Random walks on finite fields and random polynomials

Emmanuel Breuillard, *joint work with Péter Varjú*

University of Cambridge

Münster Opening Colloquium, June 21st, 2019

Plan

1. Mixing rate of linear random walks on $\mathbb{F}_p$.

2. Irreducibility of random polynomials of large degree.

Linear congruential generator

Let p a prime number, $\mathbb{F}_p$ the finite field with p elements, and $a \in \mathbb{F}_p \setminus \{0\}$.

In 1949 D.H. Lehmer, while working on the ENIAC, suggested that successive iterations of the map

$$x \mapsto ax + 1$$

on $\mathbb{F}_p$ would produce good pseudo-random numbers.

(e.g. $p = 2^{31} - 1$, $a = 48271$, see Knuth 1969 *The art of computer programming*)

Random walks on finite fields

In 1987 Chung-Graham-Diaconis (then at Bell Labs) suggested to add some randomness and consider the Markov chain on $\mathbb{F}_p$:

$$x_{n+1} = ax_n + \varepsilon_n$$

where $\varepsilon_n = \pm 1$ are independent random variables with $Proba(\varepsilon_n = 1) = Proba(\varepsilon_n = -1) = \frac{1}{2}$ and say $x_0 = 0$.

Random walks on finite fields

In 1987 Chung-Graham-Diaconis (then at Bell Labs) suggested to add some randomness and consider the Markov chain on $\mathbb{F}_p$:

$$x_{n+1} = ax_n + \varepsilon_n$$

where $\varepsilon_n = \pm 1$ are independent random variables with $Proba(\varepsilon_n = 1) = Proba(\varepsilon_n = -1) = \frac{1}{2}$ and say $x_0 = 0$.

They asked: What time does it take for the Markov chain to equidistribute?

Random walks on finite fields

In 1987 Chung-Graham-Diaconis (then at Bell Labs) suggested to add some randomness and consider the Markov chain on $\mathbb{F}_p$:

$$x_{n+1} = ax_n + \varepsilon_n$$

where $\varepsilon_n = \pm 1$ are independent random variables with $Proba(\varepsilon_n = 1) = Proba(\varepsilon_n = -1) = \frac{1}{2}$ and say $x_0 = 0$.

They asked: What time does it take for the Markov chain to equidistribute?

Theorem (Chung-Graham-Diaconis '87)

For $a = 2$ it takes $O(\log p \log \log p)$ for the chain to equidistribute and this is sharp for Mersenne primes (i.e. $p = 2^n - 1$).

Mixing time

$\pi^{(n)} \in Proba(\mathbb{F}_p) :=$ distribution of the chain at time n.

$u :=$ the uniform probability measure on $\mathbb{F}_p$, i.e. $u(x) = \frac{1}{p} \; \forall x$.

Definition (Mixing/equidistribution time)

We define the mixing time of the Markov chain as the first time n such that

$$\|\pi^{(n)} - u\|_1 < \frac{1}{10}.$$

$\|f\|_1$ is the ℓ^1-norm $\sum_{x \in \mathbb{F}_p} |f(x)|$, in particular $\|u\|_1 = 1$.

Random walk on finite fields

Theorem (Chung-Graham-Diaconis '87)

For $a = 2$ it takes $O(\log p \log\log p)$ for the chain to equidistribute and this is sharp for Mersenne primes (i.e. $p = 2^n - 1$).

Remark: It is plausible, yet not known, that $O(\log p)$ holds for most primes p.

Random walk on finite fields

Theorem (Chung-Graham-Diaconis '87)

For $a = 2$ it takes $O(\log p \log\log p)$ for the chain to equidistribute and this is sharp for Mersenne primes (i.e. $p = 2^n - 1$).

Remark: It is plausible, yet not known, that $O(\log p)$ holds for most primes p.

