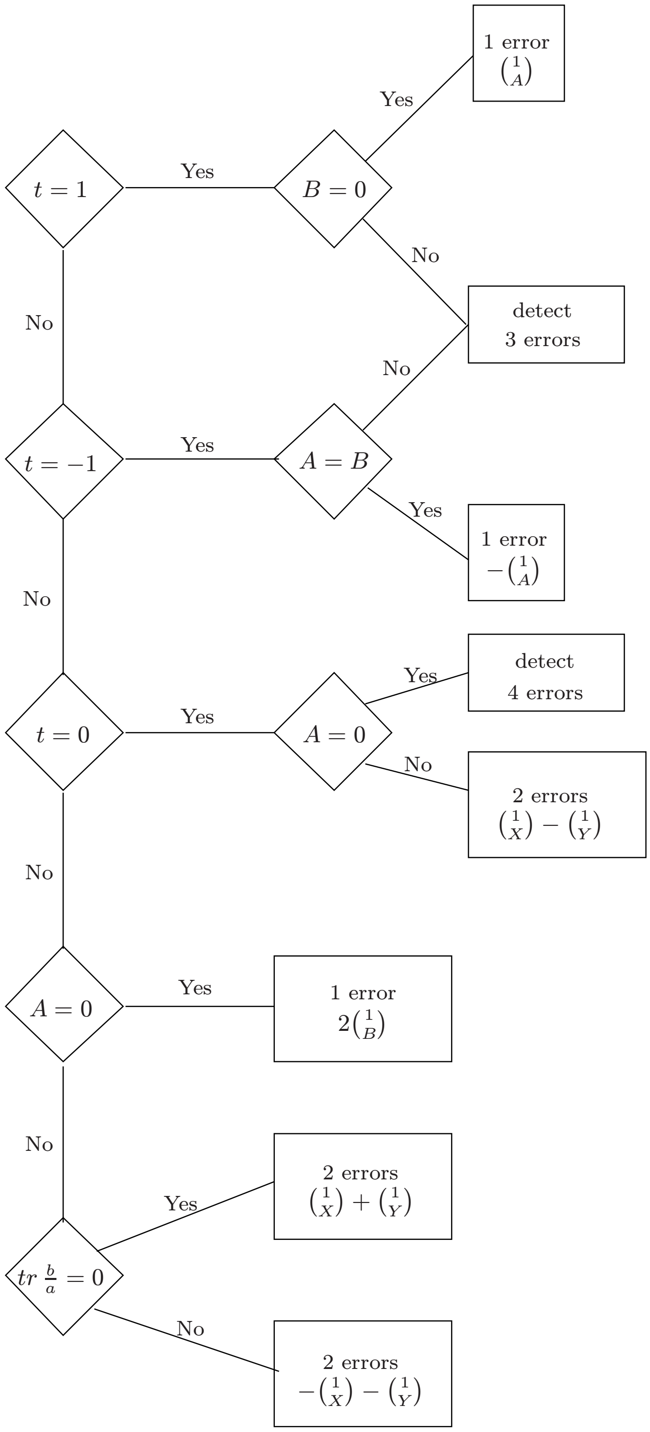


Figure 2: Decoding algorithm for 'Preparata' code

which we rewrite as

$$a = x + y \ , \quad (b + y)^2 = xy \ .$$

The unique solution to these equations is $y = b^2/a$, $x = a + b^2/a$. Note that when $b = 0$ or $b = a$, the double error involves the first column of H. Next we suppose that $t = 2$ and that

$$A + 2B = X + Y \ ,$$

where $X, Y \in \mathcal{T}$, $X \neq Y$, $A \neq 0$. Proceeding as above we find

$$a = x + y \ , \quad b^2 = xy \ ,$$

and so x and y are distinct roots of the equation

$$u^2 + au + b^2 = 0 \ .$$

A necessary and sufficient condition for this equation to have distinct roots is that

$$tr(b^2/a^2) = tr(b/a) = 0$$

(see [56], Chap. 9, Theorem 15; [51]).

Finally we suppose that $t = 2$ and

$$A + 2B = -X - Y \ ,$$

where $X, Y \in \mathcal{T}$, $X \neq Y$, $A \neq 0$. We now find that

$$a = x + y \ , \quad (b + a)^2 = xy \ ,$$

and so x and y are distinct roots of the equation

$$u^2 + au + (a^2 + b^2) = 0 \ .$$

A necessary and sufficient condition for this equation to have distinct roots is that

$$tr\left(\frac{a^2 + b^2}{a^2}\right) = tr\left(1 + \frac{b}{a}\right) = 1 + tr\left(\frac{b}{a}\right) = 0 \ .$$

5.4. Quaternary Reed-Muller codes

In §2.7 we defined a quaternary code $ZRM(r, m-1)$ whose image under the Gray map ϕ is the binary Reed-Muller code $RM(r, m)$, provided $r \in \{0, 1, 2, m-1, m\}$. In this section we define another quaternary Reed-Muller code, $QRM(r, m)$, whose image under the map α is $RM(r, m)$ for all r, and which includes the Kerdock and 'Preparata' codes as special cases.

