Max-Planck-Institut für Mathematik
Bonn

Periodic structure of the exponential pseudorandom number generator

by

Jonas Kaszián
Pieter Moree
Igor E. Shparlinski

Max-Planck-Institut für Mathematik
Preprint Series 2013 (58)

Periodic structure of the exponential pseudorandom number generator

Jonas Kaszián
Pieter Moree
Igor E. Shparlinski

Max-Planck-Institut für Mathematik
Vivatsgasse 7
53111 Bonn
Germany

Department of Mathematics
RWTH Aachen
52056 Aachen
Germany

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

MPIM 13-58

PERIODIC STRUCTURE OF THE EXPONENTIAL PSEUDORANDOM NUMBER GENERATOR

JONAS KASZIÁN, PIETER MOREE, AND IGOR E. SHPARLINSKI

Abstract. We investigate the periodic structure of the exponential pseudorandom number generator obtained from the map $x \mapsto g^x \pmod p$ that acts on the set $\{1, \dots, p-1\}$.

1. Introduction

1.1. Motivation and our results.

Given a prime p and an integer g with $p \nmid g$ and an initial value $u_0 \in \{1, \dots, p-1\}$ we consider the sequence $\{u_n\}$ generated recursively by

$$u_n \equiv g^{u_{n-1}} \pmod p, \quad 1 \le u_n \le p-1, \qquad n = 1, 2, \dots, \tag{1}$$

and then, for an integer parameter $k \ge 1$, we consider the sequence of integers $\xi_n^{(k)} \in \{0, \dots, 2^k - 1\}$ formed by the k least significant bits of u_n, $n = 0, 1, \dots$. This construction is called the *exponential pseudorandom number generator* and has numerous cryptographic applications, see [13, 16, 19, 26, 28, 30] and references therein. Certainly, for the exponential pseudorandom number generator, as for any other pseudorandom number generator, the question of periodicity is of primal interest.

More precisely, the sequence $\{u_n\}$, as any other sequence generated iterations of a function on a finite set, becomes eventually periodic with some *cycle length* t. That is, there is some integer $s \ge 0$ such that

$$u_n = u_{n+t}, \qquad n = s, s+1, \dots. \tag{2}$$

We always assume that t is the smallest positive integer with this property. Furthermore, the sequence $u_0, \dots, u_{s+t-1}$ of length $\ell = s+t$, where $t \ge 1$ and then $s \ge 0$ are chosen to be the smallest possible integers to satisfy (2), is called the *trajectory* of $\{u_n\}$ and consists of the *tail* $u_0, \dots, u_{s-1}$ and the *cycle* $u_s, \dots, u_{s+t-1}$.

Clearly, we always have $\ell \le T$ where T is the multiplicative order of g modulo p.

2010 *Mathematics Subject Classification.* 11K45, 11T71, 94A60.

Key words and phrases. finite field, exponential map, exponential pseudorandom number generator.

Since the sequence $\{u_n\}$ becomes eventually periodic with some cycle length t, so does the sequence $\{\xi_n^{(k)}\}$ and its cycle length τ_k divides t.

We further remark that if g is a primitive root modulo p, then the map $x \mapsto g^x \pmod p$ acts bijectively on the set $\{1, \ldots, p-1\}$ or in other words defines an element of the symmetric group S_{p-1}. Therefore, in this case the sequence $\{u_n\}$ is purely periodic, that is, (2) holds with $s = 0$. This also means that in this case the sequence $\{\xi_n^{(k)}\}$ is purely periodic.

As usual let φ denote Euler's totient function. Recall that there are exactly $\varphi(p-1)$ primitive roots modulo p. The above map leads to precisely $\varphi(p-1)$ different elements of S_{p-1}. The question is to what extent these $\varphi(p-1)$ permutations represent 'generic permutations of S_{p-1}'. Note that the cardinality $(p-1)!$ of S_{p-1} is vastly larger than $\varphi(p-1)$ which on average behaves as a constant times p.

Unfortunately there are essentially no theoretic results about the behaviour of either of the sequences $\{u_n\}$ and $\{\xi_n^{(k)}\}$. In fact even the distribution of t has not been properly investigated. If g is a primitive root, which is the most interesting case for cryptographic applications, then heuristically, the periodic behaviour of the sequence $\{u_n\}$ can be modelled as a random permutation on the set $\{1, \ldots, p-1\}$, see [1] for a wealth of results about random permutations. For example, by a result of [29] one expects that $t = p^{1+o(1)}$ in this case. If g is not a primitive root it is not clear what the correct statistical model describing the map $x \mapsto g^x \pmod p$ should be. Probably, if g is of order T modulo p, then one can further reduce the residue $g^x \pmod p$ modulo T and consider the associated permutation on the set $\{1, \ldots, T\}$ generated by the map

$$x \mapsto (g^x \pmod p) \pmod T.$$

This suggests that in this case one expects $t = T^{1+o(1)}$, but the sequence $\{u_n\}$ is not necessary purely periodic anymore.

