Lecturas Matemáticas
Volumen **30** *(2009), páginas 63–72*
ISSN 0120–1980

An optimal linear code over $\mathbb{F}_{11}$

Henry Chimal-Dzul & Javier Díaz-Vargas
Universidad Autónoma de Yucatán, Mérida, México.

Abstract. In this work, we design a $[7, 5, 3]$ optimal linear code over $\mathbb{F}_{11}$ and use the same method to obtain a $[7, 5, 3]$ optimal linear code over $\mathbb{F}_7$, equivalent to the one suggested in the tables of Andries E. Brouwer [1]. We construct a linear code over $\mathbb{F}_{11}$ which yields a solution to a practical problem in dealing with students entering registration information. We also describe encoding and decoding algorithms for this code.

Key words and phrases. Linear codes, optimal code, MDS code.

2000 AMS Mathematics Subject Classification. 94B05, 94B35.

Resumen. En este trabajo diseñamos un $[7; 5; 3]$ código lineal óptimo sobre $\mathbb{F}_{11}$ y usamos el mismo método para obtener un $[7; 5; 3]$ código lineal óptimo sobre $\mathbb{F}_7$, equivalente al sugerido en las tablas de Andries E. Brower [1]. Construimos un código lineal sobre $\mathbb{F}_{11}$ que da una solución a un problema práctico de registro de información por estudiantes. También describimos algoritmos de codificación y decodificación para este código.

1. Introduction

Every year, the *Facultad de Matemáticas de Yucatán* (FMAT) selects a delegation of high school students to represent the state of Yucatán in the Mexican Mathematical Olympiad.

The first phase of the selection process consists in inviting all high schools to participate. At most 100 students from each high school are allowed to participate. Each participant from that high school is issued an identification number that serves to identify that student throughout the competition. In the second phase, each student takes a common multiple choice exam and the students use their unique identification number to identify themselves. Students record their identification number on the top of the answer sheet, by filling in circles with a number 2 pencil, and below that, the students fill in the answers to the math questions. This sheet is then machine read.

Thousands of students take this exam each year and routinely about ten percent fail to record their identification number correctly. This failure causes difficulties for the examination committee. The organizing committee asked us help to solve this problem. To this end we designed an optimal, one error–correcting two error–detecting linear code over the finite field $\mathbb{F}_{11}$.

2. On block and linear codes

2.1. Block codes. The terminology and basic concepts used here can be found in [7], [8]. Let $\mathcal{A}$ be a finite set, called the alphabet, n a positive integer and let $\mathcal{A}^n$ be the set of n-tuples from the set $\mathcal{A}$. A *block code* $\mathcal{C}$ of length n over $\mathcal{A}$ is a non-empty subset of $\mathcal{A}^n$. Each element of the code will be called a codeword. The *Hamming distance* for $u = (u_1, \dots, u_n)$ and $v = (v_1, \dots, v_n)$ in $\mathcal{A}^n$ is defined by $d(u, v) = |\{i : 1 \leq i \leq n,\ u_i \neq v_i\}|$. Given a block code $\mathcal{C} \subseteq \mathcal{A}^n$ its *minimal distance* is defined as $d(\mathcal{C}) = \text{mín}\{d(u, v) : u, v \in \mathcal{C}, u \neq v\}$. An (n, M, d) code is any code of length n, M codewords and minimum distance d.

A code is said to be *t-error–detecting* if whenever $t \geq 1$ symbols in any codeword are changed, the resulting vector is not a codeword. We say that a code is *t-error–detecting exactly* if it is t-error–detecting but not $(t+1)$-error–detecting.

Theorem 1([7]) *A code $\mathcal{C}$ is t-error–detecting exactly if and only if $d(\mathcal{C}) = t+1$.*

A code is called *t-error–correcting* if for each $y \in \mathbb{F}_q^n$ there is at most an element $x \in \mathcal{C}$ such that $d(x, y) \leq t$. We say that a code is *t-error–correcting exactly* if it is t-error–correcting but not $(t+1)$-error–correcting. In the following theorem, $\lfloor \cdot \rfloor$ denotes the floor function.

