Dynamical Systems of Non-Algebraic Origins:

Fixed Points and Orbit Lengths

Igor Shparlinski

University of New South Wales

Introduction

The title "Dynamical systems of non-algebraic origins: Fixed points and orbit lengths"

. . . is a little misleading:

> We usually work in finite fields where any function is a polynomial.

Yet, we consider functions whose definitions are as *non-algebraic* as it gets.

> Besides, we will not be able to prove much about the orbit lengths.

However we will prove some results about fixed points and also give some heuristics, numerical data and ask some questions about orbit lengths.

General conventions and observations

$\mathbb{F}_q$ always denotes a finite field of q elements.

For a prime p, we assume $\mathbb{F}_p = \{0, \ldots, p-1\}$ whose elements we freely treat as integers if we need so.

We often write $A \pmod p$ to denote that integer A gets reduced modulo p and becomes an $\mathbb{F}_p$-element.

Given a map

$$f : \mathbb{F}_p \to \mathbb{F}_p$$

any orbit $u_n = f(u_{n-1})$ starting from some initial point $u_0 \in \mathbb{F}_p$ is eventually periodic: for some $s \geq 0$ and $t \geq 1$

$$u_{n+t} = u_n, \qquad n \geq s.$$

We always assume that s and t are the smallest integers with the property and call

$$s + t \leq p, \qquad s \quad \text{and} \quad t$$

the *orbit/trajectory, tail and period/cycle* lengths, respectively

Naive models

Here are two common wisdoms

- If f looks "random enough", predict s and t via the statistics of random **maps**: *Flajolet & Odlyzko* (1990).
- If f is a "permutation", predict s and t via the statistics of random **permutations**: *Goncharov* (1944) *Shepp & Lloyd* (1966);

See also *Arratia, Barbour& Tavaré* (2003).

Sometimes these approaches give good predictions, e.g.
Pollard's factoring algorithm

Sometimes they are very misleading:
We will give some examples

Maps we are going to discuss

- <u>Fermat quotients:</u>

  $$x \mapsto q_p(x) \pmod p \quad \text{and} \quad x \mapsto Q_p(x) \pmod p$$

  where

  $$q_p(x) = \frac{x^{p-1} - 1}{p} \quad \text{and} \quad Q_p(x) = \frac{x^p - x}{p}$$

  (define $q(x) = 0$ if $p \mid x$ or in any other way).

- <u>Exponential map:</u>

  $$x \mapsto g^x \pmod p$$

  where g is a fixed element of $\mathbb{F}_p^*$ (often g is a primitive root).

- <u>Self exponential map:</u>

  $$x \mapsto x^x \pmod p$$

Motivation?

Motivation?

> **'If you need a motivation, you are not a mathematician.'**
>
> *Drew Sutherland*
>
> CIRM, Luminy, Feb., 2014

Motivation?

> **'If you need a motivation, you are not a mathematician.'**
>
> *Drew Sutherland*
>
> CIRM, Luminy, Feb., 2014

. . . but also the above maps are used in cryptography for *hasing* and *pseudorandom number generation*:

- Exponential function: PRNG, *Blum, Blum & Shub* (1986),
- Self exponential function: hashing in a variant of the DSA (Digital Singnature Algorithm), *Menezes, van Oorschot & Vanstone* (1996)
- Fermat quotients: PRNG, *Woodcock & Smart* (1998)

Results and Methods

What do we typically know about these maps?

Rigorous results are rather scarce, and usually only about the number of fixed points (in some cases also for cycles of length 2 and 3).

There are **no** theoretic results about cycles of length $t \geq 4$.

There are **no** *direct* theoretic results about the distribution of elements in the trajectories and their segments

$$\{f^n(u_0) \ : \ 1 \leq n \leq N\}. \tag{1}$$

However, for the above functions we usually have reasonable control about the distribution of the elements in the images

$$\{f(n) \ : \ M+1 \leq n \leq M+N\}. \tag{2}$$

Sometimes one can use results for (2) to say something nontrivial (but very weak) about (1).

