Monatsh Math (2015) 176:241–253
DOI 10.1007/s00605-014-0697-0

Subgroups generated by rational functions in finite fields

Domingo Gómez-Pérez · Igor E. Shparlinski

Received: 29 September 2013 / Accepted: 2 October 2014 / Published online: 29 October 2014

Abstract For a large prime p, a rational function $\psi \in \mathbb{F}_p(X)$ over the finite field $\mathbb{F}_p$ of p elements, and integers u and $H \geq 1$, we obtain a lower bound on the number consecutive values $\psi(x)$, $x = u+1, \ldots, u+H$ that belong to a given multiplicative subgroup of $\mathbb{F}_p^*$.

Keywords Polynomial congruences · Finite fields · Value sets of polynomials · Multiplicative subgroups

Mathematics Subject Classification 11D79 · 11T06

1 Introduction

For a prime p, let $\mathbb{F}_p$ denote the finite field with p elements, which we always assume to be represented by the set $\{0, \ldots, p-1\}$.

Given a rational function

$$\psi(X) = \frac{f(X)}{g(X)} \in \mathbb{F}_p(X)$$

Communicated by U. Zannier.

D. Gómez-Pérez (✉)
Department of Mathematics, University of Cantabria, 39005 Santander, Spain
e-mail: domingo.gomez@unican.es

I. E. Shparlinski
Department of Pure Mathematics, University of New South Wales,
Sydney, NSW 2052, Australia
e-mail: igor.shparlinski@unsw.edu.au

where $f, g \in \mathbb{F}_p[X]$ are relatively prime polynomials, and an 'interesting' set $\mathcal{S} \subseteq \mathbb{F}_p$, it is natural to ask how the value set

$$\psi(\mathcal{S}) = \{\psi(x) \colon x \in \mathcal{S},\ g(x) \neq 0\}$$

is distributed. For instance, given another 'interesting' set $\mathcal{T}$, our goal is to obtain nontrivial bounds on the size of the intersection

$$N_\psi(\mathcal{S}, \mathcal{T}) = \#\left(\psi(\mathcal{S}) \cap \mathcal{T}\right).$$

In particular, we are interested in the cases when $N_\psi(\mathcal{S}, \mathcal{T})$ achieves the trivial upper bound

$$N_\psi(\mathcal{S}, \mathcal{T}) \le \min\{\#\mathcal{S}, \#\mathcal{T}\}.$$

Typical examples of such sets $\mathcal{S}$ and $\mathcal{T}$ are given by intervals $\mathcal{I}$ of consecutive integers and multiplicative subgroups $\mathcal{G}$ of $\mathbb{F}_p^*$. For large intervals and subgroups, a standard application of bounds of exponential and multiplicative character sums leads to asymptotic formulas for the relevant values of $N_\psi(\mathcal{S}, \mathcal{T})$, see [7,11,19]. Thus only the case of small intervals and groups is of interest.

For a polynomial $f \in \mathbb{F}_p[X]$ and two intervals $\mathcal{I} = \{u+1, \dots, u+H\}$ and $\mathcal{J} = \{v+1, \dots, v+H\}$ of H consecutive integers, various bounds on the cardinality of the intersection $f(\mathcal{I}) \cap \mathcal{J}$ are given in [7,11]. To present some of these results, for positive integers d, k and H, we denote by $J_{d,k}(H)$ the number of solutions to the system of equations

$$x_1^\nu + \dots + x_k^\nu = x_{k+1}^\nu + \dots + x_{2k}^\nu, \quad \nu = 1, \dots, d,$$

in positive integers $x_1, \dots, x_{2k} \le H$. Then by [11, Theorem 1], for any $f \in \mathbb{F}_p[X]$ of degree $d \ge 2$ and two intervals $\mathcal{I}$ and $\mathcal{J}$ of $H < p$ consecutive integers, we have

$$N_f(\mathcal{I}, \mathcal{J}) \le H(H/p)^{1/2\kappa(d)+o(1)} + H^{1-(d-1)/2\kappa(d)+o(1)},$$

as $H \to \infty$, where $\kappa(d)$ is the smallest integer κ such that for $k \ge \kappa$ there exists a constant $C(d, k)$ depending only on k and d and such that

$$J_{d,k}(H) \le C(d, k) H^{2k-d(d+1)/2+o(1)}$$

holds as $H \to \infty$, see also [7] for some improvements and results for related problems. In [7,11] the bounds of Wooley [22,23] are used that give the presently best known estimates on $\kappa(d)$ (at least for a large d), see also [24] for further progress in estimating $\kappa(d)$.

