

Algebraic properties of polynomial iterates

Alina Ostafe

Department of Computing

Macquarie University

1

Motivation

1. Better and cryptographically stronger pseudo-random number generators (PRNG) as linear constructions has been successfully attacked.

We recall that an *attack* on a PRNG is an algorithm that observes several outputs of a PRNG and then able to continue to generate the same sequence with a nontrivial probability.

2. Possible new hash functions
3. Links with traditional questions in the theory of algebraic dynamical systems such as degree growth, periods, etc.

2

General setting

p : prime number

\mathbb{F}_p : finite field of p elements.

$\mathcal{F} = \{F_1, \dots, F_m\}$: system of m polynomials/rational functions in m variables over \mathbb{F}_p

How do we iterate?

Define the k -th iteration of the functions F_i by the recurrence relation

$$F_i^{(0)} = X_i, \quad F_i^{(k)} = F_i(F_1^{(k-1)}, \dots, F_m^{(k-1)}),$$

where $i = 1, \dots, m$ and $k = 1, 2, \dots$

3

Our motivation:

Polynomial Pseudorandom Number Generators (PRNG)

Consider the PRNG defined by a recurrence relation in \mathbb{F}_p

$$u_{n+1,i} = F_i(u_{n,1}, \dots, u_{n,m}), \quad n = 0, 1, \dots,$$

with some *initial values* $u_{0,1}, \dots, u_{0,m}$, $0 \leq u_{n,i} < p$, $i = 1, \dots, m$, $n = 0, 1, \dots$

Using the vector notation

$$\mathbf{u}_n = (u_{n,1}, \dots, u_{n,m})$$

and

$$\mathbf{F} = (F_1(X_1, \dots, X_m), \dots, F_m(X_1, \dots, X_m)),$$

we have the recurrence relation

$$\mathbf{u}_{n+1} = \mathbf{F}(\mathbf{u}_n).$$

In particular, for any $n, k \geq 0$ and $i = 1, \dots, m$ we have

$$\mathbf{u}_{n+k} = \mathbf{F}^{(k)}(\mathbf{u}_n).$$

$\mathbb{F}_p =$ finite field \implies sequence of vectors (\mathbf{w}_n) is eventually periodic with some period $\tau \leq p^m$, that is, for some

$$\mathbf{u}_{n+\tau} = \mathbf{u}_n, \quad n \geq s.$$

We call

$$T = s + \tau \leq p^m$$

the *trajectory length* of the iterations of the initial vector \mathbf{u}_0 .

We assume in general that it is purely periodic, i.e.,

$$\mathbf{u}_{n+\tau} = \mathbf{u}_n, \quad n = 0, 1, \dots .$$

This is not really important, but for simplicity (mainly notational) we consider this case!

Quality of PRNG

Uniformity of distribution of PRNG



Estimating exponential sums

In our case:

$$S_{\mathbf{a}}(N) = \sum_{n=0}^{N-1} e \left(\sum_{i=1}^m a_i u_{n,i} \right),$$

where

$$e(z) = \exp(2\pi iz/p), \quad \mathbf{a} = (a_1, \dots, a_m) \in \mathbf{Z}^m.$$

Motivation: Good pseudorandom sequences should have small values of $|S_{\mathbf{a}}(N)|$!!! The smaller, the better!

Informally: Trivial bound

$$|S_{\mathbf{a}}(N)| \leq N$$

(as it is a sum on N roots of unity).

We want a **nontrivial bound**

$$|S_{\mathbf{a}}(N)| \leq N\Delta$$

with some “saving” $\Delta < 1$.

How do we estimate $|S_{\mathbf{a}}(N)|$?

Standard technique of estimating $|S_{\mathbf{a}}(N)|$ reduces it to estimating the exponential sum

$$\left| \sum_{\mathbf{v} \in \mathbb{F}_p^m} e \left(\sum_{i=1}^m a_i (F_i^{(k)}(\mathbf{v}) - F_i^{(l)}(\mathbf{v})) \right) \right|.$$

We can now try to apply [Weil](#) 1948:

$$\left| \sum_{x_1, \dots, x_m=1}^p e(F(x_1, \dots, x_m)) \right| < Dp^{m-1/2},$$

where $F \in \mathbb{F}_p[X_1, \dots, X_m]$ is a nonconstant polynomial of degree D .