Proof: Analyse the Fourier cofficients of $\pi^{(n)}$ in $\mathbb{F}_p$:

$$\|\pi^{(n)} - u\|_1^2 \leqslant p\|\pi^{(n)} - u\|_2^2 = \sum_{\xi \in \mathbb{F}_p^\times} |\widehat{\pi^{(n)}}(\xi)|^2$$

$$\widehat{\pi^{(n)}}(\xi) := \sum_{x \in \mathbb{F}_p} e^{2i\pi \frac{x\xi}{p}} \pi^{(n)}(x) = \prod_{i=0}^{n-1} \cos(2\pi \frac{2^i \xi}{p})$$

...

Random walk on finite fields

Theorem (Chung-Graham-Diaconis '87)

For $a = 2$ it takes $O(\log p \log\log p)$ for the chain to equidistribute and this is sharp for Mersenne primes (i.e. $p = 2^n - 1$).

Remark: The distribution $\pi^{(n)}$ is exactly the law of the random variable

$$P(2) \mod p$$

where $P \in \mathcal{P}_n$ is the random polynomial

$$P(X) = \sum_{i=0}^{n-1} \varepsilon_{n-i} X^i.$$

and $\mathcal{P}_n$ are the Littlewood polynomials of degree $\leqslant n-1$.

$$\mathcal{P}_n := \{P \in \mathbb{Z}[X] | \deg(P) \leqslant n-1, coeffs(P) \in \{-1, 1\}\}$$

Other values of the multiplier a

What about other values of a?, i.e. we want estimates for the mixing time of the Markov chain

$$x_{n+1} = ax_n \pm 1.$$

Other values of the multiplier a

What about other values of a?, i.e. we want estimates for the mixing time of the Markov chain

$$x_{n+1} = ax_n \pm 1.$$

Universal lower bound: For any a, mixing time $\geqslant \log_2(p)$.

Other values of the multiplier a

What about other values of a?, i.e. we want estimates for the mixing time of the Markov chain

$$x_{n+1} = ax_n \pm 1.$$

Universal lower bound: For any a, mixing time $\geqslant \log_2(p)$.

$\longrightarrow$ Indeed at most 2^n sites are visited by the chain after n steps.

Other values of the multiplier a

What about other values of a?, i.e. we want estimates for the mixing time of the Markov chain

$$x_{n+1} = ax_n \pm 1.$$

Other values of the multiplier a

What about other values of a?, i.e. we want estimates for the mixing time of the Markov chain

$$x_{n+1} = ax_n \pm 1.$$

Initial observation: When $a = 1$, mixing time $\simeq p^2$ (diffusive behavior).

Other values of the multiplier a

What about other values of a?, i.e. we want estimates for the mixing time of the Markov chain

$$x_{n+1} = ax_n \pm 1.$$

Initial observation: When $a = 1$, mixing time $\simeq p^2$ (diffusive behavior).

$\longrightarrow$ same holds when a has small multiplicative order m: mixing time is in $\Omega_m(p^2)$.

Other values of the multiplier a

Theorem (Konyagin '94)

If the multiplicative order $m(a)$ is large enough ($\geq \log p (\log\log p)^4$), then for all primes p

$$\text{mixing time } \lesssim (\log p)^2 (\log\log p)^8.$$

Other values of the multiplier a

Theorem (Konyagin '94)

If the multiplicative order $m(a)$ is large enough ($\geq \log p (\log\log p)^4$), then for all primes p

$$\text{mixing time } \lesssim (\log p)^2 (\log\log p)^8.$$

Remark: Again it is plausible that the mixing time really is in $O(\log p)$ for most primes and for all multipliers a with large enough $m(a)$. However this touches upon delicate issues $\longrightarrow$ it would *imply* the Lehmer conjecture.

Lehmer conjecture

The *Mahler measure* of a monic polynomial $P \in \mathbb{Z}[X]$ is defined as the modulus of the product of its roots located outside the unit disc, i.e.

$$M(P) := \prod_{|\theta_i|>1} |\theta_i|,$$

when

$$P(X) := \prod_{i=1}^{n} (X - \theta_i).$$

Lehmer conjecture

The *Mahler measure* of a monic polynomial $P \in \mathbb{Z}[X]$ is defined as the modulus of the product of its roots located outside the unit disc, i.e.

$$M(P) := \prod_{|\theta_i|>1} |\theta_i|,$$

when

$$P(X) := \prod_{i=1}^{n} (X - \theta_i).$$

Conjecture (Lehmer 1930's)

There is an absolute constant $\varepsilon_0 > 0$ such that for every monic polynomial $P \in \mathbb{Z}[X]$, either $M(P) = 1$ or $M(P) \geq 1 + \varepsilon_0$.

