COMPOSITIO MATHEMATICA

# Averages and moments associated to class numbers of imaginary quadratic fields

D. R. Heath-Brown and L. B. Pierce

Compositio Math. **153** (2017), 2287–2309.

doi:10.1112/S0010437X1700728X

FOUNDATION
COMPOSITIO
MATHEMATICA

LONDON
MATHEMATICAL
SOCIETY
EST. 1865

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

Compositio Math. 153 (2017) 2287–2309
doi:10.1112/S0010437X1700728X

# Averages and moments associated to class numbers of imaginary quadratic fields

D. R. Heath-Brown and L. B. Pierce

Abstract

For any odd prime $\ell$, let $h_\ell(-d)$ denote the $\ell$-part of the class number of the imaginary quadratic field $\mathbb{Q}(\sqrt{-d})$. Nontrivial pointwise upper bounds are known only for $\ell = 3$; nontrivial upper bounds for averages of $h_\ell(-d)$ have previously been known only for $\ell = 3, 5$. In this paper we prove nontrivial upper bounds for the average of $h_\ell(-d)$ for all primes $\ell \geqslant 7$, as well as nontrivial upper bounds for certain higher moments for all primes $\ell \geqslant 3$.

## 1. Introduction

Fix an imaginary quadratic field $\mathbb{Q}(\sqrt{-d})$ with square-free $-d < 0$, and let $\mathrm{Cl}(-d)$ be the corresponding class group. The size of the class group, denoted $h(-d)$, is the class number of $\mathbb{Q}(\sqrt{-d})$, a fundamental invariant that appears widely in number theory. The divisibility properties of class numbers of quadratic fields are subject to the conjectures known as the Cohen–Lenstra heuristics [CL84], which despite significant attention remain open in most cases. For any prime $\ell \geqslant 2$, let $h_\ell(-d)$ denote the $\ell$-part of the class number, that is the number of ideal classes in the class group $\mathrm{Cl}(-d)$ whose $\ell$th power is the principal ideal class. One may obtain a trivial pointwise upper bound for $h_\ell(-d)$ by noting that

$$h_\ell(-d) \leqslant h(-d) \ll d^{1/2+\varepsilon}.$$

It is conjectured that

$$h_\ell(-d) \ll d^{\varepsilon} \tag{1.1}$$

for all $d$ and any $\varepsilon > 0$. (Throughout, we will use the convention that all implied constants may depend upon $\ell$ and $\varepsilon$.)

This conjecture (and a more general version for $\ell$-torsion in class groups of number fields of any degree) is motivated by the Cohen–Lenstra heuristics [CL84], by counting elliptic curves with fixed conductor [BS96], by counting number fields of fixed degree and discriminant [Duk98], and by questions on equidistribution of CM-points on Shimura varieties [Zha05]. For $\ell = 2$, the conjecture (1.1) is known by the genus theory of Gauss. For $\ell = 3$ the currently best-known upper bound is due to Ellenberg and Venkatesh [EV07]:

$$h_3(-d) \ll d^{1/3+\varepsilon}. \tag{1.2}$$

For primes $\ell \geqslant 5$, no nontrivial upper bound for $h_\ell(-d)$ is known to hold for all $d$.

Received 15 October 2015, accepted in final form 11 February 2017, published online 14 August 2017.
*2010 Mathematics Subject Classification* 11R29, 11D45 (primary).
*Keywords:* class numbers, Cohen–Lenstra heuristics.

One may also consider averages

$$\sum_{0<d<X} h_\ell(-d).$$

In the case $\ell = 3$, Davenport and Heilbronn [DH71] established that

$$\sum_{0<d<X} h_3(-d) \sim 2 \sum_{0<d<X} 1, \tag{1.3}$$

as $X \to \infty$, in which both sums are restricted to fundamental discriminants. This asymptotic has recently been refined further to include secondary main terms (see Bhargava *et al.* [BST13], Taniguchi and Thorne [TT13], and Hough [Hou10]), but for the purposes of this paper it is sufficient that (1.3) provides an upper bound:

$$\sum_{0<d<X} h_3(-d) \ll X. \tag{1.4}$$

For $\ell = 5$, the best-known upper bound for the average is due to Soundararajan [Sou00] (also proved by Hough [Hou10]):

$$\sum_{0<d<X} h_5(-d) \ll X^{5/4+\varepsilon}. \tag{1.5}$$

For primes $\ell \geqslant 7$, the literature appears to contain no bound better than the trivial estimate

$$\sum_{0<d<X} h_\ell(-d) \ll X^{3/2+\varepsilon}.$$

However Soundararajan noted in [Sou00] that he has shown for any prime $\ell \geqslant 3$ that

$$h_\ell(-d) \ll d^{1/2-1/2\ell+\varepsilon} \tag{1.6}$$

for all but one square-free discriminant $d$ in any dyadic range $[X, 2X)$. Summing over $O(\log X)$ dyadic ranges implies the nontrivial average bound

$$\sum_{0<d<X} h_\ell(-d) \ll X^{3/2-1/2\ell+\varepsilon} \tag{1.7}$$

for any $\ell \geqslant 3$. While this is superseded by (1.4) and (1.5) for $\ell = 3$ and 5, no improvement has been given hitherto for larger values of $\ell$.

One can further consider the second moment; motivated by the conjecture (1.1) for the pointwise upper bound for $h_\ell(-d)$, one would expect that

$$\sum_{0<d<X} h_\ell(-d)^2 \ll X^{1+\varepsilon}.$$

For $\ell = 3$ and 5, one may bound the second moment by applying the best-known pointwise upper bound (respectively (1.2) and (1.6)) to one factor $h_\ell(-d)$, and then applying the best-known average upper bound to the remaining sum (respectively (1.4) and (1.5)). For $\ell \geqslant 7$, it is advantageous to apply Soundararajan's result (1.6) to both factors of $h_\ell(-d)$. This approach results in the following upper bounds for the second moment:

$$\sum_{0<d<X} h_\ell(-d)^2 \ll \begin{cases} X^{4/3+\varepsilon} & \ell = 3, \\ X^{33/20+\varepsilon} & \ell = 5, \\ X^{2-1/\ell+\varepsilon} & \ell \geqslant 7, \text{ prime}. \end{cases} \tag{1.8}$$

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

More generally, for any real number $k \geqslant 1$, known results lead to bounds for the $k$th moment of the form

$$\sum_{0<d<X} h_\ell(-d)^k \ll \begin{cases} X^{1+(k-1)/3+\varepsilon} & \ell = 3, \\ X^{5/4+(k-1)(2/5)+\varepsilon} + X^{k/2+\varepsilon} & \ell = 5, \\ X^{1+k((\ell-1)/2\ell)+\varepsilon} + X^{k/2+\varepsilon} & \ell \geqslant 7, \text{ prime.} \end{cases}$$

### 1.1 Statement of the theorems

The purpose of this paper is to improve on these bounds for the averages and moments of $h_\ell(-d)$ for $d$ square-free and $\ell$ an odd prime. (For the rest of this paper the notations $d$ and $\ell$ are reserved for square-free integers and odd primes respectively.)

Theorem 1.1. *For each prime $\ell \geqslant 5$,*

$$\sum_{0<d<X} h_\ell(-d) \ll X^{3/2-3/(2\ell+2)+\varepsilon},$$

*for any $\varepsilon > 0$.*

This recaptures Soundararajan's result (1.5) for $\ell = 5$ and improves on the bound (1.7) for all primes $\ell \geqslant 7$. (Since Davenport and Heilbronn's result (1.3) is best possible, our work provides no new information for the average of $h_3(-d)$.)

We also consider higher moments. First we consider the moments of $h_3(-d)$, for which our main result is the following.

Theorem 1.2.

$$\sum_{0<d<X} h_3(-d)^4 \ll X^{11/6+\varepsilon} \quad \textit{for any } \varepsilon > 0.$$

It may be surprising to see the fourth moment here, but it turns out to give the best results of its type, as we shall see.

By the reflection principle of Scholz [Sch32], $\log_3 h_3(-d)$ and $\log_3 h_3(+3d)$ differ by at most one. Thus the corresponding bound for the 3-part of the class number of real quadratic fields follows as a corollary, making an identical improvement over previously known bounds as in the imaginary case.

Corollary 1.3.

$$\sum_{0<d<X} h_3(d)^4 \ll X^{11/6+\varepsilon} \quad \textit{for any } \varepsilon > 0.$$

Nontrivial bounds for other moments are also an immediate corollary. For $1 \leqslant k < 4$ one merely uses Hölder's inequality in conjunction with (1.4), while for $k > 4$ one just applies (1.2) in combination with Theorem 1.2.

Corollary 1.4. *For all real $k \in [1, 4]$, and for any $\varepsilon > 0$,*

$$\sum_{0<d<X} h_3(-d)^k \ll X^{(5k+13)/18+\varepsilon},$$

$$\sum_{0<d<X} h_3(d)^k \ll X^{(5k+13)/18+\varepsilon}.$$

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

*For all real $k \geqslant 4$, and for any $\varepsilon > 0$,*

$$\sum_{0<d<X} h_3(-d)^k \ll X^{(2k+3)/6+\varepsilon},$$
$$\sum_{0<d<X} h_3(d)^k \ll X^{(2k+3)/6+\varepsilon}.$$

*In particular, for any $\varepsilon > 0$,*

$$\sum_{0<d<X} h_3(-d)^2 \ll X^{23/18+\varepsilon}.$$

This final bound improves on (1.8); we note that $23/18 = 1.2777\ldots$.

