ESTIMATES FOR REPRESENTATION NUMBERS OF QUADRATIC FORMS

VALENTIN BLOMER and ANDREW GRANVILLE

Abstract

Let f be a primitive positive integral binary quadratic form of discriminant $-D$, and let $r_f(n)$ be the number of representations of n by f up to automorphisms of f. In this article, we give estimates and asymptotics for the quantity $\sum_{n\le x} r_f(n)^\beta$ for all $\beta \ge 0$ and uniformly in $D = o(x)$. As a consequence, we get more-precise estimates for the number of integers which can be written as the sum of two powerful numbers.

Contents

DUKE MATHEMATICAL JOURNAL
Vol. 135, No. 2, © 2006
Received 18 April 2005. Revision received 3 May 2006.
2000 *Mathematics Subject Classification*. Primary 11E16; Secondary 11N56.
Blomer's work partially supported by Natural Sciences and Engineering Research Council of Canada grant A 5123.
Granville's work partially supported by Natural Sciences and Engineering Research Council of Canada Individual Discovery Grant.

1. Introduction and statement of results

1.1. Overview of the main questions and results

Let $-D < 0$ be a fundamental discriminant. Bernays [2, pages 91−92], generalizing Landau's famous result for $-D = -4$, showed that there are $\sim \kappa_D x/\sqrt{\log x}$ distinct integers $\le x$ which are represented by any given binary quadratic form $f(y, z) := ay^2 + byz + cz^2$ of discriminant $-D$, as $x \to \infty$; here κ_D depends only on D and not on f and is neither very big nor very small. (In fact, $\kappa_D = D^{o(1)}$.)

On the other hand, it is easy to show that there are $\ll x/\sqrt{D}$ pairs of integers m, n for which $|f(m, n)| \le x$; and since $x/\sqrt{D} = o(D^{o(1)}x/\sqrt{\log x})$ whenever $D > (\log x)^{1+\epsilon}$, we see that Bernays's asymptotic cannot hold until x is surprisingly large; that is, $x > \exp(D^{1-\epsilon})$. This is quite different from linear forms, in which case the formula $\#\{n : 0 \le a + nD \le x\} \sim x/D$ holds once $x/D \to \infty$. It is also quite different from the number of primes represented by f; this count should settle down to its asymptotic formula once x is larger than some fixed power of D (assuming a suitable Riemann hypothesis).

Our main concern is to give good estimates for the number of distinct integers at most x which are represented by f in all ranges of x with a particular focus on the transitional ranges, where x goes from $\exp(D^\epsilon)$ to $\exp(D^N)$, where $\epsilon > 0$ is small and N is large, determining how this count changes behaviour. Let h be the class number of $\mathbb{Q}(\sqrt{-D})$, and let g be the number of genera. We let $\ell = \ell_{-D} := L(1, \chi_{-D})(\phi(D)/D)$, and we create a parameter

$$\kappa := \frac{\log(h/g)}{(\log 2)(\log(\ell_{-D} \log x))}$$

which is suitable for measuring this transition since $h/g = D^{1/2+o(1)}$. We believe that estimates for

$$N_f(x) = \#\bigl\{n \le x : n = f(m_1, m_2) \text{ for some integers } m_1, m_2\bigr\}$$

should be split into three ranges:

$$N_f(x) \asymp \frac{L(1, \chi_D)}{\tau(D)} \frac{x}{\sqrt{\ell_{-D} \log x}} \quad \text{for } 0 \le \kappa \le 1/2, \tag{1.1}$$

an extension of the range of Bernays's result;

$$N_f(x) \asymp \frac{L(1, \chi_{-D})}{\tau(D)} \frac{(\ell_{-D} \log x)^{-1+\kappa(1-\log(2\kappa))}}{(1 + (\kappa - 1/2)(1 - \kappa)\sqrt{\log\log x})}\, x \quad \text{for } 1/2 < \kappa < 1, \tag{1.2}$$

the difficult intermediate range; and

$$N_f(x) \asymp \frac{x}{\sqrt{D}} \quad \text{for } 1 \le \kappa \ll \frac{\log D}{\log\log D}, \tag{1.3}$$

where estimates can be obtained by elementary counting arguments. We prove these estimates (in Theorem 6), except when $\kappa \in [1/2 - \varepsilon, 1/(\log 2) + \varepsilon]$, and then it is only the lower bound that yet needs to be established.

The first author studied these questions in [3] in order to deduce that there are $\asymp x/(\log x)^{(1-2^{-1/3})+o(1)}$ integers at most x which can be written as the sum of two powerful numbers (n is a powerful number if $p^2 \mid n$ whenever a prime $p \mid n$). We now can conjecture that there are

$$\asymp \frac{x(\log\log x)^{2^{2/3}-1}}{(\log x)^{(1-2^{-1/3})}}$$

such integers and prove the upper bound in this article, failing to obtain the lower conjectured bound by a power of $\log\log x$.

As can be seen from (1.1)–(1.3), we have been able to count accurately the number of distinct integers represented by f except in the difficult intermediate range. It is also of interest to understand how many times each of the distinct integers are actually represented by f. Thus we also focus on establishing sharp bounds and asymptotics for the quantity

$$\sum_{n \le x} r_f(n)^\beta \quad \text{for fixed } \beta > 0,$$

where $r_f(n)$ is the number of inequivalent representations of integer n by f, uniformly in $D = o(x)$.* To our surprise, we have been able to obtain precise results in all interesting ranges when β is a positive integer.

COROLLARY 1
Let $\beta \ge 1$ be an integer, and set $K = 2^{\beta-1}$. For a given binary quadratic form f, let u be the smallest positive integer that can be represented by some form in the coset $f\mathfrak{G}$. We have

$$\sum_{n \le x} r_f(n)^\beta = \left(a_K(\log x)^{K-1} + \frac{\pi}{\sqrt{D}}\Big(1 + \frac{2^{\beta-1}-1}{u}\Big)\right) x\big(1 + O_{\beta,\rho}((\log x)^{-\rho})\big) \tag{1.4}$$

*The proof of Lemma 3.1 shows that there is no smooth estimate for the above quantity if $x \ll c$; since the coefficient c of f can be $\asymp D$ but no larger, the range $D = o(x)$ is the natural one.

uniformly in $x \geq D(\log D)^{2\rho}/a$ *for any* $0 < \rho < 1/3$ *if* $\beta = 2$*, for* $0 < \rho < 1/2$ *if* $\beta = 3$*, and for* $0 < \rho < 1$ *for all other* β*, where* $a_K = 0$ *if* $\beta = 1$*, and otherwise,*

$$a_K = \frac{C_{D,\beta}}{\Gamma(K)} \frac{g^{\beta-1}}{h^\beta} L(1, \chi_{-D})^{2^{\beta-1}} \prod_{p|D} \Big(1 - \frac{1}{p}\Big)^{2^{\beta-1}-1} \tag{1.5}$$

with

$$C_{D,\beta} = \prod_{\chi_{-D}(p)=1} \Big(\sum_{k=0}^{\infty} \frac{(k+1)^\beta}{p^k}\Big)\Big(1 - \frac{1}{p}\Big)^{2^\beta} \prod_{\chi_{-D}(p)=-1} \Big(1 - \frac{1}{p^2}\Big)^{2^{\beta-1}-1}. \tag{1.6}$$

The range here is easily extended to $x/D \to \infty$ at the cost of a weaker error term (see Theorem 2). Note that $C_{D,\beta} \asymp_\beta 1$. The second main term on the right-hand side of (1.4) dominates when

$$(\log x)^{(2^\beta-2)/(\beta-1)+o(1)} \leq D = o(x)$$

and the first main term dominates in the complementary range. Below we give better estimates in all ranges (Theorems 1 and 2), and we prove slightly weaker estimates for arbitrary real $\beta \geq 0$ (Theorems 3, 4, and 5). Here we define $r_f(n)^0 = 0$ if $r_f(n) = 0$ and $r_f(n)^0 = 1$ otherwise; that is, the case where $\beta = 0$ corresponds to estimating $N_f(x)$.

1.2. Statement of the main theorems

It is well known (see, e.g., [6]) that there is a one-to-one correspondence between the equivalence classes of integral ideals in $\mathbb{Q}(\sqrt{-D})$ and (proper) equivalence classes of primitive positive binary quadratic forms of discriminant $-D$. We denote either set by $\mathfrak{C}$, let $h := \#\,\mathfrak{C}$ be the class number, and let $\widehat{\mathfrak{C}}$ be the set of class group characters $\chi : \mathfrak{C} \to \mathbb{C}^*$. Let $\mathfrak{G}$ be the subgroup of ambiguous classes of forms (i.e., having order at most 2) so that $\mathfrak{G} \cong \mathfrak{C}/\mathfrak{C}^2$; that is, $\mathfrak{G}$ is isomorphic to the group of genera. It is well known that

$$D^{1/2-\varepsilon} \ll h \ll D^{1/2} \log D \quad \text{and} \quad g = 2^{\omega(D)-1} \ll D^\varepsilon. \tag{1.7}$$

Two representations $\mathbf{x}_1, \mathbf{x}_2$ are equivalent if $\mathbf{x}_1 = A\mathbf{x}_2$ for an automorphism $A \in \mathrm{SL}_2(\mathbb{Z})$ of f; and the number of automorphisms of f equals the number $w \in \{2, 4, 6\}$ of units in the ring of integers $\mathfrak{o}$ of $\mathbb{Q}(\sqrt{-D})$. Therefore we define $r_f(n) = \#\{\mathbf{x} \in \mathbb{Z}^2 \mid f(\mathbf{x}) = n\}/w$.

THEOREM 1
Let $\beta \geq 1$ be an integer, and let $K = 2^{\beta-1}$. There exist constants a_k depending on β and f such that

$$\sum_{n \leq x} r_f(n)^\beta = \frac{x}{\log x} \sum_{k=1}^{K} a_k (\log x)^k + O_{\beta,\varepsilon}(D^{2^{\beta-3}/(2^{\beta-2}+1)} x^{1-1/(2^{\beta-1}+2)+\varepsilon}) \tag{1.8}$$

uniformly in $D \leq x^{1/2^{\beta-2}-\varepsilon}$ for any $\varepsilon > 0$. Precisely, a_K is given by (1.5) and (1.6), and

$$a_1 = \frac{\pi}{\sqrt{D}}\Big(1 + \frac{2^{\beta-1}-1}{u}\Big)\big(1 + O_{\beta,\varepsilon}(D^{-1/4+\varepsilon})\big) \tag{1.9}$$

for $\beta > 1$ and any $\varepsilon > 0$, where u is the smallest positive integer that can be represented by some form in the coset $f\mathfrak{G}$; and if $2^{j-1} < k \leq 2^j$, then

$$a_k \ll_\varepsilon D^{-(j+1)/2+\varepsilon}. \tag{1.10}$$

The Dirichlet series $\sum r_f(n)^\beta n^{-s}$ can be analytically continued to the region $\{s \in \mathbb{C} \setminus \{1\} \mid \mathrm{Re}\, s > 1/2\}$ and has a pole of order K at $s = 1$.

We also have more-precise estimates for small x, proved via elementary methods, which work for all $\beta \geq 0$.

THEOREM 2
For a given binary quadratic form f, let a be the smallest positive integer that is represented by f, and let u be the smallest positive integer that can be represented by some form in the coset $f\mathfrak{G}$. For any $\beta \geq 0$, we have

$$\sum_{n \leq x} r_f(n)^\beta = \pi\Big(1 + \frac{2^{\beta-1}-1}{u}\Big)\frac{x}{\sqrt{D}} + E_\beta(x, D), \tag{1.11}$$

where

$$E_\beta(x, D) \ll \begin{cases} \sqrt{\frac{x}{a}} + \tau(D)\big(\frac{x \log x}{D} + \frac{x}{D^{3/4}}\big), & 0 \leq \beta \leq 2, \\ \sqrt{\frac{x}{a}} + \tau(D)\frac{x(\log x)^{(2/q)(2^{(\beta-2)q+1}-1)+1}}{D^{(3/4)(1-1/q)}}, & \beta > 2, \end{cases}$$

for any real $q > 1$, where $\tau(D)$ denotes the number of divisors of D. The implied constants depend at most on β and q.

The proof yields that $r_f(n) = 1$ for $\pi(1 - 1/u)x/\sqrt{D} + O(E_2(x, D))$ integers $n \leq x$; that $r_f(n) = 2$ for $\pi x/(2u\sqrt{D}) + O(E_2(x, D))$ integers, $n \leq x$; and that $r_f(n) \geq 3$ for $O(E_2(x, D))$ integers, $n \leq x$.

If f is in an ambiguous class $G \in \mathfrak{G}$, then $u = 1$. Thus the constant in front of the main term of (1.11) shows that ambiguous forms represent fewer small integers (and with higher multiplicity) than nonambiguous forms.

By (1.7) and the fact that any integer n is represented by no more than $\tau(n)$ distinct quadratic forms of discriminant $-D$, we deduce that for most f, the value of u here is $\gg D^{1/2-\epsilon}$ (and must be $\ll D^{1/2}$).

The bounds on $E_\beta(x, D)$ when $\beta > 2$ can be improved with more effort. Our result yields an asymptotic for

$$(\log x)^N \leq D = o(x) \tag{1.12}$$

with $N = 2 + \varepsilon$ if $\beta \leq 2$; and $N = 4 + 2(\beta - 2)(2 + 8^{\beta-1}e)\log 2 + \varepsilon$ if $\beta > 2$ (by taking, e.g., $q = 3 + 1/((\beta - 2)\log 2)$). Since $\sum_{n \leq x} r_f(n)^\beta$ is increasing in β, we obtain

$$\sum_{n \leq x} r_f(n)^\beta \asymp \frac{x}{\sqrt{D}}, \qquad (\log x)^{((2^{\beta-2}-2)/(\beta-1))+\varepsilon} \leq D = o(x) \tag{1.13}$$

from Corollary 1 for any real $\beta \geq 1$.

For arbitrary $\beta \geq 0$, we obtain less-precise results than Theorem 1 in the following.

THEOREM 3
Fix D. For all real $\beta \geq 0$, we have

$$\sum_{n \leq x} r_f(n)^\beta \asymp x(\log x)^{2^{\beta-1}-1},$$

the implied constants being dependent on β and D.

This result only works as $x \to \infty$. However, we can show that different forms of the same discriminant behave similarly for arbitrary x with $D = o(x)$.

THEOREM 4
For any primitive binary quadratic forms f and g of discriminant $-D$ and for any real constant $\beta \geq 0$, we have

$$\sum_{n \leq x} r_f(n)^\beta \asymp_\beta \sum_{n \leq x} r_g(n)^\beta$$

whenever $D = o(x)$. In fact, the ratio of the two sides of this equation is between $2^{-|1-\beta|+o(1)}$ and $2^{|1-\beta|+o(1)}$, where the $o(1)$ term approaches 0 as $x/D \to \infty$.

The constant term $1 + (2^{\beta-1} - 1)/u$ in Theorem 2 ranges between 1 and $2^{\beta-1}$. In this sense, Theorem 4 cannot be improved, though in view of Theorem 1, we see that the ratio tends to 1 for all choices of f and g as x gets sufficiently large.

We can obtain the correct order of magnitude in all ranges for all $\beta \geq 0$; the following theorem gives for fixed $L > 0$ a uniform result in the range $D \leq (\log x)^L$. (And Theorem 2 covers the range $D \geq (\log x)^L$.) Define the numbers $\kappa_1 = \kappa_1(\beta)$ and $\kappa_2 = \kappa_2(\beta)$ by

$$\kappa_1 = \kappa_2 = \frac{2^{\beta-1} - 1}{(\beta - 1)\log 2} \quad \text{if } \beta \geq 1$$

and

$$\kappa_1 = 2^{\beta-1}, \qquad \kappa_2 = 1 \quad \text{if } 0 \leq \beta \leq 1.$$

For $\kappa, \beta \geq 0$, let

$$E(\kappa, \beta) := \begin{cases} -1 + 2^{\beta-1} - \beta\kappa \log 2, & 0 \leq \kappa \leq \kappa_1, \\ -1 + \kappa\big(1 - \log(2\kappa)\big), & \kappa_1 < \kappa < \kappa_2, \\ -\kappa \log 2, & \kappa \geq \kappa_2. \end{cases}$$

THEOREM 5

Fix $L > 0$. If x is chosen so large that when we define κ by $h/g = (\log x)^{\kappa \log 2}$ we have $\kappa \leq L$ (and $E(\kappa, \beta) \geq -1 - L \log 2$), then

$$\sum_{n \leq x} r_f(n)^\beta = x(\log x)^{E(\kappa,\beta)+o(1)}. \tag{1.14}$$

If we assume that there are no Siegel zeros, that is, that

$$L(\sigma, \chi_\delta) \neq 0, \qquad \text{and for all } \sigma \geq 1 - \frac{c_0}{\log D} \tag{1.15}$$

for all fundamental discriminants $\delta \mid D$, and for a certain constant $c_0 > 0$, then

$$\sum_{n \leq x} r_f(n)^\beta = \frac{x(\log x)^{E(\kappa,\beta)}}{g}(\log\log x)^{O(1)} \tag{1.16}$$

in the same range. Here all implicit and explicit constants depend only on L and β.