Definition. Let $QRM(0, m)$ be the quaternary repetition code of length $n = 2^m$, and for $1 \le r \le m$ let $QRM(r, m)$ be generated by $QRM(0, m)$ together with all vectors of the form

$$(0, T(\lambda_j), T(\lambda_j \xi^j), T(\lambda_j \xi^{2j}), \ldots, T(\lambda_j \xi^{(n-1)j}))$$

where j ranges over all representatives of cyclotomic cosets mod $2^m - 1$ for which $wt(j) \le r$, and λ_j ranges over R. Then $QRM(r, m)$ is a quaternary code of length $n = 2^m$ and type 4^k, where

$$k = 1 + \binom{m}{1} + \cdots + \binom{m}{r} .$$

Theorem 19.

$$\begin{aligned} QRM(1, m) &= \mathcal{K} , && (73) \\ QRM(m-2, m) &= \mathcal{P} , && (74) \\ \alpha(QRM(r, m)) &= RM(r, m) , && (75) \\ QRM(r, m)^{\perp} &= QRM(m-r-1, m) . && (76) \end{aligned}$$

Proof. (73) follows from Theorem 11(b), (74) from (76), and (75) from [56, Chap. 13, §5]. It remains to prove (76). This follows from the transform domain characterization of $QRM(r, m)$ as the set of vectors a for which $\widehat{a}(\lambda) = 0$ whenever $wt(\lambda) \le m - 1 - r$, and $QRM(r, m)^{\perp}$ as the set of vectors for which $\widehat{a}(\lambda) = 0$ whenever $wt(\lambda) \le r$. (Equivalently, we consider the cyclic codes obtained by deleting the first coordinate, and use the fact that the zeros of a code are the reciprocals of the nonzeros of the dual code.) ■

5.5. Automorphism groups.

Consider any system Ω of linear equations over $\mathbb{Z}_4$, in the variables c_x, $x \in \mathcal{T}$ (see (36))), that includes

$$\sum_{x \in \mathcal{T}} c_x = 0 , \qquad (77)$$

$$\sum_{x \in \mathcal{T}} c_x x \quad = \quad 0 \ , \tag{78}$$

together with equations of the form

$$2 \left(\sum_{x \in \mathcal{T}} c_x x^{2^j + 1} \right) = 0 \ . \tag{79}$$

Theorem 20. *The linear system Ω is invariant under the doubly transitive group G of 'affine' permutations of the form*

$$x \to (ax + b)^{2^m} \ ,$$

where $a, b \in \mathcal{T}$ and $a \neq 0$. The order of G is $2^m(2^m - 1)$.

Proof. Repeated application of the Frobenius automorphism (43) to Eq. (78) gives

$$\sum_{x \in \mathcal{T}} c_x x^{2^j} = 0 \ , \tag{80}$$

for all j. It now follows from (77), (78) and (80) that

$$\begin{aligned} \sum_{x \in \mathcal{T}} c_x (ax + b)^{2^m} &= \sum_{x \in \mathcal{T}} c_x (a^{2^m} x^{2^m} + b^{2^m} + 2a^{2^{m-1}} b^{2^{m-1}} x^{2^{m-1}}) \\ &= a \sum_{x \in \mathcal{T}} c_x x + b \sum_{x \in \mathcal{T}} c_x + 2a^{2^{m-1}} b^{2^{m-1}} \sum_{x \in \mathcal{T}} x^{2^{m-1}} \quad = \quad 0 \ . \end{aligned}$$

Finally

$$\begin{aligned} 2 \sum_{x \in \mathcal{T}} c_x [(ax + b)^{2^m}]^{2^j + 1} &= 2 \sum_{x \in \mathcal{T}} c_x (a^{2^m} x^{2^m} + b^{2^m})^{2^j + 1} \\ &= 2 \sum_{x \in \mathcal{T}} c_x (ax + b)^{2^j + 1} \\ &= 2 \sum_{x \in \mathcal{T}} c_x (a^{2^j + 1} x^{2^j + 1} + b^{2^j + 1} + a^{2^j} b x^{2^j} + b^{2^j} a x) \quad = \quad 0 \ . \end{aligned}$$

Hence Ω is invariant under G. ■

Corollary. *Our quaternary Kerdock, 'Preparata', 'Goethals', Delsarte-Goethals and 'Goethals-Delsarte' codes are invariant under a doubly transitive group of order $2^{m+1}(2^m - 1)m$ generated by G, negation, and the Frobenius map (43) acting on $\mathcal{T}$.*

Proof. The presence of negation follows from $\mathbb{Z}_4$-linearity, the action of G from Theorem 20 and the Frobenius map from Eq. (80). ■

Remarks. By the automorphism group $\mathrm{Aut}(C)$ of a binary nonlinear code C we will mean the set of all coordinate permutations that preserve the code. It is easy to see that if $C = \phi(\mathcal{C})$ is the binary image of a linear quaternary code $\mathcal{C}$, then $\mathrm{Aut}(\mathcal{C})$ is isomorphic to a subgroup of $\mathrm{Aut}(C)$.

The automorphism groups of the binary Nordstrom-Robinson, Kerdock, classical Preparata, and Delsarte-Goethals codes are known (Berlekamp [3], Carlet [14], [15], [16], Kantor [44, 45]). For odd $m \geq 5$ these groups have the same orders as those in the Corollary.

We conclude that, *for odd $m \geq 5$, the groups mentioned in the Corollary are the full automorphism groups of these quaternary codes.* (For the 'Preparata' codes we use the fact that they have the same automorphism group as their duals.)