Relation with Lehmer's conjecture

Motto: putative counter-examples to Lehmer give rise (in reduction to residue fields) to values of $a \in \mathbb{F}_p$ with *slow* mixing rate.

Relation with Lehmer's conjecture

Motto: putative counter-examples to Lehmer give rise (in reduction to residue fields) to values of $a \in \mathbb{F}_p$ with *slow* mixing rate.

Easy fact (pigeon hole): If P is irreducible and $M(P) < 2$, then

$$\exists n, \exists P_1 \neq P_2 \in \mathcal{P}_n \text{ s.t. } P | P_1 - P_2.$$

Relation with Lehmer's conjecture

Motto: putative counter-examples to Lehmer give rise (in reduction to residue fields) to values of $a \in \mathbb{F}_p$ with *slow* mixing rate.

Easy fact (pigeon hole): If P is irreducible and $M(P) < 2$, then

$$\exists n, \exists P_1 \neq P_2 \in \mathcal{P}_n \text{ s.t. } P | P_1 - P_2.$$

because

$$|\{Q(\alpha) | Q \in \mathcal{P}_n\}| \lesssim M(P)^n \lesssim 2^n$$

if α is a root of P with $M(P) < 2$.

Relation with Lehmer's conjecture

Motto: putative counter-examples to Lehmer give rise (in reduction to residue fields) to values of $a \in \mathbb{F}_p$ with *slow* mixing rate.

Relation with Lehmer's conjecture

Motto: putative counter-examples to Lehmer give rise (in reduction to residue fields) to values of $a \in \mathbb{F}_p$ with *slow* mixing rate.

Say that a prime p is δ-bad if there exists $a \in \mathbb{F}_p^\times$ with $m(a) \geq (\log p)^2$ such that for some $n \geq \frac{1}{\delta} \log p$

$$|Supp(\pi_a^{(n)})| = |\{P(a) \mod p | P \in \mathcal{P}_n\}| \leqslant p^\delta.$$

Theorem (B.-Varjú '18)

The following are equivalent:

1. *There is $\delta \in (0,1)$ s.t. almost no prime is δ-bad, i.e.*

$$|\{p \leq x | p \text{ is } \delta\text{-bad}\}| = o_{x \to +\infty}(|\{p \leq x\}|).$$

2. *The Lehmer conjecture holds.*

Relation with Lehmer's conjecture

Motto: putative counter-examples to Lehmer give rise (in reduction to residue fields) to values of $a \in \mathbb{F}_p$ with *slow* mixing rate.

Say that a prime p is δ-bad if there exists $a \in \mathbb{F}_p^\times$ with $m(a) \geq (\log p)^2$ such that for some $n \geq \frac{1}{\delta} \log p$

$$|Supp(\pi_a^{(n)})| = |\{P(a) \mod p | P \in \mathcal{P}_n\}| \leqslant p^\delta.$$

Theorem (B.-Varjú '18)

The following are equivalent:

1. *There is $\delta \in (0,1)$ s.t. almost no prime is δ-bad, i.e.*

$$|\{p \leq x | p \text{ is } \delta\text{-bad}\}| = o_{x \to +\infty}(|\{p \leq x\}|).$$

2. *The Lehmer conjecture holds.*

$\rightarrow$ hence mixing in $O(\log p)$ for all a with large $m(a)$ *implies* Lehmer.