We next consider higher moments for $h_\ell(-d)$ for primes $\ell \geqslant 5$. Theorem 1.1 combined with (1.6) implies that, for any real $k \geqslant 1$,

$$\sum_{0<d<X} h_\ell(-d)^k \ll X^{3/2-3/(2\ell+2)+(k-1)(1/2-1/2\ell)+\varepsilon} + X^{k/2+\varepsilon},$$

where the last term arises from the possible exceptions to (1.6). For purposes of comparison, we rewrite this as

$$\sum_{0<d<X} h_\ell(-d)^k \ll X^{1+k((\ell-1)/2\ell)-(2\ell-1)/(2\ell(\ell+1))+\varepsilon} + X^{k/2+\varepsilon}.$$

We will improve on this for all real $1 < k < (2\ell^2+1)/(\ell+1)$.

Theorem 1.5. *For any prime $\ell \geqslant 5$, all real $k \geqslant 1$, and any $\varepsilon > 0$,*

$$\sum_{0<d<X} h_\ell(-d)^k \ll \begin{cases} X^{1+k((\ell-2)/(2\ell+2))+\varepsilon} & \text{if } 1 \leqslant k \leqslant \dfrac{\ell^2-1}{2\ell-1}, \\ X^{1+k((\ell-1)/2\ell)-((\ell-1)/2\ell)+\varepsilon} & \text{if } \dfrac{\ell^2-1}{2\ell-1} \leqslant k \leqslant \ell+1, \\ X^{k/2+\varepsilon} & \text{if } k \geqslant \ell+1. \end{cases}$$

In particular, we single out the consequence of Theorem 1.5 for the second moment (noting that $k = 2$ lies in the first case of the theorem for $\ell \geqslant 5$).

Corollary 1.6. *For any prime $\ell \geqslant 5$, for any $\varepsilon > 0$,*

$$\sum_{0<d<X} h_\ell(-d)^2 \ll X^{2-3/(\ell+1)+\varepsilon}.$$

This improves on (1.8) in every case. Theorem 1.1 may of course be deduced from the above corollary via the Cauchy–Schwarz inequality. However we have stated and proved Theorem 1.1 separately since it is, in effect, used in the proof of Theorem 1.5.

Our approach is to develop an unconditional upper bound for $h_\ell(-d)$ that holds for almost all $d$, by using the relation between $h_\ell(-d)$ and small split primes in $\mathbb{Q}(\sqrt{-d})$. The original observation of this relation is credited to Soundararajan (and to Michel in a related context) in the work of Helfgott and Venkatesh [HV06] and Ellenberg and Venkatesh [EV07], and has been used in [HV06], for example, to prove a bound for $h_3(-d)$ for all $d$, conditional on the Generalized Riemann Hypothesis. Here we prove an unconditional version, at the cost that it only holds for 'almost all' $d$. To treat higher moments, we combine this with upper bounds for the number of

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

simultaneous representations of integers by certain polynomials; this counting problem is similar to computations performed in [Sou00] and [HB07]. Finally, we remark that the methods of § 6 may also be applied to prove upper bounds for mixed averages of the form

$$\sum_{0<d<X} h_\ell(-d) h_{\ell'}(-d)$$

for distinct odd primes $\ell, \ell'$; we leave the details to the interested reader.

We reiterate that throughout this paper we consider sums over $0 < d < X$ to be restricted to square-free integers, and $\ell$ represents an odd prime. We will frequently combine factors of size $X^\varepsilon$ for various $\varepsilon$; in all cases $\varepsilon$ may be taken to be an arbitrarily small real number, so we re-define it wherever appropriate so that the total factor remains represented by $X^\varepsilon$. We also use the notation $A \ll B$ to indicate that there is a constant $c$, possibly depending on certain allowable parameters such as $\ell$ or $\varepsilon$, such that $|A| \leqslant c|B|$, and similarly for $A \gg B$.

## 2. An unconditional pointwise upper bound

Our starting point is the following unconditional pointwise upper bound for $h_\ell(-d)$.

Proposition 2.1. *Fix any prime $\ell \geqslant 3$ and real parameters $\frac{1}{4}X^{1/2\ell} \leqslant Z \leqslant X$. There exists a small exceptional set $E(Z;X) \subset [X, 2X)$ such that for all square-free $d \in [X, 2X) \backslash E(Z;X)$,*

$$h_\ell(-d) \ll X^\varepsilon \{d^{1/2} Z^{-1} + d^{1/2} Z^{-2} S_\ell(d;Z)\},$$

*for any $\varepsilon > 0$, where $S_\ell(d;Z)$ is the cardinality of the set of pairs of primes $p, p'$ satisfying*

$$Z \leqslant p \neq p' < 2Z$$

*for which there exist $u, v \in \mathbb{Z} \backslash \{0\}$ with $(v, pp') = 1$ such that*

$$4(pp')^\ell = u^2 + dv^2.$$

*Moreover, the exceptional set satisfies*

$$\# E(Z;X) \ll X^{\varepsilon'} \tag{2.1}$$

*for any $\varepsilon' > 0$.*

Corollary 2.2. *Fix any $\varepsilon' > 0$. For all $d \in [X, 2X)$ apart from at most $O(X^{\varepsilon'})$ exceptions,*

$$h_\ell(-d) \ll d^{1/2-1/2\ell+\varepsilon}$$

*for any $\varepsilon > 0$.*

This corollary, which we will prove at the end of § 2, gives a weak form of Soundararajan's result concerning the bound (1.6).

It is clear from Proposition 2.1 that an understanding of $S_\ell(d;Z)$, both in terms of its average over $d$ and its second moment, will yield corresponding information for $h_\ell(-d)$. Our two main technical results are for the average and second moment of $S_\ell(d;Z)$.

Proposition 2.3. *For any prime $\ell \geqslant 3$ and $X^{1/2\ell} \leqslant Z \leqslant X$,*

$$\sum_{X \leqslant d < 2X} S_\ell(d;Z) \ll X^\varepsilon \{Z^2 X^{1/2} + Z^{\ell+2} X^{-1/2}\}$$

*for any $\varepsilon > 0$.*

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

Proposition 2.4. *For $\ell = 3$ and $X^{1/6} \leqslant Z \leqslant X$,*

$$\sum_{X \leqslant d < 2X} S_3(d; Z)^2 \ll X^{\varepsilon}\{Z^2 X^{1/2} + Z^{12} X^{-3/2}\}$$

*for any $\varepsilon > 0$. For any prime $\ell \geqslant 5$ and $X^{1/2\ell} \leqslant Z \leqslant X$,*

$$\sum_{X \leqslant d < 2X} S_\ell(d; Z)^2 \ll X^{\varepsilon}\{Z^2 X^{1/2} + Z^{2\ell+4} X^{-1}\}$$

*for any $\varepsilon > 0$.*

We include the case $\ell \geqslant 5$ in Proposition 2.4 as it requires little extra effort, but we will not make use of it: while it does result in a nontrivial upper bound for the second moment of $h_\ell(-d)$, a stronger result may be obtained by applying Proposition 2.3 directly.

In the remainder of this section, we prove Proposition 2.1 and its corollary. We prove Propositions 2.3 and 2.4 in §§ 3 and 4, respectively. Finally, in §§ 5 and 6 we record the consequences of these results for averages and moments of $h_\ell(-d)$.

### 2.1 Proof of Proposition 2.1

Fix a prime $\ell \geqslant 3$ and a square-free integer $X \leqslant d < 2X$. Let $H = \mathrm{Cl}(-d)$ be the class group of $\mathbb{Q}(\sqrt{-d})$, with class number $h(-d) = \#\mathrm{Cl}(-d)$. Let $H_\ell$ denote the maximal elementary abelian $\ell$-group in $H$, with $h_\ell(-d) = \#H_\ell$. Since

$$\#H/H_\ell = \frac{h(-d)}{h_\ell(-d)}, \tag{2.2}$$

in order to show that $h_\ell(-d)$ is small it suffices to show that there are many cosets of $H_\ell$ in $H$. Let $\chi_d(\cdot)$ denote the quadratic character associated to $\mathbb{Q}(\sqrt{-d})$. Picking a prime $p \nmid 2d$ such that $\chi_d(p) = 1$, it follows that $p$ splits in $\mathbb{Q}(\sqrt{-d})$ as $\mathfrak{p}\mathfrak{p}^\sigma$, say, where $\sigma$ is the nontrivial Galois automorphism of $\mathbb{Q}(\sqrt{-d})$. Suppose that two distinct primes $p, p'$ split in this manner as $\mathfrak{p}\mathfrak{p}^\sigma$ and $\mathfrak{p}'\mathfrak{p}'^\sigma$ respectively, and suppose that $\mathfrak{p}$ and $\mathfrak{p}'$ represent the same class in $H/H_\ell$, so that $\mathfrak{p}H_\ell = \mathfrak{p}'H_\ell$. It follows that $\mathfrak{p}^{-1}\mathfrak{p}' \in H_\ell$, so that $(\mathfrak{p}^{-1}\mathfrak{p}')^\ell$ is a principal ideal. Thus $(\mathfrak{p}^\sigma\mathfrak{p}')^\ell$ is also a principal ideal, say

$$(\mathfrak{p}^\sigma\mathfrak{p}')^\ell = \left(\frac{u + v\sqrt{-d}}{2}\right), \tag{2.3}$$

for some $u, v \in \mathbb{Z}$. Hence taking norms, it follows that

$$4(pp')^\ell = u^2 + dv^2. \tag{2.4}$$

Note that we may require that $\gcd(v, pp') = 1$ (and in particular that $v \neq 0$). For supposing that $p \mid v$, say, then by (2.4) we see that also $p \mid u$ so that $p \mid ((u + v\sqrt{-d})/2)$. Hence $\mathfrak{p} \mid ((u + v\sqrt{-d})/2)$, which by (2.3) implies that $\mathfrak{p} \mid (\mathfrak{p}^\sigma\mathfrak{p}')^\ell$. Since $p$ is unramified this would then imply that $\mathfrak{p} \mid \mathfrak{p}'$, which contradicts the fact that $p \neq p'$. A similar argument shows that we may require that $u \neq 0$.