Fermat Quotients

Fixed points

Let $f(p)$ and $F(p)$ denote the number of fixed points of $q_p(u)$ and $Q_p(u)$, respectively,

$$f(p) = \#\{u \in \{0, \ldots, p-1\} \ : \ q_p(u) = u\}$$

and

$$\begin{aligned} F(p) &= \#\{u \in \{0, \ldots, p-1\} \ : \ Q_p(u) = u\} \\ &= \#\{u \in \{1, \ldots, p-1\} \ : \ q_p(u) = 1\} + 1 \end{aligned}$$

Ostafe & Shparlinski (2011): $f(p) \ll p^{11/12+o(1)}$

Chen & Winterhof (2013): $f(p) \ll p^{5/6+o(1)}$

Both are based on some results/ideas of *Heath-Brown & Konyagin* (1999).

Fouche (1985): $F(p) \ll p^{1/2+o(1)}$

How do we deal with Fermat Quotients?

Observation 1: From

$$\left(u^{p-1} - 1\right)\left(v^{p-1} - 1\right) \equiv 0 \pmod{p}^2,$$

we obtain

$$(uv)^{p-1} - 1 \equiv u^{p-1} - 1 + v^{p-1} - 1 \pmod{p}^2,$$

or

$$q_p(uv) \equiv q_p(u) + q_p(v) \pmod{p}.$$

Observation 2: The distribution of $q_p(u)$ is easy to handle in the "full" interval $u = 0, \ldots, p^2 - 1$ as this essentially the distribution of monomials $u^{p-1} \pmod{p^2}$:

$$\Downarrow$$

$$\#\{u \in \{0, \ldots, p^2 - 1\} \ : \ q_p(u) = a\} = p - 1.$$

for any a with $\gcd(a, p) = 1$.

Bounding $F(p)$

Recall

$$\begin{aligned} F(p) &= \#\{u \in \{0,\ldots,p-1\} \ : \ Q_p(u) = u\} \\ &= \#\{u \in \{1,\ldots,p-1\} \ : \ q_p(u) = 1\} + 1 \end{aligned}$$

Let $u_1,\ldots,u_N \in \{0,\ldots,p-1\}$ are all such points. Then

$$q(u_iu_j) \equiv q_p(u_i) + q_p(u_j) \equiv 2 \pmod p.$$

Since an integer $w \geq 1$ has at most $w^{o(1)}$ divisors, we obtain

$$\begin{aligned} F(p)^2 &= M^2 \\ &\leq p^{o(1)} \#\{u \in \{0,\ldots,p^2-1\} \ : \ q_p(u) = 2\} \\ &= p^{1+o(1)}. \end{aligned}$$

Numerical results

Below we present numerical results for primes

$$p \in [50000, 200000].$$

$N(k) = \#$ of primes $p \in [50000, 200000]$ with $f(p) = k$ fixed points (note that we discard the "artificial" fixed point $u = 0$).
$\rho(k) = N(k)/N$, where $N = 12851$ is the total number of $p \in [50000, 200000]$.
$\rho_0(k) = (ek!)^{-1}$, expectation for a random map.

k	0	1	2	3	4	5	6
ρ_0	0.368	0.368	0.184	0.0613	0.0153	0.00306	0.000511
N	4770	4697	2327	844	174	36	3
ρ	0.371	0.365	0.181	0.0656	0.0135	0.00280	0.000233

Statistics of fixed points

In the above range $N(k) = 0$ for $k \geq 7$.

Orbit lengths and cyclic points

Random maps:

Flajolet & Odlyzko (1990): the expectations ρ_m and μ_m of the orbit and tail length for an m element set:

$$\frac{\rho_m}{\sqrt{m}} \sim \sqrt{\pi/2} = 1.2533\ldots, \qquad \frac{\mu_m}{\sqrt{m}} \sim \sqrt{\pi/8} = 0.62665\ldots$$

Fermat Quotients:

Consider the intervals

$$\mathcal{J}_i = [50000i, 50000(i+1)], \qquad i = 1, 2, 3.$$

and the whole interval $\mathcal{J} = [50000, 200000]$.

Randomly chosen initial value $u_0 \in [1, p-1]$.

