ARITHMETIC STATISTICS OF FAMILIES OF GALOIS EXTENSIONS AND APPLICATIONS

Ilaria Viglino

ETH Zürich

April 6, 2023

Overview

The goal is to deal with a question in **arithmetic statistics**, namely the distributions of the splitting primes in certain families of number fields. In particular, this allows to count on average the number of "small" totally split primes (in a suitable sense) of the splitting field of certain polynomials.
Classical result:

Theorem

(Frobenius Theorem) The density of the set of primes p for which $f \in \mathbb{Q}[X]$ has a given splitting type $r_1, \dots, r_n$ exists, and it is equal to $1/|G_f|$ times the number of $\sigma \in G_f$ with cycle pattern $r_1, \dots, r_n$.

Here G_f is the Galois group of the splitting field of f over $\mathbb{Q}$, viewed as permutation of the roots of f.
We are interested into computing the density on average over f, of the unramified primes p for which f has a given splitting type modulo p.

Notations and background

$n \geq 3$, $N > 0$ integers,

$$\mathcal{P}_{n,N} \ni f(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_0 \in \mathcal{O}_K[X].$$

We view the $a_0, \ldots, a_{n-1}$ as independent, identically uniformely distribuited taking values in $\{-N, \ldots, N\}$.
We let $N \to +\infty$. Denote by

$$\mathcal{P}^0_{n,N} = \{f \in \mathcal{P}_{n,N} : G_f \cong S_n\}$$

the set of S_n-*polynomials*.

Almost all polynomials are S_n-polynomials: $\frac{|\mathcal{P}^0_{n,N}|}{|\mathcal{P}_{n,N}|} \underset{N\to+\infty}{\longrightarrow} 1$.

(Gallagher, 1973)

$$|\mathcal{P}_{n,N} \setminus \mathcal{P}^0_{n,N}| \ll N^{n-1/2} \log N.$$

(Dietmann, 2013)

$$|\mathcal{P}_{n,N} \setminus \mathcal{P}^0_{n,N}| \ll N^{n-2+\sqrt{2}+\varepsilon},$$

for every $\varepsilon > 0$, as $N \to \infty$.

van der Waerden's conjecture (Bhargava, 2021, Chow and Dietmann, 2020)

$$|\mathcal{P}_{n,N} \setminus \mathcal{P}^0_{n,N}| \ll N^{n-1},$$

for $n \geq 3$.

Main result

Particular case of totally splitting primes.

Theorem (V.)

Let $f \in \mathcal{P}^0_{n,N}$ and let $\pi_f(x)$ the number of primes $\leq x$ splitting completely in $K_f/\mathbb{Q}$. For $x = N^{1/\log\log N}$ (so $x \ll N^{\varepsilon}$ for every $\varepsilon > 0$) and for any real numbers $a < b$ one has

$$\frac{1}{|\mathcal{P}^0_{n,N}|}\left|\left\{f \in \mathcal{P}^0_{n,N} : a \leq \left(\frac{1}{n!} - \frac{1}{n!^2}\right)^{-1/2}\left(\frac{\pi_f(x) - \frac{1}{n!}\pi(x)}{\pi(x)^{1/2}}\right) \leq b\right\}\right|$$

$$\underset{N\to+\infty}{\longrightarrow} \frac{1}{\sqrt{2\pi}}\int_a^b e^{-t^2/2}dt = \mathbb{P}(\mathcal{N}(0,1) \in [a,b]).$$

The average value $\pi_K(x)/n!$ is the one expected by the Chebotarev Density Theorem.

