PRIMES REPRESENTED BY POSITIVE DEFINITE BINARY QUADRATIC FORMS

ASIF ZAMAN

Abstract. Let $f$ be a primitive positive definite integral binary quadratic form of discriminant $-D$ and let $\pi_f(x)$ be the number of primes up to $x$ which are represented by $f$. We prove several types of upper bounds for $\pi_f(x)$ within a constant factor of its asymptotic size: unconditional, conditional on the Generalized Riemann Hypothesis (GRH), and for almost all discriminants. The key feature of these estimates is that they hold whenever $x$ exceeds a small power of $D$ and, in some cases, this range of $x$ is essentially best possible. In particular, if $f$ is reduced then this optimal range of $x$ is achieved for almost all discriminants or by assuming GRH. We also exhibit an upper bound for the number of primes represented by $f$ in a short interval and a lower bound for the number of small integers represented by $f$ which have few prime factors.

CONTENTS

1. Introduction 1
2. Notation and conventions 7
3. Elementary estimates 7
4. Uniform bounds for quadratic characters 11
5. Average bounds for quadratic characters 12
6. Congruence sum decomposition 17
7. Local densities 20
8. Application of Selberg’s sieve 22
9. Representation of primes 25
10. Representation of small integers with few prime factors 27
References 27

1. Introduction

1.1. Historical review. The distribution of primes represented by positive definite integral binary quadratic forms is a classical topic within number theory and has been intensely studied over centuries by many renowned mathematicians, including Fermat, Euler, Lagrange, Legendre, Gauss, Dirichlet, and Weber. A beautiful exposition on the subject and its history can be found in [Cox13].

We shall refer to positive definite integral binary quadratic forms as simply ‘forms’. Let

$$f(u, v) = au^2 + buv + cv^2$$
be a form of discriminant $-D = b^2 - 4ac$. An integer $n$ is said to be represented by $f$ if there exists $(u, v) \in \mathbb{Z}^2$ such that $n = f(u, v)$. A form is primitive if $a, b,$ and $c$ are relatively prime. The group $SL_2(\mathbb{Z})$ naturally acts on the set of primitive forms with discriminant $-D$ and two forms are said to be properly equivalent if they belong to the same $SL_2(\mathbb{Z})$ orbit. A primitive form is reduced if $|b| \leq a \leq c$ and $b \geq 0$ if either $|b| = a$ or $a = c$. Every primitive form is properly equivalent to a unique reduced form. Most amazingly, the set of primitive forms with discriminant $-D$ modulo proper equivalence can be given a composition law that makes it a finite abelian group $\text{Cl}(-D)$. This is the class group of $-D$ and its size $h(-D)$ is the class number. We refer to each equivalence class of the class group as a form class.

Now, assume $f$ is primitive. The central object of study is the number of primes represented by $f$ up to $x$, denoted

$$\pi_f(x) = |\{ p \leq x : p = f(u, v) \text{ for some } (u, v) \in \mathbb{Z}^2 \}|$$

for $x \geq 2$. From deep connections established by class field theory, the Chebotarev density theorem \cite{Tsc26} \cite{LO77} implies that primes are equidistributed amongst all form classes. Namely,

$$\pi_f(x) \sim \frac{\delta_f x}{h(-D) \log x}$$

as $x \to \infty$, where

$$\delta_f = \begin{cases} 1 & \text{if } f(u, v) \text{ is properly equivalent to its opposite } f(u, -v), \\ \frac{1}{2} & \text{otherwise}. \end{cases}$$

Unfortunately, the asymptotic (1.1) derived from \cite{LO77} requires $x$ to be exponentially larger than $D$ even without the existence of a Siegel zero. This is unsuitable for many applications. Assuming the Generalized Riemann Hypothesis (GRH), one can do much better. It follows from the same work of Lagarias and Odlyzko \cite{LO77} that, assuming GRH,

$$\pi_f(x) = \frac{\delta_f \text{Li}(x)}{h(-D)} + O(x^{1/2} \log(Dx))$$

for $x \geq 2$. Here $\text{Li}(x) = \int_2^x \frac{1}{\log t} \, dt \sim \frac{x}{\log x}$. Equation (1.3) is nontrivial for $x \geq D^{1+\varepsilon}$ and any $\varepsilon > 0$. This GRH range for $x$ is essentially optimal for certain but not necessarily all forms. Amongst other results, we will partially address this defect in this paper. For further discussion of range optimality and improvements (both conditional and unconditional), see Sections 1.2 and 1.3.