Theorem 2([7]) *A code $\mathcal{C}$ has minimum distance d if and only if it is $\lfloor \frac{d-1}{2} \rfloor$–error—correcting exactly.*

In accordance with this last Theorem, if we know the minimum distance d of a code, we know its error–correcting capability, which is $\lfloor \frac{d-1}{2} \rfloor$, and conversely. In general, is not easy to find the minimum distance of a code, and this is one of the principals problems in Coding Theory.

2.2. Linear codes. Let $\mathbb{F}_q$ the finite field with $q = p^\alpha$ elements, where p is a prime number, α a positive integer, and $\mathcal{C} \subseteq \mathbb{F}_q^n$ a code. The *Hamming weight* for a vector $u \in \mathbb{F}_q^n$ is defined as $w(u) = d(u, \mathbf{0})$, where $\mathbf{0}$ denotes the vector of zeros of $\mathbb{F}_q^n$. The *minimum weight* of the code $\mathcal{C}$ is defined as $w(\mathcal{C}) = \text{mín}\{w(u) : u \in \mathcal{C} \setminus \{\mathbf{0}\}\}$.

We say that $\mathcal{C} \subseteq \mathbb{F}_q^n$ is a *linear code* if $\mathcal{C}$ is a subspace of $\mathbb{F}_q^n$. From now we shall use $[n, k, d]_q$ as the notation for a linear code over $\mathbb{F}_q$ of length n, dimension k and minimum distance d. The number k is called the dimension of the code.

If $\mathcal{C} \subseteq \mathbb{F}_q^n$ is a linear code and $u, v \in \mathcal{C}$ with $u \neq v$, then $u - v \in \mathcal{C} \setminus \{\mathbf{0}\}$ and therefore $w(\mathcal{C}) \leq d(\mathcal{C})$. Conversely, since $\mathbf{0} \in \mathcal{C}$ and $w(u) = d(u, \mathbf{0})$, then $w(\mathcal{C}) \geq d(\mathcal{C})$. Hence, for linear codes, the minimum distance and the minimum weight coincide.

A *generator matrix* for an $[n, k]_q$ code $\mathcal{C}$ is any matrix whose rows form a $\mathbb{F}_q$–basis for $\mathcal{C}$. Note that if $\mathcal{C}$ is an $[n, k]_q$ code and G is a generator matrix for $\mathcal{C}$, then $\mathcal{C} = \{vG : v \in \mathbb{F}_q^k\}$. This relationship provides an encoding scheme for $\mathcal{C}$: if $v \in \mathbb{F}_q^k$ is the original message, then vG is the codeword.

The inner product between two vectors $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$ in $\mathbb{F}_q^n$ is defined as $u \cdot v = \sum_{i=1}^n u_i v_i \in \mathbb{F}_q$. If $\mathcal{C}$ is an $[n, k]_q$ code, the set $\mathcal{C}^\perp = \{u \in \mathbb{F}_q^n : u \cdot v = 0 \text{ for all } v \in \mathcal{C}\}$ is an $[n, n-k]_q$ code called the *dual code* of $\mathcal{C}$. If H is a generator matrix for $\mathcal{C}^\perp$, then $v \in \mathcal{C}$ if only if $Hv^t = \mathbf{0^t}$, where v^t is the transpose of v. Because of this, the matrix H is called the *parity check matrix* of $\mathcal{C}$. The following result permits the easy calculation of the matrix H when the generator matrix has a very special form.

Proposition 1. ([7]) *If an $[n, k]_q$ code $\mathcal{C}$ has a generator matrix $G = [I_k|A]$, then a parity check matrix for $\mathcal{C}$ is $H = [-A^t|I_{n-k}]$, and conversely.*

Let σ be a permutation of size n. For $i = 1, \ldots, n$ let α_i be a non-zero scalar in $\mathbb{F}$. Then the map $\mu : \mathbb{F}_q^n \to \mathbb{F}_q^n$ defined as

$$\mu(v_1, \ldots, v_n) = (\alpha_1 v_{\sigma_1}, \ldots, \alpha_n v_{\sigma_n})$$

is called a *monomial transformation* of degree n. We say that two $[n, k]_q$ codes $\mathcal{C}_1$ and $\mathcal{C}_2$ are *scalar multiple equivalent* if there is a monomial transformation μ of degree n for which $\{\mu(v) : v \in \mathcal{C}_1\} = \mathcal{C}_2$.

