Equidistribution of Frobenius in nilpotent extensions

Peter Koymans
University of Michigan

Joint Mathematics Meetings
Special Session on Arithmetic Statistics

5 January 2023

Malle's conjecture

Conjecture (Malle's conjecture)

Let G be a finite, non-trivial group. Then there exist numbers $a(G) \in \mathbb{Q}_{>0}$, $b(G) \in \mathbb{Z}_{\geq 0}$ and $c(G) > 0$ such that

$$\#\{K/\mathbb{Q} : D_K \leq X, \mathrm{Gal}(K/\mathbb{Q}) \cong G\} \sim c(G) X^{a(G)} (\log X)^{b(G)}.$$

Malle's conjecture

Conjecture (Malle's conjecture)

Let G be a finite, non-trivial group. Then there exist numbers $a(G) \in \mathbb{Q}_{>0}$, $b(G) \in \mathbb{Z}_{\geq 0}$ and $c(G) > 0$ such that

$$\#\{K/\mathbb{Q} : D_K \leq X, \mathrm{Gal}(K/\mathbb{Q}) \cong G\} \sim c(G) X^{a(G)} (\log X)^{b(G)}.$$

This is a generalization of the inverse Galois problem.

Malle's conjecture

Conjecture (Malle's conjecture)

Let G be a finite, non-trivial group. Then there exist numbers $a(G) \in \mathbb{Q}_{>0}$, $b(G) \in \mathbb{Z}_{\geq 0}$ and $c(G) > 0$ such that

$$\#\{K/\mathbb{Q} : D_K \leq X, \mathrm{Gal}(K/\mathbb{Q}) \cong G\} \sim c(G) X^{a(G)} (\log X)^{b(G)}.$$

This is a generalization of the inverse Galois problem.

As phrased above, this conjecture is widely believed to be correct.

Malle's conjecture

Conjecture (Malle's conjecture)

Let G be a finite, non-trivial group. Then there exist numbers $a(G) \in \mathbb{Q}_{>0}$, $b(G) \in \mathbb{Z}_{\geq 0}$ and $c(G) > 0$ such that

$$\#\{K/\mathbb{Q} : D_K \leq X, \mathrm{Gal}(K/\mathbb{Q}) \cong G\} \sim c(G) X^{a(G)} (\log X)^{b(G)}.$$

This is a generalization of the inverse Galois problem.

As phrased above, this conjecture is widely believed to be correct.

Malle proposed some explicit values $a_{\mathrm{Malle}}(G)$ and $b_{\mathrm{Malle}}(G)$. Malle's $b_{\mathrm{Malle}}(G)$ is known to be wrong in general.

Malle's conjecture

Conjecture (Malle's conjecture)

Let G be a finite, non-trivial group. Then there exist numbers $a(G) \in \mathbb{Q}_{>0}$, $b(G) \in \mathbb{Z}_{\geq 0}$ and $c(G) > 0$ such that

$$\#\{K/\mathbb{Q} : D_K \leq X, \mathrm{Gal}(K/\mathbb{Q}) \cong G\} \sim c(G) X^{a(G)} (\log X)^{b(G)}.$$

This is a generalization of the inverse Galois problem.

As phrased above, this conjecture is widely believed to be correct.

Malle proposed some explicit values $a_{\mathrm{Malle}}(G)$ and $b_{\mathrm{Malle}}(G)$. Malle's $b_{\mathrm{Malle}}(G)$ is known to be wrong in general.

Sometimes $c(G)$ is an Euler product. This is expected to be true for S_n (Malle–Bhargava principle).

Known cases

Malle's conjecture is known in the following cases:

Known cases

Malle's conjecture is known in the following cases:

- abelian G by Wright;

Known cases

Malle's conjecture is known in the following cases:

- abelian G by Wright;
- S_3 by Davenport–Heilbronn (with much follow-up work);

Known cases

Malle's conjecture is known in the following cases:

- abelian G by Wright;
- S_3 by Davenport–Heilbronn (with much follow-up work);
- S_4, S_5 by Bhargava;

Known cases

Malle's conjecture is known in the following cases:

- abelian G by Wright;
- S_3 by Davenport–Heilbronn (with much follow-up work);
- S_4, S_5 by Bhargava;
- $S_3 \subseteq S_6$ by Bhargava–Wood;