We failed to obtain an asymptotic in Theorem 5, but we did obtain an estimate that gives the correct value up to a little noise, $(\log x)^{o(1)}$. In the case where $\beta = 0$, which is not covered by Theorem 1, we can do a little better than (1.14), as see in Theorem 6.

Bernays's result (see [2]) can be given more precisely as

$$N_f(x) = \kappa_D \frac{x}{\sqrt{\log x}} + O_D\Big(\frac{x}{(\log x)^{1/2+\delta}}\Big)$$

for any $\delta < \min(1/h, 1/4)$, where $\kappa_D = a_{1/2}$ in (1.5). Thus all but a vanishingly small proportion of integers which can be represented by some form in a given genus can be represented by all forms in the genus.* Bernays's proof only gives a nontrivial estimate once $D \ll (\log\log x)^{1/2-\varepsilon}$. Here we extend the range in which we have nontrivial estimates.

THEOREM 6
Keep the notation and assumptions of Theorem 5 but with κ now defined by $h/g = (\ell_{-D} \log x)^{\kappa \log 2}$, where $\ell_{-D} = L(1, \chi_{-D})(\phi(D)/D)$. Then we obtain the upper bounds in (1.1)–(1.3) in the ranges stated there. We also get the lower bound in (1.1) for $0 \le \kappa \le 1/2 - \varepsilon$, if D is sufficiently large in terms of ε.

Note that by Theorem 2, the lower bound in (1.3) holds for $\kappa \ge (1/(\log 2)) + \varepsilon$.

Let $V(x)$ be the number of integers at most x which are the sum of two powerful numbers. We deduce the following.

COROLLARY 2
We have

$$\frac{x(\log\log x)^A}{(\log x)^{1-2^{-1/3}}} \ll V(x) \ll \frac{x(\log\log x)^{2^{2/3}-1}}{(\log x)^{1-2^{-1/3}}}$$

for some $A \in \mathbb{R}$.

We conjecture that $V(x) \asymp x(\log x)^{-1+2^{-1/3}}(\log\log x)^{2^{2/3}-1}$. At any rate, it is interesting to have a natural example of a sequence which has considerably different additive behaviour as the squares but does not behave like a typical pseudosquare sequence.

With some extra work, Theorems 1–6 can be extended to nonfundamental discriminants. Some of our results also hold for real quadratic fields: Theorem 1 without (1.9) and with D replaced by h^2 in (1.10); Theorem 3; Theorems 4 and 5 if $D = (\log x)^{O(1)}$.

To our knowledge, the only (nontrivial) results on estimates/asymptotics of $\sum_{n\le x} r_f(n)^\beta$ for generic discriminants D (in particular, with more than one form per genus) known so far are Bernays's result (see [2]), (1.14) for $\beta = 0$ in [3], and parts of (1.5) for fixed discriminant and $\beta = 2$ in [11]. We have seen that the question of

*Bernays proved this for arbitrary discriminants D; for nonfundamental discriminants D, the constant κ_D is more complicated.

obtaining asymptotics for $N_f(x)$ remains open in the intermediate range (see (1.2)), whereas we have now obtained asymptotics for positive integer moments of $r_f(n)$ in all ranges (see Corollary 1). Although these results are new, the methods are well known in principle; the novelty here is that we have succeeded in implementing these in sufficiently sharp form to obtain asymptotics. More specifically, the proofs of Theorems 1–4 are elementary and do not use any particularly new ideas, though we have not seen anything like Theorem 4 elsewhere, and it seems that no one has previously observed the straightforward Theorem 2. Theorems 5 and 6 are refinements of [3, Corollaries 1, 1.1] using a somewhat different analysis and the more refined combinatorics of Lemmata 3.2 and 3.3, which allow us, in the most troublesome ranges, to avoid certain difficult technicalities.

In many of the classical problems of analytic number theory (e.g., counting primes) the difficult range is typically when x is close to a certain power of D. Perhaps so little has been done on this very natural question because the difficult range occurs here when x is between $\exp(D^{1/2-o(1)})$ and $\exp(D^{1/\log 2+o(1)})$, that is, in a range that is exponential in certain powers of D. This difference, and the fact that there are links to questions about the existence of Siegel zeros (see, e.g., [1]), perhaps deterred previous researchers.

One expects that applications of these results may be found by researchers involved in counting questions to do with binary quadratic forms, which of course appear in many ways in number theory.

Notation. All implicit and explicit constants depend *at most* on ε and β, on D in the proof of Theorem 3, on L in the proofs of Theorems 5 and 6, and on q in the proof of Theorem 2. The dependence on ε in (1.9), (1.10), and the lower bound in (1.1) as proved in Theorem 6, on ρ in (1.4), and the constants implied in the $o(1)$-symbol in Theorems 4 and 5 are not effective, as they depend on Siegel's theorem. All other implicit and explicit constants can in principle be made effective. The letter ε denotes an arbitrarily small real number whose value may change during a calculation. As usual, let $\chi_{-D} = (-D/.)$ be the Jacobi-Kronecker symbol.

1.3. A heuristic explanation of the transition

Let $r(n)$ be the total number of representations of an integer n by forms of discriminant $-D$. We base our heuristic on a study of integers n which are squarefree and all of whose prime factors p satisfy $(-D/p) = 1$. A more accurate analysis would consider integers n which are allowed a small square factor (for which similar remarks would apply). If $r(n) \neq 0$ with n squarefree, then $(-D/p) = 0$ or 1 for every prime $p|n$, and so a more accurate analysis would consider n divisible by prime factors of D. (And, again, similar remarks would apply.)

Now almost all integers $n \leq x$ which are squarefree and all of whose prime factors p satisfy $(-D/p) = 1$ have $\sim (1/2) \log\log x$ prime factors, and usually, those prime factors come from a wide variety of classes of the class group.

Thus we assume that $n = p_1 p_2 \cdots p_k$ with $p_1 < p_2 < \cdots < p_k$ and each $(-D/p_i) = 1$, where $k \sim (1/2)\log\log x$. In this case $r(n) = 2^k$, and n can only be represented by those forms in a particular genus (which contains h/g forms).

When 2^k is significantly smaller than h/g, we might expect the 2^k representations to be mostly by distinct forms and that n is represented twice by very few forms. Thus we might expect $N_f(x)$ to be more or less the same as the total number of pairs $r, s \in \mathbb{Z}$ for which $|f(r,s)| \leq x$, up to the obvious automorphisms. To be more precise about the range, we want $2^{(1/2)\log\log x} < (h/g)^{1/2}$, which corresponds to $\kappa > 1$ as in (1.3).

When 2^k is significantly larger than h/g, then we might expect that n is represented by almost all of the forms in its genus, and so $N_f(x)$ should be roughly the same as $N_g(x)$ for any other form g in the same genus as f. The requirement on 2^k corresponds to $\kappa \leq 1/2$ as in (1.1). To take this a little further, if the number of representations of such n are roughly equal for the forms in the genus, and if the number of integers represented by each genus is about equal, then we would expect that

$$\sum_{n \leq x} r_f(n)^\beta \sim \frac{1}{g}\left(\frac{g}{h}\right)^\beta \sum_{n \leq x} r(n)^\beta,$$

which leads us to predict the main term of (1.4).

Much of what we discuss here is justified by Lemma 3.3.

2. Preliminaries

We recall the following consequence (see [6, Theorem 7.7]) of the isomorphism between classes of quadratic forms and ideal classes. Let $C_f \in \mathfrak{C}$ be the class of f in the class group, and let $f(C)$ denote the class of quadratic forms corresponding to $C \in \mathfrak{C}$.

LEMMA 2.1

There is a one-to-one correspondence between inequivalent solutions to $f(\boldsymbol{x}) = n$ and integral ideals $\mathfrak{a}$ in the class C_f corresponding to f with $N\mathfrak{a} = n$.

Thus a positive integer n is represented by a form f if and only if there is an ideal $\mathfrak{a}$ with $N\mathfrak{a} = n$ in the class C_f. In this case, we write $n \in \mathcal{R}(C_f)$. In particular, a prime p with $\chi_{-D}(p) = 1$ has exactly two inequivalent representations in classes $C, C^{-1} \in \mathfrak{C}$. (Of course, C and C^{-1} may coincide.) A prime $p \mid D$ has exactly one representation, namely, in an ambiguous class. A prime p with $\chi_{-D}(p) = -1$ cannot be represented by a primitive form of discriminant $-D$.

We introduce the following notation. For an integer $n = \prod_{j=1}^{k} p_j$ with not necessarily distinct primes p_j, and for $\mathbf{C} = (C_1, \dots, C_k) \in \mathfrak{C}^k$, we write $n \leftrightarrow \mathbf{C}$ if and only if there is a permutation $\pi \in \mathfrak{S}_k$ with $p_{\pi(j)} \in \mathcal{R}(C_j)$. Choose a fixed set of representatives of the quotient $\mathfrak{C}/\mathfrak{G}$, and for each $C \in \mathfrak{C}$, let $\tilde{C}$ be this representative. For $\mathbf{C} = (C_1, \dots, C_k) \in \mathfrak{C}^k$, let $1 \le \rho(\mathbf{C}) \le k!$ be the number of different rearrangements of the k-tuple $\tilde{\mathbf{C}} = (\tilde{C}_1, \dots, \tilde{C}_k)$, and let $0 \le \delta(\mathbf{C}) \le k$ be the number of nonambiguous entries of $\mathbf{C}$. Finally, for $\mathbf{C} = (C_1, \dots, C_k) \in \mathfrak{C}^k$, $C \in \mathfrak{C}$, let

$$N_{\mathbf{C}}(C) := \#\Big\{(\sigma_1, \dots, \sigma_k) \in \{\pm 1\}^k \,\Big|\, C = \prod_{j=1}^{k} C_j^{\sigma_j}\Big\}. \tag{2.1}$$

With this notation, the above discussion yields the following lemma.

LEMMA 2.2
Let

$$n = \prod p_j^{\alpha_j} \prod q_j^{\beta_j} \prod r_j^{\gamma_j},$$

where $\chi_{-D}(p_j) = 1$, $\chi_{-D}(q_j) = 0$, *and* $\chi_{-D}(r_j) = -1$.

(a) *The integer n can be represented in some class if and only if all γ_j are even. If $(n_1, n_2) = 1$ and $C \in \mathfrak{C}$, then*

$$r_{f(C)}(n_1 n_2) = \sum_{C_1 C_2 = C} r_{f(C_1)}(n_1) r_{f(C_2)}(n_2).$$

(b) *Assume that n consists only of split primes, and assume that $n \leftrightarrow \boldsymbol{C}$ for some $\boldsymbol{C} \in \mathfrak{C}^k$. Then there are exactly $2^{\delta(\boldsymbol{C})}\rho(\boldsymbol{C})$ different k-tuples $\boldsymbol{D} \in \mathfrak{C}^k$ with $n \leftrightarrow \boldsymbol{D}$. All of these satisfy $\rho(\boldsymbol{C}) = \rho(\boldsymbol{D})$ and $\delta(\boldsymbol{C}) = \delta(\boldsymbol{D})$.*

If n is squarefree, then $N_{\boldsymbol{C}}(C) = r_{f(C)}(n)$ for all $C \in \mathfrak{C}$.

For $\chi \in \widehat{\mathfrak{C}}$, let

$$L_K(s, \chi) = \sum_{\{0\} \ne \mathfrak{a} \text{ integral}} \frac{\chi(\mathfrak{a})}{(N\mathfrak{a})^s} = \sum_{n=1}^{\infty} \frac{\sigma(\chi, n)}{n^s},$$

say, be the class group L-function of the field $K = \mathbb{Q}(\sqrt{-D})$. The coefficients $\sigma(\chi, n)$ are multiplicative and satisfy

$$\sigma(\chi, p^\nu) = \sum_{\substack{\mathfrak{a} \text{ integral} \\ N\mathfrak{a} = p^\nu}} \chi(\mathfrak{a}) = \begin{cases} \frac{1}{2}\big((-1)^\nu + 1\big), & \big(\frac{-D}{p}\big) = -1, \\ \chi^\nu(\mathfrak{p}), & p \mid D, \\ \sum_{\mu=0}^{\nu} \chi^{\nu - 2\mu}(\mathfrak{p}), & \big(\frac{-D}{p}\big) = 1, \end{cases} \tag{2.2}$$

By Lemma 2.1 and orthogonality of characters, we have

$$r_f(n) = \frac{1}{h} \sum_{\chi \in \widehat{\mathfrak{C}}} \bar{\chi}(C_f)\sigma(\chi, n). \tag{2.3}$$

We need bounds and zerofree regions for $L_K(s, \chi)$. By the convexity principle, we have

$$L_K(s, \chi) \ll \big(D^{1/2}(1 + |t|)\big),^{1-\sigma+\varepsilon} \tag{2.4}$$

$1/2 \le \sigma \le 1 + \varepsilon/2$, provided that $|s - 1| \ge 1/8$ if χ is principal. By [3, (2.8)], we have

$$\log L_K(\sigma + it, \chi) \ll \log D + \log\log(3 + |t|) \tag{2.5}$$

uniformly in

$$1 - c_1 \min\Big(\frac{1}{\log(D(1 + |t|))}, D^{-\varepsilon}\Big) \le \sigma \le 1 + \frac{c_2}{\log(D(1 + |t|))} \tag{2.6}$$

for some absolute constants $c_1, c_2 > 0$, provided that $\chi \ne \chi_0$ if $|t| \le 1/8$. This follows essentially from a result of Fogels [8, Lemma 4] and from Carathéodory's inequality. In [3, Lemmata 4.1–4.3], Blomar provides a Siegel-Walfisz theorem for quadratic fields:

$$\epsilon(C) \sum_{\substack{p \le \xi,\, p \in \mathbb{P} \\ p \in \mathcal{R}(C)}} 1 = \frac{1}{2h} \int_2^{\xi} \frac{dt}{\log t} + O_A\big(\xi \exp(-c_3\sqrt{\log \xi})\big) \tag{2.7}$$

uniformly in

$$D \le (\log \xi)^A \tag{2.8}$$

for any fixed $A > 0$. Here and henceforth for $C \in \mathfrak{C}$, let $\epsilon(C) = 1$ if C is ambiguous, and let $\epsilon(C) = 1/2$ otherwise. If there are no Siegel zeros for real characters $\chi \in \widehat{\mathfrak{C}}$, that is, if we assume (1.15) (cf. [3, Lemma 2.1]), then (2.7) holds uniformly in

$$D \le \exp(c_4\sqrt{\log \xi}). \tag{2.9}$$

3. Lemmata

We make precise here a well-known result, counting the number of values of a binary quadratic form which are at most x. Gauss showed that every binary quadratic form of negative discriminant is equivalent to a form $ax_1^2 + bx_1x_2 + cx_2^2$, where $|b| \le a \le c$. It is easily deduced that $a \ll \sqrt{D}$ and that $c \asymp D/a$. Moreover, a is the smallest positive integer represented by the quadratic form, and c is the smallest that is not of the form au^2. Therefore if $x < c$ and $am^2 + bmn + cn^2 \le x$, then $n = 0$, and so the number of such representations is $2\sqrt{x/a} + O(1)$. In general, we have the following result.

LEMMA 3.1
If $f(\boldsymbol{x}) = ax_1^2 + bx_1x_2 + cx_2^2$ is a reduced quadratic form of discriminant $-D$, then

$$\#\{(m,n) \in \mathbb{Z}^2 \mid f(m,n) \le x\} = \frac{2\pi x}{\sqrt{D}} + O\Big(\sqrt{\frac{x}{a}}\Big). \tag{3.1}$$

The implied constant is absolute. Note that $\sqrt{x/a} = o(x/\sqrt{D})$ as $x/c \to \infty$.