It is easy to see that the argument of the proof of [11, Theorem 1] allows to consider intervals of $\mathcal{I}$ and $\mathcal{J}$ of different lengths as well and for intervals

$$\mathcal{I} = \{u+1, \dots, u+H\} \quad \text{and} \quad \mathcal{J} = \{v+1, \dots, v+K\}$$

with $1 \le H, K < p$ it leads to the bound

$$N_f(\mathcal{I}, \mathcal{J}) \le H^{1+o(1)} \left((K/p)^{1/2\kappa(d)} + (K/H^d)^{1/2\kappa(d)} \right),$$

see also a more general result of Kerr [15, Theorem 3.1] that applies to multivariate polynomials and to congruences modulo a composite number.

Furthermore, let $K_\psi(H)$ be the smallest K for which there are intervals $\mathcal{I} = \{u+1, \dots, u+H\}$ and $\mathcal{J} = \{v+1, \dots, v+K\}$ for which $N_\psi(\mathcal{I}, \mathcal{J}) = \#\mathcal{I}$. That is, $K_\psi(H)$ is the length of the shortest interval, which may contain H consecutive values of $\psi \in \mathbb{F}_p(X)$ of degree d.

Defining $\kappa^*(d)$ in the same way as $\kappa(d)$, however with respect to the more precise bound

$$J_{d,k}(H) \le C(d,k) H^{2k-d(d+1)/2}$$

[that is, without $o(1)$ in the exponent] we can easily derive that for any polynomial $f \in \mathbb{F}_p[X]$ of degree d,

$$K_f(H) \ge c(d) H^d, \tag{1}$$

for some constant $c(d) > 0$ that depends only on d. To see that the bound (1) is optimal it is enough to take $f(X) = X^d$ and $u = 0$. Note that the proof of (1) depends only on the existence of $\kappa^*(d)$ rather than on its specific bounds. However, we recall that Wooley [22, Theorem 1.2] shows that for some constant $\mathfrak{S}(d,k) > 0$ depending only on d and k we have

$$J_{d,k}(H) \sim \mathfrak{S}(d,k) H^{2k-d(d+1)/2}$$

for any fixed $d \ge 3$ and $k \ge d^2 + d + 1$. In particular, $\kappa^*(d) \le d^2 + d + 1$.

Here we concentrate on estimating $N_\psi(\mathcal{I}, \mathcal{G})$ for an interval $\mathcal{I}$ of H consecutive integers and a multiplicative subgroup $\mathcal{G} \subseteq \mathbb{F}_p^*$ of order T. This question has been mentioned in [11, Section 4] as an open problem.

We remark that for linear polynomials f the result of [4, Corollary 34] have a natural interpretation as a lower bound on the order of a subgroup $\mathcal{G} \subseteq \mathbb{F}_p^*$ for which $N_f(\mathcal{I}, \mathcal{G}) = \#\mathcal{I}$. In particular, we infer from [4, Corollary 34] that for any linear polynomials $f(X) = aX + b \in \mathbb{F}_p[X]$ and fixed integer $\nu = 1, 2, \dots$, for an interval $\mathcal{I}$ of $H \le p^{1/(\nu^2-1)}$ consecutive integers and a subgroup $\mathcal{G}$, the equality $N_f(\mathcal{I}, \mathcal{G}) = \#\mathcal{I}$ implies $\#\mathcal{G} \ge H^{\nu+o(1)}$.

We also remark that the results of [5, Section 5] have a similar interpretation for the identity $N_f(\mathcal{I}, \mathcal{G}) = \#\mathcal{I}$ with linear polynomials, however apply to almost all primes p (rather than to all primes).

Furthermore, a result of Bourgain [3, Theorem 2] gives a nontrivial bound on the intersection of an interval centered at 0, that is, of the form $\mathcal{I} = \{0, \pm 1, \dots, \pm H\}$ and a co-set $a\mathcal{G}$ (with $a \in \mathbb{F}_p^*$) of a multiplicative group $\mathcal{G} \subseteq \mathbb{F}_p^*$, provided that $H < p^{1-\varepsilon}$ and $\#\mathcal{G} \ge g_0(\varepsilon)$, for some constant $g_0(\varepsilon)$ depending only on an arbitrary $\varepsilon > 0$.

We note that several bounds on $\#(f(\mathcal{G}) \cap \mathcal{G})$ for a multiplicative subgroup $\mathcal{G} \subseteq \mathbb{F}_p^*$ are given in [19], but they apply only to polynomials f defined over $\mathbb{Z}$ and are not uniform with respect to the height (that is, the size of the coefficients) of f. Thus the question of estimating $N_f(\mathcal{G}, \mathcal{G})$ remains open. On the other hand, a number of results about points on curves and algebraic varieties with coordinates from small subgroups, in particular, in relation to the *Poonen Conjecture*, have been given in [6,8–10,17,18,20,21].