We will show that for all but a small number of 'exceptional' $d$ there are many primes $p, p'$ that split in this manner, while also showing there can only be few solutions $(u, v)$ to (2.4) with $\gcd(v, pp') = 1$ and $u, v$ in an appropriate range. This forces there to be many distinct cosets of $H_\ell$ in $H$, and provides an upper bound for $h_\ell(-d)$, as long as $d$ is not exceptional.

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

We first fix $X \leqslant d < 2X$ and count the number of primes $p$ that split appropriately, with

$$Z \leqslant p < 2Z$$

for some parameter $Z$ with $\frac{1}{4}X^{1/2\ell} \leqslant Z \leqslant X$ (to be chosen precisely in applications). We see that

$$\#\{Z \leqslant p < 2Z : \chi_d(p) = 1\} = \frac{1}{2}\sum_{Z \leqslant p < 2Z}(1 + \chi_d(p)) + O(\omega(d)),$$

where the last term reflects the contribution of the primes that divide $d$, and contributes no more than $O(\log X) = O(\log Z)$. We now separate the two terms within the sum over $p$ and apply the prime number theorem, obtaining

$$\#\{Z \leqslant p < 2Z : \chi_d(p) = 1\} = \tfrac{1}{2}Z(\log Z)^{-1} + \tfrac{1}{2}M(d;Z) + O(Z(\log Z)^{-2}),$$

say, where

$$M(d;Z) = \sum_{Z \leqslant p < 2Z}\chi_d(p).$$

Thus the number of split primes in this range is at least of order $Z(\log Z)^{-1}$, unless we have $|M(d;Z)| \geqslant \frac{1}{4}Z(\log Z)^{-1}$; we will show this exceptional scenario can occur for only a small number of $d$.

Given a character $\chi$, set

$$V(\chi) = \biggl(\sum_{Z \leqslant p < 2Z}\chi(p)\biggr)^{4\ell}.$$

Upon unfolding the product, we see that this is a character sum of the form

$$\sum_{Z^{4\ell} \leqslant n < (2Z)^{4\ell}} a_n\chi(n)$$

for some coefficients $|a_n| \ll d(n)^{4\ell} \ll Z^{\varepsilon}$. Now we note that with the particular choice $\chi = \chi_d$,

$$\sum_{X \leqslant d < 2X}|M(d;Z)|^{8\ell} = \sum_{X \leqslant d < 2X}|V(\chi_d)|^2. \tag{2.5}$$

By positivity, we can enlarge the sum on the right-hand side of (2.5) to include all primitive characters modulo $d$ and apply the large sieve (see, for example, [Dav00, Theorem 4, ch. 27]), to obtain

$$\begin{aligned}\sum_{X \leqslant d < 2X}|M(d;Z)|^{8\ell} &\leqslant \sum_{X \leqslant d < 2X}\ \sideset{}{^*}\sum_{\chi \ (\mathrm{mod}\ d)}|V(\chi)|^2\\ &\ll (X^2 + Z^{4\ell})\biggl(\sum_{Z^{4\ell} \leqslant n < (2Z)^{4\ell}}|a_n|^2\biggr)\\ &\ll Z^{4\ell+2\varepsilon}(X^2 + Z^{4\ell}) \ll Z^{8\ell+2\varepsilon},\end{aligned} \tag{2.6}$$

since $X^{1/2\ell} \ll Z$ by assumption. Let $E(Z;X)$ denote the exceptional set,

$$E(Z;X) = \{X \leqslant d < 2X : |M(d;Z)| \geqslant \tfrac{1}{4}Z(\log Z)^{-1}\}. \tag{2.7}$$

Then we may conclude from (2.6) that the exceptional set is small:

$$\#E(Z;X) \ll X^{\varepsilon},$$

for any $\varepsilon > 0$.

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

We now fix a $d$ with $X \leqslant d < 2X$ such that $d \notin E(X,Z)$; the above argument shows that there are at least of order $Z(\log Z)^{-1}$ split primes for this $d$. In particular, summing over all cosets of $H_\ell$ in $H$ shows that, for this $d$,

$$\sum_{C \in H/H_\ell} \#\{Z \leqslant p < 2Z : \chi_d(p) = 1, p = \mathfrak{p}\mathfrak{p}^\sigma, \mathfrak{p} \in C\} = \#\{Z \leqslant p < 2Z : \chi_d(p) = 1\} \gg Z(\log Z)^{-1}.$$

On the other hand, applying the Cauchy–Schwarz inequality to the left-hand side shows that

$$(\#H/H_\ell)^{1/2}(S_\ell^{(1)}(d;Z))^{1/2} \gg Z(\log Z)^{-1}, \tag{2.8}$$

where we define

$$S_\ell^{(1)}(d;Z) = \sum_{C \in H/H_\ell} \#\{Z \leqslant p < 2Z : \chi_d(p) = 1, p = \mathfrak{p}\mathfrak{p}^\sigma, \mathfrak{p} \in C\}^2.$$

By the above discussion, we know that

$$S_\ell^{(1)}(d;Z) \ll \#\{Z \leqslant p, p' < 2Z : 4(pp')^\ell = u^2 + dv^2 \text{ for some } u, v \in \mathbb{Z}\}, \tag{2.9}$$

where in the case that $p \neq p'$ we may impose the additional conditions that $u, v \neq 0$ and $(v, pp') = 1$. Combining (2.8) and (2.2), we may conclude that

$$h_\ell(-d) \ll d^{1/2+\varepsilon} Z^{-2} (\log Z)^2 S_\ell^{(1)}(d;Z),$$

still under the assumption that $d$ is not exceptional. Finally, we write

$$S_\ell^{(1)}(d;Z) = S_\ell^{(0)}(d;Z) + S_\ell(d;Z),$$

where $S_\ell^{(0)}(d;Z)$ is the contribution to the set (2.9) from pairs $p = p'$ and $S_\ell(d;Z)$ is the contribution from pairs $p \neq p'$. Trivially, $S_\ell^{(0)}(d;Z) \ll Z$, and we see that Proposition 2.1 holds.

To deduce the corollary we take $Z = \frac{1}{4}X^{1/2\ell}$, and note that any pairs of primes $p, p'$ counted by $S_\ell(d;Z)$ would satisfy

$$X \leqslant d \leqslant u^2 + dv^2 = 4(pp')^\ell \leqslant 4(4Z^2)^\ell = 4^{1-\ell}X < X.$$

Thus $S_\ell(d;Z)$ must vanish, so that $h_\ell(-d) \ll X^\varepsilon d^{1/2} Z^{-1} \ll d^{1/2-1/(2\ell)+\varepsilon}$ unless $d$ lies in $E(Z;X)$. The result then follows.

## 3. Proof of Proposition 2.3

Define the parameters

$$W = Z^2, \quad U = 2^{\ell+1}Z^\ell, \quad V = 2^{\ell+1}Z^\ell X^{-1/2}. \tag{3.1}$$

Note that $V \geqslant 2$ as long as

$$Z \geqslant X^{1/2\ell},$$

which we henceforward assume. Note also that up to a constant factor (accounting for changing signs of $u, v$), we may express $S_\ell(d;Z)$ as the quantity

$$\#\{Z \leqslant p \neq p' < 2Z : 4(pp')^\ell = u^2 + dv^2 \text{ for some } u, v \geqslant 1 \text{ with } (v, pp') = 1\}.$$

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

Furthermore, for any $X \leqslant d < 2X$, any triple $w = pp'$, $u, v$ considered in the set above certainly satisfies $W \leqslant w < 4W$, $1 \leqslant u \leqslant U$, $1 \leqslant v \leqslant V$.

We wish to bound $S_\ell(d; Z)$ on average over $d$; for this we note that

$$\sum_{\substack{X \leqslant d < 2X \\ d \notin E(Z;X)}} S_\ell(d; Z) \ll \# \{W \leqslant w < 4W, 1 \leqslant u \leqslant U, 1 \leqslant v \leqslant V : \gcd(v, w) = 1,$$
$$v^2 \mid (4w^\ell - u^2), (4w^\ell - u^2)/v^2 \in [X, 2X)\} .$$

It is convenient to work with dyadic ranges; thus for any parameter $1 \leqslant V_0 \leqslant V/2$, define

$$N(Z, X; V_0) = \# \{W \leqslant w < 4W, 1 \leqslant u \leqslant U, V_0 \leqslant v < 2V_0 : \gcd(v, w) = 1,$$
$$v^2 \mid (4w^\ell - u^2), (4w^\ell - u^2)/v^2 \in [X, 2X)\} .$$

Then certainly

$$\sum_{\substack{X \leqslant d < 2X \\ d \notin E(Z;X)}} S_\ell(d; Z) \ll \sum_{0 \leqslant j \leqslant \log_2(V)-1} N(Z, X; 2^j) = \sum_{\substack{V_0 \leqslant V/2 \\ \text{dyadic}}} N(Z, X; V_0).$$

We turn to bounding an individual term $N(Z, X; V_0)$. We first fix $w$ and $v$ and let

$$M(w; v) = \#\{u \pmod{v^2} : u^2 \equiv 4w^\ell \pmod{v^2}\}.$$

Lemma 3.1. *For any coprime $w$ and $v$,*

$$M(w; v) \leqslant 2^{\omega(v)+1} \ll v^\varepsilon, \tag{3.2}$$

*where $\omega(v)$ denotes the number of distinct prime divisors of $v$.*

*Proof.* This is proved in a standard fashion. Writing $v = q_1^{r_1} \cdots q_s^{r_s}$ in its prime decomposition, it suffices by the Chinese Remainder Theorem to count $M(w; q_i^{r_i})$ for each $q_i$. Since $(w, v) = 1$ we may assume that $(w, q_i) = 1$; we also assume for the moment that $q_i$ is odd. Then $M(w; q_i^{r_i})$ will be nonzero only if $w$ is a quadratic residue modulo $q_i$, in which case $u$ can lie in at most two residue classes modulo $q_i$; since $q_i$ is odd, each solution modulo $q_i$ lifts uniquely to a solution modulo $q_i^{2r_i}$. Thus we see that in this case

$$M(w; q_i^{r_i}) \leqslant 2.$$

If $q_i = 2$ then the relevant congruence has solutions only if $2 \mid u$, in which case we may equivalently count solutions to $(u/2)^2 \equiv w^\ell \pmod{q_i^{2r_i-2}}$. However if $n$ is odd, a congruence $x^2 \equiv n \pmod{2^r}$ has at most four solutions. We may therefore conclude that $M(w; q_i^{r_i}) \leqslant 4$, thus proving (3.2). □