Known cases

Malle's conjecture is known in the following cases:

- abelian G by Wright;
- S_3 by Davenport–Heilbronn (with much follow-up work);
- S_4, S_5 by Bhargava;
- $S_3 \subseteq S_6$ by Bhargava–Wood;
- $D_4 \subseteq S_4$ by Cohen–Diaz y Diaz–Olivier (with follow-up work by Bucur–Florea–Serrano López–Varma);

Known cases

Malle's conjecture is known in the following cases:

- abelian G by Wright;
- S_3 by Davenport–Heilbronn (with much follow-up work);
- S_4, S_5 by Bhargava;
- $S_3 \subseteq S_6$ by Bhargava–Wood;
- $D_4 \subseteq S_4$ by Cohen–Diaz y Diaz–Olivier (with follow-up work by Bucur–Florea–Serrano López–Varma);
- generalized quaternion groups and some wreath products by Klüners;

Known cases

Malle's conjecture is known in the following cases:

- abelian G by Wright;
- S_3 by Davenport–Heilbronn (with much follow-up work);
- S_4, S_5 by Bhargava;
- $S_3 \subseteq S_6$ by Bhargava–Wood;
- $D_4 \subseteq S_4$ by Cohen–Diaz y Diaz–Olivier (with follow-up work by Bucur–Florea–Serrano López–Varma);
- generalized quaternion groups and some wreath products by Klüners;
- any nilpotent group G, in the regular representation, such that all elements of order p are central, where p is the smallest prime dividing $\#G$ by K.–Pagano;

Known cases

Malle's conjecture is known in the following cases:

- abelian G by Wright;
- S_3 by Davenport–Heilbronn (with much follow-up work);
- S_4, S_5 by Bhargava;
- $S_3 \subseteq S_6$ by Bhargava–Wood;
- $D_4 \subseteq S_4$ by Cohen–Diaz y Diaz–Olivier (with follow-up work by Bucur–Florea–Serrano López–Varma);
- generalized quaternion groups and some wreath products by Klüners;
- any nilpotent group G, in the regular representation, such that all elements of order p are central, where p is the smallest prime dividing $\#G$ by K.–Pagano;
- nonic Heisenberg extensions by Fouvry–K.;
- direct products $S_n \times A$ for $n \in \{3, 4, 5\}$ and A abelian by Wang (with $\#A$ coprime to some values) and later by Masri–Thorne–Tsai–Wang.

Fair counting functions

Ordering by discriminant has some undesirable features: leading constant need not be an Euler product and subfields may occur a positive proportion of the time.

Fair counting functions

Ordering by discriminant has some undesirable features: leading constant need not be an Euler product and subfields may occur a positive proportion of the time.

Wood (2010) introduced a class of "fair counting functions".

Fair counting functions

Ordering by discriminant has some undesirable features: leading constant need not be an Euler product and subfields may occur a positive proportion of the time.

Wood (2010) introduced a class of "fair counting functions".

Important examples of fair counting functions are the conductor and the product of ramified primes.

Fair counting functions

Ordering by discriminant has some undesirable features: leading constant need not be an Euler product and subfields may occur a positive proportion of the time.

Wood (2010) introduced a class of "fair counting functions".

Important examples of fair counting functions are the conductor and the product of ramified primes.

Mäki (1993): Malle's conjecture for abelian extensions ordered by conductor.

Fair counting functions

Ordering by discriminant has some undesirable features: leading constant need not be an Euler product and subfields may occur a positive proportion of the time.

Wood (2010) introduced a class of "fair counting functions".

Important examples of fair counting functions are the conductor and the product of ramified primes.

Mäki (1993): Malle's conjecture for abelian extensions ordered by conductor.

Wood (2010): Malle's conjecture for abelian extensions ordered by any fair counting function with local conditions.

Fair counting functions

Ordering by discriminant has some undesirable features: leading constant need not be an Euler product and subfields may occur a positive proportion of the time.

Wood (2010) introduced a class of “fair counting functions”.

Important examples of fair counting functions are the conductor and the product of ramified primes.

Mäki (1993): Malle’s conjecture for abelian extensions ordered by conductor.