Chebotarev Density Theorem

- L/F normal extension of number fields of degree n,
- $[L:\mathbb{Q}] = n_L,\ [F:\mathbb{Q}] = n_F$,
- $\mathcal{O}_F \subseteq F,\ \mathcal{O}_L \subseteq L$,
- D_F = discriminant of $F/\mathbb{Q}$, D_L = discriminant of $L/\mathbb{Q}$,
- $\mathcal{C}$=fixed conjugacy class in $\mathrm{Gal}_{L/F} = G$,
- $\mathrm{Frob}_{\mathfrak{p},L/F}$=conjugacy class in G of the Frobenius automorphism corresponding to $\mathfrak{P}|\mathfrak{p}$:

$$D(\mathfrak{P}) \twoheadrightarrow \mathrm{Gal}_{\mathcal{O}_L/\mathfrak{P}|\mathcal{O}_F/\mathfrak{p}}$$

$$\mathrm{Fr}_{\mathfrak{P}|\mathfrak{p}} \longmapsto (x \mapsto x^{N(\mathfrak{p})}).$$

Let

$$\pi_{\mathcal{C},L/F}(x) = \sum_{\substack{\mathfrak{p}\subseteq\mathcal{O}_F,\ N_{F/\mathbb{Q}}\mathfrak{p}\leq x \\ \mathfrak{p}\text{ unramified in } L \\ \mathrm{Frob}_{\mathfrak{p},L/F}=\mathcal{C}}} 1,$$

Chebotarev Density Theorem

Theorem (Chebotarev, 1922)

As $x \to +\infty$,

$$\pi_{\mathcal{C},L/F}(x) \sim \frac{|\mathcal{C}|}{|G|} Li(x) \sim \frac{|\mathcal{C}|}{|G|} \frac{x}{\log x}.$$

- $L = F = \mathbb{Q} \longrightarrow$ Prime Number Theorem.
- $L = F \longrightarrow$ Prime Ideal Theorem.
- $F = \mathbb{Q}$, $L = \mathbb{Q}(e^{2\pi i/q}) \longrightarrow$ Dirichlet's Theorem mod q.

Effective versions

Theorem (Lagarias, Odlyzko, 1977)

If the GRH holds for $\zeta_L(s) = \sum_{\mathfrak{a} \subseteq \mathcal{O}_L} \frac{1}{(N_{L/\mathbb{Q}}(\mathfrak{a}))^s}$, *then for every* $x \geq 2$ *there exists* $C > 0$ *so that*

$$\left| \pi_{\mathcal{C},L/F}(x) - \frac{|\mathcal{C}|}{|G|} Li(x) \right| \leq C \frac{|\mathcal{C}|}{|G|} \sqrt{x} (\log D_L + n_L \log x).$$

Effective versions

Lagarias and Odlyzko also proved an unconditional result.
If $n_L > 1$ then ζ_L has at most one zero $s = \sigma + it$ in

$$\sigma \geq 1 - (4 \log D_L)^{-1}, \quad |t| \leq (4 \log D_L)^{-1}.$$

If it exists, it is real and simple, β_0, say.

Theorem (Lagarias, Odlyzko, 1977)

There exist effectively computable absolute constants C_1 and C_2 such that for x large enough

$$\left| \pi_{\mathcal{C},L/F}(x) - \frac{|\mathcal{C}|}{|G|} Li(x) \right| \leq \frac{|\mathcal{C}|}{|G|} Li(x^{\beta_0}) + C_1 x \exp(-C_2 n_L^{-1/2} (\log x)^{1/2}),$$

with the understanding that the β_0 term is present only if β_0 exists.

Effective versions

Theorem (Pierce, Turnage-Butterbaugh, Matchett Wood, 2017)

Fix $A \geq 2$, $0 < \delta \leq \frac{1}{2A}$. Assume that the Artin L-function $\zeta_L(s)/\zeta_F(s) = \prod_{i=1}^{s} L(s, \chi_i, L|F)^{\chi_i(1)}$ is zero-free in the region

$$[1 - \delta, 1] \times [-(\log D_L)^{2/\delta}, (\log D_L)^{2/\delta}].$$

Then for $D_L \geq D_0$, $x \geq x_0$,

$$\left| \pi_{\mathcal{C},L/F}(x) - \frac{|\mathcal{C}|}{|G|} Li(x) \right| \leq \frac{|\mathcal{C}|}{|G|} \frac{x}{(\log x)^A}.$$

For the following families of number fields, almost all of them admit a "good" zero-free region as before, so that a "good" effective CDT holds:

- G cyclic.
- $G \simeq S_3, S_4$, inertia group generated by the conjugacy class of (1 2).
- $G = D_p$ dihedral group, $p > 2$.
- $G = A_4$.
- $G = S_5, A_n$, $n \geq 5$, assuming the Strong Artin Conjecture and that the number of such fields with given discriminant is well controlled.