We recall that the notations $U = O(V)$, $U \ll V$ and $V \gg U$ are all equivalent to the statement that the inequality $|U| \le c\,V$ holds with some constant $c > 0$. Throughout the paper, any implied constants in these symbols may occasionally depend, where obvious, on $d = \deg f$ and $e = \deg g$, but are absolute otherwise.

2 Preparations

2.1 Absolute irreducibility of some polynomials

As usual, we use $\overline{\mathbb{F}}_p$ to denote the algebraic closure of $\mathbb{F}_p$ and X, Y to denote indeterminate variables. We also use $\overline{\mathbb{F}}_p(X)$, $\overline{\mathbb{F}}_p(Y)$, $\overline{\mathbb{F}}_p(X, Y)$ to denote the corresponding fields of rational functions over $\overline{\mathbb{F}}_p$.

We recall that the degree of a rational function in the variables X, Y

$$F(X, Y) = \frac{s(X, Y)}{t(X, Y)} \in \overline{\mathbb{F}}_p(X, Y), \quad \gcd(s(X, Y), t(X, Y)) = 1,$$

is $\deg F = \max\{\deg s, \deg t\}$.

It is also known that if $R(X) \in \overline{\mathbb{F}}_p(X)$ is a rational function then

$$\deg(R \circ F) = \deg R \deg F, \tag{2}$$

where $\circ$ denotes the composition.

We use the following result of Bodin [1, Theorem 5.3] adapted to our purposes. Also, see [16] for results in fields of zero characteristic.

Lemma 1 *Let $s(X, Y), t(X, Y) \in \mathbb{F}_p[X, Y]$ be polynomials such that there does not exist a rational function $R(X) \in \overline{\mathbb{F}}_p(X)$ with $\deg R > 1$ and a bivariate rational function $G(X, Y) \in \overline{\mathbb{F}}_p[X, Y]$ such that,*

$$F(X, Y) = \frac{s(X, Y)}{t(X, Y)} = R(G(X, Y)).$$

The number of elements λ such that the polynomial $s(X, Y) - \lambda t(X, Y)$ is reducible over $\overline{\mathbb{F}}_p[X, Y]$ is at most $(\deg F)^2$.

We say that a rational function $f \in \overline{\mathbb{F}}_p(X)$ is a *perfect power* of another rational function if and only if $f(X) = (g(X))^n$ for some rational function $g(X) \in \overline{\mathbb{F}}_p(X)$ and integer $n \ge 2$. Because $\overline{\mathbb{F}}_p$ is an algebraic closed field, it is trivial to see that if $f(X)$ is a perfect power, then $af(X)$ is also a perfect power for any $a \in \overline{\mathbb{F}}_p$. We need the following easy technical lemma.

Lemma 2 *Let $P_1(X), Q_1(X) \in \overline{\mathbb{F}}_p[X]$ and $P_2(Y), Q_2(Y) \in \overline{\mathbb{F}}_p[Y]$ be two pairs of relatively prime polynomials. Then the following bivariate polynomial*

$$F_{r,s}(X, Y) = rP_1(X)Q_2(Y) - sQ_1(X)P_2(Y),$$

is not divisible by any univariate polynomial for all $r, s \in \overline{\mathbb{F}}_p^*$,

Proof Suppose that this polynomial is divisible by an univariate polynomial $d(X)$. Take any root $\alpha \in \overline{\mathbb{F}}_p$ of the polynomial d and substitute $X = \alpha$ in $F_{r,s}(X, Y)$, getting

$$rP_1(\alpha)Q_2(Y) - sQ_1(\alpha)P_2(Y) = 0.$$

Here, we have two different possibilities:

- If $rP_1(\alpha) = 0$, then $Q_1(\alpha) = 0$, and we get a contradiction,
- In other case, $\gcd(Q_2(Y), P_2(Y)) \neq 1$, contradicting our hypothesis.

This finishes the proof. □

Now, we prove the following result about irreducibility.

