TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY, SERIES B
Volume 10, Pages 310–354 (March 3, 2023)
https://doi.org/10.1090/btran/140

ON MALLE'S CONJECTURE FOR NILPOTENT GROUPS

PETER KOYMANS AND CARLO PAGANO

Abstract. We develop an abstract framework for studying the strong form of Malle's conjecture [J. Number Theory 92 (2002), pp. 315–329; Experiment. Math. 13 (2004), pp. 129–135] for nilpotent groups G in their regular representation. This framework is then used to prove the strong form of Malle's conjecture for any nilpotent group G such that all elements of order p are central, where p is the smallest prime divisor of $\#G$.

We also give an upper bound for any nilpotent group G tight up to logarithmic factors, and tight up to a constant factor in case all elements of order p pairwise commute. Finally, we give a new heuristical argument supporting Malle's conjecture in the case of nilpotent groups in their regular representation.

1. Introduction

1.1. **Malle's conjectures.** Let G be a non-trivial, finite group and let K be a number field. In this paper we are interested in the counting function

$$N(X, G, K) := \#\{L/K \text{ Galois} : \mathrm{Gal}(L/K) \cong G,\ |N_{K/\mathbb{Q}}(\Delta(L/K))| \leq X\},$$

where $\Delta(L/K)$ denotes the relative discriminant of L/K. As part of a broader conjecture, Malle [37, 38] conjectured the following.

Conjecture 1 (Strong form of Malle's conjecture.)**.** *Let G be a non-trivial, finite group. Then there exist constants $a(G) \in \mathbb{Q}_{>0}$, $b(G, K) \in \mathbb{Z}_{>0}$ and $c(G, K) > 0$ such that*

$$N(X, G, K) \sim c(G, K) X^{a(G)} (\log X)^{b(G,K)-1}. \tag{1.1}$$

Malle [37] proposed the following recipe for the constant $a(G)$

$$a_{\mathrm{M}}(G) = \frac{p}{(p-1)\#G}$$

with p being the smallest prime divisor of $\#G$. In follow-up work Malle [38] also included a proposal for $b(G, K)$

$$b_{\mathrm{M}}(G, K) = \#\{C \in \mathrm{Conj}(G) : C \text{ non-trivial and } c^p = \mathrm{id}\ \forall c \in C\}/\sim,$$

where $\mathrm{Conj}(G)$ denotes the set of conjugacy classes of G and where two conjugacy classes C and C' are equivalent if there exists $\sigma \in \mathrm{Gal}(\overline{K}/K)$ such that $\sigma * C = C'$. Here the action is given by $\sigma * C = C^{\chi(\sigma)}$ with $\chi : \mathrm{Gal}(\overline{K}/K) \to \hat{\mathbb{Z}}^*$ the cyclotomic character.

Received by the editors February 18, 2022, and, in revised form, July 29, 2022, and October 25, 2022.

2020 *Mathematics Subject Classification.* Primary 11N45, 11R20, 11R29, 11R45; Secondary 11R34, 11R37.

The second author was financially supported by EPSRC Fellowship EP/P019188/1, "Cohen–Lenstra heuristics, Brauer relations, and low-dimensional manifolds".

Malle does not give a recipe to compute the constant $c(G,K)$. In some cases the constant $c(G,K)$ is a product of local densities. This is known for S_3, S_4 and S_5 (and conjectured for S_n) and is commonly referred to as the Malle–Bhargava principle.

In case G is allowed to be an arbitrary finite group, counterexamples are known to the strong form of Malle's conjecture, see the work of Klüners [27]. In the counterexample of Klüners, Malle gives the wrong value of $b_{\mathrm{M}}(G,K)$. However, the value of $a_{\mathrm{M}}(G)$ is still correct in Klüners' counterexample. It is widely believed, but unproven, that Malle's prediction for $a_{\mathrm{M}}(G)$ is correct for all non-trivial, finite groups G (for this reason we will simply write $a(G)$ from now on). Even such a result seems to be out of reach currently, since it implies a positive answer to the inverse Galois problem. Klüners' counterexample led Türkelli [51] to propose a corrected $b(G,K)$.

The strong form of Malle's conjecture (including the more general situation where G is not necessarily considered in its regular representation) has been verified in a limited amount of cases, see [15, 16] for S_3, [53] for G abelian, [10] for $D_4 \subseteq S_4$, [28] for generalized quaternion groups, [5] for S_4, [6] for S_5, [9] for $S_3 \subseteq S_6$, [22] for nonic Heisenberg extensions, [39, 52] for direct products $G \times A$ with $G = S_3, S_4, S_5$ and A abelian. There is also the recent work [4], which counts $D_4 \subseteq S_4$ when the extensions are ordered by Artin conductor instead of discriminant.

It is worth mentioning that the weak form of Malle's conjecture, which asserts that

$$X^{a(G)} \ll N(X,G,K) \ll_\epsilon X^{a(G)+\epsilon},$$

is much better understood. There are no known counterexamples to the weak form, even when G is allowed to be an arbitrary finite group. The weak form is known for nilpotent groups by the work of Klüners–Malle [30]. Alberts [1, 2] and Alberts–O'Dorney [3] made further progress in the solvable case.

1.2. **Main results.** We prove the strong form of Malle's conjecture for a large family of nilpotent groups.

Theorem 1.1. *Let K be a number field and let G be a non-trivial, finite, nilpotent group. Let p be the smallest prime divisor of $\#G$ and assume that all elements of order p are central. Then there exists a constant $c > 0$ such that*

$$N(X,G,K) \sim cX^{\frac{p}{(p-1)\#G}}(\log X)^{b_{\mathrm{M}}(G,K)-1},$$

where $b_{\mathrm{M}}(G,K)$ is the Malle constant, which equals the number of elements of order p divided by $[K(\zeta_p):K]$ in this case.

In the process, we give a parametrization of G-extensions that may prove fruitful for future investigations. We will say more about this parametrization in the next subsection.

Klüners [29] has previously proven an upper bound of the correct order of magnitude in the situation of Theorem 1.1, namely

$$N(X,G,K) \ll X^{\frac{p}{(p-1)\#G}}(\log X)^{b_{\mathrm{M}}(G,K)-1}.$$

This follows from [29, Theorem 1.4] and [29, Lemma 5.13].

Let us construct some 2-groups G satisfying the hypotheses of Theorem 1.1. Take for example any finite, abelian 2-group A. Then there is an action of $\mathbb{Z}/2\mathbb{Z}$ on A by inversion, which gives an action of $\mathbb{Z}/4\mathbb{Z}$ on A by projecting first to $\mathbb{Z}/2\mathbb{Z}$.

If one takes $G = A \rtimes \mathbb{Z}/4\mathbb{Z}$, then G fulfills all the conditions of Theorem 1.1. Note that such G can have arbitrarily large nilpotency class.

We remark that our techniques do not give an explicit handle on the constant $c > 0$ guaranteed by Theorem 1.1. Even in the case where G is cyclic of prime order it is a rather non-trivial task to provide an explicit value of the constant c, see [11]. The following classical result of Wright [53] is an immediate corollary of Theorem 1.1.

Corollary 1.2. *Let K be a number field and let A be a non-trivial, finite, abelian group. Let p be the smallest prime divisor of $\#A$. Then there exists a constant $c > 0$ such that*

$$N(X, A, K) \sim cX^{\frac{p}{(p-1)\#A}}(\log X)^{b(A,K)-1},$$

where $b(A, K)$ is the number of elements of order p divided by $[K(\zeta_p) : K]$.

Our proof is substantially shorter than Wright's proof [53] and makes limited use of class field theory. Our parametrization of G-extensions immediately implies Theorem 1.3.

Theorem 1.3. *Let K be a number field and let G be a non-trivial, finite, nilpotent group. Let p be the smallest prime divisor of $\#G$. Then*

$$N(X, G, K) \ll X^{\frac{p}{(p-1)\#G}}(\log X)^{i(G,K)-1},$$

where $i(G, K)$ is the number of elements of order p in G divided by $[K(\zeta_p) : K]$.

Theorem 1.3 improves, in case of nilpotent groups in the regular representation, on recent work of Klüners–Wang [31] and Klüners [29] (see Proposition 5.8 and the text preceding it). The result in Theorem 1.3 is sharp up to logarithmic factors. In general we have the inequality $i(G, K) \geq b_{\mathrm{M}}(G, K)$. In fact, the upper bound in Theorem 1.3 matches Malle's prediction precisely when we are in the situation of Theorem 1.1.

It should be possible to use our techniques to prove a more general version of Theorem 1.3 valid for arbitrary representations of nilpotent groups, but we shall not pursue this further here. Theorem 1.4 shows that we can achieve a sharp upper bound, up to a constant factor, provided that the elements of order p commute with each other.

Theorem 1.4. *Let K be a number field and let G be a non-trivial, finite, nilpotent group. Let p be the smallest prime divisor of $\#G$. Suppose that all elements of order p commute with each other. Then*

$$N(X, G, K) \ll X^{\frac{p}{(p-1)\#G}}(\log X)^{b_{\mathrm{M}}(G,K)-1}.$$

The corresponding lower bound for $N(X, G, K)$ follows from [1, Corollary 1.7], where we take $T := \{g \in G : g^p = \mathrm{id}\}$ (the assumptions imply that T is indeed abelian and normal). Earlier work of Klüners–Malle [30] provides a slightly weaker lower bound correct up to logarithmic factors.

To give examples of groups G where Theorem 1.4 applies (but Theorem 1.1 does not), consider an $\mathbb{F}_2$ vector space V of dimension 2^n and pick an ordered basis $\{b_0, \dots, b_{2^n-1}\}$. Let $\mathbb{Z}/2^n\mathbb{Z}$ act on V by cycling the ordered basis and extending linearly. Then we can take $G_n = V \rtimes \mathbb{Z}/2^{n+1}\mathbb{Z}$, where $\mathbb{Z}/2^{n+1}\mathbb{Z}$ acts on V by first projecting to $\mathbb{Z}/2^n\mathbb{Z}$. We remark that G_1 is isomorphic to the small group $\langle 16, 3\rangle$ from the GAP database and that G_2 is isomorphic to the small group $\langle 128, 48\rangle$. Note that G_n can have arbitrarily large nilpotency class.

1.3. **Method of proof.** The main innovation of this paper is a new parametrization of G-extensions, which is given in Section 2. The simplest case of the parametrization is that of multiquadratic fields over $\mathbb{Q}$. The straightforward way to parametrize multiquadratic fields is the following. Let $\mathbf{v} = (a_1, \dots, a_n)$ be a vector such that the a_i are squarefree and linearly independent in $\mathbb{Q}^*/\mathbb{Q}^{*2}$. To each such vector $\mathbf{v}$, we associate the field $\mathrm{Par}(\mathbf{v}) = \mathbb{Q}(\sqrt{a_1}, \dots, \sqrt{a_n})$. In this way we have created a map Par to multiquadratic fields of degree 2^n. This map has two convenient properties

- the pre-images $\mathrm{Par}^{-1}(K)$ all have the same size (and are finite);
- the map Par is surjective.

However, we will nevertheless argue that this does not provide a convenient way to count multiquadratic extensions. Indeed, the discriminant of $\mathrm{Par}(\mathbf{v})$ is equal to $\mathrm{rad}(a_1 \cdot \ldots \cdot a_n)^{2^{n-1}}$ up to factors of 2, where rad is the radical of an integer. Let us now consider the counting function

$$N'(X, \mathbb{F}_2^n, \mathbb{Q}) = \sum_{\substack{a_1,\dots,a_n \text{ squarefree} \\ \mathrm{rad}(a_1\cdot\ldots\cdot a_n)^{2^{n-1}} \le X}} 1,$$

which is a good prototype for $N(X, \mathbb{F}_2^n, \mathbb{Q})$. To evaluate $N'(X, \mathbb{F}_2^n, \mathbb{Q})$, we introduce the new variables

$$b_S = \prod_{\substack{p \\ p \mid a_i \Leftrightarrow i \in S}} p$$

for every non-empty subset S of $\{1, \dots, n\}$. Then the variables b_S are squarefree and pairwise coprime. From the variables b_S we can reconstruct the a_i using the formula

$$a_i = \prod_{\substack{S \subseteq \{1,\dots,n\} \\ i \in S}} b_S. \tag{1.2}$$

Under this change of variables, the sum becomes

$$N'(X, \mathbb{F}_2^n, \mathbb{Q}) = \sum_{\substack{b_S \text{ squarefree and pairwise coprime} \\ \prod_{\emptyset \subsetneq S \subseteq \{1,\dots,n\}} b_S^{2^{n-1}} \le X}} 1 = \sum_{d \le X^{1/2^{n-1}}} \mu^2(d) \cdot \omega(d)^{2^n - 1},$$

which can be evaluated using classical techniques in analytic number theory. The crux of this idea is that the discriminant is simply

$$\prod_{\emptyset \subsetneq S \subseteq \{1,\dots,n\}} b_S^{2^{n-1}}$$

in the variables b_S, while the discriminant is a more complicated function in terms of the variables a_i. We remark that this change of variables also plays a role in the study of 2-Selmer ranks and 4-ranks, see for instance [21, 24]. Inspired by this, we parametrize multiquadratic fields by tuples of squarefree integers $(b_S)_{\emptyset \subsetneq S \subseteq \{1,\dots,n\}}$ that are pairwise coprime. The parametrization map is then given by $\mathrm{Par}(a_1, \dots, a_n)$, where the a_i are defined through equation (1.2).

For arbitrary nilpotent extensions (but still over $\mathbb{Q}$) our parametrization keeps track of the image of the inertia subgroup I_p of $\mathrm{Gal}(\overline{\mathbb{Q}_p}/\mathbb{Q}_p)$ in our G-extension for every prime p. This involves the choice of an embedding $i_p : \mathrm{Gal}(\overline{\mathbb{Q}_p}/\mathbb{Q}_p) \to$

$\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. For arbitrary groups G it is unfortunately not true that a homomorphism $\psi : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to G$ is determined by the restrictions $\psi|_{i_p(I_p)}$.

Indeed, in general one needs to know the restriction of ψ to $j_p(I_p)$ for every embedding $j_p : \mathrm{Gal}(\overline{\mathbb{Q}_p}/\mathbb{Q}_p) \to \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ to make this conclusion. However, for nilpotent groups G it suffices to fix one embedding $i_p : \mathrm{Gal}(\overline{\mathbb{Q}_p}/\mathbb{Q}_p) \to \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, as a homomorphism $\psi : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to G$ is completely determined by the restrictions $\psi|_{i_p(I_p)}$ in this case, see Proposition 2.5. Our parametrization uses this property of nilpotent extensions in an essential way.

It is precisely for this reason that the source of our parametrization map is still tuples of squarefree integers that are pairwise coprime and are indexed by the elements in $G - \{\mathrm{id}\}$. We remark that this matches the multiquadratic case, since there is a bijection between non-empty subsets of $\{1, \dots, n\}$ and $\mathbb{F}_2^n - \{\mathrm{id}\}$. This allows us to give a similar formula as above for the discriminant, see Proposition 2.10.

The fact that our parametrization is in terms of tuples of squarefree integers is also most convenient for analytic purposes. For example, it allows us to prove Theorem 1.3 using essentially the trivial bound, namely by ignoring all the embedding problems. This demonstrates the power of the parametrization, since it improves on the existing results in the literature by bounding trivially! In some sense one may view Theorem 1.3 as the correct trivial bound for Malle's conjecture.

It also enables us to write down the various embedding problems explicitly, which we do in Section 3. We use this description to give a heuristic argument why Malle's conjecture is correct for nilpotent extensions. We believe that this may also be of value to possible future work on this topic. Indeed, it makes clear exactly what the difficulties are in proving Malle's conjecture for arbitrary nilpotent extensions. Roughly speaking, at every central extension we pinpoint a certain set of Frobenius elements that need to be equidistributed in a subgroup that we correspondingly identify. This explains how the constant $b_{\mathrm{M}}(G, K)$ arises from the construction of a nilpotent group as successive central extensions, see Section 3.

The parametrization also plays an essential role in the proof of Theorem 1.1 and Theorem 1.4. Using the parametrization and some analysis, we reduce these theorems to counting multicyclic extensions with some extra local conditions. This idea also provides a new and short proof of Wright's theorem [53].

1.4. **Related results.** There is also the related problem of counting the number of degree n extensions with bounded discriminant, which was first treated by Schmidt [43]. His upper bound was drastically improved by Ellenberg–Venkatesh [18], Couveignes [14] and Lemke Oliver–Thorne [36].

Malle's conjecture has strong ties with the Cohen–Lenstra conjectures [12]. There is the classical work of Davenport–Heilbronn [16] on 3-torsion of class groups of quadratic fields, which was later extended by [8] and [49] in the form of a secondary main term. Davenport and Heilbronn obtain their results by counting certain S_3-extensions.

Fouvry and Klüners [20, 21] dealt with the 4-rank of quadratic fields building on earlier work of Gerth [23], and Heath-Brown [24] on 2-Selmer groups. There is also a rich literature on upper bounds for 2-torsion elements in class groups of which we mention [33, 34] for multiquadratic extensions, [7] for S_n-extensions and [45, 46] for monogenic S_n-extensions. Furthermore, the average size of the 2-torsion of S_n-extensions has been determined in [25] conditional on a tail estimate. Over function

fields Malle's conjecture and the Cohen–Lenstra conjectures are better understood due to the results in [17, 19].

Recently, Smith [47, 48] dealt with the 2-part of class groups of quadratic fields, which was extended by the authors to the ℓ-part of class groups of degree ℓ cyclic fields [32]. His techniques can be adapted to give a lower bound for the number of D_{2^n}-extensions of $\mathbb{Q}$ of the correct order of magnitude: this is another instance that highlights the clear ties between Malle's conjecture and the Cohen–Lenstra conjectures. It seems plausible that this can be extended to an asymptotic once one extends the results on ray class groups of Pagano–Sofos [42].

It is natural to wonder whether it is possible to combine Smith's techniques with our parametrization to tackle Malle's conjecture for all finite, nilpotent groups. It is the authors' belief that this should be possible for a wide class of nilpotent groups, and we hope to return to this topic. However, one convenient property of D_{2^n} is that the embedding problem for odd ramified primes p often comes down to the condition that p has residue field degree 1. In terms of our parametrization this is very convenient, since it gives good inductive control over the relevant Frobenius elements.

This fails badly for arbitrary nilpotent G-extensions, where one may face "loops" in the following sense. We might simultaneously want to prove equidistribution of Frob_p in a quotient of G depending on q, while we also need equidistribution of Frob_q in a quotient of G depending on p. The authors currently do not know how to overcome this difficulty.

1.5. **Organization of the paper.** The paper is divided as follows. The core of the paper is Section 2, where we provide a new parametrization of G-extensions. In Section 3 we give a new heuristic in support of Malle's conjecture for nilpotent groups in their regular representation. Our main theorems are proven in Section 5 with Section 4 providing some analytic tools.

2. Arbitrary nilpotent groups and number fields

In our first subsection we restrict ourselves to the case that G is a finite l-group, but we work with an arbitrary number field K. This is then extended to general nilpotent G in the second subsection.

2.1. **Finite l-groups.** We use the abbreviation $[n] := \{1, \dots, n\}$ throughout the paper. For a set S and a prime number l, we write $\mathbb{F}_l^S$ for the free $\mathbb{F}_l$ vector space on the set S. Let us fix a separable closure $\mathbb{Q}^{\mathrm{sep}}$ of $\mathbb{Q}$ once and for all. All our number fields are implicitly taken inside this fixed separable closure. Furthermore, we write $G_K := \mathrm{Gal}(\mathbb{Q}^{\mathrm{sep}}/K)$ for the absolute Galois group of a number field K. Similarly, for each prime p we fix a separable closure $\mathbb{Q}_p^{\mathrm{sep}}$ of $\mathbb{Q}_p$ (which allows us to define $G_K := \mathrm{Gal}(\mathbb{Q}_p^{\mathrm{sep}}/K)$ for any extension K of $\mathbb{Q}_p$ inside $\mathbb{Q}_p^{\mathrm{sep}}$) together with an embedding

$$i_p : \mathbb{Q}^{\mathrm{sep}} \to \mathbb{Q}_p^{\mathrm{sep}}.$$

This yields an inclusion

$$i_p^* : G_{\mathbb{Q}_p} \to G_{\mathbb{Q}}.$$

Let K be a number field, picked inside $\mathbb{Q}^{\mathrm{sep}}$. Denote by Ω_K the set of all places of K. For each finite place $\mathfrak{q}$ in Ω_K, lying above a rational prime q, the restriction of

the map i_q^* provides us with an inclusion

$$i_{\mathfrak{q}}^* : G_{K_{\mathfrak{q}}} \to G_K.$$

Denote by $k_{\mathfrak{q}}$ the residue field of K at $\mathfrak{q}$. Write

$$I_{\mathfrak{q}} := \ker(G_{K_{\mathfrak{q}}} \to G_{k_{\mathfrak{q}}})$$

for the inertia subgroup.

Let now l be any prime number. In what follows the group $\mathbb{F}_l$ will always be implicitly interpreted as a Galois module with trivial action, whenever the notation suggests an implicit action of a group on $\mathbb{F}_l$. We denote by $H^1_{\mathrm{unr}}(G_{K_{\mathfrak{q}}}, \mathbb{F}_l)$ the image, via inflation, of $H^1(G_{k_{\mathfrak{q}}}, \mathbb{F}_l)$ in $H^1(G_{K_{\mathfrak{q}}}, \mathbb{F}_l)$.

For a subset $S \subseteq \Omega_K$, containing all the archimedean places of K, we consider the map

$$\Phi_K(l, S) : H^1(G_K, \mathbb{F}_l) \to \bigoplus_{\mathfrak{q} \in \Omega_K - S} \frac{H^1(G_{K_{\mathfrak{q}}}, \mathbb{F}_l)}{H^1_{\mathrm{unr}}(G_{K_{\mathfrak{q}}}, \mathbb{F}_l)}.$$

In case S consists exactly of the archimedean places, we will denote the resulting map simply by $\Phi_K(l)$. We start by recalling the following classical fact.