Applying Lemma 3.1 directly to count solutions $u \leqslant U$ to $u^2 \equiv 4w^\ell \pmod{v^2}$ would lead to the upper bound

$$N(Z, X; V_0) \ll W V_0^{1+\varepsilon}(U V_0^{-2} + 1). \tag{3.3}$$

But then summing over all dyadic ranges with $1 \leqslant V_0 \leqslant V/2$ would not allow us to take advantage of the decay with respect to $V_0$ in (3.3). Thus we return to the definition of $N(Z, X; V_0)$ and utilize the additional piece of information that

$$X \leqslant \frac{4w^\ell - u^2}{v^2} < 2X,$$

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

which we rewrite as

$$v^2X \leqslant 4w^\ell - u^2 < 2v^2X. \tag{3.4}$$

We will conclude from this that $u$ must lie within a short interval around $2w^{\ell/2}$; precisely, we write

$$\left(\frac{u}{2w^{\ell/2}}\right)^2 = 1 + E,$$

in which (3.4) shows that

$$|E| \leqslant \frac{2Xv^2}{4w^\ell} \leqslant \frac{8XV_0^2}{4W^\ell} = \frac{2XV_0^2}{Z^{2\ell}} = 2^{2\ell+3}\frac{V_0^2}{V^2}.$$

Thus $E \ll 1$ whence $\sqrt{1+E} = 1 + O(E)$. It follows that

$$u = 2w^{\ell/2} + O(w^{\ell/2}E) = 2w^{\ell/2} + O(W^{\ell/2}V_0^2V^{-2}).$$

Thus for each fixed $w, v$, in order to be counted by $N(Z, X; V_0)$, $u$ must lie in an interval $I_w$ around $2w^{\ell/2}$ of length $O(W^{\ell/2}V_0^2V^{-2})$. We apply this information along with the bound (3.2) to conclude that for each fixed $w, v$ considered in $N(Z, X; V_0)$,

$$\#\{u \in I_w : u^2 \equiv 4w^\ell \pmod{v^2}\} \ll V_0^\varepsilon\left(\frac{W^{\ell/2}V_0^2V^{-2}}{V_0^2} + 1\right) = V_0^\varepsilon(W^{\ell/2}V^{-2} + 1).$$

As a consequence,

$$\begin{aligned} N(Z, X; V_0) &\ll \sum_{\substack{W \leqslant w < 4W, V_0 \leqslant v < 2V_0 \\ (v,w)=1}} \#\{u \in I_w : u^2 \equiv 4w^\ell \pmod{v^2}\} \\ &\ll WV_0^{1+\varepsilon}(W^{\ell/2}V^{-2} + 1). \end{aligned}$$

(This improves upon (3.3) by effectively replacing $V_0^{-2}$ by $V^{-2}$; observe that up to constant factors, $U$ is the same size as $W^{\ell/2}$.) Summing over dyadic regions then shows

$$\begin{aligned} \sum_{\substack{V_0 \leqslant V/2 \\ \text{dyadic}}} N(Z, X; V_0) &\ll W^{1+\ell/2}V^{-1+\varepsilon} + WV^{1+\varepsilon} \\ &\ll X^\varepsilon\{Z^2X^{1/2} + Z^{\ell+2}X^{-1/2}\}, \end{aligned}$$

which proves Proposition 2.3.

## 4. Proof of Proposition 2.4

We define a quantity $R_\ell(d; Z)$ according to the parameters $U, V, W$ given in (3.1) as follows: set $R_\ell(d; Z) = 0$ if $d$ is not square-free, and for $d$ square-free let $R_\ell(d; Z)$ be the number of triples $(w, u, v) \in \mathbb{N}^3$ satisfying

$$\begin{gathered} W \leqslant w < 4W, \quad u \leqslant U, \quad v \leqslant V, \quad \gcd(w, v) = 1, \\ w = p_1p_2 \quad \text{with } p_1 \neq p_2 \in [Z, 2Z), \end{gathered}$$

and

$$4w^\ell = u^2 + dv^2.$$

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

Recall also the quantity $S_\ell(d;Z)$ defined in Proposition 2.1. Upon letting $w = p_1p_2$, we observe that (up to signs) any tuple $p_1, p_2, u, v$ contributing to $S_\ell(d;Z)$ must have $W \leqslant w < 4W$, $1 \leqslant u \leqslant U$, $1 \leqslant v \leqslant V$, so that $S_\ell(d;Z) \ll R_\ell(d;Z)$. Thus we may write

$$\sum_{X\leqslant d<2X} S_\ell(d;Z)^2 \ll \sum_{X\leqslant d<2X} R_\ell(d;Z) + \sum_{X\leqslant d<2X} R_\ell(d;Z)(R_\ell(d;Z)-1). \tag{4.1}$$

The advantage of separating the terms in this fashion is that in the second term on the right-hand side we may now count only distinct tuples $(u,v,w) \neq (u',v',w')$ in $R_\ell(d;Z)$.

We note that for $X^{1/2\ell} \leqslant Z \leqslant X$ the first term on the right-hand side of (4.1) satisfies

$$\sum_{X\leqslant d<2X} R_\ell(d;Z) \ll \sum_{\substack{V_0\leqslant V/2\\ \text{dyadic}}} N(Z,X;V_0) \ll X^\varepsilon\{Z^2X^{1/2} + Z^{\ell+2}X^{-1/2}\}, \tag{4.2}$$

by Proposition 2.3. The main remaining task is to treat

$$T_\ell = T_\ell(Z;X) := \sum_{X\leqslant d<2X} R_\ell(d;Z)(R_\ell(d;Z)-1).$$

We will prove the following proposition.

Proposition 4.1. *For $X^{1/2\ell} \leqslant Z \leqslant X$,*

$$T_\ell \ll Z^{2\ell+4}X^{\varepsilon-1}. \tag{4.3}$$

*Moreover when $\ell = 3$ and $X^{1/6} \leqslant Z \leqslant X$ we have*

$$T_3 \ll X^\varepsilon(Z^7X^{-1/2} + Z^{12}X^{-3/2}). \tag{4.4}$$

Combining (4.2) and (4.3), we see that

$$\sum_{X\leqslant d<2X} S_\ell(d;Z)^2 \ll X^\varepsilon(Z^2X^{1/2} + Z^{\ell+2}X^{-1/2} + Z^{2\ell+4}X^{-1}).$$

Note that

$$Z^{\ell+2}X^{-1/2} \leqslant Z^{2\ell+4}X^{-1}$$

for $Z \geqslant X^{1/(2\ell)}$, so that under this assumption

$$\sum_{X\leqslant d<2X} S_\ell(d;Z)^2 \ll X^\varepsilon(Z^2X^{1/2} + Z^{2\ell+4}X^{-1}).$$

This suffices for Proposition 2.4 for $\ell \geqslant 5$. For $\ell = 3$ we improve on this; from (4.2) and (4.4) we obtain

$$\sum_{X\leqslant d<2X} S_3(d;Z)^2 \ll X^\varepsilon(Z^2X^{1/2} + Z^5X^{-1/2} + Z^7X^{-1/2} + Z^{12}X^{-3/2}).$$

However

$$Z^5X^{-1/2} \leqslant Z^7X^{-1/2} = \{Z^2X^{1/2}\}^{1/2}\{Z^{12}X^{-3/2}\}^{1/2} \leqslant Z^2X^{1/2} + Z^{12}X^{-3/2},$$

whence the case $\ell = 3$ of Proposition 2.4 also follows.

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

### 4.1 Proof of Proposition 4.1: a first bound for $T_\ell$

We now prove (4.3). We recall the parameters $U, V, W$ of (3.1) and note that $T_\ell$ is at most the number of 6-tuples $(w_1, w_2, u_1, u_2, v_1, v_2)$ in the ranges

$$W \leqslant w_1, w_2 < 4W, \quad 1 \leqslant u_1, u_2 \leqslant U, \quad 1 \leqslant v_1, v_2 \leqslant V$$

that satisfy the conditions

$$(u_1, v_1, w_1) \neq (u_2, v_2, w_2), \tag{4.5}$$
$$\gcd(w_1, v_1) = \gcd(w_2, v_2) = 1, \tag{4.6}$$
$$v_1^2 \mid (4w_1^\ell - u_1^2) \quad \text{and} \quad v_2^2 \mid (4w_2^\ell - u_2^2), \tag{4.7}$$
$$v_1^2(4w_2^\ell - u_2^2) = v_2^2(4w_1^\ell - u_1^2) \neq 0. \tag{4.8}$$

We will obtain a first upper bound for $T_\ell$ by following the approach of [Sou00], ignoring the divisibility conditions (4.7); note that we are also now ignoring the fact that each of $w_1, w_2$ is a product of two distinct primes. We claim that for tuples satisfying the above conditions,

$$v_1^2 w_2^\ell - v_2^2 w_1^\ell \neq 0. \tag{4.9}$$

To prove this we recall that $\gcd(w_i, v_i) = 1$ for $i = 1, 2$, whence $v_1^2 w_2^\ell = v_2^2 w_1^\ell$ would imply that $v_1 = v_2$ and $w_1 = w_2$, and hence $u_1 = u_2$. This would then contradict (4.5).

