ON A CONJECTURE FOR ℓ-TORSION IN CLASS GROUPS OF NUMBER FIELDS: FROM THE PERSPECTIVE OF MOMENTS

LILLIAN B. PIERCE, CAROLINE L. TURNAGE-BUTTERBAUGH, AND MELANIE MATCHETT WOOD

Abstract. It is conjectured that within the class group of any number field, for every integer $\ell \geq 1$, the ℓ-torsion subgroup is very small (in an appropriate sense, relative to the discriminant of the field). In nearly all settings, the full strength of this conjecture remains open, and even partial progress is limited. Significant recent progress toward average versions of the ℓ-torsion conjecture has relied crucially on counts for number fields, raising interest in how these two types of question relate. In this paper we make explicit the quantitative relationships between the ℓ-torsion conjecture and other well-known conjectures: the Cohen-Lenstra heuristics, counts for number fields of fixed discriminant, counts for number fields of bounded discriminant (or related invariants), counts for elliptic curves with fixed conductor. All of these considerations reinforce that we expect the ℓ-torsion conjecture to be true, despite limited progress toward it. Our perspective focuses on the relation between pointwise bounds, averages, and higher moments, and demonstrates the broad utility of the "method of moments."

1. Introduction

Associated to each number field $K/\mathbb{Q}$ is the ideal class group Cl_K, a finite abelian group that encodes information about arithmetic in K, and can also be seen as the Galois group of the maximal unramified abelian extension of K. Correspondingly, each number field K has a class number, defined to be the cardinality $|\mathrm{Cl}_K|$ of the class group. Given a field K, it is natural to ask about the size and structure of the class group; moreover, as K varies over an infinite set, or "family," of fields, it is natural to ask how the corresponding class groups are distributed. Interest in such questions has a long history, going back to C.F. Gauss's class number conjecture, early attempts at proving Fermat's Last Theorem, and Dirichlet's development of the class number formula. In particular, although class numbers arise initially from an algebraic construction, the class number formula identifies them with values of L-functions and hence to core questions in analytic number theory, including the Generalized Riemann Hypothesis.

Our focus in this paper is a well-known conjecture for ℓ-torsion in class groups, which, in the strongest form in which it has been proposed, remains stubbornly out of reach. For each integer $\ell \geq 1$, the ℓ-torsion subgroup of Cl_K is defined by

$$\mathrm{Cl}_K[\ell] := \{[\mathfrak{a}] \in \mathrm{Cl}_K : \ell[\mathfrak{a}] = \mathrm{Id}\},$$

in which we write the class group in additive notation. Thus for example, if the ℓ-torsion is trivial, the class number is not divisible by ℓ.

How big is $\mathrm{Cl}_K[\ell]$? For any number field $K/\mathbb{Q}$ of degree n and discriminant of absolute value D_K, we may trivially bound the ℓ-torsion subgroup for any integer $\ell \geq 1$ by the size of the full class group, so that[1]

$$1 \leq |\mathrm{Cl}_K[\ell]| \leq |\mathrm{Cl}_K| \ll_{n,\varepsilon} D_K^{1/2+\varepsilon}, \tag{1.1}$$

2010 *Mathematics Subject Classification.* 11R29, 11R45.

Key words and phrases. Class groups, counting number fields, elliptic curves.

[1]We will use Vinogradov's notation: $A \ll B$ denotes that there exists a constant C such that $|A| \leq CB$, and $A \ll_\kappa B$ denotes that C may depend on κ.

for all $\varepsilon > 0$ arbitrarily small, according to Landau's upper bound for the class number via the Minkowski bound (see e.g. [Nar80, Theorem 4.4]). But conjecturally, a much stronger upper bound should hold.

Conjecture 1.1 (ℓ-torsion Conjecture)**.**
For every integer $n \geq 2$, every field extension $K/\mathbb{Q}$ of degree n satisfies the property that for all primes $\ell \geq 2$ and every $\varepsilon > 0$,

$$|\mathrm{Cl}_K[\ell]| \ll_{n,\ell,\varepsilon} D_K^{\varepsilon}. \tag{1.2}$$

This conjecture was first raised as a question of Brumer and Silverman [BS96, Question CL(ℓ, d)] in the context of counting elliptic curves with fixed conductor; we will return to this original motivation, and a possible more precise rate of growth, in §6. It is known to have implications for many other problems: for example, proving this upper bound would imply strong bounds for the ranks of elliptic curves (see [EV07, §1.2]) and for counts of certain number fields (see Remarks 1.10 and 4.1). See also Zhang [Zha05, Conjecture 3.5] in the context of equidistribution of CM points on Shimura varieties, Duke [Duk98, p. 166] in consideration of counting number fields, and Belolipetsky and Lubotzky [BL17, Thm. 7.5], who show an equivalence between certain more precise forms of Conjecture 1.1 and a conjecture on counting non-uniform lattices in semisimple Lie groups. We have stated the conjecture here for ℓ prime, as for composite ℓ the result follows from the prime cases; see §7.1.

Conjecture 1.1 is known to be true in only one case: degree $n = 2$ and the prime $\ell = 2$; this follows from the genus theory of Gauss [Gau01], which shows that for a fundamental discriminant of absolute value D_K, we have $|\mathrm{Cl}_K[2]| \leq 2^{\omega(D_K)-1}$, where $\omega(D_K)$ denotes the number of distinct prime divisors of D_K. Since $\omega(m) \ll \log m / \log\log m$ (see e.g. [HW08, Ch. 22 §10]), it follows that for any quadratic field K, for every $\varepsilon > 0$,

$$|\mathrm{Cl}_K[2]| \ll_{\varepsilon} D_K^{\varepsilon}. \tag{1.3}$$

Recent progress has focused on verifying results toward Conjecture 1.1 by making small improvements on (1.1), either for all number fields of a fixed degree, or for "almost all" fields in a family of number fields of fixed degree. Results of the first type include [HV06, Pie05, Pie06, EV07, BST$^+$17]. Results of the second type, which can also be thought of as results "on average" over a family of fields, include [DH71, Bha05, HBP17, EPW17, PTBW20, FW17, FW18]. (For an overview of recent work, see Section 7.)

The apparent difficulty of verifying Conjecture 1.1, for a fixed prime ℓ and all fields of a fixed degree n, in any case other than that due to Gauss, leads to a question: is the strong pointwise version of Conjecture 1.1 stated above too much to expect?

In this paper we describe several explicit motivations for why we expect Conjecture 1.1 to hold. The key philosophy we employ is the "method of moments," surveyed in §2: this is the principle that in certain settings, if one can control arbitrarily high moments of a function, then one can deduce that there is not even one violation of a pointwise upper bound by the function. In particular we show *quantitatively* how Conjecture 1.1 follows from three other well-known conjectures in number theory: the Cohen-Lenstra heuristics; the discriminant multiplicity conjecture for counting number fields with fixed discriminant; and a generalization of the Malle conjecture for counting number fields with bounded discriminant. Furthermore, we quantify implications of these conjectures for counting elliptic curves with fixed conductor, and we prove new unconditional counts for elliptic curves.

Many recent works on ℓ-torsion in class groups have focused on average results over families of fields, and this type of work has been significantly constrained by the difficulty of counting number fields. Our investigations in this paper confirm that counting fields is not merely a technical difficulty one encounters when proving average results, but in fact lies at the heart of understanding ℓ-torsion in class groups.

We now give an overview of this paper, organized according to four main themes.

1.1. **Cohen-Lenstra and Cohen-Martinet heuristics (§3).** First, we make quantitatively precise how Conjecture 1.1 follows from the Cohen-Lenstra [CL84] and Cohen-Martinet [CM90] heuristics. In fact, we show that even a collection of weaker upper bounds for moments, implied by the Cohen-Lenstra-Martinet heuristics, would imply Conjecture 1.1. We frame our first result as a statement about a "family" $\mathscr{F}$ of degree n field extensions of $\mathbb{Q}$; we will at the moment leave the specifications of such a family rather vague, but this could indicate that we have fixed not only the degree and the Galois group of the Galois closure, but also the signature, certain local conditions, and possibly certain ramification restrictions (see §1.6 for precise conventions).

Let $\mathscr{F}$ denote a family of fields $K/\mathbb{Q}$ of degree n and let $\mathscr{F}(X)$ denote those with $0 < D_K \leq X$, where $D_K = |\mathrm{Disc}\, K/\mathbb{Q}|$. Consider for any fixed real number $k \geq 1$, integer $\ell \geq 1$, and real number $\alpha \geq 1$ the statement that for all $X \geq 1$,

$$\sum_{K \in \mathscr{F}(X)} |\mathrm{Cl}_K[\ell]|^k \ll_{n,\ell,k,\alpha} |\mathscr{F}(X)|^\alpha. \tag{1.4}$$

Theorem 1.2. *Let n, ℓ be fixed. Suppose that for a fixed value $\alpha \geq 1$, (1.4) is known for all integers $k \geq 1$. Then for every $\varepsilon > 0$ we have $|\mathrm{Cl}_K[\ell]| \ll_\varepsilon D_K^\varepsilon$ for every field $K \in \mathscr{F}$.*

This principle can be applied to any family $\mathscr{F}$ for which appropriate moment bounds are known. Here we record that Theorem 1.2 immediately implies that for fixed n, ℓ, uniform control of arbitrarily high k-th moments is sufficient to deduce Conjecture 1.1 in its full strength:

Corollary 1.3. *Let $n \geq 1$ be fixed and let $\mathscr{F}$ be taken to be the family of all degree n extensions of $\mathbb{Q}$. For each fixed integer $\ell \geq 1$, the conclusion of Conjecture 1.1 for ℓ-torsion for all $K \in \mathscr{F}$ would follow from knowing (1.4) holds for that fixed n, ℓ and a certain fixed $\alpha \geq 1$, for arbitrarily large k.*

See Remark 3.3 for an example of how to apply this moment approach to 4-torsion in quadratic fields; while this case may be deduced by other means (§7.1), this is nevertheless an illustration of the principle.

1.2. **Discriminant multiplicity conjecture (§4).** In the second theme, we will show quantitatively how Conjecture 1.1 follows from a well-known folk conjecture about counting number fields with fixed degree and fixed discriminant. By Hermite's finiteness theorem, for any fixed integer $D \geq 1$ there are finitely many extensions $K/\mathbb{Q}$ of degree n with $D_K = D$ (see e.g. [Ser97, §4.1]). The question remains: how many? Duke [Duk98, §3] and Ellenberg and Venkatesh [EV05, Conjecture 1.3] make the following prediction, which we call the discriminant multiplicity conjecture:

Conjecture 1.4 (Discriminant Multiplicity Conjecture)**.** *For each $n \geq 2$, for every $\varepsilon > 0$ and for every integer $D \geq 1$, at most $\ll_{n,\varepsilon} D^\varepsilon$ fields $K/\mathbb{Q}$ of degree n have $D_K = D$.*

Conjecture 1.4 is known to be true for quadratic fields ($n = 2$) but is not known in full for any $n \geq 3$. See §4 for a summary of known results toward the discriminant multiplicity conjecture.

Ellenberg and Venkatesh have noted that Conjecture 1.4 implies Conjecture 1.1 (see [EV05, p. 164]). In Theorem 1.7, we make this explicit by exhibiting quantitatively how even a partial result toward Conjecture 1.4 implies progress toward Conjecture 1.1. To do so, we define two notations:

Property 1.5 (Property $\mathbf{D}_n(\varpi)$)**.** *Let $n \geq 2$ be fixed. We say that property $\mathbf{D}_n(\varpi)$ holds if for every $\varepsilon > 0$ and for every fixed integer $D > 1$, at most $\ll_{n,\varepsilon} D^{\varpi+\varepsilon}$ fields $K/\mathbb{Q}$ of degree n have $D_K = D$.*

Property 1.6 (Property $\mathbf{C}_{n,\ell}(\Delta)$)**.** *Let $n, \ell \geq 1$ be fixed. We say that property $\mathbf{C}_{n,\ell}(\Delta)$ holds if for every $\varepsilon > 0$ and for all fields $K/\mathbb{Q}$ of degree n, we have $|\mathrm{Cl}_K[\ell]| \ll_{n,\ell,\varepsilon} D_K^{\Delta+\varepsilon}$.*

Now we make the relationship between $\mathbf{D}_n(\varpi)$ and the ℓ-torsion conjecture quantitatively precise.

Theorem 1.7. *Fix an integer $n \geq 2$ and any prime ℓ. If Property $\mathbf{D}_{n\ell}(\Delta)$ holds for a real number $\Delta \geq 0$, then Property $\mathbf{C}_{n,\ell}(\ell\Delta)$ holds.*

Note that the exponent we gain for the ℓ-torsion bound is significantly weaker (that is, larger) than the exponent we assume for counting fields with fixed discriminant; in this formulation, as ℓ increases, we need to be increasingly good at counting the relevant fields, to preserve the same quality of bound on ℓ-torsion. Theorem 1.7 has an immediate corollary:

Corollary 1.8. *Let $n \geq 1$ be fixed. Then for a fixed prime ℓ, the statement of Conjecture 1.1 holds for all fields of degree n, if Conjecture 1.4 is known for degree $n\ell$.*

Remark 1.9. We highlight here a difficulty of this approach; if one fixes n and wants to use this route to prove the ℓ-torsion conjecture for all ℓ for fields of degree n, then one would need to prove the discriminant multiplicity conjecture for infinitely many degrees. The recent work of the authors in [PTBW20] to some extent eliminates this inefficiency; in that work, once the degree n is fixed, (weak) results are proved for ℓ-torsion for all ℓ by proving weak results toward the discriminant multiplicity conjecture only for degree n; see §7.5 for further remarks.

1.3. **Generalized Malle Conjecture (§5).** In the third theme, we examine how Conjecture 1.4, and hence Conjecture 1.1, follows from a generalized Malle conjecture for counting number fields ordered via other invariants than the discriminant. This observation is due to Ellenberg and Venkatesh [EV05, Prop. 4.8], and we state it here only loosely (see §5 for a precise statement):

Theorem A ([EV05, Prop. 4.8])**.** *If a generalized Malle conjecture (see (5.3)) holds for all groups G, then Conjecture 1.4 holds, and hence Conjecture 1.1 holds.*

We highlight this theorem of Ellenberg and Venkatesh because it uses not only the method of moments but also a clever strategy of reinterpreting a k-th moment of an object as an average of k-fold objects, so that one can (conjecturally) obtain bounds for arbitrarily high moments even if one only assumes bounds for averages (i.e. first moments). (This leads to speculation on whether it could be advantageous to reinterpret k-th moments of ℓ-torsion as averages of ℓ'-torsion over certain k'-fold objects, but any such precise formulation is so far elusive.)

Remark 1.10. Note also that Alberts [Alb18] has shown that Conjecture 1.1 implies the generalized Malle conjecture that we give below in (5.3), providing a converse to Theorem A.

1.4. **Counting elliptic curves with fixed conductor (§6).** In our fourth theme, we observe that for every $k \geq 1$, a k-th moment estimate for counting elliptic curves with fixed conductor can be deduced from any of the conjectures mentioned above. In addition, unconditionally, we derive improved bounds for counting elliptic curves with fixed conductor.