Proof
We need to count the number of integer solutions to $(2am + bn)^2 + Dn^2 \le 4ax$. This implies that $|n| \le \sqrt{4ax/D}$ and that

$$\frac{-\sqrt{4ax - Dn^2} - bn}{2a} \le m \le \frac{\sqrt{4ax - Dn^2} - bn}{2a}. \tag{3.2}$$

There are $\sqrt{4ax - Dn^2}/a + O(1)$ integers m in this range. Summing over the possible values of n, we see that the $O(1)$'s add up to $O(\sqrt{ax/D}) = O(\sqrt{x/a})$ since $a \ll \sqrt{D}$. The sum of $\sqrt{4ax - Dn^2}/a$ over integers n can be approximated by the corresponding integral; since the integrand is decreasing from zero to either endpoint, the error in this approximation is no more than twice the value of the integrand at zero, that is, $O(\sqrt{x/a})$. To evaluate the integral, we make the change of variable $n = t\sqrt{4ax/D}$, and the main term becomes $(4x/\sqrt{D}) \int_{-1}^{1} \sqrt{1 - t^2}\, dt$, which yields our result. $\square$

LEMMA 3.2
Let g denote the number of genera of discriminant $-D$. For fixed $\beta \ge 0$ and $\boldsymbol{C} \in \mathfrak{C}$, we have

$$\sum_{\boldsymbol{C} \in \mathfrak{C}^k} N_{\boldsymbol{C}}(C)^\beta \asymp \begin{cases} 2^k h^{k-1}, & 2^k \le \frac{h}{g}, \\ 2^{\beta k} h^{k-\beta} g^{\beta-1}, & 2^k > \frac{h}{g}, \end{cases}$$

with $N_{\boldsymbol{C}}(C)$ as in (2.1).

Proof
First note that

$$h^{\min(0,\beta-1)} \sum_{j=1}^{h} a_j^\beta \le \Big(\sum_{j=1}^{h} a_j\Big)^\beta \le h^{\max(0,\beta-1)} \sum_{j=1}^{h} a_j^\beta \tag{3.3}$$

for any sequence (a_j) of nonnegative real numbers and any real $\beta \geq 0$. Therefore, by (3.3),

$$\begin{aligned}
&\sum_{\mathbf{C}\in\mathfrak{C}^k} N_{\mathbf{C}}(C)^\beta \\
&\quad = \sum_{\mathbf{C}\in\mathfrak{C}^{k-1}} \sum_{C_k\in\mathfrak{C}} \Big(\#\Big\{(\sigma_1,\dots,\sigma_{k-1})\in\{\pm1\}^{k-1} \;\Big|\; C = C_k\prod_{j=1}^{k-1} C_j^{\sigma_j}\Big\} \\
&\qquad\qquad + \#\Big\{(\sigma_1,\dots,\sigma_{k-1})\in\{\pm1\}^{k-1} \;\Big|\; C = C_k^{-1}\prod_{j=1}^{k-1} C_j^{\sigma_j}\Big\}\Big)^\beta \\
&\quad \asymp \sum_{\mathbf{C}\in\mathfrak{C}^{k-1}} \sum_{C\in\mathfrak{C}} N_{\mathbf{C}}(C)^\beta.
\end{aligned}$$

For $\beta = 0$, the lemma is [3, Proposition 5.3]. Now assume that β is a positive integer. Expanding the βth power, we see that the right-hand side of the preceding display equals

$$\begin{aligned}
&\sum_{\sigma_{ij}\in\{\pm1\}^{\beta\times(k-1)}} \#\Big\{\mathbf{C}\in\mathfrak{C}^{k-1} \;\Big|\; \prod_{j=1}^{k-1} C_j^{\sigma_{1j}} = \cdots = \prod_{j=1}^{k-1} C_j^{\sigma_{\beta j}}\Big\} \\
&\quad = 2^{k-1} \sum_{\tau_{ij}\in\{0,2\}^{(\beta-1)\times(k-1)}} \#\Big\{\mathbf{C}\in\mathfrak{C}^{k-1} \;\Big|\; \prod_{j=1}^{k-1} C_j^{\tau_{1j}} = 1,\, 1\leq i\leq \beta-1\Big\} \\
&\quad = (2g)^{k-1} \sum_{\tau_{ij}\in\{0,1\}^{(\beta-1)\times(k-1)}} \#\Big\{\mathbf{C}\in(\mathfrak{C}^2)^{k-1} \;\Big|\; \prod_{j=1}^{k-1} C_j^{\tau_{ij}} = 1,\, 1\leq i\leq \beta-1\Big\}.
\end{aligned}$$

We bound this term from below and above. Choosing $\tau_{ij} = 0$ for all i, j, we get

$$\sum_{\mathbf{C}\in\mathfrak{C}^k} N_{\mathbf{C}}(C)^\beta \geq (2g)^{k-1}\Big(\frac{h}{g}\Big)^{k-1} = (2h)^{k-1} \tag{3.4}$$

since there are h/g elements $\mathbf{C}\in\mathfrak{C}^2$. On the other hand, assuming without loss of generality that $k \geq \beta$, the matrices $(\tau_{ij}) = (I_{\beta-1} *)$ (where I_n is the $(n\times n)$-identity matrix) give

$$\sum_{\mathbf{C}\in\mathfrak{C}^k} N_{\mathbf{C}}(C)^\beta \geq (2g)^{k-1}(2^{k-\beta})^{\beta-1}\Big(\frac{h}{g}\Big)^{k-\beta} \gg (2h)^{k-1}\Big(\frac{2^k g}{h}\Big)^{\beta-1} \tag{3.5}$$

since there are $(2^{k-\beta})^{\beta-1}$ choices of "$*$."

For the upper bound, we proceed as follows. For a prime p, let $r_p(T)$ denote the $\mathbb{F}_p$-rank of the matrix $T = (\tau_{ij})$. We claim that for $0 \le \mu \le \min(k-1, \beta-1)$, there are at most

$$\Big(\prod_{i=1}^{\mu}(\beta - i)\Big)2^{\mu(\beta-1-\mu)}2^{(k-1)\mu} \ll_\beta 2^{(k-1)\mu}$$

matrices $T \in \{0, 1\}^{(\beta-1)\times(k-1)}$ with $r_p(T) = \mu$. Indeed, we pick μ linearly independent rows and place them in $(\beta-1)(\beta-2)\cdots(\beta-\mu)$ ways in the matrix T. The remaining $\beta - 1 - \mu$ row vectors then lie in the subspace S generated by the first μ row vectors. We can pick μ-coordinates $i_1, \dots, i_\mu$ such that the map $\phi : S \subset \mathbb{F}_p^{k-1} \to \mathbb{F}_p^\mu$, $(x_1, \dots, x_{k-1}) \mapsto (x_{i_1}, \dots, x_{i_\mu})$ is an isomorphism. Since $\phi(S \cap \{0, 1\}^{k-1}) \subseteq \{0, 1\}^\mu$, we see that each of the remaining row vectors can only be chosen out of at most 2^μ possibilities. This proves the claim. Now let $r_0(T) := \min_p r_p(T)$. Any finite abelian group, and in particular, the set of square classes $\mathfrak{C}^2$, may be written as $\prod_q \mathfrak{C}_q$, where each q is the power p^α of a prime p and $\mathfrak{C}_q = \mathbb{Z}/p^\alpha\mathbb{Z}$. If $\prod_{j=1}^{k-1} C_j^{\tau_{ij}} = 1$ for $\mathbf{C} \in (\mathfrak{C}^2)^{k-1}$, then each component must satisfy the same equation, and the number of solutions within that component is at most $(\#\mathfrak{C}_q)^{k-1-r_p(T)}$ by elementary linear algebra for a given matrix T. And thus, there are at most

$$\prod_q (\#\mathfrak{C}_q)^{k-1-r_p(T)} \le \Big(\frac{h}{g}\Big)^{k-1-r_0(T)}$$

ways to choose $\mathbf{C} \in \mathfrak{C}^{k-1}$ such that $\prod_{j=1}^{k-1} C_j^{\tau_{ij}} = 1$ for $1 \le i \le \beta - 1$, and there are at most $\ll 2^{(k-1)\mu}$ matrices T with $r_0(T) = \mu$. Since $\mu \le \beta - 1$, we obtain, altogether,

$$\sum_{\mathbf{C}\in\mathfrak{C}^k} N_{\mathbf{C}}(C)^\beta \ll (2h)^{k-1}\bigg(1 + \Big(\frac{2^k g}{h}\Big)^{\beta-1}\bigg). \tag{3.6}$$

Equations (3.4)–(3.6) prove the lemma for integral β. The general case is clear for $2^k \le h/g$ and follows for $2^k > h/g$ from Hölder's inequality; let β be an arbitrary real number that is not an integer, and assume that $2^k > h/g$. Let $p := (\beta - [\beta])^{-1} > 1$, $q := (1 - 1/p)^{-1}$. Then

$$\begin{aligned}\sum_{\mathbf{C}\in\mathfrak{C}^k} N_{\mathbf{C}}(C)^\beta &\le \Big(\sum_{\mathbf{C}\in\mathfrak{C}^k} N_{\mathbf{C}}(C)^{(\beta-[\beta]/q)p}\Big)^{1/p}\Big(\sum_{\mathbf{C}\in\mathfrak{C}^k} N_{\mathbf{C}}(C)^{[\beta]}\Big)^{1/q}\\ &\asymp (2h)^{k-1}\Big(\frac{2^k g}{h}\Big)^{\beta-1}.\end{aligned}$$

Note that $(\beta - [\beta]/q)p = 1 + [\beta] \in \mathbb{N}$. On the other hand, let $p := 2 - \{\beta\} > 1$, $q := (1 - 1/p)^{-1}$. Then

$$\sum_{\mathbf{C}\in\mathfrak{C}^k} N_{\mathbf{C}}(C)^{\beta} \geq \Big(\sum_{\mathbf{C}\in\mathfrak{C}^k} N_{\mathbf{C}}(C)^{\frac{\beta}{p}+(1+\lceil\beta\rceil)/q}\Big)^{p}\Big(\sum_{\mathbf{C}\in\mathfrak{C}^k} N_{\mathbf{C}}(C)^{1+\lceil\beta\rceil}\Big)^{-p/q}$$
$$\asymp (2h)^{k-1}\Big(\frac{2^k g}{h}\Big)^{\beta-1}$$

since $\beta/p + (1 + \lceil\beta\rceil)/q = \lceil\beta\rceil \in \mathbb{N}$. This completes the proof. $\square$

LEMMA 3.3
Let $\lambda := (\log(h/g))/(\log 2)$, $k \in \mathbb{N}$, *and* $\varepsilon > 0$. *Then*

$$\sum_{C\in\mathfrak{C}} N_{\boldsymbol{C}}(C)^0 \leq \min\Big(2^k, \frac{h}{g}\Big)$$

for all $\boldsymbol{C} \in \mathfrak{C}^k$, *and*

$$\sum_{C\in\mathfrak{C}} N_{\boldsymbol{C}}(C)^0 \gg \min\Big(2^k, \frac{h}{g}\Big)$$

for all but $\ll_\varepsilon h^k D^{-\varepsilon}$ *tuples of classes* $\boldsymbol{C} \in \mathfrak{C}^k$ *if* $k \leq (1 - 4\varepsilon)\lambda$ *or* $k \geq (1 + 18\varepsilon)\lambda$.

Remark. It would be nice to prove such a result for all k.

Proof
For subsets $A_1, \ldots, A_k \subseteq \mathfrak{C}$, let

$$\prod_{j=1}^{k} A_j := \{a_1 \cdots a_k \mid a_1 \in A_1, \ldots, a_k \in A_k\}.$$

For $\mathbf{C} = (C_1, \ldots, C_k) \in \mathfrak{C}^k$, we have

$$\sum_{C\in\mathfrak{C}} N_{\mathbf{C}}(C)^0 = \#\prod_{j=1}^{k}\{C_j, C_j^{-1}\} = \#\prod_{j=1}^{k}\{1, C_j^2\}.$$

Thus the upper bound is immediate. Now let $r(\mathbf{C}) := \#\prod\{1, C_j\}$. It is not hard to see (see 3, (5.3)) that

$$\frac{1}{h}\sum_{C\in\mathfrak{C}} r(\mathbf{C}, C) = 2r(\mathbf{C}) - \frac{r^2(\mathbf{C})}{h}$$

for all $(\mathbf{C}, C) \in \mathfrak{C}^{k+1}$. (Here, of course, by slight abuse of notation, $(\mathbf{C}, C)$ is formed by appending C to $\mathbf{C}$.) Hence, given any $\eta > 1$, we have at least $h(1 - 1/\eta)$ classes $C \in \mathfrak{C}$ satisfying

$$r(\mathbf{C}, C) \geq 2r(\mathbf{C}) - \frac{\eta r(\mathbf{C})^2}{h} \geq 2r(\mathbf{C})\Big(1 - \frac{\eta 2^k}{2h}\Big).$$

Inductively, we see that for $k \in \mathbb{N}$, $\eta > 1$, $\eta 2^k < 2h$, at least $h^k(1 - 1/\eta)^k$ tuples of classes $\mathbf{C} \in \mathfrak{C}^k$ satisfy

$$\#\prod_{j=1}^{k}\{1, C_j\} \geq 2^k\Big(1 - \frac{\eta 2^k}{2h}\Big)^k.$$

Applying this result to the group of square classes and observing that each square has g square roots, we find that at least $h^k(1 - 1/\eta)^k$ tuples of classes $\mathbf{C} \in \mathfrak{C}^k$ satisfy

$$\#\prod_{j=1}^{k}\{1, C_j^2\} \geq 2^k\Big(1 - \frac{\eta 2^k}{2h/g}\Big)^k$$

(as long as $\eta 2^k < 2h/g$). Choosing $\eta = D^{3\varepsilon/2}$, we see that all but $\ll h^k D^{-\varepsilon}$ tuples of classes $\mathbf{C} \in \mathfrak{C}^k$ satisfy

$$\#\prod_{j=1}^{k}\{1, C_j^2\} \gg 2^k \tag{3.7}$$

if $k \leq (1 - 4\varepsilon)\lambda$. Now choosing $\eta = 2$, we find that for any given $\mathbf{C} \in \mathfrak{C}^k$, at least $h/2$ tuples of classes $(\mathbf{C}, C) \in \mathfrak{C}^{k+1}$ satisfy

$$r(\mathbf{C}^2, C^2) \geq \min\Big(\frac{3}{2}r(\mathbf{C}^2), \frac{h}{g}\Big).$$

(Here $\mathbf{C}^2$ means $(C_1^2, \ldots, C_k^2)$.) Hence, for any $\mathbf{C} \in \mathfrak{C}^k$ and any two positive integers $a < b$, at least

$$\frac{1}{2^b}\sum_{j=a}^{b}\binom{b}{j}h^b$$

tuples of classes $\mathbf{D} \in \mathfrak{C}^b$ satisfy

$$r(\mathbf{C}^2, \mathbf{D}^2) \geq \min\Big(\Big(\frac{3}{2}\Big)^a r(\mathbf{C}^2), \frac{h}{g}\Big).$$

Applying this result with $a = 7\varepsilon\lambda$ and $b = 22\varepsilon\lambda$ to the tuples of classes $\mathbf{C}$ found in (3.7) for $k = [(1-4\varepsilon)\lambda]$, we get $(1/2^b)\sum_{j=a}^{b}\binom{b}{j} = 1 - O(D^{-\tilde{\varepsilon}})$ for any $\tilde{\varepsilon} > 0$, and so all but $\ll h^k D^{-\varepsilon}$ tuples of classes $\mathbf{C} \in \mathfrak{C}^k$ satisfy

$$\#\prod_{j=1}^{k}\{1, C_j^2\} \gg \frac{h}{g}$$

if $k \geq (1+18\varepsilon)\lambda$. □

4. Proof of Theorem 3

Theorem 3 follows directly from the method in [4]. In view of Bernays's result (see [2]), we may assume that $\beta > 0$. For $\beta > 0$ and $m \in \mathbb{N}$, we define

$$\gamma(m, \beta) := \frac{1}{2m}\sum_{j=1}^{m}\left|2\cos\left(\frac{2\pi j}{m}\right)\right|^{\beta}.$$

