Short dominating paths and cycles in the binary hypercube

Uri Blass
Department of Electrical Engineering – Systems
Tel-Aviv University
Ramat-Aviv 69978, Israel
e-mail: blass_n@netvision.net.il

Iiro Honkala
Department of Mathematics
University of Turku
20014 Turku, Finland
e-mail: honkala@utu.fi

Mark G. Karpovsky
College of Engineering
Boston University
Boston, MA 02215 USA
e-mail: markkar@bu.edu

Simon Litsyn
Department of Electrical Engineering – Systems
Tel-Aviv University
Ramat-Aviv 69978, Israel
e-mail: litsyn@eng.tau.ac.il

February 1, 2001

Keywords: Domination, path, cycle, binary hypercube, Hamming distance

AMS classification: 05C69 (94B75, 94B65, 68R10, 05C70, 05C38)

Running head: Dominating paths and cycles

Abstract

A sequence of binary words of length n is called a cube dominating path, if the Hamming distance between two consecutive words is always one, and every binary word of length n is within Hamming distance one from at least one these words. If also the first and last words are Hamming distance one apart, the sequence is called a cube dominating cycle. Bounds on the cardinality of such sequences are given, and it is shown that asymptotically the shortest cube dominating path and cycle consist of $2^n(1+o(1))/n$ words.

1 Introduction

Denote by F the binary alphabet, and by F^n the space of binary vectors of length n endowed with the Hamming metric $d(\cdot,\cdot)$, i.e., the binary hypercube. The *covering radius* of a code $C \subseteq F^n$ is defined to be the smallest integer R such that the distance from every point to the nearest codeword does not exceed R. The covering radius of codes has been widely studied; see e.g., the book [4]. Let $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_M$ be a sequence of codewords such that the Hamming distance between any two consecutive codewords equals one. If moreover, the set $\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_M\}$ — which may consist of fewer than M codewords — has covering radius at most one, we call C a *cube dominating path* (CDP). Such dominating paths have been considered in general graphs in, e.g., [8], [10] and [18].

CDPs can be used for testing and diagnosis of multiprocessors with interconnections defined by F^n (see, e.g., [15], [9], [16]). In this case every node (processor) is labelled by an n-bit binary vector, and two nodes are connected by a bidirectional line if and only if the Hamming distance between the corresponding labels is equal to one. If $\mathbf{c}_1, \ldots, \mathbf{c}_M$ is a CDP in F^n, then one can use the following procedure for testing and diagnosis. First, node $\mathbf{c}_1$ tests itself and, if it passes the test, it tests the neighbours. After this the software ("dynamic agent") for test generation is moved to $\mathbf{c}_2$ and so on. For this approach testing time is determined by the length M of the selected CDP. Similar procedures for testing and diagnosis of multiprocessors can be found in [2].

In Section 3, we consider the situation when also $d(\mathbf{c}_M, \mathbf{c}_1) = 1$, in which case the sequence of codewords is called a *cube dominating cycle* (CDC). In general graphs such cycles have been studied, e.g., in [1], [3], [6] and [14].

Our codes are related to *circuit codes*; for circuit codes, see, e.g., [7] and [13]. A sequence $\mathbf{c}_1, \ldots, \mathbf{c}_M$ of binary vectors of length n such that $d(\mathbf{c}_1, \mathbf{c}_2) = d(\mathbf{c}_2, \mathbf{c}_3) = \ldots = d(\mathbf{c}_M, \mathbf{c}_1) = 1$ is called a *length n, spread k circuit code* ($1 \le k \le n$), if for all i and j, $d(\mathbf{c}_i, \mathbf{c}_j) = s < k$ implies that $i - j \equiv \pm s \pmod{M}$.