Proposition 2.1. *The abelian groups* $\ker(\Phi_K(l))$ *and* $\operatorname{coker}(\Phi_K(l))$ *are finite.*

Proof. By class field theory we have a canonical identification

$$\ker(\Phi_K(l)) = \mathrm{Cl}(K, m_\infty)^\vee[l],$$

where m_∞ is the modulus consisting of all archimedean places. Therefore we have that

$$\#\ker(\Phi_K(l)) = \#\,\mathrm{Cl}(K, m_\infty)^\vee[l] \leq l^{[K:\mathbb{Q}]} \cdot \#\,\mathrm{Cl}(K)^\vee[l],$$

where the first factor $l^{[K:\mathbb{Q}]}$ can be dropped when l is odd. Therefore the finiteness of $\ker(\Phi_K(l))$ follows from the finiteness of $\mathrm{Cl}(K)$. The finiteness of $\operatorname{coker}(\Phi_K(l))$ is established in [44, Theorem 5, eq. (14) and (16)]. In the notation of Shafarevich, we have

$$(\mathfrak{U}/\mathfrak{U}_S^l)^\vee = \bigoplus_{\mathfrak{q} \in \Omega_K - S} \frac{H^1(G_{K_{\mathfrak{q}}}, \mathbb{F}_l)}{H^1_{\mathrm{unr}}(G_{K_{\mathfrak{q}}}, \mathbb{F}_l)}, \quad (J/J^l K)^\vee = H^1(G_K, \mathbb{F}_l)$$

and the dual of his homomorphism ϕ_l is our $\Phi_K(l, S)$. Combining [44, Theorem 5, eq. (14) and (16)] yields

$$\dim_{\mathbb{F}_l} \operatorname{coker}(\Phi_K(l)) = \sigma(S),$$

which is a finite number defined on [44, p. 129]. □

The following important fact falls as an immediate consequence of Proposition 2.1.

Proposition 2.2. *There exists a finite set of places* S*, containing all archimedean places, such that* $\Phi_K(l, S)$ *is surjective.*

Proof. Write A for the finite subset of archimedean places of Ω_K. Thanks to Proposition 2.1 we can find a finite set of places $A \subseteq S \subseteq \Omega_K$ such that the natural map

$$\bigoplus_{\mathfrak{q} \in S - A} \frac{H^1(G_{K_{\mathfrak{q}}}, \mathbb{F}_l)}{H^1_{\mathrm{unr}}(G_{K_{\mathfrak{q}}}, \mathbb{F}_l)} \to \operatorname{coker}(\Phi_K(l))$$

is surjective. This implies that for any vector

$$v \in \bigoplus_{\mathfrak{q} \in \Omega_K - S} \frac{H^1(G_{K_\mathfrak{q}}, \mathbb{F}_l)}{H^1_{\mathrm{unr}}(G_{K_\mathfrak{q}}, \mathbb{F}_l)},$$

we can find $w \in \bigoplus_{\mathfrak{q} \in S-A} \frac{H^1(G_{K_\mathfrak{q}}, \mathbb{F}_l)}{H^1_{\mathrm{unr}}(G_{K_\mathfrak{q}}, \mathbb{F}_l)}$ and a global character $\chi \in H^1(G_K, \mathbb{F}_l)$ such that

$$\Phi_K(l)(\chi) = v + w.$$

Therefore, since w is entirely supported in S, we conclude that

$$\Phi_K(l, S)(\chi) = v.$$

Hence we have shown that the map $\Phi_K(l, S)$ is surjective with this choice of S, which is precisely the desired conclusion. □

Of course if a set S as in Proposition 2.2 works, then any larger set works as well. We fix once and for all a finite set $S_{\mathrm{clean}}(l)$ as in Proposition 2.2, making sure that it also contains all places above l. We denote by $\widetilde{\Omega}_K(l)$ the subset of $\Omega_K - S_{\mathrm{clean}}(l)$ with

$$\frac{H^1(G_{K_\mathfrak{q}}, \mathbb{F}_l)}{H^1_{\mathrm{unr}}(G_{K_\mathfrak{q}}, \mathbb{F}_l)} \neq 0. \tag{2.1}$$

For $\mathfrak{q} \in \Omega_K - S_{\mathrm{clean}}(l)$, it follows from local class field theory that equation (2.1) is equivalent to the condition

$$\#\left(\mathcal{O}_K/\mathfrak{q}\right) \equiv 1 \bmod l.$$

Write $K^{\mathrm{pro}-l}$ for the compositum of all finite Galois extensions L of K with $[L : K]$ a power of l.

Proposition 2.3. *Let $\mathfrak{q} \in \Omega_K$ be a finite place coprime to l such that $\#\left(\mathcal{O}_K/\mathfrak{q}\right) \not\equiv 1 \bmod l$. Then $\mathfrak{q}$ is unramified in every finite extension L/K inside $K^{\mathrm{pro}-l}$.*

Proof. It suffices to prove the proposition locally at $\mathfrak{q}$. Take a positive integer f and let $K_{\mathfrak{q}^f}$ be the unique unramified extension of $K_\mathfrak{q}$ of degree equal to f with residue field denoted by $k_{\mathfrak{q}^f}$. Then if we have a cyclic totally ramified degree l extension of $K_{\mathfrak{q}^f}$ it follows from local class field theory and the fact that $\mathfrak{q}$ is coprime to l

$$\#k_{\mathfrak{q}^f} = \#k_\mathfrak{q}^f \equiv 1 \bmod l.$$

Now, if f is a power of l itself, we conclude that $\#k_\mathfrak{q}$ is already congruent to 1 modulo l contrary to our assumption that $\#\left(\mathcal{O}_K/\mathfrak{q}\right) \not\equiv 1 \bmod l$. □

We remark that Proposition 2.3 certainly applies to any place $\mathfrak{q} \in \Omega_K - S_{\mathrm{clean}}(l) - \widetilde{\Omega}_K(l)$. Let $\mathfrak{q} \in \widetilde{\Omega}_K(l)$. Thanks to Proposition 2.2, there exists a character

$$\chi_\mathfrak{q} \in H^1(G_K, \mathbb{F}_l)$$

such that $\Phi_K(l, S_{\mathrm{clean}}(l))(\chi_\mathfrak{q})$ has non-trivial coordinate precisely at $\mathfrak{q}$ and at no other places in $\Omega_K - S_{\mathrm{clean}}(l)$. Fix once and for all such a choice of $\chi_\mathfrak{q}$ for each $\mathfrak{q} \in \widetilde{\Omega}_K(l)$. By construction

$$\{\chi_\mathfrak{q}\}_{\mathfrak{q} \in \widetilde{\Omega}_K(l)}$$

is a linearly independent set. Furthermore by Proposition 2.1 we obtain that the subspace

$$\langle \{\chi_\mathfrak{q}\}_{\mathfrak{q} \in \widetilde{\Omega}_K(l)} \rangle \subseteq H^1(G_K, \mathbb{F}_l)$$

has finite index. Additionally, there exist a positive integer t and a basis

$$J := \{\chi_i\}_{i=1}^{t} \subseteq \ker(\Phi_K(l, S_{\text{clean}}(l)))$$

such that $J \cup \{\chi_{\mathfrak{q}}\}_{\mathfrak{q} \in \widetilde{\Omega}_K(l)}$ is a *basis* of $H^1(G_K, \mathbb{F}_l)$. Fix once and for all such a choice of J. We denote by

$$\mathcal{B}(K, l) := J \cup \{\chi_{\mathfrak{q}}\}_{\mathfrak{q}}$$

this fixed choice of a basis. Put

$$\mathcal{G}_K^{\text{pro}-l} := \text{Gal}(K^{\text{pro}-l}/K).$$

For each finite place $\mathfrak{q} \in \Omega_K$ we denote by

$$I_{\mathfrak{q}}(l) := \text{proj}(G_K \to \mathcal{G}_K^{\text{pro}-l}) \circ i_q^*(I_{\mathfrak{q}}).$$

We have the following basic fact.

Proposition 2.4. *The group $I_{\mathfrak{q}}(l)$ is pro-cyclic for each finite place $\mathfrak{q} \in \Omega_K$ coprime to l.*

Proof. Let L be a non-archimedean local field of characteristic 0 and write p for its residue characteristic. Let d be a positive integer coprime to p. Then every finite totally ramified extension of L of degree equal to d can be obtained as $L(\sqrt[d]{\pi})$ for π a uniformizer of L. For an elementary proof of this well-known fact see [32, Proposition A.5]. Applying this repeatedly to all finite unramified extensions of L, we conclude that

$$\frac{I_L}{I_L^{\text{wild}}}$$

is a pro-cyclic group, where I_L is the inertia subgroup and I_L^{wild} is the wild inertia subgroup. Since $\mathfrak{q}$ is coprime to l, we conclude that $I_{\mathfrak{q}}(l)$ is a quotient of the pro-cyclic group $\frac{I_{\mathfrak{q}}}{I_{\mathfrak{q}}^{\text{wild}}}$, which gives in particular the desired conclusion. □

Remark 1. Since the group $I_{\mathfrak{q}}(l)$ is a pro-cyclic pro-l group, it is either isomorphic to a finite group of order a power of l or isomorphic to $\mathbb{Z}_l$. In case a primitive l-th root of unity ζ_l is in K, then we claim that $I_{\mathfrak{q}}(l)$ is infinite. Indeed, we have the infinitely ramified subextension

$$K(\sqrt[\infty]{1},\ \sqrt[\infty]{\alpha})$$

of $K^{\text{pro}-l}/K$ (which is contained in $K^{\text{pro}-l}/K$ exactly because ζ_l is in K) given by any α in K with $v_{\mathfrak{q}}(\alpha) = 1$. Therefore we conclude that if K possesses a non-trivial l-th root of unity, then

$$I_{\mathfrak{q}}(l) \simeq_{\text{top.gr.}} \mathbb{Z}_l$$

for every finite place $\mathfrak{q} \in \Omega_K$ coprime to l. Instead if ζ_l is not in K, we observe that Proposition 2.3 shows that the group $I_{\mathfrak{q}}(l)$ is trivial in case $\mathfrak{q}$ is a finite place of Ω_K that is coprime to l and satisfies $(\mathcal{O}_K/\mathfrak{q}) \not\equiv 1 \bmod l$.

We fix once and for all a topological generator $\sigma_{\mathfrak{q}}$ of $I_{\mathfrak{q}}(l)$ for all $\mathfrak{q} \in \widetilde{\Omega}_K(l)$ in the following manner. Observe that $\chi_{\mathfrak{q}}(\sigma_{\mathfrak{q}}) \neq 0$ for any topological generator of $I_{\mathfrak{q}}(l)$, since the character $\chi_{\mathfrak{q}}$ ramifies at $\mathfrak{q}$. Hence we can always pick a generator $\sigma_{\mathfrak{q}}$ with the normalization $\chi_{\mathfrak{q}}(\sigma_{\mathfrak{q}}) = 1$. We make such a choice of $\sigma_{\mathfrak{q}}$ once and for all. In case $\mathfrak{q} \in \Omega_K - S_{\text{clean}}(l) - \widetilde{\Omega}_K(l)$, then the group $I_{\mathfrak{q}}(l)$ is trivial by Proposition 2.3, and we declare $\sigma_{\mathfrak{q}} := \text{id}$.

Now it follows by construction that

$$\chi(\sigma_{\mathfrak{q}}) = \delta_{\chi_{\mathfrak{q}}}(\chi)$$

for each $\chi \in \mathcal{B}(K,l)$ and $\mathfrak{q} \in \widetilde{\Omega}_K(l)$, where δ denotes the Kronecker delta function. Therefore we can complete the set

$$\{\sigma_{\mathfrak{q}}\}_{\mathfrak{q}\in\widetilde{\Omega}_K(l)}$$

to a minimal set of generators

$$\{\sigma_i\} \cup \{\sigma_{\mathfrak{q}}\}_{\mathfrak{q}\in\widetilde{\Omega}_K(l)},$$

which is dual to the basis $\mathcal{B}(K,l)$, i.e.

$$\begin{aligned}
&\chi_i(\sigma_{\mathfrak{q}}) = 0 = \chi_{\mathfrak{q}}(\sigma_i) && \text{for each } i \in [t] \text{ and } \mathfrak{q} \in \widetilde{\Omega}_K(l),\\
&\chi_i(\sigma_j) = \delta_i(j) && \text{for each } i,j \in [t],\\
&\chi_{\mathfrak{q}}(\sigma_{\mathfrak{q}'}) = \delta_{\mathfrak{q}}(\mathfrak{q}') && \text{for every } \mathfrak{q},\mathfrak{q}' \in \widetilde{\Omega}_K(l).
\end{aligned}$$

We denote by

$$\mathcal{B}^\vee(K,l) := \{\sigma_i\}_{i=1}^t \cup \{\sigma_{\mathfrak{q}}\}_{\mathfrak{q}\in\widetilde{\Omega}_K(l)}.$$

Then we have the following.

Proposition 2.5. *The set $\mathcal{B}^\vee(K,l)$ is a set of topological generators of $\mathcal{G}_K^{\mathrm{pro}-l}$.*

Proof. By construction $\mathcal{B}^\vee(K,l)$ topologically generates $\mathrm{Gal}(L/K)$, where L is by definition the compositum of all degree l cyclic extensions of K. But since $\mathcal{G}_K^{\mathrm{pro}-l}$ is a pro-l-group, we see that $\mathrm{Gal}(L/K)$ is also the quotient of $\mathcal{G}_K^{\mathrm{pro}-l}$ by its Frattini subgroup. This implies the proposition, since a subset S topologically generates if and only if it topologically generates modulo the Frattini subgroup. □

Let L/K be a finite Galois extension inside $K^{\mathrm{pro}-l}$ with Galois group $G := \mathrm{Gal}(L/K)$. Take a 2-cocycle θ

$$\theta : G^2 \to \mathbb{F}_l$$

with the requirement that $\theta(\mathrm{id},\mathrm{id}) = 0$. By the same argument as before, every class in $H^2(G,\mathbb{F}_l)$ can be represented by such a 2-cocycle θ. Consider the group

$$(\mathbb{F}_l \times G, *_\theta),$$

where the group law is

$$(a_1,g_1) *_\theta (a_2,g_2) = (a_1 + a_2 + \theta(g_1,g_2), g_1g_2).$$

Our assumption on θ ensures that $(0,\mathrm{id})$ is the trivial element of $(\mathbb{F}_l \times G, *_\theta)$.

Proposition 2.6. *Let l be a prime number. Let K be a number field and let L be an extension with $G = \mathrm{Gal}(L/K)$ a finite l-group. Suppose that θ is non-trivial in $H^2(G,\mathbb{F}_l)$.*
(a) *The natural projection map $\pi : G_K \twoheadrightarrow G$ can be lifted to a surjective homomorphism*

$$\psi : G_K \to (\mathbb{F}_l \times G, *_\theta)$$

if and only if θ is trivial in $H^2(G_{K_v},\mathbb{F}_l)$ for each place v that ramifies in L/K. Moreover, if ψ is a lift, then the $\mathbb{F}_l$-coordinate of ψ is a continuous 1-cochain $\phi(\psi) : G_K \to \mathbb{F}_l$ with

$$\mathrm{d}(-\phi(\psi)) = \theta.$$

Conversely, given any such continuous 1*-cochain* $\phi : G_K \to \mathbb{F}_l$ *with* $\mathrm{d}(-\phi) = \theta$*, the assignment*

$$\psi(\phi)(g) = (\phi(g), \pi(g))$$

is an epimorphism lifting the canonical projection $\pi : G_K \to G$ *to an epimorphism* $G_K \to (\mathbb{F}_l \times G, *_\theta)$*. The two assignments are mutual inverses.*
(b) *In case one has a lift* ψ *as in part* (a)*, then there is a unique one satisfying*

$$\phi(\psi)(\sigma) = 0 \text{ for all } \sigma \in \mathcal{B}^\vee(K, l).$$

Proof. We start with part (a). We claim that a map $\phi(\psi) : G_K \to \mathbb{F}_l$ is the first coordinate of a homomorphism

$$\psi : G_K \twoheadrightarrow (\mathbb{F}_l \times G, *_\theta), \quad g \mapsto (\phi(\psi)(g), \pi(g))$$

if and only if

$$\mathrm{d}(-\phi(\psi)) = \theta.$$

Indeed, since ψ is a homomorphism, we obtain

$$\begin{aligned}(\phi(\psi)(g_1g_2), \pi(g_1g_2)) = \psi(g_1g_2) &= \psi(g_1)\psi(g_2) \\ &= (\phi(\psi)(g_1) + \phi(\psi)(g_2) + \theta(g_1, g_2), \pi(g_1g_2)),\end{aligned}$$

which is equivalent to

$$\mathrm{d}(-\phi(\psi))(g_1, g_2) = \phi(\psi)(g_1g_2) - \phi(\psi)(g_1) - \phi(\psi)(g_2) = \theta(g_1, g_2)$$

as claimed.

Now suppose that there exists $\phi(\psi)$ with $\mathrm{d}(-\phi(\psi)) = \theta$. We claim that $(\phi(\psi), \pi)$ is surjective. Let us first show that all characters $(\mathbb{F}_l \times G, *_\theta) \to \mathbb{F}_l$ must come from G. If not, then the kernel of such a hypothetical character provides a splitting of θ, which implies that θ is trivial contrary to our assumptions. Hence, since π is surjective, the image of $(\phi(\psi), \pi)$ generates modulo the Frattini of $(\mathbb{F}_l \times G, *_\theta)$, and therefore equals $(\mathbb{F}_l \times G, *_\theta)$.

Furthermore, we see that the lifting ψ exists if and only if the inflation of θ to $H^2(G_K, \mathbb{F}_l)$ is trivial if and only if θ is trivial in $H^2(G_{K_v}, \mathbb{F}_l)$ for every place v of K. Thanks to [32, Proposition 4.4], the vanishing at the finite places unramified in L/K is already guaranteed: notice that in [32, Section 4] the number l is assumed to be an odd prime but Proposition 4.4 of [32] also holds for $l = 2$ with an identical proof. Let us now show θ is locally trivial at any unramified, infinite place of L/K. If v is an archimedean complex place, then this is clear. If v is instead an archimedean real place and the extension L/K is unramified at v, then this means that $L_w = K_v$ for each place w of L above v and thus the embedding problem is locally trivial at v. This ends the proof of part (a).

We now prove part (b). The uniqueness follows at once from part (a) combined with the fact that $\mathcal{B}^\vee(K, l)$ is a system of topological generators for $\mathcal{G}_K^{\mathrm{pro}-l}$. Indeed, an epimorphism ψ as in part (a) is entirely determined by its values on a set of topological generators. We next show the existence: here we will take advantage of the fact that $\mathcal{B}^\vee(K, l)$ is a minimal set of topological generators. Take a map $\phi : G_K \to \mathbb{F}_l$ satisfying

$$\mathrm{d}(-\phi) = \theta.$$

The resulting epimorphism $\psi(\phi) : \mathcal{G}_K^{\mathrm{pro}-l} \twoheadrightarrow (\mathbb{F}_l \times G, *_\theta)$ corresponds to a finite extension. As such we conclude that $\phi(\sigma_{\mathfrak{q}}) = 0$ for all but finitely many $\mathfrak{q} \in \widetilde{\Omega}_K(l)$.

Therefore the sum

$$\sum_{i=1}^{t} \phi(\sigma_i) \cdot \chi_i + \sum_{\mathfrak{q} \in \widetilde{\Omega}_K(l)} \phi(\sigma_{\mathfrak{q}}) \cdot \chi_{\mathfrak{q}}$$

is a well-defined element of $H^1(G_K, \mathbb{F}_l)$. Hence, since $\mathcal{B}(K,l)$ and $\mathcal{B}^\vee(K,l)$ are dual to each other, we obtain that

$$\phi - \sum_{i=1}^{t} \phi(\sigma_i) \cdot \chi_i - \sum_{\mathfrak{q} \in \widetilde{\Omega}_K(l)} \phi(\sigma_{\mathfrak{q}}) \cdot \chi_{\mathfrak{q}}$$

vanishes at σ_i for all $i \in [t]$ and at $\sigma_{\mathfrak{q}}$ for all $\mathfrak{q} \in \widetilde{\Omega}_K(l)$. This ends the proof of part (b). □

We denote the unique 1-cochain as in part (b) of Proposition 2.6 by $\phi(G,\theta)$. In case θ is trivial as a 2-cocycle, then we choose $\phi(G,\theta) := 0$. With this choice, we see that $\phi(G,\theta)$ satisfies part (b) of Proposition 2.6, since the trivial character is the unique cyclic degree l character vanishing at all $\sigma \in \mathcal{B}^\vee(K,l)$.

Denote by $\mathcal{S}_l$ the set $\{0,1\}^{[t]} \times \mathcal{S}_l'$, where $\mathcal{S}_l'$ is the set of squarefree integral ideals in $\mathcal{O}_K$ entirely supported in $\widetilde{\Omega}_K(l)$. To an element $(T,\mathfrak{b}) \in \mathcal{S}_l$ we attach the character

$$\chi_{(T,\mathfrak{b})} := \sum_{\substack{i \in [t] \\ \pi_i(T)=1}} \chi_i + \sum_{\substack{\mathfrak{q} \mid \mathfrak{b} \\ \mathfrak{q} \in \widetilde{\Omega}_K(l)}} \chi_{\mathfrak{q}}.$$

Two pairs $(T,\mathfrak{b}), (T',\mathfrak{b}') \in \mathcal{S}_l$ are said to be *coprime* in case $\mathfrak{b}, \mathfrak{b}'$ are coprime ideals and there does not exist a $j \in [t]$ such that $\pi_j(T) = \pi_j(T') = 1$. Formulated differently, the two pairs are coprime exactly when

$$\{\sigma \in \mathcal{B}^\vee(K,l) : \chi_{(T,\mathfrak{b})}(\sigma) \neq 0\} \cap \{\sigma' \in \mathcal{B}^\vee(K,l) : \chi_{(T',\mathfrak{b}')}(\sigma') \neq 0\} = \emptyset.$$

Let V be any finite set. We denote by $\mathrm{Prim}(\mathcal{S}_l^{\mathbb{F}_l^V - \{(0,\ldots,0)\}})$ the subset of $\mathcal{S}_l^{\mathbb{F}_l^V - \{(0,\ldots,0)\}}$ consisting of vectors possessing pairwise coprime coordinates. We conclude this subsection by giving a bijection

$$\mathrm{Pow}_l(V) : \mathrm{Prim}(\mathcal{S}_l^{\mathbb{F}_l^V - \{(0,\ldots,0)\}}) \to H^1(G_K, \mathbb{F}_l)^V,$$

which sends a vector $(v_g)_{g \in \mathbb{F}_l^V - \{(0,\ldots,0)\}}$ to

$$\mathrm{Pow}_l(V)((v_g)_{g \in \mathbb{F}_l^V - \{(0,\ldots,0)\}}) := \left(\sum_{g \in \mathbb{F}_l^V - \{(0,\ldots,0)\}} \pi_j(g) \cdot \chi_{v_g} \right)_{j \in V},$$

where π_j is the projection map on the j-th coordinate. Let us prove that this map is indeed a bijection.