Lemma 3 *Given relatively prime polynomials $f, g \in \overline{\mathbb{F}}_p[X]$ and if a rational function $f(X)/g(X) \in \overline{\mathbb{F}}_p(X)$ of degree $D \geq 2$ is not a perfect power then $f(X)g(Y) - \lambda f(Y)g(X)$ is reducible over $\overline{\mathbb{F}}_p[X, Y]$ for at most $4D^2$ values of $\lambda \in \overline{\mathbb{F}}_p^*$.*

Proof First we describe the idea of the proof. Our aim is to show that the condition of Lemma 1 holds for the polynomial $f(X)g(Y) - \lambda f(Y)g(X)$. Indeed, we show that if

$$\frac{f(X)g(Y)}{g(X)f(Y)} = R(G(X, Y)), \tag{3}$$

with a rational function $R \in \overline{\mathbb{F}}_p(X)$ of degree $\deg R \geq 2$ and a bivariate rational function $G(X, Y) \in \overline{\mathbb{F}}_p(X, Y)$, then there exists another $\widetilde{R} \in \overline{\mathbb{F}}_p(X)$ and $\widetilde{G}(X, Y) \in \overline{\mathbb{F}}_p(X, Y)$

$$\frac{f(X)g(Y)}{g(X)f(Y)} = \left(\widetilde{R}\left(\widetilde{G}(X, Y)\right)\right)^m,$$

for an appropriate integer $m \geq 2$. Comparing coefficients, it is easy to arrive at the conclusion that $f(X)/g(X)$ is a perfect power.

Without loss of generality, we suppose $R(0) = 0$. Indeed, we can take any root of $R(X)$ and replace $R(X)$ with $R(X + \alpha)$ and $G(X, Y)$ with $G(X, Y) - \alpha$.

So, indeed we have

$$R(X) = a\frac{X\prod_{i=2}^{k}(X - r_i)}{\prod_{j=1}^{m}(X - s_j)}.$$

Writing $G(X,Y) = G_1(X,Y)/G_2(X,Y)$ in its lowest terms and by hypothesis, we have that the fraction on the right of this inequality,

$$\frac{f(X)g(Y)}{g(X)f(Y)} = a\frac{G_2(X,Y)^{N-k}}{G_2(X,Y)^{N-m}} \times \frac{G_1(X,Y)\prod_{i=2}^{k}(G_1(X,Y) - r_i(G_2(X,Y))}{\prod_{j=1}^{m}(G_1(X,Y) - s_j G_2(X,Y))},$$

where

$$N = \max\{k, m\}$$

is in its lowest terms. This means that $G_1(X,Y) = P_1(X)P_2(Y)$ and $G_2(X,Y) = s_1^{-1}(P_1(X)P_2(Y) - Q_1(X)Q_2(Y))$, where P_1, P_2, Q_1, Q_2 are divisors of f or g. Because $\gcd(G_1(X,Y), G_2(X,Y)) = 1$, we have that

$$\gcd(P_1(X), Q_1(X)) = \gcd(P_2(Y), Q_2(Y)) = 1.$$

Lemma 2 implies that $m = k$ as otherwise $G_2(X,Y)$ is divisible by an univariate polynomial. This implies,

$$\frac{f(X)g(Y)}{g(X)f(Y)} = a\frac{G_1(X,Y)\prod_{i=2}^{m}(G_1(X,Y) - r_i G_2(X,Y))}{\prod_{j=1}^{m}(G_1(X,Y) - s_j G_2(X,Y))}.$$

Now, suppose that there exists another value

$$s \in \{r_2, \ldots, r_m, s_2, \ldots, s_m\}, \quad s \neq 0, s_1.$$

Then, the following polynomial

$$G_1(X,Y) - sG_2(X,Y) = \left(1 - ss_1^{-1}\right)P_1(X)P_2(Y) + s_1^{-1}Q_1(X)Q_2(Y)$$

is divisible by an univariate polynomial which contradicts Lemma 2. So, this means that $R(X)$ can be written in the following form,

$$R(X) = \left(\frac{X}{X - s_1}\right)^m,$$

and this concludes the proof. □

Notice that the condition that $f(X)/g(X)$ is not a perfect power of a polynomial is necessary, indeed if $f(X) = (h(X))^n$ and $g(X) = 1$ with $f(X), h(X) \in \overline{\mathbb{F}}_p[X]$ then $f(X) - \lambda^n f(Y)$ is divisible by $h(X) - \lambda h(Y)$ for any $\lambda \in \overline{\mathbb{F}}_p$.

2.2 Integral points on affine curves

We need the following estimate of Bombieri and Pila [2] on the number of integral points on polynomial curves.

Lemma 4 *Let $\mathcal{C}$ be a plane absolutely irreducible curve of degree $n \geq 2$ and let $H \geq \exp(n^6)$. Then the number of integral points on $\mathcal{C}$ inside of the square $[0, H] \times [0, H]$ is at most $H^{1/n} \exp(12\sqrt{n \log H \log\log H})$.*

2.3 Small values of linear functions

We need a result about small values of residues modulo p of several linear functions. Such a result has been derived in [12, Lemma 3.2] from the Dirichlet pigeon-hole principle. Here use a slightly more precise and explicit form of this result which is derived in [13] from the *Minkowski theorem*.

