Large families of pseudorandom subsets formed by power residues

Cécile Dartyge, András Sárközy

Uniform Distribution Theory **2** (2007), no.2, 73–88

uniform distribution theory

LARGE FAMILIES OF PSEUDORANDOM SUBSETS FORMED BY POWER RESIDUES

Cécile Dartyge — András Sárközy

ABSTRACT. In an earlier paper the authors introduced the measures of pseudorandomness of subsets of the set of the positive integers not exceeding N, and they also presented two examples for subsets possessing strong pseudorandom properties. One of these examples included permutation polynomials $f(X) \in \mathbb{F}_p[X]$ and d-powers in $\mathbb{F}_p$. This construction is not of much practical use since very little is known on permutation polynomials and there are only very few of them. Here the construction is extended to a large class of polynomials which can be constructed easily, and it is shown that all the subsets belonging to the large family of subsets obtained in this way possess strong pseudorandom properties. The complexity of this large family is also studied.

Communicated by Christian Mauduit

1. Introduction

This paper is the third of a series of articles devoted to the study of pseudorandom subsets. Let N be a positive integer and $\mathcal{R} \subset \{1, \dots, N\}$. We associate with this subset the sequence $E_N = (e_1, \dots, e_N) \in \left\{1 - \frac{|\mathcal{R}|}{N}, \frac{|\mathcal{R}|}{N}\right\}^N$ defined by

$$e_n = \begin{cases} 1 - \frac{|\mathcal{R}|}{N} & \text{for } n \in \mathcal{R} \\ -\frac{|\mathcal{R}|}{N} & \text{for } n \notin \mathcal{R} \end{cases} \quad (n = 1, \dots, N). \tag{1.1}$$

2000 Mathematics Subject Classification: 05A05, 11Z05.
Keywords: Pseudo-randomness, subset of positive integers, complexity.
Research partially supported by the Hungarian National Foundation for Scientific Research, Grant T043623 and T049693, and by French-Hungarian EGIDE-OMKFHÁ exchange program Balaton F-2/03.

We introduced in [4] two measures of pseudorandomness for the sequence $\mathcal{R}$. The first one is the well distribution measure :

$$W(\mathcal{R},N)=\max_{a,b,t}\Big|\sum_{j=0}^{t-1}e_{aj+b}\Big|,$$

where the maximum is over all $a,b,t\in\mathbb{N}$ such that $1\le b\le b+(t-1)a\le N$.

The second measure is the correlation measure of order k :

$$C_k(\mathcal{R},N)=\max_{M,D}\Big|\sum_{n=1}^{M}e_{n+d_1}\cdots e_{n+d_k}\Big|,$$

where the maximum is taken over all $D=(d_1,\dots,d_k)$ and M such that $0\le d_1<\ldots<d_k\le N-M$. If these two measures are small, i.e., are $o(N)$, then we will say that the subset $\mathcal{R}$ has strong pseudorandom properties.

These two measures are closely related to the measures of pseudorandomness of binary sequences introduced by Mauduit and Sárközy [9] and of the p-pseudorandomness binary sequences defined by Hubert and Sárközy [6]. See [4] for a detailed bibliography.

In [4] we gave two examples of subsets with strong pseudorandom properties, but these examples generated a few number of such subsets.

Recently Elie Mosaki and the authors [3] constructed the following large family of pseudorandom subsets.

For a prime number p, a polynomial $f(x)\in\mathbb{F}_p[X]$ of degree $d\ge 2$, and some integers $r\in\mathbb{Z}$, $s\in\mathbb{N}$, such that $s<p/2$, they defined the subset $\mathcal{R}\subset\{1,\dots,p\}$ by

$$\begin{cases} n\in\mathcal{R} & \text{if } \exists h\in\{r,r+1,\dots,r+s-1\} \text{ with } f(n)\equiv h\,(\mathrm{mod}\,p)\\ n\notin\mathcal{R} & \text{otherwise.}\end{cases}\tag{1.2}$$

We proved

$$W(\mathcal{R},p)<2d\sqrt{p}\log^2 p\tag{1.3}$$

and, for $2\le k\le d-1$,

$$C_k(\mathcal{R},p)<2d\sqrt{p}(1+\log p)^{k+1}.\tag{1.4}$$

We also gave a construction of a large family of pseudorandom subsets using the multiplicative inverse.

In this paper we extend and generalize the first construction of pseudorandom subsets given in [4]. This will provide another large and rich family of subsets with strong pseudorandom properties.

Let $p \geq 2$ be a prime number, $d|p-1$ and $f \in \mathbb{F}_p[X]$. We will denote the algebraic closure of $\mathbb{F}_p$ by $\bar{\mathbb{F}}_p$. We would like to study the pseudorandom properties of the set

$$V_f = \{x \in \mathbb{F}_p, \exists y \in \mathbb{F}_p \setminus \{0\} : f(x) \equiv y^d \,(\mathrm{mod}\, p)\}.$$

In [4], we proved that if f is a permutation polynomial whose unique zero in $\mathbb{F}_p$ has a multiplicity coprime with d, then the subset V_f of $\mathbb{F}_p$ has strong pseudorandom properties. This result was inspired by a construction of Mauduit and Sárközy [9] for binary pseudorandom sequences; they considered the sequence $\left(\frac{f(n)}{p}\right)$, $n = 1, \ldots, p$, where f is a permutation polynomial whose zero in $\mathbb{F}_p$ has odd multiplicity and $\left(\frac{x}{p}\right)$ is the Legendre symbol. The weak point of these constructions is that we know only very little on permutation polynomials.

