ՀԱՅԱՍՏԱՆԻ ԳԻՏՈՒԹՅՈՒՆՆԵՐԻ ԱԶԳԱՅԻՆ ԱԿԱԴԵՄԻԱ
НАЦИОНАЛЬНАЯ АКАДЕМИЯ НАУК АРМЕНИИ
NATIONAL ACADEMY OF SCIENCES OF ARMENIA
ДОКЛАДЫ ԶԵԿՈՒՅՑՆԵՐ REPORTS

Том 110 2010 № 1

MATHEMATICS

УДК 512.624

H. K. Nurijanyan

On the Length of the Shortest Linearised Covering for "Almost All" Subsets in a Finite Field

(Submitted by the academician Yu. H. Shoukourian 8/II 2010)

Keywords: *Finite fields, system of linear equations over finite fields, linearised coverings*

1. Introduction. Throughout this paper F_q stands for a finite field with q elements, and F_q^n for an n-dimensional linear space over F_q. If L is a linear subspace in F_q^n, then the set $\bar{\alpha} + L \equiv \{\bar{\alpha} + \bar{x} | \bar{x} \in L\}$, $\bar{\alpha} \in F_q^n$ is a coset (or translate) of the subspace L and $dim(\bar{\alpha} + L)$ coincides with $dim L$. An equivalent definition: a subset $N \subseteq F_q^n$ is a coset if whenever $\bar{x}^1, \bar{x}^2, \ldots, \bar{x}^m$ are in N, so is any affine combination of them, i.e., so is $\sum_{i=1}^{m} \lambda_i \bar{x}^i$ for any $\lambda_1, \lambda_2, \ldots, \lambda_m$ in F_q such that $\sum_{i=1}^{m} \lambda_i = 1$. It can be readily verified that any k-dimensional coset in F_q^n can be represented as a set of solutions of a certain system of linear equations over F_q of rank $n - k$ and vice versa.

Let N is a subset in F_q^n.

Definition 1. *A set of cosets $\{H_1, H_2, \ldots, H_m\}$ in F_q^n forms a linearised covering of N if $N = \bigcup_{i=1}^{m} H_i$. The length (or complexity) of the covering is equal to the number of cosets, i.e. m.*

Definition 2. *A linearised covering is the shortest for the given N if it has the smallest possible length.*

Definition 3. *Let π_n be the number of subsets in F_q^n, that satisfy a certain property Π. If $\lim_{n\to\infty} \frac{\pi_n}{2^{2^n}} = 1$ then we say that "almost all" subsets of F_q^n satisfy the property Π.*

The general problem of covering with cosets is stated as follows: for a given subset in F_q^n (which is usually given as a set of solutions of a polynomial equation

with n unknowns over F_q) estimate the length of the shortest linearised covering and find an effective algorithm that constructs the shortest or "close" to the shortest linearised covering for N. This problem was originally considered in [1,2] for $q = 2$ in connection with minimization of Boolean functions. It was shown in [3] that the length of the shortest covering $L_q(N)$ for almost all subsets satisfies the following inequalities:

$$(1-\varepsilon_n)\frac{q^n}{2qn\log_q n} \le L_q(N) \le (1-\delta_n)\frac{3q^3q^n\log_q n}{2n\log_q e}, \tag{1}$$

where $\lim\limits_{n\to\infty}\varepsilon_n = \lim\limits_{n\to\infty}\delta_n = 0$.

The aim of the present paper is to improve the upper bound in (1) proving that $L_q(N) \le \text{const}\,\frac{q^n}{n}$. For that reason we use the techniques developed in [4].

2. Main Result. Theorem. *For almost all subsets in* F_q^n

$$L_q(N) \le c\frac{q^n}{n}, \tag{2}$$

where $c = \frac{q^{3-\ln 2}e^2(\ln 2+1)}{2\ln 2} \approx 18q^{3-\ln 2}$.

This theorem is a result of the following set of affirmations.

Let $T = (t_{i,j})$ be a Boolean matrix $(t_{i,j} \in \{0,1\})$.