Proposition 2.7. *Let V be a finite set. Then the map $\mathrm{Pow}_l(V)$ is a bijection.*

Proof. Assume without loss of generality that $V = [r]$. To a vector $(\chi_1, \ldots, \chi_r)$ in $H^1(G_K, \mathbb{F}_l)^r$ we attach a point

$$\Pi_l(\chi_1, \ldots, \chi_r) := (v_g(1), v_g(2))_{g \in \mathbb{F}_l^r - \{(0,\ldots,0)\}}$$

in $\mathrm{Prim}(\mathcal{S}_l^{\mathbb{F}_l^r - \{(0,\ldots,0)\}})$ as follows. For each $\mathfrak{q} \in \widetilde{\Omega}_K(l)$ we let $\mathfrak{q}$ divide the entry $v_g(2)$ if and only if

$$(\chi_1(\sigma_{\mathfrak{q}}), \ldots, \chi_r(\sigma_{\mathfrak{q}})) = g.$$

Likewise for each $j \in [t]$ we put $\pi_j(v_g(1)) = 1$ if and only if

$$(\chi_1(\sigma_j), \dots, \chi_r(\sigma_j)) = g.$$

By construction $(v_g(1), v_g(2))_{g \in \mathbb{F}_l^r - \{(0,\dots,0)\}}$ is in $\mathrm{Prim}(\mathcal{S}_l^{\mathbb{F}_l^r - \{(0,\dots,0)\}})$. Using that $\mathcal{B}^\vee(K,l)$ is a system of topological generators, we deduce that

$$\mathrm{Pow}_l([r])((v_g(1), v_g(2))_{g \in \mathbb{F}_l^r - \{(0,\dots,0)\}}) = (\chi_1, \dots, \chi_r),$$

since the equality holds by construction when evaluated in an element of $\mathcal{B}^\vee(K,l)$. Conversely let $\sigma \in \mathcal{B}^\vee(K,l)$ and let $(v_g)_{g \in \mathbb{F}_l^r - \{(0,\dots,0)\}} \in \mathrm{Prim}(\mathcal{S}_l^{\mathbb{F}_l^r - \{(0,\dots,0)\}})$. There exists at most one $g_0 \in \mathbb{F}_l^r - \{(0,\dots,0)\}$ such that $\chi_{v_{g_0}}(\sigma) \neq 0$. Suppose that such a g_0 exists. Then

$$\mathrm{Pow}_l([r])((v_g)_{g \in \mathbb{F}_l^r - \{(0,\dots,0)\}})(\sigma) = g_0,$$

which implies that

$$\Pi_l \circ \mathrm{Pow}_l([r])((v_g)_{g \in \mathbb{F}_l^r - \{(0,\dots,0)\}}) = (v_g)_{g \in \mathbb{F}_l^r - \{(0,\dots,0)\}}.$$

Hence Π_l and $\mathrm{Pow}_l([r])$ are mutual inverses, which finishes the proof of the proposition. □

We now have the necessary tools to parametrize nilpotent extensions. We carry this out in the next subsection.

2.2. **The parametrization in general.** Let $r \in \mathbb{Z}_{\geq 1}$. A sequence of pairs

$$\{(G_i, \theta_i)\}_{i \in [r]}$$

is called an *admissible sequence* if it satisfies the following inductive rules:

- G_0 is the trivial group by convention. Furthermore, G_i is an l-group and $\theta_i : G_{i-1}^2 \to \mathbb{F}_l$ is a 2-cocycle with $\theta_i(\mathrm{id}, \mathrm{id}) = 0$ for each $i \in [r]$;
- we have
$$G_i = (\mathbb{F}_l \times G_{i-1}, *_{\theta_i})$$
for all $i \in [r]$;
- θ_i is the zero map if and only if the class of θ in $H^2(G_{i-1}, \mathbb{F}_l)$ is trivial.

For the remainder of this section we fix an admissible sequence $\{(G_i, \theta_i)\}_{i \in [r]}$. Set

$$G := G_r.$$

The aim of this section is to construct a surjective map

$$P_G : \mathrm{Prim}(\mathcal{S}_l^{G - \{\mathrm{id}\}}) \twoheadrightarrow \mathrm{Epi}_{\mathrm{top.gr.}}(\mathcal{G}_K^{\mathrm{pro}-l}, G) \cup \{\bullet\},$$

which restricts to a bijection between

$$\mathrm{Prim}(\mathcal{S}_l^{G - \{\mathrm{id}\}})(\mathrm{solv.}) := P_G^{-1}(\mathrm{Epi}_{\mathrm{top.gr.}}(\mathcal{G}_K^{\mathrm{pro}-l}, G))$$

and $\mathrm{Epi}_{\mathrm{top.gr.}}(\mathcal{G}_K^{\mathrm{pro}-l}, G)$. Furthermore, we explain how to read the ramification data on the right hand side from the left hand side of this parametrization.

We start by defining a map

$$\tilde{P}_G : H^1(G_K, \mathbb{F}_l)^r \to \mathrm{Epi}_{\mathrm{top.gr.}}(\mathcal{G}_K^{\mathrm{pro}-l}, G) \cup \{\bullet\}$$

as follows. Let $v := (\chi_1, \dots, \chi_r)$ be an element of $H^1(G_K, \mathbb{F}_l)^r$. If χ_1 is the trivial character, we declare $\tilde{P}_G(v) = \bullet$. So we assume from now on that χ_1 is non-trivial. Equivalently,

$$\chi_1 \in \mathrm{Epi}_{\mathrm{top.gr.}}(\mathcal{G}_K^{\mathrm{pro}-l}, G_1).$$

Hence $\chi_1^*(\theta_2) = \theta_2(\chi_1(-), \chi_1(-))$ is now a 2-cocycle on G_K. If $\chi_1^*(\theta_2)$ is not trivial in $H^2(G_K, \mathbb{F}_l)$, then we declare $\tilde{P}_G(v) = \bullet$. Now assume that $\chi_1^*(\theta_2)$ is zero in $H^2(G_K, \mathbb{F}_l)$; we distinguish two cases. If θ_2 is already a trivial 2-cocycle on G_1, we have that $\phi(G_1, \chi_1^*(\theta_2)) = 0$ and

$$(\chi_2, \chi_1) \in \mathrm{Epi}_{\mathrm{top.gr.}}(\mathcal{G}_K^{\mathrm{pro}-l}, G_2)$$

if and only if χ_1 and χ_2 are linearly independent. In case χ_1 and χ_2 are linearly dependent, we set $\tilde{P}_G(v) = \bullet$ and otherwise we proceed. If instead θ_2 is a non-trivial 2-cocycle on G_1, we always have that

$$(\phi(G_1, \chi_1^*(\theta_2)) + \chi_2, \chi_1) \in \mathrm{Epi}_{\mathrm{top.gr.}}(\mathcal{G}_K^{\mathrm{pro}-l}, G_2)$$

by Proposition 2.6.

Now we continue in this fashion inductively. At step $i < r$ we have either already assigned v to $\bullet$ or we have obtained an epimorphism $\psi_i \in \mathrm{Epi}_{\mathrm{top.gr.}}(\mathcal{G}_K^{\mathrm{pro}-l}, G_i)$. Then we get a 2-cocycle $\psi_i^*(\theta_{i+1})$, which gives a class in $H^2(G_K, \mathbb{F}_l)$. If this class is non-trivial in $H^2(G_K, \mathbb{F}_l)$, we send v to $\bullet$.

In case $\psi_i^*(\theta_{i+1})$ is trivial in $H^2(G_K, \mathbb{F}_l)$, we distinguish two cases. If $\psi_i^*(\theta_{i+1})$ is already trivial in $H^2(G_i, \mathbb{F}_l)$, we have that $\phi(G_i, \psi_i^*(\theta_{i+1})) = 0$. Then

$$(\chi_{i+1}, \psi_i) \in \mathrm{Epi}_{\mathrm{top.gr.}}(\mathcal{G}_K^{\mathrm{pro}-l}, G_{i+1})$$

if and only if χ_{i+1} is linearly independent from the characters χ_j satisfying

$$\phi(G_{j-1}, \psi_{j-1}^*(\theta_j)) = 0.$$

Indeed, observe that in this case we have $G_{i+1} = \mathbb{F}_l \times G_i$, therefore, since we have surjectivity if and only if we have surjectivity modulo the Frattini, we need to have that χ_{i+1} is linearly independent from the characters coming from G_i. Thus our claim comes down to the claim that such characters are precisely spanned by the set

$$\{\chi_j\}_{j \leq i : \theta_j = 0}.$$

This is justified by the following observation. Let H be a finite l-group and let $\theta : H^2 \to \mathbb{F}_l$ be a 2-cocycle. Then θ is trivial in $H^2(H, \mathbb{F}_l)$ if and only if the dimension of $(\mathbb{F}_l \times H, *_\theta)$ modulo its Frattini subgroup is one larger than that of H modulo its Frattini subgroup. To see the non-trivial direction observe that if one takes a character $\chi : (\mathbb{F}_l \times H, *_\theta) \to \mathbb{F}_l$ that does not come from H, then its kernel gives a splitting of the sequence.

If χ_{i+1} is linearly dependent on these characters χ_j, we send v to $\bullet$. Otherwise we go to step $i + 1$.

Now suppose that θ_{i+1} is a non-trivial class of $H^2(G_i, \mathbb{F}_l)$. Then we always obtain by means of Proposition 2.6 a new epimorphism

$$(\phi(G_i, \psi_i^*(\theta_{i+1})) + \chi_{i+1}, \psi_i) \in \mathrm{Epi}_{\mathrm{top.gr.}}(\mathcal{G}_K^{\mathrm{pro}-l}, G_{i+1})$$

and we go to step $i+1$. Continuing in this fashion we obtain either $\bullet$ or an element of

$$\mathrm{Epi}_{\mathrm{top.gr.}}(\mathcal{G}_K^{\mathrm{pro}-l}, G),$$

which is by definition $\tilde{P}_G(v)$. We put

$$P_G := \tilde{P}_G \circ \mathrm{Pow}_l([r]).$$

We remark that $\phi(G_i, \theta)$ is only defined in case G_i is a Galois group. Fortunately, this small abuse of notation does not present any issues. Indeed, the epimorphism ψ_i realizes the implicit identification between G_i and the corresponding Galois group.

We additionally remark that in the construction of the map P_G we have implicitly used that G is set-theoretically defined to be $\mathbb{F}_l^r$ with the identity element being $(0, \dots, 0)$, which is a consequence of our convention that 2-cocycles vanish on $(\mathrm{id}, \mathrm{id})$. Hence it makes sense to invoke the map $\mathrm{Pow}_l([r])$.

Proposition 2.8. *The map*

$$P_G : \mathrm{Prim}(\mathcal{S}_l^{G-\{\mathrm{id}\}}) \twoheadrightarrow \mathrm{Epi}_{\mathrm{top.gr.}}(\mathcal{G}_K^{\mathrm{pro}-l}, G) \cup \{\bullet\}$$

is a surjection, which restricts to a bijection between $\mathrm{Prim}(\mathcal{S}_l^{G-\{\mathrm{id}\}})(\mathrm{solv.})$ *and surjective homomorphisms* $\mathrm{Epi}_{\mathrm{top.gr.}}(\mathcal{G}_K^{\mathrm{pro}-l}, G)$.

Proof. This follows upon combining Proposition 2.6 and Proposition 2.7. □

The parametrization P_G allows us to read off very neatly the image of the elements $\{\sigma_{\mathfrak{q}}\}_{\mathfrak{q} \in \widetilde{\Omega}_K(l)}$ from the tuples of squarefree ideals.

Proposition 2.9. *Let* $(v_g(1), v_g(2))_{g \in G-\{\mathrm{id}\}}$ *be an element of* $\mathrm{Prim}(\mathcal{S}_l^{G-\{\mathrm{id}\}})(\mathrm{solv.})$. *Let* $\mathfrak{q} \in \Omega_K - S_{\mathrm{clean}}(l)$. *If* $\mathfrak{q} \mid v_{g_0}(2)$ *for a (necessarily) unique* $g_0 \in G - \{\mathrm{id}\}$ *then*

$$P_G((v_g(1), v_g(2))_{g \in G-\{\mathrm{id}\}})(\sigma_{\mathfrak{q}}) = g_0.$$

If $\mathfrak{q}$ *does not divide any of the elements of the vector* $(v_g(2))_{g \in G-\{\mathrm{id}\}}$, *then*

$$P_G((v_g(1), v_g(2))_{g \in G-\{\mathrm{id}\}})(\sigma_{\mathfrak{q}}) = \mathrm{id}.$$

Proof. For the case $\mathfrak{q} \in \widetilde{\Omega}_K(l)$ the conclusion follows immediately from the fact that the map Π_l in the proof of Proposition 2.7 is inverse to the map $\mathrm{Pow}_l([r])$. Otherwise, we have that $\mathfrak{q} \in \Omega_K - S_{\mathrm{clean}}(l) - \widetilde{\Omega}_K(l)$ so that $\sigma_{\mathfrak{q}} = \mathrm{id}$ by definition. By construction of $\mathcal{S}_l$ it follows that such $\mathfrak{q}$ do not divide any $v_g(2)$. Hence we are always in the second case of the proposition. Therefore the statement also holds for such $\mathfrak{q}$. □

We next read off the value of the discriminant under the bijection P_G. For any continuous homomorphism ψ of G_K with values in some finite group, we denote by $\mathrm{Disc}(\psi)$ the relative discriminant (which is an ideal of $\mathcal{O}_K$) of the corresponding extension. For a non-zero integral ideal $\mathfrak{b}$ in $\mathcal{O}_K$ we write $\mathrm{free}_S(\mathfrak{b})$ for the largest ideal dividing $\mathfrak{b}$ and entirely supported outside of S.

Proposition 2.10. *Let* $(v_g(1), v_g(2))_{g \in G-\{\mathrm{id}\}}$ *be an element of*

$$\mathrm{Prim}(\mathcal{S}_l^{G-\{\mathrm{id}\}})(\mathrm{solv.}).$$

Then

$$\mathrm{free}_{S_{\mathrm{clean}}(l)}(\mathrm{Disc}(P_G((v_g)_{g \in G-\{\mathrm{id}\}}))) = \prod_{g \in G-\{\mathrm{id}\}} v_g(2)^{\#G \cdot (1 - \frac{1}{\#\langle g \rangle})}.$$

Proof. We show that the $\mathfrak{q}$-adic valuation matches for any prime $\mathfrak{q}$ of $\mathcal{O}_K$. This is certainly true for the places $\mathfrak{q}$ in $S_{\mathrm{clean}(l)}$, but also for the places $\mathfrak{q}$ outside $\widetilde{\Omega}_K(l)$

by Proposition 2.3. Now take a place $\mathfrak{q}$ in $\widetilde{\Omega}_K(l)$. Since $\mathfrak{q}$ is coprime to l we know that

$$v_{\mathfrak{q}}(\mathrm{Disc}(P_G((v_g)_{g\in G-\{\mathrm{id}\}}))) = \frac{\#G}{\#\langle P_G((v_g)_{g\in G-\{\mathrm{id}\}})(\sigma_{\mathfrak{q}})\rangle}\cdot\left(\#\langle P_G((v_g)_{g\in G-\{\mathrm{id}\}})(\sigma_{\mathfrak{q}})\rangle - 1\right).$$

Thanks to Proposition 2.9 we deduce that $P_G((v_g)_{g\in G-\{\mathrm{id}\}})(\sigma_{\mathfrak{q}})$ is trivial in case $\mathfrak{q}$ does not divide any v_g and equals g_0 in case $\mathfrak{q}$ divides v_{g_0}. This is precisely the desired conclusion. □

Our final goal for this subsection is to generalize Propositions 2.8, 2.9 and 2.10 to arbitrary finite, nilpotent groups. Recall that a finite group G is nilpotent if and only if it decomposes as a direct product of its Sylow subgroups. Let c be a positive integer. Let $l_1,\dots,l_c$ be distinct prime numbers. For each $j\in[c]$ fix an admissible sequence

$$\{(G_i(l_j),\theta_i(l_j))\}_{i\in[r_j]}$$

and write $G(l_j) := G_{r_j}(l_j)$ for the resulting l_j-group. We put

$$G := \prod_{j=1}^{c} G(l_j).$$

To parametrize G-extensions, we reduce to l-groups by means of Proposition 2.11.

Proposition 2.11. *We have an identification*

$$\mathrm{Epi}_{\mathrm{top.gr.}}(G_K,G) = \prod_{j\in[c]}\mathrm{Epi}_{\mathrm{top.gr.}}(\mathcal{G}_K^{\mathrm{pro}-l_j},G(l_j))$$

through the natural map.

Proof. Let H be any group and $\psi=(\psi_j)_{j\in[c]}: H\to G$ be any group homomorphism. We have to show that if ψ_j is surjective for each $j\in[c]$, then ψ is surjective. Observe that if each ψ_j is surjective, then $\#\,\mathrm{Im}(\psi)$ is divisible by $\#G(l_j)$ for each $j\in[c]$. Since these values are coprime, we find out that $\#\,\mathrm{Im}(\psi)$ is divisible by $\#G$, which means precisely that ψ is surjective. □

Thanks to Proposition 2.11 we can now bundle together the various maps $P_{G(l_j)}$ into one map

$$P_G : \prod_{j\in[c]}\mathrm{Prim}(\mathcal{S}_{l_j}^{G(l_j)-\{\mathrm{id}\}}) \twoheadrightarrow \mathrm{Epi}_{\mathrm{top.gr.}}(G_K,G)\cup\{\bullet\}$$

by simply taking the product map. Put

$$S := \bigcup_{j\in[c]} S_{\mathrm{clean}}(l_j).$$

For every $j\in[c]$, let $\{\chi_{j,i}\}_{i\in[t_j]}$ be a basis for the space of characters $G_K\to\mathbb{F}_{l_j}$ only ramified at S. Define $\mathcal{S}$ to be $\{0,1\}^{[t_1+\cdots+t_c]}\times\mathcal{S}'$, where $\mathcal{S}'$ is the set of squarefree ideals supported outside S. Let

$$\mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})$$

be the set of tuples $(v_g(1),v_g(2))_{g\in G-\{\mathrm{id}\}}$ satisfying the following properties

- writing π_i for the natural projection map $[t_1+\cdots+t_c]\to[t_i]$, we have that the $\pi_i(v_g(1))$ are pairwise coprime;

- the $v_g(2)$ are pairwise coprime;
- if $\mathfrak{p}$ divides $v_g(2)$ and l is a prime dividing the order of g, then

$$\#(\mathcal{O}_K/\mathfrak{p}) \equiv 1 \bmod l.$$

For an element

$$(v_{g,j}(1), v_{g,j}(2))_{j\in[c], g\in G(l_j)-\{\mathrm{id}\}} \in \prod_{j\in[c]} \mathrm{Prim}(\mathcal{S}_{l_j}^{G(l_j)-\{\mathrm{id}\}})$$

and for $g := (g_1, \dots, g_c) \in G - \{\mathrm{id}\}$ we define

$$v_g(1) := (v_{g_1,1}(1), \dots, v_{g_c,c}(1)), \quad v_g(2) := \prod_{\substack{\mathfrak{p} \\ \forall j\in[c]\forall h\in G(l_j)-\{\mathrm{id}\}: \mathfrak{p}\mid v_{h,j}(2)\Leftrightarrow g_j=h}} \mathfrak{p}.$$

In this way we have created a very convenient bijection between

$$\mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}}) \cong \prod_{j\in[c]} \mathrm{Prim}(\mathcal{S}_{l_j}^{G(l_j)-\{\mathrm{id}\}}).$$

We need one additional piece of notation, namely we define

$$\mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})(\mathrm{solv.}) := \prod_{j\in[c]} \mathrm{Prim}(\mathcal{S}_{l_j}^{G(l_j)-\{\mathrm{id}\}})(\mathrm{solv.}),$$

which we shall often implicitly view as a subset of $\mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})$. Proposition 2.12 generalizes Proposition 2.9.

Proposition 2.12. *Let $v := (v_{g,j}(1), v_{g,j}(2))_{j\in[c], g\in G(l_j)-\{\mathrm{id}\}}$ be an element of*

$$\mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})(\mathrm{solv.}).$$

Take some $\mathfrak{q} \in \Omega_K - S$. Let T be the subset of $[c]$ such that $j \in T$ if and only if there exists a (necessarily) unique $g_0^j \in G(l_j) - \{\mathrm{id}\}$ with $\mathfrak{q} \mid v_{g_0^j,j}(2)$. Then we have

$$P_G(v)(\sigma_{\mathfrak{q}}) = (g_0^j)_{j\in T} \times (\mathrm{id})_{k\in[c]-T}.$$

In particular if $\mathfrak{q}$ does not divide any of the elements $v_{g,j}(2)$, i.e. $T = \emptyset$, then

$$P_G(v)(\sigma_{\mathfrak{q}}) = \mathrm{id}.$$

Proof. This follows at once from Proposition 2.8, applied to each $G(l_j)$-factor. □

Proposition 2.13 generalizes Proposition 2.10. In the new coordinates we have a rather simple formula for the discriminant.

Proposition 2.13. *Notations as above. Let $(v_{g,j}(1), v_{g,j}(2))_{j\in[c], g\in G(l_j)-\{\mathrm{id}\}}$ be an element of $\mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})(\mathrm{solv.})$. Then*

$$\mathrm{free}_S(\mathrm{Disc}(P_G((v_{g,j}(1), v_{g,j}(2))_{j\in[c], g\in G(l_j)-\{\mathrm{id}\}}))) = \mathrm{free}_S\left(\prod_{g\in G-\{\mathrm{id}\}} v_g(2)^{\#G\cdot(1-\frac{1}{\#\langle g\rangle})}\right).$$

Proof. This follows from Proposition 2.12 with exactly the same argument as used to establish Proposition 2.10 as a consequence of Proposition 2.9. □

3. Local conditions and conjugacy classes

3.1. **Some group theory.** Let l be a prime number and let G be a finite l-group given by an admissible sequence $\{(G_i, \theta_i)\}_{i\in[r]}$ with $G := G_r$. Our first goal is to study the formation of a conjugacy class in G, through the various groups G_i, with $i \in [r]$. For $g \in G$ we denote by $\mathrm{Conj}_G(g)$ its conjugacy class. For each $0 \leq i \leq r$ we write

$$\pi_i : G \to G_i$$

for the natural projection map.

For now we take any finite l-group H, a 2-cocycle θ representing a class in $H^2(H, \mathbb{F}_l)$, with $\theta(\mathrm{id}, \mathrm{id}) = 0$, and an element $h \in H$. Denote the centralizer of h by $\mathrm{Cent}_H(h)$. Then lifting elements of $\mathrm{Cent}_H(h)$ to $(\mathbb{F}_l \times H, *_\theta)$ and taking the commutator with any lift h_0 of h induces a homomorphism

$$[-, h_0]_\theta : \frac{\mathrm{Cent}_H(h)}{\langle h\rangle} \to \mathbb{F}_l,$$

which does not depend on the choice of lifts. Put

$$\widetilde{\mathrm{Cent}}_H(h, \theta) := \ker([-, h_0]_\theta),$$

which is by definition a subgroup of $\frac{\mathrm{Cent}_H(h)}{\langle h\rangle}$. Let

$$\pi_\theta : (\mathbb{F}_l \times H, *_\theta) \to H$$

be the natural projection map. Proposition 3.1 describes the relationship between $\mathrm{Cent}_H(h)$ and $\widetilde{\mathrm{Cent}}_H(h, \theta)$.