First we recall some standard notions of the theory of geometric lattices.

Let $\mathbf{b}_1, \ldots, \mathbf{b}_r$ be r linearly independent vectors in $\mathbb{R}^s$. The set

$$\mathcal{L} = \{\mathbf{z} : \mathbf{z} = c_1\mathbf{b}_1 + \cdots + c_r\mathbf{b}_r, \quad c_1, \ldots, c_r \in \mathbb{Z}\}$$

is called an *r-dimensional lattice in* $\mathbb{R}^s$ with a *basis* $\{\mathbf{b}_1, \ldots, \mathbf{b}_r\}$.

To each lattice $\mathcal{L}$ one can naturally associate its *volume*

$$\operatorname{vol} \mathcal{L} = \left(\det\left(B^t B\right)\right)^{1/2},$$

where B is the $s \times r$ matrix whose columns are formed by the vectors $\mathbf{b}_1, \ldots, \mathbf{b}_r$ and B^t is the transposition of B. It is well known that $\operatorname{vol} \mathcal{L}$ does not depend on the choice of the basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_r\}$, we refer to [14] for a background on lattices.

For a vector $\mathbf{u}$, let

$$\|\mathbf{u}\|_\infty = \max\{|u_1|, \ldots, |u_s|\}$$

denote its *infinity norm* of $\mathbf{u} = (u_1, \ldots, u_s) \in \mathbb{R}^s$.

The famous *Minkowski theorem*, see [14, Theorem 5.3.6], gives an upper bound on the size of the shortest nonzero vector in any r-dimensional lattice $\mathcal{L}$ in terms of its volume.

Lemma 5 *For any r-dimensional lattice $\mathcal{L}$ we have*

$$\min\{\|\mathbf{z}\|_\infty : \ \mathbf{z} \in \mathcal{L} \setminus \{\mathbf{0}\}\} \leq (\operatorname{vol} \mathcal{L})^{1/r}.$$

For an integer a we use $\langle a \rangle_p$ to denote the smallest by absolute value residue of a modulo p, that is

$$\langle a \rangle_p = \min_{k \in \mathbb{Z}} |a - kp|.$$

The following result is essentially contained in [13, Theorem 2]. We include here a short proof.

Lemma 6 *For any real numbers* $V_1, \dots, V_s$ *with*

$$p > V_1, \dots, V_s \geq 1 \quad \textit{and} \quad V_1 \dots V_s > p^{s-1}$$

and integers $b_1, \dots, b_s$*, there exists an integer* v *with* $\gcd(v, p) = 1$ *such that*

$$\langle b_i v \rangle_p \leq V_i, \quad i = 1, \dots, s.$$

Proof Without loss of the generality, we can take $b_1 = 1$. We introduce the following notation,

$$V = \prod_{i=1}^{s} V_i \tag{4}$$

and consider the lattice $\mathcal{L}$ generated by the columns of the following matrix

$$B = \begin{pmatrix} b_s V/V_s & 0 & \dots & 0 & pV/V_s \\ b_{s-1} V/V_{s-1} & 0 & \dots & pV/V_{s-1} & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b_2 V/V_2 & pV/V_2 & \dots & 0 & 0 \\ V/V_1 & 0 & \dots & 0 & 0 \end{pmatrix}.$$

Clearly the volume of $\mathcal{L}$ is

$$\operatorname{vol} \mathcal{L} = \frac{V}{V_1} \prod_{j=2}^{s} \frac{pV}{V_j} = V^{s-1} p^{s-1} \leq V^s$$

by (4) and the conditions on the size of the product $V_1 \dots V_s$. Consider a nonzero vector with the minimum infinity norm inside $\mathcal{L}$. By the definition of $\mathcal{L}$, this vector is a linear combination of the columns of B with integer coefficients, that is, it can be written in the following way

$$\left(\frac{c_1 V}{V_1}, \frac{(c_1 b_2 + c_2 p)V}{V_2}, \dots, \frac{(c_1 b_s + c_s p)V}{V_s} \right), \quad c_1, \dots, c_s \in \mathbb{Z}.$$

By Lemma 5 and the bound on the volume of $\mathcal{L}$, the following inequality holds,

$$\max \left\{ \left| \frac{c_1 V}{V_1} \right|, \left| \frac{(c_1 b_2 + c_2 p)V}{V_2} \right|, \dots, \left| \frac{(c_1 b_s + c_s p)V}{V_s} \right| \right\} \leq V.$$