We now observe that once $v_1, v_2, w_1, w_2$ are fixed then $u_1, u_2$ are fixed up to $X^\varepsilon$ choices. For indeed, fixing $v_1, v_2, w_1, w_2$ in (4.8) gives

$$4(v_2^2 w_1^\ell - v_1^2 w_2^\ell) = (v_2 u_1 - v_1 u_2)(v_2 u_1 + v_1 u_2). \tag{4.10}$$

The left-hand side is a nonzero integer by (4.9), so that $u_1, u_2$ are fixed up to $X^\varepsilon$ choices. Thus we obtain

$$T_\ell \ll W^2 V^2 X^\varepsilon \ll Z^{2\ell+4} X^{-1+\varepsilon},$$

which is the bound given in (4.3).

### 4.2 Proof of Proposition 4.1: a second bound for $T_\ell$

We may obtain the alternative upper bound (4.4) for $T_\ell$ by following the method of [HB07], but with the addition of certain technical considerations because in the present case the variables $v_i$ are not restricted to be primes. Although it is easy enough to do this for general odd primes $\ell$ we shall confine our attention to $\ell = 3$, since this is the only case we shall use.

First we consider the contribution to $T_3$ arising from the case in which $\gcd(w_1, w_2) \neq 1$. We write $T_3^0$ for the number of 6-tuples of this type. Since each of $w_1$ and $w_2$ is a product of two primes in the interval $[Z, 2Z)$ this can happen only when there is at least one prime $p \in [Z, 2Z)$ dividing both of $w_1$ and $w_2$. The number of possible pairs $w_1, w_2$ is thus $O(Z^3)$. We now follow the argument of § 4.1. There are $O(V^2)$ pairs $v_1, v_2$, and the factorization (4.10) shows that there are $O(X^\varepsilon)$ possibilities for $u_1, u_2$ once $w_1, w_2, v_1, v_2$ are fixed. It follows that

$$T_3^0 \ll Z^3 V^2 X^\varepsilon.$$

From now on we assume that $\gcd(w_1, w_2) = 1$. For each integer $1 \leqslant \delta \leqslant V$, we will let $T_3(\delta)$ denote the contribution to $T_3$ from triples $(u_1, v_1, w_1)$ and $(u_2, v_2, w_2)$ with $w_1, w_2$ coprime, such that $\gcd(v_1, v_2) = \delta$. We will prove the following proposition.

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

Proposition 4.2. *For each integer* $1 \leqslant \delta \leqslant V$,

$$T_3(\delta) \ll X^{\varepsilon}(W^2V^{2/3}\delta^{-2/3} + V^3U\delta^{-3} + WV\delta^{-1}).$$

From this we conclude that

$$\begin{aligned} T_3 &\ll T_3^0 + \sum_{\delta=1}^{V} T_3(\delta) \\ &\ll X^{\varepsilon}\biggl\{Z^3V^2 + \sum_{\delta=1}^{V}(W^2V^{2/3}\delta^{-2/3} + V^3U\delta^{-3} + WV\delta^{-1})\biggr\} \\ &\ll X^{\varepsilon}(Z^3V^2 + VW^2 + V^3U + WV) \\ &\ll X^{\varepsilon}(Z^3V^2 + VW^2 + V^3U), \end{aligned}$$

since clearly $WV \ll VW^2$. Upon recalling the parameter definitions (3.1) this shows that

$$T_3 \ll X^{\varepsilon}\{Z^9X^{-1} + Z^7X^{-1/2} + Z^{12}X^{-3/2}\}.$$

This provides the second bound for $T_3$ given in Proposition 4.1, since $Z \geqslant X^{1/6}$.

*Proof of Proposition 4.2.* To prove Proposition 4.2, we fix $\delta$ and write $v_i = \delta y_i$ for $i = 1, 2$ so that $\gcd(y_1, y_2) = 1$. We first isolate solutions $(u_1, v_1, w_1)$ and $(u_2, v_2, w_2)$ that contribute to $T_3(\delta)$ such that $y_1, y_2$ satisfy a relation

$$y_1^2\mu_2^3 = y_2^2\mu_1^3 \tag{4.11}$$

for some integers $\mu_1, \mu_2$. Given a relation of the form (4.11), we may divide both sides by $\gcd(\mu_1, \mu_2)^3$ to obtain an equivalent relation

$$y_1^2\lambda_2^3 = y_2^2\lambda_1^3$$

in which $(\lambda_1, \lambda_2) = 1$ and $(y_1, y_2) = 1$. This implies that for each $i = 1, 2$,

$$y_i^2 = \lambda_i^3. \tag{4.12}$$

This implies that $y_i$ is itself a perfect cube, say $y_i = s_i^3$. We recall from (4.10) that once $v_1, v_2, w_1, w_2$ are fixed, $u_1, u_2$ are fixed up to $X^{\varepsilon}$ choices. Thus we count how many $v_1, v_2 \leqslant V$ with $\gcd(v_1, v_2) = \delta$ are of the type (4.12) by noting that there are at most $O((V\delta^{-1})^{1/3})$ choices for each $s_i$. We bound the number of choices for $w_1, w_2$ trivially by $O(W^2)$, and conclude that the contribution to $T_3(\delta)$ of solutions for which a relation of the form (4.11) holds is at most

$$\ll W^2V^{2/3}\delta^{-2/3}X^{\varepsilon}. \tag{4.13}$$

We now proceed to count the remaining contribution to $T_3(\delta)$; we may assume from now on that no relation of the form (4.11) holds for $y_1$ and $y_2$. Define

$$k = y_2u_1 + y_1u_2. \tag{4.14}$$

Note that if $\delta, w_1, w_2, y_1, y_2$ and $k$ are fixed, then $u_1, u_2$ are fixed uniquely by (4.10). Thus we will count the number of solutions $w_1, w_2$ contributing to $T_3(\delta)$ for each fixed $y_1, y_2, k$.

Recalling the definition of $y_1, y_2$ we see that the condition (4.8) now becomes

$$y_1^2(4w_2^{\ell} - u_2^2) = y_2^2(4w_1^{\ell} - u_1^2) \neq 0,$$

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

and since $\gcd(y_1, y_2) = 1$, this implies a system of congruences

$$
\begin{aligned}
4y_2^2w_1^3 &\equiv k^2 \pmod{y_1}, && (4.15)\\
4y_1^2w_2^3 &\equiv k^2 \pmod{y_2}, && (4.16)\\
4y_2^2w_1^3 &\equiv 4y_1^2w_2^3 \pmod{k}. && (4.17)
\end{aligned}
$$

We first reduce this to a similar system of congruences with square-free moduli. For $i = 1, 2$ let $q_i$ denote the odd square-free kernel of $y_i$, that is

$$q_i = \prod_{\substack{p \mid y_i \\ p>2}} p.$$

The congruence (4.15) implies that $4y_2^2w_1^3 \equiv k^2 \pmod{q_1}$. Since $(4y_2, q_1) = 1$ this congruence may be re-written as $w_1^3 \equiv a_1 \pmod{q_1}$ for some constant $a_1$ determined by $y_2$ and $k$. A similar observation applies to (4.16). Next, we define

$$r = \prod_{\substack{p \mid k \\ p>2}} p$$

to be the odd square-free kernel of $k$, and deduce from (4.17) an analogous congruence modulo $r$. It follows that any solutions $w_1, w_2$ of the system (4.15)–(4.17) must satisfy the congruences

$$
\begin{aligned}
w_1^3 &\equiv a_1 \pmod{q_1}, && (4.18)\\
w_2^3 &\equiv a_2 \pmod{q_2}, && (4.19)\\
y_2^2w_1^3 &\equiv y_1^2w_2^3 \pmod{r} && (4.20)
\end{aligned}
$$

for some constant $a_1$ determined by $y_2, k \pmod{q_1}$ and some constant $a_2$ determined by $y_1$, $k \pmod{q_2}$.

Certainly $(q_1, q_2) = 1$. In addition, we note that $(y_1, r) = 1$ and $(y_2, r) = 1$. For indeed, if some odd prime $p$ satisfies $p \mid k$ and $p \mid y_1$, then by (4.14) it follows that $p \mid u_1$, since by construction $(y_1, y_2) = 1$. However, by the condition $v_1^2 \mid (4w_1^3 - u_1^2)$, this would imply that $p \mid w_1$, which contradicts the fact that $(v_1, w_1) = 1$. The fact that $(y_2, r) = 1$ may be shown similarly. As a consequence of these observations,

$$(q_1, q_2) = 1, \quad (q_1, r) = 1, \quad (q_2, r) = 1. \tag{4.21}$$

The next step is to note that the conditions (4.18)–(4.20) may be interpreted as lattice conditions.

Lemma 4.3. *The congruence (4.18) requires that $w_1$ lies in one of at most $3^{\omega(q_1)}$ residue classes modulo $q_1$, and similarly (4.19) requires that $w_2$ lies in one of at most $3^{\omega(q_2)}$ residue classes modulo $q_2$.*

*Furthermore, there exists a collection of at most $3^{\omega(r)}$ lattices $\Lambda_i \subset \mathbb{Z}^2$ of determinant $r$, such that any coprime pair $(w_1, w_2)$ satisfying (4.20) must lie in $\Lambda_i$ for some $i$. Conversely any pair $(w_1, w_2)$ in any of the lattices $\Lambda_i$ will satisfy (4.20).*

*Proof.* To prove this, we first consider the congruence (4.18). Fix a prime divisor $p \mid q_1$; then $w_1$ can only be a solution to (4.18) if

$$w_1^3 \equiv a_1 \pmod{p}. \tag{4.22}$$

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

There are at most three residue classes modulo $p$ in which a solution $w_1$ to (4.22) may lie. We may conclude that $w_1$ lies in one of at most $3^{\omega(q_1)}$ residue classes modulo $q_1$. A similar argument applies to (4.19), establishing that $w_2$ may lie in at most $3^{\omega(q_2)}$ residue classes modulo $q_2$.