The case $m = 3$ is exceptional. The quaternary octacode has an automorphism group of order 1344 (Conway and Sloane [23]), whereas the group of the binary Nordstrom-Robinson code has order 80640 (Berlekamp [3], see also Conway and Sloane [21]).

5.6. A new family of distance regular graphs of diameter 4.

As before, $P = \phi(\mathcal{P}) = \phi(\mathcal{K}^{\perp})$ denotes our 'Preparata' code of length $N = 2^{m+1}$, with m odd ≥ 3.

Definition. A $\mathbb{Z}_4$-*coset* of P is the image under ϕ of a coset of $\mathcal{P}$ in $\mathbb{Z}_4^{N/2}$. We construct a graph Γ_m on the $\mathbb{Z}_4$-cosets of P by joining two cosets by an edge if they are the images of cosets $x + \mathcal{P}$, $y + \mathcal{P}$ such that $x - y + \mathcal{P}$ has minimal Lee weight 1.

Let Π denote the partition of $\mathbb{Z}_2^N$ into $\mathbb{Z}_4$-cosets of P. Then Γ_n can be thought of as the quotient graph ([9, §11.1.B]) of the N-hypercube by the partition Π.

The aim of this section is to show that Γ_m is distance regular and to compute its distance distribution diagram and eigenmatrix $\mathbf{P}$. For this purpose we need certain regularity properties of P and Π.

If C is a binary code of length N, its outer distribution matrix $B = (B_{x,j})$ is the $2^N \times (N+1)$ matrix with typical entry

$$B_{x,j} = |\{y \in C : \ d(x, y) = j\}|$$

(Delsarte [26]). In other words the rows of B are the weight distributions of the translates of C.

A *code* C of covering radius r is said to be *completely regular* [27] if B contains exactly $r+1$

distinct rows. A *partition* Π of $\mathbb{Z}_2^N$ into cosets is said to be *completely regular* if all members of the partition are completely regular with the same matrix.

Lemma 1. *The covering radius of P is 4.*

Proof. In the previous section we saw that it is at most 4. But P is contained in a code with the same weight distribution as an extended Hamming code (Theorem 15), and so by the supercode lemma [17] the covering radius is at least 4. ■

Lemma 2. *The codewords of weight 6 in P form a $3-(2^{m+1}, 6, (2^{m+1}-4)/3)$ design.*

Proof. The proof of Theorem 33 of [56], Chap. 15 can be used, since it depends only on the annihilator polynomial of P. ■

Theorem 21. *The 'Preparata' code P is completely regular.*

Proof. The well-known recurrence relation between the columns of B ([26], [56]) has order 4, by Lemma 1, and so it is sufficient to check that $B_{x,j}$ can take at most five different values for fixed $x \in \mathbb{Z}_2^N$ and $0 \le j \le 4$. If $d(x, P) \le 2$, the fact that P has minimal distance 6 shows that $B_{x,j}$ is either 0 or 1. If $d(x, P) = 3$ then Lemma 2 shows that $B_{x,3} = (N-1)/3$. Clearly $B_{x,0} = B_{x,1} = B_{x,2} = 0$. Finally if $d(x, P) = 4$ then $B_{x,0} = B_{x,1} = B_{x,2} = B_{x,3} = 0$. ■

As in [25] it will be noticed that P is neither linear, perfect, nor uniformly packed, and so (in the notation of Levenshtein [50]) is not a design of Delsarte type (i.e. $d \ge 2s' - 1$); P is a highly nontrivial example of a completely regular code. Furthermore the $\mathbb{Z}_4$-linearity of P and the properties of ϕ show that each $\mathbb{Z}_4$-coset of P is completely regular with the same outer distribution matrix. Hence Π is completely regular. The next result follows immediately from Theorems 11.1.6 and 11.1.5 of [9].

Theorem 22. *The graph Γ_m is distance regular on N^2 vertices with diameter 4 and degree N.*

We now proceed to a more detailed study of the parameters of Γ_m. Recall that the valencies v_j are the numbers of points at distance j from a given point. The intersection numbers a_j, b_j, c_j are defined in Chapter 4 of [9].

Lemma 3. *Γ_m is bipartite.*

Proof. Let us take a parity check matrix H of the form (49) for $\mathcal{P}$. For a coset $x+\mathcal{P}$ let Hx' be the associated syndrome and let $\nu(x)$ be the leading bit of Hx'. Then ν is a map from the vertices of Γ_m onto $\{0,1\}$. Let X_j be the set $\nu^{-1}(j)$, $j=0,1$. Since $\nu(x)=1$ if x has weight 1, two cosets with the same image under ν cannot have adjacent images in Γ_m. ■

Lemma 4. *If $x \in \mathbb{Z}_2^n$ is at distance 4 from P, then $B_{x,4} = N(N-1)/12$.*

Proof. From [68] and the fact that P has size $2^N/N^2$ and four nonzero dual distances d_1', d_2', d_3', d_4' we obtain

$$B_{x,4} = \frac{2^4}{4!N^2}\prod_{j=1}^{4} d_j' \ .$$

The desired result then follows from $d_1' = (N-\sqrt{N})/2$, $d_2' = N/2$, $d_3' = (N+\sqrt{N})/2$, $d_4' = N$. ■

Theorem 23. *The valencies of Γ_m are $v_0 = 1$, $v_1 = N$, $v_2 = \binom{N}{2}$, $v_3 = \frac{N(N-2)}{2}$, $v_4 = \frac{N-2}{2}$. The intersection numbers of Γ_m are $b_0 = N$, $c_1 = 1$, $b_1 = N-1$, $c_2 = 2$, $b_2 = N-2$, $c_3 = N-1$, $b_3 = 1$, $c_4 = N$. Furthermore $a_j = 0$ for $j = 0,1,2,3,4$.*

Proof. By Lemma 3, Γ_m is bipartite, hence without circuits of odd length. Therefore $a_j = 0$ for $0 \le j \le 4$.

