Polynomial Iterations: Algebraic Properties and Applications

Igor Shparlinski

Macquarie University

Sydney

1

Introduction

What do we study?

$F = fF_1, \ldots, F_m g$: system of m polynomials in m variables over a field $\mathbb{F}$ of the form. We put

$$F_i^{(0)} = X_i$$

and we study the iterations:

$$F_i^{(k)} = F_i \left(F_1^{(k \ \ 1)}, \ldots, F_m^{(k \ \ 1)} \right)$$

where $i = 1, \ldots, m$.

We discuss various **questions** about the degree grows, linear and multiplicative independence and irreducibility of these polynomial iterations.

Please note that indeed they are mainly **questions**, with very few **answers** . . .

Motivation for this talk

Both Joachim and I like polynomials

Motivation for giving this talk:

Both Joachim and I are passionate about polynomials

Both Joachim and I are passionate about iterations:

Ramenki, Moscow, 1990

Less important reasons are:

- Links with the theory of dynamical systems
- Better and cryptographically stronger pseudo-random number generators (PRNG):
- Possible new hash functions

Notation

$A \quad B$ or $B \quad A$ (I. M. Vinogradov)

m

$A = O(B)$ (E. Landau)

is more compact and easier to use and admits more informative chains like

$$A \quad B = C.$$

Now try

$$A = O(B) = C$$

Before we iterate: Polynomial decompositions

Question: Can a polynomial f over a field $\mathbb{K}$ be written as $f = g(h)$ with nontrivial polynomials g and h?

Algorithms and Charcterisation

Ritt 1922
Barton & Zippel 1976
Schinzel 1982
Giesbrecht 1988
von zur Gathen 1990
Zannier 1993
Kozen, Landau & Zippel 1996
Beardon & Ng 2000
Gutierrez & Kozen 2003
Gutierrez & Sevilla 2006
Zieve & Müller 2008

Counting

Let D_n be the number of decomposable polynomials of **composite** degree n over $\mathbb{F}_q$ nd let

$$\alpha_n = \begin{cases} 2q^{\ell+n/\ell}(1 \quad q \quad ^1), & n \neq \ell^2 \\ q^{2\ell}(1 \quad q \quad ^1), & n = \ell^2 \end{cases}$$

where ℓ is the smallest prime divisor of n.

von zur Gathen 2009
Informally: α_n is a good approximation to D_n.

E.g., if $\gcd(n, q) = 1$ then

$$jD_n \quad \alpha_n j \quad \alpha_n q \quad ^{n/3\ell^2}$$

Note:

$$0.5q^{2n^{1/2}} \quad \alpha_n \quad q^{n/2+2}.$$

8

Approximate decomposition

The following appeared on the rubbles of an attempt of a new cryptosystem

Question: Given a polynomial F *2* $\mathbb{F}_q[X]$ of degree n and an integer m 0, find polynomials f, g *2* $\mathbb{F}_q[X]$ with

$$\deg(F \quad f(g)) \quad m$$

or prove that they do not exist.

For $m = 0$, we write $F = f(g) + h$ where h *2* F_q and differentiate. Then f^0 *j* F^0.

This does not seem to work already with $m = 1$.

Degree Growth of Iterations

Multivariate vs. univariate

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

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

is fully controlled and exponential (if d 2).

Question: How about the multivariate case?

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

10

Construction with polynomial degree growth

Ostafe and Shparlinski 2009:

$F = fF_0, \ldots, F_m g$: system of $m+1$ polynomials in $m+1$ variables over $\mathbb{F}$ of the "triangular" form:

$$F_0(X_0, \ldots, X_m) = X_0 G_0(X_1, \ldots, X_m) + H_0(X_1, \ldots, X_m),$$
$$F_1(X_0, \ldots, X_m) = X_1 G_1(X_2, \ldots, X_m) + H_1(X_2, \ldots, X_m),$$
$$\ldots$$
$$F_{m\ 1}(X_0, \ldots, X_m) = X_{m\ 1} G_{m\ 1}(X_m) + H_{m\ 1}(X_m),$$
$$F_m(X_0, \ldots, X_m) = g_m X_m + h_m,$$

where

$$g_m \ 2 \ \mathbb{F} \ , \qquad h_m \ 2 \ \mathbb{F},$$

and G_i has a *unique leading monomial*:

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

which "dominates" all other terms:

$$g_i \ 6= \ 0, \quad \deg_{X_j} \widetilde{G}_i < s_{i,j}, \quad \deg_{X_j} H_i \quad s_{i,j},$$

for 0 $i < j$ m.

Lemma 1 *Let* $F_0, \ldots, F_m$ *2* $\mathbb{F}[X_0, \ldots, X_m]$ *be defined as above. Then we have*

$$F_i^{(k)} = X_i G_{i,k}(X_{i+1}, \ldots, X_m) + H_{i,k}(X_{i+1}, \ldots, X_m), \quad i = 0, \ldots, m, k = 0, 1, \ldots,$$

where

$$G_{i,k}, H_{i,k} \text{ 2 } \mathbb{F}[X_{i+1}, \ldots, X_m], \quad i = 0, 1, \ldots, m \quad 1,$$

and

$$\deg G_{i,k} = \frac{1}{(m \quad i)!} k^{m \quad i} s_{i,i+1} \cdots s_{m \quad 1,m} + \psi_i(k),$$

with

$$\psi_i(T) \text{ 2 } \mathbb{Q}[T], \ \deg \psi_i < m \quad i, \quad i = 0, 1, \ldots, m \quad 1,$$

and

$$0 \neq G_{m,k} = g_m^k \text{ 2 } \mathbb{F}.$$

Conclusion: For the polynomial systems above, 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. Thus there are no univariate analogues of our results.

Question: Why are these polynomial systems important?

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

13

Permutation Systems

For some applications it is also desirable to guarantee that the map generated by F is a permutation on $\mathbb{F}_p^{m+1}$.

We are interested in polynomial permutation systems they

- lead to better bounds on the distribution properties of the corresponding PRNG

- have some cryptographic applications

In order to get a permutation from our "triangular" polynomial systems

$$F_i = X_i G_i(X_{i+1}, \ldots, X_m) + H_1(X_{i+1}, \ldots, X_m),$$

we request that $G_i, i = 0, \ldots, m$, do not have zeros over $\mathbb{F}_p$.

Remark: "Typical" polynomial F in m 2 variables over $\mathbb{F}_p$ always has lots of zeros: *Lang and Weil* 1954, *Schmidt* 1974:

An absolutely irreducible polynomial F in m 2 variables over $\mathbb{F}_p$ always has

$$p^{m\ 1} + O\left(D^2 p^{m\ 3/2}\right)$$

zeros, where $D = \deg F$.

Dead end?

Not really, there are also "atypical" polynomials.

Example: One of the attractive choices of polynomials which would lead to a fast PRNG is

$$G_i(X_{i+1}, \ldots, X_m) = \prod_{j=1}^{m \quad i} (X_{i+j}^2 \quad a_{i,j})$$

and

$$H_i(X_{i+1}, \ldots, X_m) = b_i,$$

where $a_{i,j}$ are nonresidues modulo p and b_i are any constants in $\mathbb{F}_p$.

Even simpler, one can take

$$G_i(X_{i+1}, \ldots, X_m) = (X_{i+1}^2 \quad a_i),$$

where a_i are nonresidues.