Note that

$$\gamma(1, \beta) = \gamma(2, \beta) = 2^{\beta-1} > \gamma(m, \beta) \tag{4.1}$$

for all $m \geq 3$. From [4, pages 143–144], we conclude that

$$\sum_{n\leq x}|\sigma(\chi, n)|^{\beta} \sim c(\chi, \beta)x(\log x)^{\gamma(\mathrm{ord}\chi, \beta)-1} \tag{4.2}$$

for some constant $c(\chi, \beta) > 0$. Furthermore, it is shown in [4] that there is a subset $\mathcal{N} \subseteq \mathbb{N}$ such that $\mathrm{Re}(\chi(C_f))\sigma(\chi, n)$ is nonnegative for all $n \in \mathcal{N}$ and all $\chi \in \widehat{\mathcal{C}}$, and

$$\sum_{\substack{n\leq x\\ n\in\mathcal{N}}}|\sigma(\chi, n)|^{\beta} \sim c'(\chi, \beta)x(\log x)^{\gamma(\mathrm{ord}\chi, \beta)-1}$$

for some constant $c'(\chi, \beta) > 0$. By (3.3), (2.3), and the above remarks,

$$\frac{1}{h^{\max(1,\beta)}}\sum_{\chi\in\widehat{\mathfrak{C}}}\sum_{\substack{n\leq x\\ n\in\mathcal{N}}}\left(\mathrm{Re}(\bar{\chi}(C_f))\sigma(\chi, n)\right)^{\beta} \leq \sum_{n\leq x}r_f(n)^{\beta} \leq \frac{1}{h^{\min(1,\beta)}}\sum_{\chi\in\widehat{\mathfrak{C}}}\sum_{n\leq x}|\sigma(\chi, n)|^{\beta}.$$

By (4.1) and (4.2), we get

$$\sum_{n\leq x}r_f(n)^{\beta} \asymp_{D,\beta} x(\log x)^{2^{\beta-1}-1}.$$

This completes the proof of Theorem 3. □

5. Proof of Theorem 2

To prove Theorem 2, we recall the notion of a primitive ideal. An ideal $\mathfrak{a} \subseteq \mathfrak{o}$ is called primitive if it is not divisible by any rational integer other than 1. As before, let $\mathfrak{G} \subseteq \mathfrak{C}$ be the set of ambiguous classes, let $\mathfrak{A}$ be the set of primitive ideals coprime to D, and let

$$\mathfrak{X}_G := \{\mathfrak{c} \in G \mid \mathfrak{c} \neq \bar{\mathfrak{c}}\}$$

for $G \in \mathfrak{G}$. Note that $\mathfrak{v} \in \mathfrak{A}$ implies that $\mathfrak{v}^2 \in \mathfrak{A}$, and for two ideals $\mathfrak{v}_1, \mathfrak{v}_2 \in \mathfrak{A}$ in the same class, the principal ideal $\bar{\mathfrak{v}}_1\mathfrak{v}_2$ is generated by a rational integer if and only if $\mathfrak{v}_1 = \mathfrak{v}_2$.

We start with the following observation. By Lemma 2.1, a pair $(\mathbf{x}_1, \mathbf{x}_2) \in \mathbb{Z}^2 \times \mathbb{Z}^2$ of inequivalent solutions to $f(\mathbf{x}_1) = f(\mathbf{x}_2) = n$ corresponds to a pair $(\mathfrak{a}_1, \mathfrak{a}_2)$ of different ideals in the class C_f having norm n. These are exactly the pairs of ideals $(\mathfrak{b}\mathfrak{c}, \mathfrak{b}\bar{\mathfrak{c}})$ with $N\mathfrak{b}\mathfrak{c} = n$, where $\mathfrak{c} \in \mathfrak{X}_G$ for some $G \in \mathfrak{G}$, and $\mathfrak{b} \in \mathfrak{A}$ is in the class $C_f G$. Let $\mathfrak{u}$ be the ideal in some class $C_f G_0$ of the coset $C_f\mathfrak{G}$ having smallest possible norm $N\mathfrak{u} = u$. Then $\mathfrak{u} \in \mathfrak{A}$ since we may divide out from u any rational integer and any ideal dividing D and still have an ideal in a class $C_f G$ with even smaller norm, contradicting the definition of u. For $n \in \mathbb{N}$, $G \in \mathfrak{G}$, $x \in \mathbb{R}$, let

$$\rho_1(n, G) := \#\{\mathfrak{c} \in \mathfrak{X}_G : N\mathfrak{c} = n\}, \qquad R_1(x, G) := \sum_{n \le x} \rho_1(n),$$

$$\rho_2(n, G) := \#\{\mathfrak{b} \in C_f G \cap \mathfrak{A} : N\mathfrak{b} = n,\ \mathfrak{b} \neq \mathfrak{u}\}, \qquad R_2(x, G) := \sum_{n \le x} \rho_2(n).$$

We have the following estimates.

LEMMA 5.1
For all $G \in \mathfrak{G}$, we have

$$R_1(x, G) \le \frac{16x}{\sqrt{D}}, \qquad R_2(x, G) \begin{cases} \ll x D^{-1/2} + \sqrt{x} & \text{for all } x, \\ \le 1 & \text{for } x \le \sqrt{\frac{D}{4}}, \\ = 0 & \text{for } x \le \left(\frac{D}{4}\right)^{1/4}, \end{cases}$$

with absolute implied constants.

Proof
First, we note that if the vector $\mathbf{x} \in \mathbb{Z}^2$ corresponds to the ideal $\mathfrak{a}$ as in Lemma 2.1, then for $r \in \mathbb{Z}$ the vector $r\mathbf{x}$ corresponds to $(r)\mathfrak{a}$. Each ambiguous class $G \in \mathfrak{G}$ contains a

form of the shape

$$\begin{aligned} f_G(\mathbf{x}) &= ax_1^2 + cx_2^2, \qquad 4ac = D, \\ f_G(\mathbf{x}) &= ax_1^2 + ax_1x_2 + cx_2^2 = a\Big(x_1 + \frac{1}{2}x_2\Big)^2 + \Big(c - \frac{1}{4}a\Big)x_2^2, \\ a(4c - a) &= D, \end{aligned}$$

or

$$\begin{aligned} f_G(\mathbf{x}) &= ax_1^2 + bx_1x_2 + ax_2^2 = \Big(\frac{a}{2} + \frac{b}{4}\Big)(x_1 + x_2)^2 + \Big(\frac{a}{2} - \frac{b}{4}\Big)(x_1 - x_2)^2, \\ 4a^2 - b^2 &= D, \end{aligned}$$

with positive integers $b < a \le c$ (see [6, Lemma 3.10]). In the first case, the vectors $(0, *)$, $(*, 0)$ correspond to ideals $\mathfrak{d}(*)$ with $N\mathfrak{d} \mid D$ which are equal to their conjugate, and so, they are not in $\mathfrak{X}_G$. Therefore

$$R_1(x, G) \le \#\Big\{\mathbf{x} \in \mathbb{Z}^2 : x_1x_2 \ne 0, |x_1| \le \sqrt{\frac{x}{a}}, |x_2| \le \sqrt{\frac{x}{c}}\Big\} \le \frac{8x}{\sqrt{D}}.$$

In the second case, the vectors $\mathbf{x}$ with $x_2 = 0$ or $x_2 = -2x_1$ correspond to ideals that are equal to their conjugate. Therefore

$$\begin{aligned} R_1(x, G) &\le \#\Big\{\mathbf{y} \in \frac{1}{2}\mathbb{Z} \times \mathbb{Z} \;\Big|\; y_1y_2 \ne 0, |y_1| \le \sqrt{\frac{x}{a}}, |y_2| \le \sqrt{x}\Big(c - \frac{1}{4}a\Big)^{-1/2}\Big\} \\ &\le \frac{8x}{\sqrt{D}}. \end{aligned}$$

In the third case, the vectors $\mathbf{x}$ with $x_1 = \pm x_2$ correspond to ideals that are equal to their conjugate. Thus

$$\begin{aligned} R_1(x, G) \le \#\Big\{\mathbf{y} \in \mathbb{Z}^2 \;\Big|\; & y_1y_2 \ne 0, |y_1| \le \sqrt{x}\Big(\frac{a}{2} + \frac{b}{4}\Big)^{-1/2} \\ & |y_2| \le \sqrt{x}\Big(\frac{a}{2} - \frac{b}{4}\Big)^{-1/2}\Big\} \le \frac{16x}{\sqrt{D}}. \end{aligned}$$

This proves the first part of the lemma.

For the second part, we first note that $R_2(x, G) \ll xD^{-1/2} + \sqrt{x}$ by Lemma 3.1. For the rest, first observe that if $\mathfrak{w}$ is a principal ideal with $N\mathfrak{w} \le D/4$, then $\mathfrak{w}$ is generated by a rational integer. Therefore an ideal $\mathfrak{v} \in C_f G \cap \mathfrak{A}$ different from $\mathfrak{u}$ with $N\mathfrak{v} \le (D/4)^{1/4}$ would produce a principal ideal $\mathfrak{w} = \bar{\mathfrak{u}}^2\mathfrak{v}^2$ with $N\mathfrak{w} \le D/4$, and so, it is generated by a rational integer. However, if $\mathfrak{v}$ is different from $\mathfrak{u}$, then $\mathfrak{w}$ is, by the remarks on $\mathfrak{A}$, not generated by a rational integer. This is a contradiction. Similarly, two different ideals $\mathfrak{v}_1, \mathfrak{v}_2 \in C_f G \cap \mathfrak{A}$ with $N\mathfrak{v}_1, N\mathfrak{v}_2 \leq (D/4)^{1/2}$ would produce a principal ideal $\mathfrak{w} = \bar{\mathfrak{v}}_1\mathfrak{v}_2$ with $N\mathfrak{w} \leq D/4$, and so, it is generated by a rational integer, which is impossible by the same reason. This completes the proof of the lemma. □

We can now prove Theorem 2. We deduce from Lemma 5.1 that for large D up to a small set of exceptions, the set $\{\mathfrak{a} \in C_f \mid N\mathfrak{a} = n\}$ consists either of one element or of two elements $\{\mathfrak{u}\mathfrak{c}, \mathfrak{u}\bar{\mathfrak{c}}\}$ with $\mathfrak{c} \in \mathfrak{X}_{G_0}$. To be precise, define $A_1 = A_1(n) := \#\{\mathfrak{a} \in C_f \mid N\mathfrak{a} = n, \mathfrak{a} \notin \mathfrak{u}\mathfrak{X}_{G_0}\}$ and $A_2 = A_2(n) := \#\{\mathfrak{a} \in C_f \mid N\mathfrak{a} = n, \mathfrak{a} \in \mathfrak{u}\mathfrak{X}_{G_0}\}$ so that $A_2 = 0$ or $A_2 \geq 2$; and then,

$$r_f^*(n, \beta) := A_1(n) + 2^{\beta-1}A_2(n).$$

We define $B_3 = \{\mathfrak{c} \in G_0 : \mathfrak{c} = \bar{\mathfrak{c}}, N\mathfrak{c} \leq x/u\}$; note that if $\mathfrak{c}_0$ is the ideal in G_0 which divides D, then the elements of B_3 are simply $\mathfrak{c}_0$ times an integer, and so $|B_3| \ll \sqrt{x/(uN\mathfrak{c}_0)}$. Since $\mathfrak{u}\mathfrak{c}_0 \in C_f$, thus $uc_0 \geq a$ by the definition of a (see Corollary 1); this is $\leq \sqrt{x/a}$. By Lemma 3.1,

$$\begin{aligned}
\sum_{n\leq x} r_f^*(n, \beta) &= \sum_{n\leq x} \big(A_1(n) + A_2(n)\big) + (2^{\beta-1} - 1)\Big(|B_3| + \sum_{n\leq x} A_2(n)\Big) + O(|B_3|) \\
&= \#\{\mathfrak{a} \in C_f : Na \leq x\} + (2^{\beta-1} - 1)\#\Big\{\mathfrak{c} \in G_0 : N\mathfrak{c} \leq \frac{x}{u}\Big\} + O\Big(\sqrt{\frac{x}{a}}\Big) \\
&= \Big(1 + \frac{2^{\beta-1} - 1}{u}\Big)\frac{\pi x}{\sqrt{D}} + O\Big(\sqrt{\frac{x}{a}}\Big). \qquad (5.1)
\end{aligned}$$

Now, let us first assume that $\beta \leq 2$. A short calculation using the first derivative yields that $\xi(\beta) := r_f(n)^\beta - r_f^*(n, \beta) = (A_1 + A_2)^\beta - (A_1 + 2^{\beta-1}A_2)$ satisfies $|\xi(\beta)| \leq \xi(2)$ for $0 \leq \beta \leq 2$ and $A_1 \in \mathbb{N}_0$, $A_2 \in \mathbb{N}_0 \setminus \{1\}$, as can easily be checked. Therefore

$$\begin{aligned}
&\sum_{n\leq x} |r_f(n)^\beta - r_f^*(n, \beta)| \\
&\quad\leq \sum_{n\leq x} r_f(n)^2 - r_f^*(n, 2) \\
&\quad= \sum_{G\in\mathfrak{G}} \#\big\{(\mathfrak{b}\mathfrak{c}, \mathfrak{b}\bar{\mathfrak{c}}) : \mathfrak{c} \in \mathfrak{X}_G, \mathfrak{b} \in C_f G \cap \mathfrak{A} \setminus \{\mathfrak{u}\}, N\mathfrak{b}\mathfrak{c} \leq x\big\} \\
&\quad= \sum_{G\in\mathfrak{G}} \sum_{k\leq x} \rho_1(k, G) \sum_{l\leq x/k} \rho_2(l, G) \\
&\quad\ll \sum_{G\in\mathfrak{G}} \Big(\sum_{k\ll xD^{-1/4}} \rho_1(k, G) + \sum_{k\leq xD^{-1/2}} \rho_1(k, G)\Big(\frac{x}{k\sqrt{D}} + \sqrt{\frac{x}{k}}\Big)\Big) \\
&\quad\ll \tau(D)\Big(\frac{x \log x}{D} + \frac{x}{D^{3/4}}\Big),
\end{aligned}$$

by Lemma 5.1 and partial summation. Together with (5.1), we arrive at the theorem in the case where $\beta \le 2$.

If $\beta > 2$, we claim that $0 \le r_f(n)^\beta - r_f^*(n,\beta) \le 3r_f(n)^{\beta-2}(r_f(n)^2 - r_f^*(n,2))$; that is, $(A_1+A_2)^\beta - (A_1 + 2^{\beta-1}A_2) \le 3(A_1+A_2)^{\beta-2}((A_1+A_2)^2 - (A_1+2A_2))$. Now if $A_1 + A_2 \ge 3$, then $(A_1+A_2)^2 \ge 3(A_1+A_2) \ge (3/2)(A_1+2A_2)$, and the result follows. If $A_1 + A_2 \le 2$ with $A_2 \ne 1$, then both sides of the inequality equal zero. Therefore

$$\begin{aligned}\sum_{n\le x} |r_f(n)^\beta - r_f^*(n,\beta)| &\le 3\sum_{n\le x} \tau(n)^{\beta-2}\left(r_f(n)^2 - r_f^*(n,2)\right)\\ &\le 3\sum_{G\in\mathfrak{G}}\sum_{k\le x} \tau(k)^{\beta-2}\rho_1(k,G)\sum_{l\le x/k}\tau(l)^{\beta-2}\rho_2(l,G).\end{aligned}$$

