IJMMS 29:5 (2002) 265–270
PII. S0161171202011444
http://ijmms.hindawi.com

A DECODING SCHEME FOR THE 4-ARY LEXICODES WITH $D = 3$

D. G. KIM and H. K. KIM

Received 21 January 2001 and in revised form 8 July 2001

We introduce the algorithms for basis and decoding of quaternary lexicographic codes with minimum distance $d = 3$ for an arbitrary length n.

2000 Mathematics Subject Classification: 94Bxx.

1. Introduction. In this section, we define some particular operations and discuss q-ary lexicographic codes with minimum distance d. The game-theoretic operations of nim-addition $\oplus$ and nim-multiplication $\otimes$ which are used in the Game of Nim are introduced by Definitions 1.1 and 1.2.

The Game of Nim is played by two players, with one or more piles of counters. Each player, in turn, removes from one to all counters of a pile. The player taking the last counter wins.

Definition 1.1. Let $(\alpha_1 \cdots \alpha_r)$, $(\beta_1 \cdots \beta_r)$ be the binary representation of α, β, respectively. For each i, $\alpha \oplus \beta$ has a 0 digit in the position i where $\alpha_i = \beta_i$, and $\alpha \oplus \beta$ has a 1 in the position i where $\alpha_i \neq \beta_i$. In other words, $\alpha \oplus \beta$ is the Exclusive OR (XOR) of each digit in their binary representations.

For example, the nim-addition table for numbers less than 4 is given in Table 1.1.

Table 1.1

$\oplus$	0	1	2	3
0	0	1	2	3
1	1	0	3	2
2	2	3	0	1
3	3	2	1	0

There is a nim-multiplication $\otimes$ which, together with nim-addition $\oplus$, converts the integers into a field [1]. With nim-multiplication, we know that $0 \otimes \alpha$ must be 0 which is the zero of the field. Also $1 \otimes \alpha$ must be α. Since the elements other than 0, 1 satisfy $\alpha \otimes \alpha = \alpha \oplus 1$ in the finite field of order 4, we have $2 \otimes 2 = 3$. Next $2 \otimes 3$ cannot be one of $0, 2, 3$ and so must be 1.

In general, using the above value α we can define the following nim-multiplication.

Definition 1.2. The nim-multiplication $\alpha \otimes \beta$ is defined by $\alpha \otimes \beta = \text{mex}\{(\alpha' \otimes \beta) \oplus (\alpha \otimes \beta') \oplus (\alpha' \otimes \beta') \mid \alpha' < \alpha,\ \beta' < \beta\}$, where mex (minimal excluded number) means the least nonnegative integer not included.

For example, the nim-multiplication table for numbers less than 4 is given in Table 1.2.

TABLE 1.2

$\otimes$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	3	1
3	0	3	1	2

The following is an easy rule enabling us to compute nim-additions:

(1) the nim-sum of a number of distinct 2-powers ("2-power" means a power of 2 in the ordinary sense) is their ordinary sum;

(2) the nim-sum of two equal numbers is 0.

For finite numbers, the nim-multiplication follows from the following rules, analogous to those for nim-addition. We will use the term *Fermat* 2*-power* to denote the numbers 2^{2^a} in the ordinary sense;

(3) the nim-product of a number of distinct Fermat 2-powers is their ordinary product;

(4) the square of a Fermat 2-power is the number obtained by multiplying it by $3/2$ in the ordinary sense.

In [1], $\oplus$ and $\otimes$ convert the numbers $0,1,2,\ldots$ into a field of characteristic 2. Also, for all a, the numbers less than 2^{2^a} form a subfield isomorphic to the Galois field $\mathrm{GF}(2^{2^a})$.

Consider the lexicographic codes (for short, lexicodes) with base $B = 2^{2^a}$. A word of this code is a sequence $\mathbf{x} = \cdots x_3x_2x_1$ of elements of $\{0,1,\ldots,2^{2^a}-1\}$. The set of words is ordered lexicographically, that is, the word $\mathbf{x} = \cdots x_3x_2x_1$ is smaller than $\mathbf{y} = \cdots y_3y_2y_1$, written $\mathbf{x} < \mathbf{y}$, in case of some r we have $x_r < y_r$ and $x_s = y_s$ for all s greater than r.

