ALGEBRAIC DEPENDENCE IN GENERATING FUNCTIONS AND EXPANSION COMPLEXITY

DOMINGO GÓMEZ-PÉREZ AND LÁSZLÓ MÉRAI

ABSTRACT. In 2012, Diem introduced a new figure of merit for cryptographic sequences called expansion complexity. Recently, a series of paper has been published for analysis of expansion complexity and for testing sequences in terms of this new measure of randomness. In this paper, we continue this analysis. First we study the expansion complexity in terms of the Gröbner basis of the underlying polynomial ideal. Next, we prove bounds on the expansion complexity for random sequences. Finally, we study the expansion complexity of sequences defined by differential equations, including the inversive generator.

1. INTRODUCTION

For a sequence $\mathcal{S} = (s_n)_{n=0}^{\infty}$ over the finite field $\mathbb{F}_q$ of q elements, we define its *generating function* $G(x)$ of $\mathcal{S}$ by

$$G(x) = \sum_{n=0}^{\infty} s_n x^n,$$

viewed as a formal power series over $\mathbb{F}_q$.

A sequence $\mathcal{S}$ is called *expansion* or *automatic sequence* if its generating function satisfies an algebraic equation

$$h(x, G(x)) = 0 \tag{1}$$

for some nonzero polynomial $h(x, y) \in \mathbb{F}_q[x, y]$. Clearly, the polynomials $h(x, y) \in \mathbb{F}_q[x, y]$ satisfying (1) form an ideal in $\mathbb{F}_q[x, y]$. This ideal is called the *defining ideal* and it is a principal ideal generated by an irreducible polynomial, see [3, Proposition 4].

Expansion sequences can be efficiently computed from a relatively short subsequence via the generating polynomial of its defining ideal [3, Section 5].

Proposition 1. *Let $\mathcal{S}$ be an expansion sequence and let $h(x, y)$ be the generating polynomial of its defining ideal. The sequence $\mathcal{S}$ is uniquely determined by $h(x, y)$ and its initial sequence of length $(\deg h)^2$. Moreover, $h(x, y)$ can be computed in polynomial time (in $\log q \cdot \deg h$) from an initial sequence of length $(\deg h)^2$.*

Based on Proposition 1, Diem [3] defined the Nth expansion complexity in the following way. For a positive integer N, the *Nth expansion complexity* $E_N = E_N(\mathcal{S})$ is $E_N = 0$ if $s_0 = \ldots = s_{N-1} = 0$ and otherwise the least total degree of a nonzero polynomial $h(x, y) \in \mathbb{F}_q[x, y]$ with

$$h(x, G(x)) \equiv 0 \mod x^N. \tag{2}$$

Key words and phrases. pseudorandom sequence, expansion complexity, Gröbner basis, inversive generator.

For recent results on expansion complexity we refer to [5, 7]. For example, it was pointed out in [5], that small expansion complexity does not imply high predictability in the sense of Proposition 1.

Example. Let $\mathcal{S}$ be a sequence over the finite field $\mathbb{F}_p$ ($p \geq 3$) with initial segment $\mathcal{S} = 000001\dots$ and generating function $G(x) \equiv x^5 \mod x^6$. Then its 6th expansion complexity is $E_6(\mathcal{S}) = 2$ realized by the polynomial $h(x, y) = x \cdot y$. However, the first 4 elements do not determine the whole initial segment with length 6.

In order to achieve the predictability of sequences in terms of Proposition 1, one needs to require that the polynomial $h(x, y)$ satisfying (2) is *irreducible*. This observation leads to the *i(rreducible)-expansion complexity* of a sequence. Accordingly, for a positive integer N, the *Nth i-expansion complexity* $E^*_N = E^*_N(\mathcal{S})$ is $E^*_N = 0$ if $s_0 = \ldots = s_{N-1} = 0$ and otherwise the least total degree of an irreducible polynomial $h(x, y) \in \mathbb{F}_q[x, y]$ with (2).

See [5] for more details for expansion and i-expansion complexity.

In this paper we first give bounds on the expansion and i-expansion complexity in terms of Gröbner basis of the ideal of polynomials (2) in Section 2. It also provides an alternative method to determine the expansion and i-expansion complexity of sequences compared to [3]. In Section 3 we study the typical value of expansion complexity for random sequences. Finally, in Section 4 we study the expansion complexity of sequences defined by differential equations. An example of such sequence is the so-called explicit inversive generator.

2. Expansion complexity and Gröbner bases

In this section we determine the expansion and i-expansion complexity of a sequence in terms of the Gröbner basis of its defining ideal.

2.1. A brief introduction to Gröbner bases.

In the following section, we give a brief introduction of Gröbner bases with special emphasis in properties. For a more complete introduction, we recommend to consult the introductory book of Eisenbud [4].