Range	$\mathcal{J}_1$	$\mathcal{J}_2$	$\mathcal{J}_3$	$\mathcal{J}$
# of primes	4459	4256	4136	12851
$\rho/\sqrt{p}$ $\mu/\sqrt{p}$	1.2423 0.62179	1.2445 0.62200	1.2444 0.61806	1.2437 0.62066

Since the values $q_p(2)$ are of special interest, we also present similar data for $u_0 = 2$.

Range	$\mathcal{J}_1$	$\mathcal{J}_2$	$\mathcal{J}_3$	$\mathcal{J}$
# of primes	4459	4256	4136	12851
$\rho/\sqrt{p}$ $\mu/\sqrt{p}$	1.2381 0.61778	1.2507 0.63004	1.2401 0.62060	1.2429 0.62275

Flajolet & Odlyzko (1990): the expectations of the number C_m of cyclic nodes for an m element set:

$$\lim_{m\to\infty} C_m/\sqrt{m} = \sqrt{\pi/2} = 1.2533\ldots,$$

Let $C(p) = \#$ of cyclic points of the map $u \mapsto q_p(u)$ on $\{0,\ldots,p-1\}$.

Average values for $C(p)/\sqrt{p}$, for primes are from the same intervals $\mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3$ and $\mathcal{J}$:

Range	$\mathcal{J}_1$	$\mathcal{J}_2$	$\mathcal{J}_3$	$\mathcal{J}$
# of primes	4459	4256	4136	12851
$C(p)/\sqrt{p}$	1.2413	1.2527	1.23706	1.2437

It seems that the average values of all these parameters are *slightly* lower that those for random maps.

Probably more extensive tests would be welcome

Exponential function

Let $T_{p,g}(k)$ be the number of $u_0 \in \{1,\dots,p-1\}$ such that for the sequence

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

we have $u_k = u_0$.

Fixed points:

$$T_{p,g}(1) = \#\text{ of fixed points of } x \mapsto g^x \pmod p.$$

Trivially

$$T_{p,g}(1) \le \sqrt{2p} + 1/2$$

Let $x_i \equiv g^{x_i}$, $1 \le x_1 < \dots < x_T \le p-1$.

There exist $a \ne 0$ such that $x_i - x_j = a$ for

$$J \ge \frac{T(T-1)}{2(p-2)}$$

pairs (i,j). If $T = T_{p,g}(1) > \sqrt{2p} + 1/2$ then $J > 1$. Hence

$$x_j + a = x_i \equiv g^{x_i} \equiv g^{x_j + a} \equiv g^a x_j \pmod p$$

for 2 values of j. — Imposible!

Cobeli & Zaharescu (1999)

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

Unfortunately, the co-primality condition

$$\mathrm{gcd}(u,p-1)=1$$

is essential, thus that result does not immediately extend to counting all $u \in \{1,\ldots,p-1\}$.

Several more results of similar flavour are due to *Holden & Moree* (2004–2006)

Holden & Moree (2004–2006) made

Conjecture 1

$$(i) \quad \sum_{p \le Q} \frac{1}{p-1} \sum_{\substack{g=1 \\ g \text{ prim. root}}}^{p-1} T_{p,g}(1) \sim A\pi(Q);$$

$$(ii) \quad \sum_{p \le Q} \frac{1}{p-1} \sum_{g=1}^{p-1} T_{p,g}(1) \sim \pi(Q);$$

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 $\pi(Q) = \#\{p \text{ prime } : \ p \le Q\}$.

Bourgain, Konyagin and Shparlinski (2008):
Conjecture 1 holds.

Methods

The proof is based on a combination of several results obtained by a mix of techniques from

- the theory of exponential sum
- additive combinatorics

For example, one of the main results of *Bourgain, Konyagin and Shparlinski* (2008) is a nontrivial bound on the number of small fractions u/v, $1 \le |u|, |v| \le h$, which fall in a given subgroup $\mathcal{G} \subseteq \mathbb{F}_p^*$, that is, on

$$N_p(h, \mathcal{G}) = \#\{(u, v) \in \mathbb{Z}^2 \ : \ 1 \le |u|, |v| \le h, \ u/v \in \mathcal{G}\}.$$