Now, to remedy our lack of knowledge about the asymptotics of $\pi_f(x)$ in the unconditional case, one can settle for upper and lower bounds of the shape

$$x \frac{1}{h(-D) \log x} \ll \pi_f(x) \ll \frac{x}{h(-D) \log x}$$

in the hopes of improving the valid range of $x$. For uniform lower bounds, we refer the reader to \cite{Fog62} \cite{Wei83} \cite{KM02} and, most recently, \cite{TZ17b} wherein it was shown that there exists a prime $p$ represented by $f$ of size at most $O(D^{700})$.

The focus of this paper, however, is on upper bounds for $\pi_f(x)$. Assuming GRH, there has been no improvement beyond (1.3) itself. Unconditionally, a general result of Lagarias, Montgomery and Odlyzko \cite{LMO79} Theorem 1.4 on the Chebotarev density theorem yields some progress but the range of $x$ is still worse than exponential in $D$. Recently, a theorem...
of Thorner and Zaman [TZ17a] improves upon the aforementioned Chebotarev result and consequently implies that

\[ \pi_f(x) < \frac{2\delta_f \text{Li}(x)}{h(-D)} \quad \text{for } x \geq D^{700} \]

and \( D \) sufficiently large. Short of excluding a Siegel zero, the constant 2 is best possible and, moreover, the range of \( x \) is polynomial in \( D \) similar to (1.3). On the other hand, the quality of exponent 700 in (1.5) leaves much to be desired when compared to the GRH exponent of \( 1 + \varepsilon \). Both [LMO79] [TZ17a] carefully study the zeros of Hecke \( L \)-functions to prove their respective results whereas a more recent paper of Debaene [Deb16] uses a lattice point counting argument and Selberg’s sieve to subsequently establish another such Chebotarev-type theorem. His result implies a weaker variant of inequality (1.5) but for a greatly improved range of \( x \geq D^{9/2+\varepsilon} \). Broadly speaking, we will specialize Debaene’s strategy to positive definite binary quadratic forms.

An alternate formulation that expands the valid range of upper (and lower) bounds for \( \pi_f(x) \) can be obtained by averaging over discriminants or forms, in analogy with the famous theorems of Bombieri–Vinogradov and Barban–Davenport–Halberstam on primes in arithmetic progressions. As part of his Ph.D. thesis, Ditchen [Dit13a, Dit13b] achieved such elegant statements which we present in rough terms for simplicity’s sake. Namely, for 100\% of fundamental discriminants \( -D \not\equiv 0 \pmod{8} \) and all of their forms \( f \), he proved

\[ \pi_f(x) = \frac{\delta_f x}{h(-D) \log x} + O_x \left( \frac{x}{h(-D) (\log x)^2} \right) \]

provided \( x \geq D^{20/3+\varepsilon} \). He also showed for 100\% of fundamental discriminants \( -D \not\equiv 0 \pmod{8} \) and 100\% of their forms, equation (1.6) holds for \( x \geq D^{3+\varepsilon} \).

Finally, before we discuss an optimal range for upper bounds of \( \pi_f(x) \) and the details of our results, we give a somewhat imprecise flavour of what we have shown. One should compare with (1.3), (1.5), (1.6), and their associated papers [LO77] [TZ17a] [Deb16] [Dit13a].