2 Short dominating paths

For a vector $\mathbf{c} \in F^n$ denote by $B(\mathbf{c})$ the ball of radius one centered at $\mathbf{c}$, i.e., $B(\mathbf{c}) = \{\mathbf{x} : d(\mathbf{x}, \mathbf{c}) \le 1\}$. We will say that a vector $\mathbf{c}$ covers a vector $\mathbf{x}$ if $\mathbf{x} \in B(\mathbf{c})$.

Given n, our goal is to minimize the length M of a CDP. Let $S(n)$ stand for such minimal length.

Lemma 1 *Let $\mathbf{c}_1, \ldots, \mathbf{c}_M$ be a CDP, then*
a) $|B(\mathbf{c}_1)| = n + 1$;
b) $|B(\mathbf{c}_2) \setminus B(\mathbf{c}_1)| = n - 1$;
c) For $3 \le k \le M$, $|B(\mathbf{c}_k) \setminus \cup_{i=1}^{k-1} B(\mathbf{c}_i)| \le n - 2$.

Proof The statements a) and b) are immediate. Let $\mathbf{e}_i$ stand for the vector from F^n having 1 in the i-th coordinate and 0's elsewhere. If $\mathbf{c}_k = \mathbf{c}_{k-2}$ then $|B(\mathbf{c}_k) \setminus \cup_{i=1}^{k-1} B(\mathbf{c}_i)| \le |B(\mathbf{c}_k) \setminus B(\mathbf{c}_{k-2})| = 0$, i.e., c) is valid. So, we may assume without loss of generality that $\mathbf{c}_{k-2} = \mathbf{0}, \mathbf{c}_{k-1} = \mathbf{e}_1, \mathbf{c}_k = \mathbf{e}_1 + \mathbf{e}_2$. Then

$$|B(\mathbf{c}_k) \setminus \cup_{i=1}^{k-1} B(\mathbf{c}_i)| \le |B(\mathbf{c}_k) \setminus \cup_{i=k-2}^{k-1} B(\mathbf{c}_i)| = |B(\mathbf{e}_1 + \mathbf{e}_2) \setminus (B(\mathbf{0}) \cup B(\mathbf{e}_1))|$$

$$= (n+1) - |\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2\}| = n - 2.$$

$\square$

Theorem 1 *For $n \ge 3$,*

$$T(n) \ge \frac{2^n - 4}{n - 2}.$$

Proof Let $\mathbf{c}_1, \ldots, \mathbf{c}_M$ be a CDP with $M = T(n)$. By definition, $| \cup_{i=1}^M B(\mathbf{c}_i)| = 2^n$. Then by the previous lemma

$$2^n = | \cup_{i=1}^M B(\mathbf{c}_i)| \le (n+1) + (n-1) + (M-2)(n-2),$$

yielding

$$M - 2 \geq \frac{2^n - 2n}{n-2}.$$

□

A CDP attaining the previous bound is called *perfect*. In other words, a CDP $\mathbf{c}_1, \ldots, \mathbf{c}_M$ is perfect if and only if for every k, $3 \leq k \leq M$,

$$|B(\mathbf{c}_k) \setminus \cup_{i=1}^{k-1} B(\mathbf{c}_i)| = n - 2.$$

Lemma 2

$$T(n+1) \leq 2T(n).$$

Proof Let $\mathbf{c}_1, \ldots, \mathbf{c}_M$ be a CDP with $M = T(n)$. The following CDP can be constructed in dimension $n+1$:

$$(\mathbf{c}_1, 0), \ldots, (\mathbf{c}_M, 0), (\mathbf{c}_M, 1), \ldots, (\mathbf{c}_1, 1).$$

□

Theorem 2

$$T(1) = 1, T(2) = 2, T(3) = 4, T(4) = 6, T(5) = 10, T(6) = 16, T(7) = 28.$$

Proof This is proved using the bound from Theorem 1, computer search and the following CDPs (here we mention only the indices of the coordinates to be changed):
$n = 2$: 1
$n = 3$: 123
$n = 4$: 12341
$n = 5$: 121345121
$n = 6$: 123415361234153
$n = 7$: 121314256476165242312561764.