Lexicodes are defined by saying a word in the code in case it does not conflict with any previous codewords. That is, the lexicode with minimum distance d is defined by saying that two words do not conflict in case the Hamming distance between them is not less than d. We write $\mathcal{S}_{n,d}$ for the 4-ary lexicode consisting of the codewords with length n or less and minimum distance d.

In [2], Conway and Sloane showed that lexicodes with base $B = 2^a$ are closed under nim-addition, and if $B = 2^{2^a}$ the lexicodes are closed under nim-multiplication by scalars. Therefore if B is of the form 2^{2^a}, then the lexicode is a linear code over $\mathrm{GF}(B)$.

2. The basis and decoding for $\mathcal{S}_{4,3}$

LEMMA 2.1. *Let* $\mathbf{e}_n$ *be the basis of length* n *in* $\mathcal{S}_{4,3}$*. Then* $111 = \mathbf{e}_3$, $1012 = \mathbf{e}_4$*, and* $10013 = \mathbf{e}_5$.

PROOF. Since the weight of $\mathbf{e}_n$ must be greater than or equal to 3, the first basis has at least 3 nonzero digits, and so the smallest codeword is 111. The second basis $\mathbf{e}_4$ is the type of $10ab$, where neither a nor b is zero. Let "ab"$_n$ be the first two digits of $\mathbf{e}_n$. Since "ab"$_3$ = "11", "ab"$_4$ is lexicographically ordered "12", and then $d(\alpha \otimes \mathbf{e}_3, 1012) \geq 3$, for $\alpha \in \mathrm{GF}(4)$. Therefore, $1012 = \mathbf{e}_4$. In a similar way, we obtain $10013 = \mathbf{e}_5$. □

Theorem 2.2. *There is no basis $\mathbf{e}_n$, where $n = 6, 17s+5$ ($s \in \mathbb{N}$) in $\mathcal{S}_{4,3}$.*

Proof. Suppose that $1000ab \in \mathcal{S}_{4,3}$. Let $\alpha \in \mathrm{GF}(4)$. If neither a nor b is zero, there exists $\mathbf{e}_i$ ($3 \leq i \leq 5$) such that $d(1000ab, \alpha\mathbf{e}_i) < 3$. This contradicts the hypothesis. In all other cases, the weight of $1000ab$ is 2, and so the basis of length 6 does not exist.

Consider the basis $\mathbf{e}_7$ of length 7. Then $10000ab$ of length 7 also conflicts with any smaller basis, for all "ab". Thus $10000ab$ needs a digit 1 in the 6th position. If "ab" = "$0b$" ($b \neq 0$), then $110000b$ does not conflict with any smaller codeword. Hence 1100001 is the smallest codeword with more digits than $\mathbf{e}_5$, that is, $1100001 = \mathbf{e}_7$. Therefore, for $7 \leq n \leq 21$, "ab"$_n$ takes ordered digit from "01" to "33".

Suppose that there exists a basis of length 22, that is, $10\cdots01000ab \in \mathcal{S}_{4,3}$. Since there exists $\mathbf{e}_i$ ($7 \leq i \leq 21$) such that $d(10\cdots01000ab, \mathbf{e}_i) < 3$ for any "ab", this is a contradiction to the hypothesis. So the basis of length 22 does not exist.