Remarks:

- The degree plays an important role in the estimate of the exponential sum!
- We need linear independence of the iterations $F_i^{(k)}$, $k \geq 1$.
- In some special cases elementary methods replace the use of the Weil bound and give stronger bounds!

7

Degree growth

If f is a univariate polynomial of degree d , the degree growth

$$\deg f^{(k)} = d^k$$

is fully controlled and exponential (if $d \geq 2$).

Question: How about the multivariate case?

Clearly the growth cannot be faster than exponential, but can it be slower?

YES!!

... and for us, generally speaking, the slower the better.

What systems do we study?

Ostafe and Shparlinski 2009–2011:

$\mathcal{F} = \{F_1, \dots, F_m\}$: system of m rational functions in m variables over \mathbb{F}_p of the form:

$$F_1(X_1, \dots, X_m) = X_1^{e_1} G_1(X_2, \dots, X_m) + H_1(X_2, \dots, X_m),$$

...

$$F_{m-1}(X_1, \dots, X_m) = X_{m-1}^{e_{m-1}} G_{m-1}(X_m) + H_{m-1}(X_m),$$

$$F_m(X_1, \dots, X_m) = g_m X_m^{e_m} + h_m,$$

where

$$e_i \in \{-1, 1\}, \quad G_i, H_i \in \mathbb{F}_p[X_{i+1}, \dots, X_m], \quad g_m, h_m \in \mathbb{F}_p.$$

$e_i = 1$: G_i has a unique leading monomial,

$$G_i(X_{i+1}, \dots, X_m) = g_i X_{i+1}^{s_{i,i+1}} \dots X_m^{s_{i,m}} + \widetilde{G}_i(X_{i+1}, \dots, X_m),$$

$$g_i \neq 0, \quad \deg_{X_j} \widetilde{G}_i < s_{i,j}, \quad \deg_{X_j} H_i \leq s_{i,j}, \quad 1 \leq i < j \leq m$$

$e_i = -1$: H_i has a unique leading monomial

$$H_i(X_{i+1}, \dots, X_m) = h_i X_{i+1}^{s_{i,i+1}} \dots X_m^{s_{i,m}} + \widetilde{H}_i(X_{i+1}, \dots, X_m),$$

with

$$h_i \neq 0, \quad \deg_{X_j} \widetilde{H}_i < s_{i,j}, \quad \deg_{X_j} G_i \leq 2s_{i,j}, \quad 1 \leq i < j \leq m$$

Lemma 1 For the above systems, if $e_i = 1$,

$$F_i^{(k)} = X_i G_{i,k}(X_{i+1}, \dots, X_m) + H_{i,k}(X_{i+1}, \dots, X_m),$$

and if $e_i = -1$,

$$F_i^{(k)} = \frac{X_i R_{i,k} + S_{i,k}}{X_i R_{i,k-1} + S_{i,k-1}}, \quad i = 1, \dots, m, k = 0, 1, \dots,$$

where

$$G_{i,k}, H_{i,k}, R_{i,k}, S_{i,k} \in \mathbf{F}(X_{i+1}, \dots, X_m), \quad i = 1, \dots, m-1.$$

In both cases,

$$\deg F_i^{(k)} = \frac{1}{(m-i)!} k^{m-i} s_{i,i+1} \dots s_{m-1,m} + \psi_i(k),$$

$$\psi_i(T) \in \mathbb{Q}[T], \quad \deg \psi_i < m - i, \quad i = 1, \dots, m - 1.$$

Conclusion: The degree grows

- polynomially in the number of iterations
- monotonically (beyond a certain point).

Remark: The above effect does not occur in the univariate case.

10

Question: Why are these polynomial and rational systems important?

The above properties of the degree growth of the degrees of the iterations of \mathcal{F} has allowed to obtain rather strong results about the distribution of PRNG's, much stronger than for arbitrary polynomial generators.

Why do we win?