For the sequence $\{\xi_n^{(k)}\}$ it is probably natural to expect that $\tau_k = t$ in the overwhelming majority of the cases (and for a wide range of values of k), but this question has not been properly addressed in the literature.

The only theoretic result here seems to be the bound of [15] relating t and τ_k. First, as in [15, Section 5] we note that there are at most $p2^{-k} + 1$ integers $v \in \{1, \ldots, p-1\}$ with a given string of k least significant bits. Hence, if $2^k < p$ then obviously

$$\tau_k \geq t2^{k-1}/p. \tag{3}$$

If $k \le (1/4-\varepsilon)r$ for any fixed $\varepsilon > 0$, where r is the bit length of p, then it is shown in [15, Section 5] that using bounds of exponential sums one can improve (3) to

$$\tau_k \ge c(\varepsilon) t 2^{2k}/p, \tag{4}$$

where $c(\varepsilon) > 0$ depends only on $\varepsilon > 0$. Clearly the bound (4) trivially implies that for $k \ge r/4$ we have

$$\tau_k \ge t p^{-1/2+o(1)}, \tag{5}$$

which however is weaker than (3) for $k \ge r/2$.

In this paper we use some results of [2] on the concentration of solutions of exponential congruences to sharpen (3), (4) and (5) for $k \ge (3/8+\varepsilon)r$.

We also use the same method to establish a lower bound for the number of distinct values in the sequence $\{\xi_n^{(k)}\}$. Finally, we also show that for large values of k the modern results on the sum-product problem (see [8]) lead to better estimates.

Our results relate τ_k and t and are meaningful only when t is sufficiently large. Since no theoretic results about large values of t are known, we study the behaviour of t empirically. Our findings are consistent with the map $x \mapsto g^x \pmod p$ having a generic cycle structure. In particular, the results of our numerical tests exhibit a reasonable agreement with those predicted for random permutations, see [1].

1.2. **Previously known results.** Here we briefly review several previously known results about the cycle structure of the map $x \mapsto g^x \pmod p$. Essentially only very short cycles, such as fixed points, succumb to the efforts of getting rigorous results.

In particular, for an integer k we denote by $N_{p,g}(k)$ the number of $u_0 \in \{1,\dots,p-1\}$ such that for the sequence (1) we have $u_k = u_0$. Note that $N_{p,g}(1)$ is the number of fixed points of the map $x \mapsto g^x \pmod p$.

The quantity $N_{p,g}(k)$ for $k = 1,2,3$ has recently been studied in [5, 6, 12, 18, 21, 22, 23, 27, 31]. Fixed points with various restrictions on u have been considered as well. For example, Cobeli and Zaharescu [12] have shown that

$$\begin{aligned}\#\{(g,u)\ :\ 1 \le g,u \le p-1,\ \gcd(u,p-1)=1,\ g^u \equiv u \pmod p\}\\ = \frac{\varphi(p-1)^2}{p-1} + O\left(\tau(p-1)p^{1/2}\log p\right),\end{aligned}$$

where $\tau(m)$ is the number of positive integer divisors of $m \ge 1$. Unfortunately, the co-primality condition $\gcd(u,p-1)=1$ is essential for the method of [12], thus that result does not immediately extend to all $u \in \{1, \dots, p-1\}$. Several more results and conjectures of similar flavour are presented by Holden and Moree [23]. Furthermore, an asymptotic formula for the average value $N_{p,g}(1)$ on average over p and all primitive roots $g \in \{1, \dots, p-1\}$, as well as, over all $g \in \{1, \dots, p-1\}$ is given by Bourgain, Konyagin and Shparlinski [5, Theorems 13 and 14]:

$$\sum_{p \le Q} \frac{1}{p-1} \sum_{\substack{g=1 \\ g \text{ primitive root}}}^{p-1} N_{p,g}(1) = (A + o(1))\pi(Q)$$

and

$$\sum_{p \le Q} \frac{1}{p-1} \sum_{g=1}^{p-1} N_{p,g}(1) = (1 + o(1))\pi(Q)$$

as $Q \to \infty$, where

$$A = \prod_{p \text{ prime}} \left(1 - \frac{1}{p(p-1)}\right) = 0.373955\dots$$

is *Artin's constant* and, as usual, $\pi(Q)$ is the number of primes $p \le Q$. It is also shown in [6, Theorem 11] that

$$\sum_{g=1}^{p-1} N_{p,g}(1) = O(p),$$

however, the conjecture by Holden and Moree [23] that

$$\sum_{g=1}^{p-1} N_{p,g}(1) = (1 + o(1))p \tag{6}$$

remains open. It is known though that

$$\sum_{g=1}^{p-1} N_{p,g}(1) \ge p + O(p^{3/4+o(1)}),$$

see [6, Equation (1.15)]. It is also shown in [6, Section 5.9] that (6) may fail only on a very thin set of primes.

