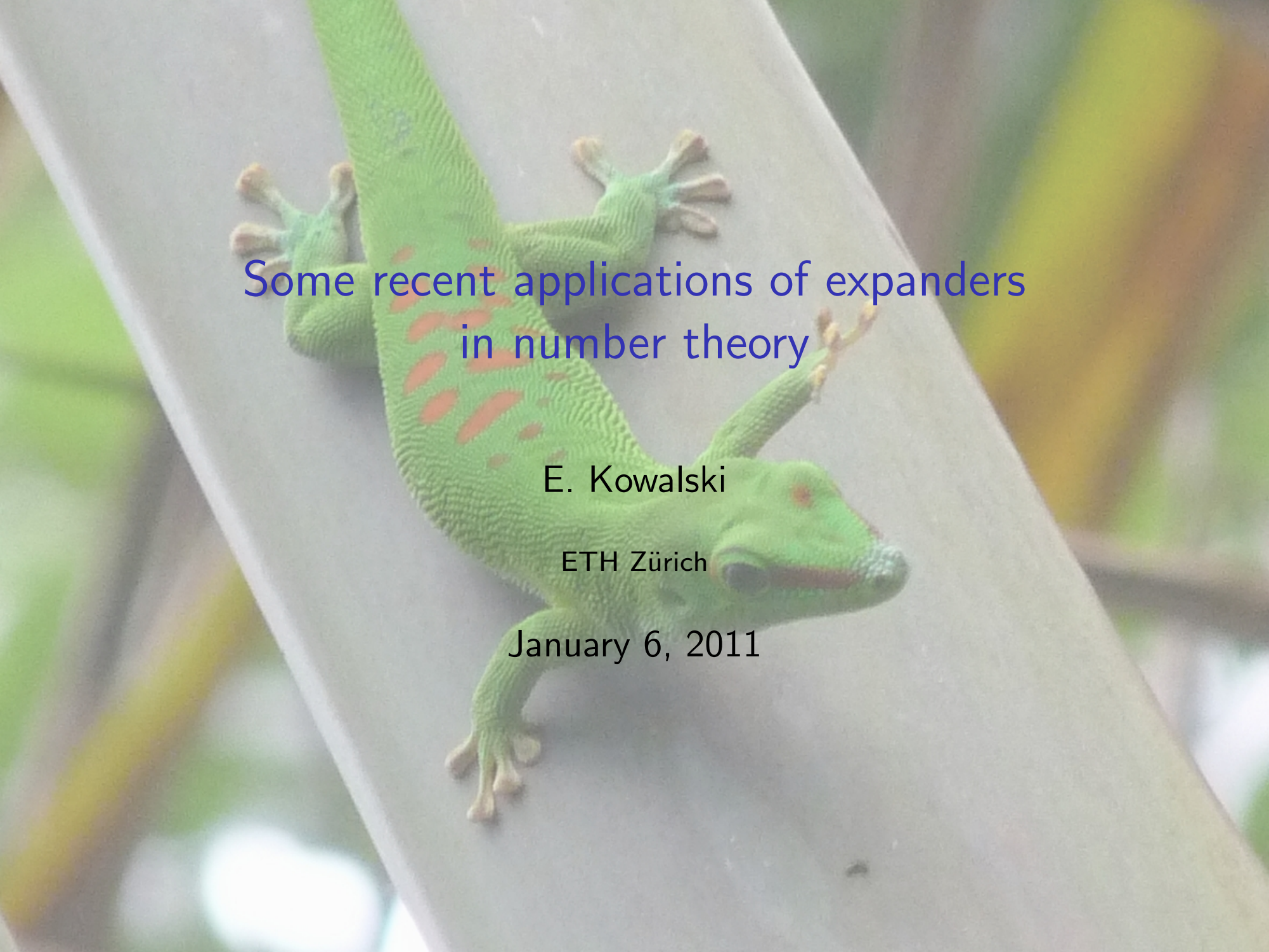

Some recent applications of expanders in number theory

E. Kowalski

ETH Zürich

January 6, 2011

Introduction

There have been links between expander graphs and number theory for a long time.
Today's topics are quite recent applications of expansion properties of graphs to number theory.

1. Sieve and expansion:
   - Sieve in orbits (Bourgain, Gamburd, Sarnak, etc);
   - Large sieve problems (Jouve, K., Zywina, etc);
2. Expander towers of coverings and arithmetic geometry (Ellenberg, Hall, K.); see Jordan's talk tomorrow since I won't say more!

(cont.)