The intersection numbers add up to the degree, so $N = b_j + c_j$ for $0 \le j \le 4$, and it only remains to calculate the c_j. The values of c_1 and c_2 are clear from the double-error-correcting character of $\mathcal{P}$. Finally c_3 and c_4 are computed from the formula of Theorem 11.1.8 of [9] by observing that $e_{l,j} = B_{x,j}$ if $d(x,P) = l$. Moreover $e_{3,3} = (N-1)/3$ by Lemma 2 and $e_{4,4} = N(N-1)/12$ by Lemma 4. The intersection numbers of the N-cube are well known to be $a_j = 0$, $b_j = N-j$, $c_j = j$. ■

Corollary. *The eigenmatrix* **P** *for Γ_m is*

$$P = \begin{pmatrix} 1 & N & \binom{N}{2} & \dfrac{N(N-2)}{2} & \dfrac{N-2}{2} \\ 1 & \sqrt{N} & 0 & -\sqrt{N} & -1 \\ 1 & 0 & -\dfrac{N}{2} & 0 & 1-\dfrac{N}{2} \\ 1 & -\sqrt{N} & 0 & \sqrt{N} & -1 \\ 1 & -N & \binom{N}{2} & -\dfrac{N(N-2)}{2} & \dfrac{N}{2}-1 \end{pmatrix}$$

Proof. See [9], §4.1.B, or [25], Proposition 3.17. ■

Remarks. 1) Let R_j denote the jth class of the association scheme corresponding to Γ_m. Then $R_1 + R_3$ has only three eigenvalues, N^2, 0, $-N^2$, and is a strongly regular graph isomorphic to the complete bipartite graph $K_{N^2/2,\, N^2/2}$. It would be interesting to see if R_3 is also distance regular.

2) Π is a 4-partition design in the sense of [12], [25].

VI. Goethals, Delsarte-Goethals, and other codes

It is natural to wonder how the constructions of $\mathcal{K}$ and $\mathcal{P}$ can be generalized. We have already seen one generalization in §5.4. Another generalization is to replace (49) by the matrix

$$\begin{bmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & \xi & \xi^2 & \cdots & \xi^{(n-1)} \\ 0 & 2 & 2\xi^3 & 2\xi^6 & \cdots & 2\xi^{3(n-1)} \\ & \cdot & \cdot & \cdot & \cdots & \\ 0 & 2 & 2\xi^{1+2^j} & 2\xi^{2(1+2^j)} & \cdots & 2\xi^{(1+2^j)(n-1)} \\ & \cdot & \cdot & \cdot & \cdots & \\ 0 & 2 & 2\xi^{1+2^r} & 2\xi^{2(1+2^r)} & \cdots & 2\xi^{(1+2^r)(n-1)} \end{bmatrix}, \tag{81}$$

where $1 \le r \le (m-1)/2$. Again we assume m is odd.

Theorem 24. *(a) The quaternary code of length 2^m with generator matrix (81) has type $4^{m+1}2^{rm}$ and minimal Lee weight $2^m - 2^{m-\delta}$, where $\delta = \frac{m+1}{2} - r$. The binary image under the Gray map (15) is the Delsarte-Goethals code $DG(m+1, \delta)$ ([28]; [56], Chap. 15). (b) The dual code, with parity check matrix (81), has a binary image with minimal distance 8 and the same weight distribution as the Goethals-Delsarte code $GD(m+1, r+2)$ defined by Hergert [40]. In particular, for $r = 1$ this produces a binary code G with the same weight distribution as the Goethals code $\mathcal{I}(m+1)$ ([31]; [32]; [56], Chap. 15).*

Proof. (a) Comparing Eqs. (37) and (34) of [56], Chap. 15, we see that the difference between the Kerdock code and the Delsarte-Goethals code comes from the vectors (c, c), where c belongs to the code defined by Eq. (31) of that chapter. We already know from Theorem 10 that the first two rows of (81) produce the Kerdock code, and it is easily seen that the remaining rows produce the required (c, c) words. (b) This follows because the Goethals-Delsarte code is by construction (see Hergert [40]) a distance invariant codes whose weight enumerator is the MacWilliams transform of the Delsarte-Goethals code. The minimal Lee distance of these dual codes is no more than 8, since they contain words of shape 2^4, corresponding to the doubles of words in the extended Hamming code defined by the binary images of the first two rows of (81). That the minimal Lee distance is at least 8 follows from Theorem 25 below. ■

Remarks. 1) There are also transform-domain characterizations of some of these codes. For the 'Goethals' codes and the dual codes defined in part (b) of Theorem 24, add to (67) the conditions

$$\tilde{a}(1+2^i)=0, \quad i=1,2,\ldots,r \ .$$

For the original Goethals codes [56, p. 477], replace (72) by

$$\begin{aligned} \tilde{a}(0) \ + \ a_\infty &= 0 \ , \\ \tilde{a}(1) &= 0 \ , \\ \tilde{b}(0) \ + \ b_\infty &= 0 \ , \\ \tilde{a}(r) &= \tilde{b}(1)^r \ , \\ \tilde{a}(s) &= \tilde{b}(1)^s \ , \end{aligned}$$

where $r=1+2^{t-1}$, $s=1+2^t$, and a, b are binary vectors of length $n=2^{2t+1}$.