To be precise, given a positive integer q, let $E(q)$ denote the number of isogeny classes of elliptic curves over $\mathbb{Q}$ with conductor q. Brumer and Silverman have put forward the following conjecture:

Conjecture 1.11 ([BS96, Conj. 3])**.** *For every $q \geq 1$, for all $\varepsilon > 0$,*

$$E(q) \ll_\varepsilon q^\varepsilon.$$

(For clarity, we note that an equivalently strong conjecture is expected to hold if we instead count the number $E'(q)$ of isomorphism classes of elliptic curves over $\mathbb{Q}$ with conductor q; the statement of our results below will hold for both $E(q)$ and $E'(q)$, up to a constant factor.) Brumer and Silverman proved the first result toward Conjecture 1.11, namely $E(q) \ll_\varepsilon q^{1/2+\varepsilon}$, by dominating this count by 3-torsion in certain quadratic fields; by combining aspects of [HV06, Thm. 4.5] and [EV07] this can now be improved to

$$E(q) \ll_\varepsilon q^{2\beta/3\log 3+\varepsilon} \ll_\varepsilon q^{0.1688\ldots+\varepsilon} \tag{1.5}$$

for an explicit constant $\beta = 0.2782...$; see §6 for an explication of this. We observe in Theorem 6.2 that Conjecture 1.11 would follow from counting sextic fields of fixed discriminant. But our main focus is on a moment statement related to Conjecture 1.11.

Brumer and Silverman's argument dominates $E(q)$ pointwise by a sup-norm of $|\mathrm{Cl}_K[3]|$ over quadratic fields; this comparison has been sharpened further in [HV06], still using a supremum over a set of certain quadratic fields. Thus proving an upper bound for 3-torsion for almost all such fields (or for an average, or for a fixed moment) does not immediately provide an analogous result for $E(q)$ (see Remark 6.4). The content of Duke and Kowalski [DK00, Prop. 1] is that a refined argument shows that nevertheless, average counts for $E(q)$ are dominated by average counts for 3-torsion in quadratic fields. (See [FNT92] for an analogue when the curves are ordered by discriminant rather than conductor.)

An upper bound for an average $\sum E(q)$ yields an upper bound for the number of q for which $E(q)$ can violate a certain pointwise upper bound; a comparable upper bound for a higher moment $\sum E(q)^k$ sharpens the upper bound on the number of possible exceptions. This motivates us to consider higher moments. We build on [BS96, DK00, HV06] to prove that the k-th moment of 3-torsion in quadratic fields dominates the (γk)-th moment of $E(q)$, for a certain numerical constant $\gamma \approx 2$ specified precisely below.

Theorem 1.12. *Let $\mathscr{F}_2(X)$ denote the family of quadratic fields with discriminant of absolute value at most X. Let $\gamma = \log 3/(2\beta) = 1.9745...$ be the numerical constant determined by $\beta = \beta(0) = 0.2782...$ as defined in* [HV06, Thm. 3.8]. *Let $k \geq 1/\gamma = 0.5064...$ be a fixed real number, and assume that there exists $\Theta_k \geq 1$ such that for every $\varepsilon > 0$ and for every $X \geq 1$,*

$$\sum_{K \in \mathscr{F}_2(X)} |\mathrm{Cl}_K[3]|^k \ll_{k,\varepsilon} X^{\Theta_k + \varepsilon}. \tag{1.6}$$

Then for that k, for every $\varepsilon > 0$ and for all $Q \geq 1$,

$$\sum_{q \leq Q} E(q)^{\gamma k} \ll_{k,\varepsilon} Q^{\Theta_k + \varepsilon}. \tag{1.7}$$

Notice that the conclusion for $E(q)$ is stronger than the assumption for $|\mathrm{Cl}_K[3]|$, since the power in the moment has nearly doubled in size. In particular, (1.6) shows that

$$\#\{K \in \mathscr{F}_2(X) : X/2 < D_K \leq X, |\mathrm{Cl}_K[3]| \geq D_K^{\varpi}\} \ll X^{\Theta_k - k\varpi},$$

while we see from (1.7) that

$$\#\{Q/2 \leq q \leq Q : E(q) \geq q^{\varpi}\} \ll Q^{\Theta_k - k\gamma\varpi}.$$

By combining Theorem 1.12 with moment bounds of the form (1.6) for 3-torsion in quadratic fields due to Heath-Brown and the first author in [HBP17, Cor. 1.4] we immediately obtain new unconditional bounds for moments of $E(q)$:

Corollary 1.13. *For each $q \geq 1$, let $E(q)$ denote the number of isogeny classes of elliptic curves over $\mathbb{Q}$ with conductor q. Let $\gamma = 1.9745...$ be the numerical constant as above. Then for every $Q \geq 1$, for all $\varepsilon > 0$,*

$$\sum_{q \leq Q} E(q)^{2\gamma} \ll_{\varepsilon} Q^{23/18+\varepsilon}.$$

More generally, for all real $1 \leq k \leq 4$,

$$\sum_{q \leq Q} E(q)^{\gamma k} \ll_{\varepsilon} Q^{(5k+13)/18+\varepsilon},$$

and for all real $k \geq 4$,

$$\sum_{q \leq Q} E(q)^{\gamma k} \ll_{\varepsilon} Q^{(2k+3)/6+\varepsilon}.$$

These results improve on all results implicitly derivable in previous literature; see §6 for details.

1.5. **Concluding sections §7 and §8.** In §7 we briefly survey certain recent works toward Conjecture 1.1, with a focus on the relation to the Generalized Riemann Hypothesis. In contrast to the impact of the other major conjectures we have mentioned, GRH does not appear to imply Conjecture 1.1 directly, due to an interesting interaction with another set of conjectures, about number fields with generating elements of small height. Finally, in §8 we include brief appendices with further remarks about moment bounds and one special case of the discriminant multiplicity conjecture.

1.6. **Notational conventions on families of fields.** A "family" (set) of fields is formally defined as follows. Given an integer $n \geq 1$ and a fixed transitive subgroup $G \subseteq S_n$, for every integer $X \geq 1$ we can define a family $\mathscr{F}(X)$ to be $\{K/\mathbb{Q} : \deg K/\mathbb{Q} = n, \mathrm{Gal}(\tilde{K}/\mathbb{Q}) \simeq G, D_K = |\mathrm{Disc}\, K/\mathbb{Q}| \leq X\}$, where all K are in a fixed algebraic closure $\overline{\mathbb{Q}}$, $\tilde{K}$ is the Galois closure of K over $\mathbb{Q}$, the Galois group is considered as a permutation group on the n embeddings of K in $\overline{\mathbb{Q}}$, and the isomorphism with G is one of permutation groups. We let $\mathscr{F} = \mathscr{F}(\infty)$. It is furthermore possible to impose additional restrictions on the definition of the family, such as fixing the signature, certain local conditions, or ramification restrictions. In some cases where we wish to be more vague (for example, considering all degree n extensions), we will just refer to a family $\mathscr{F}$ and the subset $\mathscr{F}(X)$ of those fields $K \in \mathscr{F}$ with $D_K \leq X$. Given a family $\mathscr{F}$, we will say that a set $E \subset \mathscr{F}$ is a subset of density zero if

$$\frac{|E \cap \mathscr{F}(X)|}{|\mathscr{F}(X)|} \to 0 \quad \text{as } X \to \infty.$$

If a property holds for all fields in a family except those lying in a possible exceptional set of density zero, we say the property holds for "almost all" fields in the family.

2. The method of moments

We start with a general philosophy for why one can expect control of arbitrarily high moments of a function to yield good control on the function's values. On a (σ-finite) measure space $(\mathcal{M}, \mu)$ such as $\mathbb{R}$ with Lebesgue measure or $\mathbb{Z}$ with counting measure, given a complex-valued measurable function f, define for every real $\alpha > 0$ the distribution function $\lambda(\alpha) = \mu(\{x : |f(x)| > \alpha\})$. Then for any $\alpha > 0$,

$$\alpha^p \lambda(\alpha) \leq \int_{\{|f|>\alpha\}} |f(x)|^p d\mu \leq \int_{\mathcal{M}} |f(x)|^p d\mu = \|f\|^p_{L^p(\mathcal{M},\mu)}. \tag{2.1}$$

Suppose that it is known that for a sequence of p growing arbitrarily large we have $\|f\|^p_{L^p} \leq C$ for a fixed constant C, uniformly in p. Assuming this, one intuitively sees that for fixed $\alpha > 1$, as $p \to \infty$ through such a sequence, $\lambda(\alpha)$ must be vanishingly small in (2.1). This is the principle by which uniform control of arbitrarily high moments implies pointwise (μ-a.e.) bounds. (By a "moment" we refer to the p-th power of the L^p norm, and not the L^p norm itself. Using the notation of the distribution function $\lambda(\alpha)$ defined above, we have

$$\int_{\mathcal{M}} |f(x)|^p d\mu = -\int_0^\infty \alpha^p d\lambda(\alpha) = p\int_0^\infty \alpha^{p-1}\lambda(\alpha) d\alpha,$$

which are all expressions for how "moments" are defined in different contexts, motivating the terminology.)

Let us see more concretely how this philosophy applies in a model of the discrete setting that is our concern in this paper, by taking counting measure on $\mathbb{Z}$ (denoted by $|E|$ for a set $E \subset \mathbb{Z}$) and considering integer-valued functions of finite support (that is, vanishing aside from at finitely many integers). We will consider a restricted norm $\|f\|^p_{\ell^p(A_R)} = \sum_{n \in A_R} |f(n)|^p$ where A_R is the dyadic

interval $A_R = \{n : R \le n \le 2R\}$ with $R \ge 1$. Now we will aim to show not that f is uniformly bounded but that it does not grow too quickly.

For any $1 \le p < \infty$ and any $\Delta > 0$ the following inequality holds:

$$R^{\Delta p}|\{n \in A_R : |f(n)| > n^{\Delta}\}| < \sum_{\{n \in A_R : |f(n)| > n^{\Delta}\}} |f(n)|^p \le \sum_{n \in A_R} |f(n)|^p = \|f\|^p_{\ell^p(A_R)}. \tag{2.2}$$

Assume that for a sequence of $p \ge 1$, for all $R \ge 1$ we have

$$\|f\|^p_{\ell^p(A_R)} \le |A_R|^{\beta} \tag{2.3}$$

for some fixed $\beta > 0$. Then we may conclude that for each such p, for all $R \ge 1$, we have

$$|\{n \in A_R : |f(n)| > n^{\Delta}\}| \le |A_R|^{\beta} R^{-\Delta p} \le R^{\beta - \Delta p}. \tag{2.4}$$

In particular, if the sequence of p for which we know this grows arbitrarily large, we can for each R take p sufficiently large that $R^{\beta - \Delta p} < 1$ so that the set on the left-hand side of (2.4) must be empty, that is, for all $n \in A_R$ we have $|f(n)| \le n^{\Delta}$. (Somewhat more nuanced arguments are required if the bound in (2.3) is not uniform in p; see the proof of Theorem 1.2 and the comments in §8.1.)

We will adapt this method of moments first to the setting in which our "discrete space" is a set of fields, and the function f measures the size of the ℓ-torsion subgroup (§3); and later, to the setting in which the discrete space is $\mathbb{N}$ and f counts the number of fields with a specified discriminant (§5). These ideas may be adapted to many other settings, such as other functions mapping fields (ordered by some invariant) to non-negative real numbers; for detailed remarks from the perspective of L^p spaces, which may be helpful in such generalizations, see the appendix in §8.1.

Note that in practice, even if one cannot prove uniform control of arbitrarily high moments, it can sometimes be useful to prove an upper bound for one particular moment, such as the second moment. If an upper bound for an average shows that a certain pointwise bound must hold for almost all elements in a family (that is, for all but an exceptional family of density zero), then proving an appropriate upper bound on a higher moment can force a better upper bound on the cardinality of the possible exceptional family. See Remark 3.2 for an instance of this; this also motivates the interest for Theorem 1.12 and Corollary 1.13 on elliptic curves.

The method of moments has a long history and wide applicability in number theory; we mention two additional settings. First, the method of moments shows that the Lindelöf Hypothesis for the Riemann zeta function would be a consequence of upper bounds for arbitrarily large even moments of $\zeta(1/2 + it)$ along the critical line. Precisely, arguments such as [Tit86, Thm. 13.2] show that the statement

$$\zeta(1/2 + it) \ll_{\varepsilon} t^{\varepsilon} \qquad \text{for all } \varepsilon > 0, \text{ for all } t \in \mathbb{R}$$

is equivalent to the statement that for all $k = 1, 2, 3, \ldots$,

$$\frac{1}{T}\int_1^T |\zeta(1/2 + it)|^{2k} dt \ll_{k,\varepsilon} T^{\varepsilon} \qquad \text{for all } \varepsilon > 0, \text{ for all } T \in \mathbb{R}.$$

(We remark that in this context of Lebesgue measure on $\mathbb{R}$, rather than a discrete counting measure such as we will consider, one requires some additional consideration to ensure that a result holds for all t and not just a.e. t. For this result on the zeta function, this comes from an integral expression for the analytic continuation, and the resulting notion of the derivative of ζ in the critical strip, so that any large value of $\zeta(1/2 + it)$ would imply that ζ also takes large values on a positive measure neighborhood.)

Second, the principle of the method of moments can be used to deduce the Weil-Deligne bound for character sums of prime moduli from another Riemann Hypothesis, that of L-functions of algebraic varieties over finite fields. This area is very deep (see e.g. [Kow10] for an overview). Even given Deligne's proof [Del74, Del80] of the Weil conjectures (which includes deductions for certain one-variable additive character sums), deducing the desired (square-root cancellation) bounds for

complete character sums in higher dimensions is difficult (the subject, for example, of many celebrated works of N. Katz, which fully employ tools of algebraic geometry). In contrast, "elementary" presentations proceed by the philosophy we have encapsulated above as the method of moments. The idea at the heart of this approach is, roughly speaking, that the character sum S modulo a prime p may be embedded inside a family of character sums S_ν over $\mathbb{F}_{p^\nu}$ for all $\nu \geq 1$. On the one hand, for each ν the sum S_ν may be interpreted as a sum of roots associated to a rational function (the zeta function), which are themselves ν-powers of the roots associated to the original sum S. On the other hand, for each ν the sum S_ν may be interpreted in terms of counting points on a variety over $\mathbb{F}_{p^\nu}$ with a prescribed error-term. If for all sufficiently large ν the error term is $O(B^\nu)$ then one learns (by uniting the two interpretations) that the ν-th moment of the roots corresponding to S is $O(B^\nu)$ for all sufficiently large ν, and hence by the general philosophy of the method of moments, one can deduce that each root associated to S is $O(B)$. One finally deduces that $|S|$ itself is $O(B)$. (In this application, one is of course thinking of $B = p^{n/2}$ for the n-dimensional case.) There are two general presentations of this style of argument: first, Stepanov's method, exposed in [Sch76] and [IK04, §11.7]; see in particular Lemma 11.22 of [IK04] for the role of arbitrarily high moments in closing this argument. Second, there is the "method of moments" work of Hooley [Hoo82], whose consequences are summarized in [Bro15, Lemma 3.5], and which also has a very clear, explicitly computed application in the work of [BB09, Lemma 33].