For vectors of integer components $\alpha = (\alpha_1, \alpha_2)$ define $|\alpha| = \alpha_1 + \alpha_2$. The *graded lexicographical ordering*, denoted by $<_{grlex}$, defined as $\alpha <_{grlex} \beta$ for vectors $\alpha = (\alpha_1, \alpha_2)$ and $\beta = (\beta_1, \beta_2)$ if $|\alpha| < |\beta|$ or $|\alpha| = |\beta|$ and $\alpha_1 < \beta_1$.

We will use the following notation: Let $C = \sum_{\alpha_1,\alpha_2} c_{\alpha_1,\alpha_2} x^{\alpha_1} y^{\alpha_2}$ be a nonzero polynomial with each $c_{\alpha_1,\alpha_2} \neq 0$ and $I \subset \mathbb{F}_q[x, y]$. Then,

(a) $LE(C) = leadexp(C)$ is the largest exponent vector α in C with respect to $<_{grlex}$.
(b) $LM(C)$ denotes the leading monomial of C so if $LE(C) = (\alpha_1, \alpha_2)$, then $LM(C) = x^{\alpha_1} y^{\alpha_2}$.
(c) $LC(C)$ denotes the coefficient of $LM(C)$. In other words, the so called leading term of C is $LC(C)LM(C)$.
(d) $LE(I) = \{LE(C) \mid 0 \neq C \in I\} \subseteq \mathbb{N}^2$. (Note that if $I = \{0\}$, then $LE(I) = \{\}$.)
(e) $LM(I) = \{LM(C) \mid 0 \neq C \in I\} = \{x^{\alpha_1} y^{\alpha_2} \mid (\alpha_1, \alpha_2) \in LE(I)\}$. (If $I = \{0\}$, then $LM(I) = \{\}$.)

Definition 1. *Let* $\mathcal{P} = \{P_1, \dots, P_\ell\} \subset \mathbb{F}_q[x, y]$ *and write* $I = \langle P_1, \dots, P_\ell \rangle$*.* $\mathcal{P}$ *is a* Gröbner basis *for* I *with respect to* $<_{grlex}$ *if* $\langle LM(P_1), \dots, LM(P_\ell)\rangle = \langle LM(I)\rangle$*. If* $LC(P_i) = 1$ *for* $i = 1, \dots, \ell$ *and* $LM(P_i)$ *does not divide any term of* P_j *for* $i \neq j$*, then* $\mathcal{P}$ *is a* reduced Gröbner basis *for* I *with respect to* $<_{grlex}$*.*

It is known that for any ideal I, there exists $\{P_1, \dots, P_\ell\}$ that is a reduced Gröbner basis with respect to $<_{grlex}$ and this basis is unique, apart from permutations of the elements.

2.2. **Main results on expansion complexity and Gröbner bases.** For a sequence $\mathcal{S} = (s_n)_{n=0}^{\infty}$ and $N \geq 1$, let $G_N(x) \in \mathbb{F}_q[x]$ be the generating function of the truncated sequence $(s_n)_{n=0}^{N-1}$, that is,

$$G_N(x) = \sum_{n=0}^{N-1} s_n x^n.$$

Clearly, $G(x) \equiv G_N(x) \mod x^N$.

The polynomials $h(x, y)$ satisfying (2) form an ideal I generated by $I = \langle y - G_N(x), x^N\rangle$. We prove the following result which makes a link between the expansion and i-expansion complexity and the Gröbner basis of I.

Theorem 1. *Given any sequence* $\mathcal{S}$ *over* $\mathbb{F}_q$ *let* $\mathcal{P} = \{P_1, \dots, P_\ell\}$ *be a reduced Gröbner basis for* $\langle y - G_N(x), x^N\rangle$ *with respect to* $<_{grlex}$*. Then*

$$E_N(\mathcal{S}) = \min\{|LE(P_1)|, \dots, |LE(P_\ell)|\},$$

and

$$E_N^*(\mathcal{S}) = \min\{|LE(P_i)| : P_i \in \mathcal{P} \text{ is irreducible}\}.$$

As a consequence, $E_N^*(\mathcal{S}) \leq \max\{|LE(P_1)|, \dots, |LE(P_\ell)|\}$*.*

Remark. If $Q_1, \dots, Q_r \in \mathbb{F}_q[x_1, \dots, x_n]$ are polynomials with degree $\deg Q_1, \dots, \deg Q_r \leq d$, then by using the F5 algorithm one can compute the Gröbner basis of the ideal $\langle Q_1, \dots, Q_r\rangle$ with respect to the graded lexicographical ordering in at most $O\left(rd\binom{n+d-1}{d}^{\omega}\right)$ field operation, where $\omega \leq 3$ is the exponent of matrix multiplication over $\mathbb{F}_q$, see [1]. Thus one can find the polynomials $P_1, \dots, P_\ell$ in Theorem 1, and compute the expansion and i-expansion complexity in at most $N^4 (\log q)^{O(1)}$ binary operation.