We now turn to (4.20). Since $(y_1, r) = 1$ and $(w_1, w_2) = 1$ we must have $(w_1, r) = 1$. (Indeed, otherwise, if we suppose $p$ is a prime factor of both $w_1$ and $r$, we would see in (4.20) that $p \mid y_1^2 w_2^3$, but since $(w_1, w_2) = 1$ we cannot have $p \mid w_2$, so we would conclude $p \mid y_1$. This would in turn contradict that fact we previously proved that $(y_1, r) = 1$.) Using the fact that $(w_1, r) = 1$ we see that (4.20) implies

$$w^3 \equiv a \pmod{r}, \tag{4.23}$$

where $w \equiv w_2 w_1^{-1} \pmod{r}$ and $a \equiv (y_2 y_1^{-1})^2 \pmod{r}$ is coprime to $r$. Now, just as with our analysis of (4.18), we see that there is a collection of at most $3^{\omega(r)}$ residue classes $w \equiv b_i \pmod{r}$ in which $w$ must lie. This leads to a corresponding collection of lattice conditions $w_2 \equiv b_i w_1 \pmod{r}$ which, taken together, are equivalent to (4.23). Finally we note that the resulting lattice of pairs $(w_1, w_2)$ has a basis $\{(1, b_i), (0, r)\}$, so that its determinant is just $r$. This completes the proof of the lemma. □

### 4.3 Counting lattice points

Since $q_1, q_2, r$ are coprime in pairs, we may conclude from Lemma 4.3 that $(w_1, w_2)$ must lie in one of at most $3^{\omega(q_1)+\omega(q_2)+\omega(r)}$ lattice cosets of the form $(c_1, c_2) + \Lambda$, where $\Lambda$ is a lattice with $\det(\Lambda) = q_1 q_2 r$. We note that the total number of lattices is $\ll X^{\varepsilon}$, since under the assumption $Z \leqslant X$, we have $v_i \leqslant V \ll X^{5/2}$ and $k \leqslant 2UV \ll X^{11/2}$. We now fix one of these lattices, which we will denote by $\Lambda$, and its corresponding shift $(c_1, c_2)$. Note that we may choose $(c_1, c_2)$ such that $W \leqslant c_i < 4W$ for $i = 1, 2$, since otherwise $w_1, w_2$ would lie outside the desired range $W \leqslant w_1, w_2 < 4W$. We now write $(z_1, z_2) = (w_1, w_2) - (c_1, c_2)$, and proceed to count the number of

$$(z_1, z_2) \in \Lambda, \quad |z_i| < 3W.$$

Let $\lambda_1 \leqslant \lambda_2$ be the successive minima of $\Lambda$, so that the standard Minkowski inequalities show that $\det(\Lambda) \ll \lambda_1 \lambda_2 \ll \det(\Lambda)$ (see, for example, Davenport [Dav58, Eqn (5)]). We note that in our particular case,

$$\lambda_1 \ll \sqrt{\det(\Lambda)} \ll \sqrt{q_1 q_2 r} \ll V^{3/2} U^{1/2} \delta^{-3/2}. \tag{4.24}$$

Here we have used the fact that $q_i \leqslant y_i \leqslant V\delta^{-1}$ for $i = 1, 2$ and hence $r \leqslant k \ll UV\delta^{-1}$. Moreover, by Lemma 1 of Davenport [Dav58], the number of lattice points in $\Lambda$ with $|(z_1, z_2)| \leqslant x$ is (up to a constant) at most $(1 + x/\lambda_1)(1 + x/\lambda_2)$. Thus the number of allowable $z_1, z_2$ in our case is

$$\begin{aligned}
&\ll (1 + W/\lambda_1)(1 + W/\lambda_2) \\
&\ll 1 + W^2/\det(\Lambda) + W/\lambda_1 \\
&\ll 1 + W^2/(q_1 q_2 r) + W/\lambda_1.
\end{aligned}$$

Thus we have

$$T_3(\delta) \ll X^{\varepsilon} \sum_{y_1, y_2, k} \left(1 + \frac{W^2}{q_1 q_2 r} + \frac{W}{\lambda_1}\right), \tag{4.25}$$

where we recall that $q_i$ is the odd square-free kernel of $y_i$ and for each triple $y_1, y_2, k$ we take $\lambda_1$ to be the smallest value from all the corresponding lattices $\Lambda$. Recall that $y_1, y_2 \leqslant V\delta^{-1}$ and $k \leqslant 2UV\delta^{-1}$. Then we see that the contribution of the first term in (4.25) to $T_3(\delta)$ is at most

$$\ll X^{\varepsilon} V^3 U \delta^{-3}. \tag{4.26}$$

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

The contribution to $T_3(\delta)$ from the second term in (4.25) is

$$\ll X^{\varepsilon} W^2 \bigg(\sum_{y_1 \leqslant V\delta^{-1}} \frac{1}{q_1}\bigg)\bigg(\sum_{y_2 \leqslant V\delta^{-1}} \frac{1}{q_2}\bigg)\bigg(\sum_{k \leqslant 2UV\delta^{-1}} \frac{1}{r}\bigg). \tag{4.27}$$

To bound each internal sum we apply the following minor variant of [HB07, Lemma 1].

Lemma 4.4. *Given an integer $k$, let $k^*$ denote its odd square-free kernel. For any fixed integer $\kappa \leqslant K$,*

$$\#\{k \leqslant K : k^* = \kappa\} \ll K^{\varepsilon}.$$

We defer the proof of this lemma until § 4.4, and merely apply it now to (4.27); for example the first sum is bounded by

$$\sum_{y_1 \leqslant V\delta^{-1}} \frac{1}{q_1} \leqslant \sum_{\nu \leqslant V\delta^{-1}} \frac{1}{\nu} \#\{v \leqslant V\delta^{-1} : v^* = \nu\} \ll V^{\varepsilon} \sum_{\nu \leqslant V\delta^{-1}} \frac{1}{\nu} \ll V^{\varepsilon}.$$

One may handle the second and third sums in (4.27) similarly, and deduce that the second term in (4.25) is $O(W^2 X^{\varepsilon})$ overall. Since $W^2 \leqslant W^2 V^{2/3} \delta^{-2/3}$ for $\delta \leqslant V$ we see that this error is dominated by (4.13).

Finally, the contribution, say $T_3'(\delta)$, of the third term in (4.25) may be bounded by following the same argument as in [HB07], which we sketch for completeness. For each triple $y_1, y_2, k$, let $\Lambda$ be the lattice to which $\lambda_1$ corresponds, and let $(\mu_1, \mu_2)$ be the shortest nonzero vector in $\Lambda$, so that $\lambda_1$ is the length of $(\mu_1, \mu_2)$. Then

$$T_3'(\delta) \ll X^{\varepsilon} W \sum_{\mu_1, \mu_2} \frac{\#\{y_1, y_2, k\}}{\sqrt{|\mu_1|^2 + |\mu_2|^2}},$$

where we count the number of $y_1, y_2, k$ that generate a lattice in which $(\mu_1, \mu_2)$ is a vector of minimal length. We note by (4.24) that

$$\mu_1, \mu_2 \ll V^{3/2} U^{1/2} \delta^{-3/2}. \tag{4.28}$$

Since $(\mu_1, \mu_2)$ lies in the lattice $\Lambda$, then by construction

$$q_1 \mid \mu_1, \quad q_2 \mid \mu_2 \tag{4.29}$$

and

$$r \mid (y_2^2 \mu_1^3 - y_1^2 \mu_2^3), \tag{4.30}$$

as described in Lemma 4.3.

We first consider the case where both $\mu_1, \mu_2$ are nonzero. By (4.29), once $\mu_1, \mu_2$ are fixed, they determine at most $X^{\varepsilon}$ values of $q_1, q_2$ and hence at most $X^{\varepsilon}$ values for $y_1, y_2$ by Lemma 4.4. If $y_2^2 \mu_1^3 - y_1^2 \mu_2^3$ is nonzero, then it determines at most $X^{\varepsilon}$ values for $r$ by (4.30) and hence at most $X^{\varepsilon}$ values for $k$. On the other hand, if

$$y_2^2 \mu_1^3 = y_1^2 \mu_2^3, \tag{4.31}$$

then $y_1, y_2$ would satisfy a relation of the form (4.11); pairs $y_1, y_2$ of this type have already been treated, and are excluded from the contribution we are currently calculating. We therefore see

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

that the contribution to $T_3'(\delta)$ from $\mu_1, \mu_2$ both nonzero is

$$T_3'(\delta) \ll X^{4\varepsilon} W \sum_{\mu_1,\mu_2} \frac{1}{\sqrt{|\mu_1|^2 + |\mu_2|^2}}.$$

To bound the sum, we begin by focusing on a fixed dyadic range

$$\tfrac{1}{2}B < \sqrt{|\mu_1|^2 + |\mu_2|^2} \leqslant B,$$

for any appropriate $B \geqslant 1$; we note that the restriction (4.28) implies that $B \ll V^{3/2}U^{1/2}\delta^{-3/2}$. There are $O(B^2)$ pairs $\mu_1, \mu_2$, each of which contribute $O(B^{-1})$ to the sum. Summing over dyadic $B \ll V^{3/2}U^{1/2}\delta^{-3/2}$ therefore produces a total contribution of $\ll X^{\varepsilon}WV^{3/2}U^{1/2}\delta^{-3/2}$ to $T_3'(\delta)$.