A common feature in these two applications is that the use of expanders comes through their spectral characterizations. There is however a big difference:

- Expansion fits sieve like a glove (if expander graphs didn't exist, sieve-theorists would have invented them!)
- The connection between expanders and arithmetic geometry is much more surprising (first very implicit in some works of Zograf, Abramovich).

In both cases, the connection is an invigorating brew for those who believe in the unity of mathematics.

Sieve

[**References**: A. Lubotzky's Colloquium lectures in this meeting; my survey for Bourbaki, `arXiv:1012.2793`; "The large sieve and its applications", CUP (2008)]

Sieve can be presented as a form of local-global principle. We have a set Y of "global" objects, and for every prime p, a "reduction map" $Y \longrightarrow Y_p$, giving "local" information.

Sieve is particularly concerned with subsets of Y consisting of objects restricted by local conditions

$$\mathcal{T} = \{y \in Y \mid y \,(\mathrm{mod}\, p) \text{ satisfies } \mathcal{P}_p \text{ for all/some } p\},$$

or

$$\mathcal{S} = \{y \in Y \mid y \,(\mathrm{mod}\, p) \text{ does } \textit{not} \text{ satisfy } \mathcal{P}_p \text{ for all/some } p\},$$

where the local conditions $\mathcal{P}$ are defined on Y_p.

(cont.)

Example (Classical)

Take $Y = \mathbf{Z}$, $Y_p = \mathbf{Z}/p\mathbf{Z}$, $Y \to Y_p$ reduction modulo p. For $\mathcal{P}_p = (y \,(\mathrm{mod}\, p) = 0)$, $p \leq X$, we see that $\mathcal{S}$ is the set of integers with no prime factors $\leq X$, and its first elements are 1 and the primes between X and X^2.

Example (Expansive)

Take $Y = \mathrm{SL}_m(\mathbf{Z})$, $Y_p = \mathrm{SL}_m(\mathbf{Z}/p\mathbf{Z})$, $Y \to Y_p$ reduction modulo p. For $\mathcal{P}_p = (\det(T - y \,(\mathrm{mod}\, p))$ is irreducible), we see that $\mathcal{S}$ *contains* the matrices with reducible characteristic polynomial.

(cont.)

Example (Baroque)

Take $Y = \{2^n \mid n \in \mathbf{Z}\}$, $Y_p = \mathbf{Z}/p\mathbf{Z}$, $Y \to Y_p$ reduction modulo p. For $\mathcal{P}_p = (y \pmod p = 1)$, $p \leq X$, we see that $\mathcal{S}$ is the set of integers n such that $2^n - 1$ has no prime factors $\leq X$...

Counting

The goal of sieve is really to estimate, as precisely as possible, the "size" of the sifted set $\mathcal{S}$. This requires stating how the size is measured. A general way to do it is to assume that a *finite* measure μ is given on Y, and to attempt to find $\mu(\mathcal{S})$.

Example

- Take a finite subset $Y_n \subset Y$, and the counting measure supported on Y_n. Typically the Y_n exhaust Y as $n \to +\infty$, and one considers the behavior of $\mu_n(\mathcal{S})$.
- Define a random walk (X_n) on Y, and take the distribution law μ_n of the n-th step of the walk; again, consider the limit $n \to +\infty$.

(cont.)

Example (Integers)

Here $Y = \mathbf{Z}$, $Y_p = \mathbf{Z}/p\mathbf{Z}$. Take μ_n the uniform counting measure on $\{1 \leq y \leq n\}$. If, for instance,

$$\mathcal{P}_p = (y \,(\mathrm{mod}\, p) \in \{0, 2\} \subset Y_p)$$

for $p \leq \sqrt{n}$, then $\mathcal{S}$ is the set of twin primes between $\sqrt{n}$ and n.

Example (Unimodular matrices)

Take $Y = \mathrm{SL}_m(\mathbf{Z})$. Fix a generating set S of Y, for instance elementary matrices $\mathrm{Id} \pm E_{i,j}$, $i \neq j$. Define

$$\mu_n(y) = \sum_{\substack{s_1,\ldots,s_n \in S \\ s_1 \cdots s_n = y}} \cdots \sum 1.$$

This corresponds to a random walk.

What to expect, and when

Recall

$$\mathcal{S} = \{y \in Y \mid y \,(\mathrm{mod}\, p) \text{ does } \textit{not} \text{ satisfy } \mathcal{P}_p \text{ for all/some } p\}.$$

If we assume that

- There is a definite "probability" that $\mathcal{P}_p$ holds, say τ_p;
- The different reductions behave "independently";

then a reasonable guess is that

$$\mu(\mathcal{S}) \approx \mu(Y) \prod_p (1 - \tau_p).$$

General fact. Sieve methods can confirm this guess *in part* under suitable conditions; however, in full generality, one has to be careful at interpreting $\approx$.