Proof. In order to prove the first part, observe that for any polynomial $h(x, y)$ satisfying (1) we have $LM(P_i) \leq_{grlex} LM(h)$ for some i, so $\deg P_i \leq \deg h(x, y)$.

For the second part, if $s_n = 0$ for $2 \leq n \leq N - 1$, then the result is immediate. Otherwise, we can reduce it to the case when $s_0 = s_1 = 0$. If the non-zero polynomial $h(x, y)$ satisfies (1), then $h_1(x, y) = h(x, y + s_0 + s_1 x)$ is a polynomial with $\deg h = \deg h_1$ and

$$h_1\left(x, \sum_{n=2}^{N-1} s_n x^n\right) = h\left(x, G_N(x)\right) \equiv 0 \mod x^N.$$

Now, we are going to show that one of the polynomials $P_1, \dots, P_\ell$ must be irreducible. Suppose contrary, that all the polynomials $P_1, \dots, P_\ell$ are reducible, so for all $i = 1, \dots, \ell$,

$$P_i(x, y) = R_i(x, y) T_i(x, y) \quad \text{for} \quad i = 1, \dots, \ell.$$

As P_i belongs to the reduced Gröbner basis of $\langle y - G_N(x), x^N\rangle$, we have $T_i(x, G(x)) \not\equiv 0 \mod x^N$ and so

$$R_i(x, G_N(x)) \equiv 0 \mod x.$$

Since $s_0 = s_1 = 0$, the smallest degree term of $G_N(x)$ has degree at least two, so we must have $R_i(x, y) \equiv 0 \mod x$. Similarly, we also get $T_i(x, y) \equiv 0 \mod x$. Write

$$R_i(x, y) = yq_1(x, y) + xr_1(x), \quad T_i(x, y) = yq_2(x, y) + xr_2(x).$$

Then $R_i(x, y)T_i(x, y) \in \langle y^2, yx, x^2\rangle$, so $I = \langle y - G_N(x), x^N\rangle = \langle R_1T_1, \ldots, R_\ell T_\ell\rangle \subset \langle y^2, yx, x^2\rangle$. However, $y - G_N(x) \notin \langle y^2, yx, x^2\rangle$, a contradiction. □

3. A probabilistic result

In this section we study the Nth expansion complexity for random sequences. We prove, that for such sequences the Nth expansion complexity is large.

Let μ_q be the uniform probability measure on $\mathbb{F}_q$ which assigns the measure $1/q$ to each element of $\mathbb{F}_q$. Let $\mathbb{F}_q^\infty$ be the sequence space over $\mathbb{F}_q$ and let μ_q^∞ be the complete product probability measure on $\mathbb{F}_q^\infty$ induced by μ_q. We say that a property of sequences $\mathcal{S} \in \mathbb{F}_q^\infty$ holds μ_q^∞-*almost everywhere* if it holds for a set of sequences $\mathcal{S}$ of μ_q^∞-measure 1. We may view such a property as a typical property of a random sequence over $\mathbb{F}_q$.

Theorem 2. *We have*

$$\liminf_{N\to\infty} \frac{E_N(\mathcal{S})}{(2N)^{1/3}} \geq 1 \qquad \mu_q^\infty\text{-almost everywhere.}$$

We remark, that Theorem 2 is the corrected form of [7, Theorem 4]. In [7], the authors used [3, Proposition 7], which requires the irreducibly property, and consequently, it holds for the i-expansion complexity instead for the expansion complexity, see [5, Theorem 2]. Theorem 2 gives now lower bound on the expansion complexity of typical sequences.

In order to prove the result, we need the following lemma, see [6, Section 0.22].

Lemma 2. *Let $f(x)$ be a positive decreasing function with*

$$\lim_{k\to\infty} \frac{e^k f(e^k)}{f(k)} < 1.$$

Then the series $\sum_k f(k)$ converges.