Our results for the mixing time

Theorem (Konyagin '94)

If the multiplicative order of a is large enough ($\geq \log p (\log\log p)^4$), then for all primes p

$$\text{mixing time } \lesssim (\log p)^2 (\log\log p)^8.$$

Our results for the mixing time

Theorem (Konyagin '94)

If the multiplicative order of a is large enough ($\geq \log p(\log\log p)^4$), then for all primes p

$$\text{mixing time } \lesssim (\log p)^2(\log\log p)^8.$$

Theorem 1 (B.-Varjú '19)

Let $\varepsilon > 0$. For all primes p, for at least $(1-\varepsilon)p$ values of a

$$\text{mixing time } \lesssim_\varepsilon \log p \log\log p.$$

Our results for the mixing time

Theorem (Konyagin '94)

If the multiplicative order of a is large enough ($\geq \log p (\log\log p)^4$), then for all primes p

$$\text{mixing time } \lesssim (\log p)^2 (\log\log p)^8.$$

Theorem 1 (B.-Varjú '19)

Let $\varepsilon > 0$. For all primes p, for at least $(1-\varepsilon)p$ values of a

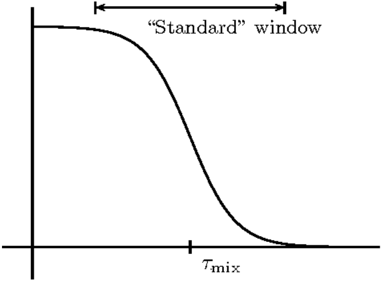
$$\text{mixing time } \lesssim_\varepsilon \log p \log\log p.$$

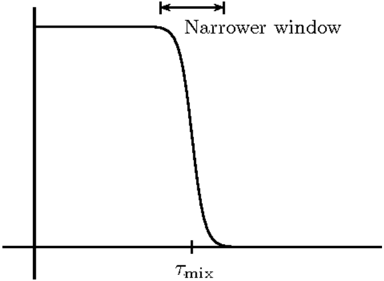
Theorem 3 (B.-Varjú '19: cut-off phenomenon)

Let $\varepsilon > 0$. Assume GRH. Then for almost all primes p, for almost all $a \in \mathbb{F}_p$,

$$\log_2(p) \leqslant \text{mixing time } \leqslant (1+\varepsilon)\log_2(p).$$

Cut-off phenomenon





y-axis: $\|\pi^{(n)} - u\|_1$
x-axis: time n

Start of proof of Thms 1 and 3

<u>Observation:</u>

$$\|\pi_a^{(n)}\|_2^2 = \mathbb{P}^{(n)}(P_1(a) = P_2(a))$$

where P_1, P_2 are independent random polynomials in $\mathcal{P}_n$.

Start of proof of Thms 1 and 3

Observation:

$$\|\pi_a^{(n)}\|_2^2 = \mathbb{P}^{(n)}(P_1(a) = P_2(a))$$

where P_1, P_2 are independent random polynomials in $\mathcal{P}_n$.

Averaging over $a \in \mathbb{F}_p$: $n \simeq \log p$

$$\begin{aligned}\mathbb{E}_a(\|\pi_a^{(n)}\|_2^2) &= \mathbb{E}^{(n)}(\#\text{roots of } P_1 - P_2 \text{ in } \mathbb{F}_p) \\ &= p\mathbb{P}^{(n)}(P_1 = P_2) + \mathbb{E}^{(n)}(\#\text{roots } | P_1 \neq P_2)\mathbb{P}^{(n)}(P_1 \neq P_2)\end{aligned}$$

Start of proof of Thms 1 and 3

Observation:

$$\|\pi_a^{(n)}\|_2^2 = \mathbb{P}^{(n)}(P_1(a) = P_2(a))$$

where P_1, P_2 are independent random polynomials in $\mathcal{P}_n$.