Wood (2010): Malle’s conjecture for abelian extensions ordered by any fair counting function with local conditions.

Altug–Shankar–Varma–Wilson (2017): Malle’s conjecture for D_4 by Artin conductor.

Main result

A group G is called *nilpotent* if it is a direct product of p-groups.

Main result

A group G is called *nilpotent* if it is a direct product of p-groups.

Theorem (K.–Pagano)

Assume GRH. Let G be a nilpotent group with $\#G$ odd. Then

$$\liminf_{X\to\infty} \frac{\#\left\{K/\mathbb{Q} : \prod_{p : I_p \neq \{\mathrm{id}\}} p \leq X, \mathrm{Gal}(K/\mathbb{Q}) \cong G\right\}}{c'(G) X (\log X)^{b'(G)}} \geq 1,$$

where $c'(G)$ is the expected Euler product and where $b'(G)$ is the naïve analogue of Malle's $b(G)$ in this situation.

Main result

A group G is called *nilpotent* if it is a direct product of p-groups.

Theorem (K.–Pagano)

Assume GRH. Let G be a nilpotent group with $\#G$ odd. Then

$$\liminf_{X \to \infty} \frac{\#\left\{K/\mathbb{Q} : \prod_{p : I_p \neq \{\mathrm{id}\}} p \leq X, \mathrm{Gal}(K/\mathbb{Q}) \cong G\right\}}{c'(G) X (\log X)^{b'(G)}} \geq 1,$$

where $c'(G)$ is the expected Euler product and where $b'(G)$ is the naïve analogue of Malle's $b(G)$ in this situation.

Surprisingly, the corresponding asymptotic

$$\lim_{X \to \infty} \frac{\#\left\{K/\mathbb{Q} : \prod_{p : I_p \neq \{\mathrm{id}\}} p \leq X, \mathrm{Gal}(K/\mathbb{Q}) \cong G\right\}}{c'(G) X (\log X)^{b'(G)}} = 1$$

is not true in general. Counterexamples exist for nilpotency class 2.

Main result

A group G is called *nilpotent* if it is a direct product of p-groups.

Theorem (K.–Pagano)

Assume GRH. Let G be a nilpotent group with $\#G$ odd. Then

$$\liminf_{X \to \infty} \frac{\#\left\{K/\mathbb{Q} : \prod_{p : I_p \neq \{\mathrm{id}\}} p \leq X, \mathrm{Gal}(K/\mathbb{Q}) \cong G\right\}}{c'(G) X (\log X)^{b'(G)}} \geq 1,$$

where $c'(G)$ is the expected Euler product and where $b'(G)$ is the naïve analogue of Malle's $b(G)$ in this situation.

Surprisingly, the corresponding asymptotic

$$\lim_{X \to \infty} \frac{\#\left\{K/\mathbb{Q} : \prod_{p : I_p \neq \{\mathrm{id}\}} p \leq X, \mathrm{Gal}(K/\mathbb{Q}) \cong G\right\}}{c'(G) X (\log X)^{b'(G)}} = 1$$

is not true in general. Counterexamples exist for nilpotency class 2.

Work in progress: asymptotic for a slightly modified counting function.

Step 1a: parametrization

To give an idea how the techniques work, we will (unconditionally!) give an overview for the proof of the asymptotic for the number of Galois D_4-extensions by product of ramified primes.

Step 1a: parametrization

To give an idea how the techniques work, we will (unconditionally!) give an overview for the proof of the asymptotic for the number of Galois D_4-extensions by product of ramified primes.

We have a central exact sequence

$$0 \to \mathbb{F}_2 \to D_4 \xrightarrow{q} \mathbb{F}_2^2 \to 0$$

and a bijection

$$\mathrm{Epi}(G_{\mathbb{Q}}, \mathbb{F}_2^2) \leftrightarrow \{(a, b) \in (\mathbb{Q}^*/\mathbb{Q}^{*2})^2 : a, b \text{ lin. ind.}\}.$$

Step 1a: parametrization

To give an idea how the techniques work, we will (unconditionally!) give an overview for the proof of the asymptotic for the number of Galois D_4-extensions by product of ramified primes.