Definition 4. *We say a column of T covers a row, if at the row's and column's intersection stands 1.*

Definition 5. *The sequence of columns* $a_1, a_2, \ldots, a_k$ *is called gradient, if for every* $i = 1,2,\ldots,k$ *column* a_i *covers maximal number of rows, which are still not covered by the columns* $a_1, a_2, \ldots, a_{i-1}$. *Number* k *is called the length of gradient sequence.*

Denote the number of rows and columns of T by $p(T)$ and $q(T)$ respectively. Let $L_\delta(T)$, $(\delta \ge 0)$ be the minimal number k, so that for every gradient sequence with length k takes place the following inequality

$$\frac{\overline{p}(T)}{p(T)} \le \frac{1}{e^\delta},$$

where $\overline{p}(T)$ stands for the number of uncovered rows in T. $\frac{\overline{p}(T)}{p(T)}$ is called "the fraction of uncovered rows".

Lemma 1. *Let* $\bar{T}$ *be such a submatrix of* T, *that every row in* $\bar{T}$ *is covered by not less than* $\chi q(\bar{T})$ *columns* $(\chi > 0)$ *and*

$$P(\bar{T}) \ge (1-\varepsilon)p(T),\ \varepsilon \in (0,1),$$

then

$$L_\delta(T) \le \frac{\delta}{\chi} + 1 + \varepsilon p(T).$$

Denote by $F(n)$ the set of all subsets of F_q^n.

Definition 6. *Let us consider a probability distribution on $F(n)$, such that random variables*

$$\xi_{\bar{x}}(N) = \begin{cases} 1; & \bar{x} \in N, \\ 0; & \bar{x} \notin N \end{cases}$$

are independent and equally distributed, and also $P(\xi_{\bar{x}} = 1) = 2^{-\lambda}$.

We denote by $\psi_k(N)$ the number of k-dimensional cosets ($0 \leq k \leq n$) in random subset N, and by $\eta_{\bar{x}}^k(N)$ the number of such cosets H, which satisfy $\bar{x} \in H$ and $H \setminus \bar{x} \subseteq N$.

Lemma 2. *1)* $M\psi_k = 2^{-\lambda q^k} q^{n-k} \begin{bmatrix} n \\ k \end{bmatrix}_q$

2) $M\psi_k^2 = 2^{-2\lambda q^k} \cdot \sum_{r=0}^{k} q^{(k-r)^2} \begin{bmatrix} n-k \\ k-r \end{bmatrix}_q \begin{bmatrix} n-r \\ k-r \end{bmatrix}_q q^{n-r} \begin{bmatrix} n \\ r \end{bmatrix}_q (2^{\lambda q^r} - 1) + (M\psi_k)^2$, *where $M\psi_k$ is the expectation value of ψ_k, and* $\begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{(q^n-1)(q^{n-1}-1)\cdot(q^{n-k+1}-1)}{(q^k-1)(q^{k-1}-1)\,(q-1)}$, *is the number of k-dimensional linear subspaces in F_q^n (*$\begin{bmatrix} n \\ k \end{bmatrix}_q$ *is called Gauss coefficient).*

Lemma 3. *1)* $M\eta_{\bar{x}}^k = \begin{bmatrix} n \\ k \end{bmatrix}_q \cdot 2^{-\lambda(q^k-1)}$

2) $M(\eta_{\bar{x}}^k)^2 = 2^{-\lambda(2q^k-1)} \begin{bmatrix} n \\ k \end{bmatrix}_q \cdot \sum_{r=0}^{k} q^{(k-r)^2} \begin{bmatrix} n-k \\ k-r \end{bmatrix}_q \begin{bmatrix} k \\ r \end{bmatrix}_q 2^{\lambda q^r}$, *where $M\eta_{\bar{x}}^k$ is the expectation value of $\eta_{\bar{x}}^k$.*

Lemma 4. *For every k ($0 \leq k \leq n$)*

$$\frac{D\psi_k}{(M\psi_k)^2} \leq \frac{2^{\lambda q^k}}{q^{n-k}}.$$

Lemma 5. *If $k = \lfloor \log_q n - \log_q \lambda - \delta - 1 \rfloor$, $\delta \in (0,1)$ then*

$$\frac{D\eta_{\bar{x}}^k}{(M\eta_{\bar{x}}^k)^2} \leq \frac{k \cdot 2^{\lambda(q-1)} \cdot q^{3k}}{q^n}.$$

Definition 7. *For every subset N let us define Boolean matrix T_N, which columns are all cosets in N, and rows are all vectors in N. If a vector belongs to a coset then at the intersection of the corresponding row and column stands 1, otherwise 0.*

Assume that $L_\delta(T_N) = L_\delta(N)$ and $\varphi_+(x) = \frac{x+|x|}{2}$.