**Corollary.** Let \( f(u,v) = au^2 + buv + cv^2 \) be a reduced positive definite integral binary quadratic form of discriminant \( -D \). If GRH holds and \( \varepsilon > 0 \) then

\[ \pi_f(x) \ll \frac{x}{h(-D) \log x} \quad \text{for } x \geq (D/a)^{1+\varepsilon}. \]

This estimate holds unconditionally for 100\% of discriminants \( -D \). Unconditionally and uniformly over all discriminants \( -D \), the same upper bound for \( \pi_f(x) \) holds for \( x \geq (D^2/a)^{1+\varepsilon} \).

**1.2. Optimal range.** What is the minimal size of \( x \) relative to \( D \) for which one can reasonably expect (1.4) to hold for all forms \( f \)? We have seen GRH implies \( x \geq D^{1+\varepsilon} \) is valid and, in one sense, this range is best possible. For example, the form

\[ f(u,v) = u^2 + \frac{D}{4} v^2 \]

with \( D \equiv 0 \pmod{4} \) does not represent any prime \( < D/4 \). The reason is simple: the coefficients are simply too large. On the other hand, what about forms

\[ f(u,v) = au^2 + buv + cv^2 \]

for which all of the coefficients are small compared to \( D \), say \( a,b,c \ll \sqrt{D} \)? The GRH range \( x \geq D^{1+\varepsilon} \) is insensitive to the size of the form’s coefficients. Since primitive forms are
properly equivalent to a reduced form and all forms in the same form class represent the same primes, it is enough to consider a reduced form $f$ so the coefficients necessarily satisfy $|b| \leq a \leq c$. Consider the sum

$$\sum_{n \leq x} \{|(u, v) \in \mathbb{Z}^2 : n = f(u, v)|\}.$$ 

If $x < c$ then, as $f$ is reduced, the only terms $n$ contributing to the above sum are of the form $n = f(u, 0) = au^2$. The sum is therefore equal to $2\sqrt{x/a} + O(1)$ and at most one prime value $n$ contributes to its size. On the other hand, if $x \geq c$ then the above sum is of size $O(x/\sqrt{D})$. One might therefore reasonably suspect that primes will appear with natural frequency for $x \geq D/a^{1+\varepsilon}$ and any $\varepsilon > 0$. As $c \asymp D/a$, it is conceivable for reduced forms $f$ to satisfy (1.3) in (what we will refer to as) the optimal range $x \geq (D/a)^{1+\varepsilon}$ for any $\varepsilon > 0$.

Can we expect to obtain any kind of unconditional or GRH-conditional bounds for $\pi_f(x)$ in the optimal range? For the sake of comparison, we turn to primes in arithmetic progressions since estimates like (1.5) are relatives of the classical Brun–Titchmarsh inequality. A version due to Montgomery and Vaughan [MV73] states for $(a, q) = 1$ and $x > q$ that

$$\pi(x; q, a) < \frac{2x}{\varphi(q) \log(x/q)}.$$ 

Here $\pi(x; q, a)$ represents the number of primes up to $x$ congruent to $a \pmod{q}$. The range $x > q$ is clearly best possible which inspires the possibility of success for a similar approach to primes represented binary quadratic forms. Our goal is to give upper bounds for $\pi_f(x)$ as close to the optimal range as possible.

1.3. Results. Our first result is a uniform upper bound for $\pi_f(x)$, both unconditional and conditional on GRH. For a discriminant $-D$, let $\chi_{-D}(\cdot) = \left(\frac{-D}{\cdot}\right)$ denote the corresponding Kronecker symbol which is a quadratic Dirichlet character.