Lemma 1. *Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be two $[n, k]_q$ codes, then $\mathcal{C}_1$ is scalar multiple equivalent to $\mathcal{C}_2$ if and only if $\mathcal{C}_1^\perp$ is scalar multiple equivalent to $\mathcal{C}_2^\perp$.*

Proof. Suppose that $\mathcal{C}_1$ is scalar multiple equivalent to $\mathcal{C}_2$. Then, there is a permutation σ of size n, and there are nonzero scalars $\alpha_1, \ldots, \alpha_n$ in $\mathbb{F}_q$ such that

$$\mathcal{C}_2 = \{(\alpha_1 v_{\sigma_1}, \ldots, \alpha_n v_{\sigma_n}) : (v_1, \ldots, v_n) \in \mathcal{C}_1\}.$$

This implies that

$$\mathcal{C}_2^\perp = \left\{(\alpha_1^{-1} u_{\sigma_1}, \ldots, \alpha_n^{-1} u_{\sigma_n}) : (u_1, \ldots, u_n) \in \mathcal{C}_1^\perp\right\}.$$

Therefore, $\mathcal{C}_1^\perp$ is scalar multiple equivalent to $\mathcal{C}_2^\perp$. The converse is similar. ☑

3. Design of the Code

In this section we will design a code that will help us solve the problem posed in section 1. We begin with the design of the key and then proceed to design a code with the proper error–correcting capabilities. Then, we give a method for encoding the keys. The decoding process will be given in the next section.

3.1. Designing the key to encode. Registration for the selection process is carried out on a website. High schools use this webpage to enter the names of the student participants. The entries on the webpage consist of two items, the name of the school and the student participant. When the high school has completed their registration, a 3 digit number (1-999) is assigned to the school and a 2 digit number (1-99) is assigned to the student. The school-student 5 digit number is thus the key to encode. Note that this key is unique to each student.

3.2. Construction of the code. The organizing committee decided that the code must have the following characteristics:

1. The data collected from previous contests showed that some students made one error in transcribing their key into the answer sheet, and rarely were there two errors. Therefore, the organizing committee asked for a 1-error correcting code. By Theorem 2, the minimum distance of the code should be 3 or 4.
2. The codeword must have the shortest length possible, since the longer it is, the more likely that the participant commits more errors.
3. The messages should be easy to encode and the codewords should be easy to decode.
4. The algorithms used for encoding and decoding must be easily implemented in software.

On the one hand, because the key of the previous section consists of digits from 0 to 9, the alphabet we will use is the field $\mathbb{F}_{11} = \{0, 1, 2, \ldots, 10\}$, where the sum and multiplication is done modulo 11. Good encoding and decoding algorithms exist for linear codes and our code will be linear. Specifying a generator or parity check matrix is enough to determine this code. We will use these matrices to encode the original message and decode messages received.

As the key consists of 5-$\mathbb{F}_{11}$ entries, the matrix generating for the code must have rank 5. Thus, two of the parameters of the code are determined: the dimension of the code, $k = 5$, and the minimum distance d, which should be 3 or 4. Thus, we need to determine the length of the code (which should be minimal). If we know two parameters of a code, there are certain bounds that allow us to determine the third, among them, the Singleton Bound.

Following [8], set $A(n, d) = \max\{\mathrm{M} : \text{ an } (\mathrm{n}, \mathrm{M}, \mathrm{d}) \text{ code exists}\}$. The study of the numbers $A(n, d)$ is considered to be central in combinatorial coding theory.

A (resp., linear) code $\mathcal{C}$ such that $|\mathcal{C}| = A(n,d)$ is called *optimal* (resp., *optimal linear*).