16

Polynomial Pseudorandom Number Generators

Construction

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

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

with some *initial values*

$$w_{0,0}, \ldots, w_{0,m}, \quad i = 0, \ldots, m.$$

Using the vector notation

$$\mathbf{w}_n = (w_{n,0}, \ldots, w_{n,m})$$

and

$$\mathbf{F} = (F_0(X_0, \ldots, X_m), \ldots, F_m(X_0, \ldots, X_m)),$$

we have the recurrence relation

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

$\mathbb{F}_p$ = finite field $=)$ sequence of vectors $(\mathbf{w}_n)$ is eventually periodic with some period $\tau \quad p^{m+1}$. We always assume that it is purely periodic, i.e.,

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

We sometimes discard the last component and define the truncated vectors

$$\mathbf{u}_n = (w_{n,0}, \ldots, w_{n,m \ 1}), \quad n = 0, 1, \ldots .$$

Quality Measure of PRNG's

Discrepancy

Informally: The discrepancy can be viewed as a quantitative measure for the deviation from the uniform distribution, or, in other words, for the irregularity of the distribution.

Formal definition: Given a set of N points:

$$\Gamma = \left\{(\gamma_{n,0}, \ldots, \gamma_{n,m\ \ 1}) j n = 0, \ldots, N \quad 1\right\} \quad [0,1)^m,$$

one defines the discrepancy

$$D_N = \sup_{B\ \ [0,1)^m} \left| \frac{T_\Gamma(B)}{N} \quad \text{vol}(B) \right|,$$

$T_\Gamma(B)$ =the number of points of Γ inside the box

$$B = [\alpha_1, \beta_1) \quad \ldots \quad [\alpha_m, \beta_m) \quad [0,1)^m$$

and the supremum is taken over all such boxes.

Our case: $\gamma_{n,i} = u_{n,i}/p,\ i = 0, \ldots, m \quad 1.$

Motivation: Good pseudorandom sequences should have a small discrepancy!!!

How do we estimate D_N?

General method:

Estimate exponential sums

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

where

$$\mathbf{e}_p(z) = \exp(2\pi i z/p)$$

and $\mathbf{a} = (a_0, \ldots, a_{m-1}) \in \mathbb{F}_p^m$.

Use Erdős-Turan-Koksma inequality:

$$D_N \ll \frac{1}{L} + \frac{1}{N} \sum_{\substack{|a_0|,\ldots,|a_{m-1}| \\ a_0^2+\ldots+a_{m-1}^2>0}} \prod_{j=0}^{m-1} \frac{1}{|a_j|+1} |S_{\mathbf{a}}(N)|,$$

for any $L > 1$ (typical choice: $L = N$).

Informally: Trivial bound

$$jS_{\mathbf{a}}(N)j \quad N$$

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

Assume we have a **nontrivial bound**

$$jS_{\mathbf{a}}(N)j \quad N\Delta$$

with some "saving" $\Delta < 1$.

Then the Erdős-Turan-Koksma inequality implies that

$$D_N \quad \Delta\,(\log(1/\Delta))^m$$

(instead of the trivial $D_N \quad 1$).

That is, the same "saving" Δ is preserved with only logarithmic losses

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

Niederreiter and Shparlinski 1999:
estimating $|S_{\mathbf{a}}(N)|$ reduces to estimating the exponential sum with polynomials:

$$\left| \sum_{\mathbf{v} \in \mathbb{F}_p^{m+1}} \mathbf{e}\left(\sum_{i=0}^{m} a_i (F_i^{(k)}(\mathbf{v}) \quad F_i^{(l)}(\mathbf{v})) \right) \right|,$$

where, as before, $\mathbf{e}(z) = \exp(2\pi i z/p)$ and

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

Why do we win?

slow degree growth of the polynomials $G_{i,k}$.

Remark: Slow, but not too slow!!! ... so that $G_{i,k} \quad G_{i,l}$ is nontrivial for $k \neq l$ (holds only for $i < m$, $=$) we discard the last component).

the linearity of the polynomials F_i in X_i

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

Previous work

Nonlinear polynomial generators in residue rings or finite fields have been considered for many years. They are hard to study as the degree of iterations grows exponentially which is detrimental for all known methods of dealing with such sequences.