- we estimate the exponential sum for $e_i = 1$ for all $i = 1, \dots, m$,

$$\left| \sum_{\mathbf{v} \in \mathbb{F}_p^m} e \left(\sum_{i=1}^{m-1} a_i (F_i^{(k)}(\mathbf{v}) - F_i^{(l)}(\mathbf{v})) \right) \right|,$$

where, as before, $e(z) = \exp(2\pi iz/p)$ and

$$F_i^{(k)} - F_i^{(l)} = X_i(G_{i,k} - G_{i,l}) + H_{i,k} - H_{i,l}.$$

- the case of rational functions will follow more or less the same... but we have to take care of the poles!!!
- slow degree growth of the polynomials

Remark: Slow, but not too slow!!! ... so that $G_{i,k} - G_{i,l}$ is nontrivial for $k \neq l$

- the linearity of the polynomials F_i in X_i

Let $a_s \neq 0$ and $a_r = 0$ for $r < s$

$$\begin{aligned} & \sum_{\mathbf{v} \in \mathbb{F}_p^m} e \left(\sum_{i=1}^{m-1} a_i (F_i^{(k)}(\mathbf{v}) - F_i^{(l)}(\mathbf{v})) \right) \\ &= p^{s-1} \sum_{v_{s+1}, \dots, v_m \in \mathbb{F}_p} E(v_{s+1}, \dots, v_m) \\ & \quad \sum_{v_s \in \mathbb{F}_p} e \left(a_s v_s [G_{i,k} - G_{i,l}] (v_{s+1}, \dots, v_m) \right) \end{aligned}$$

The last sum vanishes unless

$$[G_{i,k} - G_{i,l}] (v_{s+1}, \dots, v_m) = 0.$$

Remark: We do not use the *Weil bound*. Instead we evaluate exponential sums with linear functions and estimate the number of zeros of $G_{i,k} - G_{i,l}$, $k \neq l$: any nontrivial m -variate polynomial of degree D has at most Dp^{m-1} zeros, and thus we save p instead of $p^{1/2}$ from the Weil bound.

- If the polynomials G_i are constant, then the estimate of the exponential sum does not depend on the degree anymore!

Wish list

All results at some point are based on an estimate of exponential sums with

$$L_{\mathbf{a},k,l}(\mathbf{v}) = \sum_{i=1}^{m-1} a_i \left(F_i^{(k)}(\mathbf{v}) - F_i^{(l)}(\mathbf{v}) \right)$$

We want $L_{\mathbf{a},k,l}(\mathbf{v})$ to be

- nontrivial (i.e. not to be a constant)
- exponential sums friendly:
 - (i) of small degree (to apply the Weil bound)
 - (ii) linear in some of the variables
 - (iii) to have a low dimensional locus of singularity (to apply the Deligne bound, or Deligne-like bounds by Katz, etc.)

Properties (i) and (ii) have been exploited in the above constructions, a dream project is to find a system with well-controlled property (iii) and improve previous results.

Maximal periods

Ostafe 2010: Let (\mathbf{u}_n) be defined as above with period $\tau \leq p^m$. If $e_i = 1$ for all $i = 1, \dots, m$, then $\tau = p^m$ if and only if

$$\prod_{\mathbf{v} \in \mathbb{F}_p^{m-i}} G_i(\mathbf{v}) = 1, \quad g_m = 1, \quad \deg R_i = (m-i)(p-1),$$

where

$$R_i = H_i G_i^{(2)} \dots G_i^{(p^{m-i})} \pmod{\mathcal{I}}$$

is of degree at most $p - 1$ in each variable and \mathcal{I} is the ideal generated by $X_1^p - X_1, \dots, X_m^p - X_m$.

Remark: The maximal period reduces to having maximal period for a certain linear congruential generator.

Example: $m = 2$, $p \equiv 3 \pmod{4}$,

$$F_1 = X_1(f(X_2)^2 + (p+1)/4) + a \quad \text{and} \quad F_2 = X_2 + b,$$

$a, b \in \mathbb{F}_p^*$ and $f \in \mathbb{F}_p[X]$ generates a permutation.

Question: When do the systems of rational functions defined above achieve maximal period?

Ostafe and Shparlinski 2011: The maximal period reduces to describing the maximal period of a certain linear congruential generator or a certain Möbius transformation defined by the iterations of F_i .