Theorem 3. (Singleton Bound, [8]) *For any positive integers q, n, d, $q \geq 2$ we have $A(n,d) \leq q^{n-d+1}$.*

The Singleton bound implies that any $[n,k,d]_q$ linear code must satisfy

$$q^k \leq A(n,d) \leq q^{n-d+1}.$$

It follows that

$$k \leq n - d + 1. \tag{1}$$

Now notice that any $[n,k,d]_q$ linear code is optimal if and only if $k = n-d+1$.

Therefore, according to (1), the length of the code should be $n \geq 7$ or $n \geq 8$. As we want the shortest length possible, we question whether a $[7,5,3]_{11}$ code exists. It is worth mentioning that there are not always codes given the parameters n, k and d (for an example of this fact, see [5]). But this is not our case and we proceed to build a parity check matrix $\mathbf{H}$ as follows: for the first column of $\mathbf{H}$ choose the vector $v_1 = (1,1) \in \mathbb{F}_{11}^2$. For the second column of $\mathbf{H}$ we select any nonzero vector in $\mathbb{F}_{11}^2$ other than a multiple of v_1; we take $v_2 = (1,2)$. The third column of $\mathbf{H}$ would be a nonzero vector in $\mathbb{F}_{11}^2$ that is not a multiple of v_1 and v_2; let $v_3 = (1,3)$. We continue this way until we complete the first five columns of this matrix. For the last two columns we chose the vectors $(1,0)$ and $(0,1)$ to use Proposition 1 and obtain a dual code of dimension two. Thus, the parity check matrix is:

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & 4 & 5 & 0 & 1 \end{bmatrix}.$$

Therefore, a generator matrix $\mathbf{G}$ for a $[7,5]_{11}$ linear code is:

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 10 & 10 \\ 0 & 1 & 0 & 0 & 0 & 10 & 9 \\ 0 & 0 & 1 & 0 & 0 & 10 & 8 \\ 0 & 0 & 0 & 1 & 0 & 10 & 7 \\ 0 & 0 & 0 & 0 & 1 & 10 & 6 \end{bmatrix}.$$

The shape of these arrays will allow us to give very simple rules for encoding and decoding. Denote by $\mathscr{C}$ the $[7,5]_{11}$ code with generator matrix $\mathbf{G}$ and parity check matrix $\mathbf{H}$ given above. Note that the columns of $\mathbf{H}$ are different from zero and that any two of its columns are linearly independent, hence we deduce that $w(\mathscr{C}) \geq 3$. Since the difference of the first two rows of $\mathbf{G}$ is a codeword in $\mathscr{C}$ with Hamming weight 3, we concluded that $w(\mathscr{C}) = 3$. Hence, this is a code with the characteristics we seek, which is quite appropriate to our purposes (see section 5 of this work).

If $(x_1, x_2, x_3, x_4, x_5)$ is the key to any participant as described above, then this would be encoded as

$$(x_1, x_2, x_3, x_4, x_5, 10(x_1 + \cdots + x_5), 10x_1 + 9x_2 + \cdots + 6x_5),$$

where operations are carried out in the field $\mathbb{F}_{11}$. This codeword is the one that each participant receives and will have to write in their answer sheet. As you can see all the information of the student is in it. A program implemented in a computer is responsible for creating the key for each student and also codifying it. Thus, these processes are free from human error. We now address the process of decoding, which is presented below.