From here, it is trivial to check that if we choose $v = c_1$, then

- $\langle v \rangle_p = \langle c_1 \rangle_p \leq V_1$,
- $\langle v b_i \rangle_p = \langle c_1 b_i \rangle_p \leq V_i, \quad i = 2, \dots, s$,

which finishes the proof. □

3 Main results

Theorem 7 *Let $\psi(X) = f(X)/g(X)$ where $f, g \in \mathbb{F}_p[X]$ relatively prime polynomials of degree d and e respectively with $d + e \geq 1$. We define*

$$\ell = \min\{d, e\}, \quad m = \max\{d, e\}$$

and set

$$k = (\ell + 1)\left(\ell m - \ell^2 + m^2 + m\right) \quad \text{and} \quad s = 2m\ell + 2m - \ell^2.$$

Assume that ψ is not a perfect power of another rational function over $\overline{\mathbb{F}}_p$. Then for any interval $\mathcal{I}$ of H consecutive integers and a subgroup $\mathcal{G}$ of $\mathbb{F}_p^*$ of order T, we have

$$N_\psi(\mathcal{I}, \mathcal{G}) \ll (1 + H^\rho p^{-\vartheta}) H^{\tau + o(1)} T^{1/2},$$

where

$$\vartheta = \frac{1}{2s}, \quad \rho = \frac{k}{2s}, \quad \tau = \frac{1}{2(\ell + m)},$$

and the implied constant depends on d and e.

Proof Clearly we can assume that

$$H \leq cp^{2\vartheta/(2\rho - 1)} \tag{5}$$

for some constant $c > 0$ which may depend on d and e as otherwise one easily verifies that

$$H^{\rho+\tau} p^{-\vartheta} \geq H^\rho p^{-\vartheta} \gg H^{1/2},$$

and hence the desired bound is weaker than the trivial estimate

$$N_\psi(\mathcal{I}, \mathcal{G}) \ll \min\{H, T\} \leq H^{1/2} T^{1/2}.$$

Making the transformation $X \mapsto X + u$, we can assume that $\mathcal{I} = \{1, \dots, H\}$. Let $1 \leq x_1 < \cdots < x_r \leq H$ be all $r = N_\psi(\mathcal{I}, \mathcal{G})$ values of $x \in \mathcal{I}$ with $\psi(x) \in \mathcal{G}$.

Let Λ be the set of exceptional values of $\lambda \in \overline{\mathbb{F}}_p$ described in Lemma 3. We see that there are only at most $4m^3 r$ pairs (x_i, x_j), $1 \leq i, j \leq r$, for which $\psi(x_i)/\psi(x_j) \in \Lambda$. Indeed, if x_j is fixed, then $\psi(x_i)$ can take at most $4m^2$ values of the form $\lambda\psi(x_j)$, with $\lambda \in \Lambda$,

Furthermore, each value $\lambda\psi(x_j)$ can be taken by $\psi(x_i)$ for at most D possible values of $i = 1, \dots, r$.

We now assume that $r > 8m^3$ as otherwise there is nothing to prove. Therefore, there is $\lambda \in \mathcal{G} \setminus \Lambda$ such that

$$\psi(x) \equiv \lambda\psi(y) \pmod p \tag{6}$$

for at least

$$\frac{r^2 - 4m^3 r}{T} \geq \frac{r^2}{2T} \tag{7}$$

pairs (x, y) with $x, y \in \{1, \ldots, H\}$.

Let

$$f(X)g(Y) - \lambda f(Y)g(X) = \sum_{i=0}^{m}\sum_{j=0}^{m} b_{i,j} X^i Y^j.$$

Let

$$\mathcal{H} = \{(i, j) : i, j = 0, \ldots, m,\ i + j \geq 1, \min\{i, j\} \leq \ell\}.$$

Clearly the noncostant terms $b_{i,j} X^i Y^j$ of $f(X)g(Y) - \lambda f(Y)g(X)$ are supported only on the subscripts $(i, j) \in \mathcal{H}$. We have

$$\#\mathcal{H} = 2(m + 1)(\ell + 1) - (\ell + 1)^2 - 1 = s$$

We now apply Lemma 6 with $s = \#\mathcal{H}$ and the vector $\left(b_{i,j}\right)_{(i,j)\in\mathcal{H}}$.

We also define the quantities U and $V_{i,j}$, $(i, j) \in \mathcal{H}$ by the relations

$$V_{i,j} H^{i+j} = U, \quad (i, j) \in \mathcal{H},$$

thus

$$\prod_{(i,j)\in\mathcal{H}} V_{i,j} = 2p^{s-1}.$$

By Lemma 6 there is an integer v with $\gcd(v, p) = 1$ such that

$$\left\langle b_{i,j} v \right\rangle_p \leq V_{i,j}$$

for every $(i, j) \in \mathcal{H}$.