We have a central exact sequence

$$0 \to \mathbb{F}_2 \to D_4 \xrightarrow{q} \mathbb{F}_2^2 \to 0$$

and a bijection

$$\mathrm{Epi}(G_{\mathbb{Q}}, \mathbb{F}_2^2) \leftrightarrow \{(a, b) \in (\mathbb{Q}^*/\mathbb{Q}^{*2})^2 : a, b \text{ lin. ind.}\}.$$

Given $\pi \in \mathrm{Epi}(G_{\mathbb{Q}}, \mathbb{F}_2^2)$, this leads to the *central embedding problem*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathbb{F}_2 & \longrightarrow & D_4 & \longrightarrow & \mathbb{F}_2^2 & \longrightarrow & 0 \\ & & & & \uparrow ? & \nearrow_{\pi} & & & \\ & & & & G_{\mathbb{Q}} & & & & \end{array}$$

Step 1a: parametrization

To give an idea how the techniques work, we will (unconditionally!) give an overview for the proof of the asymptotic for the number of Galois D_4-extensions by product of ramified primes.

We have a central exact sequence

$$0 \to \mathbb{F}_2 \to D_4 \xrightarrow{q} \mathbb{F}_2^2 \to 0$$

and a bijection

$$\mathrm{Epi}(G_{\mathbb{Q}}, \mathbb{F}_2^2) \leftrightarrow \{(a, b) \in (\mathbb{Q}^*/\mathbb{Q}^{*2})^2 : a, b \text{ lin. ind.}\}.$$

Given $\pi \in \mathrm{Epi}(G_{\mathbb{Q}}, \mathbb{F}_2^2)$, this leads to the *central embedding problem*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathbb{F}_2 & \longrightarrow & D_4 & \longrightarrow & \mathbb{F}_2^2 & \longrightarrow & 0 \\ & & & & \uparrow ? & \nearrow_{\pi} & & & \\ & & & & G_{\mathbb{Q}} & & & & \end{array}$$

It is well-known that a $\mathbb{F}_2^2$-extension $\mathbb{Q}(\sqrt{a}, \sqrt{b})$ of $\mathbb{Q}$ is contained in a D_4-extension if and only if $x^2 = ay^2 + bz^2$ has a non-trivial point.

Step 1b: parametrization

If $\rho \in \mathrm{Epi}(G_{\mathbb{Q}}, D_4)$ is a lift of $\pi \in \mathrm{Epi}(G_{\mathbb{Q}}, \mathbb{F}_2)$ and $q : D_4 \twoheadrightarrow \mathbb{F}_2^2$, then

$$\{f \in \mathrm{Epi}(G_{\mathbb{Q}}, D_4) : f \circ q = \pi\} = \{\rho \cdot \chi : \chi \in \mathrm{Hom}(G_{\mathbb{Q}}, \mathbb{F}_2)\}.$$

Step 1b: parametrization

If $\rho \in \mathrm{Epi}(G_{\mathbb{Q}}, D_4)$ is a lift of $\pi \in \mathrm{Epi}(G_{\mathbb{Q}}, \mathbb{F}_2)$ and $q : D_4 \twoheadrightarrow \mathbb{F}_2^2$, then

$$\{f \in \mathrm{Epi}(G_{\mathbb{Q}}, D_4) : f \circ q = \pi\} = \{\rho \cdot \chi : \chi \in \mathrm{Hom}(G_{\mathbb{Q}}, \mathbb{F}_2)\}.$$

Therefore we have a bijection

$$\mathrm{Epi}(G_{\mathbb{Q}}, D_4) \leftrightarrow \{(a, b, c) \in (\mathbb{Q}^*/\mathbb{Q}^{*2})^3 : a, b \text{ ind.}, x^2 = ay^2 + bz^2 \text{ sol.}\}.$$

Step 1b: parametrization

If $\rho \in \mathrm{Epi}(G_{\mathbb{Q}}, D_4)$ is a lift of $\pi \in \mathrm{Epi}(G_{\mathbb{Q}}, \mathbb{F}_2)$ and $q : D_4 \twoheadrightarrow \mathbb{F}_2^2$, then

$$\{f \in \mathrm{Epi}(G_{\mathbb{Q}}, D_4) : f \circ q = \pi\} = \{\rho \cdot \chi : \chi \in \mathrm{Hom}(G_{\mathbb{Q}}, \mathbb{F}_2)\}.$$

Therefore we have a bijection

$$\mathrm{Epi}(G_{\mathbb{Q}}, D_4) \leftrightarrow \{(a, b, c) \in (\mathbb{Q}^*/\mathbb{Q}^{*2})^3 : a, b \text{ ind.}, x^2 = ay^2 + bz^2 \text{ sol.}\}.$$

Under this parametrization, the product of ramified primes maps to $\mathrm{rad}(|abc|)$ (ignoring minor issues with ramification at 2).