For any fixed integer $\nu \ge 1$ and any $h \ge 1$, we have

$$N_p(h, \mathcal{G}) \le hT^{(2\nu+1)/2\nu(\nu+1)} p^{-1/2(\nu+1)+o(1)} + h^2 T^{1/\nu} p^{-1/\nu+o(1)},$$

as $p \to \infty$, where

$$T = \max\{\#\mathcal{G}, p^{1/2}\}.$$

Remark We want to beat $N_p(h, \mathcal{G}) \le \min\{h^2, h\#\mathcal{G}\}$. Now, $h\left(T^{(2\nu+1)/\nu}/p\right)^{1/2(\nu+1)}$ gives us no trouble is ν is large and $h^2 \left(T/p\right)^{1/\nu}$ is always good if, say, $T \le p^{0.99}$.

Furthermore, *Holden & Moree* (2004–2006) also made

Conjecture 2

$$\sum_{g=1}^{p-1} T_{p,g}(1) \sim p$$

In full generality, Conjecture 2 remains open.

Bourgain, Konyagin and Shparlinski (2008,2010)

(i) $p + O\left(p^{3/4+o(1)}\right) \leq \sum_{g=1}^{p-1} T_{p,g}(1) = O(p)$,

(ii) Conjecture 2 may fail only on a very thin set of primes: for at most $O(\exp(12 \log x / \log\log x))$ primes $p \leq x$.

Methods

As above plus some results about smooth numbers.

Longer cycles

Only for $k \le 3$.

Glebsky & Shparlinski (2010)

(i) $T_{p,g}(2) \le C(g)\dfrac{p}{\log p}$,

(ii) $T_{p,g}(3) \le \dfrac{3}{4}p + \dfrac{g^{2g+1}+g+1}{4}$.

Helfgott (27 June, 2014, Mathoverflow)

(i) Sketched a proof of

$$T_{p,2}(3) = o(p)$$

(ii) Asked about $T_{p,g}(4)$.

Heuristics

Traditionally the map $x \mapsto g^x \pmod p$ has been considered as a random permutation of $\{1, \ldots, p-1\}$

Kaszián, Moree & Shparlinski (2013)
Some numerical verification

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

Shepp & Lloyd (1966): It is expected that

$$\lambda_r(N) = L_r(N)/N = G_r + o(1), \qquad N \to \infty,$$

for some constants G_r, $r = 1, 2, \ldots$. In particular,

$$G_1 \approx 0.62432, \quad G_2 \approx 0.20958, \quad G_3 \approx 0.08831.$$

Goncharov (1944): It is expected to be

$$\gamma(N) = C(N)/\log N = 1 + o(1), \qquad N \to \infty.$$

p-range \# (p,g)	$[2^{19}, 2^{20}]$ 500	$[2^{21}, 2^{22}]$ 500	$[2^{24}, 2^{25}]$ 500	$[2^{29}, 2^{30}]$ 60
Av. λ_1	0.639467	0.615087	0.631572	0.604412
Av. λ_2	0.199994	0.216876	0.204699	0.217152
Av. λ_3	0.086464	0.084508	0.090924	0.093541
Av. γ	1.038134	1.033246	1.030148	1.055669

Randomness confirmed?

The values of $\lambda_{1,2,3}$ oscillate around their predictions $G_{1,2,3} = 0.62432, 0.20958, 0.08831$, but γ seems to have a consistent bias over its prediction 1.

Question 1: Will this bias eventually disappear for large ranges and/or number pairs (p, g)?
Question 2: If the bias persists, explain it.

Kaszián, Moree & Shparlinski (2013)
Comparison of 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.
...the results are inconclusive and require further tests and investigation.

Self exponential function

This function is very far away from being a permutation:

Crocker (1969) and *Somer* (1981)

$$\left\lfloor \sqrt{\frac{p-1}{2}} \right\rfloor \le \#\{x^x \pmod p \ : \ x \in \mathbb{F}_p\} \le \frac{3}{4}p + p^{1/2+o(1)}$$

Fixed points:

$$F(p) = \#\{1 \le x \le p-1 \ : \ x^x \equiv x \pmod p\}.$$

Obviously, $x = 1$ is a *trivial* fixed point: $F(p) \ge 1$.