We have

$$\sum_{(i,j)\in\mathcal{H}} (i+j) = 2\sum_{i=0}^{m}\sum_{j=0}^{\ell}(i+j) - \sum_{i=0}^{\ell}\sum_{j=0}^{\ell}(i+j)$$
$$= 2\sum_{i=0}^{m}\left((\ell+1)i + \frac{\ell(\ell+1)}{2}\right) - \sum_{i=0}^{\ell}\left((\ell+1)i + \frac{\ell(\ell+1)}{2}\right)$$
$$= 2\left(\frac{(\ell+1)m(m+1)}{2} + \frac{\ell(\ell+1)(m+1)}{2}\right)$$
$$-\frac{\ell(\ell+1)^2}{2} - \frac{\ell(\ell+1)^2}{2} = k.$$

Certainly it is easy to evaluate $V_{i,j}$, $(i,j)\in\mathcal{H}$ explicitly, however it is enough for us to note that we have

$$U^s H^{-k} = 2p^{s-1}.$$

Hence

$$U = 2^{1/s} p^{1-1/s} H^{k/s}. \tag{8}$$

We also assume that the constant c in (5) is small enough so the condition

$$\max_{(i,j)\in\mathcal{H}} \{V_{i,j}\} = UH^{-1} < p$$

is satisfied.

Let $F(X,Y)\in\mathbb{Z}[X]$ and $G(X,Y)\in\mathbb{Z}[X]$ be polynomials with coefficients in the interval $[-p/2, p/2]$, obtained by reducing $vf(X)g(Y)$ and $v\lambda f(Y)g(X)$ modulo p, respectively. Clearly (6) implies

$$F(x,y)\equiv G(x,y) \pmod p. \tag{9}$$

Furthermore, since for $x, y\in\{1,\dots,H\}$, we see from (8) and the trivial estimate on the constant coefficients [that is, $|F(0)|, |G(0)| \le p/2$] that

$$|F(x,y)-G(x,y)| \ll U + p \ll p^{1-1/s}H^{k/s} + p,$$

which together with (9) implies that

$$F(x,y) = G(x,y) + zp \tag{10}$$

for some integer $z \ll p^{-1/s}H^{k/s} + 1$.

Clearly, for any integer z the reducibility of $F(X,Y) - G(X,Y) - pz$ over $\mathbb{C}$ implies the reducibility of $F(X,Y)-G(X,Y)$ over $\overline{\mathbb{F}}_p$, or equivalently $f(X)g(Y) - \lambda f(Y)g(X)$ over $\overline{\mathbb{F}}_p$, which is impossible because $\lambda\notin\Lambda$.

Because $F(X, Y) - G(X, Y) - pz \in \mathbb{C}[X, Y]$ is irreducible over $\mathbb{C}$ and has degree d, we derive from Lemma 4 that for every z the Eq. (10) has at most $H^{1/(d+e)+o(1)}$ solutions. Thus the congruence (6) has at most $O\left(H^{1/(d+e)+o(1)}\left(p^{-1/s}H^{k/s}+1\right)\right)$ solutions. This, together with (7), yields the inequality

$$\frac{r^2}{2T} \ll H^{1/(d+e)+o(1)}\left(p^{-1/s}H^{k/s}+1\right),$$

and concludes the proof. □

Clearly, in the case when $e = 0$, that is, $\psi = f$ is a polynomial of degree $d \geq 2$, the bound of Theorem 7 takes form

$$N_\psi(\mathcal{I}, \mathcal{G}) \ll \left(1 + H^{(d+1)/4}p^{-1/4d}\right)H^{1/2d+o(1)}T^{1/2}.$$

4 Comments

Clearly Theorem 7 also provides a bound for the case where rational function $\psi = \varphi^s$, with $\varphi \in \overline{\mathbb{F}}_p(X)$. This comes from the fact that

$$\psi(x) \in \mathcal{G} \implies \varphi(x) \in \mathcal{G}_0,$$

where $\mathcal{G}_0$ is a multiplicative subgroup of $\overline{\mathbb{F}}_p$ of order bounded by sT. However the resulting bound depends now on the degrees of the polynomials associated with φ rather than that of ψ.

Another consequence from Theorem 7 is the following: given an interval $\mathcal{I}$ and a subgroup $\mathcal{G} \subseteq \mathbb{F}_p^*$, satisfying $N_\psi(\mathcal{I}, \mathcal{G}) = \#\mathcal{I}$ then

$$\#\mathcal{G} \gg \min\{(\#\mathcal{I})^{2-2\tau+o(1)}, (\#\mathcal{I})^{1-2\rho-2\tau+o(1)}p^{2\vartheta}\}$$

where the implied constant depends only on d and e. However, we believe that this bound is very unlikely to be tight.