It is also known that $N_{p,g}(1) \le \sqrt{2p} + 1/2$ for any $g \in \{1, \dots, p-1\}$, see [18, Theorem 2].

For $N_{p,g}(2)$, the only known result is the bound

$$N_{p,g}(2) \le C(g) \frac{p}{\log p}$$

of Glebsky and Shparlinski [18, Theorem 3], where $C(g)$ depends on g.

Finally, by [18, Theorem 3] we have

$$N_g(3) \le \frac{3}{4}p + \frac{g^{2g+1}+g+1}{4}$$

(which is certainly a very weak bound).

2. Preparations

2.1. **Density of points on exponential curves.** Let p be a prime and a, b and g integers satisfying $p \nmid abg$. Given two intervals $\mathcal{I}$ and $\mathcal{J}$, we denote by $R_{a,b,g,p}(\mathcal{I},\mathcal{J})$ the number of integer solutions of the system of congruences

$$\begin{gathered} au \equiv x \pmod p \qquad \text{and} \qquad bg^u \equiv y \pmod p, \\ (u,x,y) \in \{1,\dots,p-1\} \times \mathcal{I} \times \mathcal{J}. \end{gathered}$$

Upper bounds on $R_{1,b,g,p}(\mathcal{I},\mathcal{J})$ are given in [2, Theorems 23 and 24], which in turn improve and generalise the previous estimates of [9, 10]. We need the following straightforward generalisations of the estimates of [2, Theorems 23 and 24] to an arbitrary a with $p \nmid a$.

Lemma 1. *Suppose that $p \nmid ab$ and that T is the multiplicative order of g modulo p. Let $\mathcal{I}$ and $\mathcal{J}$ be two intervals consisting of K and L consecutive integers respectively, where $L \le T$. Then*

$$R_{a,b,g,p}(\mathcal{I},\mathcal{J}) \le \left(\frac{K}{p^{1/3}L^{1/6}} + 1\right) L^{1/2+o(1)}$$

and

$$R_{a,b,g,p}(\mathcal{I},\mathcal{J}) \le \left(\frac{K}{p^{1/8}L^{1/6}} + 1\right) L^{1/3+o(1)}.$$

For intervals $\mathcal{I}$ and $\mathcal{J}$ of the same length, we derive a more explicit form of Lemma 1:

Corollary 2. *Assume that g is of multiplicative order T modulo p and that a and b are integers such that $p \nmid ab$. Let $\mathcal{I}$ and $\mathcal{J}$ be two intervals consisting of H consecutive integers respectively, where $H \le T$. Then*

$$R_{a,b,g,p}(\mathcal{I},\mathcal{J}) \le H^{o(1)} \begin{cases} H^{1/3}, & \text{if } H \le p^{3/20}, \\ H^{7/6}p^{-1/8}, & \text{if } p^{3/20} < H \le p^{3/16}, \\ H^{1/2}, & \text{if } p^{3/16} < H \le p^{2/5}, \\ H^{4/3}p^{-1/3}, & \text{if } p^{2/5} < H. \end{cases}$$

2.2. **Sum-product problem.** For a prime p, we denote by $\mathbb{F}_p$ the finite field of p elements.

Given a set $\mathcal{A} \subseteq \mathbb{F}_p$ we define the sets

$$2\mathcal{A} = \{a_1 + a_2 \ : \ a_1, a_2 \in \mathcal{A}\} \quad \text{and} \quad \mathcal{A}^2 = \{a_1 \cdot a_2 \ : \ a_1, a_2 \in \mathcal{A}\}.$$

The celebrated result of Bourgain, Katz and Tao [4] asserts that at least one of the cardinalities $\#\left(\mathcal{A}^2\right)$ and $\#\left(2\mathcal{A}\right)$ is always large.