**Theorem 1.1.** Let $f(u, v) = au^2 + buv + cv^2$ be a reduced positive definite integral binary quadratic form with discriminant $-D$ and let $\varepsilon > 0$ be arbitrary. Set

$$\phi = \phi(\chi_{-D}) := \begin{cases} 0 & \text{assuming } L(s, \chi_{-D}) \text{ satisfies GRH}, \\ \frac{1}{4} & \text{unconditionally.} \end{cases}$$

If $x \geq (D^{1+4\phi}/a)^{1+\varepsilon}$ then

$$\pi_f(x) < \frac{4}{1 - \theta} \cdot \frac{\delta_f x}{h(-D) \log x} \left\{1 + O\left(\frac{\log \log x}{\log x}\right)\right\},$$

where

$$\theta = \theta_\phi = (1 + 2\phi + \frac{\varepsilon}{2}) \frac{\log D}{\log x} - \frac{\log a}{\log x}.$$ 

**Remarks.**

(i) If $\phi = 0$ then $0 < \theta < 1$. Similarly, if $\phi = \frac{1}{4}$ then $0 < \theta < \frac{3}{4}$. The constant $\phi$ is associated with bounds for $L(s, \chi_{-D})$ in the critical strip. In particular, $\phi = \frac{1}{2}$ corresponds to the usual convexity estimate.

1Of course, one could potentially refine the factor of $\varepsilon$ but that is not our goal.
(ii) Unconditionally, one should compare this estimate with (1.5) and the related works [TZ17a, Deb16]. Theorem 1.1 achieves the upper bound in (1.4) with the range \( x \geq (D^2/a)^{1+\varepsilon} \) which improves over the prior range \( x \geq D^{9/2+\varepsilon} \) implied by [Deb16]. In fact, when \( a \gg D^{1/2} \), Theorem 1.1’s range becomes \( x \gg D^{3/2+\varepsilon} \) which is fairly close to the classical GRH range \( x \geq D^{1+\varepsilon} \). Of course, inequality (1.9) has a weaker implied constant \( 4^{1-\theta} \) instead of 2 as in (1.5).

(iii) Assuming GRH, we obtain the desired range \( x \geq (D/a)^{1+\varepsilon} \) discussed in Section 1.2. Note we only assume GRH for the quadratic Dirichlet \( L \)-function \( L(s, \chi_{-D}) \) whereas (1.3) assumes GRH for the collection of Hecke \( L \)-functions associated to the corresponding ring class field.

Since every primitive form is properly equivalent to a reduced form, we may ignore the dependence on the coefficient \( a \) in Theorem 1.1 to obtain the following simplified result.

**Corollary 1.2.** Let \( f \) be a primitive positive definite integral binary quadratic form with discriminant \( -D \) and let \( \varepsilon > 0 \) be arbitrary. If \( x \geq D^{2+\varepsilon} \) then

\[
\pi_f(x) < 8 \cdot \frac{\delta_f x}{h(-D) \log x} \left\{ 1 + O_{\varepsilon} \left( \frac{\log \log x}{\log x} \right) \right\}.
\]

We will actually prove a more general version of Theorem 1.1 that allows us to estimate the number of primes represented by \( f \) in a short interval.

**Theorem 1.3.** Let \( f(u, v) = au^2 + buv + cv^2 \) be a reduced positive definite integral binary quadratic form with discriminant \( -D \) and let \( \varepsilon > 0 \) be arbitrary. Let \( \phi \) be defined by (1.8). If \( \left( \frac{D^{1+4\phi}}{a} \right)^{1/2+\varepsilon} x^{1/2+\varepsilon} \leq y \leq x \) then

\[
\pi_f(x) - \pi_f(x - y) < \frac{2}{1 - \theta'} \cdot \frac{\delta_f y}{h(-D) \log y} \left\{ 1 + O_{\varepsilon} \left( \frac{\log \log y}{\log y} \right) \right\},
\]

where

\[
\theta' = \theta'_{\phi} = \frac{\log x}{2 \log y} + \left( \frac{1}{2} + \phi + \varepsilon \right) \frac{\log D}{\log y} - \frac{\log a}{2 \log y}.
\]

**Remarks.**

(i) Assuming GRH, (1.3) implies that

\[
\pi_f(x) - \pi_f(x - y) \ll \frac{\delta_f y}{h(-D) \log y}
\]

for \( (Dx)^{1/2+\varepsilon} \leq y \leq x \). Theorem 1.3 yields an unconditional upper bound of comparable strength and, depending on the size of the coefficients of \( f \), implies a GRH upper bound for slightly shorter intervals and smaller values of \( x \).