The "local" probability

If we put a *single* condition $\mathcal{P}_p$, we expect

$$\mu(\{y \in Y \mid \mathcal{P}_p \text{ holds}\}) \approx \mu(Y)\tau_p,$$

or

$$\frac{1}{\mu(Y)}\mu(\{y \in Y \mid \mathcal{P}_p \text{ holds}\}) \approx \tau_p.$$

If we have a sequence μ_n of measures, we may therefore expect that for each p, the image measure on Y_p of $\mu_n/\mu_n(Y)$ converges to a measure ν_p such that

$$\tau_p = \nu_p(\{y \in Y_p \mid \mathcal{P}_p \text{ holds}\}).$$

(Recall that $\mathcal{P}_p$ is a property of local objects, with no reference to Y.)

Independence of reductions

To go from a single condition to multiple condition, those must be somehow independent. Among global objects, this is not literally the case. However, one can expect asymptotic independence. This means that for a fixed set $d = \{p_1, \ldots, p_k\}$ of distinct primes, the image measure on

$$Y_d = \prod_{p \in d} Y_p$$

of $\mu_n/\mu_n(Y)$ should converge to the product measure

$$\nu_d = \prod_{i=1}^{k} \nu_{p_i}.$$

Examples

Example (Classical)

Integers in an interval become equidistributed in all residue classes: for counting measure μ_n on $\{1, \ldots, n\}$, the local limit measure ν_p is the uniform probability measure on $Y_p = \mathbf{Z}/p\mathbf{Z}$. Since $Y_d \simeq \mathbf{Z}/d\mathbf{Z}$ for $d = p_1 \cdots p_k$ (Chinese Remainder Theorem), we get independence also.

Example (Baroques can't sieve)

The image of the set of powers of 2 in $\mathbf{Z}/p\mathbf{Z}$ ($p \geq 3$) is very badly understood. In particular there is no sign of independence.

Discrete groups and random walks

Let Γ be a finitely generated discrete group (e.g. $\mathrm{SL}_m(\mathbf{Z})$) with a finite quotient $\phi : \Gamma \twoheadrightarrow G$ (e.g. $\mathrm{SL}_m(\mathbf{Z}) \twoheadrightarrow \mathrm{SL}_m(\mathbf{Z}/p\mathbf{Z})$). Let $S = S^{-1}$ be symmetric generators with $1 \in S$ and X_n the n-th step of a random walk on Γ with uniform independent steps from S.

Basic fact: $\phi(X_n)$ becomes equidistributed on G according to the uniform probability measure as $n \to +\infty$. Even better, there exists $\rho_G < 1$ such that

$$\mathbf{P}(\phi(X_n) = g_0) = \frac{1}{|G|} + O(|G|^{1/2}\rho_G^n)$$

for $n \geq 1$ and all $g_0 \in G$ (*exponentially fast convergence*). Moreover, $1 - \rho_G$ is essentially the spectral gap of the associated Cayley graph of G with respect to the generators S.

(cont.)

If Γ has two quotients G_1, G_2 (or more), then the same applies to $\Gamma \to G_1 \times G_2$, *except* that the limit measure is the uniform probability measure on the *image subgroup* of this diagonal quotient map.

In particular, there is asymptotic independence for multiple quotients *if and only if* the "simultaneous reduction" map

$$\Gamma \longrightarrow G_1 \times G_2 \times \cdots$$

is surjective.

(cont.)

Example (Where it works)

Take $\Gamma = \mathrm{SL}_m(\mathbf{Z})$, $G_i = \mathrm{SL}_m(\mathbf{Z}/p_i\mathbf{Z})$, p_i distinct primes. For each i, $\Gamma \to G_i$ is surjective (look at elementary generators). For multiple conditions, we must consider whether

$$\Gamma \to \prod_{1 \le i \le k} \mathrm{SL}_m(\mathbf{Z}/p_i\mathbf{Z})$$

is surjective. This is true: by Chinese Remainder Theorem, we have

$$\prod_{1 \le i \le k} \mathrm{SL}_m(\mathbf{Z}/p_i\mathbf{Z}) \simeq \mathrm{SL}_m(\mathbf{Z}/p_1 \cdots p_k\mathbf{Z})$$

and $\mathrm{SL}_m(\mathbf{Z}) \to \mathrm{SL}_m(\mathbf{Z}/d\mathbf{Z})$ is always surjective (same reason).

(cont.)