2) For the automorphism groups of these codes, see Section 5.5.

3) Our 'Goethals' code is thus defined as $G=\phi(\mathcal{G})$, where $\mathcal{G}$ is the quaternary code with parity check matrix

$$\begin{bmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & \xi & \xi^2 & \cdots & \xi^{(n-1)} \\ 0 & 2 & 2\xi^3 & 2\xi^6 & \cdots & 2\xi^{3(n-1)} \end{bmatrix} . \qquad (82)$$

We end by giving a direct proof that this code has minimal distance 8.

Theorem 25. *The minimal distance of the 'Goethals' code $G=\phi(\mathcal{G})$ of length 2^{m+1}, m odd ≥ 3, is 8.*

Proof. Since $\mathcal{G}\subseteq\mathcal{P}$, the minimal distance d is at least 6. Suppose, seeking a contradiction, that $c=(c_\infty,c_0,\ldots,c_{n-1})$ is a codeword of type $(\pm1)^{n_1}2^{n_2}0^{n+1-n_1-n_2}$, where $n=2^m-1$, $n_1+2n_2=6$. Write $c=2c_0+c_1$, where $2c_0$ is a vector of type $2^{n_2}0^{n+1-n_2}$ and c_1 is a vector of type $(\pm1)^{n_1}0^{n+1-n_1}$. Then c_1 is orthogonal to every row of the matrix

$$2\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & \xi & \cdots & \xi^{n-1} \\ 0 & 1 & \xi^3 & \cdots & \xi^{3(n-1)} \end{bmatrix} ,$$

and so $\alpha(c_1)$ is in the extended double-error-correcting BCH code of length 2^m. It follows that $n_2 = 0$ and $n_1 = 6$, and in fact that c must be of the type $1^3 3^3 0^{n-5}$ or $\pm(1^5 3\ 0^{n-5})$.

Case 1. c is of type $1^3 3^3 0^{n-5}$. The automorphism group of $\mathcal{G}$ is doubly transitive on the coordinate positions (Theorem 20), so we may assume $c_\infty = -1$. Thus c determines a solution to the equations

$$\begin{aligned} X_1 + X_2 + X_3 &= Z_1 + Z_2 \ , \\ X_1^3 + X_2^3 + X_3^3 &\equiv Z_1^3 + Z_2^3 \pmod 2 \ , \end{aligned}$$

where X_1, X_2, X_3, Z_1, Z_2 are distinct nonzero elements of $\mathcal{T}$. If x_1, x_2, etc., are the images of these elements in $GF(2^m)$ under μ, we have

$$\begin{aligned} x_1 + x_2 + x_3 &= z_1 + z_2 \ , \\ x_1^3 + x_2^3 + x_3^3 &= z_1^3 + z_2^3 \ , \\ x_1x_2 + x_2x_3 + x_3x_1 &= z_1z_2 \ . \end{aligned}$$

But this implies

$$\begin{aligned} x_1x_2x_3 &= (x_1 + x_2 + x_3)^3 + (x_1^3 + x_2^3 + x_3^3) \\ &\quad + (x_1 + x_2 + x_3)(x_1x_2 + x_2x_3 + x_3x_1) \\ &= (z_1 + z_2)^3 + z_1^3 + z_2^3 + z_1z_2(z_1 + z_2) \ = \ 0 \ , \end{aligned}$$

which is a contradiction.

Case 2. c is of type $1^5 3\ 0^{n-5}$. By using the automorphism group we may suppose $c_\infty = 3$, $c_0 = 1$. Thus c determines a solution to

$$\begin{aligned} X_1 + X_2 + X_3 &= -1 - Z_1 \ , \\ X_1^3 + X_2^3 + X_3^3 &\equiv -1 - Z_1^3 \pmod 2 \ , \end{aligned}$$

where X_1, X_2, X_3, 1, Z_1 are distinct nonzero elements of $\mathcal{T}$. Proceeding as before we find

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 + z_1 \ , \\ x_1^3 + x_2^3 + x_3^3 &= 1 + z_1^3 \ , \\ x_1x_2 + x_2x_3 + x_3x_1 &= 1 + z_1 + z_1^2 \ . \end{aligned}$$

For $i = 1, 2, 3$ let $y_i = x_i + 1 + z_1$. This change of variables produces the equations

$$\begin{aligned} y_1 + y_2 + y_3 &= 0 \ , \\ y_1^3 + y_2^3 + y_3^3 &= z_1(1 + z_1) \ , \\ y_1 y_2 + y_2 y_3 + y_3 y_1 &= z_1 \ . \end{aligned} \tag{83}$$