In [5], Goubin, Mauduit and Sárközy proved that if $f \in \mathbb{F}_p[X]$ is a polynomial of degree $k > 0$ and with no multiple roots in $\bar{\mathbb{F}}_p$ then the sequence $\{u_n\}_{n\in\mathbb{N}}$ defined by

$$u_n = \begin{cases} \left(\frac{f(n)}{p}\right) & \text{for } (p, f(n) = 1, \\ 1 & \text{if } p|f(n), \end{cases}$$

satisfies

$$\max_{a,b,t} \left| \sum_{j=0}^{t-1} u_{aj+b} \right| < 10k\sqrt{p}\log p,$$

where the maximum is over all $a, b, t \in \mathbb{N}$ such that $1 \leq a \leq a(t-1)+b \leq p$. For the correlations measure of order ℓ they proved that if one of the following assumptions holds:

(i) $\ell = 2$;

(ii) $\ell < p$ and 2 is a primitive root modulo p;

(iii) $(4k)^\ell < p$;

then the correlation measure of order ℓ satisfies

$$\max_{M,D} \left| \sum_{n=1}^{M} u_{n+d_1} \cdots u_{n+d_\ell} \right| < 10k\ell\sqrt{p}\log p,$$

where the maximum is taken over all $D = (d_1, \ldots, d_\ell)$ and M such that $0 \leq d_1 < \ldots < d_\ell \leq p - M$.

In the present paper we will adapt this idea for constructing pseudorandom subsets and generalize it to d-th power residues for a divisor $d|p-1$.

First we will state a general result on the cardinality of V_f.

Theorem 1.1. *Let f be a polynomial in $\mathbb{F}_p[x]$ of degree k and with m distinct zeros in $\mathbb{F}_p$. We write $f = f_1^{\alpha_1} \cdots f_r^{\alpha_r}$ for the factorization into irreducible factors in $\overline{\mathbb{F}}_p$. We suppose that the degrees $\alpha_1, \ldots, \alpha_r$ satisfy*

$$(d, \alpha_1, \ldots, \alpha_r) = 1. \tag{1.5}$$

Then the cardinality of V_f is

$$\left| |V_f| - \frac{(p-m)}{d} \right| < 9k\sqrt{p}\log p. \tag{1.6}$$

For these polynomials we will show that the well-distribution measure is small:

Theorem 1.2. *Under the same conditions as in Theorem 1.1, we have*

$$W(V_f, p) \le 20k\sqrt{p}\log p. \tag{1.7}$$

The main tool in the proofs of Theorem 1.1 and Theorem 1.2 is an estimate (Lemma 2.1 below) on short character sums with polynomial argument like $\sum_{X<n\le X+Y} \chi(g(n))$. To apply this result we have to check that the involved polynomial g is not a d-th power, more precisely we have to check a condition on the degrees of the irreducible factors of g.

When we compute the correlation, the corresponding polynomial g is complicated thus we cannot apply Lemma 2.1 directly.

Goubin, Mauduit and Sárközy [5] solved this problem by giving a condition for k, p and the order ℓ of correlation. They defined the notion of *admissible* triple (k, ℓ, p). They proved that every triple (k, ℓ, p) verifying (i), (ii) or (iii) is admissible.

In this paper we will follow this strategy. We will define the notion of d-admissible triple (k, ℓ, p) which is a generalization of the admissible triples of [5]. This new definition is a little more complicated, but the cost of this complication is not too high; we will see that the different types of admissible triples found by [5] are also d-admissible if d is a prime factor of $p-1$.

Definition 1.3. If $k, \ell, d \in \mathbb{N}$ and p is a prime number such that $d|p-1$, the triple (k, ℓ, p) is said to be d-admissible if, for all $\mathcal{A} \subset \mathbb{F}_p$, $|\mathcal{A}| \le k$, $r \le \ell$, $0 \le d_1 < \cdots < d_r < p$, $1 \le D_i < d$ for $i = 1, \ldots, r$ and $(D_1, \ldots, D_r) = 1$, we have :

$$(d, \alpha(1), \ldots, \alpha(p)) = 1,$$

where $\alpha(b)$ are the weights defined by

$$\alpha(b) = \sum_{\substack{a \in \mathcal{A} \\ 1 \le j \le r \\ a + d_j \equiv b \pmod p}} D_j.$$

The definition of admissible triples of [5] corresponds to $d = 2$. (And in this case the D_j are equal to 1.)

Under a d-admissibility condition, we will prove that the subset V_f has strong PR properties.

Theorem 1.4. *Let $f \in \mathbb{F}_p[X]$ of degree k with no multiple roots in $\bar{\mathbb{F}}_p$. If (k, ℓ, p) is a d-admissible triple, then*

$$C_\ell(V_f, p) \le \ell k\Big(1 + \frac{20k\log p}{\sqrt{p}}\Big)^\ell + 9\Big(\frac{d-1}{d}\Big)^\ell\Big(1 + \frac{10k\log p}{\sqrt{p}}\Big)^\ell \ell k\sqrt{p}\log p. \quad (1.8)$$

In Theorem 4.1 in section 4, we will give sufficient conditions for a triple being d-admissible. Combining Theorem 1.4 and Theorem 4.1 we get:

Theorem 1.5. *Let $k, \ell \in \mathbb{N}$ be such that one of the following conditions is satisfied*