Proposition 3.1. *Let H, h, θ be as above this proposition. Then*

$$\left[\frac{\mathrm{Cent}_H(h)}{\langle h\rangle} : \widetilde{\mathrm{Cent}}_H(h, \theta)\right] \in \{1, l\}.$$

The index equals 1 if and only if the elements in $\pi_\theta^{-1}(h)$ are pairwise non-conjugate in $(\mathbb{F}_l \times H, *_\theta)$. The index equals l if and only if the elements of $\pi_\theta^{-1}(h)$ sit inside a unique conjugacy class in $(\mathbb{F}_l \times H, *_\theta)$.

Proof. Indeed, take a lift $h_0 := (a, h)$ in $\pi_\theta^{-1}(h)$. Observe that if we have

$$[(b, h'), h_0] *_\theta h_0 = (b, h') *_\theta h_0 *_\theta (b, h')^{-1} = (a', h)$$

for some $a' \in \mathbb{F}_l$, then it follows that $h' \in \mathrm{Cent}_H(h)$. From the left hand side we see that if the index is l, then a' can take any possible value. If the index is instead equal to 1, then a' must be equal to a. □

We also have a similar proposition for the exponent of an element.

Proposition 3.2. *Let H, h, θ be as above. Then either $\pi_\theta^{-1}(h)$ consists entirely of elements with order equal to $l \cdot \#\langle h\rangle$ or it consists entirely of elements with order equal to $\#\langle h\rangle$.*

Proof. The class θ restricted to $\langle h\rangle$ gives an element of $H^2(\langle h\rangle, \mathbb{F}_l) = \mathrm{Ext}(\langle h\rangle, \mathbb{F}_l)$. If the class is 0, then the sequence is split and we have that all the elements of $\pi_\theta^{-1}(h)$ have the same order as h. If the class is non-zero, then the sequence has the shape

$$0 \to \mathbb{F}_l \to \mathbb{Z}/l \cdot \#\langle h\rangle\mathbb{Z} \to \langle h\rangle \to 0,$$

and hence all elements of $\pi_\theta^{-1}(h)$ have order l times bigger than that of h. □

In case an h as in Proposition 3.2 satisfies the second conclusion we say that h is θ-stable. We now return to our previous setup. To $g \in G$ we attach the following quantity

$$j_G(g, \{(G_i, \theta_i)\}_{i\in[r]}) := \# \left\{ i \in [r] : \frac{\mathrm{Cent}_{G_{i-1}}(\pi_{i-1}(g))}{\langle \pi_{i-1}(g)\rangle} \neq \widetilde{\mathrm{Cent}}_{G_{i-1}}(\pi_{i-1}(g), \theta_i) \right\}.$$

This quantity turns out to be the exponent of l in the size of the conjugacy class $\mathrm{Conj}_G(g)$. We call the i's counted by $j_G(g, \{(G_i, \theta_i)\}_{i\in[r]})$ the *breaks* for g with respect to $\{(G_i, \theta_i)\}_{i\in[r]}$.

Proposition 3.3. *We have*

$$\#\,\mathrm{Conj}_G(g) = l^{j_G(g, \{(G_i,\theta_i)\}_{i\in[r]})}.$$

Proof. We have a filtration of subgroups

$$G = \mathrm{Cent}^0_G(g) \supseteq \mathrm{Cent}^1_G(g) \supseteq \cdots \supseteq \mathrm{Cent}^r_G(g) = \mathrm{Cent}_G(g),$$

where we define

$$\mathrm{Cent}^i_G(g) := \pi_i^{-1}(\mathrm{Cent}_{G_i}(\pi_i(g)))$$

for every integer $0 \leq i \leq r$. We have that

$$\#\,\mathrm{Conj}_G(g) = \frac{\#G}{\#\,\mathrm{Cent}_G(g)} = \prod_{0\leq i\leq r-1} \frac{\#\,\mathrm{Cent}^i_G(g)}{\#\,\mathrm{Cent}^{i+1}_G(g)}.$$

Observe that

$$[\mathrm{Cent}^i_G(g) : \mathrm{Cent}^{i+1}_G(g)] = \left[\frac{\mathrm{Cent}_{G_i}(\pi_i(g))}{\langle \pi_i(g)\rangle} : \widetilde{\mathrm{Cent}}_{G_i}(\pi_i(g), \theta_{i+1}) \right].$$

Therefore the desired conclusion follows at once from Proposition 3.1. □

Proposition 3.4 provides the crucial link between group theoretic data and the local conditions imposed, through Proposition 2.6, on tuples in $\mathrm{Prim}(\mathcal{S}_l^{G-\{\mathrm{id}\}})$. We invoke the notation of Section 2. Let L/K be a finite Galois extension inside $K^{\mathrm{pro}-l}/K$. Let now $G := \mathrm{Gal}(L/K)$ and let θ be a 2-cocycle representing a class in $H^2(\mathrm{Gal}(L/K), \mathbb{F}_l)$, with $\theta(\mathrm{id}, \mathrm{id}) = 0$. Let $\mathfrak{q} \in \Omega_K$ be a finite place coprime to l. Recall that

$$\frac{G_{K_\mathfrak{q}}}{I_\mathfrak{q}}$$

is a pro-cyclic group, equipped with a canonical generator $\mathrm{Frob}_\mathfrak{q}$. We will fix once and for all the $\mathrm{proj}(G_K \to \mathcal{G}_K^{\mathrm{pro}-l}) \circ i_\mathfrak{q}^*$-image of a lift to $G_{K_\mathfrak{q}}$ of such an element. In this way we obtain an element in $\mathcal{G}_K^{\mathrm{pro}-l}$ that we will denote also by $\mathrm{Frob}_\mathfrak{q}$: this slight abuse of notation will cause no confusion. As such we have naturally an element

$$\mathrm{proj}(\mathcal{G}_K^{\mathrm{pro}-l} \to G)(\mathrm{Frob}_\mathfrak{q}) \in \frac{N_G(\langle \mathrm{proj}(\mathcal{G}_K^{\mathrm{pro}-l} \to G)(\sigma_\mathfrak{q})\rangle)}{\langle \mathrm{proj}(\mathcal{G}_K^{\mathrm{pro}-l} \to G)(\sigma_\mathfrak{q})\rangle}.$$

Here $N_G(-)$ denotes the normalizer of a subgroup in G. For any non-trivial finite group G, we denote by l_G the smallest prime divisor of $\#G$ and by $I(G)$ the subset of $g \in G - \{\mathrm{id}\}$ such that $g^{l_G} = \mathrm{id}$.

Proposition 3.4. *Let $G = \mathrm{Gal}(L/K)$ be a finite l-group and let $\mathfrak{q} \in \Omega_K$ be a finite place coprime to l. Assume that $\mathrm{proj}(\mathcal{G}_K^{\mathrm{pro}-l} \to G)(\sigma_\mathfrak{q})$ is an element of $I(G)$, which we shall also call $\sigma_\mathfrak{q}$. Then*

$$\mathrm{proj}(\mathcal{G}_K^{\mathrm{pro}-l} \to G)(\mathrm{Frob}_\mathfrak{q}) \in \frac{\mathrm{Cent}_G(\sigma_\mathfrak{q})}{\langle \sigma_\mathfrak{q} \rangle}.$$

Moreover, if $\mathrm{proj}(\mathcal{G}_K^{\mathrm{pro}-l} \to G)(\sigma_\mathfrak{q})$ is θ-stable, then

$$\mathrm{proj}(\mathcal{G}_K^{\mathrm{pro}-l} \to G)(\mathrm{Frob}_\mathfrak{q}) \in \frac{\widetilde{\mathrm{Cent}}_G(\sigma_\mathfrak{q}, \theta)}{\langle \sigma_\mathfrak{q} \rangle}$$

if and only if θ is trivial in $H^2(G_{K_\mathfrak{q}}, \mathbb{F}_l)$.

Proof. Let G be any finite non-trivial group. Then we claim that $g \in I(G)$ implies $N_G(\langle g \rangle) = \mathrm{Cent}_G(g)$. Indeed, conjugation induces a homomorphism

$$N_G(\langle g \rangle) \to \mathrm{Aut}_{\mathrm{gr.}}(\langle g \rangle) \simeq_{\mathrm{gr.}} \mathbb{F}_{l_G}^*.$$

Since the latter group has size $l_G - 1$, and l_G is the smallest prime divisor of $\#G$, we have that $\#G$ is coprime to $l_G - 1$. Therefore the above homomorphism is actually trivial. This means exactly that $N_G(\langle g \rangle) = \mathrm{Cent}_G(g)$ as claimed. Hence we have already shown the first part of this proposition, namely that

$$\mathrm{proj}(\mathcal{G}_K^{\mathrm{pro}-l} \to G)(\mathrm{Frob}_\mathfrak{q}) \in \frac{\mathrm{Cent}_G(\sigma_\mathfrak{q})}{\langle \sigma_\mathfrak{q} \rangle}.$$

Assume now that $\mathrm{proj}(\mathcal{G}_K^{\mathrm{pro}-l} \to G)(\sigma_\mathfrak{q})$ is also θ-stable. Observe that since $\sigma_\mathfrak{q}$ lands in $I(G)$, we in particular conclude that $\mathfrak{q}$ ramifies in L/K. It follows from Proposition 2.3 that

$$q := \#\left(\mathcal{O}_K/\mathfrak{q}\right) \equiv 1 \bmod l.$$

Also observe that the maximal pro-l quotient of $G_{K_\mathfrak{q}}$ is isomorphic to

$$\mathbb{Z}_l \rtimes q^{\mathbb{Z}_l},$$

where q acts by multiplication by q on $\mathbb{Z}_l$. Here $I_{K_\mathfrak{q}}$ is sent to $\mathbb{Z}_l \rtimes \{1\}$, while a lift of $\mathrm{Frob}_\mathfrak{q}$ is sent to $\{0\} \rtimes \{q\}$. Since $\mathrm{proj}(\mathcal{G}_K^{\mathrm{pro}-l} \to G)(\sigma_\mathfrak{q})$ is θ-stable and lands in $I(G)$, we have that the lifting problem imposed by θ factors through

$$l \cdot \mathbb{Z}_l \rtimes \{1\}.$$

Since q is 1 modulo l, the resulting quotient is simply

$$\mathbb{F}_l \times q^{\mathbb{Z}_l}.$$

Recalling once more that $\mathrm{proj}(\mathcal{G}_K^{\mathrm{pro}-l} \to G)(\sigma_\mathfrak{q})$ is θ-stable, we see that the lifting problem is solvable if and only if the restriction of θ to

$$\mathrm{proj}(G_K \to G) \circ i_{\mathfrak{q}^*}(G_{K_\mathfrak{q}})$$

is in

$$\mathrm{Ext}(\mathrm{proj}(G_K \to G) \circ i_{\mathfrak{q}^*}(G_{K_\mathfrak{q}}), \mathbb{F}_l).$$

This is precisely equivalent to asking

$$\mathrm{Frob}_\mathfrak{q} \in \frac{\widetilde{\mathrm{Cent}}_G(\sigma_\mathfrak{q}, \theta)}{\langle \sigma_\mathfrak{q} \rangle}$$

as was to be shown. □

The final lemma of this subsection provides a simple way to compute the Malle constant $b_{\mathrm{M}}(G,K)$ for nilpotent G.

Lemma 3.5. *Let G be a non-trivial, finite group and let K be a number field. Then we have*

$$b_{\mathrm{M}}(G,K)=\frac{\#\{C\in\mathrm{Conj}(G):C\subseteq I(G)\}}{[K(\zeta_{l_G}):K]}.$$

Proof. Recall that there is a natural action of $\mathrm{Gal}(\overline{K}/K)$ on

$$X:=\{C\in\mathrm{Conj}(G):C\subseteq I(G)\},$$

which sends a conjugacy class C to $C^{\chi(\sigma)}$ with $\chi:\mathrm{Gal}(\overline{K}/K)\to\hat{\mathbb{Z}}^*$ the cyclotomic character. By definition $b_{\mathrm{M}}(G,K)$ equals the number of orbits of this group action. But our action clearly factors through $\mathrm{Gal}(K(\zeta_{l_G})/K)$. We claim that the induced action of $\mathrm{Gal}(K(\zeta_{l_G})/K)$ on X is free, which implies the lemma.

So suppose that there exists $\sigma\in\mathrm{Gal}(K(\zeta_{l_G})/K)$ such that

$$C^{\chi(\sigma)}=C.$$

This implies that there exists a non-trivial $g\in C$ such that g and $g^{\chi(\sigma)}$ are conjugate, say

$$g=h^{-1}g^{\chi(\sigma)}h,$$

and hence $h\in N_G(\langle g\rangle)=\mathrm{Cent}_G(g)$ by the argument given at the start of Proposition 3.4. We conclude that $g=g^{\chi(\sigma)}$, which forces σ to be the identity as desired. □

3.2. **Interpretation of Malle's constant.** The goal of this subsection is to give a heuristic supporting Malle's conjecture in the nilpotent case. Our heuristic is based on a combination of the parametrization given in Proposition 2.8 and the examination of the local conditions carried out in Subsection 3.1. To simplify the notation, we shall limit ourselves to the case where G is an l-group. We leave it to the reader to generalize the material below to arbitrary nilpotent groups G.

Ignoring the finitely many bad places in $S_{\mathrm{clean}}(l)$, it follows from Proposition 2.8 and Proposition 2.10 that

$$\#\{\psi\in\mathrm{Epi}_{\mathrm{top.gr.}}(\mathcal{G}_K^{\mathrm{pro}-l},G):|N_{K/\mathbb{Q}}\,\mathrm{Disc}(\psi)|\leq X\}$$

should have order of magnitude

$$\#\{(v_g(1),v_g(2))_{g\in G-\{\mathrm{id}\}}\in\mathrm{Prim}(\mathcal{S}_l^{G-\{\mathrm{id}\}})(\mathrm{solv.}):\prod_{g\in G-\{\mathrm{id}\}}(|N_{K/\mathbb{Q}}v_g(2)|)^{\#G\cdot(1-\frac{1}{\#\langle g\rangle})}\leq X\}.$$

We now focus on the variables $(v_g(1),v_g(2))$ with $g\in I(G)$. Upon combining Proposition 3.4 and Proposition 2.9, we see that the primes $\mathfrak{q}$ dividing $v_g(2)$ impose a local condition only at the breaks for g in the admissible sequence $\{(G_i,\theta_i)\}_{i\in[r]}$. The local conditions at the points that are not breaks are automatically satisfied in virtue of Proposition 3.4. Now pretend that the values

$$\mathrm{Frob}_{\mathfrak{q}}\in\frac{\mathrm{Cent}_{G_{i-1}}(\sigma_{\mathfrak{q}})}{\langle\sigma_{\mathfrak{q}}\rangle}$$

are jointly equidistributed at every break point i. Then, in virtue of Proposition 3.4, we get the following sum

$$\approx \sum_{\prod_{g\in G-\{\mathrm{id}\}}(|N_{K/\mathbb{Q}}v_g(2)|)^{\#G\cdot(1-\frac{1}{\#\langle g\rangle})}\leq X} \left(\prod_{g\in I(G)} \frac{1}{l^{\omega(v_g(2))j_G(g,\{(G_i,\theta_i)\}_{i\in[r]})}}\right),$$

where the sum runs over all points $(v_g(1), v_g(2))$ in $\mathrm{Prim}(\mathcal{S}_l^{G-\{\mathrm{id}\}})$. Thanks to Proposition 3.3 the latter expression equals

$$\sum_{\prod_{g\in G-\{\mathrm{id}\}}(|N_{K/\mathbb{Q}}v_g(2)|)^{\#G\cdot(1-\frac{1}{\#\langle g\rangle})}\leq X} \left(\prod_{g\in I(G)} \frac{1}{\#\,\mathrm{Conj}_G(g)^{\omega(v_g(2))}}\right), \tag{3.1}$$

where the sum still ranges over all points $(v_g(1), v_g(2))$ in $\mathrm{Prim}(\mathcal{S}_l^{G-\{\mathrm{id}\}})$. Standard analytic techniques, see Theorem 4.1, show that the sum in equation (3.1) is asymptotic to

$$c(G,K)\cdot X^{a(G)}\cdot \log(X)^{\beta(G,K)-1}, \quad \beta(G,K) := \sum_{g\in I(G)} \frac{1}{\#\,\mathrm{Conj}_G(g)\cdot[K(\zeta_l):K]},$$

where $\beta(G,K)$ is the Malle constant by Lemma 3.5. We remark that to turn this simple heuristic into an argument one also has to pay careful attention to the local conditions at the primes dividing variables outside of $I(G)$ and to the local conditions at the primes in $S_{\mathrm{clean}}(l)$. These will affect the constant $c(G,K)$ in the asymptotic. We finish this section by explaining the interplay between this heuristic and the proofs of our main theorems.

During the proof of Theorem 5.1 we simply ignore the local conditions, at the cost of losing track of the conjugation in G: for this reason we get $i(G,K)-1$ instead of $b_{\mathrm{M}}(G,K)-1$.

Correspondingly for those G for which $i(G,K)-1$ and $b_{\mathrm{M}}(G,K)-1$ coincide we have that all the elements of $I(G)$ are central, and consistently with Proposition 3.4 we have no local conditions coming from the variables in $I(G)$. In this case we are able to prove an asymptotic in Theorem 5.7.

Finally in the proof of Theorem 5.3, thanks to the fact that the elements of $I(G)$ are pairwise commuting, we have to control the behavior of $\mathrm{Frob}_{\mathfrak{q}}$ in the quotient $\frac{G}{I(G)\cup\{\mathrm{id}\}}$ for each $\mathfrak{q}$ dividing a variable in $I(G)$. This is very convenient, since the corresponding field is constructed out of the variables outside of $I(G)$, and those have a very large weight in the formula for the discriminant given in Proposition 2.10. As such, they can almost be treated as fixed, and the required joint equidistribution of Frobenius elements is provable by appealing to the Chebotarev density theorem. Hence in this case we can partially turn the above heuristic into a rigorous argument: since we control only the local conditions at the places dividing variables in $I(G)$, we naturally end up with an upper bound of the correct order of magnitude.

4. Analytic considerations

In this section we provide the analytic tools used to prove our main theorems. The material in this section is a generalization of the material in Montgomery–Vaughan [41, Section 7.4] and is an application of the Selberg–Delange method.

Let K be a number field, let L be an abelian extension of K and let $S \subseteq \mathrm{Gal}(L/K)$. Write $\mathcal{I}_K$ for the group of non-zero fractional ideals of K. For a squarefree ideal I of $\mathcal{O}_K$ we define $\omega(I)$ for the number of prime divisors $\mathfrak{p}$ of I and $\omega_S(I)$ for the number of prime divisors $\mathfrak{p}$ of I that are unramified in L and satisfy $\mathrm{Frob}_{\mathfrak{p}} \in S$. Given a complex number z and a collection of prime ideals $\mathcal{P}$, we are interested in the sum

$$A_z(x) := \sum_{\substack{N_{K/\mathbb{Q}}(I) \leq x \\ \mathfrak{p} \mid I \Rightarrow \mathrm{Frob}_{\mathfrak{p}} \in S \text{ and } \mathfrak{p} \notin \mathcal{P}}} \mu^2(I) z^{\omega(I)} = \sum_{\substack{N_{K/\mathbb{Q}}(I) \leq x \\ \mathfrak{p} \mid I \Rightarrow \mathrm{Frob}_{\mathfrak{p}} \in S \text{ and } \mathfrak{p} \notin \mathcal{P}}} \mu^2(I) z^{\omega_S(I)},$$

where μ is the Möbius function of $\mathcal{O}_K$. Write

$$F(s,z) = \sum_{\substack{I \in \mathcal{I}_K \\ \mathfrak{p} \mid I \Rightarrow \mathrm{Frob}_{\mathfrak{p}} \in S \text{ and } \mathfrak{p} \notin \mathcal{P}}} \frac{\mu^2(I) z^{\omega(I)}}{N_{K/\mathbb{Q}}(I)^s} = \sum_{n=1}^{\infty} \frac{a_z(n)}{n^s} \tag{4.1}$$

for its Dirichlet series with coefficients $a_z(n)$. Then we have for $s = \sigma + it$

$$\begin{aligned} F(s,z) &= \prod_{\substack{\mathfrak{p} \notin \mathcal{P} \\ \mathrm{Frob}_{\mathfrak{p}} \in S}} \left(1 + \frac{z}{N_{K/\mathbb{Q}}(\mathfrak{p})^s}\right) \\ &= \prod_{\mathfrak{p} \notin \mathcal{P}} \left(1 + \frac{z \sum_{\sigma \in S} \frac{1}{\#\mathrm{Gal}(L/K)} \sum_{\chi \in \mathrm{Gal}(L/K)^\vee} \chi(\mathrm{Frob}_{\mathfrak{p}}) \overline{\chi(\sigma)}}{N_{K/\mathbb{Q}}(\mathfrak{p})^s}\right) \text{ for } \sigma > 1, \end{aligned}$$

where $\mathrm{Gal}(L/K)^\vee$ is by definition $\mathrm{Hom}(\mathrm{Gal}(L/K), \mathbb{C}^*)$. We assume that the Euler product

$$\prod_{\mathfrak{p} \in \mathcal{P}} \left(1 + \frac{1}{N_{K/\mathbb{Q}}(\mathfrak{p})^s}\right) \tag{4.2}$$

converges absolutely in the region $\sigma > 1 - \delta$ for some constant $\delta > 0$. Then we approximate the Dirichlet series $F(s,z)$ with

$$G(s,z) := \prod_{\sigma \in S} \prod_{\chi \in \mathrm{Gal}(L/K)^\vee} L(s,\chi)^{\frac{z\overline{\chi(\sigma)}}{\#\mathrm{Gal}(L/K)}},$$

where

$$L(s,\chi) = \prod_{\mathfrak{p}} \left(1 - \frac{\chi(\mathrm{Frob}_{\mathfrak{p}})}{N_{K/\mathbb{Q}}(\mathfrak{p})^s}\right)^{-1} \text{ for } \sigma > 1.$$

We recall that $L(s,\chi)^z$ is by definition $e^{z \log L(s,\chi)}$. Note that $\log L(s,\chi)$ exists since the region $\sigma > 1$ is simply connected and $L(s,\chi)$ does not vanish in this region. We choose our determination of the logarithm in such a way that it agrees with the real logarithm for real s. It follows from equation (4.2) that we have the fundamental relation

$$F(s,z) = G(s,z)H(s,z),$$

where $H(s,z)$ is defined by an absolutely convergent Euler product in the region $\sigma > 1 - \delta$ for some $\delta > 0$. In particular, if $|z| \leq R$, then there exists some constant $\delta(R) > 0$ such that $H(s,z)$ is a bounded non-zero holomorphic function on $\sigma > 1 - \delta(R)$.