We consider the basis of length 23, that is, $110\cdots01000ab = \mathbf{e}_{23}$. Although "$ab$"$_{23}$ = $\alpha\otimes$ "ab"$_i$ for any α, $i \leq 22$, we have $\mathrm{wt}(110\cdots01000ab \oplus (\alpha \otimes \mathbf{e}_i)) \geq 3$. Hence, $110\cdots01\ 00000$ is the smallest codeword with more digits than $\mathbf{e}_{21}$, that is, $110\cdots 0100000 = \mathbf{e}_{23}$. Therefore, for $23 \leq n \leq 38$, "ab"$_n$ takes ordered digit from "00" to "33". As a result, neither $\mathbf{e}_6$ nor $\mathbf{e}_{17s+5}$ ($s \in \mathbb{N}$) exists in $\mathcal{S}_{4,3}$. □

As we have seen in the proof of Theorem 2.2, the basis $\mathbf{e}_n$ has digit 1's in the nth, 6th, and $(17s+5)$th positions, for all $s \in \mathbb{N}$ satisfying $6 < 17s+5 < n$.

The following tables give "ab"$_n$ corresponding to the length n, where $7 \leq n \leq 21$ or $17p+6 \leq n \leq 17q+4$, for $p \in \mathbb{N}$ and $q = p+1$.

Table 2.1

ab	**00**	**01**	**02**	**03**	**10**	**11**	**12**	**13**
n		7	8	9	10	11	12	13
ab	**20**	**21**	**22**	**23**	**30**	**31**	**32**	**33**
n	14	15	16	17	18	19	20	21

Table 2.2

ab	**00**	**01**	**02**	**03**	**10**	**11**	**12**	**13**
n	23	24	25	26	27	28	29	30
n	40	41	42	43	44	45	46	47
ab	**20**	**21**	**22**	**23**	**30**	**31**	**32**	**33**
n	31	32	33	34	35	36	37	38
n	48	49	50	51	52	53	54	55

Now we may consider the basis $\mathbf{e}_n$ satisfying $n \geq 7$ and $n \neq 17s+5$, $s \in \mathbb{N}$, in the following algorithm.

Algorithm for the basis $\mathbf{e}_n$

Step 1. Suppose that $7 \leq n \leq 21$. The basis $\mathbf{e}_n$ has digit 1's in the nth and 6th positions. And "ab"$_n$ takes the $(n-6)$th lexicographically ordered digit from "01" to "33" (see Table 2.1).

Step 2. Suppose that $17p+6 \leq n \leq 17q+4$, for $p \in \mathbb{N}$ and $q = p+1$. Then $\mathbf{e}_n$ has digit 1's in the nth, 6th and $(17s+5)$th positions, for all $s \in \mathbb{N}$ satisfying $6 < 17s+5 < n$. And "ab"$_n$ takes the $(n-17p-5)$th lexicographically ordered digit from "00" to "33" (see Table 2.2).

The following table gives the basis $\mathbf{e}_n$, where $n \geq 7$, $n \neq 17s+5$, for $s \in \mathbb{N}$:

						1	1	0	0	0	0	1	$= \mathbf{e}_7$
					1	0	1	0	0	0	0	2	$= \mathbf{e}_8$
				1	0	0	1	0	0	0	0	3	$= \mathbf{e}_9$
			1	0	0	0	1	0	0	0	1	0	$= \mathbf{e}_{10}$
		1	0	0	0	0	1	0	0	0	1	1	$= \mathbf{e}_{11}$
													$\vdots$
				1	1	$\cdots$	1	0	0	0	0	0	$= \mathbf{e}_{23}$
			1	0	1	$\cdots$	1	0	0	0	0	1	$= \mathbf{e}_{24}$
		1	0	0	1	$\cdots$	1	0	0	0	0	2	$= \mathbf{e}_{25}$
	1	0	0	0	1	$\cdots$	1	0	0	0	0	3	$= \mathbf{e}_{26}$
1	0	0	0	0	1	$\cdots$	1	0	0	0	1	0	$= \mathbf{e}_{27}$
													$\vdots$

Example 2.3. We take $n = 19$ as the length. Since $7 \leq n \leq 21$,

$$100000000\ 00001000ab = \mathbf{e}_{19}, \tag{2.1}$$

by Step 1. Then "ab"$_{19}$ takes the 13th order "31" from "01". Therefore, we have $100000000\ 0000100031 = \mathbf{e}_{19}$.