□

For $n = 3, 4$ these CDPs are perfect; for $n = 5, 6, 7$ they are not. For other optimal CDPs of lengths 3 and 7, see the proof of Theorem 5.

3 Short dominating cycles

Consider now the cyclic case, in which $c_M = c_1$. Given n, our goal is again to minimize M. Let $S(n)$ stand for the minimum.

Theorem 3 *For $n \geq 4$,*

$$S(n) \geq \frac{2^n}{n-2}.$$

Proof Let $\mathbf{c}_1, \ldots, \mathbf{c}_M$ be a CDC with $M = S(n)$ codewords. Because $n \geq 4$, $S(n) \geq K(4,1) = 4$. Here $K(n,R)$ denotes the smallest cardinality of any binary code of length n and covering radius R. Consider the codeword indices $1, 2, \ldots, M$ cyclically, so that $M+1$ is interpreted as 1. Then for each i, the codeword $\mathbf{c}_i$ covers at most $n-2$ such vectors that are not covered by $\mathbf{c}_{i-1}$ and $\mathbf{c}_{i-2}$. Indeed, without loss of generality $\mathbf{c}_{i-2} = 000\ldots0$ and $\mathbf{c}_{i-1} = 100\ldots0$. If $\mathbf{c}_i = \mathbf{c}_{i-2}$, there is nothing to prove. So, assume without loss of generality that $\mathbf{c}_i = 110\ldots0$. Of the $n+1$ vectors that $\mathbf{c}_i$ covers, the vectors $\mathbf{c}_{i-1}$, $\mathbf{c}_{i-2}$ and $010\ldots0$ are already covered by $\mathbf{c}_{i-1}$ and $\mathbf{c}_{i-2}$. We claim that the union of the sets $A_i := B(\mathbf{c}_i) \setminus (B(\mathbf{c}_{i-1}) \cup B(\mathbf{c}_{i-2}))$ is the whole space F^n. Then we obtain the inequality $2^n \leq (n-2)S(n)$ and the theorem follows.

It remains to show that every vector $\mathbf{x} \in F^n$ is in at least one of the sets A_i. Because $n \geq 4$, the distance between $\mathbf{x}$ and its complement $\overline{\mathbf{x}}$ is at least four. Let $\mathbf{c}_j$ be any codeword in C that covers $\overline{\mathbf{x}}$. By the triangle inequality, $d(\mathbf{x}, \mathbf{c}_j) \geq 3$, $d(\mathbf{x}, \mathbf{c}_{j-1}) \geq 2$ and $d(\mathbf{x}, \mathbf{c}_{j+1}) \geq 2$. Consequently, none of the codewords $\mathbf{c}_{j-1}$, $\mathbf{c}_j$ and $\mathbf{c}_{j+1}$ cover $\mathbf{x}$. Let now i be the smallest index exceeding j such that $\mathbf{c}_i$ covers $\mathbf{x}$. Then i is as required. $\square$

If the binary vectors $\mathbf{c}_1, \ldots, \mathbf{c}_M$ of length n form a CDC and $M = 2^n/(n-2)$, it is again natural to call this sequence a *perfect* CDC.

Theorem 4 *Assume that $\mathbf{c}_1$, $\mathbf{c}_2$, $\ldots$, $\mathbf{c}_M$ is a sequence of codewords of length $n \geq 5$ and $M = 2^n/(n-2)$. Then the following two properties are equivalent:*

i) the sequence forms a CDC;

ii) the sequence forms a length n, spread 3 circuit code.

Proof Assume that i) holds. From the previous proof we see that the sets A_i must be disjoint and each must have cardinality $n-2$. In particular,

$d(\mathbf{c}_i, \mathbf{c}_{i+2}) = 2$ for all i (cyclically). Moreover, all the $\mathbf{c}_i$ must be different codewords: if $\mathbf{c}_i = \mathbf{c}_j = \mathbf{0}$, say, then A_i and A_j both contain $n-2 > \frac{1}{2}n$ vectors of weight 1 and hence have a nonempty intersection.