Now write $y_2 = a y_1$, $y_3 = (1 + a) y_1$, so that

$$\begin{aligned} y_1^3(a + a^2) &= z_1 + z_1^2 \ , \\ y_1^2(1 + a + a^2) &= z_1 \ , \end{aligned}$$

and also $y_1 \neq 0$. It follows that

$$y_1^2(1 + a^2 + a^4) + y_1(a + a^2) + (1 + a + a^2) = 0 \ .$$

Setting $s = a + a^2$, we obtain the quadratic equation

$$s^2 + \frac{(1 + y_1)}{y_1^2} s + \frac{1 + y_1^2}{y_1^2} = 0 \ . \tag{84}$$

We shall prove that this equation has no solution. First observe that $y_1 \neq 1$, so the equation does not have a double root. Suppose the equation has two distinct roots. It follows from (83) that there exist distinct nonzero elements Y_1, Y_2, Y_3, Y_2', Y_3', $Z_1^{1/2}$ of $\mathcal{T}$ such that

$$Y_1 + Y_2 + Y_3 = 2Z_1^{1/2} = Y_1 + Y_2' + Y_3' \ .$$

However, this implies the existence of codewords in the 'Preparata' code of type $1^2 3^2 \ 0^{n-3}$, which is not the case. Hence (84) has no solutions and the proof is complete. ■

We are presently investigating other generalizations of (49).

VII. Conclusions

The classical theory of cyclic codes, which includes BCH, Reed-Solomon, Reed-Muller codes, etc., regards these codes as ideals in polynomial rings over finite fields. Some famous nonlinear codes found by Nordstrom-Robinson, Kerdock, Preparata, Goethals and others, more powerful than any linear codes, cannot be handled by this machinery. We have shown that when suitably defined all these codes are ideals in polynomial rings over the ring of integers mod 4. This new point of view should completely transform the study of cyclic codes.

Acknowledgements

We thank Claude Carlet, Pascale Charpin, Dave Forney, Tor Helleseth, Vladimir Levenshtein, Kyeongcheol Yang and Viktor Zinoviev for helpful discussions and comments.

References

[1] J. P. Adoul, *Fast ML decoding algorithm for the Nordstrom-Robinson code*, IEEE Trans. Inform. Theory, **33** (1987), 931–933.

[2] R. D. Baker, J. H. van Lint and R. M. Wilson, *On the Preparata and Goethals codes*, IEEE Trans. Inform. Theory, **29** (1983), 342–345.

[3] E. R. Berlekamp, *Coding theory and the Mathieu groups*, Information and Control, **18** (1971), 40–64.

[4] M. R. Best, *Binary codes with a minimum distance of four*, IEEE Trans. Inform. Theory, **26** (1980), 738–742.

[5] I. M. Boyarinov, *Parallel decoding of Preparata codes* (in Russian), Voprosy Kibernet. No. 135 (1988), 170–179. (See Math. Rev. 89g: 94034.)

[6] S. Boztaş, *Near-optimal* 4ϕ *(4-phase) sequences and optimal binary sequences for CDMA*, Ph.D. dissertation, Univ. Southern Calif., Los Angeles, 1990.

[7] S. Boztaş, A. R. Hammons, Jr., and P. V. Kumar, *4-Phase sequences with near-optimum correlation properties*, IEEE Trans. Inform. Theory, **38** (1992), 1101–1113.

[8] S. Boztaş and P. V. Kumar, *Binary sequences with Gold-like correlation but larger linear span*, IEEE Trans. on Inform. Theory, March 1994, to appear.

[9] A. E. Brouwer, A. M. Cohen and A. Neumaier, *Distance-regular graphs*, Springer-Verlag, NY, 1989.

[10] A. E. Brouwer and L. M. G. M. Tolhuizen, *A sharpening of the Johnson bound for binary linear codes and the nonexistence of linear codes with the Preparata properties*, preprint, 1992.

[11] A. R. Calderbank, A. R. Hammons, Jr., P. V. Kumar, N. J. A. Sloane and P. Solé, *A linear construction for certain Kerdock and Preparata codes*, Bull. Amer. Math. Soc., **29** (1993), 218–222.

[12] P. Camion, B. Courteau and P. Delsarte, *On r-partition designs in Hamming space*, AAECC, **2** (1992), 147–162.

[13] C. Carlet, *A simple description of Kerdock codes*, Lect. Notes Computer Science, **388** (1989), 202–208.

[14] C. Carlet, *The automorphism groups of the Kerdock codes*, J. Information and Optimization Sciences, **12** (No. 3, 1991), 387–400.

[15] C. Carlet, *Les groupes d'automorphism des codes de Delsarte-Goethals*, C.R. Acad. Sci. Paris Série I, **315** (1992), 475–478.

[16] C. Carlet, *The automorphism groups of the Delsarte-Goethals codes*, Designs, Codes and Cryptography, **3** (1993), 237–249.

[17] G. D. Cohen, M. R. Karpovsky, H. F. Mattson, Jr., and J. R. Schatz, *Covering radius-survey and recent results*, IEEE Trans. Inform. Theory, **31** (1985), 328–343.

[18] M. Cohn and A. Lempel, *On fast m-sequence transforms*, IEEE Trans. Inform. Theory, **23** (1977), 135–137.