Proof of Theorem 2. First we fix an ε with $0 < \varepsilon < 1$ and we put

$$b_N = \lfloor (1-\varepsilon)(2N)^{1/3} \rfloor \qquad \text{for } N = 1, 2, \ldots.$$

Then

$$b_N \geq 1 \quad \text{and} \quad b_N \binom{b_N + 2}{2} \leq (1 - \varepsilon_0)N \tag{3}$$

for some positive ε_0 if N is large enough. For such N put

$$A_N = \{\mathcal{S} \in \mathbb{F}_q^\infty : E_N(\mathcal{S}) \leq b_N\}.$$

Since $E_N(\mathcal{S})$ depends only on the first N terms of $\mathcal{S}$, the measure $\mu_q^\infty(A_N)$ is given by

$$\mu_q^\infty(A_N) = q^{-N} \cdot \#\{\mathcal{S} \in \mathbb{F}_q^N : E_N(\mathcal{S}) \leq b_N\}. \tag{4}$$

If $\mathcal{S} \in \mathbb{F}_q^N$ is a sequence with $E_N(\mathcal{S}) \le b_N$, there is a polynomial $h(x,y)$ with degree at most b_N of form $h(x,y) = h_1(x,y)\cdots h_k(x,y)$ such that all the factors $h_i(x,y)$ are irreducible and

$$h_i(x, G(x)) \equiv 0 \mod x^{N_i}, \quad \text{and } N_1 + \cdots + N_k = N. \tag{5}$$

Clearly, we can assume, that $N_1 \ge N/b_N$. We estimate the cardinality of A_N by the number of such sequences that (2) holds modulo x^{N_1} with an irreducible polynomial $h_1(x,y)$ of degree at most b_N and $N_1 \ge N/b_N$. Write $\mathcal{S} = (\mathcal{S}_1, \mathcal{S}_2) \in \mathbb{F}_q^N$ with $\mathcal{S}_1 \in \mathbb{F}_q^{N_1}$ and $\mathcal{S}_2 \in \mathbb{F}_q^{N-N_1}$. For a fixed irreducible polynomial of degree d there are at most d choices for $\mathcal{S}_1$ (see [3, p. 332]) and q^{N-N_1} choices for $\mathcal{S}_2$. If two irreducible polynomials are constant multiples of each other, they define the same sequences $\mathcal{S}_1$.

Let a polynomial $f(x,y)$ of degree d be called *normalized* if in the coefficient vector $(a_0, a_1, \dots, a_d)$ of the homogeneous part with degree d of f, i.e.,

$$a_0x^d + a_1x^{d-1}y + \cdots + a_dy^d,$$

the first nonzero element is 1.

Let $I_2(d)$ be the number of *normalized* irreducible polynomials (with two variables) in $\mathbb{F}_q[x,y]$ of total degree d. Then by [2] we have

$$I_2(d) = \frac{1}{q-1}q^{\binom{d+2}{2}} + O\left(q^{\binom{d+1}{2}}\right).$$

Thus

$$\begin{aligned}\#\{\mathcal{S} \in \mathbb{F}_q^N : E_N(\mathcal{S}) \le b_N\} &\le \sum_{d_1 \le b_N} \sum_{n/b_N \le N_1 \le N} d_1 I_2(d_1) q^{N-N_1} \\ &\ll N b_n^2 q^{\binom{b_N+2}{2} + N - \frac{N}{b_N} - 1}\end{aligned}$$

By the choice of b_N, we have that it is at most $q^{N-\delta N^{2/3}}$ for some positive δ. Then by Lemma 2, $\sum_N \mu_q^\infty(A_N) < \sum_N q^{-\delta N^{2/3}} < \infty$. The Borel-Cantelli lemma shows that the set of all $\mathcal{S} \in \mathbb{F}_q^\infty$ for which $\mathcal{S} \in A_N$ for infinitely many N has μ_q^∞-measure 0. In other words, μ_q^∞-almost everywhere we have $\mathcal{S} \in A_N$ for at most finitely many N. It follows then from the definition of A_N that μ_q^∞-almost everywhere we have

$$E_N(\mathcal{S}) > b_N > (1-\varepsilon)(2N)^{1/3}$$

for all sufficiently large N. Therefore μ_q^∞-almost everywhere,

$$\liminf_{N\to\infty} \frac{E_N(\mathcal{S})}{(2N)^{1/3}} \ge (1-\varepsilon).$$

By applying this for $\varepsilon = 1/r$ with $r = 1, 2, \dots$ and noting that the intersection of countably many sets of μ_q^∞-measure 1 has again μ_q^∞-measure 1, we obtain the result of the theorem. □

4. Sequences defined by differential equations

In this section we study the expansion complexity of sequences characterized by the property that their generating function satisfies certain differential equation, namely sequences $\mathcal{S} = (s_n)$ with

$$f_{k+1}(x)G^{(k)}(x) + \cdots + f_2(x)G^{(1)}(x) + f_1(x)G(x) + f_0(x) = 0 \tag{6}$$

with polynomials $f_{k+1}(x), \dots, f_0(x) \in \mathbb{F}_q[x]$, where $G^{(i)}(x)$, $i \geq 0$, stands for the ith derivative of the generating function $G(x)$.