On the other hand if $\mu_1$ vanishes, then there are $V\delta^{-1}$ choices for $y_1$ and $O(X^{2\varepsilon})$ choices for $q_2, r$, hence $O(X^{4\varepsilon})$ choices for $y_2, k$. (In particular, (4.31) cannot occur, since it would force $\mu_1 = \mu_2 = 0$.) Thus the contribution from these terms to $T_3'(\delta)$ is

$$\ll X^{5\varepsilon} V W \delta^{-1} \sum_{\mu_2 \ll V^{3/2}U^{1/2}\delta^{-3/2}} \frac{1}{|\mu_2|} \ll X^{6\varepsilon} V W \delta^{-1}.$$

The case where $\mu_2$ vanishes may be treated by an analogous argument. We may conclude that

$$T_3'(\delta) \ll X^{\varepsilon}(WV^{3/2}U^{1/2}\delta^{-3/2} + VW\delta^{-1}).$$

Combining this with the contributions (4.13) and (4.26) shows that

$$T_3(\delta) \ll X^{\varepsilon}(W^2V^{2/3}\delta^{-2/3} + V^3U\delta^{-3} + WV^{3/2}U^{1/2}\delta^{-3/2} + VW\delta^{-1}).$$

Since

$$\begin{aligned} WV^{3/2}U^{1/2}\delta^{-3/2} &= \{W^2\}^{1/2}\{V^3U\delta^{-3}\}^{1/2} \\ &\leqslant \{W^2V^{2/3}\delta^{-2/3}\}^{1/2}\{V^3U\delta^{-3}\}^{1/2} \end{aligned}$$

for $\delta \leqslant V$, the third term above is dominated by the first two, so that Proposition 4.2 follows. □

### 4.4 Proof of Lemma 4.4

We now prove Lemma 4.4, in the following more general form. Given any finite set $\mathcal{P}$ of primes (possibly empty), let

$$k(\mathcal{P}) = \prod_{\substack{p \mid k \\ p \notin \mathcal{P}}} p.$$

Consider the set $\{k \leqslant K : k(\mathcal{P}) = \kappa\}$ for a fixed positive integer $\kappa$. The set is empty unless $\kappa \leqslant K$ is square-free and satisfies $(\kappa, \prod_{p \in \mathcal{P}} p) = 1$, which we now assume. Then for any $\eta > 0$,

$$\begin{aligned} \#\{k \leqslant K : k(\mathcal{P}) = \kappa\} &\leqslant \sum_{\substack{k=1 \\ k(\mathcal{P})=\kappa}}^{K} \left(\frac{K}{k}\right)^{\eta} \\ &\leqslant K^{\eta} \sum_{\substack{k=1 \\ k(\mathcal{P})=\kappa}}^{\infty} k^{-\eta} \\ &= K^{\eta} \prod_{p \in \mathcal{P}} \left(\sum_{e=0}^{\infty} p^{-e\eta}\right) \prod_{p \mid \kappa} \left(\sum_{e=1}^{\infty} p^{-e\eta}\right). \end{aligned}$$

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

Setting $A(\eta)=\sum_{e=0}^{\infty}2^{-e\eta}$ we then see that

$$\#\{k\leqslant K: k(\mathcal{P})=\kappa\}\leqslant K^{\eta}A(\eta)^{\omega(\kappa)+\#\mathcal{P}}\leqslant K^{\eta}A(\eta)^{(\#\mathcal{P}+1)\omega(\kappa)}.$$

Upon recalling that $\omega(\kappa)\ll(\log 3\kappa)(\log\log 3\kappa)^{-1}$ and $\kappa\leqslant K$ we may conclude that

$$\#\{k\leqslant K: k(\mathcal{P})=\kappa\}\ll_{\eta}K^{(\#\mathcal{P}+2)\eta}$$

for any $\eta>0$, which proves Lemma 4.4. □

## 5. Average of $h_\ell(-d)$

We now turn to applications of the key propositions. We first apply Proposition 2.1 to derive a nontrivial upper bound for the average of $h_\ell(-d)$. Fix a dyadic region $X\leqslant d<2X$ and assume that $X^{1/(2\ell)}\leqslant Z\leqslant X$. Then Proposition 2.1 implies that

$$\sum_{X\leqslant d<2X}h_\ell(-d)\ll X^{\varepsilon}\bigg\{X^{1/2}\#E(Z;X)+X^{3/2}Z^{-1}+X^{1/2}Z^{-2}\sum_{\substack{X\leqslant d<2X\\ d\notin E(Z;X)}}S_\ell(d;Z)\bigg\}.$$

We apply the upper bound (2.1) to the exceptional set $E(Z;X)$ and Proposition 2.3 to the average of $S_\ell(d;Z)$ to conclude that

$$\sum_{X\leqslant d<2X}h_\ell(-d)\ll X^{\varepsilon}\{X^{3/2}Z^{-1}+X+Z^{\ell}\}.$$

It is optimal to choose $Z=X^{3/(2\ell+2)}$, resulting in

$$\sum_{X\leqslant d<2X}h_\ell(-d)\ll X^{3/2-3/(2\ell+2)+\varepsilon}.$$

Summing over $O(\log X)$ dyadic intervals to cover the full range $0<d<X$ then yields the result of Theorem 1.1.

## 6. Higher moments of $h_\ell(-d)$

We now consider higher moments. For any odd prime $\ell$, define for any real $H\geqslant 1$ the set

$$A_\ell(H;X)=\{X\leqslant d<2X: h_\ell(-d)>H\},$$

with corresponding counting function

$$N_\ell(H;X)=\#A_\ell(H;X).$$

We also define for any $\frac{1}{4}X^{1/2\ell}\leqslant Z\leqslant X$ the set

$$A_\ell^0(H,Z;X)=\{X\leqslant d<2X: h_\ell(-d)>H\}\backslash E(Z;X),$$

where $E(Z;X)$ is as usual the exceptional set provided by Proposition 2.1. We define the corresponding counting function

$$N_\ell^0(H,Z;X)=\#A_\ell^0(H,Z;X).$$

We note that for any fixed choice of $Z$ in the above range,

$$N_\ell(H;X)\leqslant\#E(Z;X)+N_\ell^0(H,Z;X)\ll X^{\varepsilon}+N_\ell^0(H,Z;X). \tag{6.1}$$

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

### 6.1 The case $\ell = 3$

Restricting to the case $\ell = 3$, we see that (1.4) implies that

$$N_3(H;X) \ll XH^{-1}. \tag{6.2}$$

We also note that $A_3(H;X)$ is empty by (1.2) unless $H \leqslant X^{1/3+\varepsilon}$ for some small $\varepsilon > 0$. In general we have the following.

PROPOSITION 6.1. *For* $1 \leqslant H \leqslant X^{1/3+\varepsilon}$,

$$N_3(H;X) \ll X^{\varepsilon}(X^{1/2} + X^{7/2}H^{-10}).$$

*Proof.* To prove this we consider $A_3^0(H,Z;X)$ with the choice $Z = X^{1/2+2\varepsilon}H^{-1}$; note in particular $Z \geqslant X^{1/6}$ when $H \leqslant X^{1/3+\varepsilon}$. Moreover we will have

$$h_3(-d) > H \gg d^{1/2+\varepsilon}Z^{-1}$$

for all $d$ in $A_3^0(H,Z;X)$, whence Proposition 2.1 shows that

$$h_3(-d) \ll d^{1/2+\varepsilon}Z^{-2}S_3(d;Z).$$

We therefore have

$$S_3(d;Z) \gg d^{-1/2-\varepsilon}Z^2h_3(-d) \gg X^{-1/2-\varepsilon}Z^2h_3(-d) > X^{-1/2-\varepsilon}Z^2H,$$

for all $d \in A_3^0(H,Z;X)$. This leads to the bound

$$N_3^0(H;X)(X^{-1/2-\varepsilon}Z^2H)^2 \ll \sum_{d\in A_3^0(H,Z;X)} S_3(d;Z)^2 \ll \sum_{X\leqslant d<2X} S_3(d;Z)^2.$$

We can now apply the case $\ell = 3$ of Proposition 2.4 to obtain

$$N_3^0(H,Z;X)(X^{-1/2-\varepsilon}Z^2H)^2 \ll X^{\varepsilon}\{Z^2X^{1/2} + Z^{12}X^{-3/2}\},$$

so that

$$N_3^0(H,Z;X) \ll X^{3\varepsilon}H^{-2}\{Z^{-2}X^{3/2} + Z^8X^{-1/2}\} \ll X^{19\varepsilon}\{X^{1/2} + X^{7/2}H^{-10}\}$$

in view of our choice of $Z$. This is sufficient to prove Proposition 6.1, by (6.1). □

*Proof of Theorem 1.2.* We may now derive Theorem 1.2 from Proposition 6.1. It will suffice to consider a dyadic range $X \leqslant d < 2X$. Then

$$\begin{aligned}\sum_{X\leqslant d<2X} h_3(-d)^k &\ll \sum_{\substack{H\leqslant X^{1/3+\varepsilon}\\ \text{dyadic}}} \sum_{\substack{X\leqslant d<2X\\ H<h_3(-d)\leqslant 2H}} h_3(-d)^k \\ &\leqslant \sum_{\substack{H\leqslant X^{1/3+\varepsilon}\\ \text{dyadic}}} N_3(H;X)(2H)^k.\end{aligned}$$

In view of (6.2) we have

$$N_3(H;X)(2H)^k \ll XH^{k-1}.$$

On the other hand, Proposition 6.1 yields

$$N_3(H;X)(2H)^k \ll X^{\varepsilon}(X^{1/2}H^k + X^{7/2}H^{k-10}).$$

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

In particular for $k = 4$ we deduce that

$$\begin{aligned} N_3(H;X)(2H)^4 &\ll X^{\varepsilon}\min\{XH^3, X^{1/2}H^4 + X^{7/2}H^{-6}\} \\ &\ll X^{\varepsilon}\min\{XH^3, X^{1/2}H^4\} + \min\{XH^3, X^{7/2}H^{-6}\}. \end{aligned}$$

For $H \leqslant X^{1/3+\varepsilon}$ the first term is at most

$$X^{1/2}H^4 \leqslant X^{11/6+4\varepsilon}$$

while the second term is at most

$$\{XH^3\}^{2/3}\{X^{7/2}H^{-6}\}^{1/3} = X^{11/6}.$$

It follows that $N_3(H;X)(2H)^4 \ll X^{11/6+4\varepsilon}$, whence

$$\sum_{X\leqslant d<2X} h_3(-d)^4 \ll X^{11/6+5\varepsilon}.$$

This suffices to prove Theorem 1.2. □

As noted in the introduction, one can deduce estimates for other moments from the fourth moment. The reader may check that a direct application of the methods of this section to the general moment only reproduces these consequences of the special case $k = 4$.