General cases:

Niederreiter and Shparlinski 1999 and *Niederreiter and Winterhof* 2008: nonlinear univariate polynomial generator

Gutierrez and Gomez 2001: nonlinear multivariate polynomial generators

$$D_N = O(N^{1/2} p^{m/2} (\log p)^{1/2} (\log\log p)^s)$$

Nontrivial: $N \quad p^m / \log p$

+

Results are nontrivial in microscopic ranges

Special cases:

Niederreiter and Shparlinski 2001: inversive generator ($x \mapsto a + b/x$)

$$D_N = O(N^{1/2} p^{1/4} (\log p)^s)$$

Nontrivial: $N \geq p^{1/2} (\log p)^{2s}$

Friedlander and Shparlinski 2001: power generator $x \mapsto x^e$

Gomez, Gutierrez and Shparlinski 2006: Dickson generator $x \mapsto D_e(x)$
Gutierrez and Winterhof 2007: Redei generator

+

Better results

Remark: The proofs of all these results are "custom" made and do not work for any other case. They also depend on the Weil bound of exponential sums.

General systems

Ostafe and Shparlinski 2009: polynomial generators described above, truncated vectors $\mathbf{u}_n$

$$D_N = O\left(p^{\alpha_{m,\nu}} N^{\ \beta_{m,\nu}} (\log p)^m\right)$$

where

$$\alpha_{m,\nu} = \frac{m^2 + m\nu + m}{2\nu(m+\nu)} \qquad \text{and} \qquad \beta_{m,\nu} = \frac{1}{2\nu}$$

Nontrivial: $N \quad p^{m+\varepsilon}$ for some $\varepsilon > 0$
Max. Range: N could be up to p^{m+1}.

Ostafe and Shparlinski 2009: In the same range of N, polynomial generators described above, for full vectors $\mathbf{w}_n$

Permutation systems

Ostafe 2009: For any $\varepsilon > 0$ there exists $\delta > 0$ such that almost all initial vectors $\mathbf{v}$, polynomial generators described above, truncated vectors $\mathbf{u}_n$

$$D_N(\mathbf{v}) \quad N^{\ \delta}$$

provided $N \quad (\log p)^{2+\varepsilon}$.

Ostafe and Shparlinski 2009:
In the same range of N, one can estimate the discrepancy for full vectors $\mathbf{w}_n$ (with a slightly weaker estimate)

General Comments

The method does **not** depend on the Weil bound of exponential sums— a standard tool in all previous approaches.

+

The method extends to arbitrary residue rings, where the Weil bound doesn't work which stops other approaches from being efficient in those rings.

E.g. we can get similar results in "computer friendly" residue rings modulo 2^r.

. . . but keep in mind that we still need to estimate the number of zeros of some polynomials, so the results are still a little weaker.

What is next?

to get better bounds on the number of zeros of polynomials we might need absolute irreducibility of the iterations of polynomials

Remark:The polynomials $G_{i,k} \quad G_{i,l}$ are never irreducible!

Remark: This leads to a question about the algebraic structure of polynomial iterates, which is of independent interest.

Multiplicative character sums

Similar approach also applies to multiplicative character sums with the same sequences:
Ostafe, Shparlinski and Winterhof 2010-??

However:

We need a more subtle version of Lemma 1 as the polynomials $H_{i,k}$ now matter (just the degree argument is not enough).

Optimisation of the right strategy is more difficult and we need the Weil bound (for some parameter ranges).

Multiplicative independence

For our applications, the linear independance of iterates is the keystone.

Gao 1999:
Multiplicative independence of iterations of univariate polynomoials over $\mathbb{F}_q$ (excluding monomials and a few other obvious exceptions).

Question: Are there any analogues of these results for multivariate polynomials?

Let's come back to 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 $F_0, \ldots, F_m$.

It is natural to start with studying the same iterates:

Question 1: Is the variety

$$F_0^{(k)}(X) = \ldots = F_m^{(k)}(X) = 0$$

absolutely irreducible?
Question 2: Are the polynomials $F_i^{(k)}(X)$ absolutely irreducible?
NO RESULTS!