Acknowledgments D. G-P. would like to thank Macquarie University for the support and hospitality during his stay in Australia. During the preparation of this paper D. G-P. was supported by the Ministerio de Economia y Competitividad project TIN2011-27479-C04-04 and I. S. by the Australian Research Council Grants DP130100237 and DP140100118.

References

1. Bodin, A.: Reducibility of rational functions in several variables. Isr. J. Math. **164**, 333–347 (2008)
2. Bombieri, E., Pila, J.: The number of integral points on arcs and ovals. Duke Math. J. **59**, 337–357 (1989)
3. Bourgain, J.: On the distribution of the residues of small multiplicative subgroups of $\mathbb{F}_p$. Isr. J. Math. **172**, 61–74 (2009)
4. Bourgain, J., Garaev, M.Z., Konyagin, S.V., Shparlinski, I.E.: On the hidden shifted power problem. SIAM J. Comput. **41**, 1524–1557 (2012)

5. Bourgain, J., Garaev, M.Z., Konyagin, S.V., Shparlinski, I.E.: Multiplicative congruences with variables from short intervals. J. d'Analyse Math. (To appear)
6. Burkhart, J.F., Calkin, N.J., Gao, S., Hyde-Volpe, J.C., James, K., Maharaj, H., Manber, S., Ruiz, J., Smith, E.: Finite field elements of high order arising from modular curve. Des. Codes Cryptogr. **51**, 301–314 (2009)
7. Chang, M.-C., Cilleruelo, J., Garaev, M.Z., Hernández, J., Shparlinski, I.E., Zumalacárregui, A.: Points on curves in small boxes and applications. Mich. Math. J. **63**, 503–534 (2014)
8. Chang, M.-C.: Order of Gauss periods in large characteristic. Taiwan. J. Math. **17**, 621–628 (2013)
9. Chang, M.-C.: Elements of large order in prime finite fields. Bull. Aust. Math. Soc. **88**, 169–176 (2013)
10. Chang, M.-C., Kerr, B., Shparlinski, I.E., Zannier, U.: Elements of large order on varieties over prime finite fields. J. Théorie Nombres Bordeaux (To appear)
11. Cilleruelo, J., Garaev, M.Z., Ostafe, A., Shparlinski, I.E.: On the concentration of points of polynomial maps and applications. Math. Z. **272**, 825–837 (2012)
12. Cilleruelo, J., Shparlinski, I.E., Zumalacárregui, A.: Isomorphism classes of elliptic curves over a finite field in some thin families. Math. Res. Lett. **19**, 335–343 (2012)
13. Gómez-Pérez, D., Gutierrez, J.: On the linear complexity and lattice test of nonlinear pseudorandom number generators. In: Applied Algebra and Number Theory, Cambridge University Press, Cambridge (2014)
14. Grötschel, M., Lovász, L., Schrijver, A.: Geometric Algorithms and Combinatorial Optimization. Springer, Berlin (1993)
15. Kerr, B.: Solutions to polynomial congruences in well shaped sets. Bull. Aust. Math. Soc. **88**, 435–447 (2013)
16. Lorenzini, D.: Reducibility of polynomials in two variables. J. Algebra **156**, 65–75 (1993)
17. Popovych, R.: Elements of high order in finite fields of the form $\mathbb{F}_q[x]/\Phi_r(x)$. Finite Fields Appl. **18**, 700–710 (2012)
18. Popovych, R.: Elements of high order in finite fields of the form $\mathbb{F}_q[x]/(x^m - a)$. Finite Fields Appl. **19**, 86–92 (2013)
19. Shparlinski, I.E.: Groups generated by iterations of polynomials over finite fields. Proc. Edinb. Math. Soc. (To appear)
20. Voloch, J.F.: On the order of points on curves over finite fields. Integers **7**(A49), 1–4 (2007)
21. Voloch, J.F.: Elements of high order on finite fields from elliptic curves. Bull. Aust. Math. Soc. **81**, 425–429 (2010)
22. Wooley, T.D.: Vinogradov's mean value theorem via efficient congruencing. Ann. Math. **175**, 1575–1627 (2012)
23. Wooley, T.D.: Vinogradov's mean value theorem via efficient congruencing, II. Duke Math. J. **162**, 673–730 (2013)
24. Wooley, T.D.: Multigrade efficient congruencing and Vinogradov's mean value theorem (2011). Preprint. arXiv:1310.8447