Without loss of generality, $\mathbf{c}_i = 000\ldots0$, $\mathbf{c}_{i-1} = 100\ldots0$ and $\mathbf{c}_{i-2} = 110\ldots0$ and $\mathbf{c}_s \neq \mathbf{c}_i$. We show that $d(\mathbf{c}_i, \mathbf{c}_s) > 2$ whenever $|i-s| > 2$.

Denote by $w(\mathbf{c})$ the weight of $\mathbf{c}$, i.e., the number of 1's in it. Let $\mathbf{c}_j$ be any codeword covering $111\ldots1$. Then $w(\mathbf{c}_j) \geq 4$. Let k be the first index among $j, j+1, \ldots, i-3$ such that $d(\mathbf{c}_k, \mathbf{c}_i) \leq 2$, i.e., $w(\mathbf{c}_k) = 2$. Then A_k contains two vectors of weight 1, and since $\mathbf{c}_k \neq \mathbf{c}_{i-2}$, the intersection of A_k and A_i is nonempty, a contradiction. Assume that k is the last index among $i+3, i+4, \ldots, j-1, j$ such that $d(\mathbf{c}_k, \mathbf{c}_i) \leq 2$. Then we have the same contradiction as before, but now for the perfect CDC in which we go through the same codewords in the reverse order.

Conversely, if ii) holds, then the sets A_i in the previous proof are disjoint. Indeed, $A_i \cap A_j \neq \emptyset$ implies $d(\mathbf{c}_i, \mathbf{c}_j) \leq 2$, but the only codewords within distance two from $\mathbf{c}_i$ are $\mathbf{c}_{i-1}$, $\mathbf{c}_{i-2}$, $\mathbf{c}_{i+1}$ and $\mathbf{c}_{i+2}$ and by the definition of the sets A_j, A_i has an empty intersection with A_{i-1}, A_{i-2}, A_{i+1} and A_{i+2}. Each A_i has cardinality $n-2$, and $M = 2^n/(n-2)$ therefore implies that the union of A_i's is the whole space. Hence the sequence forms a CDC. □

So, if a sequence of codewords of length n attains the upper bound $2^n/(n-2)$ on length n, spread 3 circuit codes, or the lower bound $2^n/(n-2)$ on CDCs of length n, it is simultanously a CDC and a spread 3 circuit code. Notice that the result is not valid for $n = 4$: in the proof of Theorem 5 we have a sequence of codewords of length 4 and cardinality $2^4/(4-2) = 8$, which is a CDC but not a spread 3 circuit code. However, it is known that there is a length 4 spread 3 circuit code with 8 codewords (see, e.g., the table in [13]), and therefore the converse part of the previous proof (which is valid also for $n = 4$) gives the result $S(4) = 8$ from circuit codes. In the same way we get the bound $S(6) = 16$.

It is immediately clear from the definition that in a CDC each coordinate must all in all be changed an even number of times, and therefore the cardinality of every CDC is even.

Lemma 3 *For $n \geq 2$, the cardinality of a CDC is even.* □

Corollary 1 *For* $n \geq 4$,

$$S(n) \geq 2\lceil \frac{1}{2} \lceil \frac{2^n}{n-2} \rceil \rceil.$$

Lemma 4

$$S(n+1) \leq 2S(n).$$

Proof Again, if $\mathbf{c}_1, \ldots, \mathbf{c}_M$ is a CDC of length n then

$$(\mathbf{c}_1, 0), \ldots, (\mathbf{c}_M, 0), (\mathbf{c}_M, 1), \ldots, (\mathbf{c}_1, 1)$$

is a CDC of length $n+1$. $\square$

Theorem 5 $S(1) = 1, S(2) = 2, S(3) = 4, S(4) = 8, S(5) = 12, S(6) = 16, S(7) = 28$.