By Hölder's inequality and $\rho_1(n,G) \le \tau(n)$, we get

$$\begin{aligned}\sum_{k\le x}\tau(k)^{\beta-2}\rho_1(k,G) &\le \Big(\sum_{k\le x}\rho_1(k,G)\Big)^{1/p}\Big(\sum_{k\le x}\tau(k)^{((p-1)/p+\beta-2)q}\Big)^{1/q}\\ &\ll \frac{x}{D^{(1/2)(1-1/q)}}(\log x)^{(1/q)\left(2^{(\beta-2)q+1}-1\right)},\end{aligned}$$

and similarly,

$$\sum_{k\le x}\tau(k)^{\beta-2}\rho_2(k,G) \ll x/D^{(1/4)(1-1/q)}(\log x)^{(1/q)(2^{(\beta-2)q+1}-1)},$$

where we used the crude bound $\sum_{n\le x}\rho_2(n,G) \ll xD^{-1/4}$, which follows from Lemma 5.1. Collecting these estimates, we find by partial summation that

$$\sum_{n\le x} r_f(n)^\beta - r_f^*(n,\beta) \ll \tau(D)\frac{x(\log x)^{(2/q)\left(2^{(\beta-2)q+1}-1\right)+1}}{D^{(3/4)(1-1/q)}}$$

for any $q > 1$. This completes the proof of Theorem 2. □

6. Proof of Theorem 4

Theorem 4 follows from Theorem 2 for $(\log x)^N \le D = o(x)$ with N sufficiently large (see (1.12)). We now prove it in the complementary range. By Theorem 5, we know that

$$\sum_{n\le x} r_f(n)^\beta \gg x(\log x)^{-2-N\log 2}$$

in this range. Let us first note that if $\mathcal{N}$ is a set of integers up to x containing $\ll_B x/(\log x)^B$ elements for any given B, then

$$\sum_{\substack{n\le x\\ n\in\mathcal{N}}} r_f(n)^\beta \le \Big(\sum_{\substack{n\le x\\ n\in\mathcal{N}}} 1 \sum_{n\le x} \tau(n)^{2\beta}\Big)^{1/2} = o\big(x(\log x)^{-3-N\log 2}\big),$$

by the Cauchy-Schwarz inequality; thus their contribution is negligible. We let $\mathcal{N}$ be the set of integers not exceeding x which have all of their prime factors less than $z = x^{1/\log\log x}$ or which are divisible by the square of a prime greater than z. The number of integers up to x, all of whose prime factors are $< x^{1/u}$, is $\ll x/u^{u+o(u)}$ for $u \le \log x/(\log\log x)^2$ (see [9]); therefore

$$\#\mathcal{N} \ll_B \frac{x}{(\log x)^B} + \sum_{p>z} \frac{x}{p^2} \ll_B \frac{x}{(\log x)^B}$$

for any constant B, as desired. It remains to sum over integers $n = pm \le x$, where $p \ge z$ and $P(m)$, the largest prime factor of m, is $< p$. This implies that $m \le x/p \le M := x/z$. For sufficiently large x, we have $z > D$, so p must be split in $\mathbb{Q}(\sqrt{-D})$ and hence is represented by classes C_p, C_p^{-1}, say (which may be the same). If, as before, C_f denotes the class corresponding to f, and $f(C)$ denotes the form corresponding to C, we get, by (3.3) with $h = 2$ and Lemma 2.2(a),

$$\begin{aligned}\sum_{n\le x} r_f(n)^\beta \sim \sum_{\substack{n\le x\\ n\notin\mathcal{N}}} r_f(n)^\beta &= \sum_{\substack{m\le M\\ mP(m)\le x}} \sum_{\substack{z\le p\le x/m\\ P(m)<p}} \big(r_{f(C_fC_p^{-1})}(m) + r_{f(C_fC_p)}(m)\big)^\beta \\ &\asymp \sum_{\substack{m\le M\\ mP(m)\le x}} \sum_{\substack{z\le p\le x/m\\ P(m)<p}} r_{f(C_fC_p^{-1})}(m)^\beta + r_{f(C_fC_p)}(m)^\beta \\ &= 2\sum_{C\in\mathfrak{C}} \sum_{\substack{m\le M\\ mP(m)\le x}} r_{f(C)}(m)^\beta \sum_{\substack{z\le p\le x/m\\ P(m)<p\\ p\in\mathcal{R}(C_fC^{-1})}} 1.\end{aligned}$$

Since we assumed that $D \le (\log z)^N$, the innermost sum above can be evaluated asymptotically, by (2.7), as

$$\sum_{\substack{z\le p\le x/m\\ P(m)<p\\ p\in\mathcal{R}(C_fC^{-1})}} 1 = \frac{1}{2h}\int_{\max(P(m),z)}^{x/m} \frac{dt}{\log t} + O\Big(\frac{x}{m\exp(c_6\sqrt{\log z})}\Big).$$

Writing

$$T := \frac{1}{h}\sum_{C\in\mathfrak{C}}\sum_{\substack{m\le M\\ mP(m)\le x}} r_C(m)^\beta \int_{\max(P(m),z)}^{x/m}\frac{dt}{\log t},$$

we have that

$$\sum_{\substack{n\le x\\ n\notin\mathcal{N}}} r_f(n)^\beta \asymp T + O\Big(\frac{x}{\exp((\log x)^{1/3})}\sum_{C\in\mathfrak{C}}\sum_{m\le M}\frac{r_{f(C)}(m)^\beta}{m}\Big).$$

Finally, using the bound $r_f(m)\le\tau(m)$ and the fact that $h\ll D<(\log x)^N$, we find that the error term in the preceding display is negligible. Since T is independent of the particular form chosen, this implies Theorem 4. □

7. Proofs of Theorem 1 and Corollary 1

Let us assume that β is a positive integer and write $K=2^{\beta-1}$. Expanding the character sum (2.3) yields

$$r_f(n)^\beta = \frac{1}{h^\beta}\sum_{\chi_1,\dots,\chi_\beta\in\widehat{\mathfrak{C}}}\prod_{j=1}^{\beta}\bar\chi_j(C_f)\sigma(\chi_j,n). \tag{7.1}$$

Comparing Euler products, we see that for a fixed β-tuple $(\chi_1,\dots,\chi_\beta)$ of characters, we have

$$\sum_{n=1}^{\infty}\frac{\prod_j\sigma(\chi_j,n)}{n^s} = G(s;\chi_1,\dots,\chi_\beta)\prod_{\tau\in\{1\}\times\{\pm1\}^{\beta-1}} L_K\Big(s,\prod_j\chi_j^{\tau_j}\Big). \tag{7.2}$$

Here G is holomorphic in $\mathrm{Re}\,s\ge 1/2+\varepsilon$ with $G(s)=G_1(s)H(s)$, where H is an Euler product, convergent in $\mathrm{Re}\,s>1/2+\varepsilon$, and

$$G_1(s)=\prod_{p\mid D}\prod_{\tau}\Big(1-\frac{\prod_j\chi_j^{\tau_j}(p)}{p^s}\Big)$$

with τ running through the elements of $(\{1\}\times\{\pm1\}^{\beta-1})\setminus\{1\}^\beta$. Thus $G(s)\ne0$, and for $\mathrm{Re}\,s\ge 2/3$, we have

$$\Big|\frac{d^\mu}{ds^\mu}\log G(s;\chi_1,\dots,\chi_\beta)\Big| \ll_\mu \sum_{p\mid D}\frac{(\log p)^\mu}{p^\sigma}+O(1)=o(\log D)$$

for any integer $\mu \geq 0$, which implies that

$$D^{-\varepsilon} \ll_{\varepsilon,\mu} \left|\frac{d^{\mu}}{ds^{\mu}} G(s; \chi_1, \dots, \chi_\beta)\right| \ll_{\varepsilon,\mu} D^{\varepsilon} \tag{7.3}$$

for any integer $\mu \geq 0$. Thus we see that $\sum r_f(n)^\beta n^{-s}$ can be continued holomorphically to the region $\{s \in \mathbb{C} \setminus \{1\} \mid \operatorname{Re} s > 1/2\}$ and has a pole of order at most K at $s = 1$. (We see in a moment that the order of the pole is K.)

Now we use the usual truncated version of Perron's formula (e.g., [5, page 28]):

$$\sum_{n \leq x} d_n = \frac{1}{2\pi i} \int_{v-iT}^{v+iT} \Big(\sum_{n=1}^{\infty} \frac{d_n}{n^s}\Big) \frac{x^s}{s}\, ds + O\Big(x^v \sum_{n=1}^{\infty} \frac{|d_n|}{n^v} \frac{1}{1 + T \log(x/n)}\Big) \tag{7.4}$$

with $d_n = r_f(n)^\beta$, $v = 1 + (\log x)^{-1}$, and

$$T = x^{1/(2^{\beta-1}+2)} D^{-2^{\beta-3}/(2^{\beta-2}+1)}.$$

By (7.1) and (7.2), the function $\sum r_f(n)^\beta n^{-s}$ is a linear combination of terms of the right-hand side of (7.2). We shift the contour to the line $\operatorname{Re} s = 1/2 + \varepsilon$, and we pick up the pole at $s = 1$, which gives the main term in (1.8). By (2.4), the remaining integral and the error term in (7.4) are bounded by

$$\begin{aligned} &x^{1/2+\varepsilon} (D^{1/2} T)^{2^{\beta-2}+\varepsilon} + T^{-1} \int_{1/2+\varepsilon}^{v} x^{\sigma} \big((D^{1/2} T)^{2^{\beta-1}+\varepsilon}\big)^{1-\sigma}\, d\sigma + \frac{x^{1+\varepsilon}}{T} \\ &\quad \ll x^{1/2+\varepsilon} (D^{1/2} T)^{2^{\beta-2}+\varepsilon} + x^{1+\varepsilon} T^{-1}, \end{aligned}$$

which gives the error term in (1.8) for the considered range of D.

Now, let us investigate the coefficients of the main term more closely. We start with the leading coefficient a_K. It is easy to see that exactly the $g^{\beta-1}$ distinct β-tuples $(\chi_1, \dots, \chi_\beta)$ with only real characters χ_j satisfying $\prod \chi_j = \chi_0$ (χ_0 being the principal character) contribute to the coefficient a_K. In fact, to obtain a pole of order K, we need to have $\prod_j \chi_j^{\tau_j} = \chi_0$ for all τ, and so, $\chi_j^2 = \chi_0$ for each $j \geq 2$ (comparing the two cases, where $\tau_i = 1$ for each $i \neq j$) and $\chi_1 = \chi_2 \cdots \chi_\beta$, so that $\chi_1^2 = \chi_0$; on the other hand, if these hold, then $\prod_j \chi_j^{\tau_j} = \prod_j \chi_j \prod_{j:\tau_j=-1} \chi_j^{-2} = \chi_0$ for all τ.

For real characters, (2.2) simplifies to

$$\sigma(\chi, p^\nu) = \begin{cases} \frac{1}{2}\big((-1)^\nu + 1\big), & \big(\frac{-D}{p}\big) = -1, \\ \chi^\nu(\mathfrak{p}), & p \mid D, \\ (\nu + 1)\chi^\nu(\mathfrak{p}), & \big(\frac{-D}{p}\big) = 1. \end{cases}$$

Therefore $\prod_{j=1}^{\beta} \sigma(\chi_j, n) = \sigma(\prod_{j=1}^{\beta} \chi_j, n)^{\beta}$ if all χ_j are real, and hence, we have to study the residue of

$$\frac{g^{\beta-1}}{h^{\beta}} G(s; \chi_0, \ldots, \chi_0) L_K(s, \chi_0)^{2^{\beta-1}} \frac{x^s}{s}$$

at $s = 1$, where $L_K(s, \chi_0) = \zeta(s) L(s, \chi_{-D})$ and

$$\begin{aligned} G(s; \chi_0, \ldots, \chi_0) = &\prod_{\chi_{-D}(p)=-1} \Big(1 - \frac{1}{p^{2s}}\Big)^{2^{\beta-1}-1} \prod_{p \mid D} \Big(1 - \frac{1}{p^s}\Big)^{2^{\beta-1}-1} \\ &\times \prod_{\chi_{-D}(p)=1} \Big(\sum_{k=0}^{\infty} \frac{(k+1)^{\beta}}{p^{ks}}\Big)\Big(1 - \frac{1}{p^s}\Big)^{2^{\beta}}. \end{aligned}$$

This gives (1.5) and (1.6).

We proceed to prove (1.10). To this end, let $2^{j-1} < k \le 2^j$, and fix k distinct tuples $\tau \in \{1\} \times \{\pm 1\}^{\beta-1}$. We claim that there are at most $\ll h^{\beta-1-j+\varepsilon}$ tuples $(\chi_1, \ldots, \chi_\beta)$ such that (7.2) has a pole of order at least k at $s = 1$. Note that these are the only terms in (7.2) that contribute to $a_{\beta,k}$. Indeed, writing the group $\widehat{\mathfrak{C}}$ additively, we have to solve

$$S\mathbf{x} = 0 \tag{7.5}$$

for some $S \in \{\pm 1\}^{k \times \beta}$ having distinct rows and $\mathbf{x} \in \widehat{\mathfrak{C}}^{\beta}$. Let $r_p(S)$ denote the $\mathbb{F}_p$-rank of S. From the proof of Lemma 3.2, we know that a subspace of $\mathbb{F}_p^K$ of dimension d can have at most 2^d distinct vectors with entries only ± 1. Therefore the k row vectors from S generate a subspace of dimension at least j; that is, $r_p(S) \ge j$ for all $p > 2$. By elementary linear algebra and the Chinese remainder theorem, we conclude, as in the proof of Lemma 3.2, that the number of solutions to (7.4) is bounded by $h^{\beta-1-j}$ (the 2-part of $\widehat{\mathfrak{C}}$). From (1.7) and the theorem on finite abelian groups, we conclude that the 2-part of $\widehat{\mathfrak{C}} \ll h^{\varepsilon}$, which establishes the claim. Now we obtain (1.10) by using (1.7) and (7.3) and observing that $\frac{d^{\mu}}{ds^{\mu}} L(s, \chi_{-D})|_{s=1}$, $\frac{d^{\mu}}{ds^{\mu}} L_K(s, \chi)|_{s=1} \ll_{\mu,\varepsilon} D^{\varepsilon}$ for any $\varepsilon > 0$, so that the holomorphic part at $s = 1$ in (7.2) is harmless.

Equation (1.9) can be obtained by selecting $x = \exp((\log D)^2)$, say. For given $\varepsilon > 0$, we choose q in the error term of (1.11) sufficiently large, so that $E_\beta(x, D) \ll x D^{-3/4+\varepsilon}$ and $\sum_{k=2}^{K} a_k (\log x)^k \ll (\log x)^2 D^{-1+\varepsilon}$, by (1.10). Equating (1.8) and (1.11) now gives

$$a_1 x + O\Big(\frac{x \log x}{D^{1-\varepsilon}}\Big) + O(x^{1-\delta}) = \pi\Big(1 + \frac{2^{\beta-1}-1}{u}\Big)\frac{x}{\sqrt{D}} + O\Big(\frac{x}{D^{3/4-\varepsilon}}\Big)$$

for some $\delta > 0$, and we obtain (1.9). This completes the proof of Theorem 1. $\square$

Finally, let us prove the corollary. This follows immediately from Lemma 3.1 for $\beta = 1$. The statement follows from Theorem 2 if $x \le \exp(D^{c_5})$ for sufficiently small c_5. It follows from Theorem 1 (using (1.8) and (1.10)) if $x \ge \exp(D^{c_6})$ for sufficiently large c_6. So, let us now assume that $x = \exp(D^c)$ for some constant $c_5 \le c \le c_6$. We want to show that the terms with $2 \le k \le K-1$ in (1.8) save some power of $\log x$ compared to the terms $k = 1$ and $k = K$, namely,

$$D^{1/2} \sum_{k=2}^{K-1} a_k (\log x)^k \ll \bigl(\log x + D^{-(\beta-1)/2} (\log x)^K\bigr) \frac{1}{(\log x)^\rho}$$

with ρ as in Corollary 1. There is nothing to show for $\beta = 2$, so let us assume that $\beta > 2$. Since $\rho < 1$, we can, by (1.10), ignore the terms $K/2 < k \le K-1$, and using the bound (1.10), this reduces to showing

$$\sum_{j=1}^{\beta-2} \frac{(\log x)^{2^j}}{D^{j/2}} = \sum_{j=1}^{\beta-2} D^{c2^j - j/2} \ll \Bigl(D^c + D^{c2^{\beta-1} - ((\beta-1)/2)}\Bigr) D^{-c\rho + O(\epsilon)}.$$

Since the left-hand side is bounded by $\ll D^{2c-1/2} + D^{c2^{\beta-2} - ((\beta-2)/2)}$, the statement can be checked by a simple calculation. Let us finally note that the error in (1.9) can be absorbed into the error term in (1.4). Indeed, this is clear for $D^{1/4} \ge (\log x)^\rho$. If $D^{1/4} \le (\log x)^\rho$, then a simple calculation shows that $a_1 = O(a_K (\log x)^{K-1} (\log x)^{-\rho+\varepsilon})$. This completes the proof of Corollary 1.

