CONNECTED COMPONENTS OF THE GRAPH GENERATED BY POWER MAPS IN PRIME FINITE FIELDS

CARL POMERANCE AND IGOR E. SHPARLINSKI

For Jeffrey Outlaw Shallit on his 60th birthday

ABSTRACT. Consider the power pseudorandom-number generator in a finite field $\mathbb{F}_q$. That is, for some integer $e \geq 2$, one considers the sequence $u, u^e, u^{e^2}, \ldots$ in $\mathbb{F}_q$ for a given seed $u \in \mathbb{F}_q^\times$. This sequence is eventually periodic. One can consider the number of cycles that exist as the seed u varies over $\mathbb{F}_q^\times$. This is the same as the number of cycles in the functional graph of the map $x \mapsto x^e$ in $\mathbb{F}_q^\times$. We prove some estimates for the maximal and average number of cycles in the case of prime finite fields.

1. INTRODUCTION

1.1. **Set up.** For a prime power q, we use $\mathbb{F}_q$ to denote the finite field of q elements. For a fixed integer $e \geq 2$ we denote by $\mathcal{G}_{e,q}$ the functional graph of the map $x \mapsto x^e$ with vertices formed by the elements of $\mathbb{F}_q^\times$. We also denote by $N(e,q)$ the total number of cycles in $\mathcal{G}_{e,q}$. Alternatively, $N(e,q)$ can be defined as the number of connected components of $\mathcal{G}_{e,q}$ when it is considered as an undirected graph.

By a result of [3, Theorem 1] for prime fields (see also [14] for $e = 2$), which can easily be extended to arbitrary finite fields, we have

$$N(e,q) = \sum_{d \mid \rho} \frac{\varphi(d)}{\ell_e(d)}, \tag{1.1}$$

where ρ is the largest divisor of $q-1$ which is relatively prime to e and, for relatively prime integers a and b, we use $\ell_a(b)$ to denote the multiplicative order of a modulo b.

Here we are interested in the extreme and average values of $N(e,q)$ when e is fixed and q varies over primes.

We remark that under the Generalised Riemann Hypothesis, the orders $\ell_a(b)$ tend to be large (of magnitude b in a logarithmic scale); we refer to [12] in the case of primes. Hence one expects that for most primes we have $N(e,p) \leq p^{o(1)}$. On the other hand, we show that the average value of $N(e,p)$ is quite large.

1.2. **Notation.** Throughout the paper, the letters p and r always denote prime numbers while the letters a, e, k, m, and n denote positive integers.

As usual, for a positive real number x we use $\pi(x)$ to denote the number of primes $p \le x$. Furthermore, for integers a and $k \ge 1$ we define $\pi(x;k,a)$ as the number of primes $p \le x$ in the arithmetic progression $p \equiv a \pmod k$.

We also use $P(k)$ and $\varphi(k)$ to denote the largest prime divisor and the Euler function of k, respectively, with $P(1) = 1$.

We recall that the statements $U = O(V)$, $U \gg V$ and $U \ll V$ are all equivalent to the inequality $|U| \le cV$ with some constant c. In this note, implied constants may depend on the exponent e unless stated otherwise.

1.3. **New results.** First, combining a recent result of Chang [2] with a result of Harman [10], we show that $N(e,p)$ is rather large for infinitely many primes p.

Theorem 1.1. *For any fixed integer $e \ge 2$, there are infinitely many primes p with*

$$N(e,p) \ge p^{0.472+o(1)}.$$

We also show the following lower bound on the average value of $N(e,p)$.

Theorem 1.2. *For any fixed integer $e \ge 2$ and all sufficiently large real numbers x, we have*

$$\frac{1}{\pi(x)} \sum_{p \le x} N(e,p) \ge x^{0.332}.$$

2. Preliminaries

2.1. **Primes in arithmetic progressions modulo smooth integers.** The following result can be derived if one combines the bound of Chang [2, Theorem 10] on the zero-free region of L-functions of characters with smooth moduli, with a result of Harman [10, Theorem 1.2]. One only needs to check that the exponent $3/4$ in [10, Equation (1.2)] can be replaced by any constant $c < 1$, see also the remark after [9, Theorem 1.2].