Example (Where it doesn't)

Let $\Gamma = O(2,1)(\mathbf{Z})$ (orthogonal group of $X^2 + Y^2 - Z^2$) and its image G_p in $O(2,1)(\mathbf{Z}/p\mathbf{Z})$. Then because the determinant ± 1 of $\gamma \in \Gamma$ is "the same" as that of any of its reductions modulo p, the simultaneous reduction map

$$O(2,1)(\mathbf{Z}) \to G_{p_1} \times G_{p_2}$$

is *never* onto. In such cases one has to adapt the sieve settings.

Why sieve and expanders match

We can now connect sieve with expanders. We consider μ_n arising by random walk on a suitable subgroup $Y = \Gamma \subset \mathrm{SL}_m(\mathbf{Z})$ with respect to reductions $\Gamma \twoheadrightarrow \Gamma_p$, which we assume to be independent. Hence

$$\Gamma \longrightarrow \Gamma_d = \prod_{1 \le i \le k} \Gamma_{p_i}$$

is always surjective for $d = \{p_1, \ldots, p_k\}$. To implement sieve results one must control the sum of error terms in the equidistribution

$$\mu_n(\mathcal{P}_{p_i} \text{ holds}, \ 1 \le i \le k) = \prod_i \nu_{p_i}(\mathcal{P}_{p_i}) + r_d(n)$$

when summed over all d in a range as large as possible.

(cont.)

We illustrate by going back to the classical case.

Example (Classics return)

Take $Y = \mathbf{Z}$ again and the "twin-prime" conditions. Then

$$\frac{1}{n}|\{1 \leq y \leq n \mid y \equiv 0, 2 \,(\mathrm{mod}\, p_i)\}| = \frac{2^k}{d} + O(2^k n^{-1})$$

uniformly. The sum of error terms over d is then smaller than the number n of integers counted as long as $d < n^{1-\delta}$. This leads (after much work) to existence of infinitely many integers n such that n and $n+2$ both have a small number of prime factors (V. Brun). In fact it would be enough, qualitatively, to control error terms up to n^{α} for some $\alpha > 0$.

(cont.)

For the discrete groups, one has

$$r_d(n) = O(|\Gamma_d|^{3/2}\rho_{\Gamma_d}^n).$$

To sum over d, one needs *uniform control* of the spectral radius $\rho_{\Gamma_d} < 1$; if we sum up to $d < D$, the best we can get seems to be

$$\sum_{d<D} r_d(n) \ll D(\max_{d<D}|\Gamma_d|)^{3/2}(\max_{d<D}\rho_{\Gamma_d})^n$$

which we want to tend to 0 as $n \to +\infty$ with D as large as possible. (Think also of the size of Γ_d being polynomial in d.)

(cont.)

In fact, because of the underlying *exponential growth*, we want $D = \beta^n$ for some $\beta > 1$. The only reasonable way this may happen is if

$$\sup_{d \geq 1} \rho_{\Gamma_d} < 1.$$

This is exactly to say that the family of Cayley graphs of $(\Gamma_d, S \,(\mathrm{mod}\, d))$ *is an expander.*

The large sieve inequality

Adapting sieve ideas (going back to Linnik, Montgomery, etc) one can then prove, among other things, the following:

Theorem ("Large sieve" inequality)

Let $\Gamma \subset \mathrm{SL}_m(\mathbf{Z})$ with reductions Γ_p, generating set S, and μ_n the distribution measures of the corresponding random walks. Assume

- *Independence of the reductions;*
- *The Cayley graphs of Γ_d form an expander family;*

Then there exists $\beta > 1$ such that for any conditions $\mathcal{P}_p$ for $p \leq \beta^n$, we have

$$\mu_n(\mathcal{S}) \leq 2\Delta^{-1},$$

where

$$\Delta = \sum_{d \leq \beta^n} \prod_{p \mid d} \frac{\nu_p(\mathcal{P}_p)}{1 - \nu_p(\mathcal{P}_p)} \geq \sum_{p \leq \beta^n} \nu_p(\mathcal{P}_p).$$

Applicability

Example

This result, and the sieve in orbit of Bourgain-Gamburd-Sarnak, is applicable (in decreasing order of "softness") to

- $\mathrm{SL}_m(\mathbf{Z})$, $m \geq 3$, or $\mathrm{Sp}_{2g}(\mathbf{Z})$, $g \geq 2$, or finite index subgroups thereof (Property (T) of Kazhdan);
- $\mathrm{SL}_2(\mathbf{Z})$ (Selberg's $\lambda_1 \geq 3/16$ inequality);
- Any Zariski-dense finitely generated subgroup of $\mathrm{SL}_m(\mathbf{Z})$ (Helfgott, Gill, Pyber-Szabó, Breuillard-Green-Tao, Bourgain-Gamburd, Varjú, Salehi Golsefidy, ...)