In Theorem 3 below, we give bounds on the Nth expansion complexity of sequences whose generating function satisfies a first order differential equation (6) with small degree coefficient polynomials.

One of the most important example for such a sequence is the *explicit inversive generator* over $\mathbb{F}_p$, with some prime $p \geq 3$, defined by

$$s_n = \begin{cases} (an-b)^{-1} & \text{if } an-b \not\equiv 0 \mod p \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

with some $a, b \in \mathbb{F}_p$, $a \neq 0$. Its generating function $G_{a,b}(x)$ satisfies

$$ax(1-x)^p G'_{a,b}(x) - b(1-x)^p G_{a,b}(x) - a(1-x)^{p-1} + ax^{b/a \bmod p} = 0,$$

see Corollary 1 below.

Theorem 3. *Let $\mathcal{S} = (s_n)$ be a sequence over $\mathbb{F}_p$. Assume, that its generating function $G(x)$ satisfies*

$$f_2(x)G'(x) + f_1(x)G(x) + f_0(x) \equiv 0 \mod x^M \tag{8}$$

with $M \geq 1$ for some polynomials $f_0(x), f_1(x), f_2(x) \in \mathbb{F}_p[x]$ with

(i) $f_2(x) = \bar{f}_2(x)f_1(x)$ *for some* $\bar{f}_2(x) \in \mathbb{F}_p[x]$*;*
(ii) $f_1(x)$ *is a non-constant polynomial without multiple zeros;*
(iii) $\gcd(f_2, f_1, f_0) = 1$.

Let $F = \max\{\deg f_2 - 1, \deg f_1, \deg f_0 - 1\}$. Then

$$E_N(\mathcal{S})(E_N(\mathcal{S}) + F) \geq N \quad \textit{for } \deg f_0 + 1 < N \leq M.$$

Previously, only a few example for sequences were known with large expansion complexity, namely the sequences of binomial coefficients $\mathcal{A} = (a_n)_{n=0}^{\infty}$, defined by

$$a_n = \binom{n+k}{k} \mod p, \quad n = 0, 1, \dots$$

for some $k \geq 0$, see [7], and the explicit inversive generator defined by (7) with $b = 0$, see [5]. Theorem 3 allows us to control the expansion complexity of a much wider family of sequences.

We also remark, that (8) defines a regular linear equation system for the coefficients of $G(x)$, that is, for the sequence elements (s_n). Hence, sequences satisfying (8) can be effectively computable.

In order to prove Theorem 3, we need the following result, see [3, Lemma 6].

Lemma 3. *Let $h(x, y) \in \mathbb{F}_q[x, y]$ be an irreducible polynomial of degree d and let $\mathcal{S}$ be an expansion sequences defined by $h(x, y)$. Let $f(x, y) \in \mathbb{F}_q[x, y]$ be a nonzero polynomial with*

$$f(x, G(x)) \equiv 0 \mod x^{d \cdot \deg f}.$$

Then $f(x, y)$ is a multiple of $h(x, y)$.

Proof of Theorem 3. Put $K = \deg f_0(x)$. There is a nonzero element among $s_0, \dots, s_{K+1}$ and thus $E_{K+2}(\mathcal{S}) \geq 1$. Indeed, if $G(x) \equiv 0 \mod x^{K+2}$, then $G(x) \equiv 0 \mod x^{K+1}$. so by (8) we have $f_0(x) = 0$, a contradiction.

If $s_0 = 0$, consider the sequence $\bar{\mathcal{S}} = (\bar{s}_n)$ with $\bar{s}_0 = 1$ and $\bar{s}_n = s_n$ for $n \geq 1$. Let $\bar{G}(x) = G(x) + 1$ be the generating function of $\bar{\mathcal{S}}$. Then $h(x, \bar{G}(x)) \equiv 0 \mod x^N$

if and only if $h(x, G(x)+1) \equiv 0 \mod x^N$. Thus $E_N(\mathcal{S}) = E_N(\bar{\mathcal{S}})$ whenever $E_N(\mathcal{S}) > 0$. As it holds for $N \geq K+2$, we can assume that $s_0 \neq 0$ and $E_1(\mathcal{S}) = 1$.