Balog, Broughan & Shparlinski (2011)
Methods of additive combinatorics:

$$F(p) \le p^{1/3+o(1)}$$

Let

$$\mathcal{A}(N) = \{p \le N \text{ prime } : F(p) = 1\}.$$

For an integer $k \ge 2$, we define recursively $\log_k x = \log \log_{k-1} x$.

Kurlberg, Luca & Shparlinski (2013)

(i) $\#\mathcal{A}(N) \le \dfrac{N}{\log N (\log_3 N)^{\vartheta + o(1)}}$,
where

$$\vartheta = \frac{1}{\zeta(2)} - \frac{1}{2\zeta(2)^2} = \frac{6\pi^2 - 18}{\pi^4} \simeq 0.4231\ldots,$$

and where $\zeta(s)$ is the Riemann zeta-function.

(ii) Naive heuristically suggests

$$\#\mathcal{A}(N) \ge c\frac{N}{(\log N)^2}$$

(iii) Improved heuristically suggests

$$\#\mathcal{A}(N) \ge \frac{N}{(\log N)^2} \exp\left(\left(\frac{1}{\log 2} + o(1)\right) \log_3 N \log_4 N\right)$$

as $N \to \infty$.

It is very unlikely one will ever be able to distinguish between (ii) and (iii) numerically.

Method

Observe that a nontrivial fixed point corresponds to a solution of the congruence

$$x^{x-1} \equiv 1 \pmod{p}, \qquad x \in \{2, 3, \ldots, p-1\}. \tag{3}$$

We wish to show that for almost all p there is a solution to (3).

For a "small" prime $q \mid p-1$, we write $p-1 = qr$.

For $x = 1 + r(q-u)$, with $u \in \{1, \ldots, q-1\}$. Note that

$$x = 1 + r(q-u) \equiv -ru \equiv -(p-1)u/q \equiv u/q \pmod{p}$$

Hence

$$x^{x-1} \equiv (u/q)^{u(p-1)/q} \pmod{p}.$$

$$\Downarrow$$

We obtain a solution to (3) if u/q is a q-th power modulo p for some $u \in \{1, \ldots, q-1\}$.

We control

- the density of primes p for which $p-1$ has a small divisor via Brun's sieve — Easy part
- the existence of q-powers via effective Chebotarev's Density Theorem, due to *Lagarias & Odlyzko* (1977), applied to the Kummer extension

  $$\mathbb{K}_{q,n} = \mathbf{L}_q(\sqrt[q]{n/q}),$$

  where $\mathbf{L}_q = \mathbb{Q}(\zeta_q)$ is the cyclotomic extension generated by unity $\zeta_q = \exp(2\pi i/q)$.

  Main difficulty: we cannot just use one prime $q \mid p-1$ as

  $\Pr[u \in \{1,\ldots,q-1\} \ : \ u/q \pmod p$ is a qth power$]$

  is small: about $1/q$.

We have to work with several values of q at the same time. In fact with all primes q in a certain interval, dictated by:

$\boxed{\text{Brun's sieve}}$ and $\boxed{\text{Chebotarev's Density Theorem}}$

Heuristics

Note that $x = 1$ is a trivial fixed point and $x = p - 1$ is never a fixed point. So, we are only interested in $x \in \{2, \ldots, p-2\}$.

Assumption 1: The exponent $x - 1$ is "independent" of the base x

$$\Downarrow$$

If $\mathcal{G}_d^* =$ set of primitive dth roots of unity, then

$$\begin{aligned}\Pr_{x \in \mathcal{G}_d^*}[x^{x-1} \equiv 1 \pmod{p}] &= \Pr_{x \in \mathcal{G}_d^*}[d \mid x - 1] \\ &= \Pr_{x \in \{2,\ldots,p-2\}}[d \mid x - 1] = \frac{\lfloor (p-3)/d \rfloor}{p-3}\end{aligned}$$

$$\Downarrow$$

$$\Pr[x^{x-1} \not\equiv 1 \pmod{p},\ \forall\ x \in \mathcal{G}_d^*] = \left(1 - \frac{\lfloor (p-3)/d \rfloor}{p-3}\right)^{\varphi(d)}.$$

Assumption 2: Independence of the above probabilities when d ranges over divisors of $p-1$.