3. Relation to the Cohen-Lenstra-Martinet moment predictions

We start from the perspective of the heuristics of Cohen-Lenstra [CL84] and Cohen-Martinet [CM90]. We will show in this section that a collection of moment asymptotics predicted by these heuristics, and indeed even a collection of much weaker upper bounds, implies Conjecture 1.1. We will furthermore note that another well-known conjecture formulated under the heuristics (the density statement) does not suffice alone to imply Conjecture 1.1.

The Cohen-Lenstra and Cohen-Martinet heuristics predict, for "good" primes ℓ, the distribution of class groups and their ℓ-torsion subgroups as K varies over a family of number fields of fixed degree. For example, the Cohen-Lenstra heuristics for the family $\mathscr{F}_2^-(X)$ of imaginary quadratic fields imply a moment asymptotic for each odd prime ℓ: namely that for every positive integer $k \geq 1$,

$$\sum_{K \in \mathscr{F}_2^-(X)} |\mathrm{Cl}_K[\ell]|^k \sim c_{\ell,k}^- \sum_{K \in \mathscr{F}_2^-(X)} 1, \tag{3.1}$$

where $c_{\ell,k}^-$ is the number of $\mathbb{Z}/\ell\mathbb{Z}$ subspaces of the vector space $(\mathbb{Z}/\ell\mathbb{Z})^k$ (see, e.g. [Woo15, Lemma 3.2]). For real quadratic fields, the Cohen-Lenstra heuristics imply similar moment asymptotics with different values of the constant, say $c_{\ell,k}^+$.

Given a permutation group G, for primes $\ell \nmid |G|$, the Cohen-Martinet heuristics imply moment asymptotics analogous to (3.1) for the family of fields with Galois group G and fixed signature, again with different values of the constant playing the role of $c_{\ell,k}^-$. (This follows from [WW19, Theorem 6.1], after expressing $|\mathrm{Cl}_K[\ell]|^k$ as $|\operatorname{Hom}(\mathrm{Cl}_K, (\mathbb{Z}/\ell\mathbb{Z})^k)|$ and then noting each homomorphism ϕ corresponds to a Γ module homorphism from Cl_K to $\operatorname{Ind}_1^\Gamma \operatorname{im}\phi$.)

It is worth mentioning the few known asymptotic results that verify cases of (3.1) or its analogues for higher degree fields. First, Davenport and Heilbronn [DH71] proved (3.1) for $\ell = 3$ and $k = 1$, with $c_{3,1}^- = 2$, as well as its real analogue with $c_{3,1}^+ = 4/3$. Second, Bhargava [Bha05] proved the conjectured asymptotic for cubic fields and 2-torsion, for the first moment ($k = 1$). Although not originally included in the Cohen-Lenstra heuristics, Fouvry and Klüners [FK06] proved an analogous result related to 4-torsion when K is quadratic, for all $k \geq 1$ (see Remark 3.3).

There is also an increasing body of recent work toward related asymptotics, though not exactly of the kind mentioned above. For example, recent work of Klys [Kly16] gives certain asymptotic results

on 3-torsion in cyclic cubic fields; see also the recent work on 16-rank in quadratic fields, e.g. Milovic [Mil17], Koymans and Milovic [KM16, KM17], or work of Bhargava and Varma [BV15, BV16] elaborating on [DH71, Bha05]. See also results of Klys [Kly16] on moments of p-torsion in cyclic degree p fields (conditional on GRH for $p \geq 5$) and of Smith [Smi16, Smi17] on the distribution of the 2^k-class groups in imaginary quadratic fields (recently extended by Koymans and Pagano [KP18] to ℓ^k-class groups of degree ℓ cyclic fields).

3.1. **Relation to Theorem 1.2.** We will now show that if a moment asymptotic such as (3.1) holds, as long as we assume information is known for *arbitrarily high* moments, we can deduce pointwise upper bounds on $|\mathrm{Cl}_K[\ell]|$ for each field K, with no exceptions.

In fact our assumption in Theorem 1.2 is weaker than (3.1), taking the form of an upper bound: we assume the upper bound (1.4) holds for one fixed α and all integers $k \geq 1$. For any $\alpha \geq 1$ the case of the inequality (1.4) is implied by (3.1) or its appropriate analogue, and for all choices of $\alpha \geq 1$ the inequality is weaker than the conjectured Cohen-Lenstra-Martinet asymptotic for $\ell \nmid |G|$. For the cases of ℓ dividing $|G|$, for higher degree fields, in analogy with the quadratic case, we might also guess that for a given permutation group G, (1.4) holds for ℓ dividing $|G|$ with $\alpha = 1 + \varepsilon$, and that a proof of this in the case of ℓ dividing $|G|$ is probably much more accessible than for good primes (see e.g. [Kly16]).

3.2. **Proof of Theorem 1.2.** Recall the setting of Theorem 1.2, which we introduced in §1.1. We will use the fact that there exist constants C, B (not necessarily sharp) such that

$$|\mathscr{F}(X)| \leq CX^B \tag{3.2}$$

for all $X \geq 1$. In fact this is known for each $n \geq 2$ if we consider the set of all degree n extensions of $\mathbb{Q}$. By Malle's conjectures [Mal02, Mal04], it is expected that $B = 1$ should suffice, and this is known for $n = 2$ by a classical computation, $n = 3$ by [DH71], $n = 4$ by [CDyDO02], [Bha05], and $n = 5$ by [Bha10]. For each $n \geq 6$ it is known that for the set of all degree n extensions of $\mathbb{Q}$,

$$|\mathscr{F}(X)| \leq a_n X^{c_0(\log n)^2}, \tag{3.3}$$

for a constant a_n depending only on n and a certain absolute constant c_0 ($c_0 = 1.564$ suffices). This is due to Lemke Oliver and Thorne [OT20], and represents the current state of a progression of works with ever-improving bounds: Couveignes [Cou20], Ellenberg and Venkatesh [EV06], Schmidt [Sch95]. Of course for certain more restrictive collections of degree n fields, we would expect the sharp value for B to be even smaller, but we do not require this specificity here. We will note the dependence on B in the initial steps of our argument below, but later use that B can be taken to depend only on the degree n.

We now prove Theorem 1.2. Let $n, \ell \geq 1$ be fixed, and assume that for a fixed $\alpha \geq 1$, (1.4) is known for every integer $k \geq 1$. Let $\Delta > 0$ be given, and then let $\mathscr{B}_{\mathscr{F}}(X, \Delta)$ denote the "bad" set of fields $K \in \mathscr{F}(X)$ such that $|\mathrm{Cl}_K[\ell]| > D_K^{\Delta}$. Then (1.4) implies that for every $X \geq 1$,

$$\begin{aligned}(X/2)^{\Delta k} \cdot |\mathscr{B}_{\mathscr{F}}(X,\Delta) \cap \{K \in \mathscr{F}(X) : X/2 < D_K \leq X\}| &< \sum_{\substack{K \in \mathscr{B}_{\mathscr{F}}(X,\Delta) \\ X/2 < D_K \leq X}} |\mathrm{Cl}_K[\ell]|^k \\ &\leq \sum_{\substack{K \in \mathscr{F}(X) \\ X/2 < D_K \leq X}} |\mathrm{Cl}_K[\ell]|^k \ll_{n,\ell,k,\alpha} |\mathscr{F}(X)|^{\alpha}.\end{aligned}$$

We deduce that for each $k \geq 1$, for every $X \geq 1$,

$$|\{K \in \mathscr{F}(X) : X/2 < D_K \leq X, |\mathrm{Cl}_K[\ell]| > D_K^{\Delta}\}| \ll_{n,\ell,k,\alpha,\Delta} X^{B\alpha - \Delta k}.$$

Applying this for some k sufficiently large that $B\alpha - \Delta k < 0$, we conclude that for all X sufficiently large (determined by n, ℓ, α, Δ),

$$|\{K \in \mathscr{F}(X) : X/2 < D_K \leq X, |\mathrm{Cl}_K[\ell]| > D_K^{\Delta}\}| < 1$$

so that the set must be empty. Consequently, given any $\Delta > 0$, there exists $X_0 = X_0(n, \ell, \alpha, \Delta)$ such that for all fields $K \in \mathscr{F}$ with $D_K \geq X_0$, we have $|\mathrm{Cl}_K[\ell]| \leq D_K^{\Delta}$. Of course, for all fields $K \in \mathscr{F}$ with $D_K < X_0$ we can apply the trivial bound

$$|\mathrm{Cl}_K[\ell]| \leq |\mathrm{Cl}_K| \ll_{\varepsilon} D_K^{1/2+\varepsilon} \ll_{\varepsilon} X_0^{1/2+\varepsilon},$$

for all $\varepsilon > 0$. Upon fixing $\varepsilon = 1/2$, say, this is a bounded constant (depending on n, ℓ, α, Δ). Combining these two results, we conclude that for the given Δ, for all $K \in \mathscr{F}$,

$$|\mathrm{Cl}_K[\ell]| \ll_{n,\ell,\alpha,\Delta} D_K^{\Delta}.$$

Since $\Delta > 0$ may be taken arbitrarily small in this argument, this concludes the proof of Theorem 1.2.

3.3. **Further remarks.**

Remark 3.1 (Equivalent formulations). This proof may clearly be adapted to functions $V : \mathscr{F} \to \mathbb{R}_{>0}$ other than $V(K) = |\mathrm{Cl}_K[\ell]|$, or in settings where $\mathscr{F}$ is ordered by invariants other than the discriminant. Also, note that in the hypothesis of Theorem 1.2, we have not ruled out the possibility that α is large. In fact, the statement that (1.4) holds (for all $X \geq 1$) for a particular fixed α and an infinite sequence of arbitrarily large k is equivalent to the statement that (1.4) holds (for all $X \geq 1$) for $\alpha = 1 + \varepsilon_0$ with any $\varepsilon_0 > 0$, and all real $k \geq 1$. We provide details on this in Remark 8.1 in the appendix in §8.1.

Remark 3.2 (Density zero exceptions). If one only knows (1.4) for one value of k, one can still conclude that an upper bound holds for $|\mathrm{Cl}_K[\ell]|$, except for certain possible exceptional fields. To be precise, let n, ℓ be fixed. Suppose that for a single fixed value $k \geq 1$, (1.4) is known for every $\alpha > 1$. Then for each $\Delta > 0$, one can deduce that $|\mathrm{Cl}_K[\ell]| \leq D_K^{\Delta}$ holds for every field $K \in \mathscr{F}$ except for at most a subset of density zero. To see this, we argue as above to show that for any $\Delta > 0$ of our choice, for all $\alpha > 1$, and for that fixed k,

$$|\mathscr{B}_{\mathscr{F}}(X, \Delta) \cap \{K \in \mathscr{F}(X) : X/2 < D_K \leq X\}| \ll_{n,\ell,k,\alpha,\Delta} |\mathscr{F}(X)|^{\alpha} X^{-\Delta k} \quad \text{for all } X \geq 1. \tag{3.4}$$

To remove the dyadic restriction on the discriminant, we sum over at most $\log_2 X + 1 \leq 2 \log_2 X$ dyadic ranges. For the purposes of this discussion we suppose that $|\mathscr{F}(X)|$ can be expressed as a power of X, and then we break the discussion into two cases: when $|\mathscr{F}(X)|^{\alpha} X^{-\Delta k}$ is expressed as a power of X, the exponent is (i) non-negative; (ii) negative. In case (i), after summing over the dyadic ranges we can conclude that $|\mathscr{B}_{\mathscr{F}}(X, \Delta)| \ll_{n,\ell,k,\alpha,\Delta} |\mathscr{F}(X)|^{\alpha} X^{-\Delta k} \log X$ for all $X \geq 1$. Consequently there exists $\alpha > 1$ sufficiently close to 1 that $|\mathscr{B}_{\mathscr{F}}(X, \Delta)|/|\mathscr{F}(X)| \ll |\mathscr{F}(X)|^{\alpha-1} X^{-\Delta k} \log X$ has a negative exponent on the right-hand side. Then $\mathscr{B}_{\mathscr{F}}(X, \Delta)$ has density zero in $\mathscr{F}$. In case (ii), then there exists a threshold $X_0 = X_0(n, \ell, k, \alpha, \Delta)$ such that for $X \geq X_0$, the set on the left-hand side of (3.4) is empty. Then there are only finitely many fields in $\mathscr{F}$ with $|\mathrm{Cl}_K[\ell]| > D_K^{\Delta}$, so that certainly these fields are density zero in $\mathscr{F}$.

Remark 3.3 (Illustrative case). We note a case in which moment bounds sufficient for the hypotheses of Theorem 1.2 have been proved: Fouvry and Klüners have proved that for $\mathscr{F}_2^{\pm}$ defined to be the family of real (respectively imaginary) quadratic fields with fundamental discriminants, for every integer $k \geq 1$,

$$\sum_{K \in \mathscr{F}_2^{\pm}(X)} 2^{k \, \mathrm{rk}_2(2\mathrm{Cl}_K)} \sim C_k^{\pm} |\mathscr{F}_2^{\pm}(X)| \tag{3.5}$$

as $X \to \infty$, for a specific constant $C_k^{\pm}$. (We use additive notation for the finite abelian group Cl_K, so that we write $2\mathrm{Cl}_K$ where Fouvry and Klüners write Cl_K^2; see [FK06, Conj. 2] and Theorems 6–11 therein.) By Theorem 1.2, (3.5) shows that $|(2\mathrm{Cl}_K)[2]| \ll_{\varepsilon} D_K^{\varepsilon}$ for all $\varepsilon > 0$. From this it follows that $|\mathrm{Cl}_K[4]| \ll_{\varepsilon} D_K^{\varepsilon}$ for all $K \in \mathscr{F}_2^{\pm}(X)$. (Indeed, $|\mathrm{Cl}_K[4]| = |2(\mathrm{Cl}_K[4])| \cdot |\mathrm{Cl}_K[4]/2(\mathrm{Cl}_K[4])|$.

Within the first factor, $|2(\mathrm{Cl}_K[4])| = |(2\mathrm{Cl}_K)[2]|$, which we have just bounded. Within the second factor, by the structure theorem of finite abelian groups, $|\mathrm{Cl}_K[4]/2(\mathrm{Cl}_K[4])| = |\mathrm{Cl}_K[2]|$, which is bounded by (1.3).) Although this is an illustration of how Theorem 1.2 can be applied, the result for 4-torsion holds trivially for group-theoretic reasons once (1.3) is known; see §7.1.