Step 1b: parametrization

If $\rho \in \mathrm{Epi}(G_{\mathbb{Q}}, D_4)$ is a lift of $\pi \in \mathrm{Epi}(G_{\mathbb{Q}}, \mathbb{F}_2)$ and $q : D_4 \twoheadrightarrow \mathbb{F}_2^2$, then

$$\{f \in \mathrm{Epi}(G_{\mathbb{Q}}, D_4) : f \circ q = \pi\} = \{\rho \cdot \chi : \chi \in \mathrm{Hom}(G_{\mathbb{Q}}, \mathbb{F}_2)\}.$$

Therefore we have a bijection

$$\mathrm{Epi}(G_{\mathbb{Q}}, D_4) \leftrightarrow \{(a, b, c) \in (\mathbb{Q}^*/\mathbb{Q}^{*2})^3 : a, b \text{ ind.}, x^2 = ay^2 + bz^2 \text{ sol.}\}.$$

Under this parametrization, the product of ramified primes maps to $\mathrm{rad}(|abc|)$ (ignoring minor issues with ramification at 2).

It turns out to be more convenient to work with seven variables α_S for $\emptyset \subset S \subseteq \{a, b, c\}$, where α_S is the product over all primes p dividing the variables in S and not dividing the variables in $\{a, b, c\} - S$.

Step 1b: parametrization

If $\rho \in \mathrm{Epi}(G_{\mathbb{Q}}, D_4)$ is a lift of $\pi \in \mathrm{Epi}(G_{\mathbb{Q}}, \mathbb{F}_2)$ and $q : D_4 \twoheadrightarrow \mathbb{F}_2^2$, then

$$\{f \in \mathrm{Epi}(G_{\mathbb{Q}}, D_4) : f \circ q = \pi\} = \{\rho \cdot \chi : \chi \in \mathrm{Hom}(G_{\mathbb{Q}}, \mathbb{F}_2)\}.$$

Therefore we have a bijection

$$\mathrm{Epi}(G_{\mathbb{Q}}, D_4) \leftrightarrow \{(a, b, c) \in (\mathbb{Q}^*/\mathbb{Q}^{*2})^3 : a, b \text{ ind.}, x^2 = ay^2 + bz^2 \text{ sol.}\}.$$

Under this parametrization, the product of ramified primes maps to $\mathrm{rad}(|abc|)$ (ignoring minor issues with ramification at 2).

It turns out to be more convenient to work with seven variables α_S for $\emptyset \subset S \subseteq \{a, b, c\}$, where α_S is the product over all primes p dividing the variables in S and not dividing the variables in $\{a, b, c\} - S$.

The variables α_S are squarefree and pairwise coprime, and we have $\mathrm{rad}(|abc|) = \prod_{\emptyset \subset S \subseteq \{a,b,c\}} |\alpha_S|$.

Step 2: character sums

Define $T(a)$ to be the subsets of $\{a, b, c\}$ containing a. Then we have

$$a = \prod_{S \in T(a)} \alpha_S$$

and similarly for b, c.