This suggests

$$\#\mathcal{A}(N) \sim H(N)$$

as $N \to \infty$, where

$$H(N) = \sum_{p<N} \prod_{\substack{d|p-1 \\ 2<d<p-1}} \left(1 - \frac{\lfloor (p-3)/d \rfloor}{p-3}\right)^{\varphi(d)}.$$

Some rearrangements, neglecting error terms, and hand-waving, lead us to

$$H(N) \geq \frac{N}{(\log N)^2} \exp\left((1/\log 2 + o(1)) \log_3 N \log_4 N\right)$$

Similar argument, also suggests that

$$\sum_{p \le N} F(p) = (1 + o(1))K(N)$$

where

$$K(N) = \sum_{p \le N} \sum_{\substack{d \mid p-1 \\ d > 2}} \frac{\varphi(d)}{d} = \sum_{d=3}^{N} \frac{\varphi(d)}{d} \sum_{\substack{p \le N \\ p \equiv 1 \pmod{d}}} 1.$$

Using the approximation

$$\sum_{\substack{p \le N \\ p \equiv 1 \pmod{d}}} 1 = (1 + o(1)) \frac{N}{\varphi(d) \log N},$$

it seems reasonable to expect that

$$K(N) = (1 + o(1))N.$$

Numerical results

We compare the observed data for

- $A(N)$ with $N = 100000 \cdot k$, $1 \leq k \leq 10$, with the heuristically predicted value $H(N)$.
- $G(N) = \sum_{p \leq N} F(p)$ with $N = 50000 \cdot k$, $1 \leq k \leq 9$, and compare it with $K(N)$.

N	$A(N)$	$H(N)$	Relative error
100000	567	585.6	-0.0318
200000	1007	1020.6	-0.0134
300000	1358	1421.4	-0.0446
400000	1715	1790.1	-0.0419
500000	2068	2151.8	-0.0389
600000	2404	2490.0	-0.0345
700000	2725	2826.7	-0.0360
800000	3053	3151.0	-0.0311
900000	3350	3479.5	-0.0372
1000000	3632	3796.2	-0.0433

N	$G(N)$	K(N)	Relative error
500000	465413	410686.1	0.1333
1000000	936280	831872.7	0.1255
1500000	1408964	1256499.5	0.1213
2000000	1883411	1683081.9	0.1190
2500000	2357781	2110954.9	0.1169
3000000	2832933	2539862.9	0.1154
3500000	3306597	2968852.5	0.1138
4000000	3780495	3398836.9	0.1123
4500000	4256757	3829903.3	0.1115

There seems to be a consistent negative bias in the prediction for $A(N)$ and a consistent positive bias in in the prediction for $G(N)$.

We have no satisfactory explanation of this phenomenon

Orbit length model

Question 1: Is it reasonable to model the map $\psi_p : x \mapsto x^x \pmod{p}$ as a random map?

Question 2: Do we expect that the "Birthday Paradox" will force the orbits to be of size $p^{1/2+o(1)}$ with probability exponentially close to one?

In fact, it is easy to see that the orbit of ψ_p are shorter than expected from a random map:
once $x \in \mathcal{G}$ for a multiplicative subgroup $\mathcal{G}$ of $\mathbb{F}_p^*$, then also $\psi_p(x) \in \mathcal{G}$, and the remaining part of the orbit never leaves $\mathcal{G}$.

So, the behavior of orbits of ψ_p, is ruled by two (apparently independent) factors:

- random map-like behaviour inside of a subgroup of $\mathbb{F}_p^*$ which eventually leads to a cycle formed by the "Birthday Paradox";

- reducing the size of the multiplicative subgroup where the iterations of ψ_p get locked in as they progress along the trajectory.

For example, if the initial point x_0 is not a primitive root of $\mathbb{F}_p$, this immediately puts all elements of the corresponding trajectory in a nontrivial multiplicative subgroup of $\mathbb{F}_p^*$.

<u>Question 3:</u> Develop a reliable heuristic model of the orbit length that matches numerical data below.