Theorem 4.1. *Let K, L, S and $\mathcal{P}$ as above. Then we have for all positive real numbers R and all $|z| \leq R$*

$$A_z(x) = Cx(\log x)^{\frac{z\#S}{\#\operatorname{Gal}(L/K)}-1} + O_{R,K,L,S,\mathcal{P}}\left(x(\log x)^{\frac{\operatorname{Re}(z)\#S}{\#\operatorname{Gal}(L/K)}-2}\right),$$

where $C > 0$ is a real constant depending only on z, K, L, S and $\mathcal{P}$.

Proof. Since the proof is similar to Montgomery–Vaughan [41, Theorem 7.17], we shall only sketch the necessary modifications. Set $a = 1 + 1/\log x$. An effective version of Perron's formula, see [41, Corollary 5.3], shows that

$$\begin{aligned} A_z(x) - \frac{1}{2\pi i}\int_{a-iT}^{a+iT} F(s,z)\frac{x^s}{s}ds \\ \ll \sum_{\frac{1}{2}x<n<2x} |a_z(n)| \min\left(1, \frac{x}{T|x-n|}\right) + \frac{x^a}{T}\sum_{n=1}^{\infty} |a_z(n)| n^{-a}, \end{aligned}$$

where we recall that $a_z(n)$ is defined by equation (4.1) and T is a parameter at our disposal. We choose $T = \exp(\sqrt{\log x})$ and estimate the error terms as in [41]. To do so, we need to have a good estimate for the sum

$$\sum_{|n-x|\leq \frac{x}{(\log x)(2R)^{[K:\mathbb{Q}]}+R+1}} |a_z(n)| \leq \sum_{|n-x|\leq \frac{x}{(\log x)(2R)^{[K:\mathbb{Q}]}+R+1}} (2R)^{[K:\mathbb{Q}]\omega(n)},$$

where we assume without loss of generality that R is an integer greater than 1. The latter sum is estimated in [41, Theorem 7.17] with Dirichlet's hyperbola method.

We next move the path of integration. Note that $F(s,z)$ has a branch point at $s = 1$ if z is not an integer. For this reason, we move the path of integration in such a way to avoid this branch point. Put $b = 1 - c/\log T$, where c is a small positive constant. Let $\mathcal{C}_1$ be the polygonal path with vertices $a - iT, b - iT, b - i/\log x$, let $\mathcal{C}_2$ be the line segment from $b - i/\log x$ to $1 - i/\log x$, followed by a semicircle $\{1+e^{i\theta}/\log x : -\pi/2 \leq \theta \leq \pi/2\}$, and a line segment from $1+i/\log x$ to $b+i/\log x$, and finally let $\mathcal{C}_3$ be the polygonal path with vertices $b + i/\log x, b + iT, a + iT$.

Let $\mathcal{D}$ be the region enclosed by $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$ and the line segment from $a - iT$ to $a + iT$. If c is sufficiently small, then $L(s,\chi)$ has no zeroes in the region $\mathcal{D}$ by [26, Theorem 5.10]. Clearly, $H(s,z)$ also has no zeroes in $\mathcal{D}$ provided that c is sufficiently small. Since the union of $\mathcal{D}$ with the region $\operatorname{Re}(s) > 1$ is still simply connected, $\log L(s,\chi)$ and $\log H(s,z)$ are also well-defined in this region.

The main term comes from the integral over $\mathcal{C}_2$, and is extracted in exactly the same way as in the proof of [41, Theorem 7.17]. Finally, Montgomery–Vaughan estimate the integrals on the paths $\mathcal{C}_1$ and $\mathcal{C}_3$ by appealing to bounds for $\zeta(s)$, see their [41, Theorem 6.7]. Hence we need to supply similar bounds for $\zeta_K(s)$ and $L(s,\chi)$. These can be derived by following the proof of [41, Theorem 6.7], where we use [26, Proposition 5.7, (2)] as a replacement for [41, Lemma 6.4]. □

Write $*$ for the Dirichlet convolution on $\mathcal{I}_K$. In Section 5 we will combine Theorem 4.1 with the following general lemma on convolutions.

Lemma 4.2. *Let $f, g : \mathcal{I}_K \to \mathbb{R}$ be functions such that*

$$\sum_{N_{K/\mathbb{Q}}(I)\leq x} f(I) = C_1 x(\log x)^A + O(x(\log x)^{A-\delta}),$$
$$\sum_{N_{K/\mathbb{Q}}(I)\leq x} g(I) = C_2 x(\log x)^B + O(x(\log x)^{B-\delta})$$

for some real numbers $A, B > -1$, $C_1, C_2 > 0$ and $0 < \delta < 1$. Then there is $C_3 > 0$ such that

$$\sum_{N_{K/\mathbb{Q}}(I)\leq x} (f * g)(I) = C_3 x(\log x)^{A+B+1} + O(x(\log x)^{A+B+1-\delta}).$$

Proof. It follows from Dirichlet's hyperbola method that

$$\sum_{N_{K/\mathbb{Q}}(I)\leq x} (f * g)(I) = \sum_{N_{K/\mathbb{Q}}(IJ)\leq x} f(I)g(J)$$

equals

(4.3)
$$\sum_{N_{K/\mathbb{Q}}(I)\leq \sqrt{x}} \sum_{N_{K/\mathbb{Q}}(J)\leq \frac{x}{N_{K/\mathbb{Q}}(I)}} f(I)g(J) + \sum_{N_{K/\mathbb{Q}}(J)\leq \sqrt{x}} \sum_{N_{K/\mathbb{Q}}(I)\leq \frac{x}{N_{K/\mathbb{Q}}(J)}} f(I)g(J) - \sum_{N_{K/\mathbb{Q}}(I)\leq \sqrt{x}} \sum_{N_{K/\mathbb{Q}}(J)\leq \sqrt{x}} f(I)g(J).$$

The latter sum is at most $O(x(\log x)^{A+B})$. Since the first two sums in equation (4.3) play a symmetric role, we shall only treat the first sum. The first sum equals

(4.4)
$$\sum_{N_{K/\mathbb{Q}}(I)\leq \sqrt{x}} \frac{C_2 x f(I)}{N_{K/\mathbb{Q}}(I)} \left(\log \frac{x}{N_{K/\mathbb{Q}}(I)}\right)^B = C_2 x \sum_{N_{K/\mathbb{Q}}(I)\leq \sqrt{x}} \frac{f(I)}{N_{K/\mathbb{Q}}(I)} \big(\log x - \log N_{K/\mathbb{Q}}(I)\big)^B$$

up to an error of size bounded by

$$O\left(x \sum_{N_{K/\mathbb{Q}}(I)\leq \sqrt{x}} \frac{f(I)}{N_{K/\mathbb{Q}}(I)} \big(\log x - \log N_{K/\mathbb{Q}}(I)\big)^{B-\delta}\right). \tag{4.5}$$

We shall give an asymptotic formula for equation (4.4), from which it will also be clear how to treat the error term in equation (4.5). Put

$$F(t) := \sum_{N_{K/\mathbb{Q}}(I)\leq t} f(I),$$

so that we have the formula

$$F(t) = C_1 t(\log t)^A + O(t(\log t)^{A-\delta}) \tag{4.6}$$

by assumption. Partial summation shows that

$$\sum_{N_{K/\mathbb{Q}}(I)\leq \sqrt{x}} \frac{f(I)}{N_{K/\mathbb{Q}}(I)} \big(\log x - \log N_{K/\mathbb{Q}}(I)\big)^B = \frac{F(\sqrt{x}) \cdot (\frac{1}{2}\log x)^B}{\sqrt{x}} - \int_1^{\sqrt{x}} F(t) d\left(\frac{(\log x - \log t)^B}{t}\right).$$

The first term is $O((\log x)^{A+B})$. Plugging in equation (4.6) shows that the second term above equals

$$C_1 \int_1^{\sqrt{x}} \frac{(\log t)^A (\log x - \log t)^B}{t} dt + O\left((\log x)^{A+B+1-\delta}\right).$$

Recall the Taylor expansion, valid for $-\log x < \log t < \log x$,

$$(\log x - \log t)^B = (\log x)^B \left(1 - \frac{\log t}{\log x}\right)^B = (\log x)^B \sum_{k=0}^{\infty} \binom{B}{k} \left(\frac{-\log t}{\log x}\right)^k,$$

where $\binom{B}{k}$ is the generalized binomial coefficient. Since the Taylor expansion converges uniformly for $1 \le t \le \sqrt{x}$, we may switch the infinite sum and the integral to obtain

$$(\log x)^B \sum_{k=0}^{\infty} \frac{(-1)^k \binom{B}{k}}{(\log x)^k} \int_1^{\sqrt{x}} \frac{(\log t)^{A+k}}{t} dt = (\log x)^{A+B+1} 2^{-A-1} \sum_{k=0}^{\infty} \frac{(-1)^k \binom{B}{k}}{2^k (A+k+1)},$$

where we used that $A > -1$ to compute the integral. We conclude that equation (4.4) equals

$$C_1 C_2 2^{-A-1} \sum_{k=0}^{\infty} \frac{(-1)^k \binom{B}{k}}{2^k (A+k+1)} x (\log x)^{A+B+1} + O\left(x (\log x)^{A+B+1-\delta}\right).$$

Set

$$C_3 := C_1 C_2 \left(2^{-A-1} \sum_{k=0}^{\infty} \frac{(-1)^k \binom{B}{k}}{2^k (A+k+1)} + 2^{-B-1} \sum_{k=0}^{\infty} \frac{(-1)^k \binom{A}{k}}{2^k (B+k+1)}\right).$$

It remains to show that $C_3 > 0$. But we have the lower bound

$$\int_1^{\sqrt{x}} \frac{(\log t)^A (\log x - \log t)^B}{t} dt \gg \int_1^{\sqrt{x}} \frac{(\log t)^{A+B}}{t} dt \gg (\log x)^{A+B+1},$$

and this completes the proof. □

Finally, we will need the following version of the Siegel–Walfisz theorem.

Theorem 4.3. *Let $A > 0$ be a given real number and let K be a fixed number field. Then we have for all $X > 2$, all Galois extensions L/K with $[L : K] < A$ and $N_{K/\mathbb{Q}}(\Delta(L/K)) \le (\log X)^A$ and all conjugacy classes C of $\mathrm{Gal}(L/K)$*

$$\frac{\#\{\mathfrak{p} \in \Omega_K : N_{K/\mathbb{Q}}(\mathfrak{p}) \le X, \mathfrak{p} \text{ unr. in } L, \mathrm{Frob}_{\mathfrak{p}} = C\}}{\#\{\mathfrak{p} \in \Omega_K : N_{K/\mathbb{Q}}(\mathfrak{p}) \le X\}} = \frac{\#C}{\#\mathrm{Gal}(L/K)} \mathrm{Li}(X) + O\left(\frac{X}{(\log X)^A}\right),$$

where the implied constant depends only on A and K.

Proof. This follows immediately from [50, Theorem 1.1], were it not for potential Siegel zeroes. To control a potential Siegel zero of $\zeta_L(s)$, we apply the ineffective Brauer–Siegel theorem, which yields for every $\epsilon > 0$

$$\mathrm{res}_{s=1} \zeta_L(s) \gg_{\epsilon,K} \frac{1}{\Delta(L/\mathbb{Q})^{\epsilon}}.$$

Picking ϵ sufficiently small in terms of A gives the desired lower bound for $1 - \beta$ by [35, Theorem 1]. □

5. Proof of main theorems

In this section we prove our three main results in complete generality. Recall that l_G is the smallest prime divisor of $\#G$ and that $I(G)$ is the subset of $g \in G$ with order equal to l_G. Set $H(G) := I(G) \cup \{\mathrm{id}\}$. Recall that

$$i(G,K) := \frac{\#I(G)}{[K(\zeta_{l_G}) : K]}.$$

5.1. Proof of Theorem 1.3 and Theorem 1.4.

In this subsection we prove Theorem 1.3 and Theorem 1.4.

Theorem 5.1. *Let G be a finite non-trivial nilpotent group and let K be a number field. Then there exists a constant $c \in \mathbb{R}_{>0}$ such that*

$$\#\{\psi \in \mathrm{Epi}_{\mathrm{top.gr.}}(G_K, G) : |N_{K/\mathbb{Q}}(\mathrm{Disc}(\psi))| \leq X\} \leq c \cdot X^{a(G)} \cdot \log(X)^{i(G,K)-1}$$

for all $X \in \mathbb{R}_{>2}$.

Proof. Write

$$e_g := \#G\left(1 - \frac{1}{\#\langle g \rangle}\right).$$

It follows from Proposition 2.13 and Proposition 2.11 that it suffices to bound

$$\#\left\{(v_{g,j}(1), v_{g,j}(2))_{j\in[c], g\in G(l_j)-\{\mathrm{id}\}} : \left|N_{K/\mathbb{Q}}\left(\prod_{g\in G-\{\mathrm{id}\}} v_g(2)^{e_g}\right)\right| \leq X\right\},$$

where

$$v_g(2) := \prod_{\substack{\mathfrak{p} \\ \forall j\in[c]\forall h\in G(l_j)-\{\mathrm{id}\}:\mathfrak{p}\mid v_{h,j}(2)\Leftrightarrow g_j=h}} \mathfrak{p}$$

with $g = (g_1, \ldots, g_c)$. Here the variables $(v_{g,j}(1), v_{g,j}(2))_{g\in G(l_j)-\{\mathrm{id}\}}$ are in

$$\mathrm{Prim}(\mathcal{S}_{l_j}^{G(l_j)-\{\mathrm{id}\}})$$

for every $j \in [c]$. We now drop the following conditions

- we drop the condition that $v_{g,j}(2)$ is supported outside $S_{\mathrm{clean}}(l_j)$. We also drop the condition that $v_{g,j}(2)$ is supported in $\widetilde{\Omega}_K(l_j)$ except if $j = 1$ and $g \in I(G)$;
- we drop the coprimality conditions between the $v_{g,j}(2)$, except that we remember that $v_{g,1}(2)$ and $v_{g',1}(2)$ are squarefree and coprime for $g, g' \in I(G)$.

Since there are only finitely many possibilities (depending on G and K) for $v_{g,j}(1)$, the above set has size bounded by

$$\ll_{G,K} \sum_{|N_{K/\mathbb{Q}}(\prod_{g\in G-\{\mathrm{id}\}} v_g(2)^{e_g})|\leq X} 1.$$

Here the sum runs over variables $v_g(2)$ for $g \in G - \{\mathrm{id}\}$, where

- for $g \in G - H(G)$, the variable $v_g(2)$ is an arbitrary integral ideal of $\mathcal{O}_K$;
- for $g \in I(G)$, the variables $v_g(2)$ are squarefree and supported in $\widetilde{\Omega}_K(l_G)$;
- for distinct $g, g' \in I(G)$, the variables $v_g(2)$ and $v_{g'}(2)$ are coprime.

We pull out the variables $v_g(2)$ for which g is not in $I(G)$. For such variables $v_g(2)$ we know that

$$e_g > \#G\left(1-\frac{1}{l_G}\right) = a(G)^{-1}. \tag{5.1}$$

Hence we get an upper bound

$$\ll \sum_{|N_{K/\mathbb{Q}}(\prod_{g\in G-H(G)} v_g(2)^{e_g})|\leq X} \sum_{\substack{|N_{K/\mathbb{Q}}(\prod_{h\in I(G)} v_h(2)^{a(G)^{-1}})|\leq \frac{X}{|N_{K/\mathbb{Q}}(\prod_{g\in G-H(G)} v_g(2)^{e_g})|} \\ v_h(2) \text{ squarefree and pairwise coprime} \\ v_h(2) \text{ supported in } \widetilde{\Omega}_K(l_G)}} 1. \tag{5.2}$$

Letting I be the product of $v_h(2)$ with $h \in I(G)$, we see that I is a squarefree ideal supported in $\widetilde{\Omega}_K(l_G)$. Equivalently, all prime divisors of I split in the extension $K(\zeta_{l_G})/K$. By Theorem 4.1, the inner sum is bounded by

$$\begin{aligned} &\sum_{\substack{|N_{K/\mathbb{Q}}(I^{a(G)^{-1}})|\leq \frac{X}{|N_{K/\mathbb{Q}}(\prod_{g\in G-H(G)} v_g(2)^{e_g})|} \\ I \text{ supported in } \widetilde{\Omega}_K(l_G)}} \mu^2(I)\#I(G)^{\omega(I)} \\ &\ll_{G,K} \frac{X^{a(G)}(\log X)^{i(G,K)-1}}{|N_{K/\mathbb{Q}}(\prod_{g\in G-H(G)} v_g(2)^{e_g a(G)})|}. \end{aligned}$$

Plugging this back in equation (5.2) yields

$$\begin{aligned} &X^{a(G)}\cdot\log(X)^{i(G,K)-1} \sum_{|N_{K/\mathbb{Q}}(\prod_{g\in G-H(G)} v_g(2)^{e_g})|\leq X} \frac{1}{|N_{K/\mathbb{Q}}(\prod_{g\in G-H(G)} v_g(2)^{e_g a(G)})|} \\ &\ll X^{a(G)}\cdot\log(X)^{i(G,K)-1} \prod_{g\in G-H(G)} \left(\sum_{|N_{K/\mathbb{Q}}(v_g(2)^{e_g})|\leq X} \frac{1}{|N_{K/\mathbb{Q}}(v_g(2)^{e_g a(G)})|} \right). \end{aligned}$$

The inner sum converges since $e_g a(G) > 1$ by equation (5.1), and this completes the proof. □

Now let $G := G(l_1)\times\cdots\times G(l_c)$ be a finite non-trivial nilpotent group such that the elements of $I(G)$ are pairwise commuting. In this case we are going to prove an upper bound matching the prediction of Malle's conjecture. As before we assume that the elements of $\{l_1,\dots,l_c\}$ are increasingly ordered, so that l_1 equals to l_G. Under these assumptions $H(G) = I(G)\cup\{\mathrm{id}\}$ is an $\mathbb{F}_{l_1}$ vector space and furthermore characteristic (hence normal) in $G(l_1)\subseteq G$. Write $h(G)$ for the $\mathbb{F}_{l_1}$-dimension of $H(G)$. Observe that $\frac{G}{H(G)}$ naturally acts on $H(G)$ by conjugation. For each $h\in H(G)$ we denote by

$$\mathrm{Stab}_{\frac{G}{H(G)}}(h)$$

the stabilizer of h under the above action, which is a subgroup of $\frac{G}{H(G)}$. Observe that

$$\#\,\mathrm{Conj}_G(h) = \left[\frac{G}{H(G)} : \mathrm{Stab}_{\frac{G}{H(G)}}(h)\right].$$

For each $2\leq j\leq c$, we filter $G(l_j)$ by any admissible sequence

$$\{(G_{i_j}(l_j),\theta_{i_j})\}_{i_j\in[r_j]}.$$

Instead for l_1 we filter $G(l_1)$ by an admissible sequence

$$\{(G_{i_1}(l_1), \theta_{i_1})\}_{i_1 \in [r_1]}$$

such that the kernel of the projection map from $G(l_1) = G_{r_1}(l_1)$ to $G_{r_1 - h(G)}(l_1)$ coincides with $H(G)$. In other words $H(G)$ equals the subset of vectors in $\mathbb{F}_{l_1}^{r_1}$ with last $r_1 - h(G)$ coordinates equal to 0.

Fix now

$$\psi \in \mathrm{Epi}_{\mathrm{top.gr.}}(G_K, \frac{G}{H(G)}).$$

Let us denote by

$$\mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})(\mathrm{solv.})(\psi)$$

the subset of

$$\mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})(\mathrm{solv.})$$

such that the induced epimorphism to $\frac{G}{H(G)}$, by means of the canonical projection, coincides with ψ. Observe that if $v := (v_{g,j}(1), v_{g,j}(2))_{j \in [c], g \in G(l_j) - \{\mathrm{id}\}} \in \mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})(\mathrm{solv.})(\psi)$, and $h \in H(G)$ and $\mathfrak{q} \mid v_h(2)$, then $\mathfrak{q}$ is unramified in the $\frac{G}{H(G)}$-extension given by ψ. Hence $\psi(\mathrm{Frob}_{\mathfrak{q}})$ is well-defined; we remind the reader that $\mathrm{Frob}_{\mathfrak{q}}$ depends on the choice of the embedding $i_{\mathfrak{q}}$ fixed so far in the paper.

Proposition 5.2. *Notation as immediately above this proposition. Then for each*

$$v := (v_{g,j}(1), v_{g,j}(2))_{j \in [c], g \in G(l_j) - \{\mathrm{id}\}} \in \mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})(\mathrm{solv.})(\psi),$$

for each $h \in H(G)$ and each $\mathfrak{q} \mid v_h(2)$ we have that $\psi(\mathrm{Frob}_{\mathfrak{q}}) \in \mathrm{Stab}_{\frac{G}{H(G)}}(h)$.

Proof. This is an immediate consequence of Proposition 2.11 and Proposition 3.4. □

As mentioned above $\mathrm{Frob}_{\mathfrak{q}}$ depends on the choice of embedding $i_{\mathfrak{q}}$. Unfortunately this means that $\mathrm{Frob}_{\mathfrak{q}}$ might not be equidistributed (as $\mathfrak{q}$ varies) for some choices of the embeddings $i_{\mathfrak{q}}$. But since we are free to choose the embeddings as we like, we are able to work around this.

Theorem 5.3. *Let G be a finite non-trivial nilpotent group and let K be a number field. Suppose that all elements of $I(G)$ commute with each other. Then there exists a constant $c \in \mathbb{R}_{>0}$ such that*

$$\#\{\psi \in \mathrm{Epi}_{\mathrm{top.gr.}}(G_K, G) : |N_{K/\mathbb{Q}}(\mathrm{Disc}(\psi))| \leq X\} \leq c \cdot X^{a(G)} \cdot \log(X)^{b_{\mathrm{M}}(G,K)-1}$$

for all $X \in \mathbb{R}_{>2}$.

Proof. We start as in the proof of Theorem 5.1 and we get the same upper bound as in equation (5.2) except that the $v_h(2)$ are now also such that $\mathfrak{q} \mid v_h(2)$ implies $\psi'(\mathrm{Frob}_{\mathfrak{q}}) \in \mathrm{Stab}_{\frac{G}{H(G)}}(h)$, where $\psi' : G_K \to G/H(G)$ is the composition of ψ with the quotient map $G \twoheadrightarrow G/H(G)$. More precisely, we see that

$$\#\{\psi \in \mathrm{Epi}_{\mathrm{top.gr.}}(G_K, G) : |N_{K/\mathbb{Q}}(\mathrm{Disc}(\psi))| \leq X\}$$

is bounded by

$$\sum_{|N_{K/\mathbb{Q}}(\prod_{g \in G - H(G)} v_g(2)^{e_g})| \leq X} \ \sum_{\substack{|N_{K/\mathbb{Q}}(\prod_{h \in I(G)} v_h(2)^{a(G)^{-1}})| \leq \frac{X}{|N_{K/\mathbb{Q}}(\prod_{g \in G - H(G)} v_g(2)^{e_g})|} \\ v_h(2) \text{ squarefree and pairwise coprime} \\ v_h(2) \text{ supported in } \widetilde{\Omega}_K(l_G) \\ \mathfrak{p} \mid v_h(2) \Rightarrow \psi'(\mathrm{Frob}_{\mathfrak{p}}) \in \mathrm{Stab}_{G/H(G)}(h)}} 1, \tag{5.3}$$

where ψ' is the map associated to the tuple $(v_g(1), v_g(2))$ as g runs through the elements that are zero on the coordinates corresponding to $H(G)$ (of course ignoring all tuples $(v_g(1), v_g(2))$ that map to $\bullet$).