Example 2.4. Let $n = 27$. Since $6 < 17s+5 < n$ for $s = 1$, $\mathbf{e}_{27}$ has digit 1's in the 27th, 22nd, and 6th positions, by Step 2. So we have

$$1000010\ 0000000000\ 00001000ab = \mathbf{e}_{27}. \tag{2.2}$$

Since $17p+6 \leq n \leq 17q+4$ for $p = 1$ and $q = 2$, "ab"$_{27}$ takes the 5th order "10" from "00". Therefore, $1000010\ 0000000000\ 0000100010 = \mathbf{e}_{27}$.

Example 2.5. Let $n = 62$. Since $6 < 17s+5 < n$ for $s = 1,2,3$, $\mathbf{e}_{62}$ has digit 1's in the 62nd, 56th, 39th, 22nd, and 6th positions, by Step 2. So we have $10\ 0000100000\ 0000000000\ 0100000000\ 0000000010\ 0000000000\ 00001000ab = \mathbf{e}_{62}$. Since $17p+6 \leq n \leq 17q+4$ for $p = 3$ and $q = 4$, "ab"$_{62}$ takes the 6th order "11" from "00". Therefore, we have $10\ 0000100000\ 0000000000\ 0100000000\ 0000000010\ 0000000000\ 0000100011 = \mathbf{e}_{62}$.

Below we discuss a decoding algorithm for $\mathcal{S}_{4,3}$.

Definition 2.6. For a given received vector $\mathbf{r} = a_n a_{n-1} \cdots a_2 a_1$, $a_i \in \mathrm{GF}(4)$, the *testing vector*, denoted by $\mathbf{t}$, in $\mathcal{S}_{4,3}$ is defined by $\mathbf{t} = (a_n \otimes \mathbf{e}_n) \oplus \cdots \oplus (a_3 \otimes \mathbf{e}_3)$, where $n \neq 6,\ 17s+5$, for $s \in \mathbb{N}$.

In the following remark we explain a decoding algorithm of $\mathcal{S}_{4,3}$ in more detail.

Remark 2.7. For a given received vector $\mathbf{r} = a_n a_{n-1} \cdots a_2 a_1$, we obtain the testing vector $\mathbf{t} = b_n b_{n-1} \cdots b_2 b_1$, by Definition 2.6. Let $s \in \mathbb{N}$ and $\alpha \in \mathrm{GF}(4)$, and let "$f_2 f_1$"$_i$ be the first two digits of $\mathbf{e}_i$ in $\mathbf{t}$.

(A) Certainly, the codeword $\mathbf{c}$ is a linear combination of some bases by scalar nim-multiplication. From the given received vector, we can guess the bases which may generate the codeword.

If $d(\mathbf{t},\mathbf{r}) = 1$, we have the following two cases. First, one of a_1, a_2 is not correct. In the second case, one of the 6th, $(17s+5)$th digit is not correct. In all cases, $\mathbf{t}$ is obtained by bases which do not depend on errored digit. Therefore, we have the desired codeword $\mathbf{c} = \mathbf{t}$.

(B) Suppose that $d(\mathbf{t},\mathbf{r}) > 1$. This means that both a_1 and a_2 are correct. Hence, we have to find "$d_2 d_1$" ($d_1, d_2 \in \mathrm{GF}(4)$) such that "$b_2 b_1$" $\oplus$ "$d_2 d_1$"= "$a_2 a_1$" because $\mathbf{t}$ is more added by a component vector $(a_p \otimes \mathbf{e}_p)$ with "$d_2 d_1$" of $\mathbf{t}$. Therefore, if such a vector exists, we have the desired codeword $\mathbf{c} = \mathbf{t} \oplus (a_p \otimes \mathbf{e}_p)$.