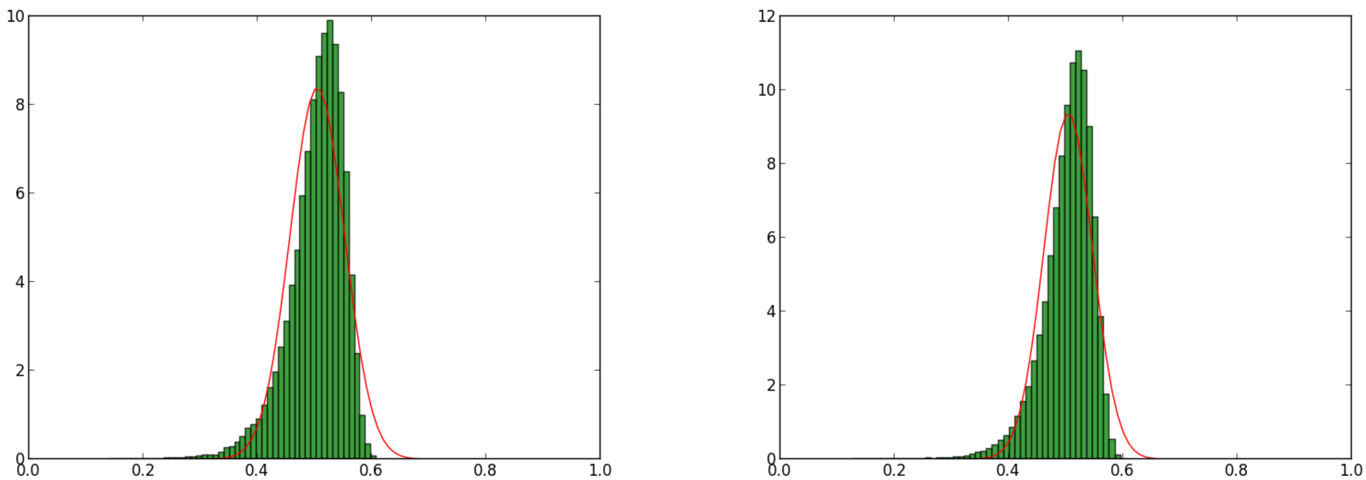
Orbit length statistics

We give histograms of the ratios $\log T_{\eta,p}(x_0)/\log p$ for various maps $\eta : \mathbb{F}_p \to \mathbb{F}_p$ over all initial points $x_0 \in \mathbb{F}_p$.

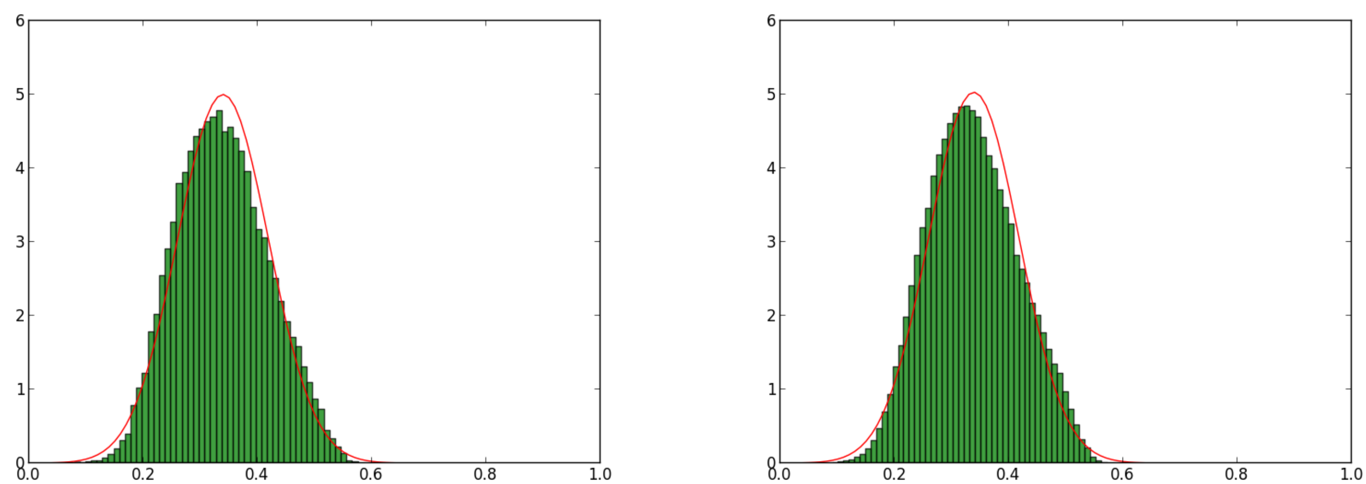
To model a random map we use $\eta(x) = x^2 + 1$ which is well known to illustrate how a random maps behave, which also forms the basis of the so-called *Pollard's rho-factorisation* algorithm.