We split the sum depending on

$$\left| N_{K/\mathbb{Q}} \left(\prod_{g \in G - H(G)} v_g(2)^{e_g} \right) \right| \leq (\log X)^{A_1}. \tag{5.4}$$

If we pick $A_1 > 0$ large enough (depending on G and K), the terms with

$$\left| N_{K/\mathbb{Q}} \left(\prod_{g \in G - H(G)} v_g(2)^{e_g} \right) \right| > (\log X)^{A_1}$$

can be bounded as in the proof of Theorem 5.1. Hence it remains to bound the terms satisfying equation (5.4). This implies that $\mathrm{Disc}(\psi') \leq (\log X)^{A_2}$ for a constant A_2 depending only on G and K. The Rosser–Iwaniec sieve [13, Lemma 3] gives some $A_3 > 0$ such that

$$\sum_{\substack{N_{K/\mathbb{Q}}(I) \leq X \\ \mathfrak{p} | I \Rightarrow \mathfrak{p} \in \widetilde{\Omega}_K(l_G) \\ \mathfrak{p} | I \Rightarrow \psi'(\mathrm{Frob}_\mathfrak{p}) \in \mathrm{Stab}_{G/H(G)}(h)}} \mu^2(I) \ll_{G,K} \frac{X}{\log X} \cdot \prod_{\substack{N_{K/\mathbb{Q}}(\mathfrak{p}) \leq X^{A_3} \\ \mathfrak{p} \in \widetilde{\Omega}_K(l_G) \\ \psi'(\mathrm{Frob}_\mathfrak{p}) \in \mathrm{Stab}_{G/H(G)}(h)}} \left(1 + \frac{1}{N_{K/\mathbb{Q}}(\mathfrak{p})}\right). \tag{5.5}$$

Let M be the compositum of all the extensions corresponding to a map $\psi' : G_K \to G/H(G)$ with $\mathrm{Disc}(\psi') \leq (\log X)^{A_2}$. Let $\mathrm{Emb}(X)$ be the set of functions f that send a place $\mathfrak{p} \in \Omega_K$ with $N_{K/\mathbb{Q}}(\mathfrak{p}) \leq X$ to a place of M above $\mathfrak{p}$. If a function f is given, it makes sense to speak of $\mathrm{Frob}_\mathfrak{p}$ as an element of $\mathrm{Gal}(M/K)$ and its quotients.

Before we proceed, let us quickly explain the main strategy. In equation (5.5) it would be particularly nice if we were to have equidistribution of the element $\mathrm{Frob}_\mathfrak{p}$. Observe that we indeed have an element $\mathrm{Frob}_\mathfrak{p}$, and not merely a conjugacy class of elements, because we have chosen an embedding $i_\mathfrak{p}$ in Section 2. Theorem 4.3 is the first step towards equidistribution, since it provides equidistribution of $\mathrm{Frob}_\mathfrak{p}$ over the conjugacy classes of G.

Although $N(X, G, K)$ obviously does not depend on the choice of embeddings $i_\mathfrak{p}$, it turns out that our proof does become substantially easier for certain choices of embeddings. More precisely, we want a choice of embeddings where we do not just have equidistribution of $\mathrm{Frob}_\mathfrak{p}$ over the conjugacy classes, but equidistribution as an element of G. The transition from conjugacy classes to elements will be an entirely combinatorial problem, while the equidistribution over conjugacy classes relies on deep analytic machinery, namely Theorem 4.3.

We now call a function $f \in \mathrm{Emb}(X)$ A_4-unfavorable in case there exists $\psi' : G_K \to G/H(G)$ with $\mathrm{Disc}(\psi') \leq (\log X)^{A_1}$ and there exists $g \in G/H(G)$ with

$$\left| \frac{\#\{\mathfrak{p} \in \widetilde{\Omega}_K(l_G) : N_{K/\mathbb{Q}}(\mathfrak{p}) \leq X, \psi'(\mathrm{Frob}_\mathfrak{p}) = g\}}{\#\{\mathfrak{p} \in \Omega_K : N_{K/\mathbb{Q}}(\mathfrak{p}) \leq X\}} - \frac{\mathrm{Li}(X)}{[K(\zeta_{l_G}) : K] \cdot \#\,\mathrm{Gal}(L/K)} \right| \geq \frac{X}{(\log X)^{A_4}}, \tag{5.6}$$

where L is the field corresponding to ψ'. Recall that $\mathfrak{p} \in \widetilde{\Omega}_K(l_G)$ is equivalent to $\mathfrak{p}$ splitting in $K(\zeta_{l_G})$, except for finitely many bad primes. Our first goal is to show that the set of A_4-unfavorable f is small.

Since the set of $\psi' : G_K \to G/H(G)$ with $\mathrm{Disc}(\psi') \leq (\log X)^{A_2}$ is bounded by $(\log X)^{A_5}$ for some $A_5 > 0$ depending only on G and K (see Theorem 5.1 for example), it suffices to bound the number of f failing equation (5.6), where we now treat ψ' and g as fixed. We proceed to bound the LHS of equation (5.6) by

$$\begin{aligned}(5.7)\quad &\left|\frac{\#\{\mathfrak{p} \in \widetilde{\Omega}_{K,X}(l_G) : \psi'(\mathrm{Frob}_{\mathfrak{p}}) = g\}}{\#\{\mathfrak{p} \in \Omega_{K,X}\}} - \frac{\mathrm{Li}(X)}{[K(\zeta_{l_G}) : K] \cdot \#\operatorname{Gal}(L/K)}\right| \leq \\ &\left|\frac{\#\{\mathfrak{p} \in \widetilde{\Omega}_{K,X}(l_G) : \psi'(\mathrm{Frob}_{\mathfrak{p}}) = g\}}{\#\{\mathfrak{p} \in \Omega_{K,X}\}} - \frac{\#\{\mathfrak{p} \in \widetilde{\Omega}_{K,X}(l_G) : \psi'(\mathrm{Frob}_{\mathfrak{p}}) \in C\}}{\#C \cdot \#\{\mathfrak{p} \in \Omega_{K,X}\}}\right| + \\ &\left|\frac{\#\{\mathfrak{p} \in \widetilde{\Omega}_{K,X}(l_G) : \psi'(\mathrm{Frob}_{\mathfrak{p}}) \in C\}}{\#C \cdot \#\{\mathfrak{p} \in \Omega_{K,X}\}} - \frac{\mathrm{Li}(X)}{[K(\zeta_{l_G}) : K] \cdot \#\operatorname{Gal}(L/K)}\right|,\end{aligned}$$

where C is the unique conjugacy class of G containing g and $\Omega_{K,X}(l_G)$ and $\widetilde{\Omega}_{K,X}(l_G)$ are respectively $\Omega_K(l_G) \cap \{\mathfrak{p} : N_{K/\mathbb{Q}}(\mathfrak{p}) \leq X\}$ and $\widetilde{\Omega}_K(l_G) \cap \{\mathfrak{p} : N_{K/\mathbb{Q}}(\mathfrak{p}) \leq X\}$. Now we apply Theorem 4.3 to the compositum of (the fixed field of the kernel of) ψ' and $K(\zeta_{l_G})$ with a very large A. This shows that the latter expression in equation (5.7) is small. Therefore if f is A_4-unfavorable, we deduce that

$$\begin{aligned}&\left|\frac{\#\{\mathfrak{p} \in \widetilde{\Omega}_{K,X}(l_G) : \psi'(\mathrm{Frob}_{\mathfrak{p}}) = g\}}{\#\{\mathfrak{p} \in \Omega_{K,X}\}} - \frac{\#\{\mathfrak{p} \in \widetilde{\Omega}_{K,X}(l_G) : \psi'(\mathrm{Frob}_{\mathfrak{p}}) \in C\}}{\#C \cdot \#\{\mathfrak{p} \in \Omega_{K,X}\}}\right| \\ &\quad \geq \frac{X}{2(\log X)^{A_4}}.\end{aligned}$$

We have now reduced our problem to an entirely combinatorial problem. Fix a conjugacy class C, which we view as a probability space by giving every element $c \in C$ probability equal to $1/\#C$. Write

$$Y := \prod_{\substack{\mathfrak{p} \in \widetilde{\Omega}_{K,X}(l_G) \\ \psi'(\mathrm{Frob}_{\mathfrak{p}}) \in C}} C,$$

which is also a probability space as it is a finite, cartesian product of probability spaces. The natural map $\mathrm{Emb}(X) \to Y$, sending f to $\psi'(\mathrm{Frob}_{\mathfrak{p}})$, is surjective with fibers all of equal cardinality. Note that $\mathrm{Frob}_{\mathfrak{p}}$ (implicitly) depends on the choice of f, so this map indeed depends on the choice of $f \in \mathrm{Emb}(X)$.

Now also fix some choice of $g \in C$. We let $Y_{\mathfrak{p}}$ be the random variable on Y given by $\mathbf{1}_{\pi_{\mathfrak{p}}(y)=g}$, where $\pi_{\mathfrak{p}}$ is projection on the $\mathfrak{p}$-th coordinate of Y. An application of Hoeffding's inequality with the random variables $Y_{\mathfrak{p}}$ and our fixed g and C shows that

$$\frac{\#\{f \in \mathrm{Emb}(X) : f \text{ is } A_4\text{-unfavorable}\}}{\#\{f \in \mathrm{Emb}(X)\}}$$

is small for any fixed $A_4 > 0$.

In particular, we can fix one choice of embeddings that is not, say, 100-unfavorable. It follows from partial summation and equations (5.5) and (5.6) that

$$\sum_{\substack{N_{K/\mathbb{Q}}(I)\le X\\ \mathfrak{p}|I\Rightarrow \mathfrak{p}\in\widetilde{\Omega}_K(l_G)\\ \mathfrak{p}|I\Rightarrow \psi'(\mathrm{Frob}_{\mathfrak{p}})\in \mathrm{Stab}_{G/H(G)}(h)}} \mu^2(I) \ll_{G,K} \frac{X}{(\log X)^{1-\frac{\#\,\mathrm{Stab}_{G/H(G)}(h)}{[K(\zeta_{l_G}):K]\cdot\#(G/H(G))}}}. \tag{5.8}$$

We will now bound each inner sum of equation (5.3). We drop the condition that the $v_h(2)$ are pairwise coprime in equation (5.3). Define for each h the multiplicative function $f_h(I)$ as follows

$$f_h(\mathfrak{p}^k) := \begin{cases} 1 & \text{if } k=1, \mathfrak{p}\in\widetilde{\Omega}_K(l_G) \text{ and } \psi'(\mathrm{Frob}_{\mathfrak{p}})\in \mathrm{Stab}_{G/H(G)}(h),\\ 0 & \text{otherwise.}\end{cases}$$

We have a good estimate for the average of $f_h(I)$ by equation (5.8). If we let g be the convolution of the $f_h(I)$ as h runs over all elements of $I(G)$, then the inner sum of equation (5.3), without the pairwise coprime condition, is equal to

$$\sum_{|N_{K/\mathbb{Q}}(I^{a(G)^{-1}})|\le \frac{X}{|N_{K/\mathbb{Q}}(\prod_{g\in G-H(G)} v_g(2)^{e_g})|}} g(I).$$

The theorem then follows from Lemma 3.5 after repeated application of Dirichlet's hyperbola method to the above sum and inserting the bound (5.8). □

5.2. **Poitou–Tate duality.** In order to prove Theorem 1.1 it will be convenient to pick a favorable choice of the characters $\chi_{\mathfrak{q}}$. In particular we would like to show the following. Let K be a number field, let l be a prime number and let $S_{\mathrm{clean}}(l)$ be as in Section 2. Then there exists an extension L/K such that $\mathrm{Frob}_{L/K}(\mathfrak{q}) = \mathrm{Frob}_{L/K}(\mathfrak{q}')$ implies that the characters $\chi_{\mathfrak{q}}, \chi_{\mathfrak{q}'} : G_K \to \mathbb{F}_l$ can be chosen in such a way that their restriction to

$$\bigoplus_{\mathfrak{q}\in S_{\mathrm{clean}}(l)} \frac{H^1(G_{K_{\mathfrak{q}}}, \mathbb{F}_l)}{H^1_{\mathrm{unr}}(G_{K_{\mathfrak{q}}}, \mathbb{F}_l)}$$

is the same. Here we recall that $\chi_{\mathfrak{q}}$ is a character satisfying the following two properties: the place $\mathfrak{q}$ ramifies in the field corresponding to $\chi_{\mathfrak{q}}$, and furthermore any other ramified place must be in $S_{\mathrm{clean}}(l)$. We will use the following form of Poitou–Tate duality to achieve our goal. Let us start by giving some background material on Selmer groups.

Let M be a finite, discrete G_K-module. We define for each place v the unramified classes to be

$$H^1_{\mathrm{unr}}(G_{K_v}, M) := \ker\left(H^1(G_{K_v}, M) \to H^1(G_{K_v^{\mathrm{unr}}}, M)\right)$$

with K_v^{unr} the maximal, unramified extension of K_v. A Selmer structure for M is then a collection $\mathcal{L} = \{\mathcal{L}_v\}_v$, where each $\mathcal{L}_v$ is a subgroup of $H^1(G_{K_v}, M)$ such that $\mathcal{L}_v = H^1_{\mathrm{unr}}(G_{K_v}, M)$ for all but finitely many places. The associated Selmer group $\mathrm{Sel}_{\mathcal{L}}(G_K, M)$ is then the kernel of the map

$$H^1(G_K, M) \to \prod_{v\in\Omega_K} H^1(G_{K_v}, M)/\mathcal{L}_v.$$

Define $M^* = \mathrm{Hom}(M, \mathbb{Q}/\mathbb{Z}(1))$, where $\mathbb{Q}/\mathbb{Z}(1)$ is the Tate twist of $\mathbb{Q}/\mathbb{Z}$. We have the local Tate pairing

$$H^1(G_{K_v}, M) \times H^1(G_{K_v}, M^*) \to H^2(G_{K_v}, \mathbb{Q}/\mathbb{Z}(1)) \cong \mathrm{Br}(K_v) \to \mathbb{Q}/\mathbb{Z}$$

given by the cup product and the local invariant map. The dual Selmer structure is then defined to be the orthogonal complement of $\mathcal{L}_v$ under the local Tate pairing, which gives subspaces $\{\mathcal{L}_v^*\}_v$ of $H^1(G_{K_v}, M^*)$. If v does not divide $|M|$ and the inertia group I_{K_v} acts trivially on M, then it is known that $H^1_{\mathrm{unr}}(G_{K_v}, M)$ and $H^1_{\mathrm{unr}}(G_{K_v}, M^*)$ are orthogonal complements under the local Tate pairing. The dual Selmer group is defined as the kernel of the map

$$H^1(G_K, M^*) \to \prod_{v \in \Omega_K} H^1(G_{K_v}, M^*)/\mathcal{L}_v^*.$$

Theorem 5.4 (Poitou–Tate duality)**.** *Let $\mathcal{L}$ and $\mathcal{F}$ be Selmer structures such that $\mathcal{L}_v \subseteq \mathcal{F}_v$ for each v. Let Ω be a finite set of places such that $\mathcal{L}_v = \mathcal{F}_v$ for all $v \notin \Omega$. Then we have exact sequences*

$$0 \to \mathrm{Sel}_{\mathcal{L}}(G_K, M) \to \mathrm{Sel}_{\mathcal{F}}(G_K, M) \to \bigoplus_{v \in \Omega} \mathcal{F}_v/\mathcal{L}_v$$

and

$$0 \to \mathrm{Sel}_{\mathcal{F}^*}(G_K, M^*) \to \mathrm{Sel}_{\mathcal{L}^*}(G_K, M^*) \to \bigoplus_{v \in \Omega} \mathcal{L}_v^*/\mathcal{F}_v^*.$$

Now consider the pairing

$$\bigoplus_{v \in \Omega} \mathcal{F}_v/\mathcal{L}_v \times \bigoplus_{v \in \Omega} \mathcal{L}_v^*/\mathcal{F}_v^* \to \mathbb{Q}/\mathbb{Z},$$

which is by definition the sum of the local Tate pairings at each $v \in \Omega$. Then the images of $\mathrm{Sel}_{\mathcal{F}}(G_K, M)$ in $\oplus_{v \in \Omega} \mathcal{F}_v/\mathcal{L}_v$ and $\mathrm{Sel}_{\mathcal{L}^*}(G_K, M^*)$ in $\oplus_{v \in \Omega} \mathcal{L}_v^*/\mathcal{F}_v^*$ are orthogonal complements.

Proof. See [40, Theorem 2.3.4]. □

We remark that it is only the last part of the theorem that is deep. With Theorem 5.4 in hand, it is now easy to control the local behavior at the places in $S_{\mathrm{clean}}(l)$.

Theorem 5.5. *There exists an extension L/K with the following property. Take any two primes $\mathfrak{p}, \mathfrak{p}' \in \widetilde{\Omega}_K(l)$. Assume that*

$$\mathrm{Frob}_{L/K}(\mathfrak{p}) = \mathrm{Frob}_{L/K}(\mathfrak{p}').$$

Then we can choose characters $\chi_{\mathfrak{p}}$ and $\chi_{\mathfrak{p}'}$ such that they have the same restriction to

$$\bigoplus_{\mathfrak{q} \in S_{\mathrm{clean}}(l)} \frac{H^1(G_{K_{\mathfrak{q}}}, \mathbb{F}_l)}{H^1_{\mathrm{unr}}(G_{K_{\mathfrak{q}}}, \mathbb{F}_l)}.$$

Remark 2. By a choice for $\chi_{\mathfrak{p}}$ we mean a character from G_K to $\mathbb{F}_l$ that is ramified at $\mathfrak{p}$, and unramified at all other places except possibly for those in $S_{\mathrm{clean}}(l)$.

Proof. We apply Theorem 5.4 as follows. We take $M = \mathbb{F}_l$ so that $M^* \cong \langle \zeta_l \rangle$. For now fix a finite place w outside $S_{\mathrm{clean}}(l)$ but in $\widetilde{\Omega}_K(l)$. We take $\mathcal{L} = \{\mathcal{L}_v\}_v$ with $\mathcal{L}_v = H^1_{\mathrm{unr}}(G_{K_v}, M)$ for all places v. Furthermore, we take $\mathcal{F}^w = \{\mathcal{F}_v^w\}_v$ with

$$\mathcal{F}_v^w = \begin{cases} H^1(G_{K_v}, M) & \text{if } v \in S_{\mathrm{clean}}(l) \cup \{w\}, \\ H^1_{\mathrm{unr}}(G_{K_v}, M) & \text{otherwise.} \end{cases}$$

Finally, we take Ω to be the union of $S_{\mathrm{clean}}(l)$ with $\{w\}$.

By Kummer theory we know that $H^1(G_K, M^*)$ can be identified with K^*/K^{*l}. A computation then shows that $\mathrm{Sel}_{\mathcal{L}^*}(G_K, M^*)$ is given by the elements $\alpha \in K^*/K^{*l}$ that have valuation divisible by l at all finite places v. Now take the field L to be

$$L := K(\zeta_l, \{\sqrt[l]{\alpha} : \alpha \in \mathrm{Sel}_{\mathcal{L}^*}(G_K, M^*)\}).$$

Suppose now that we are given two primes $\mathfrak{p}, \mathfrak{p}' \in \widetilde{\Omega}_K(l)$ with

$$\mathrm{Frob}_{L/K}(\mathfrak{p}) = \mathrm{Frob}_{L/K}(\mathfrak{p}'). \tag{5.9}$$

Take two non-zero elements

$$x_{\mathfrak{p}} \in \frac{H^1(G_{K_{\mathfrak{p}}}, \mathbb{F}_l)}{H^1_{\mathrm{unr}}(G_{K_{\mathfrak{p}}}, \mathbb{F}_l)}, \quad x_{\mathfrak{p}'} \in \frac{H^1(G_{K_{\mathfrak{p}'}}, \mathbb{F}_l)}{H^1_{\mathrm{unr}}(G_{K_{\mathfrak{p}'}}, \mathbb{F}_l)}.$$

Observe that any character $\chi \in \mathrm{Sel}_{\mathcal{F}^{\mathfrak{p}}}(G_K, M)$ restricting to $x_{\mathfrak{p}}$ is a valid choice of $\chi_{\mathfrak{p}}$, and similarly for $x_{\mathfrak{p}'}$ and $\chi_{\mathfrak{p}'}$. By construction of $S_{\mathrm{clean}}(l)$ we can find some

$$y \in \bigoplus_{\mathfrak{q} \in S_{\mathrm{clean}}(l)} \frac{H^1(G_{K_{\mathfrak{q}}}, \mathbb{F}_l)}{H^1_{\mathrm{unr}}(G_{K_{\mathfrak{q}}}, \mathbb{F}_l)}$$

such that the pair $(x_{\mathfrak{p}}, y)$ is in the image of $\mathrm{Sel}_{\mathcal{F}^{\mathfrak{p}}}(G_K, M)$.

To complete the proof, we will show that there exists $\lambda \in \mathbb{F}_l^*$ such that $(\lambda x_{\mathfrak{p}'}, y)$ is in the image of $\mathrm{Sel}_{\mathcal{F}^{\mathfrak{p}'}}(G_K, M)$. By Poitou–Tate duality this is equivalent to showing that $(\lambda x_{\mathfrak{p}'}, y)$ is orthogonal to all $\alpha \in \mathrm{Sel}_{\mathcal{L}^*}(G_K, M^*)$ under the sum of the local Tate pairings. Consider the linear functionals $\varphi_{\mathfrak{p}}, \varphi_{\mathfrak{p}'} : \mathrm{Sel}_{\mathcal{L}^*}(G_K, M^*) \to \frac{1}{l}\mathbb{Z}_l/\mathbb{Z}_l$ given respectively by

$$\alpha \mapsto \mathrm{inv}_{\mathfrak{p}}(x_{\mathfrak{p}} \cup \mathrm{res}_{\mathfrak{p}}(\alpha)), \quad \alpha \mapsto \mathrm{inv}_{\mathfrak{p}'}(x_{\mathfrak{p}'} \cup \mathrm{res}_{\mathfrak{p}'}(\alpha)),$$

where res denotes the natural restriction map. We claim that $\ker(\varphi_{\mathfrak{p}}) = \ker(\varphi_{\mathfrak{p}'})$.