Step 2: character sums

Define $T(a)$ to be the subsets of $\{a, b, c\}$ containing a. Then we have

$$a = \prod_{S \in T(a)} \alpha_S$$

and similarly for b, c. So to count D_4-extensions, must evaluate

$$\sum_{\substack{\prod_{\emptyset \subset S \subseteq \{a,b,c\}} |\alpha_S| \leq X \\ a,b \text{ lin. ind.}}} \mu^2\left(\prod_S |\alpha_S|\right) \cdot \mathbf{1}_{x^2 = \alpha_a \alpha_{a,b} \alpha_{a,c} \alpha_{a,b,c} y^2 + \alpha_b \alpha_{a,b} \alpha_{b,c} \alpha_{a,b,c} z^2 \text{ sol.}}.$$

Step 2: character sums

Define $T(a)$ to be the subsets of $\{a, b, c\}$ containing a. Then we have

$$a = \prod_{S \in T(a)} \alpha_S$$

and similarly for b, c. So to count D_4-extensions, must evaluate

$$\sum_{\substack{\prod_{\emptyset \subset S \subseteq \{a,b,c\}} |\alpha_S| \leq X \\ a,b \text{ lin. ind.}}} \mu^2\left(\prod_S |\alpha_S|\right) \cdot \mathbf{1}_{x^2 = \alpha_a \alpha_{a,b} \alpha_{a,c} \alpha_{a,b,c} y^2 + \alpha_b \alpha_{a,b} \alpha_{b,c} \alpha_{a,b,c} z^2 \text{ sol.}}.$$

Hasse-Minkowski: detect solubility of conic locally at primes dividing α_S.

Step 2: character sums

Define $T(a)$ to be the subsets of $\{a, b, c\}$ containing a. Then we have

$$a = \prod_{S \in T(a)} \alpha_S$$

and similarly for b, c. So to count D_4-extensions, must evaluate

$$\sum_{\substack{\prod_{\emptyset \subset S \subseteq \{a,b,c\}} |\alpha_S| \leq X \\ a,b \text{ lin. ind.}}} \mu^2\left(\prod_S |\alpha_S|\right) \cdot \mathbf{1}_{x^2 = \alpha_a \alpha_{a,b} \alpha_{a,c} \alpha_{a,b,c} y^2 + \alpha_b \alpha_{a,b} \alpha_{b,c} \alpha_{a,b,c} z^2 \text{ sol.}}.$$

Hasse-Minkowski: detect solubility of conic locally at primes dividing α_S.

Now rewrite the above sum as a sum over Legendre symbols involving the variables α_S.

Step 3: equidistribution

Evaluate the resulting character sum using Chebotarev and the large sieve.

Step 3: equidistribution

Evaluate the resulting character sum using Chebotarev and the large sieve.

How does this process generalize?

Step 3: equidistribution

Evaluate the resulting character sum using Chebotarev and the large sieve.

How does this process generalize?

- Build a nilpotent extension by iterated central extensions. This yields a parametrization of G-extensions by tuples of squarefree integers satisfying central embedding problems.

Step 3: equidistribution

Evaluate the resulting character sum using Chebotarev and the large sieve.

How does this process generalize?

- Build a nilpotent extension by iterated central extensions. This yields a parametrization of G-extensions by tuples of squarefree integers satisfying central embedding problems.
- These central embedding problems get much more complicated, but still satisfy local-to-global and are certainly determined by Frob_p for p dividing the variables of the parametrization.

Step 3: equidistribution

Evaluate the resulting character sum using Chebotarev and the large sieve.

How does this process generalize?

- Build a nilpotent extension by iterated central extensions. This yields a parametrization of G-extensions by tuples of squarefree integers satisfying central embedding problems.
- These central embedding problems get much more complicated, but still satisfy local-to-global and are certainly determined by Frob_p for p dividing the variables of the parametrization.
- In our chosen ordering, a typical extension is a rather large twist of a "minimally ramified central extension". Getting equidistribution of Frobenius in minimally ramified extensions is *very hard*. The key idea of the proof is to exploit the twisting.

Step 3: equidistribution

Evaluate the resulting character sum using Chebotarev and the large sieve.

How does this process generalize?

- Build a nilpotent extension by iterated central extensions. This yields a parametrization of G-extensions by tuples of squarefree integers satisfying central embedding problems.
- These central embedding problems get much more complicated, but still satisfy local-to-global and are certainly determined by Frob_p for p dividing the variables of the parametrization.
- In our chosen ordering, a typical extension is a rather large twist of a "minimally ramified central extension". Getting equidistribution of Frobenius in minimally ramified extensions is *very hard*. The key idea of the proof is to exploit the twisting.
- Proof can most likely be made unconditional with a suitably strong large sieve for nilpotent extensions.

Questions?

Thank you for your attention!