Lemma 2.1. *Let $\psi(t) \downarrow 0$ as $t \to \infty$ be arbitrary. There is a real number K_ψ such that if $k \ge K_\psi$ and $P(k) \le k^{\psi(k)}$, then for all real*

numbers x with $k < x^{0.472}$ we have uniformly over integers a with $\gcd(a,k) = 1$*, that*

$$\pi(x;k,a) \gg_\psi \frac{1}{\varphi(k)}\pi(x).$$

It is useful to note that

$$0.472 = \frac{59}{125}.$$

In particular, this means that one can take $125/59 = 2.1186\ldots$ as the Linnik constant for the special moduli of [2, Theorem 12] (instead of $12/5 = 2.4$ given in [2]).

2.2. **Shifted primes with prescribed smoothness.** We also need the following result, which follows from the work of Baker and Harman [1, Theorem 1], which improves the estimate in [6]. We recall our convention that r always denotes a prime number

Lemma 2.2. *There is an absolute positive constant κ with the following property. Let $u > 10$,*

$$v = \frac{\log u}{\log_2 u}, \quad w = v^{1/0.2961},$$

and let

$$\mathcal{Q} = \{r \in [w/(\log w)^\kappa, w] \ : \ r - 1 \mid M_v\},$$

where M_v is the least common multiple of the integers in $[1, v]$. Then for u sufficiently large, we have

$$\#\mathcal{Q} \geq w/(\log w)^\kappa.$$

3. Proofs of main results

3.1. **Proof of Theorem 1.1.** We fix some integer $e \geq 2$. For each positive integer s we define m_s as the product of the first s primes and set $k_s = e^{m_s} - 1$. Since, by Mertens' theorem, we have $\varphi(m_s) \ll m_s/\log(s+1)$, factoring the polynomial $X^{m_s} - 1$ into a product of cyclotomic polynomials, we obtain

$$\log P(k_s) \ll \varphi(m_s) \ll \frac{m_s}{\log(s+1)} \ll \frac{\log k_s}{\log\log\log k_s} = o(\log k_s),$$

as $s \to \infty$. Hence, by Lemma 2.1, there exists a prime

$$p \ll k_s^{125/59} \tag{3.1}$$

with

$$p \equiv 1 \pmod{k_s}.$$

Since $\gcd(k_s, e) = 1$, we have $k_s \mid \rho$, where ρ is the part of $p-1$ coprime to e. Thus, using $\ell_e(k_s) = m_s$,

$$N(e,p) = \sum_{d \mid \rho} \frac{\varphi(d)}{\ell_e(d)} \ge \frac{\varphi(k_s)}{m_s} \gg \frac{\varphi(k_s)}{\log k_s}.$$

Using the minimal order of the Euler function, see [8, Theorem 328], we thus have

$$N(e,p) \gg \frac{k_s}{\log k_s \log\log k_s},$$

which together with (3.1) concludes the proof.

3.2. **Proof of Theorem 1.2.** We follow the construction from the proof of [13, Theorem 1] which in turn is based on some ideas of Erdős [4] .

Let x be large. For

$$u = x^{0.472}$$

we consider the set $\mathcal{Q}$ and parameters v and w as in Lemma 2.2. Put

$$m = \left\lfloor \frac{\log u}{\log w} \right\rfloor$$

and consider the set $\mathcal{S}$ of all products of m distinct primes from $\mathcal{Q}$. Clearly

$$u \ge w^m \ge d \ge (w/(\log w)^\kappa)^m = u^{1+o(1)} \tag{3.2}$$

for every $d \in \mathcal{S}$.

Furrthermore, using Lemma 2.2, an easy calculation shows that

$$\#\mathcal{S} = \binom{\#\mathcal{Q}}{m} = u^{0.7039+o(1)}. \tag{3.3}$$

For every $d \in \mathcal{S}$ we have

$$\ell_e(d) \mid M_v$$

and so by the prime number theorem, we obtain that

$$\ell_e(d) \le \exp((1+o(1))v) = u^{o(1)} = x^{o(1)}. \tag{3.4}$$

We also have $P(d) \le w = d^{o(1)}$. Recalling the choice of u and the upper bound in (3.2) we see that by Lemma 2.1 we have