Define the sets

$$I_+ = \{1 \leq i \leq m : e_i = 1\}, \quad I_- = \{1 \leq i \leq m : e_i = -1\}.$$

Assume that the sequence generated by the lower $m - i$ rational functions F_{i+1}, \dots, F_m in \mathbb{F}_q^{m-i} is purely periodic with period τ_{i+1} , $i = 1, \dots, m - 1$. Then

$$F_i^{(k\tau_{i+1})}(\mathbf{u}_0) = f_i^{(k)}(u_{0,i}), \quad k \geq 1,$$

where

$$f_i(Y) = G_{i,\tau_{i+1}}(\mathbf{u}_0)Y + H_{i,\tau_{i+1}}(\mathbf{u}_0)$$

if $i \in I_+$, and f_i is the Möbius transformation

$$f_i(Y) = \frac{R_{i,\tau_{i+1}}(\mathbf{u}_0)Y + S_{i,\tau_{i+1}}(\mathbf{u}_0)}{R_{i,\tau_{i+1}-1}(\mathbf{u}_0)Y + S_{i,\tau_{i+1}-1}(\mathbf{u}_0)}$$

if $i \in I_-$.

We have maximal period when:

1. $i \in I_+, i < m,$

$$\prod_{\mathbf{v} \in \mathbb{F}_p^{m-i}} G_i(\mathbf{v}) = 1 \quad \text{and} \quad \sum_{\mathbf{v} \in \mathbb{F}_p^{m-i}} R_i(\mathbf{v}) \neq 0;$$

2. $i \in I_-, i < m,$

(a) if $R_{i,q^{m-i-1}}(\mathbf{u}_0) = 0,$ then

$$R_{i,q^{m-i}}(\mathbf{u}_0) = S_{i,q^{m-i-1}}(\mathbf{u}_0),$$

$$S_{i,q^{m-i}}(\mathbf{u}_0)S_{i,q^{m-i-1}}(\mathbf{u}_0) \neq 0;$$

(b) if $R_{i,q^{m-i-1}}(\mathbf{u}_0) \neq 0,$ then

$$X^2 - \frac{R_{i,q^{m-i}}(\mathbf{u}_0)}{R_{i,q^{m-i-1}}(\mathbf{u}_0)}X - \frac{\prod_{\mathbf{v} \in \mathbb{F}_q^{m-i}} G_1(\mathbf{v})}{R_{i,q^{m-i-1}}(\mathbf{u}_0)}$$

is a primitive polynomial over \mathbb{F}_q ;

3. if $m \in I_+,$ then $g_m = 1;$

4. if $m \in I_-,$ then $X^2 - h_m X - g_m$ is a primitive polynomial over $\mathbb{F}_q.$

Hash functions and other algebraic properties

- Hash functions from polynomial dynamical systems

The idea comes from *Charles, Goren and Lauter* 2008: Hash functions from expander graphs

n, r positive integers, p = random n -bit prime

2^r permutation polynomial systems \mathcal{F}_ℓ , $\ell = 0, \dots, 2^r - 1$, not necessary distinct

$\mathbf{w}_0 \in \mathbb{F}_p^{m+1}$ random initial vector

Input: bit string Σ of length L

- pad Σ with at most $r - 1$ zeros on the left to make sure that L is a multiple of r
- split Σ into blocks σ_j , $j = 1, \dots, L/r$, of length r and see each block as an integer $\ell \in [0, 2^r - 1]$
- Starting at the vector \mathbf{w}_0 , apply the polynomial systems \mathcal{F}_ℓ iteratively \Rightarrow sequence of vectors $\mathbf{w}_j \in \mathbb{F}_p^{m+1}$
- Output $\mathbf{w}_{L/r}$ the value of the hash function

Question: What can we say about the collision and preimage resistance?

- Can we find nontrivial lower bounds for the number of monomials of the polynomials $F_i^{(k)}$, for any $i = 1, \dots, m$, $k = 1, 2, \dots$ over finite fields?

Fuchs and Zannier 2009:

The number of monomials of the k th iterations of a univariate rational function over \mathbb{C} tends to infinity with k .

Remark: This is not necessarily the case of multivariate polynomials ...