[19] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker and R. A. Wilson, *ATLAS of finite groups*, Oxford Univ. Press, 1985.

[20] J. H. Conway and N. J. A. Sloane, *Soft decoding techniques for codes and lattices, including the Golay code and the Leech lattice*, IEEE Trans. Inform. Theory, **32** (1986), 41–50.

[21] J. H. Conway and N. J. A. Sloane, *Orbit and coset analysis of the Golay and related codes*, IEEE Trans. Inform. Theory, **36** (1990), 1038–1050.

[22] J. H. Conway and N. J. A. Sloane, *Sphere-packings, lattices and groups*, 2nd ed., Springer-Verlag, NY 1992.

[23] J. H. Conway and N. J. A. Sloane, *Self-dual codes over the integers modulo 4*, J. Comb. Theory, Ser. A, **62** (1993), 30–45.

[24] J. H. Conway and N. J. A. Sloane, *Quaternary constructions for the binary single-error correcting codes of Julin, Best, and others*, Designs, Codes and Cryptography, **41** (1994), 31–42.

[25] B. Courteau and A. Montpetit, *A class of codes admitting at most three nonzero dual distances*, Discrete Math. **90** (1991), 267–280.

[26] P. Delsarte, *Four fundamental parameters of a code and their combinatorial significance*, Inform. Control, **23** (1973), 407–438.

[27] P. Delsarte, *An algebraic approach to the association schemes of coding theory*, Philips Research Reports Supplements, No. 10 (1973).

[28] P. Delsarte and J. M. Goethals, *Alternating bilinear forms over* $GF(q)$, J. Combin. Theory, Series A, **19** (1975), 26–50.

[29] M. Elia, C. Losana, and F. Neri, *A note on the complete decoding of Kerdock codes*, IEEE Internat. Symp. Inform. Theory (June 1988), Kobe, Japan.

[30] G. D. Forney, Jr., N. J. A. Sloane and M. D. Trott, *The Nordstrom-Robinson code is the binary image of the octacode*, Proceedings DIMACS/IEEE Workshop on Coding and Quantization, Amer. Math. Soc., 1993, to appear.

[31] J. M. Goethals, *Two dual families of nonlinear binary codes*, Electronics Letters, **10** (1974), 471–472.

[32] J. M. Goethals, *Nonlinear codes defined by quadratic forms over* $GF(2)$, Inform. Control, **31** (1976), 43–74.

[33] J. M. Goethals, *The extended Nadler code is unique*, IEEE Trans. Inform. Theory, **23** (1977), 132–135.

[34] A. R. Hammons, Jr., *On four-phase sequences with low correlation and their relation to Kerdock and Preparata codes*, Ph.D. Dissertation, Univ. Southern California, Nov. 1992.

[35] A. R. Hammons, Jr. and P. V. Kumar, *Applications of Galois ring theory to sequence design and related topics*, Technical Report CSI-91-05-05, Communication Sciences Institute, University of Southern California, Los Angeles, CA, May 1991.

[36] A. R. Hammons, Jr. and P. V. Kumar, *A linear quadriphase interpretation of the binary Kerdock code*, Technical Report CSI-92-04-01, Communication Sciences Institute, University of Southern California, Los Angeles, CA, April 1992.

[37] A. R. Hammons, Jr. and P. V. Kumar, *Majority-logic decoding of certain quadriphase linear block codes*, Technical Report CSI-92-04-02, Communication Sciences Institute, University of Southern California, Los Angeles, CA, April 1992.

[38] A. R. Hammons, Jr. and P. V. Kumar, *On the apparent duality of the Kerdock and Preparata codes*, IEEE 1993 International Symposium on Information Theory, January 17–22, San Antonio, TX.

[39] A. R. Hammons, Jr., P. V. Kumar, A. R. Calderbank, N. J. A. Sloane and P. Solé, *On the apparent duality of the Kerdock and Preparata codes*, Proc. 10th Internat. Sympos. AAECC-10, San Juan, Puerto Rico, May 1993; Lecture Notes in Computer Science, **673**, ed. G. Cohen, T. Mora and O. Moreno, Springer-Verlag, NY, 1993, pp. 13–24.

[40] F. B. Hergert, *On the Delsarte-Goethals codes and their formal duals*, Discrete Math., **83** (1990), 249–263.

[41] T. W. Hungerford, *Algebra*, Springer-Verlag, NY, 1974.

[42] D. Julin, *Two improved block codes*, IEEE Trans. Inform. Theory, **11** (1965), 459.

[43] W. M. Kantor, *An exponential number of generalized Kerdock codes*, Inform. Control, **53** (1982), 74–80.

[44] W. M. Kantor, *Spreads, translation planes and Kerdock sets*, SIAM J. Alg. Discr. Math., **3** (1982), 151–165 and 308–318.

[45] W. M. Kantor, *On the inequivalence of generalized Preparata codes*, IEEE Trans. Inform. Theory, **29** (1983), 345–348.

[46] A. M. Kerdock, *A class of low-rate nonlinear binary codes*, Inform. Control, **20** (1972), 182–187.

[47] M. Klemm, *Über die Identität von MacWilliams für die Gewichtsfunktion von Codes*, Arch. Math., **49** (1987), 400–406.

[48] M. Klemm, *Selbstduale codes über dem Ring der ganzen Zahlen modulo 4*, Arch. Math., **53** (1989), 201–207.