Remark 3.4 (Density statements). It is worth pointing out that aside from moment asymptotics, there is another type of asymptotic, a density statement, which also follows from the Cohen-Lenstra and Cohen-Martinet heuristics. For example, the density statement conjectured by the Cohen-Lenstra heuristics for imaginary quadratic fields is that for every odd prime p and finite abelian p-group G,

$$\lim_{X\to\infty} \frac{\sum_{K\in\mathscr{F}_2^-(X)} 1}{\sum_{K\in\mathscr{F}_2^-(X)} 1} = \frac{\prod_{i\geq 1}(1-p^{-i})}{|\operatorname{Aut} G|}, \tag{3.6}$$

where for any abelian group A we denote by A_p the Sylow p-subgroup of A. (For $p = 2$, Smith [Smi17] has proved that as K varies over all imaginary quadratic fields, the distribution of $(2\mathrm{Cl}_K)_2$ obeys the heuristic expectation (3.6) for every finite abelian 2-group G; see also the extension [KP18].) In fact the Cohen-Lenstra and Cohen-Martinet conjectures make claims about averages over class groups of any "reasonable" real-valued function of groups, where reasonable is not precisely defined, but is meant to include both characteristic functions as in (3.6), and functions like $G \mapsto |G[p]|^k$ that are averaged in our moments, e.g. (3.1).

We have shown that moments such as (3.1) that are predicted by the Cohen-Lenstra-Martinet heuristics imply Conjecture 1.1. In contrast, it is easy to see that density statements such as (3.6) do not formally imply Conjecture 1.1, unless some additional information is used as input. For suppose we had a function C from number fields to finite abelian groups such that for every odd prime p and finite abelian p-group G,

$$\lim_{X\to\infty} \frac{\sum_{\substack{K\in\mathscr{F}(X)\\ C(K)_p\simeq G}} 1}{\sum_{K\in\mathscr{F}(X)} 1} = \frac{\prod_{i\geq 1}(1-p^{-i})}{|\operatorname{Aut} G|}.$$

Then such averages would also hold if we replaced C by any function C' that only differs on a set of number fields that is of density zero in $\mathscr{F}(X)$. In particular, while the densities above still held for C', we could impose $|C'(K)[\ell]| > D_K$ for infinitely many K.

4. Relation to the discriminant multiplicity conjecture

Our next goal is to make precise the relationship of the discriminant multiplicity conjecture (Conjecture 1.4) to the ℓ-torsion conjecture (Conjecture 1.1).

4.1. **Context of the discriminant multiplicity conjecture.** Recall the definition of Property $\mathbf{D}_n(\varpi)$ (Property 1.5). The discriminant multiplicity conjecture posits that $\mathbf{D}_n(0)$ should hold for each $n \geq 2$. (In the other direction, Ellenberg and Venkatesh [EV05, p. 164] have noted that genus theory shows that the number of degree extensions $K/\mathbb{Q}$ with a fixed Galois group and fixed discriminant D can be as large as $D^{c/\log\log D}$.)

For $n = 3, 4, 5$ the best-known results are as follows. For cubic fields, $\mathbf{D}_3(1/3)$ is known, by [EV07, Cor. 3.7]. (That corollary in fact states that $|\mathrm{Cl}_K[3]| \ll D_K^{1/3+\varepsilon}$ for all quadratic fields $K/\mathbb{Q}$; this implies that $\mathbf{D}_3(1/3)$ by work of Hasse [Has30]; see also Remark 4.1.)

For quartic fields, $\mathbf{D}_4(1/2)$ holds. To see this, let us momentarily denote by $\mathbf{D}_4(G;\varpi)$ the veracity of Property 1.5 restricted to those fields of degree 4 and Galois group of the Galois closure isomorphic to G. Then $\mathbf{D}_4(S_4;1/2)$ is due to [Klü06b], $\mathbf{D}_4(A_4;0.2784....)$ is due to [PTBW20, Prop. 2.5], $\mathbf{D}_4(K_4;0)$ follows from class field theory, as does $\mathbf{D}_4(C_4;0)$ (see e.g. techniques in [Wri89] also

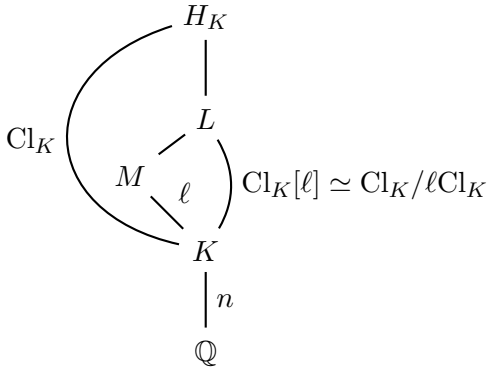
illustrated in [PTBW20, Prop. 2.1]), and finally $\mathbf{D}_4(D_4;0)$ holds; we include a brief proof of this last fact in an appendix in §8.2.

For quintic fields, $\mathbf{D}_5(199/200)$ is deduced in [PTBW20, Prop. 2.4] as an immediate consequence of a count for S_5-fields with a power-saving error term (in particular the version in [EPW17, Thm. 2.4], which builds on [ST14, Thm. 1] and [Bha10, Thm. 1]).

At present, for $n \geq 6$, nothing better than the bound trivially available from (3.3) is known for $\mathbf{D}_n(\varpi)$. For $n \geq 6$, if one further restricts to counting only those degree n fields K with $D_K = D$ and with Galois group of the Galois closure $\simeq G$ for a certain fixed transitive subgroup $G \subseteq S_n$, then in some cases better results are known; see [PTBW20, §2].

Conjecture 1.4 has been known for some time to be closely connected to questions about ℓ-torsion in class groups, for primes ℓ. For example, Duke [Duk95] showed that quartic fields of fixed discriminant $-q$ (q prime) can be explicitly classified by odd octahedral Galois representations of conductor q, and the number of such fields can be expressed as in [Hei71] as an appropriate average of the number of 2-torsion elements in the class groups of cubic number fields of discriminant $-q$. More generally, in [EV05, p. 164], Ellenberg and Venkatesh stated that Conjecture 1.4 implies Conjecture 1.1 in the case of ℓ prime; these two ideas are also implicitly linked in the exposition of [Duk98].

4.2. **Proof of Theorem 1.7.** Let $K/\mathbb{Q}$ be a degree n extension with Hilbert class field H_K, so that $\mathrm{Gal}(H_K/K) \simeq \mathrm{Cl}_K$. Let ℓ be a fixed prime and write Cl_K additively, so that we have $\mathrm{Cl}_K[\ell] \simeq \mathrm{Cl}_K/\ell\mathrm{Cl}_K$ and a corresponding extension $K \subset L \subset H_K$ given by the fixed field $L = H_K^{\ell\mathrm{Cl}_K}$.



Suppose $|\mathrm{Cl}_K[\ell]| = \ell^r$. Consider for each surjection $\phi : \mathrm{Cl}_K[\ell] \to \mathbb{Z}/\ell\mathbb{Z}$ the corresponding extension $K \subset M \subset L$ with M/K of degree ℓ. There are $N(\ell, r) := (\ell^r - 1)/(\ell - 1)$ such extensions M/K, each of which is degree $n\ell$ over $\mathbb{Q}$ and naturally satisfies

$$\mathrm{Disc}\, M/\mathbb{Q} = (\mathrm{Nm}_{K/\mathbb{Q}}\mathrm{Disc}\, M/K) \cdot (\mathrm{Disc}\, K/\mathbb{Q})^\ell.$$

But since M/K lives inside the unramified extension H_K/K, the first factor on the right-hand side is 1, so that $D_M = D_K^\ell$. Now supposing we have assumed Property $\mathbf{D}_{n\ell}(\Delta)$ holds, we know that for every $\varepsilon > 0$ there exists a constant $C_{n\ell,\varepsilon}$ such that the number of degree $n\ell$ extensions of $\mathbb{Q}$ with discriminant precisely D_K^ℓ is $\leq C_{n\ell,\varepsilon}(D_K^\ell)^{\Delta+\varepsilon}$. Then we can conclude that

$$|\mathrm{Cl}_K[\ell]| = (\ell - 1)N(\ell, r) + 1 \leq (\ell - 1)C_{n\ell,\varepsilon}(D_K^\ell)^{\Delta+\varepsilon} + 1.$$

Thus for this fixed pair n, ℓ, Property $\mathbf{C}_{n,\ell}(\Delta')$ holds with $\Delta' = \ell\Delta$, proving our claim.

Remark 4.1. For ℓ an odd prime, there is a partial converse to Theorem 1.7. Let $\mathbf{D}_\ell(D_\ell;\Delta)$ denote the veracity of Property 1.5 for those degree ℓ fields with Galois group of the Galois closure isomorphic to $D_\ell \subset S_\ell$, the dihedral group of 2ℓ elements (symmetries of a regular ℓ-gon). Then for every odd prime ℓ, for every $\Delta > 0$, $\mathbf{D}_\ell(D_\ell;\Delta)$ holds if $\mathbf{C}_{2,\ell}(\Delta)$ holds. For $\ell = 3$ this is the conclusion of Hasse [Has30], and in particular shows that the upper bound for 3-torsion in

quadratic fields controls the number of cubic extensions of $\mathbb{Q}$. In general, for this deduction see Klüners [Klü06a, Lemma 2.3], which applies to degree ℓ D_ℓ-extensions of any fixed number field. Recently [CT16, Thm. 1.1] and [FW18, Cor. 1.6] have used this relation to capitalize on the recent improvements of [EPW17, PTBW20] on ℓ-torsion in quadratic fields in order to bound from above the number of degree ℓ D_ℓ-fields of bounded discriminant.

5. Relation to a generalized Malle conjecture

In this section we recall the formal statement of a generalized Malle conjecture as in Ellenberg and Venkatesh [EV05], and state how via the method of moments it implies the discriminant multiplicity conjecture (Conjecture 1.4). Counting number fields of bounded discriminant is an average version of the pointwise bound conjectured in the discriminant multiplicity conjecture. Of course, if $\mathbf{D}_n(\varpi)$ is known then we may immediately deduce an upper bound for the average. At first glance, it is surprising that an implication in the other direction also exists; this would seem to contradict the philosophy of §2, which relied on bounding arbitrarily high moments, and not just an average, in order to deduce that a pointwise bound holds at *all* points in a discrete space. The strategy of [EV05] is that by converting k-th moments into averages of k-products, one can deduce pointwise bounds only under assumptions on averages.

5.1. **Context of the Malle conjecture.** Recall that it is conjectured that for every $n \geq 2$, if we define $Z_n(\mathbb{Q}; X)$ to be the set of fields $K/\mathbb{Q}$ of degree n (within a fixed closure $\overline{\mathbb{Q}}$), then

$$|Z_n(\mathbb{Q}; X)| \sim c_n X$$

for a particular constant c_n (this is recorded for example as a "folk conjecture" in [EV05, Conj. 1.1]). This is easy to show for $n = 2$ (in which case one is effectively counting square-free numbers, see e.g. the appendix in [EPW17]). It has been proved in deep work of Davenport-Heilbronn [DH71] for $n = 3$; Cohen, Diaz y Diaz, and Olivier [CDyDO02] and Bhargava [Bha05] for $n = 4$ (the D_4 and non-D_4 cases, respectively); and Bhargava [Bha10] for $n = 5$. For any degree $n \geq 6$, such an asymptotic, or even an upper bound better than (3.3), is not known. (Lower bounds are also difficult; [BSW16] holds the current record $\gg X^{1/2+1/n}$.)

If we further define $Z_n(\mathbb{Q}; G, X)$ to restrict to those fields in $Z_n(\mathbb{Q}; X)$ with $\mathrm{Gal}(\tilde{K}/K) \simeq G$, then Malle [Mal02], [Mal04] has presented heuristics for upper and lower bounds

$$X^{a_G} \ll |Z_n(\mathbb{Q}; G, X)| \ll X^{a_G+\varepsilon} \tag{5.1}$$

as $X \to \infty$ with $a_G^{-1} = \min\{\mathrm{ind}(g) : g \in G, g \neq \mathrm{id}\}$. (Here the index of an element $g \in S_n$ is defined as $\mathrm{ind}(g) = n - j$ where j is the number of orbits of g acting on a set of n elements.) These ideas can also be adapted, at least heuristically according to the so-called Malle-Bhargava principle [Woo16, §10], to take into account finitely or infinitely many local conditions. Proving that this principle holds in appropriate situations is currently of central interest in number theory.

5.2. **Ellenberg and Venkatesh's moment argument.** We first recall, somewhat informally, how a standard method of moments argument would approach the discriminant multiplicity conjecture. Let us consider (2.2) and (2.4), adapted to the setting where for each $m \in \mathbb{N}$, $f(m)$ denotes the number of fields $K \in Z_n(\mathbb{Q}; G, X)$ such that $D_K = m$. Then after adapting (2.2) to this setting, as long as we can prove that for a sequence of arbitrarily large k we have for all $X \geq 1$ that

$$\sum_{m \leq X} |f(m)|^k \leq X^\beta \tag{5.2}$$

for some fixed β, we will be able to conclude similarly to (2.4) that for all $X \geq 1$,

$$|\{X/2 < m \leq X : f(m) > m^\alpha\}| \ll X^{\beta - \alpha k}.$$

Then no matter the fixed value of β, and for any $\alpha > 0$ arbitrarily small, by applying this with sufficiently large k, we could deduce that for sufficiently large X the set on the left-hand side is empty. From this we could deduce that there are at most $O_\alpha(m^\alpha)$ fields $K \in \mathbb{Z}_n(\mathbb{Q}; G, X)$ with $D_K = m$, as desired. A significant obstacle to this approach is that in general, bounding even the first moment

$$|Z_n(\mathbb{Q}; G, X)| = \sum_{m \leq X} f(m)$$

remains completely out of reach at present (e.g. for $G \simeq S_n$ with $n \geq 6$), so that proving (5.2) may seem an unrealistic starting point.

This is where the ideas of Ellenberg and Venkatesh [EV05, Remark 4.7, Prop. 4.8] play a role, by cleverly translating the k-th moment estimate (5.2) into merely an average estimate, but for a different set of fields than $Z_n(\mathbb{Q}; G, X)$. Consequently, they show that a generalized version of the *average* upper bound in (5.1) would imply Conjecture 1.4.

The formulation of Ellenberg and Venkatesh is in terms of a field invariant called an f-discriminant, with f a rational class function (not to be confused with the function f used above). Precisely, if G is a fixed transitive subgroup of S_n, then a map f from the set $\mathcal{C}_G$ of nontrivial conjugacy classes of G to $\mathbb{Z}_{\geq 0}$ is said to be a rational class function if it is invariant under the $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-action and has the property that for $g \in G$, $f(g) = 0$ if and only if $g = \{\mathrm{id}\}$. One can define a corresponding f-discriminant of any Galois extension $L/\mathbb{Q}$ with Galois group G. Precisely, for any $p \nmid |G|$ (so that certainly p is at most tamely ramified in L and the inertia group at p is cyclic), let $c_p \in \mathcal{C}_G$ be the image of a generator of tame inertia at p. Then the f-discriminant $D_f(L)$ may be defined by

$$D_f(L) = \prod_{p \nmid |G|} p^{f(c_p)}.$$

In particular, if $f(g) = \mathrm{ind}(g)$, then $D_f(L)$ is the prime-to-$|G|$ part of the standard absolute discriminant of any of the n conjugate degree n extensions $L_0/\mathbb{Q}$ with Galois closure L, corresponding via Galois theory to the element stabilizers in G.