8. Proof of Theorem 5

8.1. The generating function

The proof of (1.14) uses ideas from the article [3], which considers the case where $\beta = 0$. It ultimately relies on the fact that, roughly speaking, numbers with many prime factors can be represented by many classes. This idea has been used in many articles on quadratic forms (see, e.g., [10], which is in a slightly different setting) and was made more precise, on average, in Lemma 3.2. We now define some Dirichlet series. Let

$$Q := \begin{cases} \exp\bigl((\log x)^\varepsilon\bigr) & \text{unconditionally,} \\ \exp\bigl(c_7 (\log\log x)^2\bigr) & \text{if we assume (1.15),} \end{cases} \tag{8.1}$$

and let

$$\mathbb{P} = \mathbb{P}_Q = \bigl\{p \bigm| \chi_{-D}(p) = 1,\, p > Q\bigr\}.$$

With L as in Theorem 5, we may, by (1.7), always assume that

$$D \le (\log x)^{2L \log 2 + 1}. \tag{8.2}$$

Define (for $\mathrm{Re} s > 1$) a modified L-function:

$$\tilde{L}_K(s, Q, \chi) := \prod_{p \in \mathbb{P}} \prod_{\mathfrak{p} | (p)} \exp\Big(\frac{\chi(\mathfrak{p})}{p^s}\Big).$$

Then we define

$$\begin{aligned} P_{C,Q}(s) &:= \frac{1}{2h} \sum_{\chi \in \widehat{\mathfrak{C}}} \bar{\chi}(C) \log \tilde{L}_K(s, Q, \chi) = \frac{1}{2h} \sum_{\chi \in \widehat{\mathfrak{C}}} \bar{\chi}(C) \sum_{p \in \mathbb{P}} \sum_{\mathfrak{p}|(p)} \frac{\chi(\mathfrak{p})}{p^s} \\ &= \frac{1}{2} \sum_{p \in \mathbb{P}} \frac{1}{p^s} \#\big\{\mathfrak{p} \,\big|\, (p) : \mathfrak{p} \in C\big\} = \epsilon(C) \sum_{\substack{p \in \mathcal{R}(C) \\ p \in \mathbb{P}}} \frac{1}{p^s} \\ &=: \frac{1}{2h} \log \zeta(s) + T(s, C), \end{aligned} \tag{8.3}$$

where $\epsilon(C)$ was defined at the end of Section 2. For $k \in \mathbb{N}$, let

$$A_{C,k}(s) := \frac{1}{k!} \sum_{\mathbf{C}=(C_1,\ldots,C_k) \in \mathfrak{C}^k} N_{\mathbf{C}}(C)^{\beta} \prod_{\nu=1}^{k} P_{C_\nu,Q}(s) = \sum_{n=1}^{\infty} \frac{a_{C,k}(n)}{n^s}, \tag{8.4}$$

say, with $N_{\mathbf{C}}(C)$ as in (2.1). If $n = \prod p_j^{e_j}$ is the canonical prime factorization, then the series $(1/k!) \prod_{\nu=1}^{k} P_{C_\nu,Q}(s)$ for some $\mathbf{C} \in \mathfrak{C}^k$ counts a number $n \leftrightarrow \mathbf{C}$ (cf. Section 2 for the notation) with multiplicity $2^{-\delta(\mathbf{C})} \rho(\mathbf{C})^{-1} \prod (e_j!)^{-1}$. Thus Lemma 2.2(b) implies that $a_{C,k}(n) = r_{f(C)}(n)^{\beta}$ if all $p_j \in \mathbb{P}$, n is squarefree and $\Omega(n) = k$, and we always have

$$a_{C,k}(n) \le \frac{2^{\Omega(n)\beta}}{\prod_j e_j!} \le \big(\exp(2^{\beta})\big)^{\omega(n)} \le \tau(n)^{2^{\beta}/\log 2} \le \tau(n)^{2^{\beta+1}}$$

since $\max\{N_{\mathbf{C}}(C)^{\beta} \mid \mathbf{C} \in \mathfrak{C}^k\} \le 2^{k\beta}$. Let us define

$$B_C(s) := \sum_{p|m \Rightarrow p \notin \mathbb{P}} \frac{\tau(m)^{\max(0,\beta-1)} r_{f(C)}(m)^{\beta}}{m^s}. \tag{8.5}$$

If we write $n = lm$, where l contains only prime factors from $\mathbb{P}$ and m contains only prime factors not in $\mathbb{P}$, then

$$r_{f(C)}(n)^{\beta} \le \tau(m)^{\max(0,\beta-1)} \sum_{C_1 C_2 = C} r_{f(C_1)}(l)^{\beta} r_{f(C_2)}(m)^{\beta},$$

by (3.3) and Lemma 2.2(a), since m is represented by at most $\tau(m)$ classes. For $k = 0$, let us set $A_{\tilde{C},0}(s) = 1$ if $\tilde{C} = 1$ is the principal class and set $A_{\tilde{C},0}(s) = 0$ otherwise, and

$$\sum_{k=0}^{K}\sum_{\tilde{C}\in\mathfrak{C}} A_{\tilde{C},k}(s)B_{\tilde{C}^{-1}C}(s) = \sum_{n=1}^{\infty}\frac{b_{C,K}(n)}{n^s}$$

for $K \in \mathbb{N}$. Then $b_{C,K}(n) \geq r_{f(C)}(n)^\beta$ if n is not divisible by the square of a prime $p \in \mathbb{P}$ and $\Omega(n) \leq K$. Furthermore, $0 \leq a_{C,k}(n) \leq b_{C,K}(n) \leq h\tau(n)^{c_8}$, where $c_8 = 2^{\beta+1} + \beta + \max(0, \beta - 1) + 1$ for all $n \in \mathbb{N}$. Now let x be sufficiently large,

$$v := 1 + (\log x)^{-1}, \qquad T := \exp(c_9\sqrt{\log x}), \tag{8.6}$$

for a sufficiently small constant c_8. Perron's formula (see (7.4)) gives

$$\begin{aligned}\sum_{n\leq x} r_f(n)^\beta \geq \frac{1}{2\pi i}\int_{v-iT}^{v+iT} A_{C,k}(s)\frac{x^s}{s}\,ds + O\Big(\frac{x}{\sqrt{T}}\Big)\\ + O\Big(\sum_{p>Q}\sum_{n\leq x,\, p^2|n} \tau(n)^{2^{\beta+1}}\Big)\end{aligned} \tag{8.7}$$

for any $k > 0$, and

$$\begin{aligned}\sum_{n\leq x} r_f(n)^\beta \leq \frac{1}{2\pi i}\int_{v-iT}^{v+iT}\sum_{k=0}^{K}\sum_{\tilde{C}\in\mathfrak{C}} A_{\tilde{C},k}(s)B_{\tilde{C}^{-1}C}(s)\frac{x^s}{s}\,ds + O\Big(\frac{x}{\sqrt{T}}\Big)\\ + O\Big(\sum_{p>Q}\sum_{n\leq x,\, p^2|n}\tau(n)^\beta\Big) + O\Big(\sum_{n\leq x,\,\Omega(n)\geq K}\tau(n)^\beta\Big)\end{aligned} \tag{8.8}$$

for any $K > 0$. The second error term in (8.7) and (8.8) is $\ll x(\log x)^{O(1)}Q^{-1}$, and the third term error term in (8.8) can be estimated by

$$\begin{aligned}\sum_{\substack{n\leq x\\ \Omega(n)\geq K}}\tau(n)^\beta &\leq \sum_{\substack{n\leq x\\ \Omega(n)\geq K\\ \tau(n)\leq S}}\tau(n)^\beta + \sum_{\substack{n\leq x\\ \tau(n)\geq S}}\tau(n)^\beta \leq S^\beta\sum_{\substack{n\leq x\\ \Omega(n)\geq K}} 1 + \frac{1}{S}\sum_{n\leq x}\tau(n)^{\beta+1}\\ &\ll S^\beta\frac{x(\log x)^2}{2^K} + \frac{x(\log x)^{2^{\beta+1}-1}}{S}\end{aligned}$$

for any $S > 0$, where we use [7, Corollary 1] in the last step. Choosing

$$S = (\log x)^{2^{\beta+1}+L\log 2} \quad \text{and} \quad K = \big(3 + (\beta 2^{\beta+1} + 1)L\log 2\big)\log\log x,$$

we can bound the third error term in (8.8) by $O(x(\log x)^{-L\log 2-1})$, which is acceptable since $E(\kappa, \beta) \geq -1 - L\log 2$, by hypothesis. By the choices of T and Q in (8.6) and (8.1), all error terms in (8.7) and (8.8) are admissible.

8.2. Useful estimates

We want to shift the contour in (8.7) and (8.8). Let us write

$$w := 1 - (\log x)^{-1/2}, \qquad r := M(\log x)^{-1} \tag{8.9}$$

for a sufficiently large constant M. Define the path $\Gamma = \Gamma_1^- \Gamma_2^- \Gamma_3^- \Gamma_4 \Gamma_3^+ \Gamma_2^+ \Gamma_1^+$ by

$$\begin{aligned} &\Gamma_1^- = [v - iT, w - iT], \qquad \Gamma_2^- = [w - iT, w], \qquad \Gamma_3^- = [w, -r], \\ &\Gamma_4 = \{re^{it} \mid -\pi \le t \le \pi\}, \\ &\Gamma_3^+ = [-r, w], \qquad \Gamma_2^+ = [w, w + iT], \qquad \Gamma_1^+ = [w + iT, v + iT]. \end{aligned} \tag{8.10}$$

By (2.6) and (8.2), the functions $P_{C,Q}(s)$ extend holomorphically to Γ.

LEMMA 8.1

For $k, K \ll \log\log x$, the integral over $\Gamma_{1,2}^{\pm}$ contributes an error of at most $x \exp(-(\log x)^{1/3})$ to (8.7) and (8.8).

Proof

Let us first observe that $\tilde{L}_K(s, Q, \chi) = L_K(s, \chi) H(s)$, where $H(s)$ is holomorphic in $\mathrm{Re} s > 1/2$ and satisfies $H(s) \ll (\log Q)^2$ if $\Re s \ge 1 - (\log Q)^{-1}$. Thus we obtain, by (2.5) for $s \in \Gamma$,

$$\log \tilde{L}_K(\sigma + it, Q, \chi) \ll \log\log Q + \log D + \log\log(3 + |t|). \tag{8.11}$$

By (8.1)–(8.3), we conclude that $P_{C,Q}(s) \ll \log\log x$ on $\Gamma_{1,2}^{\pm}$, so that

$$A_{C,k}(s) \ll (c_{10} h \log\log x)^k \ll \exp\bigl(c_{11} (\log\log x)^2\bigr) \tag{8.12}$$

on $\Gamma_{1,2}^{\pm}$ for $k \ll \log\log x$. Furthermore, we observe that

$$\begin{aligned} \sum_{C \in \mathfrak{C}} |B_C(s)| &\le \sum_{m : p \mid m \Rightarrow p \notin \mathbb{P}} \frac{\tau(m)^\beta (\sum_C r_{f(C)}(m))^{\beta+1}}{m^\sigma} \\ &\ll \prod_{p \le Q} \Bigl(1 + \frac{2^{\beta+1}}{p}\Bigr) \ll (\log Q)^{c_{12}} \end{aligned} \tag{8.13}$$

on Γ, by (8.1) and (8.5). The lemma now follows easily from (8.12) and (8.13).

For the remaining parts of the integral, we need the following lemma. □

LEMMA 8.2

On $\Gamma_3^- \Gamma_4 \Gamma_3^+$, we have $T(s, C) = T(1, C) + O((\log x)^{-1/3}/h)$ with $T(1, C) \ll \log\log Q/h$.

Proof

For any $\mu \geq 0$, the Dirichlet series for $T^{(\mu)}(s, C)$ converges (conditionally) at $s = 1$, and

$$T^{(\mu)}(s, C) = \epsilon(C) \sum_{\substack{p \in \mathcal{R}(C) \\ p \in \mathbb{P}}} \frac{(-\log p)^{\mu}}{p^s} - \frac{1}{2h} \frac{d^{\mu}}{ds^{\mu}} \log \zeta(s) = \sum_{m=1}^{\infty} \frac{t_{\mu}(m)}{m^s},$$

say. Choosing $A = (2L \log 2 + 1)/\varepsilon$ in (2.8) and observing (8.1) and (8.2), we conclude for $\xi \geq Q$, from (2.7)–(2.9) by partial summation,

$$\sum_{m \leq \xi} \frac{t_{\mu}(m)}{m} \ll \frac{1}{2h} \int_2^Q \frac{(\log y)^{\mu-1}}{y}\, dy + O\Big(\frac{1}{h}\Big) + O\Big(\exp\Big(-\frac{c_3}{2}\sqrt{\log Q}\Big)\Big).$$

Since the right-hand side is independent of ξ, we get

$$T(1, C) \ll \frac{\log\log Q}{h} \qquad \text{and} \qquad T^{(\mu)}(1, C) \ll \frac{(\log Q)^{\mu}}{h}$$

for $\mu \geq 1$. Since $|s - 1| \leq (\log x)^{-1/2}$ on $\Gamma_3^- \Gamma_4 \Gamma_3^+$, we can apply Taylor's formula about $s = 1$ up to degree $\mu_0 := \lceil 2(\log 2)L \rceil + 1$ to estimate $T(s, C)$. We use the trivial estimation

$$\begin{aligned} |T^{(\mu_0)}(s, C)| &\leq \max_{\chi \neq \chi_0} \Big| \frac{d^{\mu_0}}{ds^{\mu_0}} \log \tilde{L}_K(s, Q, \chi) \Big| + \frac{1}{h} \Big| \frac{d^{\mu_0}}{ds^{\mu_0}} \log \frac{\tilde{L}_K(s, Q, \chi_0)}{\zeta(s)} \Big| \\ &\ll (\log x)^{\varepsilon} \end{aligned}$$

on $\Gamma_3^- \Gamma_4 \Gamma_3^+$, which follows easily from (2.6), (8.2), (8.11), and Cauchy's integral formula (see, e.g., [3, (2.9)]). This gives

$$T(s, C) - T(1, C) \ll \frac{(\log x)^{\varepsilon}}{h\sqrt{\log x}} + \frac{(\log x)^{\varepsilon}}{(\log x)^{\mu_0/2}} \ll \frac{(\log x)^{-1/3}}{h}.$$

Hence the lemma is concluded. □

8.3. The upper bound

By (8.3), (8.4), Lemma 3.2, and the two preceding lemmata, we have

$$A_{C,k}(s) \ll \begin{cases} \frac{1}{k!} \frac{1}{h} \big(\log|\zeta(s)| + c_{13} \log\log Q \big)^k, & k \leq \frac{\log h/g}{\log 2}, \\ \frac{1}{k!} \frac{2^{(\beta-1)k}}{h^{\beta}} \big(\log|\zeta(s)| + c_{13} \log\log Q \big)^k, & k \geq \frac{\log h/g}{\log 2}, \end{cases}$$

on $\Gamma_3^- \Gamma_4 \Gamma_3^+$, so that by (8.8) and (8.13) up to an admissible error,

$$
\begin{aligned}
&\sum_{n\le x} r_f(n)^\beta \\
&\ll (\log Q)^{c_{12}} \int_{\Gamma_3^-\Gamma_4\Gamma_3^+} x^\sigma \Big(\sum_{k\le(\log h/g)/(\log 2)} \frac{1}{k!}\frac{1}{h} (\log|\zeta(s)| + c_{13}\log\log Q)^k \\
&\qquad + \sum_{(\log h/g)/(\log 2)\le k\le K} \frac{1}{k!} \frac{2^{(\beta-1)k} g^{\beta-1}}{h^\beta} \big(\log|\zeta(s)| + c_{13}\log\log Q\big)^k \Big) \, |ds| \\
&\ll \frac{(\log Q)^{c_{12}}}{g} \Big(\int_{\Gamma_3^\pm} x^\sigma \, |d\sigma| + \frac{x}{\log x}\Big)\Big(\frac{g}{h} \sum_{k\le(\log h/g)/(\log 2)} \frac{(\log\log x + c_{13}\log\log Q)^k}{k!} \\
&\qquad + \frac{g^\beta}{h^\beta} \sum_{(\log h/g)/(\log 2)\le k\le K} \frac{(2^{\beta-1}(\log\log x + c_{13}\log\log Q))^k}{k!}\Big) \\
&\ll \frac{x(\log Q)^{c_{14}}}{g\log x} \bigg((\log x)^{-\kappa\log 2} \max_{k\le\kappa\log\log x} \Big(\frac{e\log\log x}{k}\Big)^k \\
&\qquad + (\log x)^{-\beta\kappa\log 2} \max_{\kappa\log\log x\le k\le K} \Big(\frac{e2^{\beta-1}\log\log x}{k}\Big)^k \bigg),
\end{aligned}
$$

where we use Stirling's formula and the definition of κ in Theorem 5 together with (1.7) and (8.1). Since

$$
\max_{k\le\kappa\log\log x} \Big(\frac{e\log\log x}{k}\Big)^k \ll \begin{cases} \log x, & \kappa\ge 1, \\ (\log x)^{\kappa(1-\log\kappa)}, & \kappa\le 1, \end{cases}
$$

and

$$
\max_{\kappa\log\log x\le k\le K} \Big(\frac{e2^{\beta-1}\log\log x}{k}\Big)^k \ll \begin{cases} (\log x)^{2^{\beta-1}}, & \kappa\le 2^{\beta-1}, \\ (\log x)^{\kappa(1+(\beta-1)\log 2-\log\kappa)}, & \kappa\ge 2^{\beta-1}, \end{cases}
$$

the upper bound of (1.14) and (1.16) follows after a short calculation.