Galois groups

Definition

Let $g \in \mathrm{GL}_m(\mathbf{Q})$. The *Galois group* Gal_g of g is the Galois group of the splitting field of the polynomial $\det(T - g) \in \mathbf{Q}[T]$.

Theorem (Jouve, K., Zywina)

Let $\Gamma \subset \mathbf{G}(\mathbf{Q}) \cap \mathrm{GL}_m(\mathbf{Z})$ be an arithmetic subgroup of a split reductive group $\mathbf{G}/\mathbf{Q}$. Let S be a finite symmetric generating set of Γ, $1 \in S$, (X_n) the corresponding random walk. Then

$$\mathbf{P}(\mathrm{Gal}_{X_n} \text{ is not } W(\mathbf{G}))$$

tends to zero exponentially fast as $n \to +\infty$.

[**Reference**: F. Jouve, E. K. and D. Zywina: *Splitting fields of characteristic polynomials of random elements in arithmetic groups*, to appear in Israel J. of Math.; `arXiv:1008.3662`]

(cont.)

Here, the Weyl group $W(\mathbf{G})$ is, as abstract group, isomorphic to $N(\mathbf{T})/\mathbf{T}$ where $\mathbf{T}$ is a maximal torus of $\mathbf{G}$. For SL_m, this is the symmetric group on m letters.

Basic idea of the proof: use the large sieve! More precisely:

- First reduce to walks on cosets of a suitable subgroup of Γ to avoid non-independence problems and problems related to examples like $2^{\mathbf{Z}}$;
- Prove the existence of an a-priori injection

$$\mathrm{Gal}_g \hookrightarrow W(\mathbf{G});$$

(cont.)

- Once independence is achieved, define maps

$$\theta_p : \Gamma_p \to W(\mathbf{G})^\sharp$$

such that $\theta_p(X_n \pmod p)$ "is" the conjugacy class of the Frobenius at p in the splitting field of g (ideas go back to Steinberg);

- Show that the image of θ_p also becomes equidistributed:

$$\frac{1}{|\Gamma_p|}|\{\gamma \in \Gamma_p \mid \theta_p(\gamma) = c\}| = \frac{|c|}{|W(\mathbf{G})|} + O(p^{-1/2})$$

for any fixed conjugacy class c of $W(\mathbf{G})$;

(cont.)

- If Gal_{X_n} does not intersect c, then for all p we have the local condition

$$X_n \,(\mathrm{mod}\, p) \notin \{\gamma \in \Gamma_p \mid \theta_p(\gamma) = c\}\ ;$$

- Combining using the "large" sieve, we find for each c that

$$\mathbf{P}(\mathrm{Gal}_{X_n} \text{ does not intersect } c) \ll 1/\pi(\beta^n)$$

  which tends to 0 exponentially fast;
- And we conclude because no proper subgroup of the finite group $W(\mathbf{G})$ can miss a fixed conjugacy class (due to another Jordan).

Application: relations between roots

Having "large" Galois group can have interesting applications. Here is an example: take $\Gamma = \mathrm{Sp}_{2g}(\mathbf{Z})$, the integral symplectic group. The previous result applies (that case was already proved by Rivin and myself independently).

Corollary

Let (X_n) be a random walk as before on Γ. Then

$$\mathbf{P}(\textit{the roots of } P_{X_n} = \det(T - X_n) \textit{ are } \mathbf{Q}\textit{-linearly dependent}),$$
$$\mathbf{P}(\textit{the roots of } P_{X_n} \textit{ satisfy non-trivial multiplicative relations})$$

both go to zero exponentially fast as $n \to +\infty$.

The unreasonable effectiveness of fast equidistribution

The results described previously are asymptotic. However, in practice, the limits are approached quite fast.
To give an example: consider $\Gamma = \mathbf{E}_8(\mathbf{Z})$; can we *write down* an element of Γ with Galois group $W(\mathbf{E}_8)$ (a subgroup of $\mathfrak{S}_{248}$ with $696,729,600$ elements)?
Yes we can! (*Il peut le faire !*)

$$\gamma = x_{\alpha_1} \cdots x_{\alpha_8} x_{-\alpha_1} \cdots x_{-\alpha_8}$$

[**Reference**: F. Jouve, E. K. and D. Zywina: *An explicit integral polynomial whose splitting field has Galois group* $W(\mathbf{E}_8)$, J. Théor. des Nombres de Bordeaux; `arXiv:0801.1733`]