Lemma 6. *If $\lambda \geq 1$ and $\delta \in (0,1)$, then*

$$M\varphi_+\left(L_\delta(N) - q^{2+\delta}\lambda\delta\frac{q^n 2^{-\lambda}}{n}\right) \leq \frac{q^n}{n\log_q^2 n}.$$

As in [4], we divide the set of variables $X = \{x_1, x_2, \ldots, x_n\}$ to not intersecting subsets $X = X^1 \cup X^2 \cup \ldots \cup X^k \cup Y$.

$$|X^i| = m;\ i = 1, \ldots, k;\ k = [\log_q n];\ m = \left[\frac{n}{\log_q n}\right].$$

Suppose ν_δ is an operator associating every $N \subseteq F_q^n$ with a set of cosets which form such gradient sequence that the fraction of uncovered rows does not exceed $e^{-\delta}$ and removing the last member of the gradient sequence breaks the relation.

Let us associate every subset $N \subseteq F_q^n$ with a sequence of subsets $N_0, N_1, \ldots, N_k$ in the following way:

1) $N_0 = N$.

2) Suppose that N_{i-1}, $i \leq k$ are already constructed. N_{i-1}^j $j = 1, \ldots, q^{n-m}$ are subsets of N_{i-1} dependent only from variables of X^j. Let $\nu_\delta^i(N) = \nu_\delta(N_{i-1}^1) \cup \ldots \cup \nu_\delta(N_{i-1}^{q^{n-m}})$, then N_i is constructed as follows: $\bar{x} \in N_i \Leftrightarrow \bar{x} \notin \nu_\delta^i(N)$ and $\bar{x} \in N_{i-1}$.

Denote by $\nu^k(N)$ the largest gradient sequence of N_k, and by $L_{\nu,\delta}(N)$

$$L_{\nu,\delta}(N) = \bigcup_{i=1}^{k} \nu_\delta^i(N) \cup \nu^k(N)$$

Lemma 7. *If $\lambda \geq 1$ and $\delta \geq 1 - \ln 2$ then*

$$M\varphi_+\left(L_{\nu,\delta} - \frac{q^3 e^2}{2q^{\ln 2}} \frac{q^n}{n} 2^{-\lambda} \frac{\lambda \ln 2 + 1}{\ln 2}\right) \leq \frac{q^n}{n \log_q n} \tag{3}$$

Finally, putting $\lambda = 1$ in (3) proves the assertion of the theorem.

Yerevan State University
E-mail address: hovikn@gmail.com (H.Nurijanyan)

H. K. Nurijanyan

On the Length of the Shortest Linearised Covering for "Almost All" Subsets in a Finite Field

Given an n-dimensional linear space over a finite field for an arbitrary number n. For a given subset in the linear space. It is necessary to determine the minimal number of systems of linear equations over the same field such that the union of all vectors representing the solutions of the systems covers that subset. The problem is solved for "almost all" subsets of the linear space.

Ո. Ք. Նուրիջանյան

Վերջավոր դաշտի "համարյա բոլոր" ենթաբազմությունների գծայնացվող ծածկույթի երկարության մասին

Կամայական n բնական թվի համար տրված է վերջավոր դաշտի վրա կառուցված n-չափանի գծային տարածություն: Պահանջվում է որոշել նույն վերջավոր դաշտի վրա տրված գծային հավասարումների համակարգերի նվազագույն քանակը, որոնց լուծումների բազմությունը ծածկում է գծային տարածությունում տրված ենթաբազմությանը: Խնդիրը լուծված է գծային տարածության "համարյա բոլոր" ենթաբազմությունների համար:

О. К. Нуриджанян

О длине кратчайшего линеаризированного покрытия для "почти всех" подмножеств конечного поля

Для произвольного натурального числа задано n-мерное линейное над конечным полем пространство. Требуется определить минимальное количество систем линейных уравнений, заданных над тем же полем, таких, что множество всех решений этих систем покрывает данное подмножество из линейного пространства. Задача решена для "почти всех" подмножеств линейного пространства.

References

1. *Алексанян А.* Дизъюнктивные нормальные формы над линейными функциями (Теория и приложения). Изд. ЕГУ. 1990. 201 с.

2. *Aleksanyan A.* Soviet Math. Dokl. 1989. V.39. N1. P. 131-135.

3. *Габриелян В.* О метрических характеристиках, связанных с покрытиями подмножеств конечных полей смежными классами линейных подпространств. Институт проблем информатики и автоматизации. Ереван. 2004. 50 с.

4. *Андреев А.* Вестник МГУ. 1985. N3. С. 29-35.