8.4. The lower bound

By (8.7) and Lemma 8.1 we have, for any $\log\log x \ll k \ll \log\log x$ up to an admissible error,

$$
\begin{aligned}
\sum_{n\le x} r_f(n)^\beta \ge \frac{x}{2\pi}\frac{1}{k!} \sum_{\mathbf{C}\in\mathfrak{C}^k} N_{\mathbf{C}}(C_f)^\beta \bigg(\int_{-\pi}^{\pi} \prod_{\nu=1}^k P_{C_\nu,Q}(1+re^{it}) \frac{re^{it}x^{re^{it}}}{1+re^{it}}\,dt \\
+ O\Big(\int_r^{(\log x)^{-1/2}} x^{-t} \prod_{\nu=1}^k |P_{C_\nu,Q}(1-t)|\,dt\Big)\bigg).
\end{aligned}
\tag{8.14}
$$

By (8.3) and Lemma 8.2 there are some real numbers $a_\nu = 2hT(1, C_\nu)$, independent of t, such that $|a_\nu| \le c_{13} \log\log Q$. Since $\zeta(1+s) = 1/s + O(1)$, we get

$$(2h)^k \prod_{\nu=1}^{k} P_{C_\nu,Q}(1+re^{it}) = \prod_{\nu=1}^{k} \Big(\log\frac{1}{r} - it + O(r) + a_\nu + O\Big(\frac{1}{(\log x)^{1/3}}\Big)\Big),$$

the O-constants being absolute. Since $r \asymp (\log x)^{-1}$, the right-hand side has absolute value

$$\prod_{\nu=1}^{k} \Big(\log\frac{1}{r} + a_\nu + O\big((\log\log x)^{-1}\big)\Big) = \prod_{\nu=1}^{k} \Big(\Big(\log\frac{1}{r} + a_\nu\Big)\big(1 + O((\log\log x)^{-2})\big)\Big)$$

and argument

$$\arctan\Big(-\frac{t}{\log(1/r)+a_\nu} + O\big((\log x)^{-1/3}\big)\Big) = -\sum_{\nu=1}^{k}\Big(\frac{t}{\log 1/r + a_\nu} + O\big((\log\log x)^{-2}\big)\Big).$$

For brevity we write $B := \sum_{\nu=1}^{k} (1/\log(1/r) + a_\nu)$. Note that $B \asymp 1$ for $\log\log x \ll k \ll \log\log x$. This gives

$$\prod_{\nu=1}^{k} P_{C_\nu,Q}(1+re^{it}) = (2h)^{-k}\Big(1 + O\Big(\frac{1}{\log\log x}\Big)\Big) \prod_{\nu=1}^{k}\Big(\log\frac{1}{r} + a_\nu\Big) \exp(-itB).$$

Clearly, $(1+re^{it})^{-1} = 1 + O(r)$. Thus the first integral in (8.14) equals

$$\Big(1 + O\Big(\frac{1}{\log\log x}\Big)\Big) \frac{r}{(2h)^k} \prod_{\nu=1}^{k}\Big(\log\frac{1}{r} + a_\nu\Big) \int_{-\pi}^{\pi} \exp\big(it(1-B) + r(\log x)e^{it}\big)\, dt.$$

This last integral can be interpreted as the contour integral

$$\frac{(r\log x)^{B-1}}{i} \int e^s s^{-B}\, ds \gg_B (r\log x)^{B-1} \tag{8.15}$$

over the circle $\{r(\log x)e^{it} \mid -\pi \le t \le \pi\}$. Recalling that $r = M(\log x)^{-1}$, we see that the first integral in (8.14) is bounded below by

$$\gg \frac{M^B}{(2h)^k \log x} \prod_{\nu=1}^{k}\Big(\log\frac{1}{r} + a_\nu\Big). \tag{8.16}$$

To estimate the error term in (8.14), we note that

$$\prod_{\nu=1}^{k} |P_{C_\nu,Q}(1-t)| \ll (2h)^{-k} \prod_{\nu=1}^{k} \Big(\log \frac{1}{t} + a_\nu \Big) \le (2h)^{-k} \prod_{\nu=1}^{k} \Big(\log \frac{1}{r} + a_\nu \Big)$$

on $\Gamma_3^{\pm}$ and that $\int_r^{(\log x)^{-1/2}} x^{-t}\, dt \ll e^{-M} (\log x)^{-1}$. Selecting M sufficiently large, we see that the error term is dominated by (8.16), and by Lemma 8.2, we obtain

$$\sum_{n \le x} r_f(n)^\beta \gg \frac{x}{(2h)^k k! \log x} \sum_{\mathbf{C} \in \mathfrak{C}^k} N_{\mathbf{C}}(C)^\beta (\log\log x - c_{13} \log\log Q)^k$$

for $\log\log x \ll k \ll \log\log x$. Using Lemma 3.2 and Stirling's formula, we obtain the lower bound by exactly the same calculation that led to the upper bound in Section 8.3. This completes the proofs of (1.14), (1.16), and Theorem 5. $\square$

Is the statement of the conclusion of the proof of Th. 5 at the end of Sec. 8.4 Okay?

9. Proofs of Theorem 6 and Corollary 2

The proof of Theorem 6 is a variant of the proof of Theorem 5; here we use Lemma 3.3 instead of Lemma 3.2, and we also use slightly different generating Dirichlet series.

9.1. The upper bound

Let $\mathcal{M} = \{m_1 m_2 : m_1 \text{ powerful}, m_2 \mid D\}$. First, let us observe that we can write each positive integer n as $n = lm$ with $(l, Dm) = 1$, $\mu^2(l) = 1$ and $m \in \mathcal{M}$. From Lemma 2.2 we conclude, similarly as in the proof of Theorem 5 (cf. (8.4)), that the coefficients a_n, say, of the Dirichlet series

$$\sum_{C \in \mathfrak{C}} \sum_{k \le K} \frac{1}{k!} \sum_{\mathbf{C} \in \mathfrak{C}^k} N_{\mathbf{C}}(C)^0 \prod_{\nu=1}^{k} P_{C_\nu,1}(s) \sum_{m \in \mathcal{R}(C^{-1}C_0) \cap \mathcal{M}} \frac{1}{m^s}$$

satisfy

$$\sum_{n \le x} a_n \ge \sum_{n \le x} r_{f(C_0)}(n)^0 + O\Big(\sum_{n \le x, \Omega(n) > K} 1 \Big) \tag{9.1}$$

for any $K \in \mathbb{N}$, $C_0 \in \mathfrak{C}$. (Note that we have now chosen $Q = 1$.) Clearly,

$$\begin{aligned} \sum_{C \in \mathfrak{C}} \sum_{m \in \mathcal{R}(C) \cap \mathcal{M}} \frac{1}{m^s} &\ll \prod_{p \mid D} \Big(1 + \frac{1}{p^\sigma} \Big) \prod_{p \nmid D} \Big(1 + \frac{O(1)}{p^{2\sigma}} \Big) \\ &\ll \prod_{p \mid D} \Big(\Big(1 + \frac{1}{p} \Big) \Big(1 + \frac{O(|1-\sigma| \log p)}{p} \Big) \Big) \ll \prod_{p \mid D} \Big(1 + \frac{1}{p} \Big) \end{aligned} \tag{9.2}$$

on Γ. As in (8.8), we see that the error term in (9.1) is admissible if $K > (L+3) \log\log x$, say. By Theorem 4, (9.1) also holds—up to a constant—for the coefficients of

$$\frac{1}{h}\sum_{C_0\in\mathfrak{C}}\Big(\sum_{C\in\mathfrak{C}}\sum_{k\le K}\frac{1}{k!}\sum_{\mathbf{C}\in\mathfrak{C}^k}N_{\mathbf{C}}(C)^0\prod_{\nu=1}^{k}P_{C_\nu,1}(s)\sum_{m\in\mathcal{R}(C^{-1}C_0)\cap\mathcal{M}}\frac{1}{m^s}\Big)$$
$$=\frac{1}{h}\sum_{k\le K}\frac{1}{k!}\sum_{C\in\mathfrak{C}}\sum_{\mathbf{C}\in\mathfrak{C}^k}N_{\mathbf{C}}(C)^0\prod_{\nu=1}^{k}P_{C_\nu,1}(s)\sum_{C\in\mathfrak{C}}\sum_{m\in\mathcal{R}(C)\cap\mathcal{M}}\frac{1}{m^s}.$$

Thus by Lemma 3.3, (9.1) holds, a fortiori, for the coefficients of

$$\sum_{k\le K}\frac{1}{k!}\min\Big(\frac{2^k}{h},\frac{1}{g}\Big)\Big(\sum_{C\in\mathfrak{C}}P_{C,1}(s)\Big)^k\sum_{C\in\mathfrak{C}}\sum_{m\in\mathcal{R}(C)\cap\mathcal{M}}\frac{1}{m^s}$$
$$=\sum_{k\le K}\frac{1}{k!}\min\Big(\frac{2^k}{h},\frac{1}{g}\Big)\Big(\frac{1}{2}\log\tilde{L}_K(s,1,\chi_0)\Big)^k\sum_{C\in\mathfrak{C}}\sum_{m\in\mathcal{R}(C)\cap\mathcal{M}}\frac{1}{m^s}. \tag{9.3}$$

Let us now apply Perron's formula (see (7.4)) to the series (9.3) with the contour given by (8.6), (8.9), and (8.10). As in Lemma 8.1, we see that (8.12) holds for (9.3) on $\Gamma_{1,2}^{\pm}$, so this part of the path is negligible. To estimate the contribution of $\Gamma_3^-\Gamma_4\Gamma_3^+$, we proceed exactly as in Section 8.3. Up to an admissible error, we have, with the definition $\lambda := \log(h/g)/(\log 2)$ as in Lemma 3.3,

$$\sum_{n\le x}r_f(n)^0\ll\frac{D}{\phi(D)}\int_{\Gamma_3^-\Gamma_4\Gamma_3^+}x^\sigma\Big(\sum_{k\le\lambda}\frac{1}{k!}\frac{1}{h}\big(\log|\tilde{L}_K(s,1,\chi_0)|\big)^k$$
$$+\sum_{\lambda\le k\le K}\frac{1}{k!}\frac{1}{g}\Big(\frac{1}{2}\log|\tilde{L}_K(s,1,\chi_0)|\Big)^k\Big)\,|ds|.$$

Note that

$$\frac{\tilde{L}_K(s,1,\chi_0)}{\zeta(s)}\,\Big|_{s=1}\asymp L(1,\chi_{-D})\frac{\phi(D)}{D}=\ell$$

with the notation as in Theorem 6, so that $\log\tilde{L}_K(s,1,\chi_0)\le\log\log x+\log\ell+O(1)$ on $\Gamma_3^-\Gamma_4\Gamma_3^+$. By Stirling's formula,

$$\sum_{k\le\kappa U}\frac{1}{k!}U^k\asymp\begin{cases}e^U, & \kappa\ge 1,\\ (e^U)^{\kappa(1-\log\kappa)}\min\Big(1,\frac{1}{(1-\kappa)\sqrt{\kappa U}}\Big), & \kappa<1,\end{cases}$$

and

$$\sum_{k\ge\kappa U}\frac{1}{k!}U^k\asymp\begin{cases}e^U, & \kappa\le 1,\\ (e^U)^{\kappa(1-\log\kappa)}\min\Big(1,\frac{1}{(\kappa-1)\sqrt{\kappa U}}\Big), & \kappa>1,\end{cases}$$

for any $U>1$, $\kappa>0$, which gives the upper bounds in (1.1), (1.2), and (1.3).

9.2. The lower bound

Fix a small number $\varepsilon > 0$ as in Theorem 6, and assume that $\kappa \leq 1/2 - \varepsilon$. Let us define $\mathcal{K} = [(1/2 - \varepsilon^2) \log\log x, \log\log x] \cap \mathbb{N}$, and $Q = \exp(D^{\varepsilon^*})$ with $\varepsilon^* = \min((\varepsilon/18)/(6c_{14} + 1), \varepsilon^2/4)$. Note that with this choice, we are still able to apply Lemma 8.2. Let $\mathcal{Q}$ be the set of squarefree integers whose prime factors p satisfy $\chi_{-D}(p) \neq -1$ and $p \leq Q$, and for $m \in \mathcal{Q}$, let $r^*(m)$ be the number of classes $C \in \mathfrak{C}$ which can represent m. As in Sections 8.1 and 9.1 (cf. (8.4)), we see that the coefficients of

$$\sum_{k \in \mathcal{K}} \sum_{\tilde{C} \in \mathfrak{C}} A_{\tilde{C},k}(s) \sum_{m \in \mathcal{R}(C\tilde{C}^{-1}) \cap \mathcal{M}} \frac{1/r^*(m)}{m^s}$$

with $\beta = 0$ minorize $r_{f(C)}(n)^0$ for any $C \in \mathfrak{C}$. Using Theorem 4 as before, we find that the coefficients a_n, say, of

$$\begin{aligned} &\frac{1}{h} \sum_{C \in \mathfrak{C}} \sum_{k \in \mathcal{K}} \frac{1}{k!} \sum_{\mathbf{C} \in \mathfrak{C}^k} \sum_{\tilde{C} \in \mathfrak{C}} N_{\mathbf{C}}(C)^0 \prod_{\nu=1}^{k} P_{C_\nu, Q}(s) \sum_{m \in \mathcal{R}(\tilde{C}C^{-1}) \cap \mathcal{Q}} \frac{1/r^*(m)}{m^s} \\ &= \sum_{k \in \mathcal{K}} \frac{1}{k!} \sum_{\mathbf{C} \in \mathfrak{C}^k} \frac{1}{h} \sum_{C \in \mathfrak{C}} N_{\mathbf{C}}(C)^0 \prod_{\nu=1}^{k} P_{C_\nu, Q}(s) \sum_{m \in \mathcal{Q}} \frac{1}{m^s} \end{aligned}$$

satisfy

$$\sum_{n \leq x} a_n \leq \frac{1}{h} \sum_{C \in \mathfrak{C}} \sum_{\substack{n \leq x \\ n \in \mathcal{R}(C)}} 1 \ll \sum_{n \leq x} r_{f(C_0)}(n)^0 \tag{9.4}$$

for any $C_0 \in \mathfrak{C}$. Now let us apply Lemma 3.3 with $\tilde{\varepsilon} = \varepsilon/18$. Then all $k \in \mathcal{K}$ satisfy

$$(1 + 18\tilde{\varepsilon})\Big(\frac{1}{2} - \varepsilon\Big)(\log\log x + \log \ell) \leq (1 + \varepsilon)\Big(\frac{1}{2} - \varepsilon\Big)\Big(1 + \frac{1}{2}\varepsilon\Big) \log\log x \leq k,$$

so that we can decompose $\mathfrak{C}^k = \mathcal{C}_1 \cup \mathcal{C}_2$ such that

$$\sum_{C \in \mathfrak{C}} N_{\mathbf{C}}(C)^0 \gg \min\Big(2^k, \frac{h}{g}\Big) = \frac{h}{g}$$

for all $\mathbf{C} \in \mathcal{C}_1$ and $\#\mathcal{C}_2 \ll h^k D^{-\tilde{\varepsilon}}$. Therefore the coefficients of the series

$$\begin{aligned} &\frac{1}{g} \sum_{k \in \mathcal{K}} \frac{1}{k!} \sum_{\mathbf{C} \in \mathfrak{C}^k} \prod_{\nu=1}^{k} P_{C_\nu, Q}(s) \sum_{m \in \mathcal{Q}} \frac{1}{m^s} \\ &\quad - \frac{1}{g} \sum_{k \in \mathcal{K}} \frac{1}{k!} \sum_{\mathbf{C} \in \mathcal{C}_2} \prod_{\nu=1}^{k} P_{C_\nu, Q}(s) \sum_{m \in \mathcal{Q}} \frac{1}{m^s} =: A_1(s) - A_2(s), \end{aligned} \tag{9.5}$$

say, satisfy, a fortiori, (9.4) for all $C_0 \in \mathfrak{C}$. Clearly,

$$A_1(s) = \frac{1}{g} \sum_{k \in \mathcal{K}} \frac{1}{k!} \Big(\frac{1}{2} \log \tilde{L}_K(s, Q, \chi_0) \Big)^k \sum_{m \in \mathcal{Q}} \frac{1}{m^s},$$

by orthogonality. Let us now apply Perron's formula (see (7.4)) to both terms in (9.5) with the contour given by (8.6), (8.9), and (8.10). As in Lemma 8.1, we see that (8.12) holds for (9.5) on $\Gamma^{\pm}_{1/2}$, so this part of the path is negligible. As in (9.2), we see that

$$\sum_{m \in \mathcal{Q}} \frac{1}{m^s} \ll \prod_{p \le Q} \Big(1 + \frac{1}{p} \Big) \ll \log Q$$

for the considered s, so exactly the same calculation as in Section 8.3 shows that the contribution of $A_2(s)$ is at most

$$\frac{(\log Q)^{c_{14}+1}}{g D^{\tilde{\varepsilon}}} \frac{x}{\sqrt{\log x}} \ll \frac{(\log Q)^{c_{14}+1}}{D^{\tilde{\varepsilon}/2}} \frac{\sqrt{D L(1, \chi_{-D})}}{g \sqrt{\phi(D)}} \frac{x}{\sqrt{\log x}} \ll \frac{\text{main term}}{D^{\tilde{\varepsilon}/3}}.$$