Example: Let $m = 4$, $F_1 = X_1 - X_4(X_3X_1 + X_4X_2)$, $F_2 = X_2 + X_3(X_3X_1 + X_4X_2)$, $F_3 = X_3$, $F_4 = X_4$. Then, for any $k \geq 1$,

$$F_1^{(k)} = X_1 - kX_4(X_3X_1 + X_4X_2)$$

$$F_2^{(k)} = X_2 + kX_3(X_3X_1 + X_4X_2).$$

So, the degree growth is constant, as the number of monomials at every iteration (over any field)!!!

19

There are situations (with a bit of dirty tricks with the characteristic) when

- the degree grows exponentially,
- the number of monomials is constant, or even decreases to a monomial

... at least over finite fields.

Example: Let $m = 2$, $F_1 = X_1^p - X_2^p$, $F_2 = X_1^p + X_2^p$ over \mathbb{F}_q , $q = p^m$. Then

$$F_i^{(k)} = \begin{cases} c_{i,k} X_{j_k}^{p^k}, & k = \text{even}, j_k \in \{1, 2\} \\ d_{i,k} (X_1^{p^k} \pm X_2^{p^k}), & k = \text{odd} \end{cases},$$

where $c_{i,k}, d_{i,k} \in \mathbb{F}_q$, $i = 1, 2$.

Question: Can we achieve this without cheating?

- The *additive complexity* of the polynomials $F_i^{(k)}$, for any $i = 1, \dots, m$, $k = 1, 2, \dots$, is the smallest number of '+' signs in the formulas evaluating these polynomials.

Note that the additive complexity can be much smaller than the number of monomials.

Example:

$$f(X, Y) = (X^2 + 2Y)^{100} (3X + Y^3)^{200} + (X^{300} + Y)^{10}$$

is of total degree 3000 and thus has a very long representation via the list of coefficients. However, the additive complexity is 4 and thus has a very concise representation/evaluation.

Construct polynomial systems with exponentially degree growth, but with polynomial additive complexity growth of their iterations. It is possible...

Example: $f(X) = (x - b)^d + b$, $f^{(k)}(X) = (X - b)^{d^k} + b$, $b \in \mathbb{F}_q$, $d \geq 2$

21

Let's talk about irreducibility!!!

At some point of the argument we need to estimate the number of zeros of some linear combinations of several distinct iterates $F_i^{(k)}(X)$ of F_1, \dots, F_m .

It is natural to start with studying the same iterates. E.g.: can we say anything interesting about the polynomials $F_i^{(k)}(X)$? Are they irreducible for any k ?

NO RESULTS!



Let's start with the univariate case
... still no general results yet

... except *Gomez, Nicolas, Ostafe and Sadornil* 2011, work in progress.



Let's start with quadratic polynomials
... not too many results but there are some!!

Irreducibility of iterations

For a field \mathbb{F} , $f \in \mathbb{F}[X]$ is called **stable** if $f^{(k)}$ is irreducible for all k .

- Irreducibility is very common over \mathbb{Q} . \implies over $\mathbb{F} = \mathbb{Q}$ a “random” polynomial is expected to be stable. *Ahmadi, Luca, Ostafe and Shparlinski* 2010: proved this for quadratic polynomials.
- Over \mathbb{F}_q irreducibility is rare: prob. $\sim 1/d$ for a random polynomial of degree d . \implies We expect very few stable polynomials (recall that $\deg f^{(k)}$ grows fast).
 - *Gomez Perez and Piñera Nicolas* 2010: estimate on the number of stable quadratic polynomials for odd q .
 - *Ahmadi, Luca, Ostafe and Shparlinski* 2010: No stable quadratic polynomial over \mathbb{F}_{2^n} ; it has nothing to do with $\text{char} = 2$ as $x^2 + t$ is stable over $\mathbb{F}_2(t)$.