Ellenberg and Venkatesh ask whether for such f-discriminants, an appropriate analogue to (5.1) holds. Given G, Ellenberg and Venkatesh define for a rational class function f the invariant

$$a_G(f) = \max_{c \in \mathcal{C}_G, c \neq \{\mathrm{id}\}} f(c)^{-1}.$$

They then propose [EV05, Ques. 4.3, Remark 4.7] that if $Z^{(f)}(\mathbb{Q}, G; X)$ counts Galois extensions of $\mathbb{Q}$ with Galois group G and f-discriminant at most X, then for every $\varepsilon > 0$,

$$|Z^{(f)}(\mathbb{Q}, G; X)| \ll_{G,f,\varepsilon} X^{a_G(f)+\varepsilon}. \tag{5.3}$$

(In fact they propose a more refined asymptotic, but this upper bound version suffices for our considerations.)

Under these definitions, we can see that (5.3) suffices to imply Conjecture 1.4, and hence Conjecture 1.1.

Theorem B ([EV05, Prop. 4.8])**.** *Given any group G, define on $G^k = G \times G \times \cdots \times G$ the rational class function that is identically 1 away from the identity, that is $f(c_1, c_2, \ldots, c_k) = 1$ for any choice of conjugacy classes $c_1, c_2, \ldots, c_k$ in $\mathcal{C}_G$ that are not all trivial. If a generalized Malle conjecture (5.3) holds for all G, for this choice of f, then Conjecture 1.4 holds.*

We merely sketch the ideas, directly following [EV05, Prop. 4.8]. Since isomorphism classes of degree n G-fields are in finite-to-one correspondence (that is, m_G-to-one, where m_G depends only on G) with isomorphism classes of degree $|G|$ Galois G-fields, to verify Conjecture 1.4 it suffices to prove that for any fixed Galois group G, the number of Galois extensions $K/\mathbb{Q}$ with discriminant

$D_K = D$ is $\ll_\varepsilon D^\varepsilon$ for every ε. The proof hinges on the fact that for f as chosen in the theorem statement, the f-discriminant of a field K is of the shape

$$\prod_{p|D_K,\, p\nmid |G|} p,$$

in which the power of the prime is always 1. Moreover, for the particular choice of f in Theorem B, we have $a_{G^k}(f) = 1$ for all k; this fact will play a key role.

For each fixed integer $D \geq 1$ let $m(D)$ denote the number of Galois fields $K_1, \dots, K_{m(D)}$, each with Galois group $\simeq G$ and each with $D_{K_i} = D$; our goal is to show that $m(D) \ll D^\varepsilon$. For any fixed $k \leq m(D)$ we can generate on the order of $\binom{m(D)}{k} \gg_k m(D)^k$ composite fields $L = K_{j_1} \cdots K_{j_k}$ each with Galois group $\simeq G^k$, and each with f-discriminant $D_f(L/\mathbb{Q}) \leq D$, with f as defined as above.

Now suppose that there exists some $\alpha > 0$ such that for a sequence of arbitrarily large D, we have $m(D) \geq D^\alpha$. Then for a sequence of arbitrarily large D we would obtain at least $\gg_k D^{\alpha k}$ elements in $Z^{(f)}(\mathbb{Q}, G^k; D)$. Comparing this to the conjectured statement (5.3) in the case of the group G^k, we see that we must have $\alpha k \leq a_{G^k}(f)$, but by construction, we have $a_{G^k}(f) = 1$ for all k. Since k could be taken arbitrarily large, we would conclude that α must be arbitrarily small, implying Property $\mathbf{D}_n(0)$ must hold.

6. Counting elliptic curves

6.1. **Context of counting elliptic curves.** Given a positive integer q, let $E(q)$ denote the number of isogeny classes of elliptic curves over $\mathbb{Q}$ with conductor q. Let us clarify the relation between $E(q)$ and the number $E'(q)$ of isomorphism classes of elliptic curves over $\mathbb{Q}$ with conductor q. As used in [DK00, p. 12], given any elliptic curve $E/\mathbb{Q}$, at most 8 $\mathbb{Q}$-isomorphism classes are $\mathbb{Q}$-isogenous to E; thus $E(q) \leq E'(q) \leq 8E(q)$ (see [Sil09, Ch.IX, Example 6.4] for this fact). Thus an analogous conjecture holds for $E'(q)$, and we may proceed by considering either $E(q)$ or $E'(q)$; in particular the statement of our results will hold for both $E(q)$ and $E'(q)$, up to a constant factor.

As we recalled in Conjecture 1.11, Brumer and Silverman [BS96, Conj. 3] conjectured that for every q, $E(q) \ll_\varepsilon q^\varepsilon$ for all $\varepsilon > 0$. Indeed, Brumer and Silverman make a more precise conjecture: that there exist absolute constants κ, κ' such that for any finite set S of primes, the number of elliptic curves with good reduction outside of S is at most $\kappa M_S^{\kappa'/\log\log M_S}$, where $M_S = \prod_{p\in S} p$.

Remark 6.1. Correspondingly, Brumer and Silverman's original formulation of Conjecture 1.1 is the more precise question of whether for every field K of degree n one has

$$\log_\ell |\mathrm{Cl}_K[\ell]| \ll_{n,\ell} \log D_K / \log\log D_K.$$

See also [BL17, Conj. 3.3.A, Conj. 3.3.B, Thm. 7.5] for other possible specific rates of growth, and their relevance for counting the number of conjugacy classes of non-uniform lattices in semisimple Lie groups.

Upper bounds for $E(q)$ follow from bounding 3-torsion in quadratic fields. Define for every $X \geq 1$ the quantity

$$H_3(X) = \max_{K\in\mathscr{F}_2(X)} |\mathrm{Cl}_K[3]|,$$

in which $\mathscr{F}_2(X)$ denotes the family of quadratic fields (real or imaginary) with absolute discriminant at most X (in absolute value). Then Brumer and Silverman showed in [BS96, p. 99-100] that for every q,

$$E(q) \ll_\varepsilon q^\varepsilon H_3(1728q), \tag{6.1}$$

for every $\varepsilon > 0$. This was then sharpened in the proof of [HV06, Thm. 4.5] to

$$E(q) \ll_\varepsilon q^\varepsilon (H_3(1728q))^{2\beta/\log 3}, \tag{6.2}$$

for every $\varepsilon > 0$, where $\beta = 0.2782...$ and hence $2\beta/\log 3 = 0.5064...$, so that $\mathbf{C}_{2,3}(\Delta)$ implies $E(q) \ll_\varepsilon q^{(0.5064...)\Delta+\varepsilon}$. Originally the only bound available to Brumer and Silverman was the trivial bound $H_3(X) \ll X^{1/2+\varepsilon}$ via (1.1); now that $\mathbf{C}_{2,3}(1/3)$ is known by [EV07], it follows that

$$E(q) \ll_\varepsilon q^{2\beta/(3\log 3)+\varepsilon} \ll_\varepsilon q^{0.1688...+\varepsilon}$$

for all q. This is currently the best result known toward Conjecture 1.11.

In addition, Brumer and Silverman have proved that Conjecture 1.11 is true under the assumption of the Generalized Riemann Hypothesis and a weak version of the Birch-Swinnerton–Dyer Conjecture ([BS96, Thm. 4]). This is an interesting counterpart to the situation for bounding ℓ-torsion, for which it appears that GRH implies partial results, but stops far short of proving Conjecture 1.1 in full (see §7.4).

From Theorem 1.7 and (6.2), we make a simple observation, recalling the notation of Property 1.5.

Theorem 6.2. *If Property* $\mathbf{D}_6(\Delta)$ *holds, then*

$$E(q) \ll_\varepsilon q^{6\beta\Delta/\log 3+\varepsilon} \ll_\varepsilon q^{(1.5193...)\Delta+\varepsilon}$$

holds for all $q \geq 1$. *In particular, if it is known that for every* D *there are at most* $\ll_\varepsilon D^\varepsilon$ *sextic fields* $K/\mathbb{Q}$ *of discriminant* D, *then Conjecture 1.11 holds.*

6.2. **Average results.** One can also ask about an average or asymptotic conjecture related to Conjecture 1.11. Here we distinguish briefly between ordering elliptic curves by discriminant or by conductor. We first consider ordering by discriminant. Brumer and McGuinness [BM90, §5] have conjectured that the number of elliptic curves over $\mathbb{Q}$ whose minimal discriminant has absolute value less than X is asymptotically of size $cX^{5/6}$ for an explicitly predicted constant c (see also Watkins [Wat08, Conj. 2.1]). Each elliptic curve over $\mathbb{Q}$ in its Weierstrass form $y^2 = x^3 - 27c_4x - 54c_6$ with integers c_4, c_6 has discriminant Δ with

$$1728\Delta = c_4^3 - c_6^2. \tag{6.3}$$

(Here the discriminant Δ and c_4, c_6 are the standard invariants as in [Sil09, p. 42 Ch. III §1].) Very roughly speaking, we may think of the numerology as suggesting that for Δ fixed, most of the solutions to the equation (6.3) must come from $c_4 \ll \Delta^{1/3}$ and $c_6 \ll \Delta^{1/2}$. Counting elliptic curves ordered by discriminant, Fouvry, Nair and Tenenbaum [FNT92] proved the weaker upper bound $\ll X^{1+\varepsilon}$ for every $\varepsilon > 0$ and a lower bound of the order of magnitude $\gg X^{5/6}$. Any upper bound that is $o(X)$ would imply that the integers that are discriminants of elliptic curves are zero-density among all integers (a difficult open problem; see [You15]).

We next consider ordering by conductor q. Recall that the same primes divide the discriminant and the conductor, and the conductor is a divisor of the discriminant; the precise relationship between the two is related to Szpiro's conjecture [Szp90] (for example predicting that $\Delta \ll_\varepsilon q^{6+\varepsilon}$). Watkins [Wat08, Heuristic 4.1] adapted the argument of Brumer-McGuinness [BM90, §5] by analyzing heuristically how often a large power of a prime divides the discriminant, and proposed the following expectation:

Conjecture 6.3. *Let* $E'(q)$ *denote the number of isomorphism classes of elliptic curves of conductor* q. *There exists a constant* $c' > 0$ *such that*

$$\sum_{1\leq q\leq Q} E'(q) \sim c'Q^{5/6}$$

as $Q \to \infty$.

Duke and Kowalski [DK00, Prop. 1] have proved the weaker result that for $Q \geq 1$,

$$\sum_{1\leq q\leq Q} E(q) \ll_\varepsilon Q^{1+\varepsilon}. \tag{6.4}$$

See also [Won01a, Won01b], which further proves a lower bound $\gg_\varepsilon Q^{5/6-\varepsilon}$. All these works [DK00, FNT92, Won01a] argue by applying the asymptotic of Davenport and Heilbronn [DH71] for the average of $|\mathrm{Cl}_K[3]|$.

Remark 6.4. It is worth noting that an argument truly is needed to deduce (6.4) or our theorem on moments. Armed only with (6.1), the statement we can trivially deduce is that

$$\sum_{q\leq Q} E(q)^k \ll_\varepsilon Q^\varepsilon \sum_{q\leq Q} H_3(1728q)^k \ll_\varepsilon Q^{1+\varepsilon} H_3(1728Q)^k.$$

Since H_3 is a sup-norm, even one violation of a pointwise upper bound $|\mathrm{Cl}_K[3]| \leq D_K^\Delta$ would prevent us from concluding that the expression above is bounded by $\ll_\varepsilon Q^{1+k\Delta+\varepsilon}$. Similar considerations apply if we use (6.2).

6.3. **Moment results: proof of Theorem 1.12.** The proof of Theorem 1.12 builds on the arguments in [BS96], [DK00] and [HV06]; we begin by recalling the set-up of [BS96]. An elliptic curve of conductor q has good reduction outside of the primes dividing q, and so it suffices to count the number $E'_g(q)$ of elliptic curves with good reduction outside of the set $S_q = \{p|q, p \text{ prime}\} \cup \{2,3\}$, and for any fixed $\kappa \geq 1$ we write

$$\sum_{q\leq Q} E'(q)^\kappa \leq \sum_{q\leq Q} E'_g(q)^\kappa.$$

Let us fix q and the corresponding set S_q. Then any elliptic curve $E/\mathbb{Q}$ counted by $E'_g(q)$ has

$$1728\Delta_E = ad^6 \tag{6.5}$$

for integers a, d with a being 6-th power free, and such that $(c_6(E)/d^3, c_4(E)/d^2)$ is an S_q-integral point on the elliptic curve $\mathscr{E}_a$ defined by

$$\mathscr{E}_a : Y^3 = X^2 + a.$$

Let us denote by $A(q)$ the set of values a that arise in this manner as E ranges over $E'_g(q)$. Furthermore, given any S_q-integral point P on one such $\mathscr{E}_a$, let us denote by $C(P)$ the set of elliptic curves E counted by $E'_g(q)$ that give rise to the point P by following the procedure above. Brumer and Silverman's argument (see also [DK00]) shows that

$$\sum_{q\leq Q} E'_g(q)^\kappa \leq \sum_{q\leq Q} \Big(\sum_{a\in A(q)} \sum_{P} |C(P)| \Big)^\kappa,$$

in which P varies over the S_q-integral points on $\mathscr{E}_a$. For each P, Brumer and Silverman show that $|C(P)| \leq 2^{\omega(q)+5} \ll_\varepsilon q^\varepsilon$ for every $\varepsilon > 0$ (see [BS96, p. 100], as usual recalling that $\omega(q) \ll \log q/\log\log q$). Moreover, they show that for each such a, q, the number of S_q-integral points on $\mathscr{E}_a$ is $\ll_\varepsilon q^\varepsilon |\mathrm{Cl}_{K_a}[3]|$ for every $\varepsilon > 0$, where K_a is the quadratic field $\mathbb{Q}(\sqrt{-a})$. (See the line leading to [BS96, Eqn. (4)]; in their equation we note that $\#S = \#S_q \leq \omega(q) + 2$, so that for any constant c we have $c^{\#S} \ll q^\varepsilon$ for all $\varepsilon > 0$.) More recently, due to improved counts for integral points on a fixed Mordell elliptic curve (such as $\mathscr{E}_a$), we can improve this by [HV06, Thm. 4.5] to write that the number of S_q-integral points on $\mathscr{E}_a$ is

$$\ll_\varepsilon a^\varepsilon \exp\{2\log_3 |\mathrm{Cl}_{K_a}[3]|(\beta+\varepsilon)\} \ll_\varepsilon a^\varepsilon |\mathrm{Cl}_{K_a}[3]|^{2\beta/\log 3},$$

for any $\varepsilon > 0$, where $\beta = 0.2782...$ is a numerical constant defined explicitly as $\beta(0)$ in [HV06, Thm. 3.8]. We note that $2\beta/\log 3 = 0.5064...$ is strictly smaller than 1, so that this is advantageous.