The current state of affairs regarding quantitative versions of this result, due to several authors, has been summarised by Bukh and Tsimerman [8] as follows:

Lemma 3. *For an arbitrary set $\mathcal{A} \subseteq \mathbb{F}_p$, we have*

$$\max\{\#\left(\mathcal{A}^2\right), \#\left(2\mathcal{A}\right)\} \\ \geq (\#\mathcal{A})^{o(1)} \begin{cases} (\#\mathcal{A})^{12/11}, & \text{if } \#\mathcal{A} \leq p^{1/2}, \\ (\#\mathcal{A})^{7/6} p^{-1/24}, & \text{if } p^{1/2} \leq \#\mathcal{A} \leq p^{35/68}, \\ (\#\mathcal{A})^{10/11} p^{1/11}, & \text{if } p^{35/68} \leq \#\mathcal{A} \leq p^{13/24}, \\ (\#\mathcal{A})^{2} p^{-1/2}, & \text{if } p^{13/24} \leq \#\mathcal{A} \leq p^{2/3}, \\ (\#\mathcal{A})^{1/2} p^{1/2}, & \text{if } \#\mathcal{A} \geq p^{2/3}. \end{cases}$$

3. Main Results

3.1. **Period length.** For any $k \leq r$ we now obtain an improvement of (3)

Theorem 4. *For any r-bit prime p and g with $p \nmid g$, we have*

$$\tau_k \geq t p^{o(1)} \begin{cases} (2^k/p)^{1/3}, & \text{if } k/r \geq 17/20, \\ 2^{7k/6} p^{-25/24}, & \text{if } 17/20 > k/r \geq 13/16, \\ (2^k/p)^{1/2}, & \text{if } 13/16 > k/r \geq 3/5, \\ 2^{4k/3} p^{-1}, & \text{if } 3/5 > k/r. \end{cases}$$

Proof. Recall that we have the divisibility $\tau_k \mid t$ and consider the sequence $u_{s\tau_k}$ for $s = 1, \ldots, t/\tau_k$. By the definition of τ_k, all these numbers end with the same string of k least significant bits. Furthermore, this is also true for $u_{s\tau_k+1} \equiv g^{u_{s\tau_k}} \pmod p$. Therefore, there are some integers $\lambda, \mu \in [0, 2^k - 1]$ so that

$$u_{s\tau_k} = 2^k v_s + \lambda \qquad \text{and} \qquad u_{s\tau_k+1} = 2^k w_s + \mu$$

for some integers $v_s, w_s \in [0, 2^{r-k} - 1]$.

Hence, defining $\alpha \in [1, p-1]$ by the congruence $\alpha 2^k \equiv 1 \pmod p$, we see that the residues modulo p of $\alpha u_{s\tau_k}$ and of $\alpha g^{u_{s\tau_k}}$ belong to some intervals of $\mathcal{I}$ and $\mathcal{J}$, respectively, of length 2^{r-k} each. Since $t \leq T$, where T is the multiplicative order of g, for these intervals $\mathcal{I}$ and $\mathcal{J}$ we have

$$t/\tau_k \le R_{\alpha,\alpha,g,p}(\mathcal{I},\mathcal{J}).$$

Using Corollary 2 with $H = 2^{r-k}$, we conclude the proof. □

Combining Theorem 4 with (4) and (5) we derive

Corollary 5. *For any r-bit prime p and g with $p \nmid g$, we have*

$$\tau_k \ge t p^{o(1)} \begin{cases} (2^k/p)^{1/3}, & \text{if } k/r \ge 17/20, \\ 2^{7k/6} p^{-25/24}, & \text{if } 17/20 > k/r \ge 13/16, \\ (2^k/p)^{1/2}, & \text{if } 13/16 > k/r \ge 3/5, \\ 2^{4k/3} p^{-1}, & \text{if } 3/5 > k/r \ge 3/8, \\ p^{-1/2}, & \text{if } 3/8 > k/r \ge 1/4, \\ 2^{2k} p^{-1}, & \text{if } 1/4 > k/r. \end{cases}$$

3.2. **The number of distinct values.** We now obtain a lower bound on the number $\nu_k(N)$ of distinct values which appear among the elements $\xi_n^{(k)}$, $n = 0, \ldots, N-1$. Let $\ell = s + t$ be the trajectory length of the sequence $\{u_n\}$, see (2).

Note that if $2^k < p$ then the following analogue of (3) holds:

$$\nu_k(N) \ge N 2^{k-1}/p. \tag{7}$$

In fact for $N = \ell = p^{1+o(1)}$ the bound (7) is asymptotically optimal as we obviously have $\nu_k(N) \le 2^k$. However for smaller values of ℓ we obtain a series of other bounds.