4. Decoding $\mathscr{C}$

Suppose that we send $x \in \mathscr{C}$ and that we receive y, which (possibly) was corrupted in at most one coordinate, ie, $w(y - x) \leq 1$. If $w(y - x) = 0$, then $y = x \in \mathscr{C}$ and this happens if and only if $\mathbf{H}y^t = \mathbf{0}^t$. If $w(y - x) = 1$, $y - x = \alpha e_i$, where $\alpha \in \mathbb{F}_{11} \setminus \{0\}$ and e_i has a one in the position i, $1 \leq i \leq 7$, and zeros in the others. Then $y = x + \alpha e_i$ and hence $\mathbf{H}y^t = \alpha h_i$, where h_i is the column i of $\mathbf{H}$. We distinguish between two cases. First, if $i = 6$ or $i = 7$, then $y - x = (0, 0, 0, 0, 0, \alpha, 0)$ or $y - x = (0, 0, 0, 0, 0, 0, \alpha)$ respectively. For $1 \leq i \leq 5$, the h_i column has all its coordinates no zero. The first is 1 and the second is between 1 and 5. Then, the first coordinate of the vector $(\alpha, \beta)^t = \mathbf{H}y^t$ is the no zero coordinate of $y - x$ and $\beta\alpha^{-1}$, where α^{-1} is the inverse of α in $\mathbb{F}_{11}$, tell us the position where the error has occurred when $1 \leq \beta\alpha^{-1} \leq 5$. If $\beta\alpha^{-1} \geq 6$, there must have occurred more than one error. Our decoding scheme can be summarized in the following algorithm:

```
1. Input:The information received $y = (y_1, y_2, y_3, y_4, y_5, y_6, y_7)$
2. Calculate  $\mathbf{H}y^t = (\alpha, \beta)^t$
3. IF $(\alpha, \beta)^t = (0, 0)^t$ or $\alpha = 0$ or $\beta = 0$
      $(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = (y_1, y_2, y_3, y_4, y_5, y_6 - \alpha, y_7 - \beta)$
   ELSE
      IF    $1 \leq \beta\alpha^{-1} \leq 5$, set $i = \beta\alpha^{-1}$
            $e = \alpha e_i = (0, \ldots, \alpha, \ldots, 0)$
            $(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = y - e$
      ELSE
            print "More than one error has occurred"
4. Output:$(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$ or "More than one error has occurred".
```

Example 1. Let $c = (0, 0, 1, 1, 1)$ be a key as described in the section 3. We encode this key as $x = (0, 0, 1, 1, 1, 8, 10)$. Suppose that $y = (0, 1, 1, 1, 1, 8, 10)$ is the key written on the answer sheet rather than x. We can see that an error has occurred but in practice we do not know that this happened. Then, we work with the only thing we know: the vector y. We have that $\mathbf{H}y^t = (1, 2)^t$ and because $(1, 2)^t \neq (0, 0)^t$, we calculate the inverse of 1 in the field $\mathbb{F}_{11}$

(it is 1) and we see that $1 \cdot 2 = 2$. Therefore $e = (0,1,0,0,0,0,0)$. Hence, $y - e = (0,0,1,1,1,8,10)$, which coincides with x since in this case one error has occurred.

Example 2. Let x as in the example above and we suppose that the student recorded his key as the vector $y = (1,0,1,1,1,9,10)$ (A vector in which two errors have occurred.) Therefore $\mathbf{H}y^t = (2,1)^t \neq (0,0)^t$. The inverse of 2 in $\mathbb{F}_{11}$ is 6 and therefore, $6 \cdot 1 = 6 > 5$. Hence the output of the algorithm is "More than one error has occurred".

It is also possible that more than two errors occurred, but in this case the code $\mathscr{C}$ is not always able to detect it and indeed, it may be that the resulting vector be a codeword. This does not contradict the detecting and correcting capability of code $\mathscr{C}$ since the Theorems 1 and 2 establish clearly these capabilities.

5. Some properties of $\mathscr{C}$

The most important property of the code $\mathscr{C}$ is that its parameters satisfy the equality $d = n - k + 1$. In particular, *$\mathscr{C}$ is an optimal linear code.* Usually $[n, k, n-k+1]$ linear codes are called *Maximum Distance Separable* because the codewords are the most distant possible (with respect to the Hamming metric). This favors the decoding of such codes (for details see [7], [8]).