$$\pi(x;d,1) \gg \frac{1}{\varphi(d)}\pi(x) = x^{1+o(1)}u^{-1}$$

for every $d \in \mathcal{S}$. Thus, using (3.3) we obtain

$$\sum_{d \in \mathcal{S}} \pi(x;d,1) \ge x^{1+o(1)}u^{-0.2961} \tag{3.5}$$

Now, let $\mathcal{P}$ be the union of all primes $p \le x$ with $d \mid p-1$ for some $d \in \mathcal{S}$. Since, by the classical bound on the divisor function, each prime $p \in \mathcal{P}$ can come from at most $x^{o(1)}$ integers $d \in \mathcal{S}$, we obtain from (3.5) that

$$\#\mathcal{P} \ge x^{1+o(1)} u^{-0.2961}. \tag{3.6}$$

For every p with $d \mid p-1$ for some $d \in \mathcal{S}$, using (3.4) and then (3.2), we have

$$N(e,p) \ge \frac{\varphi(d)}{\ell_e(d)} = d^{1+o(1)} = u^{1+o(1)}.$$

Therefore, using (3.6),

$$\sum_{p \le x} N(e,p) \ge \sum_{p \in \mathcal{P}} N(e,p) \ge u^{1+o(1)} \#\mathcal{P} \ge x^{1+o(1)} u^{0.7039}.$$

Recalling the choice of u, we conclude the proof.

4. Further comments

Hypothetically the exponents in Theorems 1.1 and 1.2 may be replaced with any fixed number smaller than 1. This is true for Theorem 1.1 on the assumption that we have exponent $1+\varepsilon$ in Linnik's theorem; that is, for each integer $k > k_0(\varepsilon)$ and residue class $a \pmod k$ coprime to k, the least prime in this residue class is smaller than $k^{1+\varepsilon}$. The proof that $N(e,p) > p^{1-\varepsilon}$ for infinitely many primes p then follows the same lines as our proof of Theorem 1.1.

To prove a $1-\varepsilon$ analogue of Theorem 1.2 we need in addition to the strong Linnik constant as above, the conjecture that in Lemma 2.2 we may replace the number 0.2961 with ε. This conjecture of Erdős is known to follow from the Elliott–Halberstam conjecture. The proof that the average of $N(e,p)$ for $p \le x$ exceeds $x^{1-\varepsilon}$ is then the same as our proof of Theorem 1.2.

In [11, Theorem 2] lower bounds are given for the order of e modulo the part of $p-1$ coprime to e that translate to upper bounds for $N(e,p)$. Indeed, we have for any function $\varepsilon(p) \downarrow 0$ that $N(e,p) < p^{1/2-\varepsilon(p)}$ for almost all primes p and on the generalized Riemann Hypothesis, $N(e,p) < p^{\varepsilon(p)}$ for almost all p. (These normal-order results are in stark contrast to the above extremal and average-order results.)

One can also consider the average cycle length. For a positive integer n, let $\ell_e^*(n)$ denote the order of e modulo the prime-to-e part of n. The average cycle length is then

$$C(e,p) = \frac{1}{p-1} \sum_{d \mid p-1} \varphi(d) \ell_e^*(d).$$

Note that $\ell_e^*(p-1) = \ell_e(\rho)$, so we have

$$\frac{\varphi(p-1)}{p-1}\ell_e(\rho) \le C(e,q) \le \ell_e(\rho).$$

One then sees that results on $\ell_e(\rho)$ immediately translate to results on $C(e,p)$. So, it follows from [11, Theorem 2] that for any $\varepsilon(p) \downarrow 0$, we have that for almost all primes p, $C(e,p) > p^{1/2+\varepsilon(p)}$. Further, the average of $C(e,p)$ for $p \le x$ exceeds $x^{0.592}$ for all sufficiently large values of x. And on the Generalised Riemann Hypothesis, the average exceeds $x^{1-\epsilon}$. An upper bound for the minimal order of $C(e,p)$ follows from the proof of Theorem 1.1. In particular, we have $C(e,p) < p^{0.472+o(1)}$ for infinitely many primes p.