[49] P. V. Kumar, T. Helleseth, A. R. Calderbank and A. R. Hammons, Jr., *Large families of quaternary sequences with low correlation*, preprint, 1993.

[50] V. I. Levenshtein, *Designs as maximum codes in polynomial metric spaces*, Acta Applicandae Mathematicae **29** (1992), 1–82.

[51] R. Lidl and H. Niederreiter, *Finite fields*, Encyclopedia of Mathematics and Its Applications, vol. 20, Addison-Wesley, Reading, MA, 1983.

[52] R. A. Liebler and R. A. Mena, *Certain distance-regular digraphs and related rings of characteristic 4*, J. Combin. Theory, Series A, **47** (1988), 111-123.

[53] S. Lin and D. J. Costello, *Error control coding fundamentals and applications*, Prentice-Hall, NJ, 1983.

[54] J. H. van Lint, *Kerdock and Preparata codes*, Congressus Numerantium, **39** (1983), 25–41.

[55] B. R. MacDonald, *Finite rings with identity*, Marcel Dekker, NY, 1974.

[56] F. J. MacWilliams and N. J. A. Sloane, *The theory of error-correcting codes*, North-Holland, Amsterdam, 1977.

[57] A. A. Nechaev, *The Kerdock code in a cyclic form*, Diskret. Mat. **1** (1989), 123–139. English translation in Discrete Math. Appl., **1** (1991), 365–384.

[58] A. W. Nordstrom and J. P. Robinson, *An optimum nonlinear code*, Inform. Control, **11** (1967), 613–616.

[59] K. T. Phelps, *A general product construction for error correcting codes*, SIAM J. Algeb. and Discr. Methods, **5** (1984), 224–228.

[60] K. T. Phelps, *Every finite group is the automorphism group of some perfect 1-code*, J. Comb. Theory, Series A, **43** (1986), 45–51.

[61] F. P. Preparata, *A class of optimum nonlinear double-error correcting codes*, Inform. Control, **13** (1968), 378–400.

[62] N. V. Semakov and V. A. Zinoviev, *Perfect and quasi-perfect equal-weight codes*, Problemy Peradachi Informatsii, **5** (No. 2, 1969), 14–18. English translation in Problems of Information Transmission, **5** (No. 2, 1969), 11–13.

[63] N. V. Semakov, V. A. Zinoviev and G. V. Zaitsev, *Uniformly packed codes*, Problemy Peradachi Informatsii **7** (No. 1, 1971), 38–50. English translation in Problems of Information Transmission, **7** (No. 1, 1971), 30–39.

[64] P. Shankar, *On BCH codes over arbitrary integer rings*, IEEE Trans. Inform. Theory, **25** (1979), 480–483.

[65] N. J. A. Sloane and D. S. Whitehead, *A new family of single-error correcting codes*, IEEE Trans. Inform. Theory, **16** (1970), 717–719.

[66] S. L. Snover, *The uniqueness of the Nordstrom-Robinson and the Golay binary codes*, Ph.D. Dissertation, Math. Dept., Michigan State Univ., 1973.

[67] P. Solé, *A quaternary cyclic code, and a family of quadriphase sequences with low correlation properties*, Lect. Notes Computer Science, **388** (1989), 193–201.

[68] P. Solé, *An inversion formula for Krawtchouk polynomials with application to coding theory*, J. Inform. Optimization Sciences, **11** (1990), 207–213.

[69] P. Solé, *Completely regular codes and completely transitive codes*, Discrete Math., **81** (1990), 193–201.

[70] J. V. Uspensky, *Theory of equations*, McGraw-Hill, NY, 1948.

[71] M. Yamada, *Distance-regular digraphs of girth 4 over an extension ring of* $Z/4Z$, Graphs and Combinatorics, **6** (1990), 381–394.

[72] P. Udaya and M. U. Siddiqi, *Large linear complexity sequences over* $\mathbb{Z}_4$ *for quadriphase modulated communication systems having good correlation properties*, IEEE Internat. Symp. Inform. Theory, Budapest, Hungary, June 1991.

[73] J. L. Vasil'ev, *On nongroup close-packed codes*, Probl. Kibern., **8** (1962), 337–339. English translation in Probleme der Kybernetik, **8** (1965), 92–95.

[74] K. Yang, A. R. Hammons, Jr., and P. V. Kumar, *Constructions for lattices based on 4-phase sequences*, Technical Report CSI-92-06-02, Communication Sciences Institute, University of Southern California, Los Angeles, CA, June 1992.

[75] G. V. Zaitzev, V. A. Zinoviev and N. V. Semakov, *Interrelation of Preparata and Hamming codes and extension of Hamming codes to new double-error correcting codes*, in Proc. Second Internat. Symp. Information Theory, Tsahkadsor, Armenia, 1971; Akadémiai Kiadó, Budapest, 1973, pp. 257–263.

[76] O. Zariski and P. Samuel, *Commutative algebra*, Van Nostrand, Princeton, NJ, 1960, 2 vols.

Key words and phrases

Kerdock code, Preparata code, Nordstrom-Robinson code, Goethals code, Delsarte-Goethals code, Goethals-Delsarte code, octacode, nonlinear codes, quaternary codes, Reed-Muller codes, cyclic codes, completely regular codes.

List of Figure Captions