Once the claim is proven, it follows that there exists $\lambda \in \mathbb{F}_l^*$ such that $\varphi_{\mathfrak{p}} = \lambda\varphi_{\mathfrak{p}'}$. We write y as a tuple $(y_v)_{v \in S_{\mathrm{clean}}(l)}$. Since

$$\varphi_{\mathfrak{p}}(\alpha) = \mathrm{inv}_{\mathfrak{p}}(x_{\mathfrak{p}} \cup \mathrm{res}_{\mathfrak{p}}(\alpha))$$

is precisely the local Tate pairing of $x_{\mathfrak{p}}$ with α at the place $\mathfrak{p}$ and $x_{\mathfrak{p}}$ is already known to be orthogonal to all $\alpha \in \mathrm{Sel}_{\mathcal{L}^*}(G_K, M^*)$ (again by Poitou–Tate duality), we get that

$$\varphi_{\mathfrak{p}}(\alpha) + \sum_{v \in S_{\mathrm{clean}}(l)} \mathrm{inv}_v(y_v \cup \mathrm{res}_{\mathfrak{p}}(\alpha)) = 0.$$

Therefore we get that

$$\lambda\varphi_{\mathfrak{p}'}(\alpha) + \sum_{v \in S_{\mathrm{clean}}(l)} \mathrm{inv}_v(y_v \cup \mathrm{res}_{\mathfrak{p}}(\alpha)) = 0.$$

It follows that $(\lambda x_{\mathfrak{p}'}, y)$ is also orthogonal to all $\alpha \in \mathrm{Sel}_{\mathcal{L}^*}(G_K, M^*)$. We conclude that the claim implies the theorem, so it remains to prove the claim.

In order to prove the claim, we observe that

$$\mathrm{inv}_{\mathfrak{p}}(x_{\mathfrak{p}} \cup \mathrm{res}_{\mathfrak{p}}(\alpha)) = 0 \iff \mathfrak{p} \text{ splits completely in } K(\zeta_l, \sqrt[l]{\alpha}).$$

Then the claim is a consequence of equation (5.9). □

5.3. **Proof of Theorem 1.1.** Recall that the quantities $i(G,K)$ and $b_{\mathrm{M}}(G,K)$ coincide if and only if $I(G)$ is entirely contained in the center of G. In this subsection we shall establish an asymptotic for $N(G,K,X)$ for such groups G.

Recall that for a finite group we denote by $Z(G)$ the center of G. Let now $G := G(l_1)\times\cdots\times G(l_c)$ be a finite non-trivial group with each $G(l_i)$ an l_i-group. We assume that $I(G)\subseteq Z(G)$ and we order the primes $l_1,\dots,l_c$ such that $l_1<\cdots<l_c$, so that $l_1 = l_G$.

In particular we see that $H(G) := I(G)\cup\{\mathrm{id}\}$ is a vector space over $\mathbb{F}_{l_1}$, we denote by $h(G)$ its dimension. For each $2\leq j\leq c$, we filter $G(l_j)$ by any admissible sequence

$$\{(G_{i_j}(l_j),\theta_{i_j})\}_{i_j\in[r_j]}.$$

Instead for l_1 we filter $G(l_1)$ by an admissible sequence

$$\{(G_{i_1}(l_1),\theta_{i_1})\}_{i_1\in[r_1]}$$

such that the kernel of the projection map from $G(l_1) = G_{r_1}(l_1)$ to $G_{r_1-h(G)}(l_1)$ coincides with $H(G)$. In other words $H(G)$ equals the subset of vectors in $\mathbb{F}_{l_1}^{r_1}$ with last $r_1-h(G)$ coordinates equal to 0. Let us denote by

$$\mathrm{Prim}(\mathcal{S}^{G-H(G)})(\mathrm{solv.})\subseteq \mathrm{Prim}(\mathcal{S}_{l_1}^{G(l_1)-H(G)})\times\prod_{2\leq j\leq c}\mathrm{Prim}(\mathcal{S}_{l_j}^{G(l_j)-\{\mathrm{id}\}})$$

the image of the projection map π from $\mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})(\mathrm{solv.})$ that drops the coordinates in $H(G)$. We also denote by $\left(\prod_{j\in[c]}\mathrm{Prim}\left(\mathcal{S}_{l_j}^{G(l_j)-\{\mathrm{id}\}}\right)\right)^{\circ}$ the subset of vectors

$$(v_{g,j}(1),v_{g,j}(2))_{g\in G(l_j)-\{\mathrm{id}\},j\in[c]}$$

such that $v_{g,1}(2)$ is not the trivial ideal for all $g\in H(G)-\{\mathrm{id}\}$.

Proposition 5.6. *We have*

$$\mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})(\mathrm{solv.})\supseteq$$
$$\left(\mathrm{Prim}\left(\mathcal{S}_{l_1}^{H(G)-\{\mathrm{id}\}}\right)\times\mathrm{Prim}(\mathcal{S}^{G-H(G)})(\mathrm{solv.})\right)\bigcap\left(\prod_{j\in[c]}\mathrm{Prim}\left(\mathcal{S}_{l_j}^{G(l_j)-\{\mathrm{id}\}}\right)\right)^{\circ}.$$

Furthermore, we have

$$\mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})(\mathrm{solv.})\subseteq(\mathrm{Prim}(\mathcal{S}_{l_1}^{H(G)-\{\mathrm{id}\}})\times\mathrm{Prim}(\mathcal{S}^{G-H(G)})(\mathrm{solv.}))$$
$$\cap\prod_{j\in[c]}\mathrm{Prim}\left(\mathcal{S}_{l_j}^{G(l_j)-\{\mathrm{id}\}}\right).$$

Remark 3. In case $G = H(G)$, the set $\mathrm{Prim}(\mathcal{S}^{G-H(G)})(\mathrm{solv.})$ is by definition the one element set containing the empty tuple, so that

$$\mathrm{Prim}(\mathcal{S}^{H(G)-\{\mathrm{id}\}})\times\mathrm{Prim}(\mathcal{S}^{G-H(G)})(\mathrm{solv.}) = \mathrm{Prim}(\mathcal{S}^{H(G)-\{\mathrm{id}\}}).$$

Proof. Let $\mathcal{G}$ be any profinite group and let $\psi\in\mathrm{Hom}_{\mathrm{top.gr.}}(\mathcal{G},G)$ such that

$$\pi_{\mathrm{mod}\ H(g)}\circ\psi$$

is in $\mathrm{Epi}_{\mathrm{top.gr.}}(\mathcal{G},\frac{G}{H(G)})$. Let $\chi\in\mathrm{Hom}_{\mathrm{top.gr.}}(\mathcal{G},H(G))$. Note that the assignment

$$\chi\cdot\psi:\mathcal{G}\to G,\quad g\mapsto\chi(g)\cdot\psi(g)$$

is an element of $\mathrm{Hom}_{\mathrm{top.gr.}}(\mathcal{G}, G)$. Indeed, this map is clearly continuous and furthermore

$$\begin{aligned}\chi(g_1g_2)\psi(g_1g_2) &= \chi(g_1)\chi(g_2)\psi(g_1)\psi(g_2) = \chi(g_1)\psi(g_1)\chi(g_2)\psi(g_2)\\ &= (\chi\cdot\psi)(g_1)(\chi\cdot\psi)(g_2).\end{aligned}$$

Here the second equality uses that $H(G) \subseteq Z(G)$.

Next observe that ψ and χ induce natural maps

$$\psi^* : \mathrm{Hom}(G, \mathbb{F}_{l_1}) \to \mathrm{Hom}_{\mathrm{top.gr.}}(\mathcal{G}, \mathbb{F}_{l_1})$$

and $\chi^* : \mathrm{Hom}(H(G), \mathbb{F}_{l_1}) \to \mathrm{Hom}_{\mathrm{top.gr.}}(\mathcal{G}, \mathbb{F}_{l_1})$. We define V to be the image of ψ^* and W to be the image of χ^*, so that V and W are naturally $\mathbb{F}_{l_1}$ vector spaces. We claim that if

$$V \cap W = \{0\}$$

and if χ^* is injective, then

$$\chi\cdot\psi \in \mathrm{Epi}_{\mathrm{top.gr.}}(\mathcal{G}, G).$$

Since G is nilpotent, it is enough to show that $\chi\cdot\psi$ surjects modulo the Frattini subgroup or equivalently

$$\chi' \circ (\chi\cdot\psi) \neq 0 \tag{5.10}$$

for every non-trivial character $\chi' : G \to \mathbb{F}_{l_i}$. We aim to establish equation (5.10). Let us distinguish two cases. First assume that

$$\chi'(H(G)) = \{0\}.$$

Then $\chi'\circ(\chi\cdot\psi) = \chi'\circ\psi$ and the claim follows from the assumption that $\pi_{\mathrm{mod}\ H(G)}\circ\psi$ is surjective. Suppose now that $\chi'(H(G)) \neq \{0\}$, which implies $l_i = l_1$. Since χ^* is injective it follows that $\chi' \circ \chi \neq 0$. On the other hand we also know that $V \cap W = \{0\}$. This means that there exists $g \in \mathcal{G}$ with $(\chi' \circ \psi)(g) = 0$ and $(\chi' \circ \chi)(g) = 1$. Therefore we find that

$$\chi' \circ (\chi\cdot\psi)(g) = 1$$

and we have established equation (5.10).

We apply the above to

$$\mathcal{G} = \mathcal{G}_K^{\mathrm{pro}-l_1} \times \cdots \times \mathcal{G}_K^{\mathrm{pro}-l_c}.$$

Fix $(y_g)_{g\in G-\{\mathrm{id}\}} \in \mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})(\mathrm{solv.})$. Take now any vector

$$(y'_g)_{g\in G-\{\mathrm{id}\}} \in \mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})$$

with $y'_g = y_g$ for each $g \in G - H(G)$ and $y'_h(2) \neq (1)$ for each $h \in H(G) - \{\mathrm{id}\}$, where we write $y'_h = (y'_h(1), y'_h(2))$. We must show that

$$(y'_g)_{g\in G-\{\mathrm{id}\}} \in \mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})(\mathrm{solv.}).$$

Through the map $\mathrm{Pow}([h(G)])$, we see that $(y_h)_{h\in H(G)-\{\mathrm{id}\}}$ corresponds uniquely to a character $\chi \in \mathrm{Hom}_{\mathrm{top.gr.}}(\mathcal{G}, H(G))$. We have that the map

$$\psi := \chi^{-1}\cdot P_G((y_g)_{g\in G-\{\mathrm{id}\}})$$

is an element of $\mathrm{Hom}_{\mathrm{top.gr.}}(\mathcal{G}, G)$ such that $\pi_{\mathrm{mod}\ H(G)} \circ \psi$ is surjective.

Let χ' be the character from $\mathcal{G}$ to $H(G)$ corresponding to $(y'_h)_{h\in H(G)-\{\mathrm{id}\}}$. Since $y'_h \neq 1$ for each $h \in H(G) - \{\mathrm{id}\}$, it follows that χ' is surjective and hence χ'^* is injective. We claim that $V \cap W = \{0\}$.

Take a non-trivial character $\rho \in \mathrm{Hom}(H(G), \mathbb{F}_{l_1})$. Then there exists some $h \in H(G) - \{\mathrm{id}\}$ such that $\rho(h) \neq 0$. Since $y'_h(2) \neq (1)$ by assumption, there exists a prime ideal $\mathfrak{p}$ such that $\mathfrak{p} \mid y'_h(2)$. It follows that $\chi'^*(\rho)(\sigma_\mathfrak{p}) = \rho(\chi'(\sigma_\mathfrak{p})) = \rho(h) \neq 0$. However, observe that ψ is only ramified at primes in S or primes dividing $y_g(2)$ for some $g \in G - H(G)$ by construction. Therefore, by the coprimality conditions on the variables, ψ is unramified at $\mathfrak{p}$ and thus $\psi(\sigma_\mathfrak{p}) = 0$, which readily implies the claim.

Hence

$$\chi' \cdot \psi \in \mathrm{Epi}_{\mathrm{top.gr.}}(\mathcal{G}, G),$$

and furthermore

$$P_G((y'_g)_{g \in G - \{\mathrm{id}\}}) = \chi' \cdot \psi.$$

This establishes the first part of the proposition. The second part is straightforward. □

We are now ready to prove our main theorem.

Theorem 5.7. *Let G be a finite non-trivial nilpotent group such that $I(G)$ is entirely contained in the center of G. Then there exists a constant $c > 0$ such that*

$$\#\{\psi \in \mathrm{Epi}_{\mathrm{top.gr.}}(G_K, G) : |N_{K/\mathbb{Q}}(\mathrm{Disc}(\psi))| \leq X\} \sim c \cdot X^{a(G)} \cdot \log(X)^{b_{\mathrm{M}}(G,K)-1}.$$

Proof. We start by labelling the elements of $\mathrm{Prim}(\mathcal{S}^{G-H(G)})(\mathrm{solv.})$ as $x_1, x_2, x_3, \ldots$ and we write L for the length of the sequence, where L is possibly infinite. Recall that π denotes the natural projection map from $\mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})(\mathrm{solv.})$ to $\mathrm{Prim}(\mathcal{S}^{G-H(G)})(\mathrm{solv.})$. We have the decomposition

$$\#\{\psi \in \mathrm{Epi}_{\mathrm{top.gr.}}(G_K, G) : |N_{K/\mathbb{Q}}(\mathrm{Disc}(\psi))| \leq X\} = \sum_{i=1}^{L} \sum_{\substack{y \in \mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})(\mathrm{solv.}) \\ \pi(y) = x_i \\ \mathrm{Disc}(y) \leq X}} 1,$$

where $\mathrm{Disc}(y)$ is the discriminant of $P_G(y)$. We claim that for all i there is a constant $b_{G,x_i} > 0$ such that

$$\sum_{\substack{y \in \mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})(\mathrm{solv.}) \\ \pi(y) = x_i \\ \mathrm{Disc}(y) \leq X}} 1 \sim b_{G,x_i} \cdot X^{a(G)} \cdot \log(X)^{b_{\mathrm{M}}(G,K)-1}. \tag{5.11}$$

We will first prove the theorem assuming the claim. Let us recall the statement of Tannery's theorem, which, in modern terms, is just the dominated convergence theorem on ℓ^1. Let $f_i : \mathbb{Z}_{\geq 1} \to \mathbb{C}$ be functions and let

$$S(X) = \sum_{i=1}^{\infty} f_i(X)$$

and suppose that $\lim_{X \to \infty} f_i(X) = b_i$. If $|f_i(X)| \leq M_i$ for sufficiently large X and

$$\sum_{i=1}^{\infty} M_i < \infty, \tag{5.12}$$

then $\lim_{X \to \infty} S(X)$ exists, $\sum_{i=1}^{\infty} b_i$ converges absolutely and

$$\lim_{X \to \infty} S(X) = \sum_{i=1}^{\infty} b_i. \tag{5.13}$$

We apply Tannery's theorem with

$$f_i(X) := \frac{1}{X^{\frac{l_G}{(l_G-1)\#G}} \cdot \log(X)^{b_{\mathrm{M}}(G,K)-1}} \cdot \sum_{\substack{y \in \mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})(\mathrm{solv.}) \\ \pi(y)=x_i \\ \mathrm{Disc}(y)\le X}} 1,$$

so that $\lim_{X\to\infty} f_i(X) = b_{G,x_i}$ by equation (5.11). It follows from the proof of Theorem 5.1 that equation (5.12) is satisfied with

$$M_i = \frac{C_G}{\prod_{g\in G-H(G)} y_g^{\frac{l_G e_g}{(l_G-1)\#G}}},$$

where C_G is a constant. Then equation (5.13) shows that

$$\lim_{X\to\infty} \frac{1}{X^{\frac{l_G}{(l_G-1)\#G}} \cdot \log(X)^{b_{\mathrm{M}}(G,K)-1}} \cdot \sum_{\substack{y \in \mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})(\mathrm{solv.}) \\ \mathrm{Disc}(y)\le X}} 1 = \lim_{X\to\infty} S(X) = \sum_{i=1}^{L} b_{G,x_i}.$$

Since $b_{G,x_i} > 0$ and the sum is non-empty by Shafarevich's theorem, the theorem follows.

To establish the claim, we take a more close look at the conditions imposed on y for a given x_i. Write x_i as a tuple $(v_{g'}(1), v_{g'}(2))_{g'\in G-H(G)}$. Let $(v_g(1), v_g(2))_{g\in I(G)}$ be in $\mathrm{Prim}(\mathcal{S}_{l_1}^{H(G)-\{\mathrm{id}\}})$. Fix, for every $g = (g_1, \dots, g_c) \in I(G)$, any choice of

$$v_g(1) = (v_{g_1,1}(1), \dots, v_{g_c,c}(1))$$

such that $v_{h_j,j}(1)$ and $v_{h'_j,j}(1)$ are coprime (as defined in Section 2) for all distinct $h_j, h'_j \in G(l_j)$ and for all $j \in [c]$.

Then by Proposition 5.6 we see that $\{v_g(2)\}_{g\in I(G)}$ (together with x_i and the variables $v_g(1)$ for $g \in I(G)$) is in $\mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})(\mathrm{solv.})$ if the $v_g(2)$ are squarefree non-trivial ideals and pairwise coprime, coprime to the $v_{g'}(2)$ for $g' \in G - H(G)$ and all prime divisors of $v_g(2)$ lie in $\widetilde{\Omega}_K(l_G)$. Conversely, if $\{v_g(2)\}_{g\in I(G)}$ (together with x_i and the variables $v_g(1)$ for $g \in I(G)$) is in $\mathrm{Prim}(\mathcal{S}^{G-\{\mathrm{id}\}})(\mathrm{solv.})$, then the $v_g(2)$ are squarefree and pairwise coprime, coprime to the $v_{g'}(2)$ for $g' \in G - H(G)$ and all prime divisors of $v_g(2)$ lie in $\widetilde{\Omega}_K(l_G)$.

However, we would like to achieve some finer control. Indeed, Proposition 2.13 does not give us full control over the discriminant, but only over the part outside of some finite set of bad places S. To remedy this, we apply Theorem 5.5 with this set of places S. Let L/K be the extension guaranteed by Theorem 5.5; tracing through the proof, we see that we can take the field L to be an elementary abelian extension of $K(\zeta_{l_G})$ of degree a power of l_G. Let $S_1, \dots, S_k$ be the conjugacy classes of $\mathrm{Gal}(L/K)$ that project trivially to the identity in $\mathrm{Gal}(K(\zeta_{l_G})/K)$. Now choose the characters $\chi_{\mathfrak{p}}$ as in Theorem 5.5. Concretely this means that

$$\mathrm{Frob}_{L/K}(\mathfrak{p}) = \mathrm{Frob}_{L/K}(\mathfrak{p}')$$

implies that the characters $\chi_{\mathfrak{p}}$ and $\chi_{\mathfrak{p}'}$ have the same restriction to

$$\bigoplus_{\mathfrak{q}\in S} \frac{H^1(G_{K_{\mathfrak{q}}}, \mathbb{F}_{l_G})}{H^1_{\mathrm{unr}}(G_{K_{\mathfrak{q}}}, \mathbb{F}_{l_G})}.$$

We recall that we have already fixed the finitely many possibilities of $\{v_g(1)\}_{g\in I(G)}$. Motivated by Theorem 5.5, we further split the sum in equation (5.11) depending on the values of

$$a_{j,g} \coloneqq \omega_{S_j}(v_g(2)) \bmod l_G$$

for $g \in I(G)$, where we remind the reader that $\omega_{S_j}(I)$ denotes the number of primes $\mathfrak{p}$ dividing I such that $\mathrm{Frob}_{L/K}(\mathfrak{p}) = S_j$. Indeed, we claim that if two vectors $\{v_g(2)\}_{g\in I(G)}, \{w_g(2)\}_{g\in I(G)}$ are such that

$$\omega_{S_j}(v_g(2)) \equiv \omega_{S_j}(v_g(2)) \bmod l_G$$

for all j and all g, then the S-part of the discriminant is the same for the two extensions corresponding to respectively $\{v_g(2)\}_{g\in I(G)}$ and $\{w_g(2)\}_{g\in I(G)}$. Because all elements of order l_G are central, we see that the ϕ maps appearing in the parametrization depend only on x_i (and not on $v_g(1)$ or $v_g(2)$ for $g \in I(G)$). Because of our choice of $\chi_{\mathfrak{p}}$, we deduce that $\omega_{S_j}(v_g(2)) \bmod l_G$ together with x_i determines the image of the inertia subgroup $I_{\mathfrak{q}}$ for every $\mathfrak{q} \in S$, which establishes the claim.

Hence the sum in equation (5.11) can be upper bounded by finitely many sums of the shape

$$\sum_{\substack{|\prod_{g\in I(G)} N_{K/\mathbb{Q}}(v_g(2))|\leq C(x_i,\mathbf{a})X^{a(G)} \\ \omega_{S_j}(v_g(2))\equiv a_{j,g} \bmod l_G \\ v_g(2) \text{ squarefree and pairwise coprime} \\ \mathfrak{p}|v_g(2)\Rightarrow \mathfrak{p}\in\widetilde{\Omega}_K(l_G) \\ \gcd(v_g(2),\mathfrak{p})=1 \ \forall \mathfrak{p}\in\mathcal{P}(x_i)}} 1, \tag{5.14}$$

where $\mathbf{a}$ is any vector in $\mathbb{F}_{l_G}^{[k]\times I(G)}$ with entries $a_{j,g}$, $C(x_i, \mathbf{a})$ is a positive real number depending only on K, G, x_i, $\{v_g(1)\}_{g\in I(G)}$, $\mathbf{a}$ and $\mathcal{P}(x_i)$ is a finite set depending only on K, G, x_i, $\{v_g(1)\}_{g\in I(G)}$. Similarly, the sum in equation (5.11) can be lower bounded by finitely many sums of the shape

$$\sum_{\substack{|\prod_{g\in I(G)} N_{K/\mathbb{Q}}(v_g(2))|\leq C(x_i,\mathbf{a})X^{a(G)} \\ v_g(2)\neq(1) \\ \omega_{S_j}(v_g(2))\equiv a_{j,g} \bmod l_G \\ v_g(2) \text{ squarefree and pairwise coprime} \\ \mathfrak{p}|v_g(2)\Rightarrow \mathfrak{p}\in\widetilde{\Omega}_K(l_G) \\ \gcd(v_g(2),\mathfrak{p})=1 \ \forall \mathfrak{p}\in\mathcal{P}(x_i)}} 1. \tag{5.15}$$

We shall give an asymptotic for equation (5.14). From the proof it shall be clear how to extract a matching asymptotic for equation (5.15), which implies the claimed equation (5.11). It remains to give an asymptotic for equation (5.14).