(ii) If \( f \) is only assumed to be primitive then the same statement holds by setting \( a = 1 \) in the condition on \( y \) and the value of \( \theta' \).

(iii) Theorem 1.1 follows by setting \( y = x \).

For any given discriminant \( -D \), the unconditional versions (\( \phi = \frac{1}{4} \)) of Theorems 1.1 and 1.3 fall slightly shy of their GRH counterparts (\( \phi = 0 \)). Averaging over all discriminants, we show that the GRH quality estimates hold almost always. For \( Q \geq 3 \), define

\[
\mathcal{D}(Q) := \{ \text{discriminants } -D \text{ with } 3 \leq D \leq Q \}.
\]
Here and throughout, a discriminant $-D$ is that of a positive definite integral binary quadratic form. Thus, $-D$ is a negative integer $\equiv 0$ or $1$ (mod $4$).

**Theorem 1.4.** Let $Q \geq 3$ and $0 < \varepsilon < \frac{1}{2}$. For all except at most $O_\varepsilon(Q^{1-\varepsilon})$ discriminants $-D \in \mathcal{D}(Q)$, the statements in Theorems 1.1 and 1.3 hold unconditionally with $\phi = 0$.

**Remark.** When considering upper bounds for $\pi_f(x)$, this improves over (1.6) in several aspects. First, the desired range discussed in Section 1.2 is achieved on average. Second, we did not utilize any averaging over forms, only their discriminants. Finally, there are no restrictions on the family of discriminants; they need not be fundamental or satisfy any special congruence condition.

The underlying strategy to establish Theorems 1.1, 1.3 and 1.4 rests on a natural two-step plan. First, estimate the congruence sums

$$|A_\ell| = \sum_{\ell \mid n \leq x} r_f(n),$$

where $\ell$ is squarefree and $r_f(n) = \{|(u,v) \in \mathbb{Z}^2 : n = f(u,v)\}$ is the number of representations of the integer $n$ by the form $f$. Second, apply Selberg’s upper bound sieve. Our application of the sieve is fairly routine but calculating the congruence sums with sufficient precision poses some difficulties.

Inspired by a beautiful paper of Granville and Blomer [BG06] wherein they carefully study the moments of $r_f(n)$, we determine the congruence sums via geometry of numbers methods. However, for their purposes, only a simple well-known estimate [BG06, Lemma 3.1] for the first moment $\sum_{n \leq x} r_f(n)$ was necessary. We execute a more refined analysis of the first moment and related quantities $|B_\ell(m)|$ (see (6.4) for a definition) using standard arguments with the sawtooth function. Afterwards, the main technical hurdle is to carefully decompose the congruence sums $|A_\ell|$ into a relatively small number of disjoint quantities $|B_\ell(m)|$. This allows us to apply our existing estimates for $|B_\ell(m)|$ (see Lemma 6.2) while simultaneously controlling the compounding error terms. We achieve this in Proposition 7.1; see Sections 6 and 7 for details on this argument. When finalizing the proofs of Theorems 1.1, 1.3 and 1.4 the various valid ranges for $x$ are determined by relying on character sum estimates like Burgess’s bound, Heath-Brown’s mean value theorem for quadratic characters [HB95], and Jutila’s zero density estimate [Jut75].