Finding optimal codes is one of the main problems of Coding Theory and many efforts have been made with the aim of identifying such codes in certain fields [2], [4], [6]. In the chapter "*Bounds on linear codes*" of [3] there are Brower's tables which contain bounds on the parameters of various types of codes over the field $\mathbb{F}_q$ for $q \in \{2,3,4,5,7,8,9\}$. More updated versions of these tables can be found in [1]. Thus, the code $\mathscr{C}$ is not in these tables, although can be obtained from the well-known general theory of optimal linear codes.

The code $\mathscr{C}$ can be used to transmit at most 161051 different messages. This amount is more than what we need but this excess allows that $\mathscr{C}$ can be used in other situations with increased demand, where one error is not uncommon but two errors are rare. This will also justify the choice of field $\mathbb{F}_{11}$. If it is required a code with fewer codewords and with the same parameters that $\mathscr{C}$, we suggest the $[7,5,3]_7$ code $\mathscr{C}_1$ with generator matrix and parity check matrix given below,

$$\mathbf{H_1} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & 4 & 5 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{G_1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 6 & 6 \\ 0 & 1 & 0 & 0 & 0 & 6 & 5 \\ 0 & 0 & 1 & 0 & 0 & 6 & 4 \\ 0 & 0 & 0 & 1 & 0 & 6 & 3 \\ 0 & 0 & 0 & 0 & 1 & 6 & 2 \end{bmatrix}.$$

The matrix $\mathbf{H}_1$ was built in the same manner as the matrix $\mathbf{H}$. This is not a coincidence and the shape of these matrices is very general as shows the following

Theorem 4. *Each $[7,5,3]_q$ code $\mathcal{C}$ (up to scalar multiple equivalence) has generator matrix $G = [I_5| - A^t]$, where A is the $\mathbb{F}_q$–matrix*

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_5 \end{bmatrix},$$

with $\alpha_i \neq \alpha_j$,and α_i a unit in $\mathbb{F}_q$ for $1 \leq i, j \leq 5$. Moreover, these codes exist for $q \geq 7$ and none of them is equivalent to a Hamming code.

Proof. Suppose there is an $[7,5,3]_q$ code with parity check matrix H. Applying elementary operations the matrix H can be transformed to

$$H' = \begin{bmatrix} c_1 & c_2 & c_3 & c_4 & c_5 & 1 & 0 \\ d_1 & d_2 & d_3 & d_4 & d_5 & 0 & 1 \end{bmatrix}.$$

By Lemma 2, the code $\mathcal{C}$ is scalar multiple equivalent to the code with parity check matrix H'. As any two columns of H are linearly independent and equivalent codes have the same minimum distance, then the scalars $c_1, \ldots, c_5, d_1, \ldots, d_5$ are nonzero in $\mathbb{F}_q$. Thus, we can transform the matrix H' to the matrix

$$\overline{H} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ d_1c_1^{-1} & d_2c_2^{-1} & d_3c_3^{-1} & d_4c_4^{-1} & d_5c_5^{-1} & 0 & 1 \end{bmatrix}.$$

Again, by Lemma 2, the code with parity check matrix H' is scalar multiple equivalent to the code with parity check matrix parity $\overline{H}$. This implies that $\mathcal{C}$ is scalar multiple equivalent to the code with generator matrix of the form asked in the theorem, where $\alpha_i = d_ic_i^{-1}$, for $1 \leq i \leq 5$. Note that the $\alpha_i's$ are different, because the columns of H are linearly independent. From this observation we can conclude that the finite fields for which these codes exist are those with group of units of order ≥ 5. The smallest number with this property is $q = 7$. Now, remember that a Hamming code of order r on $\mathbb{F}_q$ and length n has parameters $[n = (q^r-1)/(q-1), n-r, 3]_q$. Being $\mathcal{C}$ a $[7,5,3]_q$ code, we must have necessarily that $r = 2$. Then $[(q^2-1)/(q-1) = (q+1), 5, 3]_q = [q+1, 5, 3]_q$. But as we have proved that $[7,5,3]_q$ codes exist for $q \geq 7$, then $q+1 \geq 8$. Therefore, none $[7,5,3]_q$ code can be equivalent to a Hamming code as the length of the blocks do not match. ☑

Theorem 4 is a concrete instance of the following theorem [7, Theorem 5.3.2].