Proof The following CDCs attain the given bounds (here we mention only the indices of the coordinates to be changed):
$n = 2$: 1
$n = 3$: 121
$n = 4$: 1231243
$n = 5$: 12341523141
$n = 6$: 123415361234153
$n = 7$: 123425647616524231256176465.

The lower bounds follow from Corollary 1 for $n = 4, 5, 6$ and from $S(n) \geq T(n)$ for $n = 7$ and for the easy case $n = 3$. $\square$

4 Asymptotically optimal CDPs

As we have seen from the lower bound of Theorem 1, the length of CDPs cannot be less than

$$\frac{2^n}{n}(1 + o(1)).$$

In what follows we will show that asymptotically this can be attained. Let $M(n)$ be the length of the shortest CDPs in dimension n, and define

$$\delta_n = \frac{M(n)n}{2^n}$$

and

$$\delta = \lim_{n\to\infty} \delta_n.$$

We will show that the limit exists and equals 1.

Although we will later consider the case of arbitrary n, we start from a subsequence of dimensions illustrating the idea.

Let $\mathcal{G}^m = (\mathbf{g}_1^m, \mathbf{g}_2^m, \ldots, \mathbf{g}_{2^m}^m)$ stand for a Gray code of length m, i.e. an ordering of all the 2^m binary m-vectors satisfying $d(\mathbf{g}_i^m, \mathbf{g}_{i+1}^m) = 1$ for $i = 1, \ldots, 2^m - 1$, and $d(\mathbf{g}_{2^m}^m, \mathbf{g}_1^m) = 1$. A possible way to construct such a code is to use reflection, i.e., given a Gray code of length m we produce another one of length $m+1$ as follows:

$$\mathcal{G}^{m+1} = \{(\mathbf{g}_1^m, 0), (\mathbf{g}_2^m, 0), \ldots, (\mathbf{g}_{2^m}^m, 0), (\mathbf{g}_{2^m}^m, 1), (\mathbf{g}_{2^m-1}^m, 1), \ldots, (\mathbf{g}_1^m, 1)\}.$$

Consider now the Hamming code C^m of length $2^m - 1$ having $2^k = 2^{2^m-m-1}$ codewords. The generator matrix of the code, say V_m, can be composed of vectors of weight 3 (see, e.g., [17] for a proof of this fact), say, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. Using the construction of van Zanten [19], we obtain the following ordering of the Hamming code:

$$\begin{aligned} C^m &= (\mathbf{c}_1^m, \mathbf{c}_2^m, \ldots, \mathbf{c}_{2^k}^m) \\ &= (\mathbf{g}_1^k V_m, \mathbf{g}_2^k V_m, \ldots, \mathbf{g}_{2^k}^k V_m) \end{aligned}$$

which satisfies $d(\mathbf{c}_i^m, \mathbf{c}_{i+1}^m) = 3$ for $i = 1, \ldots, 2^k - 1$, and $d(\mathbf{c}_{2^k}^m, \mathbf{c}_1^m) = 3$. Thus we can construct another code

$$\hat{C}^m = (\hat{\mathbf{c}}_1^m, \hat{\mathbf{c}}_2^m, \ldots, \hat{\mathbf{c}}_{3\times 2^k}^m)$$

$$= (\mathbf{c}_1^m, \mathbf{c}_1^m + \mathbf{e}_{1,1}, \mathbf{c}_1^m + \mathbf{e}_{1,1} + \mathbf{e}_{1,2}, \mathbf{c}_2^m, \mathbf{c}_2^m + \mathbf{e}_{2,1}, \mathbf{c}_2^m + \mathbf{e}_{2,1} + \mathbf{e}_{2,2}, \ldots,$$
$$\mathbf{c}_{2^k}^m, \mathbf{c}_{2^k}^m + \mathbf{e}_{2^k,1}, \mathbf{c}_{2^k}^m + \mathbf{e}_{2^k,1} + \mathbf{e}_{2^k,2}),$$