Now suppose that the result does not hold for some $N \geq K+3$, and fix N as a minimal such value and put $d = E_N(\mathcal{S})$. Then

$$d(d+F) < N. \tag{9}$$

Let $h(x, y) \in \mathbb{F}_p[x, y]$ such that $\deg h(x, y) = d$ and $h(x, G(x)) \equiv 0 \mod x^N$. First we prove, that $h(x, y)$ is irreducible. Suppose, that $h(x, y) = h_1(x, y)h_2(x, y)$ and

$$h_1(x, G(x)) \equiv 0 \mod x_1^N, \quad h_2(x, G(x)) \equiv 0 \mod x_2^N, \quad N_1 + N_2 \geq N.$$

Then by the minimality of N we have

$$\deg h_1(\deg h_1 + F) \geq N_1 \text{ and } \deg h_2(\deg h_2 + F) \geq N_2.$$

Thus

$$N_1 + N_2 \leq \deg h_1(\deg h_1 + F) + \deg h_2(\deg h_2 + F) \leq d(d+F) < N, \tag{10}$$

a contradiction.

Taking the derivative of the equation $h(x, G(x)) \equiv 0 \mod x^N$ we get

$$\frac{\partial h}{\partial x}(x, G(x)) + \frac{\partial h}{\partial y}(x, G(x))G'(x) \equiv 0 \mod x^{N-1},$$

thus multiplying it with $f_2(x)$ the we get by (8) that

(11)

$$f_2(x)\frac{\partial h}{\partial x}(x, G(x)) - f_1(x)G(x)\frac{\partial h}{\partial y}(x, G(x)) - f_0(x)\frac{\partial h}{\partial y}(x, G(x)) \equiv 0 \mod x^{N-1}.$$

The degree of

$$g(x, y) = f_2(x)\frac{\partial h}{\partial x}(x, y) - f_1(x)y\frac{\partial h}{\partial y}(x, y) - f_0(x)\frac{\partial h}{\partial y}(x, y) \in \mathbb{F}_p[x, y] \tag{12}$$

is $\deg g(x, y) \leq d + F$.

Let $\bar{\mathcal{S}} = (\bar{s}_n)$ be an expansion sequence defined $h(x, y)$ with $\bar{s}_n = s_n$ for $0 \leq n < N$. As $d^2 < N$, $\bar{S}$ is unique. Then by (9), (11) and by Lemma 3 we get that $g(x, y)$ is a multiple of $h(x, y)$,

$$g(x, y) = c(x, y)h(x, y) \tag{13}$$

for some nonzero $c(x, y) \in \mathbb{F}_p[x, y]$. Comparing the degrees of $g(x, y)$ and $c(x, y)h(x, y)$ with respect to y, we get $c(x, y) = c(x) \in \mathbb{F}_p[x]$. Write

$$h(x, y) = \sum_{i=0}^{k} r_k(x)y^k, \quad r_i(x) \in \mathbb{F}_p[x], \quad 0 \leq i \leq k.$$

The coefficient of y^k in $c(x)h(x, y)$ is

$$f_2(x)r_k'(x) - kf_1(x)r_k(x) = c(x)r_k(x). \tag{14}$$

By (13) we have $\gcd(f_1, c) \mid g$ and by (12) and (13) we get $\gcd(f_1, c) \mid \frac{\partial h}{\partial y}(x, y)$. Consequently $\gcd(f_1, c) \mid r_k$. As $f_1(x)$ has no multiple zeros and it divides the right-hand side of (14), we get that f_1 divides r_k.

Let $\alpha \in \overline{\mathbb{F}}_p$ be a root of f_1 and let $t > 0$ be its multiplicity in $r_k(x)$. As its multiplicity in $f_1(x)r_k(x)$ is $t+1$ and in $f_2(x)r_k'(x)$ is t, we have that it is not a zero of $c(x)$ by (14). Substituting $x = \alpha$ in (14), we get

$$c(\alpha)h(\alpha, y) = f_0(\alpha)\frac{\partial h}{\partial y}(\alpha, y)$$

with $c(\alpha) \neq 0$. By assumption $f_0(\alpha) \neq 0$, thus comparing the degrees of both sides with respect to y, we get $h(\alpha, y) = 0$, i.e. the minimal polynomial of α divides $h(x, y)$, a contradiction. □

Theorem 3 allows to us to control the expansion complexity of the explicit inversive generator defined by (7). We remark, that for $b = 0$ it was shown by Gómez-Pérez, Mérai and Niederreiter that the sequence has optimal expansion complexity, see [5]. Now we deal with the general case.

Corollary 1. *Let $\mathcal{S} = (s_n)$ the explicit inversive generator defined by (7) with $a, b \in \mathbb{F}_p$, $a \neq 0$. Then we have*

$$E_N(\mathcal{S}) \geq cN^{1/4} \quad \text{for } 2 \leq N < p$$

for some absolute constant $c > 0$.