(i) *$\ell = 2$;*

(ii) *d is a prime divisor of $p-1$, and $(4\ell)^k < p$;*

(iii) *the polynomial $X^{p-1}+\cdots+X+1$ is irreducible in $\mathbb{F}_d[X]$ and $\max(k, \ell) < p$.*

Let $f \in \mathbb{F}_p[X]$ be a polynomial of degree k with no multiple roots in $\bar{\mathbb{F}}_p$. Then we have

$$C_\ell(V_f, p) \le \ell k\Big(1 + \frac{20k\log p}{\sqrt{p}}\Big)^\ell + 9\Big(\frac{d-1}{d}\Big)^\ell\Big(1 + \frac{10k\log p}{\sqrt{p}}\Big)^\ell \ell k\sqrt{p}\log p. \quad (1.9)$$

Note that Johnsen [7] and Choi and Zaharescu [2] (in particular, see Lemma 2) studied the distribution of the d-powers in finite fields $\mathbb{F}_q$. The $f(X) = X$ special case of our results covers the distribution of the d-powers in $\mathbb{F}_p$, thus some of ours results (in particular (1.9) above) are extensions and generalizations of the $q = p$ special case of some results in [7] and [2].

In cryptography we need to know how rich is the structure of a large family of pseudorandom subsets. In particular we would like to be sure that a subset is not determined by just a few elements of it. To study this problem, Ahlswede, Khachatrian, Mauduit and Sárközy [1] defined a complexity measure for families of pseudorandom sequences. They applied it to the family of the sequences $\left(\frac{f(n)}{p}\right)$ defined in [5] and they showed that the complexity of this family is large.

Mosaki and the authors [3] introduced a similar definition of complexity of families subsets. They applied it to the families of subsets defined by (1.2) and proved that the complexity is large.

In the last section of this paper we will use the definition of complexity given in [3] to study the complexity of the family of subsets V_f for all polynomial f of degree at most k and with no multiple roots in $\bar{\mathbb{F}}_p$.

Definition 1.6. Let $\mathcal{F}$ be a family subsets of $\{1,\dots,N\}$. Then the family complexity $K(\mathcal{F})$ of $\mathcal{F}$ is defined as the greatest $k \in \mathbb{N}$ such that for every $\mathcal{A} \subset \{1,\dots,N\}$ with $|\mathcal{A}| = k$ and every partition $\mathcal{A} = \mathcal{B} \cup \mathcal{C}$, $\mathcal{B} \cap \mathcal{C} = \emptyset$, there is an $\mathcal{R} \in \mathcal{F}$ such that $\mathcal{R} \cap \mathcal{A} = \mathcal{B}$.

(This implies that writing $\mathcal{R}^c = \{1,\dots,N\} \setminus \mathcal{R}$, we have $\mathcal{R}^c \cap \mathcal{A} = \mathcal{C}$.) We will prove that our family of pseudorandom subsets is rich.

Theorem 1.7. *Let $\mathcal{F}(k)$ denote the family of the subsets V_f formed with all polynomials $f \in \mathbb{Z}[X]$ of degree $\le k$ and with no multiple roots in $\bar{\mathbb{F}}_p$. For all $t \le k+1$ such that*

$$(2t-3)p^{t-1} < \Big(\frac{p-1}{d}\Big)^t, \tag{1.10}$$

the complexity of $\mathcal{F}(k)$ is $\ge t$.

To prove this theorem, we use the Lagrange interpolation polynomials. The difficulty is that we have to find some suitable polynomials with no multiple roots. In [1], the authors solved this problem by showing that any polynomial $f \in \mathbb{F}_p[X]$ may be written in the form $f = h^2 f^*$ with $h, f^* \in \mathbb{F}_p[X]$ and such that f^* has no multiple roots in $\bar{\mathbb{F}}_p$; so that $\Big(\frac{f(n)}{p}\Big) = \Big(\frac{f^*(n)}{p}\Big)$.

This argument doesn't work in our problem. We will use a different approach. We will apply the Lagrange interpolation method to many suitable sequences and we will use an algebraic lemma of Ore [11] (see also [8] Theorem 6.13 p. 275) to prove that at least one of these polynomials has no multiple roots.

In [3] we already met this difficulty of finding interpolation polynomial without multiplicative roots. We worked in a general situation and the result we obtained by combinatorial arguments could be applied here. It would give Theorem 1.7 with (1.10) replaced by $t \le \frac{p-1}{d}\big(1 - \frac{p-1}{dp}\big)$.

We decided to present here our alternative proof because it is a little shorter and gives a different point of view.

2. The cardinality of V_f and the distribution measure: proofs of Theorem 1.1 and Theorem 1.2

The main tool of the proof of these two theorems is the following upper bound for short character sums.

Lemma 2.1. *Suppose that p is a prime number, χ is a non-principal character modulo p of order d (so that $d|p-1$), $f(x) \in \mathbb{F}_p[X]$ has the factorization $f(X) = b(X-x_1)^{d_1}\cdots(X-x_s)^{d_s}$ (where $x_i \neq x_j$ for $i \neq j$) in $\overline{\mathbb{F}}_p$ with*

$$(d, d_1, \dots, d_s) = 1. \tag{2.1}$$

Let X, Y be real numbers with $0 < Y \leq p$. Then

$$\left| \sum_{X<n\leq X+Y} \chi(f(n)) \right| < 9s\sqrt{p}\log p.$$

This is Lemma 2 in [12], it is a slightly modified form of Theorem 2 in [9], and it was derived from A. Weil's theorem [13].