Theorem 6. *For any r-bit prime p and g with $p \nmid g$, we have*

$$\nu_k(N) \ge N^{1/2} p^{o(1)} \begin{cases} (2^k/p)^{1/6}, & \text{if } 1 \ge k/r \ge 17/20, \\ 2^{7k/12} p^{-25/48}, & \text{if } 17/20 > k/r \ge 13/16, \\ (2^k/p)^{1/4}, & \text{if } 13/16 > k/r \ge 3/5, \\ 2^{2k/3} p^{-1/2}, & \text{if } 3/5 > k/r, \end{cases}$$

for all $N \le \ell$.

Proof. Consider the pairs $(\xi_n^{(k)}, \xi_{n+1}^{(k)})$, $n = 0, \ldots, N-1$. Then at least one pair (λ, μ) appears at least $N/\nu_k^2(N)$ times. Since $N \le \ell < T$, where T is the multiplicative order of g, as in the proof of Theorem 4 we obtain

$$N/\nu_k^2(N) \le R_{\alpha,\alpha,g,p}(\mathcal{I},\mathcal{J})$$

for some intervals $\mathcal{I}$ and $\mathcal{J}$ of length 2^{r-k} each and some integer $\alpha \in \{1, \ldots, p-1\}$. Using Corollary 2 with $H = 2^{r-k}$, we conclude the proof. □

Using the same technique as in [15, Section 5], it is easy to show that any fixed pair (λ, μ) occurs amongst the pairs $(\xi_n^{(k)}, \xi_{n+1}^{(k)})$, $n = 0, \ldots, \ell - 1$, at most $O\left(p2^{-2k} + p^{1/2}(\log p)^2\right)$ times. So, we also have

$$N/\nu_k^2(N) = O\left(p2^{-2k} + p^{1/2}(\log p)^2\right),$$

and thus, after simple calculations, we derive the following estimate.

Corollary 7. *For any r-bit prime p and any integer g with $p \nmid g$, we have*

$$\nu_k(N) \geq N^{1/2}p^{o(1)} \begin{cases} (2^k/p)^{1/6}, & \text{if } k/r \geq 17/20, \\ 2^{7k/12}p^{-25/48}, & \text{if } 17/20 > k/r \geq 13/6, \\ (2^k/p)^{1/4}, & \text{if } 13/16 > k/r \geq 3/5, \\ 2^{2k/3}p^{-1/2}, & \text{if } 3/5 > k/r \geq 3/8, \\ p^{-1/4}, & \text{if } 3/8 > k/r \geq 1/4, \\ 2^k p^{-1/2}, & \text{if } 1/4 > k/r, \end{cases}$$

for all $N \leq \ell$.

We now obtain a different bound which is stronger than Corollary 7 in a wide range of values of k and ℓ.

Theorem 8. *For any r-bit prime p and any integer g with $p \nmid g$, we have*

$$\nu_k(N) \geq N^{o(1)} \begin{cases} N^{6/11}(2^k/p)^{1/2}, & \text{if } N \leq p^{1/2}, \\ N^{7/12}2^{k/2}p^{-13/24}, & \text{if } p^{1/2} < N \leq p^{35/68}, \\ N^{5/11}2^{k/2}p^{-9/22}, & \text{if } p^{35/68} < N \leq p^{13/24}, \\ N2^{k/2}p^{-1}, & \text{if } p^{13/24} < N \leq p^{2/3}, \\ N^{1/4}2^{k/2}p^{-1/4}, & \text{if } N > p^{2/3}, \end{cases}$$

for all $N \leq \ell$.

Proof. Consider the set $\mathcal{A} = \{u_n \ : \ n = 0, \ldots, N-1\}$. Clearly $\#\mathcal{A} = N$ as the first $N \leq \ell$ elements of the sequence $\{u_n\}$ are pairwise distinct.

Since $u_n = 2^k w_n + \xi_n^{(k)}$ for some integer $w_n \in [0, 2^{r-k}-1]$, $n = 0, 1, \ldots$, we see that

$$\#(2\mathcal{A}) \leq \nu_k^2(N)2^{r-k+1} \tag{8}$$

(even if the addition of the elements of $\mathcal{A}$ is considered in $\mathbb{Z}$ without the reduction modulo p).

Furthermore, from the definition of the sequence $\{u_n\}$ we see that

$$\mathcal{A}^2 = \{g^{a_1+a_2} \ : \ a_1, a_2 \in A\}$$

(where g^b is computed in $\mathbb{F}_p$), thus we also have

$$\#(\mathcal{A}^2) \leq \nu_k^2(N)2^{r-k+1}. \tag{9}$$

Comparing (8) and (9) with Lemma 3, we conclude the proof. □

In particular, if $N = p^{1/2+o(1)}$ then Theorem 8 improves Corollary 7 for $k \geq (41/44 + \varepsilon)r$, with arbitrary $\varepsilon > 0$.