+

Let's start with the univariate case (Q.1 = Q.2)
. . . still no results.

+

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

Stability

For a field $\mathbb{F}$, f *2* $\mathbb{F}[X]$ is called **stable** if $f^{(k)}$ irreducible for all k.
Very few known results (only for $\deg f = 2$):

> Irreducibility is very common over $\mathbb{Q}$. $=)$ over $\mathbb{Q}$ a "random" polynomial is expected to be stable. *Ahmadi, Luca, Ostafe and Shparlinski* 2009: proved this for monic and arbitrary quadratic polynomials.

Over $\mathbb{F}_q$ irreducibility is rare: prob. $1/d$ for a random polynomial of degree d. $=)$ We expect very few stable polynomials (recall that deg $f^{(k)}$ grows fast).

– *Gomez Perez and Piñera Nicolas* (2009): For odd q, there are $O(q^{14/5})$ stable polynomials over $\mathbb{F}_q$.

– *Ahamdi Luca, Ostafe and Shparlinski* 2009: No stable quadratic polynomial over $\mathbb{F}_{2^n}$.

Remark: It has nothing to do with char $= 2$ as $x^2 + t$ is stable over $\mathbb{F}_2(t)$.

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

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

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

Critical orbit of f:

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

$\exists t$ 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:

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

Jones and Boston 2009:
$f \in \mathbb{F}_q[X]$ is stable if and only if the *adjusted critical orbit*

$$\overline{\mathrm{Orb}}(f) = \{f \quad f(\gamma)\} \bigcup \mathrm{Orb}(f)$$

contains no squares.

Trivially $\#\overline{\mathrm{Orb}}(f)$ q and thus one can test f 2 $\mathbb{F}_q[X]$ for stability in q steps.

Ostafe and Shparlinski 2009:

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

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

Corollary 3 *For any odd q, a quadratic polynomial f 2 $\mathbb{F}_q[X]$ can be tested for stability in time $q^{3/4+o(1)}$.*

Questions:

What about polynomials of higher degree? Can the stability be tested in finitely many steps (even over $\mathbb{F}_q$)?

Is there an algorithm to check a quadratic polynomial for stability in $\mathbb{Q}[X]$ in finitely many steps?

Sparse Polynomials

Let

$$f(X) = \sum_{i=1}^{n} a_i X^{k_i} \; 2 \; \mathbb{R}[X]$$

Descartes Rule: f has a most n distinct real positive roots.

Khovanski, Risler, . . . 1980–. . .
Multidimensional generalisations.

Question: What about finite fields??

Waarning: Beware of $X^{q \;\; 1} \;\; 1$ and $X^{(q \;\; 1)/2} \;\; 1$.

Let

$$f(X) = \sum_{i=1}^{n} a_i X^{k_i} \in \mathbb{F}_q[X]$$

Canetti, Friedlander, Konyagin, Larsen, Lieman and Shparlinski 1999:

$$\#\{x \in \mathbb{F}_q : f(x) = 0\} \ll q^{1-1/(t-1)} D^{1/(t-1)}$$

where

$$D = \min_{1 \le i \le t} \max_{j \ne i} \gcd(k_j - k_i, q - 1).$$

Question: How tight is it??

... probably not all.

Let $N_n(q)$ be the number of sequences

$$0 \quad k_1 < \ldots < k_n \quad q \quad 2,$$

for which there exist $a_1, \ldots, a_n \; 2 \; \mathbb{F}_q$ such that

$$f(X) = \sum_{i=1}^{n} a_i X^{k_i} \; 2 \; \mathbb{F}_q[X]$$

split into linear factors over F_q.

Shparlinski 1999:

$$N_n(p) = O\left(p^{n \;\; 1} \log\log p\right)$$

Question: Obtain an analogue of this result for arbitraty fields.

Question: Any lower bounds on the degree of the splitting field of "typical" sparse polynomials.

Holy Grail: Estimate the number of irreducible sparse polynomials.