We denote by χ_0 the principal character over $\mathbb{F}_p$. By the orthogonality of characters of order d, we have for $x \in \mathbb{F}_p$:

$$\sum_{\chi^d=\chi_0} \chi(x) = \begin{cases} d & \text{if } \exists\ y \in \mathbb{F}_p \setminus \{0\} : x = y^d \\ 0 & \text{otherwise.} \end{cases} \tag{2.2}$$

By Lemma 2.1 we have

$$\begin{aligned} \text{card } V_f = \sum_{x\in V} 1 &= \frac{1}{d} \sum_{\chi^d=\chi_0} \sum_{x\in\mathbb{F}_p} \chi(f(x)) \\ &= \frac{1}{d} \sum_{x\in\mathbb{F}_p} \chi_0(f(x)) + \frac{1}{d} \sum_{\substack{\chi^d=\chi_0 \\ \chi\neq\chi_0}} \sum_{x\in\mathbb{F}_p} \chi(f(x)) \\ &= \frac{p-m}{d} + O(k\sqrt{p}\log p), \end{aligned}$$

where the implicit constant in the Landau symbol is in absolute value less than 9.

This proves Theorem 1.1.

The basic ideas of the proof of Theorem 1.2 are the same as the proof of (3.5) of Theorem 3.1 in [4] but in some steps there are differences arising from the fact that here f is not assumed to be a permutational polynomial. The associated sequence $\{e_n\}_{1\leq n\leq p}$ defined by (1.1) is

$$e_n = \begin{cases} 1-\alpha & \text{if } n \in V_f \\ -\alpha & \text{if } n \notin V_f, \end{cases}$$

with

$$\alpha = \frac{\text{card } V_f}{p} = \frac{1}{d} - \frac{m}{dp} + 9\theta k p^{-1/2}\log p, \tag{2.3}$$

with some θ satisfying $|\theta| \leq 1$.

We also define

$$\beta := \frac{1}{d} - \alpha \text{ so that } |\beta| \le \frac{m}{dp} + \frac{9k\log p}{\sqrt{p}}.$$

Let a, b, t be positive integers such that $b + a(t-1) \le p$. Using the orthogonality formula (2.2) we proved in [4] the equality

$$\sum_{j=0}^{t-1} e_{aj+b} = \frac{1}{d} \sum_{\chi^d=\chi_0} \sum_{0\le j\le t-1} \chi(f(aj+b)) - \alpha t.$$

The contribution of χ_0 is t/d minus the number of $0 \le j \le t-1$ such that $f(aj+t) \equiv 0 \pmod p$. There are at most m such integers j, thus we have

$$\Big|\frac{1}{d}\sum_{j=0}^{t-1} \chi_0(f(aj+b)) - \alpha t\Big| \le \frac{m}{d} + t\beta \le 11k\sqrt{p}\log p. \tag{2.4}$$

It remains to compute the contribution of the non-trivial characters. By Lemma 2.1 we have

$$\Big|\frac{1}{d} \sum_{\substack{\chi^d=\chi_0\\ \chi\ne\chi_0}} \sum_{0\le j\le t-1} \chi(f(aj+b))\Big| \le \frac{1}{d} \sum_{\substack{\chi^d=\chi_0\\ \chi\ne\chi_0}} 9k\sqrt{p}\log p \le 9k\frac{(d-1)}{d}\sqrt{p}\log p. \tag{2.5}$$

The upper bounds (2.4) and (2.5) give (1.7).

3. The correlation: proof of Theorem 1.4

In this section we will denote by m the number of zeros of f in $\mathbb{F}_p$. The beginning of the computation of the correlation is nearly the same as in [4].

Let $\ell \ge 2$. We have to compute

$$C_\ell(V,p) = \max_{M,D} \Big|\sum_{n=1}^{M} e_{n+d_1}\cdots e_{n+d_\ell}\Big|,$$

where the maximum is over all M and $D = (d_1, \dots, d_\ell)$, $0 \le d_1 \le \ldots \le d_\ell$ such that $M + d_\ell \le p$.

By (2.2)

$$e_{n+d_1}\cdots e_{n+d_\ell} =$$
$$= \prod_{j=1}^{\ell}\Big[(1-\alpha)\frac{1}{d}\sum_{\chi^d=\chi_0}\chi(f(n+d_j)) - \alpha\Big(1-\frac{1}{d}\sum_{\chi^d=\chi_0}\chi(f(n+d_j))\Big)\Big].$$
$$= \prod_{j=1}^{\ell}\Big[\frac{1}{d}\sum_{\chi^d=\chi_0}\chi(f(n+d_j)) - \alpha\Big].$$

In the following step we will compute the perturbations arising from the integers n such that $\prod_{1\le j\le \ell} f(n+d_j)\equiv 0 \pmod p$. If $f(n+d_j)\not\equiv 0 \pmod p$ for all $1\le j\le \ell$ then

$$e_{n+d_1}\cdots e_{n+d_\ell} = \prod_{j=1}^{\ell}\Big[\frac{1}{d}\sum_{\substack{\chi^d=\chi_0\\ \chi\neq\chi_0}}\chi(f(n+d_j)) + \beta\Big]$$
$$= \frac{1}{d^\ell}\prod_{j=1}^{\ell}\Big[\sum_{\substack{\chi^d=\chi_0\\ \chi\neq\chi_0}}\chi(f(n+d_j)) + d\beta\Big].$$