Averaging over $a \in \mathbb{F}_p$: $n \simeq \log p$

$$\begin{aligned}\mathbb{E}_a(\|\pi_a^{(n)}\|_2^2) &= \mathbb{E}^{(n)}(\#\text{roots of } P_1 - P_2 \text{ in } \mathbb{F}_p) \\ &= p\mathbb{P}^{(n)}(P_1 = P_2) + \mathbb{E}^{(n)}(\#\text{roots } | P_1 \neq P_2)\mathbb{P}^{(n)}(P_1 \neq P_2)\end{aligned}$$

- for Thm 1: if $P_1 \neq P_2$, use $\#roots \leqslant n - 1 \simeq \log p$ and some further analysis as in C-D-G.
- for Thm 3: if $P_1 - P_2$ is irreducible, on average over p

$$\#\text{roots of } P_1 - P_2 \simeq 1.$$

Irreducibility of random polynomials

Consider a random polynomial:

$$P = \sum_{i=0}^{n} a_i X^i$$

where, say, the $a_i \in \mathbb{Z}$ are independent and distributed in an interval $[-H, H]$.

Irreducibility of random polynomials

Consider a random polynomial:

$$P = \sum_{i=0}^{n} a_i X^i$$

where, say, the $a_i \in \mathbb{Z}$ are independent and distributed in an interval $[-H, H]$.

Question: Is P irreducible over $\mathbb{Q}$? What are its irreducible factors? its Galois group?

Irreducibility of random polynomials

Consider a random polynomial:

$$P = \sum_{i=0}^{n} a_i X^i$$

where, say, the $a_i \in \mathbb{Z}$ are independent and distributed in an interval $[-H, H]$.

Question: Is P irreducible over $\mathbb{Q}$? What are its irreducible factors? its Galois group?

Two different regimes:

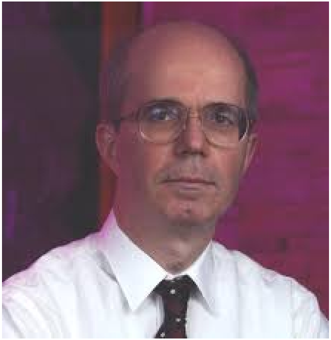
- fixed degree n, but $H \to +\infty$ (known for uniform distribution since van der Waerden '30s, Gallagher '60s)
- H fixed, but $n \to +\infty$: open problem put forth by Odlyzko and Poonen (1993).

Irreducibility of random polynomials

Odlyzko and Poonen '93 conjectured that most polynomials of the form

$$P = 1 + \sum_{i=1}^{n} a_i X^i$$

where $a_i \in \{0, 1\}$ are irreducible.



Irreducibility of random polynomials: our result

Fix H. Assume the a_i's are independent and distributed according to a common law on $[-H, H] \subset \mathbb{Z}$ and set:

$$P = \sum_{i=0}^{n} a_i X^i$$

Irreducibility of random polynomials: our result

Fix H. Assume the a_i's are independent and distributed according to a common law on $[-H, H] \subset \mathbb{Z}$ and set:

$$P = \sum_{i=0}^{n} a_i X^i$$

Theorem 2 (B.-Varjú '18)

Assume GRH. Then with probability at least $1 - \exp(-O(\frac{\sqrt{n}}{\log n}))$

$$P = \Phi \widetilde{P} \text{ where}$$

(i) $\widetilde{P}$ *is irreducible,*
(ii) $d^0(\Phi) = O(\sqrt{n})$ *and* Φ *is a product of cyclotomic factors,*
(iii) moreover the Galois group of P *contains* $Alt(n)$*.*

Irreducibility of random polynomials: our result

Fix H. Assume the a_i's are independent and distributed according to a common law on $[-H, H] \subset \mathbb{Z}$ and set:

$$P = \sum_{i=0}^{n} a_i X^i$$

Theorem 2 (B.-Varjú '18)

Assume GRH. Then with probability at least $1 - \exp(-O(\frac{\sqrt{n}}{\log n}))$

$$P = \Phi \widetilde{P} \text{ where}$$

(i) $\widetilde{P}$ *is irreducible,*
(ii) $d^0(\Phi) = O(\sqrt{n})$ *and* Φ *is a product of cyclotomic factors,*
(iii) moreover the Galois group of P *contains* $Alt(n)$*.*

Corollary (Irreducibility of 0, 1 polynomials)

GRH implies the Odlyzko-Poonen conjecture.