However, the orbit sizes of ψ_p behaves very differently.

Red curves indicate normal distributions with mean and variance fitted to the data.

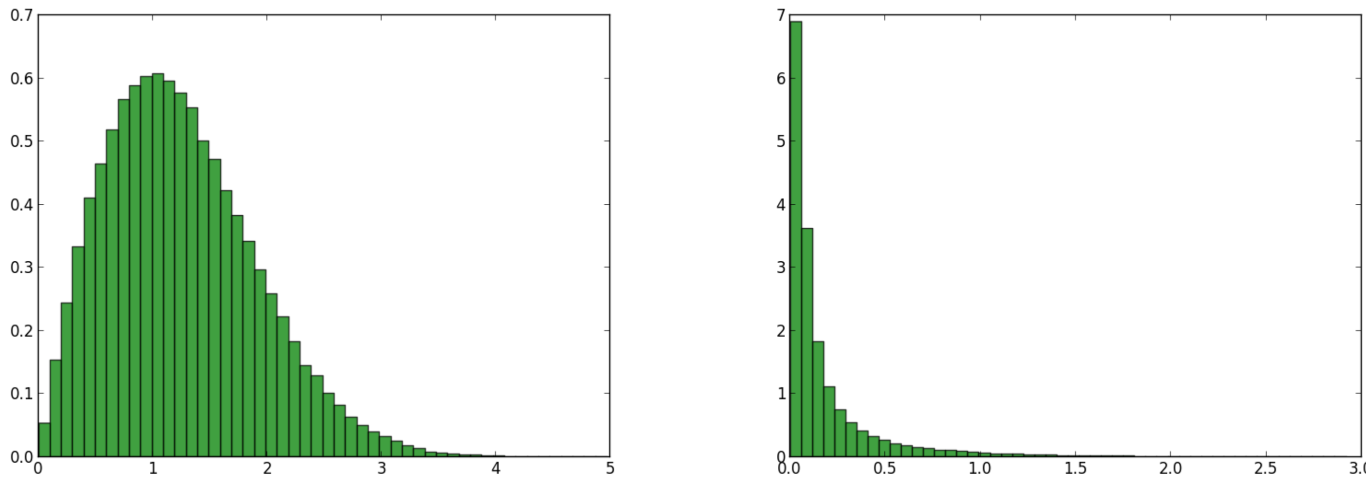


Histograms of $\log T_{\eta,p}(x_0)/\log p$ with $\eta(x) = x^2 + 1$ for $p \leq 1000000$ (left) and $p \leq 5000000$ (right)



Histograms of $\log T_{\psi_p,p}(x_0)/\log p$, $p \leq 1000000$ (left) and $p \leq 5000000$ (right).

To further show the difference in orbit statitics, it is also interesting to compare statisticics when normalized by dividing by $\sqrt{p}$



Histograms of $T_{\eta,p}(x_0)/\sqrt{p}$ with $\eta(x) = x^2+1$ (left) and $T_{\psi_p,p}(x_0)/\sqrt{p}$ (right) for $p \leq 5000000$.

Note that if $\mathbb{F}_p$ has very few subgroups, e.g. $p = 2q+1$ is a Sophie Germain Prime, ψ_p behaves like a random map.

. . . however "typical" primes have a lot of subgroups:

$$\tau(p-1) \sim (\log p)^{\log 2}$$

of all possible sizes (on a logarithmic scale).