Back to the family $\mathcal{P}^0_{n,N}$

Let $f \in \mathcal{P}^0_{n,N}$, p rational prime.
$r = (r_1, r_2, \ldots, r_n)$ is the **splitting type** of f mod p if f mod p splits into distinct monic irreducible factors (square-free factorization), with r_1 linear factors, r_2 quadratic factors and so on.
$p \nmid D_f$, r corresponds to the cycle structure of the Frobenius element $\mathrm{Frob}_{K_f/\mathbb{Q},p} =: \mathrm{Frob}_{f,p}$ acting on the roots of f.
For each r we have

$$\sum_{i=1}^{n} i r_i = n.$$

Let $\mathcal{C}_r$ be the conjugacy class in S_n of elements of cycle type r;

$$|\mathcal{C}_r| = n!\delta(r) = n! \prod_{i=1}^{n} \frac{1}{i^{r_i} r_i!}$$

First goal: density on average of primes p unramified in $K_f/\mathbb{Q}$ for which f has a given splitting type modulo p. This is a special case of the Chebotarev Density Theorem, for which we want an explicit asymptotic, with an effective error term, for "almost all" polynomials in our family.
For $x \geq 1$, let

$$\pi_{f,r}(x) = \sum_{\substack{p \leq x \\ f \text{ of splitting type } r \bmod p}} 1 = \sum_{p \leq x} 1_{f,r}(p).$$

$\mathcal{P}^0_{n,N} \subseteq [-N, N]^n$ subset $\longrightarrow$ sum of random variables on $\mathcal{P}^0_{n,N}$:

$$1_{f,r}(p) : \mathcal{P}^0_{n,N} \longrightarrow \{0, 1\}.$$

$\mathbb{P}_N$=uniform probability measure on $[-N, N]^n$
$\mathbb{E}_N$=expectation, σ^2_N=variance of a random variable.

Prime splitting densities

Effective average version of the Chebotarev Density Theorem.

Proposition (V.)

One has, for all primes $\wp$ of norm $q_\wp < N^{1/(n+1)}$,

(1) $\mathbb{P}_N(1_{f,r}(p) = 1) = \mathbb{E}_N(1_{f,r}(\wp)) = \delta(r) + \frac{C_r}{p} + O\Big(\frac{1}{p^2} + p^n N^{-1}\Big)$,
for some explicit constant C_r;

(2) $\sigma_N^2(1_{f,r}(p)) = (\delta(r) - \delta(r)^2) + \frac{C_r(1-2\delta(r))}{p} + O\Big(\frac{1}{p^2} + p^n N^{-1}\Big)$.

It follows that, for $x < N^{1/(n+1)}$,

(3) $\mathbb{E}_N(\pi_{f,r}(x)) = \delta(r)\pi(x) + C_r \log\log x + O_n(1)$,

as $x, N \to +\infty$.

Higher moments

The proof of the below Proposition is motivated by the method of Granville-Soundararajan to compute the moments

$$\sum_{n \leq x} (\omega(n) - \log\log x)^k$$

of the prime divisor function, uniformly in a wide range of k. It has been used to prove the Erdős-Kac Theorem.

Sketch: $\pi_{f,r}(x)$ can be written as a sum of random variables on $\mathcal{P}^0_{n,N}$ as a subset of $[-N, N]^n$. We can approximate with *independent* discrete random variables. At some point we obtain in the main term a function evaluated at products $p_1 \dots p_k$ of primes which is 0 a lot of times.