Proof. As $G_{a,b}(x) = a^{-1}G_{1,b/a}(x)$, we can assume, that $a = 1$. Write $G(x) = G_{1,b}(x)$. Then

$$G(x) = \sum_{\substack{n=0 \\ n \not\equiv b \bmod p}}^{\infty} \frac{1}{n-b}x^n = x^b \sum_{\substack{n=0 \\ n \not\equiv b \bmod p}}^{\infty} \frac{1}{n-b}x^{n-b}$$

Now

$$\left(x^{-b}G(x)\right)' = -bx^{-b-1}G(x) + x^{-b}G'(x). \tag{15}$$

On the other hand

(16)

$$\left(x^{-b}G(x)\right)' = \left(\sum_{\substack{n=0 \\ n \not\equiv b \bmod p}}^{\infty} \frac{1}{n-b}x^{n-b}\right)' = \sum_{\substack{n=0 \\ n \not\equiv b \bmod p}}^{\infty} x^{n-b-1} = \frac{1}{x^{b+1}} \sum_{\substack{n=0 \\ n \not\equiv b \bmod p}}^{\infty} x^n$$
$$= \frac{1}{x^{b+1}}\left(\sum_{n=0}^{\infty} x^n - \sum_{n=0}^{\infty} x^{pn+b}\right) = \frac{1}{x^{b+1}}\left(\frac{1}{1-x} - x^b\frac{1}{1-x^p}\right).$$

Then by (15) and (16) we get

$$x(1-x)^{p+1}G'(x) - b(1-x)^{p+1}G(x) - (1-x)^p + x^b(1-x) = 0. \tag{17}$$

For $N \leq b$ we apply Theorem 3 with $M = b$. In this case we get

$$E_N(\mathcal{S})(E_N(\mathcal{S}) + 1) \geq N \quad \text{for} \quad 2 \leq N \leq b. \tag{18}$$

For $N \geq b$ we choose $M = p$, and by Theorem 3 we get

$$E_N(\mathcal{S})(E_N(\mathcal{S}) + b) \geq N \quad \text{for} \quad b+3 \leq N \leq p-1. \tag{19}$$

If $N \ll b$, $E_N(\mathcal{S}) \gg \sqrt{N}$ by (18) and if $N \gg b^2$, we get $E_N(\mathcal{S}) \gg \sqrt{N}$. Finally, using $E_{N+1}(\mathcal{S}) \geq E_N(\mathcal{S})$, we get $E_N(\mathcal{S}) \gg \sqrt{b}$ for $b \ll N \ll b^2$ which gives the result. □

Remark. The proof gives the stronger bounds on expansion complexity of the explicit inversive generator $\mathcal{S}_{a,b}$ with parameters $a \in \mathbb{F}_p^*$, $b \in \mathbb{F}_p$

$$E_N(\mathcal{S}_{a,b}) \gg \sqrt{N} \quad \text{for} \quad N \ll b \text{ or } N \gg b^2.$$

If the parameters (a, b) are chosen uniformly from $\mathbb{F}_p^* \times \mathbb{F}_p$, then it provides a square-root bound for almost all parameters (a, b) which is optimal, see [5, Theorem 1].

In Theorem 3 we gave lower bounds on the Nth expansion complexity of sequences whose generating function satisfies a first order differential equation (6). However, we conjecture that sequences with higher order differential equation (6) have also large expansion complexity.

Problem 1. *Let $\mathcal{S} = (s_n)$ be a sequence in $\mathbb{F}_q$ such that its generating function $G(x)$ satisfies* (6)*. Estimate the Nth expansion complexity $E_N(\mathcal{S})$ of the sequence $\mathcal{S}$ in terms of the coefficient polynomials of* (6)*.*

In [7], Mérai, Niederreiter and Winterhof studied the connection between the expansion and linear complexity of sequences. We recall, that the Nth linear complexity $L_N(\mathcal{S})$ of a sequence $\mathcal{S}$ over a finite field $\mathbb{F}_q$ is zero if $s_0 = \cdots = s_{N-1} = 0$, otherwise the least positive L such that there exist $c_0, \ldots, c_{L-1} \in \mathbb{F}_q$ such that

$$s_{n+L} = c_{L-1}s_{n+L-1} + \cdots + c_0 s_n, \quad 0 \le n \le N - L - 1. \tag{20}$$

They proved, that large expansion complexity implies large linear complexity

$$L_N(\mathcal{S}) \ge \min\left\{E_N(\mathcal{S}) - 1, \frac{N+3}{2}\right\}.$$

They also provided lower bound on the expansion complexity in terms on the linear complexity, however the bound also depends on the linear recurrence relation (20).

Here we give lower bounds on the Nth linear complexity of sequences with (6). This result along with [7] establishes Problem 1.