If there exists some j, $1\le j\le \ell$ such that $f(n+d_j)\equiv 0 \pmod p$ then

$$e_{n+d_j} = \begin{cases} -\alpha & \text{if } f(n+d_j)\equiv 0 \pmod p,\\ \frac{1}{d}\Big[\sum_{\substack{\chi^d=\chi_0\\ \chi\neq\chi_0}}\chi(f(n+d_j)) + d\beta\Big] & \text{otherwise.}\end{cases}$$

In this last case we have

$$\Big|e_{n+d_1}\cdots e_{n+d_\ell} - \frac{1}{d^\ell}\prod_{j=1}^{\ell}\Big[\sum_{\substack{\chi^d=\chi_0\\ \chi\neq\chi_0}}\chi(f(n+d_j)) + d\beta\Big]\Big| =$$
$$= \prod_{\substack{1\le j\le \ell\\ f(n+d_j)\not\equiv 0 \pmod p}}\Big|\frac{1}{d}\sum_{\substack{\chi^d=\chi_0\\ \chi\neq\chi_0}}\chi(f(n+d_j)) + \beta\Big| \prod_{\substack{1\le j\le \ell\\ f(n+d_j)\equiv 0 \pmod p}} |-\alpha+\beta|.$$

We have, for all $n\in\mathbb{N}$,

$$\Big|\frac{1}{d}\sum_{\substack{\chi^d=\chi_0\\ \chi\neq\chi_0}}\chi(f(n+d_j)) + \beta\Big| \le 1-\frac{1}{d}+|\beta| \le 1+\frac{m}{dp}+\frac{9k\log p}{\sqrt{p}},$$

and

$$|-\alpha+\beta| = |-\frac{1}{d}+2\beta| \le \frac{1}{d}+\frac{2m}{dp}+\frac{18k\log p}{\sqrt{p}}.$$

Finally we have

$$\Big|e_{n+d_1}\cdots e_{n+d_\ell}-\frac{1}{d^\ell}\prod_{j=1}^{\ell}\Big[\sum_{\substack{\chi^d=\chi_0\\ \chi\neq\chi_0}}\chi(f(n+d_j))+d\beta\Big]\Big|\le\left(1+\frac{20k\log p}{\sqrt{p}}\right)^\ell.$$

For any $1\le j\le\ell$, there are m integers in $\mathbb{F}_p$ such that $f(n+d_j)\equiv 0 \pmod p$. Thus there are at most ℓm integers $n\in\{1,\dots,p\}$ such that $\prod_{j=1}^{\ell} f(n+d_j)\equiv 0 \pmod p$.

So we have

$$\Big|\sum_{n=1}^{M}e_{n+d_1}\cdots e_{n+d_\ell}-Z\Big|\le\ell m\left(1+\frac{20k\log p}{\sqrt{p}}\right)^\ell, \tag{3.1}$$

with

$$Z=\frac{1}{d^\ell}\sum_{n=1}^{M}\prod_{j=1}^{\ell}\Big[\sum_{\substack{\chi^d=\chi_0\\ \chi\neq\chi_0}}\chi(f(n+d_j))+\frac{1}{p}\Big]. \tag{3.2}$$

In the rest of the proof we will give an upper bound for Z. We compute this product :

$Z=$

$$\frac{1}{d^\ell}\sum_{r=0}^{\ell}\frac{1}{p^{\ell-r}}\sum_{1\le j_1<\dots<j_r\le\ell}\ \sum_{\substack{\chi_{j_1}\neq\chi_0\\ \chi_{j_1}^d=\chi_0}}\cdots\sum_{\substack{\chi_{j_r}\neq\chi_0\\ \chi_{j_r}^d=\chi_0}}\sum_{n=1}^{M}\chi_{j_1}(f(n+d_{j_1}))\cdots\chi_{j_r}(f(n+d_{j_r})).$$

We have to obtain an upper bound for the innermost sums of type

$$\sum_{n=1}^{M}\chi_{j_1}(f(n+d_{j_1}))\cdots\chi_{j_r}(f(n+d_{j_r})),$$

where $0\le r\le\ell$ and $\chi_{j_1},\dots\chi_{j_r}$ are non-principal characters of order dividing d.

A simpler form of the following step was already done in [4] and it is in fact an adaptation of some ideas of [12]. We rewrite here the argument because we need precise information on the order of the characters.

Since $\mathbb{F}_p^*$ is cyclical, each χ_{j_h} is of form $\chi_{j_h}=\chi^{\delta_h}$, where χ is a character of order $p-1$ and $1\le\delta_h<p-1$. Let $\delta=(\delta_1,\dots,\delta_r)$, and $\delta_i=\delta D_i$ for $1\le i\le r$; we have $(D_1,\dots,D_r)=1$. Since $\chi_{j_i}^d=\chi_0$, we have $d\delta_i\equiv 0 \pmod{p-1}$.

We write $\chi^*=\chi^\delta$. It is proved in (16) in [12] that $\chi^*\neq\chi_0$, more precisely, the order D of χ^* is $D=(p-1)/(p-1,\delta)$ and in our case, $D|d$. Furthermore, there exists μ such that $\delta=\mu(p-1)/d$. We have $\delta_i=\mu(p-1)D_i/d<p-1$ for all $1\le i\le r$. Thus $1\le D_i<d$, for all $1\le i\le r$.