The integral over $A_1(s)$ can be estimated exactly as in Section 8.4 with

$$a_\nu := \log \big(\operatorname{res}_{s=1} \tilde{L}_K(s, Q, \chi_0) \big) + O\Big(\frac{D^\varepsilon}{r} \Big) = \log \ell - \sum_{p \le Q} \frac{1 + \chi_{-D}(p)}{p} + O(1).$$

The estimate (8.15) becomes

$$\frac{(r \log x)^{B-1}}{i} \int e^s s^{-B} \prod_{p \in \mathcal{Q}} \Big(1 + \frac{1}{p^{1+s/(\log x)}} \Big) \, ds \gg (r \log x)^{B-1} \prod_{p \in \mathcal{Q}} \Big(1 + \frac{1}{p} \Big),$$

and we arrive as in Section 8.4 at

$$\begin{aligned} &\sum_{n \le x} r_f(n)^0 \\ &\quad \gg \frac{x}{\log x} \prod_{p \in \mathcal{Q}} \Big(1 + \frac{1}{p} \Big) \sum_{k \in \mathcal{K}} \frac{1}{k! g} \Big(\frac{1}{2} \Big(\log \log x + \log \ell - \sum_{p \le Q} \frac{1 + \chi_{-D}(p)}{p} + O(1) \Big) \Big)^k \\ &\quad \asymp \frac{1}{\tau(D)} \prod_{p \in \mathcal{Q}} \Big(1 + \frac{1}{p} \Big) \frac{x}{\log x} \Big(\ell (\log x) \prod_{p \in \mathcal{Q}} \Big(1 - \frac{1 + \chi_{-D}(p)}{p} \Big) \Big)^{1/2}, \end{aligned}$$

which gives the lower bound in (1.1). For the last step, note that by our choice of Q, we have

$$\frac{1}{2} \Big(\log \log x + \log \ell - \sum_{p \le Q} \frac{2}{p} + O(1) \Big) \ge \Big(\frac{1}{2} - \frac{\varepsilon^2}{2} \Big) \log \log x + O_\varepsilon(1),$$

which is larger than the lower bound of $\mathcal{K}$ for large x.

9.3. Sums of two powerful numbers

Corollary 2 can be proved as in [3, Section 5] by using our refined estimates (1.16) and the upper bound in (1.2). The main idea is that a powerful number n can be written uniquely as $n = a^3b^2$ with $\mu^2(a) = 1$. Thus

$$V(x) = \#\bigl\{1 \le n \le x \bigm| a_1^3x_1^2 + a_2^3x_2^2 \text{ represents } n \text{ for some } (a_1, a_2) = 1\bigr\}.$$

All results from Section 2 as well as (1.7) also hold for nonfundamental discriminants Df^2, except that numbers $m \mid f^\infty$ behave differently; a prime $p \mid f$ cannot be represented, while [12, main theorem] states that p^α can be represented by at most $p^{[\alpha/2]}$ classes. Thus (1.16) carries over to nonfundamental discriminants, as do (1.1)–(1.3) since (9.2) remains unchanged.

The lower bound in Corollary 2 follows from (1.16), as in [3, Section 5], by considering the quadratic forms $x_1^2 + p^3x_2^2$ for primes $p \equiv 3 \pmod 4$ so that $L(s, \chi_{-4p})$ has no Siegel zero and $(\log x)^{(2^{2/3}/3)\log 2} \le p \le 2(\log x)^{(2^{2/3}/3)\log 2}$.

Let us now turn toward the upper bound and use the notation from Theorem 6. Using the trivial bound

$$\sum_{n \le x} r_f(n)^0 \ll \min\Bigl(\frac{x}{(\log x)^{1/2-\varepsilon}}, \frac{x}{\sqrt{D}} + \sqrt{D}\Bigr),$$

it is easy to show that the contribution of forms $a_1^3x_1^2 + a_2^3x_2^2$ with $\kappa \le 1/2$ or $\kappa \ge 1$, that is, $(a_1a_2)^3 \le (\log x)^{\log 2+\varepsilon}$ or $(a_1a_2)^3 \ge (\log x)^{2\log 2+\varepsilon}$, is negligible; in fact, it is $\ll x(\log x)^{\log 2/3+\varepsilon}$ (cf. [3]). Let us now consider the contribution of the remaining forms. Since

$$(\ell \log x)^{\kappa \log 2} = \frac{h}{g} \asymp \frac{D^{1/2}L(1, \chi_{-D})}{\tau(D)},$$

we have, by the upper bound in (1.2), that

$$\begin{aligned}\sum_{n \le x} r_f(n)^0 &\ll \frac{x}{\ell \log x\sqrt{\log\log x}}\Bigl(\frac{L(1, \chi_{-D})}{\tau(D)}\Bigr)^{1/3}\frac{1}{D^{1/3}}(\ell \log x)^{\kappa(1-\log(2^{1/3}\kappa))}\\ &\ll \Bigl(\frac{D}{\phi(D)}\Bigr)^{1-2^{-1/3}} L(1, \chi_{-D})^{2^{-1/3}-2/3}\frac{x(\log x)^{-2^{-1/3}+\kappa(1-\log(2^{1/3}\kappa))}}{(\log x)^{1-2^{-1/3}}(\tau(D)D)^{1/3}\sqrt{\log\log x}}\\ &\le \bigl(1 + L(1, \chi_{-D})\bigr)\frac{x(\log x)^{-2^{-1/3}+\kappa(1-\log(2^{1/3}\kappa))}}{(\log x)^{1-2^{-1/3}}(\tau(D)\phi(D))^{1/3}\sqrt{\log\log x}}\end{aligned}$$

if $1/2 < \kappa < 1$. Here we used $0 < 2^{-1/3} - 2/3 < 1$. Thus the contribution of the remaining forms is at most

$$\sum_{(\log x)^{\log 2/3-\varepsilon} \le d \le (\log x)^{2\log 2/3+\varepsilon}} \big(1 + L(1, \chi_{-d})\big) \frac{\tau(d)^{2/3}}{\phi(d)} \frac{x(\log x)^{-2^{-1/3}+\kappa(1-\log(2^{1/3}\kappa))}}{(\log x)^{1-2^{-1/3}}\sqrt{\log\log x}}, \tag{9.6}$$

where $\kappa = \kappa_d$ refers to the discriminant $-4d^3$. One can easily see that

$$-2^{-1/3} + \kappa\big(1 - \log(2^{1/3}\kappa)\big) \le -\frac{(\kappa - 2^{-1/3})^2}{2}$$

for all $1/2 \le \kappa \le 1$; by definition and the class number formula,

$$\kappa = \frac{3}{2\log 2} \frac{(\log d + O(\log L(1, \chi_{-d}) + \log \tau(d)))}{(\log\log x + O(\log \ell))},$$

so that

$$(\log x)^{-2^{-1/3}+\kappa(1-\log(2^{1/3}\kappa))} \le \exp\Big(-\big(1 + o(1)\big)\frac{\log(d/(\log x)^{\theta})^2}{2^{5/3}\theta^2 \log\log x}\Big), \tag{9.7}$$

where $\theta = (2^{2/3}/3)\log 2$, provided that $\log d \asymp \log\log x$ and $|\log L(1, \chi_{-d})| + \log\tau(d) + \log(d/\phi(d)) = o(\sqrt{\log\log x})$. We always have $1 \le d/\phi(d) \ll \log\log d \ll \log\log\log x$. The only way that $|\log L(1, \chi_{-d})| \gg \log\log\log x$ is if we have a Siegel zero, and this happens for just $O(1)$ values of d in such a range. However, such terms contribute $\ll 1/d^{1/2}$ to the sum, so they are negligible. The exceptional d are those with $\tau(d)$ large, which we take to mean greater than $(\log\log x)^A$. By the Cauchy-Schwarz inequality and the known bounds for moments of $L(1, \chi_{-D})$,

$$\begin{aligned}
&\Big(\sum_{\substack{\tau(d)>(\log\log x)^A \\ d=(\log x)^{O(1)}}} \big(1 + L(1, \chi_{-d})\big)\frac{\tau(d)^{2/3}}{\phi(d)}\Big)^2 \\
&\quad\le \sum_{d=(\log x)^{O(1)}} \frac{(1 + L(1, \chi_{-d}))^2}{d} \sum_{\substack{d=(\log x)^{O(1)} \\ \tau(d)>(\log\log x)^A}} \frac{d\tau(d)^{4/3}}{\phi(d)^2} \\
&\quad\ll \frac{\log\log x}{(\log\log x)^A} \sum_{d=(\log x)^{O(1)}} \frac{d\tau(d)^{7/3}}{\phi(d)^2} \ll \frac{1}{(\log\log x)^{A-2^{7/3}-1}}.
\end{aligned}$$

We thus get a negligible contribution to the above sum if A is chosen sufficiently large since $-2^{-1/3} + \kappa(1 - \log(2^{1/3}\kappa)) \le 0$.

For the remaining terms d, we split the sum up into dyadic intervals $(y, 2y]$ (with y a power of 2). By the Polya-Vinogradov and the Cauchy-Schwarz inequalities, we have

$$\sum_{y<d\le 2y}\big(1+L(1,\chi_{-d})\big)\frac{\tau(d)^{2/3}}{\phi(d)}=\sum_{y<d\le 2y}\bigg(\Big(\sum_{n=1}^{y^{1/2+\varepsilon}}\frac{(-d/n)}{n}\Big)+O(1)\bigg)\frac{\tau(d)^{2/3}}{\phi(d)}$$
$$\ll\sum_{n=1}^{\log y}\frac{1}{n}\sum_{y<d\le 2y}\frac{(-d/n)\tau(d)^{2/3}}{\phi(d)}+\Big(\sum_{y<d\le 2y}\frac{\tau(d)^{4/3}}{\phi(d)^2}\cdot\sum_{y<d\le 2y}\Big|\sum_{n=\log y}^{y^{1/2+\varepsilon}}\frac{(-d/n)}{n}\Big|^2\Big)^{1/2}$$
$$+(\log y)^{2^{2/3}-1}. \tag{9.8}$$

Let us consider the second term of the right-hand side. The first sum inside the brackets is $\ll y^{-1}(\log y)^{2^{4/3}-1}$. For the second, we expand to get

$$\sum_{m,n=\log y}^{N}\frac{1}{mn}\sum_{y<d\le 2y}\Big(-\frac{d}{mn}\Big).$$

For the terms where mn is not a square, the final term is $\ll (mn)^{1/2}\log y$, by the Polya-Vinogradov inequality, and thus, their total contribution is

$$\ll\Big(\sum_{n\le y^{1/2+\varepsilon}}\frac{1}{\sqrt{n}}\Big)^2\log y=y^{1/2+\varepsilon}.$$

When mn is a square, write $a=(m,n)$ and $m=ar^2, n=as^2$, so that the total contribution of these terms is

$$\le y\sum_{a\le y^{1/2+\varepsilon}}\sum_{r,s=\sqrt{(\log y)/a}}^{\sqrt{y^{1/2+\varepsilon}/a}}\frac{1}{(ars)^2}$$
$$\ll y\sum_{a\le\log y}\frac{1}{a^2}\Big(\frac{1}{\sqrt{(\log y/a)}}\Big)^2+y\sum_{\log y\le a\le y^{1/2+\varepsilon}}\frac{1}{a^2}$$
$$\ll\frac{y\log\log y}{\log y}.$$

Finally, let us consider the first term on the right-hand side of (9.8). The contribution of n that are squares is at most $\ll(\log y)^{2^{2/3}-1}$. If n is a nonsquare, then $(-d/n)\tau(d)^{2/3}d/\phi(d)$ is the coefficient of $L(s,\chi)^{2^{2/3}}A_n(s)$, where $A_n(s)$ is absolutely convergent in $\mathrm{Re}(s)>1/2$ and χ is a nonprincipal character. By the Siegel-Walfisz theorem, the d-sum is at most $\exp(-c_{15}\sqrt{\log y})$ for some $c_{15}>0$, so that the total contribution of nonsquare n is negligible. Altogether, we find that (9.8) is bounded above by

$$(\log y)^{2^{2/3}-1} = (\log\log x)^{2^{2/3}-1}. \tag{9.9}$$

The contribution of the $d \in (y, 2y]$ to the sum (9.6) thus depends by (9.7) only on the value of $\log(y/(\log x)^{\theta})^2/\log\log x$ and is thus bounded for $\asymp \sqrt{\log\log x}$ values of y (which are each powers of 2). The contribution of values of y further away from $(\log x)^{\theta}$ decays rapidly and, in total, does not contribute as much as the values of y nearby, and thus, we deduce from (9.6) and (9.9) the upper bound

$$\sqrt{\log\log x}\,\frac{x(\log\log x)^{2^{2/3}-1}}{(\log x)^{1-2^{-1/3}}\sqrt{\log\log x}},$$

which is Corollary 2.

References

[1] N. C. ANKENY and S. CHOWLA, *The relation between the class number and the distribution of primes*, Proc. Amer. Math. Soc. **1** (1950), 775–776. MR 0038376

[2] P. BERNAYS, *Über die Darstellung von positiven, ganzen Zahlen durch die primitiven, binären quadratischen Formen einer nicht quadratischen Diskriminante*, Ph.D. dissertation, Georg-August-Universität, Göttingen, Germany, 1912.

[3] V. BLOMER, *Binary quadratic forms with large discriminants and sums of two squareful numbers*, J. Reine Angew. Math. **569** (2004), 213–234; *II*, J. London Math. Soc. (2) **71** (2005), 69–84. MR 2057239; MR 2108246

[4] ———, *On cusp forms associated with binary theta series*, Arch. Math. (Basel) **82** (2004), 140–146. MR 2047667

[5] J. BRÜDERN, *Einführung in die analytische Zahlentheorie*, Springer, Berlin, 1995.

[6] D. A. COX, *Primes of the form $x^2 + ny^2$: Fermat, Class Field Theory and Complex Multiplication*, Wiley-Intersci. Publ., Wiley, New York, 1989. MR 1028322

[7] P. ERDÖS and A. SÁRKÖZY, *On the number of prime factors of integers*, Acta Sci. Math. (Szeged) **42** (1980), 237–246. MR 0603312

[8] E. FOGELS, *On the zeros of Hecke's L-functions, I*, Acta Arith. **7** (1961/1962), 87–106. MR 0136585

[9] A. HILDEBRAND, *On the number of positive integers $\leq x$ and free of prime factors $> y$*, J. Number Theory **22** (1986), 289–307. MR 0831874

[10] H. IWANIEC, *The generalized Hardy-Littlewood's problem involving a quadratic polynomial with coprime discriminants*, Acta Arith. **27** (1975), 421–446. MR 0366847

[11] R. MURTY and R. OSBURN, *Representations of integers by certain positive definite binary quadratic forms*, preprint, arXiv:math.NT/0412237

[12] G. PALL, *The structure of the number of representations function in a positive binary quadratic form*, Math. Z. **36** (1933), 321 – 343. MR 1545349

Blomer
Department of Mathematics, Bahen Center, 40 St. George Street, Toronto, University of Toronto, M5S 2E4, Ontario Canada; vblomer@math.toronto.edu

Granville
Départment de Mathématiques et Statistique, Université de Montréal, CP 6128 succ. Centre-Ville, Montréal, Québec, H3C 3J7, Canada; andrew@dms.umontreal.ca