Theorem 4. *Let $\mathcal{S} = (s_n)_{n=0}^{\infty}$ be a sequence over $\mathbb{F}_q$. Assume, that its generating function $G(x)$ satisfies*

$$f_{k+1}(x)G^{(k)}(x) + \cdots + f_2(x)G^{(1)}(x) + f_1(x)G(x) + f_0(x) \equiv 0 \mod x^M \tag{21}$$

for some polynomials $f_{k+1}(x), \ldots, f_0(x) \in \mathbb{F}_q[x]$ and some integer $M \ge 1$. Let $F = \max\{\deg f_{k+1}(x), \ldots, \deg f_0(x)\}$. Then

$$L_N(\mathcal{S}) \ge \frac{N - F + k + 1}{2^{k+1} - k - 1} \quad \text{for } N \le M.$$

Proof. For an $N \le M$ put $L = L_N(\mathcal{S})$. Then there exist polynomials $g(x), h(x) \in \mathbb{F}_q[x]$, $\deg g(x) < L$, $\deg h(x) \le L$, $h(x) \ne 0$ such that

$$h(x)G(x) \equiv g(x) \mod x^N.$$

One can choose

$$h(x) = \sum_{i=0}^{L-1} c_i x^{L-i} \quad \text{and} \quad g(x) = \sum_{m=0}^{L-1}\left(\sum_{\ell=L-m}^{L} c_\ell s_{m+\ell-L}\right)x^m,$$

where $c_L = -1$ and $c_0, \ldots, c_{L-1}$ are the coefficients of the linear recurrence relation (20). Then

$$h^2(x)G'(x) \equiv g_1(x) \mod x^{N-1} \tag{22}$$

for some polynomial $\deg g_1(x) \le 2L-2$. Indeed, writing

$$G(x) = \frac{g(x)}{h(x)} + \frac{K(x)x^N}{h(x)}$$

and taking its derivative we get (22). Then by induction we get

$$h^{2^\ell}(x)G^{(\ell)}(x) \equiv g_\ell(x) \mod x^{N-\ell}, \quad \deg g_\ell(x) \le 2^\ell L - \ell - 1, \quad 0 \le k \le N. \tag{23}$$

Then multiplying (21) by $h^{2^k}(x)$ we get

$$\begin{aligned} 0 &\equiv f_{k+1}(x)h^{2^k}(x)G^{(k)}(x) + \dots + f_1(x)h^{2^k}(x)G(x) + f_0(x)h^{2^k}(x) \\ &\equiv f_{k+1}(x)g_k(x) + \dots + f_1(x)h^{2^k-1}(x)g_1(x) + f_0(x)h^{2^k}(x) \mod x^{N-L} \end{aligned}$$

Then

$$f_{k+1}(x)g_k(x) + \dots + f_1(x)h^{2^k-1}(x)g_1(x) + f_0(x)h^{2^k}(x) = J(x)x^{N-L}$$

Comparing the degrees of both sides we get

$$\max_{0\le\ell\le k}\left\{\deg f_{\ell+1}(x) + \deg g_\ell(x) + \left(2^k - \ell\right)\deg h(x)\right\} \ge N - L$$

which gives the result.

□

Acknowledgement

D. G-P. is partially supported by project MTM2014-55421-P from the Ministerio de Economia y Competitividad and L. M. is partially supported by the Austrian Science Fund FWF Project I1751-N26.

References

[1] Magali Bardet, Jean-Charles Faugère, and Bruno Salvy. On the complexity of the F_5 Gröbner basis algorithm. *J. Symbolic Comput.*, 70:49–70, 2015.

[2] Leonard Carlitz et al. The distribution of irreducible polynomials in several indeterminates. *Illinois Journal of Mathematics*, 7(3):371–375, 1963.

[3] Claus Diem. On the use of expansion series for stream ciphers. *LMS Journal of Computation and Mathematics*, 15:326–340, 2012.

[4] David Eisenbud. *Commutative algebra, volume 150 of Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995.

[5] Domingo Gómez-Pérez, László Mérai, and Harald Niederreiter. On the expansion complexity of sequences over finite fields. *IEEE Trans. Inform. Theory*. http://dx.doi.org/10.1109/TIT.2018.2792490.

[6] Izrail Solomonovich Gradshteyn and Iosif Moiseevich Ryzhik. *Table of integrals, series, and products*. Academic press, 2014.

[7] László Mérai, Harald Niederreiter, and Arne Winterhof. Expansion complexity and linear complexity of sequences over finite fields. *Cryptography and Communications*, 9(4):501–509, 2017.

D.G.-P.: Department of Mathematics, University of Cantabria, Santander 39005, Spain

E-mail address: domingo.gomez@unican.es

Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences and Institute of Financial Mathematics and Applied Number Theory, Johannes Kepler University, Altenberger Strasse 69, A-4040 Linz, Austria

E-mail address: laszlo.merai@oeaw.ac.at