(C) Suppose that $d(\mathbf{t},\mathbf{r}) > 1$. If there is not any component vector $(a_p \otimes \mathbf{e}_p)$ with "$d_2 d_1$" in $\mathbf{t}$, then one of the nonzero digits in $\mathbf{r}$ is not correct, let a_q, for $q \neq 1, 2, 6, 22, \ldots$. Such a digit is obtained from the equation $\alpha \otimes (a_q \otimes$ "$f_2 f_1$"$_q)$ = "$d_2 d_1$". Next, if we obtain a digit a'_q $(\neq a_q)$ satisfying $(a_n \otimes$ "$f_2 f_1$"$_n) \oplus \cdots \oplus (a'_q \otimes$ "$f_2 f_1$"$_q) \oplus \cdots \oplus (a_3 \otimes$ "$f_2 f_1$"$_3)$ = "$a_2 a_1$", then the desired codeword $\mathbf{c}$ is $(a_n \otimes \mathbf{e}_n) \oplus \cdots \oplus (a'_q \otimes \mathbf{e}_q) \oplus \cdots \oplus (a_3 \otimes \mathbf{e}_3)$.

(D) Suppose that $d(\mathbf{t},\mathbf{r}) > 1$ and there is no component vector $(a_p \otimes \mathbf{e}_p)$ with "$d_2 d_1$" in $\mathbf{t}$. For all α, a_q such that $q \neq 6,\ 17s+5$, if it does not satisfy the equation $\alpha \otimes (a_q \otimes$ "$f_2 f_1$"$_q)$ = "$d_2 d_1$", then $\mathbf{r}$ has a nonzero leading digit in the 6th or $(17s+5)$th position. If $\mathbf{r}$ has a nonzero leading digit in the 6th position, then we have the desired codeword $\mathbf{c} = \mathbf{t} \oplus (a_k \otimes \mathbf{e}_k)$, for some a_k $(7 \le k \le 21)$. If $\mathbf{r}$ has a nonzero leading digit in the $(17s+5)$th position, then we have the desired codeword $\mathbf{c} = \mathbf{t} \oplus (a_k \otimes \mathbf{e}_k)$, for some a_k $(17s+6 \le k \le 17s+21)$. In fact, we can obtain a_k satisfying $(a_k \otimes$ "$f_2 f_1$"$_k)$ = "$d_2 d_1$".

Decoding algorithm of $\mathcal{S}_{4,3}$

Step 1. Suppose that $d(\mathbf{t},\mathbf{r}) = 1$. Then $\mathbf{c} = \mathbf{t}$.

Step 2. Suppose that $d(\mathbf{t},\mathbf{r}) > 1$ and there is $(a_p \otimes \mathbf{e}_p)$ with "$d_2 d_1$" in $\mathbf{t}$. Then $\mathbf{c} = \mathbf{t} \oplus (a_p \otimes \mathbf{e}_p)$.

Step 3. Suppose that $d(\mathbf{t},\mathbf{r}) > 1$ and there is no $(a_p \otimes \mathbf{e}_p)$ with "$d_2 d_1$" in $\mathbf{t}$. If there exist α, q such that $\alpha \otimes (a_q \otimes$ "$f_2 f_1$"$_q)$ = "$d_2 d_1$", then $\mathbf{c} = \mathbf{t} \oplus ((a_q \oplus a'_q) \otimes \mathbf{e}_q)$, where a'_q $(\neq a_q)$ satisfies $(a_q \oplus a'_q) \otimes$ "$f_2 f_1$"$_q$ = "$a_2 a_1$" $\bigoplus_{i=3}^{n} (a_i \otimes$ "$f_2 f_1$"$_i)$.

(Note that $\bigoplus_{i=3}^{n} (a_i \otimes$ "$f_2 f_1$"$_i)$ is the first two digits of $\mathbf{t}$.)

Step 4. Suppose that $d(\mathbf{t},\mathbf{r}) > 1$ and there is no $(a_p \otimes \mathbf{e}_p)$ with “d_2d_1” in $\mathbf{t}$. If there is no q such that $\alpha \otimes (a_q \otimes \text{“}f_2f_1\text{”}_q) = \text{“}d_2d_1\text{”}$ for all α, then $\mathbf{c} = \mathbf{t} \oplus (a_k \otimes \mathbf{e}_k)$, where a_k satisfies $(a_k \otimes \text{“}f_2f_1\text{”}_k) = \text{“}d_2d_1\text{”}$ for $7 \le k \le 21$ or $17s+6 \le k \le 17s+21$.