Thus we may conclude that for every $\varepsilon > 0$,

$$\sum_{q\leq Q} E'_g(q)^\kappa \ll_\varepsilon Q^\varepsilon \sum_{q\leq Q} \Big(\sum_{a\in A(q)} a^\varepsilon |\mathrm{Cl}_{K_a}[3]|^{2\beta/\log 3} \Big)^\kappa.$$

Now, as in the argument in [DK00, §3], for each fixed q we will pass from a sum over $a \in A(q)$ to a sum over the set of (distinct) corresponding square-free kernels a'; we set $A'(q) = \{a' : a \in A(q)\}$. For any elliptic curve $E \in E'_g(q)$ associated to a, so that the discriminant relation (6.5) holds, we see from (6.5) and the fact that the corresponding square-free kernel a' is *square-free* that $|a'| \leq 1728q$ and $a'|\delta q$, where

$$\delta = 6/\gcd(6, q).$$

In general, denoting by $s(a) = a'$ the square-free kernel of an integer, we have for any a' that

$$\#\{a \leq X : s(a) = a'\} \ll_\varepsilon X^\varepsilon$$

for all $X \geq 1$ and for all $\varepsilon > 0$; see e.g. [HBP17, Lemma 4.4]. Since a is 6-th power free, we may apply this in our setting with $X \leq 2^5 3^5 q^5$, say. Thus there are at most $\ll_\varepsilon q^\varepsilon$ such values a corresponding to each $a' \in A'(q)$.

Thus we have shown that

$$\sum_{q \leq Q} E'_g(q)^\kappa \ll_\varepsilon Q^\varepsilon \sum_{q \leq Q} (\sum_{a'|\delta q} |\mathrm{Cl}_{K_{a'}}[3]|^{2\beta/\log 3})^\kappa,$$

where we have used the fact that $a^\varepsilon \ll q^\varepsilon$ for any $\varepsilon > 0$. On the right-hand side, we initially only consider square-free a', but we can enlarge the right-hand side by considering any such divisors a'. Applying Hölder's inequality to the right-hand side, this becomes

$$\sum_{q \leq Q} E'_g(q)^\kappa \ll_\varepsilon Q^\varepsilon \sum_{q \leq Q} (d(\delta q))^{\kappa - 1} \sum_{a'|\delta q} |\mathrm{Cl}_{K_{a'}}[3]|^{2\beta\kappa/\log 3} \ll_\varepsilon Q^{2\varepsilon} \sum_{q \leq Q} \sum_{a'|\delta q} |\mathrm{Cl}_{K_{a'}}[3]|^{2\beta\kappa/\log 3},$$

since for all integers m, the divisor function satisfies $d(m) \ll_\varepsilon m^\varepsilon$ for all $\varepsilon > 0$. Note that we may trivially enlarge the right-most quantity to

$$\ll_\varepsilon Q^{2\varepsilon} \sum_{q \leq 6Q} \sum_{a'|q} |\mathrm{Cl}_{K_{a'}}[3]|^{2\beta\kappa/\log 3}. \tag{6.6}$$

We apply the following elementary lemma.

Lemma 6.5. *For any sequence α_n of non-negative real numbers, define $A_t = \sum_{n \leq t} \alpha_n$. Suppose it is known that for all $t \geq 1$ we have $A_t \leq t^\Theta$, for some $\Theta \geq 1$. Then for all $Q \geq 1$,*

$$\sum_{q \leq Q} \sum_{n|q} \alpha_n \ll_\Theta Q^\Theta \mathscr{L}(\Theta)$$

in which $\mathscr{L}(\Theta) = \log Q$ if $\Theta = 1$ and $\mathscr{L}(\Theta) = 1$ if $\Theta > 1$.

Proof. First note that (using positivity)

$$\sum_{q \leq Q} \sum_{n|q} \alpha_n = \sum_{n \leq Q} \alpha_n \sum_{\substack{q \leq Q \\ n|q}} 1 = \sum_{n \leq Q} \alpha_n \lfloor Q/n \rfloor \leq Q \sum_{n \leq Q} \frac{\alpha_n}{n}.$$

Then by partial summation and the assumption on A_t we have

$$\sum_{n \leq Q} \frac{\alpha_n}{n} = A_Q Q^{-1} + \int_1^Q A_t t^{-2} dt \leq A_Q Q^{-1} + \int_1^Q t^{\Theta - 2} dt \ll_\Theta Q^{\Theta - 1} \mathscr{L}(\Theta),$$

which suffices to prove the lemma. □

We apply the lemma in (6.6) with $\alpha_n = |\mathrm{Cl}_{K_n}[3]|^{2\beta\kappa/\log 3}$ and $\Theta = \Theta_k$ the exponent assumed to hold in (1.6) for the value $k = 2\beta\kappa/\log 3$. Note that for any $k > 0$, $\Theta_k \geq 1$ since $|\mathscr{F}_2(X)| \gg X$ and $|\mathrm{Cl}_K[3]| \geq 1$ for all fields K. We may conclude from (6.6) that

$$\sum_{q \leq Q} E'_g(q)^\kappa \ll_{k,\varepsilon} Q^{2\varepsilon} Q^{\Theta_k} \mathscr{L}(\Theta) \ll_{k,\varepsilon} Q^{\Theta_k + 2\varepsilon}$$

for all $\varepsilon > 0$. This suffices to prove the theorem, upon noting that

$$\kappa = \frac{k \log 3}{2\beta} = k \cdot \gamma = k \cdot 1.9745...$$

Corollary 1.13 follows from Theorem 1.12 immediately, upon applying the results of [HBP17, Cor. 1.4]. To situate the strength of Corollary 1.13, we briefly observe that for any k such that $\gamma k \geq 1$, by combining (6.4) with the pointwise bound $E(q) \ll_\varepsilon q^{1/(3\gamma)+\varepsilon}$ in (1.5),

$$\sum_{q \leq Q} E(q)^{k\gamma} \leq \max_{q \leq Q} E(q)^{k\gamma - 1} \sum_{q \leq Q} E(q) \ll Q^{1+(k\gamma-1)/(3\gamma)+\varepsilon}. \tag{6.7}$$

For example, in the case $k = 2/\gamma \geq 1$ the exponent (ignoring ε) in (6.7) is $1 + 1/(3\gamma) = 1.1688...$ which is larger than the exponent $1+(2/\gamma-1)/3 = 1.0043...$ obtained by applying Theorem 1.12, and applying $|\mathrm{Cl}_K[3]| \ll D_K^{1/3+\varepsilon}$ to one factor and using Davenport-Heilbronn to bound the remaining average.

As another example, in the case $k = 2$, (6.7) gives the exponent $5/3 - 1/(3\gamma) = 1.4978...$ while Corollary 1.13 gives the exponent $23/18 = 1.2777....$

7. Known results toward Conjecture 1.1, and relation to GRH

We have focused thus far on heuristics to support Conjecture 1.1. To conclude, we provide an overview of recent progress toward the conjecture, focusing first on results that hold for *all* fields of a given degree. Then we focus on a relationship to the Generalized Riemann Hypothesis. We will use the notation $\mathbf{C}_{n,\ell}(\Delta)$ defined in Property 1.6.

7.1. **Composite ℓ.** First we remark that while we have stated Conjecture 1.1 only for primes ℓ, we expect it to hold for all integers $\ell \geq 1$. Indeed, let $K/\mathbb{Q}$ be a number field. Since Cl_K is a finite abelian group, it is isomorphic to

$$C_{\ell_1^{a_{1,1}}} \times \cdots \times C_{\ell_1^{a_{1,r_1}}} \times \cdots \times C_{\ell_s^{a_{s,1}}} \times \cdots \times C_{\ell_s^{a_{s,r_s}}},$$

for distinct primes $\ell_1, \dots, \ell_s$, and cyclic groups $C_{\ell^a} \simeq \mathbb{Z}/\ell^a\mathbb{Z}$. For any prime $\ell \,|\, |\mathrm{Cl}_K|$, letting r be the number of cyclic groups in this decomposition corresponding to ℓ, then $|\mathrm{Cl}_K[\ell]| = \ell^r$. (If $\ell \nmid |\mathrm{Cl}_K|$ then this remains true, with $r = 0$ and $|\mathrm{Cl}_K[\ell]| = 1$.)

Furthermore, by the above decomposition, $|\mathrm{Cl}_K[\ell]|$ is multiplicative as a function of ℓ, that is, for integers $(\ell_1, \ell_2) = 1$, $|\mathrm{Cl}_K[\ell_1\ell_2]| = |\mathrm{Cl}_K[\ell_1]| \cdot |\mathrm{Cl}_K[\ell_2]|$. Consequently if $(\ell_1, \ell_2) = 1$ and $\mathbf{C}_{n,\ell_1}(\varpi_1)$ and $\mathbf{C}_{n,\ell_2}(\varpi_2)$ are known then $\mathbf{C}_{n,\ell_1\ell_2}(\varpi_1+\varpi_2)$ holds. This is of course only interesting if $\varpi_1+\varpi_2 < 1/2$.

Finally, for any prime $\ell \,|\, |\mathrm{Cl}_K|$, we have $|\mathrm{Cl}_K[\ell^k]| \leq |\mathrm{Cl}_K[\ell]|^k$ for any integer $k \geq 1$. (If $\ell \nmid |\mathrm{Cl}_K|$ we can say this trivially holds, since both sides are 1.) Indeed, more generally, for any $k' \geq k \geq 1$, $|\mathrm{Cl}_K[\ell^{k'}]| \leq |\mathrm{Cl}_K[\ell^k]|^{k'/k}$. Thus if $\mathbf{C}_{n,\ell^k}(\varpi)$ is known for some $k \geq 1$, then $\mathbf{C}_{n,\ell^{k'}}(k'\varpi/k)$ holds for all $k' \geq k$. Again, this is only interesting if $k'\varpi/k < 1/2$. As a consequence, if for a given degree n, Conjecture 1.1 holds for a prime ℓ, then for all $k \geq 1$, $|\mathrm{Cl}_K[\ell^k]| \ll_{n,\ell,k,\varepsilon} D_K^\varepsilon$ for all $\varepsilon > 0$.

7.2. **Pointwise results, prime and prime power ℓ.** We remarked that Gauss proved $\mathbf{C}_{2,2}(0)$; from the discussion above, $\mathbf{C}_{2,2^k}(0)$ follows for all $k \geq 1$. For $n = 2$ and $\ell = 3$, initial progress occurred in [HV06, Pie05, Pie06] and the current record of $\mathbf{C}_{2,3}(1/3)$ is held by [EV07], which also proved $\mathbf{C}_{3,3}(1/3)$ and $\mathbf{C}_{4,3}(1/2-\delta)$ for a small $\delta > 0$ (Ellenberg, personal communication, computed one can take $\delta = 1/168$ if K is non-D_4). Bhargava et al. [BST+17] have established $\mathbf{C}_{n,2}(0.2784...)$ for $n = 3, 4$ and $\mathbf{C}_{n,2}(1/2 - 1/2n)$ where $n \geq 5$, the first results to hold for all fields of arbitrarily high degree.

7.3. **Pointwise results, composite ℓ.** By multiplicativity, combining Gauss's work with $\mathbf{C}_{2,3}(1/3)$ thus implies $\mathbf{C}_{2,2^k3}(1/3)$ for all $k \geq 1$.

7.4. **Results conditional on GRH.** Much recent progress has relied on finding sufficiently many "small" rational primes that split completely within K, based on an insight credited to K. Soundararajan and P. Michel, and codified in work of Helfgott and Venkatesh [HV06] and Ellenberg and Venkatesh [EV07], from which we quote the following:

Theorem C ([EV07, Lemma 2.3])**.** *If a field K of degree n simultaneously has the properties that*

(1) the minimum multiplicative Weil height of a generating element of K is large, say

$$\inf\{H_K(\alpha) : K = \mathbb{Q}(\alpha)\} \gg D_K^{\eta}, \tag{7.1}$$

(2) and $\gg M$ rational primes $p \leq D_K^{\eta/\ell}$ split completely in K,

then $|\mathrm{Cl}_K[\ell]| \ll_{n,\ell,\eta,\varepsilon} D_K^{1/2+\varepsilon} M^{-1}$, for every $\varepsilon > 0$.

The idea underlying this theorem can be most easily seen in the case of K an imaginary quadratic field of discriminant $-D$ and ℓ an odd prime. Letting $H = \mathrm{Cl}_K[\ell]$, we of course have

$$|H| = |\mathrm{Cl}_K|/[\mathrm{Cl}_K : H], \tag{7.2}$$

so that to show that $|H| = |\mathrm{Cl}_K[\ell]|$ is small it suffices to show that H has many cosets in Cl_K. Suppose that p_1, p_2 are distinct rational primes with $p \nmid 2D$ that split in K as $p_1 = \mathfrak{p}_1\mathfrak{p}_1^{\sigma}$, $p_2 = \mathfrak{p}_2\mathfrak{p}_2^{\sigma}$. If $\mathfrak{p}_1 H = \mathfrak{p}_2 H$ then it follows that $\mathfrak{p}_1\mathfrak{p}_2^{\sigma} \in H$, so that $(\mathfrak{p}_1\mathfrak{p}_2^{\sigma})^{\ell}$ is a principal ideal, say $((u+v\sqrt{-D})/2)$ for some rational integers u, v. Upon taking norms, this shows that

$$4(p_1p_2)^{\ell} = u^2 + Dv^2 \tag{7.3}$$

for some rational integers u, v. Now if we had specified to begin with that $p_1, p_2 < (1/4)D^{1/(2\ell)}$, such a relation could not hold (since it would impose $v = 0$ and $4(p_1p_2)^{\ell}$ cannot be a square of an integer), and we must have that $\mathfrak{p}_1$ and $\mathfrak{p}_2$ represent different cosets of H.

Now suppose we have M distinct rational primes $p_1, \dots, p_M < (1/4)D^{1/(2\ell)}$ with $p \nmid 2D$, each of which splits in K. By the above argument, the cosets $\mathfrak{p}_j H$ must all be distinct for $j = 1, \dots, M$, so that by (7.2), for every $\varepsilon > 0$,

$$|\mathrm{Cl}_K[\ell]| \leq |\mathrm{Cl}_K| M^{-1} \ll_{\varepsilon} D^{1/2+\varepsilon} M^{-1}.$$

The fully general case of considering all degree n fields requires much more sophisticated machinery; see [EV07].

How can we apply Theorem C in practice? Regarding the parameter η in (7.1), the lower bound $\eta \geq 1/(2(n-1))$ is known for all degree n fields (see [EV07, Lemma 2.2] or [Sil84, Thm. 1]). Although the Chebotarev density theorem shows that a positive proportion of primes split completely in K, it is notoriously difficult to find *small* split primes (i.e. to show $M \gg D_K^{\eta/\ell}/\log D_K^{\eta/\ell}$, which is expected to be true). In general this can only be done effectively within a certain field K by assuming GRH for the Dedekind zeta function associated to the Galois closure $\tilde{K}$, which implies an effective version of the Chebotarev density theorem, with a good explicit error term. Taken together, these considerations lead to Ellenberg and Venkatesh's result in [EV07] that $\mathbf{C}_{n,\ell}(1/2 - 1/(2\ell(n-1)))$ holds for all n, ℓ if GRH is assumed; we will call this GRH-quality savings. This opens several avenues for possible progress, such as: first, try to remove the dependence on GRH, and second, try to improve the benchmark of GRH-quality savings, possibly by showing that η can be taken to be larger in (7.1).