3.3. **Frequency of values.** We now give an upper bound on the frequency $V_k(\omega)$ of a given k-bit string ω that appears in the full trajectory $\xi_n^{(k)}$, $n = 0, \ldots, \ell - 1$.

More precisely, let $\Omega_k(U)$ be the set of k-bit strings ω for which $V_k(\omega) \geq U$.

Theorem 9. *For any r-bit prime p and g with $p \nmid g$, we have*

$$\#\Omega_k(U) \leq U^{-1} p^{o(1)} \begin{cases} 2^{2k/3} p^{1/3}, & \text{if } k/r \geq 17/20, \\ 2^{k/6} p^{25/24}, & \text{if } 17/20 > k/r \geq 13/16, \\ 2^{k/2} p^{1/2}, & \text{if } 13/16 > k/r \geq 3/5, \\ 2^{-k/3} p, & \text{if } 3/5 > k/r. \end{cases}$$

Proof. Consider the pairs

$$(\xi_n^{(k)}, \xi_{n+1}^{(k)}), \qquad \xi_n^{(k)} \in \Omega_k(U),\ n = 0, \ldots, \ell - 1. \tag{10}$$

Clearly, there are

$$W = \sum_{\omega \in \Omega_k(U)} V_k(\omega) \geq \#\Omega_k(U) U$$

such pairs.

Since $\xi_{n+1}^{(k)}$ can take at most 2^k possible values, we see that at least one pair (ω, σ) of two k-bit strings occurs at least $W/2^k$ times amongst the pairs (10). Now, the same argument as used in the proof of Theorem 4 implies that

$$W/2^k \leq R_{\alpha,\alpha,g,p}(\mathcal{I}, \mathcal{J})$$

for some intervals $\mathcal{I}$ and $\mathcal{J}$ of lengths 2^{r-k} each and some integer $\alpha \in \{1, \ldots, p-1\}$. Using Corollary 2 with $H = 2^{r-k}$, we conclude the proof. □

Examining the value of U for which the bound of Theorem 9 implies that $\#\Omega_k(U) < 1$, we derive

Corollary 10. *For any r-bit prime p and g with $p \nmid g$, we have*

$$V_k(\omega) \leq p^{o(1)} \begin{cases} 2^{2k/3} p^{1/3}, & \text{if } k/r \geq 17/20, \\ 2^{k/6} p^{25/24}, & \text{if } 17/20 > k/r \geq 13/16, \\ 2^{k/2} p^{1/2}, & \text{if } 13/16 > k/r \geq 3/5, \\ 2^{-k/3} p, & \text{if } 3/5 > k/r. \end{cases}$$

4. Numerical Results on Cycles in Exponential Map

Here we present results of some numerical tests concerning the cycle structure of the permutation on the set $\{1, \ldots, p-1\}$ generated by the map $x \mapsto g^x \pmod p$.

We use $\mathcal{I}_m$ to denote the dyadic interval $\mathcal{I}_m = [2^{m-1}, 2^m - 1]$.

We test 500 pairs (p, g) of primes p and primitive roots g modulo p selected using a pseudorandom number generator separately each of the interval $p \in \mathcal{I}_{20}$ and $p \in \mathcal{I}_{22}$ and $p \in \mathcal{I}_{25}$.

We also repeat this for 60 pairs (p, g) in the larger range $p \in \mathcal{I}_{30}$.

Let $L_r(N)$ and $C(N)$ be the length of the rth longest cycle and the number of disjoint cycles in a random permutation on N symbols, respectively.

We now recall that by the classical result of Shepp and Lloyd [29] the ratios $\lambda_r(N) = L_r(N)/N$ is expected to be

$$\lambda_r(N) = G_r + o(1),$$

as $N \to \infty$, for some constants G_r, $r = 1, 2, \ldots$, explicitly given in [29] via some integral expressions. In particular, we find from [29, Table 1] that

$$G_1 = 0.624329\ldots, \quad G_2 = 0.209580\ldots, \quad G_3 = 0.088316\ldots,$$

(we note that values reported in [25] slightly deviate from those of [29], but they agree over the approximations given here). Interestingly, the constants G_r also occur when one considers the size (in terms of number of digits) of the rth largest prime factor of an integer n, see Knuth and Trabb Pardo [25]. For example, de Bruijn [7] has shown that

$$\sum_{n \le x} \log P(n) = G_1 x \log x + O(x),$$

with $P(n)$ the largest prime factor of n, thus establishing a claim by Dickman. The constant G_1 is now known as the Golomb-Dickman constant. For further information and references see the book by Finch [14, Section 5.4].

We also recall that Goncharov [20] has shown that the ratio $\gamma(N) = C(N)/\log N$, is expected to be

$$\gamma(N) = 1 + o(1) \qquad \text{as } N \to \infty.$$

The above asymptotic results can also be found in [1, Section 1.1].