where $\mathbf{e}_{i,1}$ is the binary vector having zeros in all coordinates except the first one where the vectors $\mathbf{c}_i^m$ and $\mathbf{c}_{i+1}^m$ differ, where it has 1. The definition of $\mathbf{e}_{i,2}$ is similar, except that we place the only 1 in the second position where the corresponding vectors differ. The words of this new code $\hat{C}^m$ possess the property that $d(\hat{\mathbf{c}}_i^m, \hat{\mathbf{c}}_{i+1}^m) = 1$.

Now, we construct a new code P^m of length $2^m + m - 1$ in the following way:

$$P^m = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_M)$$

$$= ((\hat{\mathbf{c}}_1^m, \mathbf{g}_1^m), (\hat{\mathbf{c}}_1^m, \mathbf{g}_2^m), \dots, (\hat{\mathbf{c}}_1^m, \mathbf{g}_{2^m}^m), (\hat{\mathbf{c}}_2^m, \mathbf{g}_{2^m}^m), (\hat{\mathbf{c}}_3^m, \mathbf{g}_{2^m}^m),$$
$$(\hat{\mathbf{c}}_4^m, \mathbf{g}_{2^m}^m), (\hat{\mathbf{c}}_4^m, \mathbf{g}_{2^m-1}^m), \dots, (\hat{\mathbf{c}}_4^m, \mathbf{g}_1^m), (\hat{\mathbf{c}}_5^m, \mathbf{g}_1^m), (\hat{\mathbf{c}}_6^m, \mathbf{g}_1^m),$$
$$\dots$$
$$(\hat{\mathbf{c}}_{3\times 2^k-5}^m, \mathbf{g}_1^m), (\hat{\mathbf{c}}_{3\times 2^k-5}^m, \mathbf{g}_2^m), \dots, (\hat{\mathbf{c}}_{3\times 2^k-5}^m, \mathbf{g}_{2^m}^m), (\hat{\mathbf{c}}_{3\times 2^k-4}^m, \mathbf{g}_{2^m}^m), (\hat{\mathbf{c}}_{3\times 2^k-3}^m, \mathbf{g}_{2^m}^m),$$
$$(\hat{\mathbf{c}}_{3\times 2^k-2}^m, \mathbf{g}_{2^m}^m), (\hat{\mathbf{c}}_{3\times 2^k-2}^m, \mathbf{g}_{2^m-1}^m), \dots, (\hat{\mathbf{c}}_{3\times 2^k-2}^m, \mathbf{g}_1^m), (\hat{\mathbf{c}}_{3\times 2^k-1}^m, \mathbf{g}_1^m), (\hat{\mathbf{c}}_{3\times 2^k}^m, \mathbf{g}_1^m)).$$

Here we concatenate the vectors of $\hat{C}^m$ with vectors of $\mathcal{G}^m$ in such a way that with $\hat{\mathbf{c}}_{6i+1}^m$ the vectors of $\mathcal{G}^m$ are listed in the direct order, while with $\hat{\mathbf{c}}_{6i+4}^m$ they are listed in the inverse order. The vectors $\hat{\mathbf{c}}_{6i+2}^m$ and $\hat{\mathbf{c}}_{6i+3}^m$ are concatenated only with the vector $\mathbf{g}_{2^m}^m$, while $\hat{\mathbf{c}}_{6i+5}^m$ and $\hat{\mathbf{c}}_{6i+6}$ are concatenated with $\mathbf{g}_1^m$. The code P^m is a CDC (and CDP even without the last two vectors): the claim about the distances between the consecutive vectors is trivial, and to see that P^m is a covering it is enough to notice that P^m contains $C^m \oplus F^m$ as its subcode.