7.5. **Average and "almost all" results.** Recent work has been able to obtain savings on $|\mathrm{Cl}_K[\ell]|$ as strong as the GRH-quality savings described above (or stronger), but without assuming GRH, at the cost of showing that it holds only for almost all fields within a certain family. This includes remarks of Soundararajan [Sou00] on imaginary quadratic fields, and also recent work of Heath-Brown and the first author [HBP17]. The latter work even does better than GRH-quality savings, by proving for each prime $\ell \geq 5$ the unconditional bound $|\mathrm{Cl}_K[\ell]| \ll_{\ell,\varepsilon} D_K^{1/2-3/(2\ell+2)+\varepsilon}$ for all but

a possible density zero family of imaginary quadratic fields. This is achieved by nontrivial use of averaging over the discriminant.

For each degree $n \leq 5$, Ellenberg and the first and third authors [EPW17] developed a new combination of field-counting and sieving to show that GRH-quality savings on $|\mathrm{Cl}_K[\ell]|$ holds unconditionally for all but a possible density zero exceptional family of degree n extensions of $\mathbb{Q}$ (non-D_4 in the case $n = 4$). This work relied in particular on precise counts for fields with certain properties, such as a count for fields in which a fixed prime splits completely, where the count has a power-saving error term with explicit dependence on the prime. The strategy of [EPW17] has also been adapted in work of Frei and Widmer [FW17] to prove results on average over the special family of totally ramified cyclic extensions of any fixed number field. In general, for degrees $n \geq 6$, methods do not yet exist to count all number fields of degree n precisely enough for the strategies of [EPW17] to be carried out.

Another approach is to tackle the effective Chebotarev density theorem directly, without assuming GRH. This has been carried out by the authors in [PTBW20], which proves a new effective Chebotarev density theorem valid for almost all members of certain families of fields, and consequently proves GRH-quality savings hold for ℓ-torsion, for all $\ell \geq 1$, in almost all such fields. (The families of fields are too numerous to describe here, but include for example: totally ramified cyclic extensions of $\mathbb{Q}$, in which case the result is unconditional; degree n S_n-fields with squarefree discriminant, conditional for $n \geq 5$ on the strong Artin conjecture and certain counts for number fields; degree n A_n-fields, conditional on the strong Artin conjecture.) We note that this approach has recently been adapted by An [An18] to certain families of D_4-quartic fields. We note that an interesting contrast of [PTBW20] to [EPW17] is that while it also depends heavily on counting number fields, it only requires rough counts. That is, given a family of fields, at a key step in [PTBW20] it suffices that one can show that at most $\ll X^\beta$ fields in the family have any fixed discriminant X, and at least $\gg X^\alpha$ fields in the family have bounded discriminant $\leq X$, where $\alpha > \beta$. The fact that such rough counts suffice allows [PTBW20] to prove results for many new families of fields. See also Remark 1.9 on the surprising efficiency of this method.

As mentioned above, the bound $\eta \geq 1/(2(n-1))$ for the exponent η in the expression (7.1) in property (1) is uniformly true for all degree n extensions. For fields in which η can be taken to be larger, one could also increase the savings in the ℓ-torsion count by increasing M in property (2). This interesting strategy for improving the very notion of "GRH-quality savings" has been proposed by Ellenberg [Ell08] and taken up in recent work of Widmer [Wid18]. One underlying question is which families can admit an improved uniform lower bound for η, or an improved lower bound on η for almost all fields in the family. Very recently Frei and Widmer [FW18] have strengthened this approach, and their ideas combined with the methods of [EPW17, PTBW20] improve the average upper bounds in certain results of those papers. For example, they can improve the relevant exponent $1/2 - 1/(2\ell(n-1))$ in [EPW17] for $n \geq 5$ to $\approx 1/2 - 1/(\ell n)$; see [FW18, Thm. 1.2].

The above results apply to average or "almost all" results. But if we turn to pointwise results, is it even true that all fields of degree n should have η strictly larger than $1/(2(n-1))$? In some restricted families of degree n fields, no such improvement is possible; see [FW17]. Such considerations of enlarging η relate also to conjectured uniform upper bounds such as $\eta \leq 1/2$, suggested by Ruppert [Rup98] (and proved to hold for "almost all" fields in many families of fields, by work of the authors [PTBW20]). If true, this restriction on η would imply a limit for how far the "small split prime approach" can go in bounding ℓ-torsion (e.g. the best possible would be $\mathbf{C}_{n,\ell}(1/2 - 1/2\ell)$), without additional information or new ideas.

7.6. **Higher moments.** Theorem 1.2 makes a clear case for studying higher moments of ℓ-torsion over families of fields. The first paper to do so was by Heath-Brown and the first author [HBP17], which proved nontrivial upper bounds for k-th moments of ℓ-torsion in imaginary quadratic fields for odd primes ℓ. Very recent work of Frei and Widmer [FW18] now makes progress on upper bounds

for k-th moments in certain other settings, e.g. for ℓ-torsion in (imaginary and real) quadratic fields, or cyclic cubic fields, by simultaneously incorporating ideas for improving (on average) the lower bound in (7.1) into the settings of certain methods in [EPW17, PTBW20].

7.7. **Asymptotics.** For recent work on asymptotics, see §3.

8. Appendices

8.1. **Further remarks on moment bounds in a general setting.** In the context of Theorem 1.2 we argued explicitly using the function $V : K \mapsto |\mathrm{Cl}_K[\ell]|$ mapping fields (ordered by discriminant) to positive real numbers. This was to make the argument clear in a setting where there could be more than one field with a fixed discriminant, and where the underlying finite set $\mathscr{F}(X)$ changes as X grows. We include further general remarks here, which clarify whether one needs uniformity in the constants appearing in the moments, as we anticipate there will be further applications for this perspective.

Let $(\mathcal{M}, \mu)$ be a (σ-finite) measure space as before and consider the spaces $L^p(E)$ for $\mu(E) < \infty$ and $1 \le p \le \infty$. If $f \in L^q(E)$ then $f \in L^p(E)$ for all $1 \le p \le q$; indeed Hölder's inequality shows that

$$\|f\|_{L^p(E)} \le \mu(E)^{1/p-1/q}\|f\|_{L^q(E)}.$$

In particular, if $f \in L^\infty(E)$ then $f \in L^p(E)$ for all $1 \le p \le \infty$ with

$$\|f\|_{L^p(E)} \le \mu(E)^{1/p}\|f\|_{L^\infty(E)}. \tag{8.1}$$

We are interested in the converse direction. Let us make the strong assumption that there exist constants B, M such that for an infinite sequence of arbitrarily large $p \ge 1$ there exists a constant $c_p \le M$ such that for all $f \in L^p(E)$,

$$\|f\|_{L^p(E)} \le c_p\mu(E)^{B/p}. \tag{8.2}$$

Then we claim $f \in L^\infty(E)$ and moreover $\|f\|_{L^\infty(E)} \le M$.

Indeed, to show $f \in L^\infty(E)$ we must show there exists C such that $\mu(\{x : |f(x)| > C\}) = 0$. By Chebyshev's inequality (2.1) and our assumption (8.2), for every $\alpha > 0$,

$$\mu(\{x \in E : |f(x)| > \alpha\}) \le \frac{c_p^p}{\alpha^p}\mu(E)^B$$

for an infinite sequence of arbitrarily large $p \ge 1$. Let us choose α_0 sufficiently large such that $\alpha_0 > M$. Then given any $\varepsilon' > 0$, we may take p sufficiently large (depending on $M, \alpha_0, \mu(E), B$) that

$$\mu(\{x \in E : |f(x)| > \alpha_0\}) \le (M/\alpha_0)^p\mu(E)^B < \varepsilon'.$$

Thus we may conclude that $f \in L^\infty(E)$, with $\|f\|_{L^\infty(E)} \le M$. Consequently, (8.1) holds for all $1 \le p \le \infty$. In particular, this shows that if the apparently weaker estimate (8.2) holds for arbitrarily large p, this implies the apparently stronger estimate (8.1); that is to say, assuming that (8.2) holds for some B and an infinite sequence of arbitrarily large p is equivalent to assuming it holds for $B = 1$ and all $1 \le p \le \infty$.

These ideas apply to the case where $E = \{1, 2, 3, \dots, X\}$ and μ is counting measure, so that $\mu(E) = X$. They can also be adapted to the setting in which we are studying moments of a function V mapping fields (ordered by discriminant, for example) to positive real numbers. Let X be fixed. If it is known that there is some α and some $d_k \le M^k$ for a uniform constant M such that

$$\|V\|_{\ell^k(\mathscr{F}(X))}^k = \sum_{K \in \mathscr{F}(X)} V(K)^k \le d_k|\mathscr{F}(X)|^\alpha$$

holds for a sequence of arbitrarily large k, then this holds with $\alpha = 1$ for all $1 \le k \le \infty$, and indeed these two statements are equivalent to each other, and equivalent to the statement that $\|V\|_{\ell^\infty(\mathscr{F}(X))} \le M$.

Above, we made the assumption in (8.2) that there exists a uniform upper bound M such that $c_p \leq M$ for all p in the sequence of arbitrarily large p. (Without this assumption, the L^p norms can blow up and the claim $f \in L^\infty(E)$ need not hold.) However, the reader may note that in our use of (1.4) in Theorem 1.2 we made no such assumption of an upper bound on the constant $c_{n,\ell,k,\alpha}$ being uniform in k. Indeed, we do not require such an assumption because in Theorem 1.2 we prove not a uniform upper bound but merely show the growth with respect to X is slow as $X \to \infty$.

In the language of the model setting above, let $\mathcal{M} = \mathbb{Z}$, $E = (X/2, X] \cap \mathbb{Z}$ for $X \geq 1$ and let μ be counting measure. Let $f \in \ell^k((X/2, X] \cap \mathbb{Z})$ be given. Assume that there is some non-decreasing function $F \geq 1$ such that $|f(x)| \leq F(x)$ for all x. Also assume that there exists a constant B and an infinite sequence of arbitrarily large $k \geq 1$ with corresponding constants c_k such that for all $X \geq 1$,

$$\|f\|_{\ell^k((X/2,X]\cap\mathbb{Z})} \leq c_k X^{B/k}.$$

Then for any fixed $\delta > 0$, for every $X \geq 1$,

$$|\{x \in (X/2, X] \cap \mathbb{Z} : |f(x)| > x^\delta\}| \leq \frac{c_k^k}{(X/2)^{\delta k}} X^{B/k}.$$

Thus for any fixed δ, regardless of how big each constant c_k is, we may take a fixed k_0 sufficiently large (determined only by δ, B) that $B/k_0 - \delta k_0 < 0$, and consequently $c_{k_0}^{k_0} 2^{\delta k_0} X^{B/k_0 - \delta k_0} < 1$ for X sufficiently large with respect to δ, B, c_{k_0}, k_0. The set of constants c_k was determined by the function f, so we will denote this dependence as a dependence on f. This means that there exists some constant $X_0(f, \delta, B)$ such that for $X \geq X_0(f, \delta, B)$, the set $\{x \in (X/2, X] \cap \mathbb{Z} : |f(x)| > x^\delta\}$ is empty. In other words, for any $\delta > 0$, the only x such that $|f(x)| > x^\delta$ must have $x \leq X_0(f, \delta, B)$ and hence $|f(x)| \leq F(x) \leq F(X_0(f, \delta, B))$, which is a constant ≥ 1 depending only on f, δ, B. In total, we have proved that given any $\delta > 0$ there exists a constant $C_\delta = C_\delta(f, B)$ such that for all $x \in \mathbb{Z}_{\geq 1}$, $|f(x)| \leq C_\delta x^\delta$ (it suffices to take $C_\delta = F(X_0(f, \delta, B))$). This argument, adapted to families of fields, underlies Theorem 1.2.

Remark 8.1. The above discussion also verifies our earlier remark that the following statements are equivalent: (i) (1.4) holds (for all $X \geq 1$) for a particular fixed α and an infinite sequence of arbitrarily large k; (ii) (1.4) holds (for all $X \geq 1$) for all $\alpha = 1 + \varepsilon_0$ with $\varepsilon_0 > 0$ and all integers $k \geq 1$. Certainly the statement (ii) implies statement (i). In the other direction, the proof we gave for Theorem 1.2 (with k ranging over all integers $k \geq 1$) adapts easily to the case where k ranges over an infinite sequence of arbitrarily large k, and by this method we can prove that statement (i) implies that for every $\varepsilon_0 > 0$, $|\mathrm{Cl}_K[\ell]| \leq C_{\varepsilon_0} D_K^{\varepsilon_0}$ holds uniformly for $K \in \mathscr{F}$. Averaging this pointwise upper bound then implies statement (ii), completing the equivalence.

8.2. **Counting quartic D_4-fields with fixed discriminant.** We briefly prove that property $\mathbf{D}_4(D_4; 0)$ holds; that is, for every $D \geq 1$ and for every $\varepsilon > 0$ there are at most $\ll_\varepsilon D^\varepsilon$ quartic D_4-fields of discriminant D. We follow [Bai80, §4], except we fix a discriminant D instead of considering all discriminants up to X. Any quartic D_4-field K_4 is a quadratic extension of a quadratic field K_2. We have that K_4 corresponds to a quadratic ray class character for K_2 of conductor $\mathfrak{d}$ with finite part $\mathfrak{d}^* = \mathrm{Disc}\,(K_4/K_2)$. The form of such characters has been determined in [Bai80, Lemma 17], and each character is a product of a character on $(\mathcal{O}_{K_2}/\mathfrak{d}^*)^\times$, a character on the class group of K_2, and a character on signature (see [Bai80][Eqn. (4)]). There are at most 4 characters of signature, and $h_2(K_2)$ characters of the class group, which is $\ll_\varepsilon (\mathrm{Disc}\, K_2)^\varepsilon$ for every $\varepsilon > 0$ by genus theory (1.3). It follows from the proof of [Bai80, Lemma 18] that given any C, there are at most $O(2^{\omega(C)})$ ideals $\mathfrak{d}^*$ and possible characters on $(\mathcal{O}_{K_2}/\mathfrak{d}^*)^\times$ with $N_{K_2/\mathbb{Q}}(\mathfrak{d}^*) = C$. Note $\mathrm{Disc}\, K_4 = (\mathrm{Disc}\, K_2)^2 \cdot N_{K_2/\mathbb{Q}}(\mathrm{Disc}\,(K_4/K_2))$. Thus, given D, in order to count the quartic D_4-fields of discriminant D, we sum for each $d|D$, the number of quartic D_4-fields K_4 that contain

the quadratic field of discriminant d, and we find there are at most

$$O_\varepsilon\left(\sum_{d|D} d^\varepsilon 2^{\omega(D/d^2)}\right) = O_\varepsilon(D^\varepsilon)$$

quartic D_4-fields of discriminant D.