Proposition (V.)

*For k **even**, and $x < N^{1/k(n+1/2)}$ one has*

$$\mathbb{E}_N\left(\left(\pi_{f,r}(x) - \delta(r)\pi(x)\right)^k\right) = C_{k,r}\pi(x)^{k/2}\left(1 + O\left(\frac{\log\log x}{\pi(x)^{1/2}}\right)\right) + O_n(\pi(x)^{k(n+1)}N^{-1}),$$

*while for k **odd**,*

$$\mathbb{E}_N\left(\left(\pi_{f,r}(x) - \delta(r)\pi(x)\right)^k\right) \ll_n C_{k,r}\pi(x)^{k/2}\frac{\log\log x}{\pi(x)^{1/2}} + \pi(x)^{k(n+1)}N^{-1},$$

as $x, N \to +\infty$.

The main theorem

Theorem (V.)

For $x = N^{1/\log\log N}$ *and for any* $b \in \mathbb{R}$,

$$\mathbb{P}_N\left(\frac{\pi_{f,r}(x) - \delta(r)\pi(x)}{(\delta(r) - \delta(r)^2)^{1/2}\pi(x)^{1/2}} \leq b\right) \longrightarrow \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{b} e^{-t^2/2}dt,$$

as $N \to +\infty$.

Question. In this Central Limit Theorem, what is the "right" range of x and N for which it holds?

Sketch of the proof

Since the gaussian $\mathcal{N}(0,1)$ is determined by its moments $\mathbb{E}(\mathcal{N}(0,1)^k)$ for $k \geq 0$, it is enough to prove that

$$\mathbb{E}_N\left(\left(\pi_{f,r}(x) - \delta(r)\pi_(x)\right)\right)^k \underset{N\to+\infty}{\longrightarrow} \mu_k,$$

where

$$\mu_k = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} x^k e^{-x^2/2} dx = \begin{cases} \frac{k!}{2^{k/2}(k/2)!} & \text{if } k \text{ is even} \\ 0 & \text{if } k \text{ is odd.} \end{cases}$$

One can see that the constants $C_{k,r}$ of the above Proposition agree with the moments μ_k.

Main application

Bounds for ℓ-torsion of class groups

- $L/\mathbb{Q}$ number field of degree d and discriminant D_L,
- Cl_L=ideal class group, finite abelian group that encodes information about the arithmetic of L,
- $h_L = |Cl_L|$=class number.

(Landau, 1915)

$$h_L \leq \frac{s!}{s^s} \frac{4^{r_2}}{\pi^{r_2}} D_L^{1/2} (\log D_L)^{s-1}$$

If L is an imaginary quadratic field, the bounds for h_L in terms of the discriminant lie at the heart of solving the Gauss class number problem: to provide for each $n \geq 1$ a complete list of imaginary quadratic fields having class number n.

We focus on the ℓ-torsion subgroup of Cl_L for $\ell \geq 1$:

$$Cl_L[\ell] = \{[\mathfrak{a}] \in Cl_L : [\mathfrak{a}]^\ell = \mathrm{Id}\}$$

$$h_L[\ell] = |Cl_L[\ell]|.$$

We may trivially apply Landau's upper bound,

$$h_L[\ell] \ll_{s,\varepsilon} D_L^{1/2+\varepsilon},$$

for all $\varepsilon > 0$. This trivial bound is thought to be far from the truth, leading to the following conjecture.

The ℓ-torsion conjecture

Conjecture

One has

$$h_L[\ell] \ll_{s,\ell,\varepsilon} D_L^{\varepsilon}, \qquad \forall \varepsilon > 0.$$

Motivations: It follows from the Cohen-Lenstra heuristics, the discriminant multiplicity conjecture and a generalization of the Malle conjecture.
Furthermore Brumer and Silverman motivate it by counting elliptic curves with fixed conductor.