The total length of the CDC is $2^{k+m} + 2^{k+1}$. Thus, for $n = 2^m + m - 1$,

$$\delta_n = \frac{n(2^{k+m} + 2^{k+1})}{2^n} = 1 + O(m2^{-m}) = 1 + o(1).$$

To extend the idea of the construction we need the following result from [11] (for a simplified proof see [4]).

Lemma 5 *For every, big enough, length n there exists a radius 1 covering code with $\frac{2^n}{n}(1 + o(1))$ codewords.*

Employing the words of such codes in the previous construction instead of the words of Hamming codes gives us the sought result. Assume that we have a sequence of codes C_n where C_n has length n, covering radius one and cardinality $M_n = 2^n(1 + f(n))/n$ where $f(n) \to 0$ when $n \to \infty$. Without loss of generality, M_n is even for all n.

It is easy to check that for our purposes it is enough that there exists an ordering of the codewords $(\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_{M_n})$ of each C_n such that $\sum d(\mathbf{c}_i, \mathbf{c}_{i+1}) \leq M_n c$, where the constant c does not depend on n. In other words, the average distance between two consecutive words in the ordering should be a constant not depending on n. Indeed, if this is the case, then we consider the resulting CDCs of lengths $n + m$ where $m = \lceil \log_2 n \rceil$ and

$m = \lceil \log_2 n \rceil - 1$. They have cardinality at most $M_n 2^m + (c-1)M_n$ and density

$$\begin{aligned}\delta_{n+m} &\leq \frac{(n+m)\frac{2^n(1+f(n))}{n}(2^m+c-1)}{2^{n+m}} \\ &= (1+\frac{m}{n})(1+f(n))(1+\frac{c-1}{2^m}) \\ &= 1+o(1).\end{aligned}$$

Moreover, $n+m$ gets all integers values (when we use all n and our two choices for m).

It remains to prove the following lemma.

Lemma 6 *Let C_n be a sequence of codes, where C_n has length n and size $M_n = \frac{2^n}{n}(1+o(1))$. For all n, there exists an ordering $\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_{M_n}\}$ of the codewords of C_n such that $\sum d(\mathbf{c}_i, \mathbf{c}_{i+1}) \leq M_n c$, where c is a constant not depending on n.*

Proof We employ an idea from [5] and [12] of a greedy construction path. Let $d(n, M)$ stand for the maximal possible minimum distance of any binary code of length n and size M. It is proved in [5, Theorem 1, part II] that any code of size M can be ordered in such a way that

$$\frac{1}{M}\sum_{i=1}^{M} d(\mathbf{c}_i, \mathbf{c}_{i+1}) \leq \frac{1}{M}\sum_{i=1}^{M} d(n,i).$$

Taking into account that $d(n,i)$ is a non-increasing function in i and evident $d(n,i) \leq n$ we conclude

$$\frac{1}{M}\sum_{i=1}^{M} d(n,i) \leq \frac{1}{M}\sum_{i=\lceil 2^n/n^3 \rceil}^{M} d(n,i) + \frac{2^n n}{Mn^3}$$

$$\leq d\left(n, \lceil \frac{2^n}{n^3} \rceil\right) + o(1) \leq 8 + o(1).$$

The last inequality follows from the classical Hamming bound

$$MV(d(n,M), n) \leq 2^n,$$

where $V(r,n) = \sum_{j=0}^{r} \binom{n}{j}$ is the volume of the Hamming ball of radius r. $\square$

This lemma completes our proof of the following result.

Theorem 6

$$\delta = 1.$$

□

The previous argument is in fact valid for all alphabets whose cardinality is a prime power. It is easy to construct nonbinary Gray codes. For length $n = 1$ the construction is trivial. If $\mathbf{c}_1$, $\mathbf{c}_2$, ..., $\mathbf{c}_m$ is a Gray code of length n over the alphabet $\{0, 1, 2, \ldots, q-1\}$, then

$$(\mathbf{c}_1, 0), (\mathbf{c}_1, 1), \ldots, (\mathbf{c}_1, q-2), (\mathbf{c}_2, q-2), (\mathbf{c}_2, q-3), \ldots, (\mathbf{c}_2, 0),$$

$$\ldots, (\mathbf{c}_m, 0 \text{ or } q-2), (\mathbf{c}_m, q-1), (\mathbf{c}_{m-1}, q-1), \ldots, (\mathbf{c}_1, q-1)$$

is a Gray code of length $n+1$. The necessary modifications for the proof are very minor and are omitted.