Our first step is to pass to $K(\zeta_{l_G})$. We define $\mathcal{P}'(x_i)$ as the set of finite places w of $K(\zeta_{l_G})$ such that v is in $\mathcal{P}(x_i)$ or v does not split completely in $K(\zeta_{l_G})/K$, where v denotes the unique place of K below w. Then equation (5.14) becomes

$$\sum_{\substack{|\prod_{g\in I(G)} N_{K(\zeta_{l_G})/\mathbb{Q}}(v_g(2))|\leq C(x_i,\mathbf{a})X^{a(G)} \\ \omega_{S_j}(v_g(2))\equiv a_{j,g} \bmod l_G \\ v_g(2) \text{ squarefree and pairwise coprime} \\ \gcd(v_g(2),\mathfrak{p})=1 \ \forall \mathfrak{p}\in\mathcal{P}'(x_i)}} \frac{1}{[K(\zeta_{l_G}):K]^{\sum_{g\in I(G)}\omega(v_g(2))}}, \tag{5.16}$$

where the $v_g(2)$ are now ideals of $K(\zeta_{l_G})$. To evaluate this sum, define $f_j(I)$ to be the function on $\mathcal{I}_{K(\zeta_{l_G})}$ that sends I to zero if I is divisible by a square, by a prime

$\mathfrak{p} \in \mathcal{P}'(x_i)$ or any $\mathfrak{p}$ with $\mathrm{Frob}_{\mathfrak{p}} \notin S_j$. If instead I is a squarefree ideal entirely supported on primes $\mathfrak{p}$ with $\mathrm{Frob}_{\mathfrak{p}} \in S_j$ and $\mathfrak{p} \notin \mathcal{P}'(x_i)$, we define

$$f_j(I) = \frac{\#\{(I_g)_{g\in I(G)} \in \mathcal{I}^{I(G)}_{K(\zeta_{l_G})} : \prod_{g\in I(G)} I_g = I, \omega_{S_j}(I_g) \equiv a_{j,g} \bmod l_G\}}{[K(\zeta_{l_G}) : K]^{\omega(I)}}.$$

Having defined $f_j(I)$, we see that equation (5.16) is simply

$$\sum_{N_{K/\mathbb{Q}}(I) \le C(x_i,\mathbf{a})X^{a(G)}} (f_1 * \cdots * f_k)(I). \tag{5.17}$$

We will now approximate $f_j(I)$ using some probability theory. Assume, for now, that I is a squarefree ideal entirely supported on primes $\mathfrak{p}$ with $\mathrm{Frob}_{\mathfrak{p}} \in S_j$ and $\mathfrak{p} \notin \mathcal{P}'(x_i)$. First define for $(b_{j,g})_{g\in I(G)} \in \mathbb{Z}^{I(G)}_{\ge 0}$

$$g_{j,(b_{j,g})_{g\in I(G)}}(I) = \frac{\#\{(I_g)_{g\in I(G)} \in \mathcal{I}^{I(G)}_{K(\zeta_{l_G})} : \prod_{g\in I(G)} I_g = I, \omega_{S_j}(I_g) = b_{j,g}\}}{\#I(G)^{\omega(I)}}.$$

Let us recall the multinomial distribution. As input it takes an integer n, which is the sample size, an integer k, which are the number of mutually exclusive events $E_1, \dots, E_k$, and real numbers $0 \le p_1, \dots, p_k \le 1$ such that p_i is the probability of the event E_i. We further demand

$$p_1 + \cdots + p_k = 1.$$

The multinomial distribution is then a vector $X = (X_1, \dots, X_k)$, where X_i indicates the number of outcomes of event E_i in n independent samples. We now take $n = \omega(I)$, $k = \#I(G)$ and $p_1 = \cdots = p_k = 1/k$. Then $g_{j,(b_{j,g})_{g\in I(G)}}(I)$ is the probability density function of the resulting multinomial distribution. Concretely,

$$\mathbb{P}(X_g = b_{j,g} \text{ for all } g \in I(G)) = g_{j,(b_{j,g})_{g\in I(G)}}(I) = \frac{\omega(I)!}{\prod_{g\in I(G)} b_{j,g}!} \cdot \frac{1}{\#I(G)^{\omega(I)}},$$

where we have relabelled the variables X_i as X_g with $g \in I(G)$. Now observe that

$$\frac{\omega(I)!}{\prod_{g\in I(G)} b_{j,g}!}$$

are precisely the coefficients when one expands

$$\left(\sum_{g\in I(G)} p_g\right)^{\omega(I)}$$

as a polynomial in the variables p_g. Evaluating this polynomial at $p_g = \zeta_{l_G}^{c_g}/\#I(G)$ as c_g runs through all vectors in $\mathbb{F}_{l_G}^{I(G)}$ shows that

$$\begin{aligned}
&\left(\frac{[K(\zeta_{l_G}) : K]}{\#I(G)}\right)^{\omega(I)} f_j(I) \\
&\quad = l_G^{-\#I(G)} \sum_{(c_g)_{g\in I(G)} \in \mathbb{F}_{l_G}^{I(G)}} \left(\prod_{g\in I(G)} \zeta_{l_G}^{-c_g a_{j,g}}\right) \cdot \left(\sum_{g\in I(G)} \frac{\zeta_{l_G}^{c_g}}{\#I(G)}\right)^{\omega(I)} \\
&\quad = l_G^{-\#I(G)+1} \cdot \mathbf{1}_{\omega(I) \equiv \sum_{g\in I(G)} a_{j,g} \bmod l_G} + O(\delta^{\omega(I)})
\end{aligned} \tag{5.18}$$

with $\delta < 1$. We apply Theorem 4.1 with

$$z = \frac{\#I(G)}{[K(\zeta_{l_G}) : K]} \cdot \zeta_{l_G}^a$$

as a runs through $0, \dots, l_G - 1$. We conclude that

$$\sum_{N_{K/\mathbb{Q}}(I) \leq X} f_j(I) = CX(\log X)^{\frac{\#I(G)\#S_j}{[K(\zeta_{l_G}):K]\#\operatorname{Gal}(L/K(\zeta_{l_G}))}-1} + O\left(X(\log X)^{\frac{\#I(G)\#S_j}{[K(\zeta_{l_G}):K]\#\operatorname{Gal}(L/K(\zeta_{l_G}))}-1-\delta'}\right) \tag{5.19}$$

for some constant $C > 0$ and some $\delta' > 0$. The theorem now follows from equation (5.19), equation (5.17) and repeated application of Lemma 4.2. □

5.4. **Comparison with previous results.** The upper bound in Theorem 1.3 is always at least as good as the upper bound appearing in [29], where an upper bound of the shape

$$c_{G,K} \cdot X^{a(G)} \cdot \log(X)^{d(G,K)-1}$$

is established. For 2-groups we can compute $d(G, K)$ as follows. Take a refinement R of the upper central series of G

$$\{\mathrm{id}\} = G_r \subseteq G_{r-1} \subseteq \cdots \subseteq G_1 \subseteq G_0 = G$$

such that each quotient is of size 2 (we caution the reader that these G_i are closely related, but different than the ones in the definition of admissible sequence). Then define $A_i := G_{i-1} - G_i$ for $1 \leq i \leq r$ and

$$d(R, K) = \sum_{\substack{1 \leq i \leq r \\ A_i \cap \mathrm{inv}(G) \neq \emptyset}} |A_i|,$$

where $\mathrm{inv}(G) := \{g \in G - \{\mathrm{id}\} : g^2 = \mathrm{id}\}$. Now $d(G, K)$ is simply the minimum of $d(R, K)$ over all refinements R. We conclude this section by showing that for some groups G, Theorem 1.3 provides a strictly better upper bound. In particular the next example is a group of size 64, where [29] always gives (i.e. for any choice of the filtration $\{G_i\}$ as above) at least 4 logarithms more than Theorem 1.3.

Proposition 5.8. *Let*

$$G := \frac{\mathbb{F}_2[x_1, x_2]}{(x_1^2, x_2^2)} \rtimes \mathbb{F}_2^2,$$

where $(1, 0)$ *acts as multiplication by* $1 + x_1$ *and* $(0, 1)$ *as multiplication by* $1 + x_2$. *Then for any choice of the filtration* $\{G_i\}$ *as in* [29], *one has*

$$d(G, K) \geq \#\operatorname{Inv}(G) + 4.$$

Proof. Let

$$\{\mathrm{id}\} = G_6 \subseteq G_5 \subseteq \cdots \subseteq G_1 \subseteq G_0 = G$$

be a refinement of the upper central series where each group is of index 2 in the next one. We say that an element $g \in G$ has *weight* i in case $g \in G_i - G_{i+1}$. We claim that there are at least 4 elements of order 4 with the same weight as an involution.

We start by computing

$$[G, G] = (x_1, x_2) \rtimes \{0\}, \quad \frac{G}{[G, G]} \cong \mathbb{F}_2^3,$$

where (x_1, x_2) is the ideal generated by x_1 and x_2. From this we deduce that any surjective homomorphism

$$\pi : G \twoheadrightarrow \mathbb{F}_2^2$$

does not vanish identically on $\{0\} \rtimes \mathbb{F}_2^2$. Next observe that each of the G_i's above is normal in G. Therefore, since $[G : G_2] = 4$, we conclude that $\frac{G}{G_2}$ is an abelian group, which implies that

$$G_2 \supseteq [G, G] = (x_1, x_2) \rtimes \{0\}, \quad \frac{G}{G_2} \cong \mathbb{F}_2^2. \tag{5.20}$$

Looking at the surjective homomorphism $G \twoheadrightarrow \frac{G}{G_2} \cong \mathbb{F}_2^2$, we conclude that at least one of the involutions in the set

$$\{(0, (1, 0)), (0, (1, 1)), (0, (0, 1))\}$$

has weight at most 1. On the other hand equation (5.20) shows that every commutator is necessarily of weight at least 2. Hence at least one of the following three sets

$$\{(x_2, (1, 0)), (x_1 + x_2, (1, 0)), (x_2 + x_1x_2, (1, 0)), (x_1 + x_2 + x_1x_2, (1, 0))\},$$
$$\{(x_2, (1, 1)), (x_1, (1, 1)), (x_1 + x_1x_2, (1, 1)), (x_2 + x_1x_2, (1, 1))\},$$
$$\{(x_1, (0, 1)), (x_1 + x_2, (0, 1)), (x_1 + x_2 + x_1x_2, (0, 1)), (x_1 + x_1x_2, (0, 1))\}$$

consists entirely of elements with order 4 and weight 1. This shows that the set of elements having the same weight as an involution contains at least 4 elements that are not involutions, which is precisely the desired conclusion. □

Acknowledgments

A large part of this work was done while the authors held postdoctoral positions at the Max Planck Institute. The authors wish to thank the Max Planck Institute for Mathematics in Bonn for its great work conditions and an inspiring atmosphere. We are grateful to user 2734364041 of MathOverflow for answering a question related to Theorem 4.3, to Alex Bartel, Jürgen Klüners and Jeffrey Lagarias for comments on an earlier version of this paper and to Adam Morgan for introducing us to Poitou–Tate duality. We also thank the anonymous referee for many valuable comments.

References

[1] Brandon Alberts, *Statistics of the first Galois cohomology group: a refinement of Malle's conjecture*, Algebra Number Theory **15** (2021), no. 10, 2513–2569, DOI 10.2140/ant.2021.15.2513. MR4377858

[2] Brandon Alberts, *The weak form of Malle's conjecture and solvable groups*, Res. Number Theory **6** (2020), no. 1, Paper No. 10, 23, DOI 10.1007/s40993-019-0185-7. MR4047213

[3] Brandon Alberts and Evan O'Dorney, *Harmonic analysis and statistics of the first Galois cohomology group*, Res. Math. Sci. **8** (2021), no. 3, Paper No. 50, 16, DOI 10.1007/s40687-021-00283-2. MR4298097

[4] S. Ali Altuğ, Arul Shankar, Ila Varma, and Kevin H. Wilson, *The number of D_4-fields ordered by conductor*, J. Eur. Math. Soc. (JEMS) **23** (2021), no. 8, 2733–2785, DOI 10.4171/jems/1070. MR4269426

[5] Manjul Bhargava, *The density of discriminants of quartic rings and fields*, Ann. of Math. (2) **162** (2005), no. 2, 1031–1063, DOI 10.4007/annals.2005.162.1031. MR2183288

[6] Manjul Bhargava, *The density of discriminants of quintic rings and fields*, Ann. of Math. (2) **172** (2010), no. 3, 1559–1591, DOI 10.4007/annals.2010.172.1559. MR2745272

[7] M. Bhargava, A. Shankar, T. Taniguchi, F. Thorne, J. Tsimerman, and Y. Zhao, *Bounds on 2-torsion in class groups of number fields and integral points on elliptic curves*, J. Amer. Math. Soc. **33** (2020), no. 4, 1087–1099, DOI 10.1090/jams/945. MR4155220

[8] Manjul Bhargava, Arul Shankar, and Jacob Tsimerman, *On the Davenport-Heilbronn theorems and second order terms*, Invent. Math. **193** (2013), no. 2, 439–499, DOI 10.1007/s00222-012-0433-0. MR3090184

[9] Manjul Bhargava and Melanie Matchett Wood, *The density of discriminants of S_3-sextic number fields*, Proc. Amer. Math. Soc. **136** (2008), no. 5, 1581–1587, DOI 10.1090/S0002-9939-07-09171-X. MR2373587

[10] Henri Cohen, Francisco Diaz y Diaz, and Michel Olivier, *Enumerating quartic dihedral extensions of $\mathbb{Q}$*, Compositio Math. **133** (2002), no. 1, 65–93, DOI 10.1023/A:1016310902973. MR1918290

[11] Henri Cohen, Francisco Diaz y Diaz, and Michel Olivier, *On the density of discriminants of cyclic extensions of prime degree*, J. Reine Angew. Math. **550** (2002), 169–209, DOI 10.1515/crll.2002.071. MR1925912

[12] H. Cohen and H. W. Lenstra Jr., *Heuristics on class groups of number fields*, Number theory, Noordwijkerhout 1983 (Noordwijkerhout, 1983), Lecture Notes in Math., vol. 1068, Springer, Berlin, 1984, pp. 33–62, DOI 10.1007/BFb0099440. MR756082

[13] M. D. Coleman, *The Rosser-Iwaniec sieve in number fields, with an application*, Acta Arith. **65** (1993), no. 1, 53–83, DOI 10.4064/aa-65-1-53-83. MR1239243

[14] Jean-Marc Couveignes, *Enumerating number fields*, Ann. of Math. (2) **192** (2020), no. 2, 487–497, DOI 10.4007/annals.2020.192.2.4. MR4151082

[15] Boris Datskovsky and David J. Wright, *Density of discriminants of cubic extensions*, J. Reine Angew. Math. **386** (1988), 116–138, DOI 10.1515/crll.1988.386.116. MR936994

[16] H. Davenport and H. Heilbronn, *On the density of discriminants of cubic fields. II*, Proc. Roy. Soc. London Ser. A **322** (1971), no. 1551, 405–420, DOI 10.1098/rspa.1971.0075. MR491593

[17] J. Ellenberg, T. Tran, and C. Westerland, *Fox-Neuwirth-Fuks cells, quantum shuffle algebras, and Malle's conjecture for function fields*, Preprint, `arXiv:1701.04541`, 2017.

[18] Jordan S. Ellenberg and Akshay Venkatesh, *The number of extensions of a number field with fixed degree and bounded discriminant*, Ann. of Math. (2) **163** (2006), no. 2, 723–741, DOI 10.4007/annals.2006.163.723. MR2199231

[19] Jordan S. Ellenberg, Akshay Venkatesh, and Craig Westerland, *Homological stability for Hurwitz spaces and the Cohen-Lenstra conjecture over function fields*, Ann. of Math. (2) **183** (2016), no. 3, 729–786, DOI 10.4007/annals.2016.183.3.1. MR3488737

[20] Étienne Fouvry and Jürgen Klüners, *Cohen-Lenstra heuristics of quadratic number fields*, Algorithmic number theory, Lecture Notes in Comput. Sci., vol. 4076, Springer, Berlin, 2006, pp. 40–55, DOI 10.1007/11792086_4. MR2282914

[21] Étienne Fouvry and Jürgen Klüners, *On the 4-rank of class groups of quadratic number fields*, Invent. Math. **167** (2007), no. 3, 455–513, DOI 10.1007/s00222-006-0021-2. MR2276261

[22] É. Fouvry and P. Koymans, *Malle's conjecture for nonic Heisenberg extensions*, Preprint, `arXiv:2102.09465`, 2021.

[23] Frank Gerth III, *The 4-class ranks of quadratic fields*, Invent. Math. **77** (1984), no. 3, 489–515, DOI 10.1007/BF01388835. MR759260

[24] D. R. Heath-Brown, *The size of Selmer groups for the congruent number problem. II*, Invent. Math. **118** (1994), no. 2, 331–370, DOI 10.1007/BF01231536. With an appendix by P. Monsky. MR1292115

[25] Wei Ho, Arul Shankar, and Ila Varma, *Odd degree number fields with odd class number*, Duke Math. J. **167** (2018), no. 5, 995–1047, DOI 10.1215/00127094-2017-0050. MR3782066

[26] Henryk Iwaniec and Emmanuel Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004, DOI 10.1090/coll/053. MR2061214

[27] Jürgen Klüners, *A counterexample to Malle's conjecture on the asymptotics of discriminants* (English, with English and French summaries), C. R. Math. Acad. Sci. Paris **340** (2005), no. 6, 411–414, DOI 10.1016/j.crma.2005.02.010. MR2135320

[28] J. Klüners, *Über die Asymptotik von Zahlkörpern mit vorgegebener Galoisgruppe*, Shaker Verlag, Aachen, 2005.

[29] Jürgen Klüners, *The asymptotics of nilpotent Galois groups*, Acta Arith. **204** (2022), no. 2, 165–184, DOI 10.4064/aa211207-16-5. MR4458889

[30] Jürgen Klüners and Gunter Malle, *Counting nilpotent Galois extensions*, J. Reine Angew. Math. **572** (2004), 1–26, DOI 10.1515/crll.2004.050. MR2076117

[31] Jürgen Klüners and Jiuya Wang, *ℓ-torsion bounds for the class group of number fields with an ℓ-group as Galois group*, Proc. Amer. Math. Soc. **150** (2022), no. 7, 2793–2805, DOI 10.1090/proc/15882. MR4428868

[32] Peter Koymans and Carlo Pagano, *On the distribution of* $\mathrm{Cl}(K)[l^\infty]$ *for degree l cyclic fields*, J. Eur. Math. Soc. (JEMS) **24** (2022), no. 4, 1189–1283, DOI 10.4171/JEMS/1112. MR4397040

[33] Peter Koymans and Carlo Pagano, *Higher genus theory*, Int. Math. Res. Not. IMRN **4** (2022), 2772–2823, DOI 10.1093/imrn/rnaa196. MR4381932

[34] Peter Koymans and Carlo Pagano, *A sharp upper bound for the 2-torsion of class groups of multiquadratic fields*, Mathematika **68** (2022), no. 1, 237–258, DOI 10.1112/mtk.12123. MR4405977

[35] Stéphane Louboutin, *Explicit upper bounds for residues of Dedekind zeta functions and values of L-functions at $s = 1$, and explicit lower bounds for relative class numbers of CM-fields*, Canad. J. Math. **53** (2001), no. 6, 1194–1222, DOI 10.4153/CJM-2001-045-5. MR1863848

[36] Robert J. Lemke Oliver and Frank Thorne, *Upper bounds on number fields of given degree and bounded discriminant*, Duke Math. J. **171** (2022), no. 15, 3077–3087, DOI 10.1215/00127094-2022-0046. MR4497223

[37] Gunter Malle, *On the distribution of Galois groups*, J. Number Theory **92** (2002), no. 2, 315–329, DOI 10.1006/jnth.2001.2713. MR1884706

[38] Gunter Malle, *On the distribution of Galois groups. II*, Experiment. Math. **13** (2004), no. 2, 129–135. MR2068887

[39] R. Masri, F. Thorne, W.-L. Tsai, and J. Wang, *Malle's conjecture for $G \times A$, with $G = S_3, S_4, S_5$*, Preprint, `arXiv:2004.04651v2`, 2020.

[40] Barry Mazur and Karl Rubin, *Kolyvagin systems*, Mem. Amer. Math. Soc. **168** (2004), no. 799, viii+96, DOI 10.1090/memo/0799. MR2031496

[41] Hugh L. Montgomery and Robert C. Vaughan, *Multiplicative number theory. I. Classical theory*, Cambridge Studies in Advanced Mathematics, vol. 97, Cambridge University Press, Cambridge, 2007. MR2378655

[42] C. Pagano and E. Sofos, *4-Ranks and the general model for statistics of ray class groups of imaginary quadratic fields*, Preprint, `arXiv:1710.07587`, 2017.

[43] Wolfgang M. Schmidt, *Number fields of given degree and bounded discriminant*, Astérisque **228** (1995), 4, 189–195. Columbia University Number Theory Seminar (New York, 1992). MR1330934

[44] I.R. Shafarevich. Extensions with given points of ramification (Russian). *Inst. Hautes Études Sci. Publ. Math.* 18, 71-95, 1963. English translation in: Collected mathematical papers. Springer-Verlag, Berlin, 1989.

[45] A. Siad, *Monogenic fields with odd class number Part I: odd degree*, Preprint, `arXiv:2011.08834`, 2020.

[46] A. Siad, *Monogenic fields with odd class number Part II: even degree*, Preprint, `arXiv:2011.08842`, 2020.

[47] A. Smith, *Governing fields and statistics for 4-Selmer groups and 8-class groups*, Preprint, `arXiv:1607.07860`, 2016.

[48] A. Smith, *2^∞-Selmer groups, 2^∞-class groups, and Goldfeld's conjecture*, Preprint, `arXiv:1702.02325v2`, 2017.

[49] Takashi Taniguchi and Frank Thorne, *Secondary terms in counting functions for cubic fields*, Duke Math. J. **162** (2013), no. 13, 2451–2508, DOI 10.1215/00127094-2371752. MR3127806

[50] Jesse Thorner and Asif Zaman, *A unified and improved Chebotarev density theorem*, Algebra Number Theory **13** (2019), no. 5, 1039–1068, DOI 10.2140/ant.2019.13.1039. MR3981313

[51] Seyfi Türkelli, *Connected components of Hurwitz schemes and Malle's conjecture*, J. Number Theory **155** (2015), 163–201, DOI 10.1016/j.jnt.2015.03.005. MR3349443

[52] Jiuya Wang, *Malle's conjecture for $S_n \times A$ for $n = 3, 4, 5$*, Compos. Math. **157** (2021), no. 1, 83–121, DOI 10.1112/s0010437x20007587. MR4219215

[53] David J. Wright, *Distribution of discriminants of abelian extensions*, Proc. London Math. Soc. (3) **58** (1989), no. 1, 17–50, DOI 10.1112/plms/s3-58.1.17. MR969545

Department of Mathematics, University of Michigan, Ann Arbor, Michigan 48109
Email address: koymans@umich.edu

School of Mathematics and Statistics, University of Glasgow, University Place, Glasgow G12 8SQ, United Kingdom
Current address: Department of Mathematics and Statistics, Montreal, Quebec H3G 1M8, Canada
Email address: carlein90@gmail.com