Example 2.8. Let $\mathbf{r} = 3001202011$. Then $\mathbf{t}$ is $(3 \otimes \mathbf{e}_{10}) \oplus (1 \otimes \mathbf{e}_7) \oplus (2 \otimes \mathbf{e}_4) = 3001202012$. Since $d(\mathbf{r},\mathbf{t}) = 1$, therefore, $\mathbf{c} = \mathbf{t}$.

Example 2.9. Let $\mathbf{r} = 3011202012$. Then $\mathbf{t}$ is $(3 \otimes \mathbf{e}_{10}) \oplus (1 \otimes \mathbf{e}_8) \oplus (1 \otimes \mathbf{e}_7) \oplus (2 \otimes \mathbf{e}_4) = 3011302010$, and “$d_2d_1$”=“02”. Since $d(\mathbf{r},\mathbf{t}) > 1$ and there is $(1 \otimes \mathbf{e}_8)$ with “02” in $\mathbf{t}$, therefore, $\mathbf{c} = \mathbf{t} \oplus (1 \otimes \mathbf{e}_8) = 3001202012$.

Example 2.10. Let $\mathbf{r} = 3002202012$. We have $\mathbf{t} = (3 \otimes \mathbf{e}_{10}) \oplus (2 \otimes \mathbf{e}_7) \oplus (2 \otimes \mathbf{e}_4) = 3002102011$, and “$d_2d_1$” = “03”. Then $d(\mathbf{r},\mathbf{t}) > 1$ and there is no $(a_p \otimes \mathbf{e}_p)$ with “03” in $\mathbf{t}$. Since there are $\alpha = 2$, $a_7 = 2$ satisfying $\alpha \otimes (a_7 \otimes \text{“}f_2f_1\text{”}_7) = \text{“03”}$, a_7 is not correct. We obtain a_7' $(= 1)$ satisfying $(2 \oplus a_7') \otimes \text{“01”}_7 = \text{“12”} \oplus \text{“11”}$, by Step 3. Therefore, $\mathbf{c} = \mathbf{t} \oplus ((2 \oplus 1) \otimes \mathbf{e}_7) = 3001202012$.

Example 2.11. Let $\mathbf{r} = 1202012$. We have $\mathbf{t} = (1 \otimes \mathbf{e}_7) \oplus (2 \otimes \mathbf{e}_4) = 1102022$, and “$d_2d_1$”=“30”. Then $d(\mathbf{t},\mathbf{r}) > 1$ and there is no $(a_p \otimes \mathbf{e}_p)$ with “30” in $\mathbf{t}$. Also, there is no q such that $\alpha \otimes (a_q \otimes \text{“}f_2f_1\text{”}_q) = \text{“30”}$ for all α. By Step 4, we have to obtain a_k $(7 \le k \le 21)$ because a_6 is nonzero. Since $(3 \otimes \text{“10”}_{10}) = \text{“30”}$, we obtain a_{10} $(= 3)$. Therefore, $\mathbf{c} = \mathbf{t} \oplus (3 \otimes \mathbf{e}_{10}) = 3001202012$.

Acknowledgements. The first author was supported by the KRF. The second author was supported by the Combinatorial and Computational Mathematics Center (Com^2 MaC-KOSEF).

References

[1] J. H. Conway, *On Numbers and Games*, London Mathematical Society Monographs, no. 6, Academic Press, London, 1976.
[2] J. H. Conway and N. J. A. Sloane, *Lexicographic codes: error-correcting codes from game theory*, IEEE Trans. Inform. Theory **32** (1986), no. 3, 337–348.

D. G. Kim: Department of Internet and Computer, Chungwoon University, Hongsung, Chungnam, 350-701, South Korea
E-mail address: codekim@chungwoon.ac.kr

H. K. Kim: Department of Mathematics, Pohang University of Science and Technology, Pohang 790-784, South Korea
E-mail address: hkkim@euclid.postech.ac.kr