Stability testing of quadratic polynomials over \mathbb{F}_q

$$f(X) = aX^2 + bX + c \in \mathbb{F}_q[X], \quad a \neq 0$$

$\gamma = -b/2a$ the unique critical point of f

Critical orbit of f :

$$\text{Orb}(f) = \{f^{(n)}(\gamma) : n = 2, 3, \dots\}$$

$\exists t$ such that $f^{(t)}(\gamma) = f^{(s)}(\gamma)$ for some positive integer $s < t$.

t_f = the smallest value of t with the above condition. Then:

$$\text{Orb}(f) = \{f^{(n)}(\gamma) : n = 2, \dots, t_f\}$$

Jones and Boston 2009:

$f \in \mathbb{F}_q[X]$ is stable if and only if the *adjusted critical orbit*

$$\overline{\text{Orb}}(f) = \{-f(\gamma)\} \cup \text{Orb}(f)$$

contains no squares.

What is behind? Just *Capelli*'s Lemma:

\mathbb{K} field, $f, g \in \mathbb{K}[X]$, $\beta \in \overline{\mathbb{K}}$ any root of g . Then $g(f)$ irreducible over $\mathbb{K} \iff g$ irreducible over \mathbb{K} and $f - \beta$ irreducible over $\mathbb{K}(\beta)$.

Trivially $\#\overline{\text{Orb}}(f) \leq q$ and thus one can test $f \in \mathbb{F}_q[X]$ for stability in q steps.

Ostafe and Shparlinski 2009:

Theorem 2 *For any odd q and any stable quadratic polynomial $f \in \mathbb{F}_q[X]$ we have*

$$t_f = O\left(q^{3/4}\right).$$

Corollary 3 *For any odd q , a quadratic polynomial $f \in \mathbb{F}_q[X]$ can be tested for stability in time $q^{3/4+o(1)}$.*

25

Stability of arbitrary univariate polynomials over \mathbb{F}_q

Ahmadi, Luca, Ostafe and Shparlinski 2010: even degree polynomials of the form $F = g(f)$ where f is a quadratic polynomial, and g is any polynomial of degree d

Theorem 4 *For any odd q and any stable polynomial $F = g(f) \in \mathbb{F}_q[X]$, where $f = aX^2 + bX + c \in \mathbb{F}_q[X]$ and $g \in \mathbb{F}_q[X]$ of degree d , we have*

$$t_F = O\left(q^{1-\alpha_d}\right),$$

where

$$\alpha_d = \frac{\log 2}{2 \log(4d)}.$$

We can say even more ...

Gomez, Nicolas, Ostafe and Sadornil 2011 (work in progress):

q odd, $f \in \mathbb{F}_q[X]$ stable with leading coefficient $a \in \mathbb{F}_q^*$, $\deg f \geq 2$

f' nonconstant, $\deg f' = k$, a_{k+1} coefficient of x^{k+1} in f , γ_i roots of f' , $i = 1, \dots, k$

Then

1. if $d = \deg f$ is even,

$$\left\{ a^k \prod_{i=1}^k f^{(n)}(\gamma_i) \mid n > 1 \right\} \cup \left\{ (-1)^{\frac{d}{2}} a^k \prod_{i=1}^k f(\gamma_i) \right\}$$

contains only non squares in \mathbb{F}_q ;

2. if $d = \deg f$ is odd,

$$\left\{ (-1)^{\frac{(d-1)}{2} + k} (k+1) a_{k+1} a \prod_{i=1}^{d-1} f^{(n)}(\gamma_i) \mid n \geq 1 \right\}$$

contains only squares in \mathbb{F}_q .

Question: Do we have the other implication too?
We don't know yet ...

27

What is behind this result?

... new ideas involving resultants of polynomials and

Stickelberg 1897: $f \in \mathbb{F}_q[X]$, q odd, is a polynomial of degree at least 2 and is the product of r pairwise distinct irreducible polynomials over \mathbb{F}_q . Then $r \equiv \deg f \pmod{2}$ if and only if $\text{Disc}(f)$ is a square in \mathbb{F}_q .

Questions:

- Is there an algorithm to check a quadratic polynomial for stability in $\mathbb{Q}[X]$ in finitely many steps?
- What about $p = 2$?
- Estimate the orbit length for an arbitrary polynomial.
- Do there exist stable polynomials of even degree over a field of characteristic 2?
- Can we say something about the stability of multivariate polynomials?

Zaharescu 2005: some results about the irreducibility of iterates of multivariate polynomials, but not in the usual way, only with respect to one variable.