### 6.2 The case $\ell \geqslant 5$

We now consider the $k$th moment of $h_\ell(-d)$ for primes $\ell \geqslant 5$ and any real $k \geqslant 1$. By Corollary 2.2 we see that

$$N_\ell(H;X) \ll X^{\varepsilon} \quad \text{if } H \geqslant X^{1/2-1/2\ell+\varepsilon}.$$

We also record the trivial bound

$$N_\ell(H;X) \ll X, \tag{6.3}$$

valid for all $H$. In addition, we claim the following.

Proposition 6.2. *For any prime $\ell \geqslant 3$ and $1 \leqslant H \leqslant X^{1/2-1/(2\ell)+\varepsilon}$,*

$$N_\ell(H;X) \ll X^{\varepsilon}(XH^{-1} + X^{\ell/2}H^{-(\ell+1)}).$$

With Proposition 6.2 in hand, we will prove the following.

Proposition 6.3. *For any prime $\ell \geqslant 5$ and any real number $k \geqslant 1$,*

$$\sum_{X\leqslant d<2X} h_\ell(-d)^k \ll X^{\sigma+\varepsilon},$$

*where*

$$\sigma = \max\{\sigma_1, \sigma_2, \sigma_3\}$$

*and*

$$\begin{aligned} \sigma_1 &= 1 + k\left(\frac{\ell-2}{2\ell+2}\right), \\ \sigma_2 &= 1 + k\left(\frac{\ell-1}{2\ell}\right) - \left(\frac{\ell-1}{2\ell}\right), \\ \sigma_3 &= \frac{k}{2}. \end{aligned}$$

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

We note that the maximum is $\sigma_1$ in the range $1 \leqslant k \leqslant (\ell^2-1)/(2\ell-1)$; it is $\sigma_2$ in the range $(\ell^2-1)/(2\ell-1) \leqslant k \leqslant \ell+1$; and it is $\sigma_3$ for $k \geqslant \ell+1$. This leads immediately to the statement of Theorem 1.5. We note that Proposition 6.2 does not imply any new results in the case of $h_3(-d)$.

*Proof of Proposition 6.2.* The proof of Proposition 6.2 follows similar lines to that of Proposition 6.1. As before we set $Z = X^{1/2+2\varepsilon}H^{-1}$, so that $Z \geqslant X^{1/(2\ell)}$ for $H \leqslant X^{1/2-1/(2\ell)+\varepsilon}$. We deduce that

$$S_\ell(d;Z) \gg d^{-1/2-\varepsilon}Z^2h_\ell(-d) \gg X^{-1/2-\varepsilon}Z^2h_\ell(-d) > X^{-1/2-\varepsilon}Z^2H,$$

again under the assumption that $d \in A_\ell^0(H,Z;X)$. As a result,

$$N_\ell^0(H,Z;X)X^{-1/2-\varepsilon}Z^2H \ll \sum_{d\in A_\ell^0(H,Z;Z)} S_\ell(d;Z) \ll \sum_{X\leqslant d<2X} S_\ell(d;Z).$$

Upon applying Proposition 2.3 we obtain

$$N_\ell^0(H,Z;X)X^{-1/2-\varepsilon}Z^2H \ll X^\varepsilon(Z^2X^{1/2}+Z^{\ell+2}X^{-1/2}),$$

so that

$$N_\ell^0(H,Z;X) \ll X^{2\varepsilon}(XH^{-1}+Z^\ell H^{-1}) \ll X^{(2+2\ell)\varepsilon}(XH^{-1}+X^{\ell/2}H^{-(\ell+1)}),$$

upon recalling the choice of $Z$. This is sufficient to prove Proposition 6.2, by (6.1). □

*Proof of Proposition 6.3.* We turn finally to Proposition 6.3, for which we initially fix any real number $k \geqslant 1$. We have already observed that $N_\ell(H;X) \ll X^\varepsilon$ if

$$X^{1/2-1/(2\ell)+\varepsilon} \leqslant H \leqslant X^{1/2+\varepsilon},$$

which shows that for such $H$,

$$N_\ell(H;X)H^k \ll X^{k/2+\varepsilon}. \tag{6.4}$$

Thus we now instead assume that

$$H \leqslant X^{1/2-1/(2\ell)+\varepsilon}. \tag{6.5}$$

Then by the trivial bound (6.3) and Proposition 6.2 we have

$$\begin{aligned}N_\ell(H;X)H^k &\ll X^\varepsilon \min\{XH^k, XH^{k-1}+X^{\ell/2}H^{k-\ell-1}\}\\ &\ll X^\varepsilon(XH^{k-1}+\min\{XH^k, X^{\ell/2}H^{k-\ell-1}\}).\end{aligned}$$

Under (6.5), the first term is $\ll X^{\sigma_2}$. As long as $k \leqslant \ell+1$, the second term is largest when $XH^k = X^{\ell/2}H^{k-\ell-1}$, namely when

$$H = X^{(\ell-2)/(2\ell+2)} = X^{1/2-3/(2\ell+2)}.$$

We may conclude that if $k \leqslant \ell+1$ and $H \leqslant X^{1/2-1/(2\ell)+\varepsilon}$ then

$$N_\ell(H;X)H^k \ll X^\varepsilon(X^{\sigma_1}+X^{\sigma_2}),$$

with the notation of Proposition 6.3. On the other hand, if $k \geqslant \ell+1$ then

$$X^{\ell/2}H^{k-\ell-1} \leqslant X^{\ell/2}H^{k-\ell} \leqslant X^{\ell/2}(X^{1/2})^{k-\ell} = X^{k/2}.$$

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

Thus $N_\ell(H;X)H^k \ll X^\varepsilon(X^{\sigma_2} + X^{k/2})$ in this case; note that the second term dominates in the range $k \geqslant \ell + 1$. To conclude, we have

$$N_\ell(H;X)H^k \ll X^\varepsilon(X^{\sigma_1} + X^{\sigma_2} + X^{\sigma_3}) \tag{6.6}$$

for all $k \geqslant 1$.

Combining (6.4) and (6.6) shows that

$$\begin{aligned}\sum_{X \leqslant d < 2X} h_\ell(-d)^k &\ll \sum_{\substack{H \ll X^{1/2+\varepsilon} \\ \text{dyadic}}} \sum_{\substack{X \leqslant d < 2X \\ H < h_\ell(-d) \leqslant 2H}} h_\ell(-d)^k \\ &\leqslant \sum_{\substack{H \ll X^{1/2+\varepsilon} \\ \text{dyadic}}} N_\ell(H;X)(2H)^k \\ &\ll X^\varepsilon(X^{\sigma_1} + X^{\sigma_2} + X^{\sigma_3}).\end{aligned}$$

We note that $k/2 \leqslant \max\{\sigma_1, \sigma_2\}$ in the range $k \leqslant \ell + 1$. This proves Proposition 6.3, and hence Theorem 1.5. □

The reader may verify that a similar computation based on Proposition 2.4 yields no improvements.

## Acknowledgements

The authors thank Peter Sarnak for asking a question that spurred this line of enquiry. The first author was supported by EPSRC grant number EP/K021132X/1. The second author was partially supported by NSF DMS-1402121, and thanks the Hausdorff Center for Mathematics for a very pleasant working environment.

## References

BST13 M. Bhargava, A. Shankar and J. Tsimerman, *On the Davenport–Heilbronn theorem and second order terms*, Invent. Math. **193** (2013), 439–499.

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

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

Dav58 H. Davenport, *Indefinite quadratic forms in many variables II*, Proc. Lond. Math. Soc. (3) **8** (1958), 109–126.

Dav00 H. Davenport, *Multiplicative number theory*, Graduate Texts in Mathematics, vol. 74, third edition (Springer, New York, 2000).

DH71 H. Davenport and H. Heilbronn, *On the density of discriminants of cubic fields II*, Proc. R. Soc. Lond. A **322** (1971), 405–420.

Duk98 W. Duke, *Bounds for arithmetic multiplicities*, in *Proc. Int. Congress of Mathematicians, Berlin, 1998*, Doc. Math., Extra Volume II (1998), 163–172.

EV07 J. S. Ellenberg and A. Venkatesh, *Reflection principles and bounds for class group torsion*, Int. Math. Res. Not. IMRN **2007** (2007), rnm002.

HB07 D. R. Heath-Brown, *Quadratic class numbers divisible by 3*, Funct. Approx. Comment. Math. **37** (2007), 203–211.

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

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press

Hou10 R. Hough, *Average equidistribution of Heegner points associated to the 3-part of the class group of imaginary quadratic fields*, Preprint (2010), arXiv:1005.1458v2.

Sch32 A. Scholz, *Über die Beziehung der Klassenzahlen quadratischer Körper*, J. Reine Angew. Math. **166** (1932), 201–203.

Sou00 K. Soundararajan, *Divisibility of class numbers of imaginary quadratic fields*, J. Lond. Math. Soc. (2) **61** (2000), 681–690.

TT13 T. Taniguchi and F. Thorne, *The secondary term in the counting function for cubic fields*, Duke Math. J. **162** (2013), 2451–2508.

Zha05 S.-W. Zhang, *Equidistribution of CM-points on quaternion Shimura varieties*, Int. Math. Res. Not. IMRN **2005** (2005), 3657–3689.

D. R. Heath-Brown rhb@maths.ox.ac.uk
Mathematical Institute, Radcliffe Observatory Quarter, Woodstock Road,
Oxford OX2 6GG, UK

L. B. Pierce pierce@math.duke.edu
Department of Mathematics, Duke University, Durham NC 27708, USA

https://doi.org/10.1112/S0010437X1700728X Published online by Cambridge University Press