In Table 1 we present the average value, over the tested primes p in each group, of the lengths of the 1st, 2nd and 3rd longest cycles normalised by dividing by the size of the set, that is, by $p-1$.

We also calculate the number of cycles for the above pairs (p, g), normalised by dividing by $\log(p-1)$, and then present the average value for each of the ranges.

Range # of (p, g)	$\mathcal{I}_{20}$ 500	$\mathcal{I}_{22}$ 500	$\mathcal{I}_{25}$ 500	$\mathcal{I}_{30}$ 60
Aver. λ_1	0.63946789	0.61508766	0.63157252	0.60441217
Aver. λ_2	0.19999487	0.21687612	0.20469932	0.21715242
Aver. λ_3	0.08646438	0.08450844	0.09092497	0.09354165
Aver. γ	1.03813497	1.03324650	1.03014896	1.05566909

TABLE 1. Numbers of connected components

We note that we have also tried to compare the length of the smallest cycle with the expected length $e^{-\gamma}\log p$ for a random permutation on $\{1, \ldots, p-1\}$, where $\gamma = 0.5772\ldots$ is the Euler-Mascheroni constant. However the results are inconclusive and require further tests and investigation.

5. Comments

It is certainly interesting to study similar questions over arbitrary finite fields, although in this case there is no canonical way to interpret field elements as integer numbers and thus to extract bits from field elements. Probably the most interesting and natural case is the case of binary fields $\mathbb{F}_{2^r}$ of 2^r elements with a sufficiently large r. First, we use the isomorphism $\mathbb{F}_{2^r} = \mathbb{F}_2(\alpha)$, where α is a root of an irreducible polynomial over $\mathbb{F}_2$ of degree r. Now we can represent each element of $\mathbb{F}_{2^r}$ as an r-dimensional binary vector of coefficients in the basis $1, \alpha, \ldots, \alpha^{r-1}$, and the bit extraction is now apparent. For example, the proof of [18, Theorem 2] can easily be adjusted to give a square-root bound for the number of fixed points (when we identify elements of $\mathbb{F}_{2^r}$ with r-dimensional binary vectors). It is also quite likely that using the results and methods of [11] one can obtain some variants of our results in these settings.

Furthermore, for cryptographic applications it is also interesting to study the relation between t and τ_k and, in particular, obtain improvements of Corollaries 7 and 10 for almost all p and almost all initial values u_0. It is quite likely that the method of [3], combined with the ideas of [2], can be used to derive such results.

Finally we note that exponential maps have also been considered modulo prime powers, see [17, 24]. Although many computational problems, such as the discrete logarithm problem, are easier modulo prime powers, the corresponding exponential pseudorandom number generator does not seem to have any immediate weaknesses.

Acknowledgements

The authors would like to thank Daniel Panario for useful discussions and references and to Arne Winterhof for a careful reading of the manuscript.

This work was finished during a very enjoyable internship of the first author and research stay of the third author at the Max Planck Institute for Mathematics, Bonn. The third author was also supported in part by ARC grants DP110100628 and DP130100237.

References

[1] R. Arratia, A. D. Barbour and S. Tavaré, *Logarithmic combinatorial structures: A probabilistic approach*, EMS Monographs in Mathematics. European Math. Soc., Zürich, 2003.

[2] J. Bourgain, M. Z. Garaev, S. V. Konyagin and I. E. Shparlinski, 'On congruences with products of variables from short intervals and applications', *Proc. Steklov Math. Inst.*, **280** (2013), 67–96.

[3] J. Bourgain, M. Z. Garaev, S. V. Konyagin and I. E. Shparlinski, 'Multiplicative congruences with variables from short intervals', *J. d'Analyse Math.*, (to appear).

[4] J. Bourgain, N. Katz and T. Tao, 'A sum product estimate in finite fields and applications', *Geom. Funct. Analysis*, **14** (2004), 27–57.

[5] J. Bourgain, S. V. Konyagin and I. E. Shparlinski, 'Product sets of rationals, multiplicative translates of subgroups in residue rings and fixed points of the discrete logarithm', *Intern. Math. Research Notices*, **2008** (2008), Article ID rnn090, 1–29 (Corrigenda *Intern. Math. Research Notices*, **2009** (2009), 3146-3147).

[6] J. Bourgain, S. V. Konyagin and I. E. Shparlinski, 'Distribution of elements of cosets of small subgroups and applications', *Intern. Math. Research Notices*, **2012** (2012), Article rnn097, 1968–2009.