The computations in p. 383 of [12] yield

$$\sum_{n=1}^{M} \chi_{j_1}(f(n+d_{j_1}))\cdots\chi_{j_r}(f(n+d_{j_r})) = \sum_{n=0}^{M-1} \chi^*\big(f(n+d_{j_1})^{D_1}\cdots f(n+d_{j_r})^{D_r}\big). \tag{3.3}$$

We write $f(X) = bg(X)$, $b \in \mathbb{F}_p$, $g \in \mathbb{F}_p[X]$ a unitary polynomial. Let $B = \prod_{i=1}^{r} b^{D_i}$. The roots of f and g are the same. The sum (3.3) becomes :

$$\sum_{n=1}^{M} \chi_{j_1}(f(n+d_{j_1}))\cdots\chi_{j_r}(f(n+d_{j_r})) = \\ = \chi^*(B)\sum_{n=0}^{M-1} \chi^*\big(g(n+d_{j_1})^{D_1}\cdots g(n+d_{j_r})^{D_r}\big).$$

We would like to apply Lemma 2.1 to the polynomial

$$F(X) = \prod_{\lambda=1}^{r} g(n+d_{j_\lambda})^{D_\lambda}.$$

We have to check whether F satisfies condition (2.1). For this, we follow the proof of Lemma 2 in [5] to show that if (k,ℓ,p) is d-admissible then we can apply Lemma 2.1. Like in the proof of this Lemma 2 in [5] we will say that two polynomials $\varphi, \psi \in \mathbb{F}_p[X]$ are equivalent and write $\varphi \sim \psi$ if there exists $a \in \mathbb{F}_p$ such that $\psi(X) = \varphi(X+a)$. Let $f_1, \ldots, f_t$ be the irreducible factors of g over $\mathbb{F}_p$. This irreducible factors are distinct because we have supposed that f has no multiple roots. We group these factors so that in each group the equivalent factors are collected.

Let $\varphi_1, \ldots, \varphi_t$ be the representants of the each equivalence classes of factors. Then g can be written as

$$g(X) = \prod_{i=1}^{t}\prod_{j=1}^{s_i} \varphi_i(X + a_{i,j}),$$

s_i being the number of irreducible factors of g equivalent to φ_i $(1 \le i \le t)$ so that $\sum_{i=1}^{t} s_i = k$.

Since f has no multiple roots, condition (2.1) in Lemma 2.1 is not satisfied if and only if there exists q, a prime factor of d, dividing the multiplicity of each root of F. Let u be a zero of F. There exists one and only one $1 \le i \le t$ such that u belongs to the equivalence class of φ_i. There exists $c \in \mathbb{F}_p$ such that u is

a zero of the polynomial $\varphi_i(X+c)$. Thus the multiplicity of u is

$$\sum_{j=1}^{s_i} \sum_{\substack{1\le\lambda\le r\\ a_{i,j}+d_\lambda\equiv c \pmod p}} D_\lambda.$$

Finally we cannot apply Lemma 2.1 if and only if, there exists a prime $q|d$ such that for all $c\in\mathbb{F}_p$ and $1\le i\le t$

$$\sum_{j=1}^{s_i} \sum_{\substack{1\le\lambda\le r\\ a_{i,j}+d_\lambda\equiv c \pmod p}} D_\lambda \equiv 0 \pmod q.$$

This does not hold if (k,ℓ,p) is d-admissible.

Thus we may apply Lemma 2.1 and end the computation as in [4] to obtain the result.

4. Admissible triples

In this section we will generalize the ideas of [5] to provide sufficient conditions for d-admissibility.

Theorem 4.1. (i) *For every prime p and d such that $d|p-1$, and for all $k\in\mathbb{N}$ the triple $(k,2,p)$ is d-admissible.*

(ii) *If $k,\ell\in\mathbb{N}^*$ and d is a prime divisor of $p-1$, such that $(4\ell)^k<p$, then (k,ℓ,p) is d-admissible.*

(iii) *If the polynomial $X^{p-1}+\cdots+X+1$ is irreducible in $\mathbb{F}_q[X]$ for all prime factors q of d then (k,ℓ,p) is d-admissible if* $\max(k,\ell)<p$.

Proof. (i) We suppose that the result is false. There exist $\mathcal{A}\subset\mathbb{F}_p$ with cardinality $\le k$, $1\le d_1<d_2\le p$, $1\le D_1,D_2<d$ with $(D_1,D_2)=1$ and a prime factor q of d such that for all $c\in\mathbb{F}_p$ we have $\alpha(c)\equiv 0 \pmod q$.

If $q|D_1$ then $q\nmid D_2$ since $(D_1,D_2)=1$. In this case, for all $c\in\mathbb{F}_p$,

$$\alpha(c)\equiv\begin{cases} D_2 & \text{if } c\in\mathcal{A}+d_2\\ 0 & \text{otherwise}\end{cases} \pmod q.$$

In particular $\alpha(c)\not\equiv 0 \pmod q$ if $c\in\mathcal{A}+d_2$. Thus $q\nmid D_1D_2$.

We write $b=d_2-d_1$. Since $q\nmid D_1$, every element of $\mathcal{A}+d_1$ must belong to $\mathcal{A}+d_1+b$. It follows that $\mathcal{A}+d_1=\mathcal{A}+d_1+b$. In [5] it is shown that this implies $\mathcal{A}+d_1=\mathbb{F}_p$.