Theorem 5. *An $[n,k]_q$ code with parity check matrix H is optimal if and only if $n-k$ columns of H are linearly independent.*

On the other hand, we must emphasize that in [1] is provided a code which have the same parameters of $\mathscr{C}_1$ and is constructed in the following way:

1. Construct the $[57,54,3]_7$ Hamming code ($r = 3$);
2. Shorten the $[57,54,3]_7$ Hamming code in the coordinates $51, \ldots, 57$. We obtain a $[50,47,3]$ code;
3. Delete from the $[50,47,3]_7$ code the (at most) 42 coordinates of a word in the dual;

4. Shortening the last $[8, 6, 3]$ code in the eighth coordinate code gives the parity check matrix of the desired code.

It should be noted that our construction is simpler than that suggested in [1]. In addition, the construction indicated in [1], in step 3 does not specify which coordinates must be eliminated. Using the *Magma*© computer algebra system, removing the 42 *last* coordinates of step 3, we find the following parity check matrix,

$$\widetilde{H} = \begin{bmatrix} 1 & 0 & 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Obviously, we obtain the matrix $\mathbf{H_1}$ permuting the columns of $\widetilde{H}$. Hence, the code suggested in [1] is scalar multiple equivalent to the code $\mathscr{C}_1$ by Lemma 1.

In general, the weight distribution of an optimal linear code is known [7]. Using this result and MacWilliams Identities for linear codes, we calculate the weight distribution w_i and $w_i^{\perp}$ of $\mathscr{C}$ and $\mathscr{C}^{\perp}$, respectively.

i	w_i	$w_i^{\perp}$
0	1	1
3	350	-
4	2800	-
5	17430	-
6	57820	70
7	82650	20

A linear code is a *projective code* if $w(\mathcal{C}^{\perp}) \geq 3$. Therefore, $\mathscr{C}$ is a projective code.

Finally, note that the algorithms for encoding and decoding data given are still valid, with slight modifications, for any code $\mathcal{C}$ described above in Theorem 4.

Acknowledgement. We thank Horacio Tapia–Recillas and J. Noé Gutiérrez–Herrera, both from the Department of Mathematics of the UAM–I, D. F., México, and William Y. Vélez, from the Department of Mathematics of The University of Arizona, USA, for helpful discussions and comments.

References

[1] Grassi, Markus. "Bounds on the minimum distance of linear codes" Online available at http://www.code.de. Accessed on 2009–09–11.

[2] Greenough, P.; Hill, R. *Optimal linear codes over GF(4).* Discrete Mathematics. Volume 125. pp. 187–199. 1994.

[3] Huffman, W. C. ; Pless, V. (Editors). *Handbook of Coding Theory.* Volume I. Elseiver. 1998.

[4] Landjev, I. *Optimal Linear Codes of Dimension 4 over GF(5).* Lecture Notes In Computer Science. Vol. 1255. *Proceedings of the 12th International Symposium on Applied Algebra, Algebraic Algorithms and Error-Correcting Codes.* 212–220, 1997.

[5] Landjev, I. *The Nonexistence of Some Optimal Ternary Codes of Dimension Five.* Designs, Codes and Cryptography **15** (1998), 245–258.

[6] Landjev, I.; Rousseva, A.; Maruta, T.; Hill, R. *On Optimal Codes Over the Field with Five Elements.* Designs, Codes and Cryptography, **29** (1–3) (2003), 165–175.

[7] Roman, S. *Coding and Information Theory.* First Edition. Springer–Verlag: New York, 1992.

[8] Van Lint, J. H. *Introduction to Coding Theory.* Second Edition. Springer–Verlag: Berlin–Heidelberg, 1992.

(Recibido en julio de 2009. Aceptado para publicación en octubre de 2009)

Henry Chimal–Dzul & Javier Díaz-Vargas
Facultad de Matemáticas, Universidad Autónoma de Yucatán
97110, Mérida, Yucatán, México
e-mail: henrychimal@gmail.com, jdvargas@uady.mx