Results in the direction of the ℓ-torsion conjecture

Most relevant to this work are upper bound results for averages, or equivalently results that hold for "almost all" fields in a certain family.

Here: $\mathcal{F}_d$ family of number fields of fixed degree d, $\mathcal{E} \subseteq \mathcal{F}_d$ has relative density zero if there exists $\beta > 0$

$$|\mathcal{F}_{d,\mathsf{disc} \leq x}| \gg x^{\beta}, \quad \forall x \geq 1$$

while there exists $0 \leq \alpha < \beta$

$$|\mathcal{E}_{\mathsf{disc} \leq x}| \ll x^{\alpha}, \quad \forall x \geq 1.$$

(Soundararajan, 2000) Imaginary quadratic fields

$$h_L[\ell] \ll_{\ell,\varepsilon} D_L^{1/2-1/2\ell+\varepsilon}, \quad \forall \varepsilon > 0$$

for all but a possible family of exceptional fields, with density zero.

(Heath-Brown, Pierce, 2017) Imaginary quadratic fields:

$$h_L[\ell] \ll_{\ell,\varepsilon} D_L^{1/2-3/(2\ell+2)+\varepsilon}, \quad \forall \varepsilon > 0$$

for almost all fields.
(Ellenberg-Venkatesh, 2005) Conditional on the GRH for the Dedekind zeta function of L,

$$h_L[\ell] \ll_{s,\ell,\varepsilon} D_L^{1/2-1/(2\ell(s-1))+\varepsilon}, \quad \forall \varepsilon > 0$$

The main point is the existence of many primes splitting completely of "small" norm. The GRH guarantees the existence of many such primes. But we want to proceed unconditionally. Recall that in our case we have the presence of "small" primes splitting completely.

Splitting primes and torsion in the class group

Theorem (Ellenberg, Venkatesh, 2005)

Let $L/\mathbb{Q}$ be a field extension of degree s. Set $\delta < \frac{1}{2\ell(s-1)}$ and let

$$M = |\{p \leq D_L^{\delta} : p \text{ splits completely in } L/\mathbb{Q}\}|.$$

Then, for any $\varepsilon > 0$,

$$h_L[\ell] \ll_{s,\ell,\varepsilon} \frac{D_L^{1/2+\varepsilon}}{M}.$$

Result for almost all $f \in \mathcal{P}^0_{n,N}$

Corollary (V.)

For every positive integer ℓ, $\varepsilon > 0$ and for almost all $f \in \mathcal{P}^0_{n,N}$ (outside of a set of size $o(N^n)$) we have

$$h_f[\ell] \ll_{n,\ell,\varepsilon} D_f^{\frac{1}{2} - \frac{1}{(2n-2)(n-1)! \log\log |d_f|} + \varepsilon},$$

as $N \to +\infty$.

Sketch: Use the main theorem to lower bound the number of primes splitting completely in $K_f/\mathbb{Q}$.

Corollary (V.)

The average of the number of ramified primes in $K_f/\mathbb{Q}$ is

$$\mathbb{E}_N(|\{p : p|D_f\}|) \ll_n \log\log N,$$

as $N \to +\infty$.

Further applications

Discriminant of f and D_f

Let α be a root of f, and let $\mathbb{Q}(\alpha)/\mathbb{Q}$ be the extension generated by α over $\mathbb{Q}$. The relation between the discriminant d_f of f and D_f is given by

$$d_f = a_f^2 \cdot D_{\mathbb{Q}(\alpha)}$$

for some $a_f \in \mathbb{Z}$, so

$$d_f = D_f^{1/(n-1)!} a_f^2 (N_{\mathbb{Q}(\alpha)/\mathbb{Q}} \mathfrak{D}_{K_f/\mathbb{Q}(\alpha)})^{1/(n-1)!}.$$

Corollary (V.)

The average of the number of primes dividing a_f is

$$\mathbb{E}_N(|\{p : p|a_f\}|) \ll_n 1,$$

as $N \to +\infty$.

Thank you for your attention!