(ii) Goubin, Mauduit and Sárközy proved that if (k, ℓ, p) satisfies the condition (ii), then there exists $c \in \mathbb{F}_p$ such that the equation $a + d_j \equiv c \pmod p$ has only one solution. For this c, we have $\alpha(c) = D_j \not\equiv 0 \pmod p$.

It remains to prove (iii). We adapt the proof of Theorem 3 of [5]. Let q be a prime factor of d.

We consider the polynomial in $\mathbb{F}_q[X]$ defined by $P(X) = \sum_{c\in\mathcal{C}} D_c X^{s(c)}$ where $s(c)$ denotes the least non negative element of the residue class c modulo p.

For $0 \le u < p$ we have $u+s(c) \equiv s(c+u) \pmod p$. More precisely, $s(u+c) = u + s(c)$ or $u + s(c) - p$. Since $X^p \equiv 1 \pmod{X^p - 1}$, we have $X^u P(X) \equiv \sum_{c\in\mathcal{C}} D_c X^{s(u+c)} \pmod{X^p - 1}$.

We obtain :

$$\sum_{a\in\mathcal{A}} X^a \sum_{j=1}^{r} D_j X^{s(d_j)} \equiv \sum_{\substack{1\le j\le r\\ a\in\mathcal{A}}} D_j X^{s(a+d_j)} \pmod{X^p - 1}$$
$$\equiv \sum_{b\in\mathbb{F}_p} X^{s(b)}\alpha(b) \pmod{X^p - 1}.$$

We have $q|(d, \alpha(b) : b \in \mathbb{Z})$ if and only $\sum_{b\in\mathbb{F}_p} \alpha(b)X^b = 0$ in $\mathbb{F}_q[X]$. When P is a polynomial of degree less than p we see that $P = 0$ if and only if $P \equiv 0 \pmod{X^p - 1}$.

We suppose that $1 + X + \cdots + X^{p-1}$ is irreducible in $\mathbb{F}_q[X]$ and that there exists $\mathcal{A} \subset \mathbb{F}_p$, $|\mathcal{A}| \le k$, $\mathcal{D} = (d_1, \ldots, d_r)$ with $0 \le d_1 < \ldots < d_r < p$, $1 \le D_1, \ldots, D_r < d$ such that $q|(\alpha(b) : b \in \mathbb{F}_p)$. We write $P_{\mathcal{A}}(X) = \sum_{a\in\mathcal{A}} X^a$ and $P_{\mathcal{D}}(X) = \sum_{j=1}^{r} D_j X^{s(d_j)}$. If q divides all the weights $\alpha(b)$ then $(X^p-1)|P_{\mathcal{A}}(X)P_{\mathcal{D}}(X)$. In particular $\Phi_p(X) = 1+X+\cdots+X^{p-1}|P_{\mathcal{A}}(X)P_{\mathcal{D}}(X)$. Since Φ_p is irreducible in $\mathbb{F}_q[X]$ this implies $\Phi_p|P_{\mathcal{A}}$ or $\Phi_p|P_{\mathcal{D}}$. In the first case we should have $\mathcal{A} = \mathbb{F}_p$ and in the second case $\ell = p$. These two cases are impossible. This ends the proof of Theorem 4.1. □

5. The complexity of the family V_f : proof of Theorem 1.7

Let $\mathcal{A} \subset \mathbb{F}_p$ be a set of cardinality $t \le k + 1$. If $t = 1$, $\mathcal{A} = \{a\}$ we take the polynomial $f = 1$ if we impose that $f(a)$ is a d-power residue and $f = 0$ otherwise.

Now we suppose that $t \ge 2$. Let $\mathcal{B} \cup \mathcal{C} = \mathcal{A}$ be a partition of $\mathcal{A}$. We write

$$\mathcal{A} = \{a_1, \ldots, a_r, a_{r+1}, \ldots, a_t\},$$

with $\mathcal{B} = \{a_1, \dots, a_r\}$ and $\mathcal{C} = \{a_{r+1}, \dots, a_t\}$. We have to find a polynomial f of degree $\le k$ with no multiple roots such that for $1 \le i \le r$, $f(a_i)$ is a non-zero d-power residue and for $r+1 \le i \le t$, $f(a_i)$ is not a non-zero d-power residue.

Let G denote the set of the different non-zero d-power residues and G^c its complementary in $\mathbb{F}_p$.

The set G is $G = \varphi(\mathbb{F}_p^*)$ where φ is the homomorphism $\mathbb{F}_p^* \to \mathbb{F}_p^*$ defined by $\varphi(x) = x^d$.

Thus $|G| = \operatorname{card} \mathbb{F}_p^* / |\ker \varphi| = (p-1)/d$ and $|G^c| = p - (p-1)/d$.

Let $\mathbf{u} = (u_1, \dots, u_t) \in \mathbb{F}_p^t$ with $u_1, \dots, u_r \in G$ and $u_{r+1}, \dots, u_t \in G^c$. There exists a polynomial $f \in \mathbb{Z}[X]$ of degree $\le t-1 \le k$ such that $f(a_i) = u_i$ for all $1 \le i \le t$.

Using the Lagrange interpolation formula this polynomial is

$$f_{\mathbf{u}}(X) = \sum_{i=1}^{t} u_i \prod_{\substack{1 \le j \le t \\ j \ne i}} (X - a_j)\overline{(a_i - a_j)}, \tag{5.1}$$

where $\overline{(a_i - a_j)}$ denotes the multiplicative inverse of $a_i - a_j$ in $\mathbb{F}_p$.