Acknowledgements

The authors thank J. Ellenberg, J. Wang, B. Alberts, and the referee for insightful comments that improved and clarified the content and exposition. Pierce has been partially supported by NSF CAREER grant DMS-1652173, a Sloan Research Fellowship, a Birman Fellowship, and as a von Neumann Fellow at the Institute for Advanced Study, by the Charles Simonyi Endowment and NSF Grant No. 1128155. Pierce thanks the Hausdorff Center for Mathematics for hospitality during visits as a Bonn Research Fellow. Turnage-Butterbaugh is partially supported by NSF DMS-1902193 and NSF DMS-1854398 FRG, and thanks the Mathematical Sciences Research Institute (NSF DMS-1440140) and the Max Planck Institute for Mathematics for support and hospitality during portions of this work. Wood was partially supported by an American Institute of Mathematics Five-Year Fellowship, a Packard Fellowship for Science and Engineering, a Sloan Research Fellowship, National Science Foundation grant DMS-1301690 and CAREER grant DMS-1652116, and a Vilas Early Career Investigator Award. Wood thanks Princeton University for its hospitality during Fall 2018.

References

[Alb18] B. Alberts. The weak form of malle's conjecture and solvable groups. *arXiv:1804.11318*, 2018.

[An18] C. An. ℓ-torsion in class groups of certain families of D_4-quartic fields. *arXiv:1808.02148*, 2018.

[Bai80] A. M. Baily. On the density of discriminants of quartic fields. *J. Reine Angew. Math.*, 315:190–210, 1980.

[BB09] E. Bombieri and J. Bourgain. On Kahane's ultraflat polynomials. *J. Eur. Math. Soc.*, 11:627–703, 2009.

[Bha05] M. Bhargava. The density of discriminants of quartic rings and fields. *Annals of Mathematics*, 162(2):1031–1063, September 2005.

[Bha10] M. Bhargava. The density of discriminants of quintic rings and fields. *Annals of Mathematics*, 172(3):1559–1591, October 2010.

[BL17] M. Belolipetsky and A. Lubotzky. Counting non-uniform lattices. *arXiv:1706.02180*, 2017.

[BM90] A. Brumer and O. McGuinness. The behavior of the Mordell-Weil group of elliptic curves. *Bull. Amer. Math. Soc. (N.S.)*, 23(2):375–382, 1990.

[Bro15] Timothy Daniel Browning. The polynomial sieve and equal sums of like polynomials. *Int. Math. Res. Not. IMRN*, (7):1987–2019, 2015.

[BS96] A. Brumer and J. H. Silverman. The number of elliptic curves over $\mathbb{Q}$ with conductor N. *Manuscripta Math.*, 91(1):95–102, 1996.

[BST+17] 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. *arXiv:1701.02458*, 2017.

[BSW16] M. Bhargava, A. Shankar, and X. Wang. Squarefree values of polynomial discriminants. *arXiv:1611.09806v2*, 2016.

[BV15] M. Bhargava and I. Varma. On the mean number of 2-torsion elements in the class groups, narrow class groups, and ideal groups of cubic orders and fields. *Duke Math. J.*, 164:1911–1933, 2015.

[BV16] M. Bhargava and I. Varma. The mean number of 3-torsion elements in the class groups and ideal groups of quadratic orders. *Proc. London Math. Soc.*, 112:235–266, 2016.

[CDyDO02] H. Cohen, F. Diaz y Diaz, and M. Olivier. Enumerating quartic dihedral extensions of $\mathbb{Q}$. *Compositio Math.*, 133(1):65–93, 2002.

[CL84] H. Cohen and H. W. Lenstra, Jr. Heuristics on class groups of number fields. In *Number theory, Noordwijkerhout 1983 (Noordwijkerhout, 1983)*, volume 1068 of *Lecture Notes in Math.*, pages 33–62. Springer, Berlin, 1984.

[CM90] H. Cohen and J. Martinet. Étude heuristique des groupes de classes des corps de nombres. *Journal für die Reine und Angewandte Mathematik*, 404:39–76, 1990.

[Cou20] Jean-Marc Couveignes. Enumerating number fields. *Ann. of Math. (2)*, 192(2):487–497, 2020.

[CT16] H. Cohen and F. Thorne. On D_ℓ-extensions of odd prime degree ℓ. *arXiv:1609.09153*, 2016.

[Del74] P. Deligne. La conjecture de Weil I. *Inst. Hautes Études Sc. Publ. Math. No.*, 43:273–307, 1974.

[Del80] P. Deligne. La conjecture de Weil II. *Inst. Hautes Études Sc. Publ. Math. No.*, 52:137–252, 1980.

[DH71] H. Davenport and H. Heilbronn. On the density of discriminants of cubic fields II. *Proc. Roy. Soc. Lond. A.*, 322:405–420, 1971.

[DK00] W. Duke and E. Kowalski. A problem of Linnik for elliptic curves and mean-value estimates for automorphic representations. *Invent. math.*, 139:1–39, 2000.

[Duk95] W. Duke. The dimension of the space of cusp forms of weight one. *Internat. Math. Res. Notices*, (2):99–109, 1995.

[Duk98] W. Duke. Bounds for arithmetic multiplicities. In *Proceedings of the International Congress of Mathematicians, Vol. II (Berlin, 1998)*, 1998.

[Ell08] Jordan S. Ellenberg. Points of low height on $\mathbb{P}^1$ over number fields and bounds for torsion in class groups. In *Computational arithmetic geometry*, volume 463 of *Contemp. Math.*, pages 45–48. Amer. Math. Soc., Providence, RI, 2008.

[EPW17] J. S. Ellenberg, L. B. Pierce, and M. M. Wood. On ℓ-torsion in class groups of number fields. *Algebra and Number Theory*, 11:1739–1778, 2017.

[EV05] J. S. Ellenberg and A. Venkatesh. Counting extensions of function fields with bounded discriminant and specified Galois group. In *Geometric Methods in Algebra and Number Theory*, volume 235 of *Progr. Math.*, pages 151–168. Birkhäuser Boston, Boston, MA, 2005.

[EV06] J. S. Ellenberg and A. Venkatesh. The number of extensions of a number field with fixed degree and bounded discriminant. *Ann. of Math. (2)*, 163(2):723–741, 2006.

[EV07] J. S. Ellenberg and A. Venkatesh. Reflection principles and bounds for class group torsion. *Int. Math. Res. Not. IMRN*, (1):Art. ID rnm002, 18, 2007.

[FK06] É. Fouvry and J. Klüners. On the 4-rank of class groups of quadratic number fields. *Inventiones mathematicae*, 167(3):455–513, November 2006.

[FNT92] É. Fouvry, M. Nair, and G. Tenenbaum. L'ensemble exceptionnel dans la conjecture de Szpiro. *Bull. Soc. Math. France*, 120(4):485–506, 1992.

[FW17] C. Frei and M. Widmer. Average bounds for the ℓ-torsion in class groups of cyclic extensions. *arXiv:1709.09934*, 2017.

[FW18] C. Frei and M. Widmer. Averages and higher moments for the ℓ-torsion in class groups. *arXiv:1810.04732*, 2018.

[Gau01] C. F. Gauss. *Disquisitiones Arithmeticae, 1801.* trans. Arthur A. Clarke, Yale University Press (1965), 1801.

[Has30] H. Hasse. Arithmetische Theorie der kubischen Zahlkörper auf klassenkörpertheoretischer Grundlage. *Math. Z.*, 31(1):565–582, 1930.

[HBP17] D.R. Heath-Brown and L. B. Pierce. Averages and moments associated to class numbers of imaginary quadratic fields. *Compositio Math.*, 153:2287–2309, 2017.

[Hei71] H. Heilbronn. On the 2-class group of cubic fields. In *Studies in Pure Mathematics (Presented to Richard Rado)*, pages 117–119. Academic Press, London, 1971.

[Hoo82] C. Hooley. On exponential sums and certain of their applications. In *Number theory days, 1980 (Exeter, 1980)*, volume 56 of *London Math. Soc. Lecture Note Ser.*, pages 92–122. Cambridge Univ. Press, Cambridge-New York, 1982.

[HV06] H. A. Helfgott and A. Venkatesh. Integral points on elliptic curves and 3-torsion in class groups. *J. Amer. Math. Soc.*, 19(3):527–550, 2006.

[HW08] G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers.* D. R. Heath-Brown and J. H. Silverman. Oxford University Press, 6th edition, 2008.

[IK04] H. Iwaniec and E. Kowalski. *Analytic Number Theory*, volume 53. Amer. Math. Soc. Colloquium Publications, Providence RI, 2004.

[Klü06a] J. Klüners. Asymptotics of number fields and the Cohen-Lenstra heuristics. *J. Théor. Nombres Bordeaux*, 18:607–615, 2006.

[Klü06b] J. Klüners. The number of S_4-fields with given discriminant. *Acta Arith.*, 122(2):185–194, 2006.

[Kly16] J. Klys. The distribution of p-torsion in degree p cyclic fields. *arXiv:1610.00226 [math]*, October 2016.

[KM16] P. Koymans and D. Milovic. On the 16-rank of class groups of $\mathbb{Q}(\sqrt{-p})$. *arXiv:1611.10337*, 2016.

[KM17] P. Koymans and D. Milovic. On the 16-rank of class groups of $\mathbb{Q}(\sqrt{-2p})$ for primes $p \equiv 1 \bmod 4$. *arXiv:1705.05505*, 2017.

[Kow10] E. Kowalski. Some aspects and applications of the Riemann hypothesis over finite fields. *Milan J. Math.*, 78(1):179–220, 2010.

[KP18] P. Koymans and C. Pagano. On the distribution of $\mathrm{Cl}(K)[\ell^\infty]$ for degree ℓ cyclic fields. *arXiv:1812.06884*, 2018.

[Mal02] G. Malle. On the distribution of Galois groups. *J. Number Theory*, 92:315–219, 2002.

[Mal04] G. Malle. On the distribution of Galois groups II. *Experiment. Math.*, 13:129–135, 2004.

[Mil17] D. Milovic. On the 16-rank of class groups of $\mathbb{Q}(\sqrt{-8p})$ for $p \equiv -1 \mod 4$. *Geom. Funct. Anal.*, 27(4):973–1016, 2017.
[Nar80] W. Narkiewicz. *Elementary and Analytic Theory of Algebraic Numbers, edition 2*. Springer-Verlag, Berlin, 1980.
[OT20] Robert J. Lemke Oliver and Frank Thorne. Upper bounds on number fields of given degree and bounded discriminant, 2020.
[Pie05] L. B. Pierce. The 3-part of class numbers of quadratic fields. *J. London Math. Soc.*, 71:579–598, 2005.
[Pie06] L. B. Pierce. A bound for the 3-part of class numbers of quadratic fields by means of the square sieve. *Forum Math.*, 18:677–698, 2006.
[PTBW20] L. B. Pierce, C. Turnage-Butterbaugh, and M. Matchett Wood. An effective Chebotarev density theorem for families of number fields, with an application to ℓ-torsion in class groups. *Inventiones Math.*, 219:701–778, 2020.
[Rup98] W. M. Ruppert. Small generators of number fields. *Manuscripta Math.*, 96(1):17–22, 1998.
[Sch76] W. Schmidt. *Equations over Finite Fields: An Elementary Approach*. Lecture Notes in Mathematics 536. Springer-Verlag, 1976.
[Sch95] Wolfgang M. Schmidt. Number fields of given degree and bounded discriminant. Number 228, pages 4, 189–195. 1995. Columbia University Number Theory Seminar (New York, 1992).
[Ser97] J.-P. Serre. *Lectures on the Mordell-Weil Theorem (3rd ed.)*. Friedr. Vieweg and Sohn, Braunschweig, 1997.
[Sil84] J. H. Silverman. Lower bounds for height functions. *Duke Math. J.*, 51(2):395–403, 1984.
[Sil09] Joseph H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics*. Springer, Dordrecht, second edition, 2009.
[Smi16] A. Smith. Governing fields and statistics for 4-Selmer groups and 8-class groups. *arXiv:1607.07860*, 2016.
[Smi17] A. Smith. 2^∞-Selmer groups, 2^∞-class groups, and Goldfeld's conjecture. *arXiv:1702.02325v2*, 2017.
[Sou00] K. Soundararajan. Divisibility of class numbers of imaginary quadratic fields. *J. London Math. Soc. (2)*, 61(3):681–690, 2000.
[ST14] A. Shankar and J. Tsimerman. Counting S_5-fields with a power saving error term. *Forum of Mathematics, Sigma*, 2, 2014.
[Szp90] L. Szpiro. Discriminant et conducteur des courbes elliptiques. *Astérisque*, (183):7–18, 1990. Séminaire sur les Pinceaux de Courbes Elliptiques (Paris, 1988).
[Tit86] E. C. Titchmarsh. *The theory of the Riemann zeta-function*. The Clarendon Press, Oxford University Press, New York, second edition, 1986. Edited and with a preface by D. R. Heath-Brown.
[Wat08] M. Watkins. Some heuristics about elliptic curves. *Experiment. Math.*, 17(1):105–125, 2008.
[Wid18] M. Widmer. Bounds for the ℓ-torsion in class groups. *Bull. L.M.S.*, (to appear), 2018.
[Won01a] S. Wong. Exponents of class groups and elliptic curves. *J. Number Theory*, 89(1):114–120, 2001.
[Won01b] S. Wong. Exponents of class groups and elliptic curves, corrigendum. *J. Number Theory*, 90(2):376–377, 2001.
[Woo15] M. M. Wood. Random integral matrices and the Cohen-Lenstra Heuristics. *arXiv:1504.04391*, 2015.
[Woo16] M. M. Wood. *Asymptotics for number fields and class groups*. Springer, 2016.
[Wri89] D. Wright. Distribution of discriminants of abelian extensions. *Proc. London Math. Soc.*, 58:17–50, 1989.
[WW19] Weitong Wang and Melanie Matchett Wood. Moments and interpretations of the Cohen-Lenstra-Martinet heuristics. *arXiv:1907.11201 [math]*, July 2019.
[You15] Matthew P. Young. The number of solutions to Mordell's equation in constrained ranges. *Mathematika*, 61(3):708–718, 2015.
[Zha05] S.-W. Zhang. Equidistribution of CM-points on quaternion Shimura varieties. *Int. Math. Res. Not.*, (59):3657–3689, 2005.

Department of Mathematics, Duke University, 120 Science Drive, Durham, NC 27708 USA
Email address: pierce@math.duke.edu

Department of Mathematics & Statistics, Carleton College, 1 North College Street, Northfield, MN 55057 USA
Email address: cturnageb@carleton.edu

Department of Mathematics, Harvard University, 1 Oxford Street, Cambridge, MA 02138 USA
Email address: mmwood@math.harvard.edu