Irreducibility of random polynomials: previous results

- Konyagin (1999) showed that for $0, 1$ polynomials

$$\mathbb{P}(P \text{ is irreducible }) \gg 1/\log n.$$

Irreducibility of random polynomials: previous results

• Konyagin (1999) showed that for $0, 1$ polynomials

$$\mathbb{P}(P \text{ is irreducible }) \gg 1/\log n.$$

• Bary-Soroker and Kozma (2017) showed that if the distribution of coefficients is uniform over $[1, H]$ and H is divisible by at least 4 distinct primes, then

$$\mathbb{P}(P \text{ is irreducible }) \to_{n \to +\infty} 1.$$

Irreducibility of random polynomials: proof method

• It is a sieve argument: we reduce modulo p and average over all primes p in a window $[X, 2X]$ with $X \simeq \exp(\sqrt{n})$.

• Prime Ideal Theorem: For any given $P \in \mathbb{Z}[X]$ monic,

(1) $\#$ irreducible factors of $P = \mathbb{E}_p(\#$ roots of $P \mod p) +$ error

Irreducibility of random polynomials: proof method

- It is a sieve argument: we reduce modulo p and average over all primes p in a window $[X, 2X]$ with $X \simeq \exp(\sqrt{n})$.

- Prime Ideal Theorem: For any given $P \in \mathbb{Z}[X]$ monic,

(1) $\#$ irreducible factors of $P = \mathbb{E}_p(\#\text{ roots of } P \mod p) + \text{error}$

(2) $\mathbb{E}_P(\#\text{ roots of } P \mod p)) = \sum_{a \in \mathbb{F}_p} \mathbb{P}_P(P(a) = 0) = \sum_{a \in \mathbb{F}_p} \pi_a^{(n)}(0)$

Irreducibility of random polynomials: proof method

- It is a sieve argument: we reduce modulo p and average over all primes p in a window $[X, 2X]$ with $X \simeq \exp(\sqrt{n})$.

- Prime Ideal Theorem: For any given $P \in \mathbb{Z}[X]$ monic,

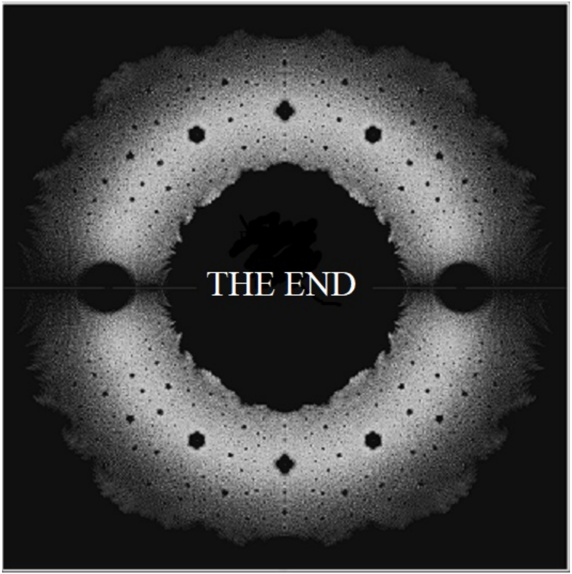
(1) $\#$ irreducible factors of $P = \mathbb{E}_p(\#$ roots of $P \mod p) +$ error

(2) $\mathbb{E}_P(\#\text{ roots of } P \mod p)) = \sum_{a \in \mathbb{F}_p} \mathbb{P}_P(P(a) = 0) = \sum_{a \in \mathbb{F}_p} \pi_a^{(n)}(0)$

- Use Konyagin's $(\log p)^{2+o(1)}$ mixing time estimate to conclude that for $n \geq (\log p)^{2+o(1)}$ we get $\pi_a^{(n)}(0) \simeq \frac{1}{p}$ and hence

$$\mathbb{E}(\#\text{ roots of } P \mod p) \simeq 1.$$

- GRH is used in controlling the error term in the Prime Ideal Theorem: $O(X^{\frac{1}{2}+o(1)} \log Disc(P))$ (Stark, Odlyzko)



Roots of $-1, 0, 1$ polynomials (picture: R. Vanderbei)