The difficulty is that these polynomials $f_{\mathbf{u}}$ may have multiple roots in $\bar{\mathbb{F}}_p$.

Anyway in this way we find $((p-1)/d)^r (p - (p-1)/d)^{t-r}$ polynomials $f_{\mathbf{u}}$ such that

$$\begin{cases} a_i \in V_{f_{\mathbf{u}}} & \text{for } 1 \le i \le r \\ a_i \notin V_{f_{\mathbf{u}}} & \text{for } r+1 \le i \le t. \end{cases}$$

We will show that at least one of these polynomials has no multiple roots.

For $\mathbf{u} = (u_1, \dots, u_t) \in G^r \times (G^c)^{t-r}$, let $D(\mathbf{u})$ denote the discriminant of the interpolation polynomial $f_{\mathbf{u}}$ defined by (5.1). Its degree is at most $2t-3$. This polynomial belongs to $\mathbb{F}_p[u_1, \dots, u_t]$.

We will use the following theorem of Ore [11]. A proof may be also found in the book of Lidl and Niederreiter [8] theorem 6.13 p. 275.

Theorem 5.1. *Let $f \in \mathbb{F}_q[x_1, \dots, x_n]$ a polynomial with degree $d \ge 0$. Then the equation $f(x_1, \dots, x_n) = 0$ has at most dq^{n-1} solutions in $\mathbb{F}_q^n$.*

The polynomial D is not identically equal to 0. If we take

$$u_1 = \prod_{j=2}^{t}(a_1 - a_j), \quad u_2 = \ldots = u_t = 0,$$

then $f_{\mathbf{u}} = \prod_{j=2}^{t}(X - a_j)$, has no multiple roots, thus $D(\mathbf{u}) \neq 0$.

By Theorem 5.1 and since (1.10) is assumed,

$$|\{\mathbf{u} \in \mathbb{F}_p : D(\mathbf{u}) \equiv 0 \pmod p\}| \leq (2t-3)p^{t-1} \leq |G|^t.$$

Thus there exists $\mathbf{u} \in \mathbb{F}_p$ such that $D(\mathbf{u}) \not\equiv 0 \pmod p$, *i. e.*, such that $f_{\mathbf{u}}$ has no multiple roots.

Remark. We did not really use the special nature of the set G. In the above argument we could replace G by an arbitrary subset of $\mathbb{F}_p$. It would be possible to find a interpolation polynomial $f_{\mathbf{u}}$ with no multiple roots, if

$$(2t-3)p^{t-1} \leq (\min(|G|, p - |G|))^t.$$

REFERENCES

[1] AHLSWEDE, R. – KHACHATRIAN, L.H. – MAUDUIT, CH. – SÁRKÖZY, A.: *A complexity measure for families of binary sequences*, Periodica Math. Hungar. **46** (2003), 107-118.

[2] CHOI, G. – ZAHARESCU, A.: *Additive patterns in a multiplicative group in a finite field*, Manuscripta Math. **111** (2003), 187-194.

[3] DARTYGE, C. – MOSAKI, E. – SÁRKÖZY, A.: *On large families of subsets of the set of the integers not exceeding N*, to appear.

[4] DARTYGE, C. – SÁRKÖZY, A.: *On pseudo-random subsets of the set of the integers not exceeding N*, Period. Math. Hungar. **54** (2007), no. 2, 183–200.

[5] GOUBIN, L. – MAUDUIT, CH. – SÁRKÖZY, A.: *Construction of large families of pseudorandom binary sequences*, J. Number Theory **106** (2004), 56-69.

[6] HUBERT, P. – SÁRKÖZY, A.: *On p-pseudorandom binary sequences*, Periodica Math. Hungar. **49** (2004), 73-91.

[7] JOHNSEN, J.: *On the distribution of powers in finite fields*, J. Reine Angew. Math. **25** (1971), 10-19.

[8] LIDL, R. – NIEDERREITER, H.: *Finite Fields*, Encyclopedia of Mathematics and its Applications, vol. **20**, Addison-Wesley Publishing Company, Reading, MA, 1983.

[9] MAUDUIT, CH. – SÁRKÖZY, A.: *On finite pseudorandom binary sequences, I. Measure of pseudorandomness, the Legendre symbol*, Acta Arith. **82** (1997), no. 4, 365-377.

[10] MAUDUIT, CH. – SÁRKÖZY, A.: *Construction of pseudorandom binary sequences by using the multiplicative inverse*, Acta Math. Hungar. **108** (2005), 239-252.

[11] ORE, Ö.: *Über höhere Kongruenzen*, Norsk Mat. Forenings Skrifter Ser I, no. **7**, (1922), pp. 15.

[12] SÁRKÖZY, A.: *A finite pseudorandom binary sequence*, Studia Sci. Math. Hungar. **38** (2001), 377-384.
[13] WEIL, A.: *Sur les courbes algébriques et les variétés qui s'en déduisent*, Publ. Inst. Math. Univ. Strasbourg **7** (1945), Hermann, Paris, 1948.

Received Mart 30, 2007
Accepted October 9, 2007

Cécile Dartyge
Institut Élie Cartan
Université Henri Poincaré–Nancy 1
54506 Vandœuvre Cedex
FRANCE
E-mail: dartyge@iecn.u-nancy.fr

András Sárközy
Department of Algebra and Number Theory
Eötvös Loránd University
H-1518 Budapest
HUNGARY
E-mail: sarkozy@cs.elte.hu