It would be interesting to generalize the results of this paper to arbitrary finite fields, or perhaps to consider quantities such as

$$N(e,p^k), \quad k = 1,2,\dots\ .$$

For example, we can show that for any fixed choice of e and p, for infinitely many k we have

$$N(e,p^k) > \exp(k^{c/\log\log k}), \tag{4.1}$$

where c is a positive constant. Indeed, from [5, Theorem 1] there are infinitely many positive integers m with $\lambda(m) \le (\log m)^{O(\log\log\log m)}$, where $\lambda(m)$ is the maximum order of an element in $(\mathbb{Z}/m\mathbb{Z})^\times$. Further, with an easy argument, one can insure that m is coprime to ep. Let $k = \ell_p(m) \le \lambda(m)$. We have

$$N(e,p^k) \ge \varphi(\rho)/\ell_e(\rho) \ge \varphi(\rho)/\lambda(\rho) \ge m/\lambda(m),$$

using [7, Lemma 2]. Hence $N(e,p^k) \ge m^{1+o(1)}$. The small size of $\lambda(m)$ in comparison to m implies that m is large in comparison to $\lambda(m)$. In particular, we have $m \ge \exp(\lambda(m)^{c/\log\log\lambda(m)})$ for some $c > 0$. The bound (4.1) follows using $\lambda(m) \ge k$.

Acknowledgements

The authors are very grateful to Glyn Harman for some clarifications concerning the possible relaxation of the conditions of [10, Theorem 1.2].

The first-named author was supported in part by NSF grant number DMS-1440140 at the Mathematical Sciences Research Institute. He thanks MSRI for their hospitality.

The second-named author thanks the Max Planck Institute for Mathematics, Bonn for the generous support and hospitality. He was also supported by ARC Grant DP140100118.

References

[1] R. C. Baker and G. Harman, 'Shifted primes without large prime factors', *Acta Arith.*, **83** (1998), 331–361.

[2] M.-C. Chang, 'Short character sums for composite moduli', *J. d'Analyse Math.* , **2** (2014), 1–33.

[3] W.-S. Chou and I. E. Shparlinski, 'On the cycle structure of repeated exponentiation modulo a prime', *J. Number Theory*, **107** (2004), 345–356.

[4] P. Erdős, 'On the normal number of prime factors of $p-1$ and some other related problems concerning Euler's φ-function', *Quart. J. Math. (Oxford Ser.)*, **6** (1935), 205–213.

[5] P. Erdős, C. Pomerance, and E. Schmutz, 'Carmichael's lambda function', *Acta Arith.*, **58** (1991), 363–385.

[6] J. B. Friedlander, 'On shifted primes without large prime factors', *Number Theory and Applications*, R. A. Mollin, ed., Kluwer NATO ASI, 1989, pp. 393–401.

[7] J. B. Friedlander, C. Pomerance, and I. E. Shparlinski, 'Period of the power generator and small values of the Carmichael function', *Math. Comp.*, **70** (2001), 1591–1605. 'Corrigendum', *op. cit.*, **71** (2002), 1803–1806.

[8] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, Oxford Univ. Press, Oxford, 1979.

[9] G. Harman, 'On the number of Carmichael numbers up to x', *Bull. London Math. Soc.*, **37** (2005), 641–650.

[10] G. Harman, 'Watt's mean value theorem and Carmichael numbers', *Int. J. Number Theory*, **4** (2008), no. 2, 241–248.

[11] P. Kurlberg and C. Pomerance, 'On the period of the linear congruential and power generators', *Acta Arith.*, **119** (2005), 149–169.

[12] S. Li and C. Pomerance, 'On generalizing Artin's conjecture on primitive roots to composite moduli', *J. Reine Angew. Math.*, **556** (2003), 205–224.

[13] C. Pomerance and I. E. Shparlinski, 'Rank statistics for a family of elliptic curves over a function field', *Pure and Applied Mathem. Quart.*, **6** (2010), 21–40.

[14] T. Vasiga and J. O. Shallit, 'On the iteration of certain quadratic maps over GF(p)', *Discr. Math.*, **277** (2004), 219–240.

Mathematics Department, Dartmouth College, Hanover, NH 03755, USA

E-mail address: carl.pomerance@dartmouth.edu

Department of Pure Mathematics, University of New South Wales, Sydney, NSW 2052, Australia

E-mail address: igor.shparlinski@unsw.edu.au