[7] N. G. de Bruijn, 'On the number of positive integers $\leq x$ and free of prime factors $> y$', *Nederl. Acad. Wetensch. Proc. Ser. A.* **54** (1951), 50–60.

[8] B. Bukh and J. Tsimerman, 'Sum-product estimates for rational functions', *Proc. Lond. Math. Soc.*, **104** (2012), 1–26.

[9] T. H. Chan and I. E. Shparlinski, 'On the concentration of points on modular hyperbolas and exponential curves', *Acta Arith.*, **142** (2010), 59–66.

[10] J. Cilleruelo and M. Z. Garaev, 'Concentration of points on two and three dimensional modular hyperbolas and applications', *Geom. and Funct. Anal.*, **21** (2011), 892–904.

[11] J. Cilleruelo and I. E. Shparlinski, 'Concentration of points on curves in finite fields', *Monatsh. Math.*, **171** (2013), 315–327.

[12] C. Cobeli and A. Zaharescu, 'An exponential congruence with solutions in primitive roots', *Rev. Roumaine Math. Pures Appl.*, **44** (1999), 15–22.

[13] R. R. Farashahi, B. Schoenmakers and A. Sidorenko, 'Efficient pseudorandom generators based on the DDH assumption', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **4450**, (2007), 426–441.

[14] S. R. Finch, *Mathematical constants*, Encyclopedia of Mathematics and its Applications **94**, Cambridge University Press, Cambridge, 2003.

[15] J. B. Friedlander and I. E. Shparlinski, 'On the distribution of the power generator', *Math. Comp.*, **70** (2001), 1575–1589.

[16] R. Gennaro, 'An improved pseudo-random generator based on discrete logarithm problem', *J. Crypto.*, **18** (2006), 91–110.

[17] L. Glebsky, 'Cycles in repeated exponentiation modulo p^n', *Integers*, **13** (2013), #A66.

[18] L. Glebsky and I. E. Shparlinski, 'Short cycles in repeated exponentiation modulo a prime', *Designs, Codes and Cryptography* **56** (2010), 35–42.

[19] O. Goldreich and V. Rosen, 'On the security of modular exponentiation with application to the construction of pseudorandom generators', *J. Cryptology*, **16** (2003), 71–93.

[20] V. Goncharov, 'Du domaine d'analyse combinatoire', *Bull. Acad. Sei. USSR Ser. Mat. (Izv. Akad. Nauk SSSR)*, **8** (1944), 3–48.

[21] J. Holden, 'Fixed points and two cycles of the discrete logarithm', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **2369** (2002), 405–416.

[22] J. Holden and P. Moree, 'New conjectures and results for small cycles of the discrete logarithm', *High Primes and Misdemeanours: Lectures in Honour of the 60th Birthday of Hugh Cowie Williams*, Fields Institute Communications **41**, Amer. Math. Soc., 2004, 245–254.

[23] J. Holden and P. Moree, 'Some heuristics and and results for small cycles of the discrete logarithm', *Math. Comp.*, **75** (2006), 419–449.

[24] J. Holden and M. M. Robinson, 'Counting fixed points, two-cycles, and collisions of the discrete exponential functions using p-adic methods', *J. Aust. Math. Soc.* , **92** (2012), 163–178.

[25] D. E. Knuth and L. Trabb Pardo, 'Analysis of a simple factorization algorithm', *Theoret. Comput. Sci.*, **3** (1976), 321–348.

[26] J. C. Lagarias, 'Pseudorandom number generators in cryptography and number theory', *Proc. Symp. in Appl. Math.*, Amer. Math. Soc., Providence, RI, **42** (1990), 115–143.

[27] M. Levin, C. Pomerance and K. Soundararajan, 'Fixed points for discrete logarithms', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **6197** (2010), 6–15.

[28] S. Patel and G. S. Sundaram, 'An efficient discrete log pseudo random generator', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **1462** (1999), 35–44.

[29] L. A. Shepp and S. P. Lloyd, 'Ordered cycle lengths in a random permutation', *Trans. Amer. Math. Soc.*, **121** (1966), 340–357.

[30] H. Shi, S. Jiang and Z. Qin, 'More efficient DDH pseudorandom generators', *Des. Codes Crypto.*, **55** (2010), 45–64.

[31] W. P. Zhang, 'On a problem of Brizolis', *Pure Appl. Math.*, **11** (1995), suppl., 1–3 (in Chinese).

Department of Mathematics, RWTH Aachen, 52056 Aachen, Germany
E-mail address: jonas.kaszian@rwth-aachen.de

Max-Planck-Institut für Mathematik, Vivatsgasse 7, D-53111 Bonn, Germany
E-mail address: moree@mpim-bonn.mpg.de

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