Theorem 7 *Assume that q is a prime power. Then*

$$\frac{(q-1)nM(n)}{q^n} \to 1$$

when $n \to \infty$. □

References

[1] P. Ash and B. Jackson, Dominating cycles in bipartite graphs, in: Progress in Graph Theory, J. A. Bondy, Ed., Academic Press, Waterloo, Ontario, 1982, pp. 81-87.

[2] K. Chakrabarty, M. G. Karpovsky, L. B. Levitin, Fault isolation and diagnosis in multiprocessor systems with point-to-point connections, in: Fault Tolerant Parallel and Distributed Systems, Kluwer, Boston, 1998, pp. 285–301.

[3] B. N. Clark, C. J. Colbourn, and P. Erdös, A conjecture on dominating cycles, Congr. Numer. 47 (1985) 189-197.

[4] G. Cohen, I. Honkala, S. Litsyn, and A. Lobstein, Covering Codes, Elsevier, Amsterdam, 1997.

[5] G. Cohen, S. Litsyn and G. Zémor, On the traveling salesman problem in binary Hamming spaces, IEEE Trans. Info. Theory 42 (1996) 1274–1276.

[6] C. J. Colbourn and L. K. Stewart, Dominating cycles in series-parallel graphs, Ars Combin. 19A (1985) 107-112.

[7] K. Deimer, Some new bounds for the maximum length of circuit codes, IEEE Trans. Inform. Theory 30 (1984) 754–756.

[8] P. Fraisse, Dominating cycles and paths, Notices Amer. Math. Soc. 5 (1984) 231.

[9] J. P. Hayes, T. N. Mudge, Q. F. Stout, S. Colley, J. Palmer, Architecture of a hypercube supercomputer, in: Proc. Int. Conf. Parallel Processing, 1986, pp. 653–660.

[10] S. M. Hedetniemi, S. T. Hedetniemi, R. C. Laskar, Domination in trees: models and algorithms, in: Graph Theory with Applications to algorithms and Computer Science, Y. Alavi et. al., Eds., Wiley, New York, 1985, pp. 423-442.

[11] G. A. Kabatyanskii, V. I. Panchenko, Unit sphere packings and coverings of the Hamming space, Problemy Peredachi Informatsii 24 (1988) 3–16.

[12] L. B. Levitin, M. G. Karpovsky, Traveling salesman problem in the space of binary vectors, in: Proceedings of 1994 IEEE Symposium on Information Theory, 1994, p. 316.

[13] K. G. Paterson, J. Tuliani, Some new circuit codes, IEEE Trans. Inform. Theory 44 (1998) 1305–1309.

[14] A. Proskurowski, Minimum dominating cycles in 2-trees, Internat. J. Comput. Inform. Sci. 8 (1979) 405-417.

[15] C. L. Seitz, The cosmic cube, Commun. ACM, 28(1), 1985.

[16] Y. Shih, J. Fier, Hypercube systems and key applications, in: Parallel Processing for Supercomputing and Artificial Intelligence, McGraw-Hill, New York, 1989, pp. 203-244.

[17] J. Simonis, On generator matrices of codes, IEEE Trans. Inform. Theory 38 (1992) 516.

[18] H. J. Veldman, Existence of dominating cycles and paths, Discrete Math. 43 (1983) 281-296.

[19] A. J. van Zanten, Minimal-change order and separability in linear codes, IEEE Trans. Inform. Theory 39 (1993) 1988-9.