Since we have calculated the congruence sums (1.12) in Proposition 7.1 we thought it may be of independent interest to examine its performance in conjunction with a lower bound sieve. By a direct application of the beta sieve, we show:

**Theorem 1.5.** Let $f(u,v) = au^2 + buv + cv^2$ be a reduced positive definite integral binary quadratic form with discriminant $-D$. For every integer $k \geq 10$, the number of integers represented by $f$ with at most $k$ prime factors is

$$\gg \frac{x}{\sqrt{D}(\log x)^2} \quad \text{for } x \geq \left(\frac{D}{a}\right)^{1+\frac{\varepsilon}{\log D}}.$$

**Remarks.**

(i) One can formulate this statement in a slightly weaker alternative fashion: for every $\varepsilon > 0$ and $x \geq (D/a)^{1+\varepsilon}$, the number of integers represented by $f$ with at most $O_\varepsilon(1)$ prime factors is $\gg \frac{x}{\sqrt{D}(\log x)^2}$. 

6
(ii) It is unsurprising that $f$ represents many integers with few prime factors as this follows from standard techniques in sieve theory. The form $f$ even represents primes of size $O(D^{700})$ by [TZ17b]. However, the key feature of Theorem 1.5 is that the size of these integers with few prime factors is very small (in a best possible sense per Section 1.2).

Acknowledgements. I would like to thank John Friedlander and Jesse Thorner for their encouragement and many helpful comments on an earlier version of this paper. I am also grateful to Kannan Soundararajan for several insightful conversations and for initially motivating me to pursue this approach. Part of this work was completed with the support of an NSERC Postdoctoral Fellowship.

2. Notation and conventions

For each of the asymptotic inequalities $F \ll G$, $F = O(G)$, or $G \gg F$, we mean there exists a constant $C > 0$ such that $|F| \leq CG$. We henceforth adhere to the convention that all implied constants in all asymptotic inequalities are absolute with respect to all parameters and are effectively computable. If an implied constant depends on a parameter, such as $\varepsilon$, then we use $\ll_{\varepsilon}$ and $O_{\varepsilon}$ to denote that the implied constant depends at most on $\varepsilon$.

Throughout the paper,

- A form refers to a positive definite binary integral quadratic form.
- $-D$ is the discriminant of a positive definite integral binary quadratic form, so $-D$ is any negative integer $\equiv 0$ or 1 (mod 4). It is not necessarily fundamental.
- $f(u,v) = au^2 + buv + cv^2 \in \mathbb{Z}[u,v]$ is a positive definite integral binary quadratic form of discriminant $-D$. It is not necessarily primitive or reduced.
- $\chi_{-D}(\cdot) = (\frac{-D}{\cdot})$ is the Kronecker symbol attached to $-D$.
- $\mathcal{D}(Q)$ is the set of discriminants $-D$ with $3 \leq D \leq Q$.
- $\Delta$ is a (negative) fundamental discriminant of a form.
- $\tau_k$ is the $k$-divisor function and $\tau = \tau_2$ is the divisor function.
- $\varphi$ is the Euler totient function.
- $[s,t]$ is the least common multiple of integers $s$ and $t$.
- $(s,t)$ is the greatest common divisor of integers $s$ and $t$. However, we may abuse notation and sometimes refer to a lattice point $(u,v) \in \mathbb{Z}^2$ but this will be made clear from context (e.g. with the set membership symbol).

3. Elementary estimates

First, we establish a standard result employing the Euler-Maclaurin summation formula [IK04, Lemma 4.1], which will later allow us to counts lattice points inside an ellipse.

**Lemma 3.1.** For $W \geq 1$,

$$\sum_{1 \leq w \leq W} \sqrt{W^2 - w^2} = \frac{\pi W^2}{4} - \frac{W}{2} + O(\sqrt{W}).$$

**Proof.** Set $G(w) = \sqrt{W^2 - w^2}$. By partial summation, observe that

$$\sum_{1 \leq w \leq W} G(w) = -\int_0^W tG'(t)dt + \int_0^W \psi(t)G'(t)dt - \frac{1}{2}G(0),$$

where $\psi(t)$ is a smooth function that vanishes at the endpoints. Then

$$\sum_{1 \leq w \leq W} \sqrt{W^2 - w^2} = \frac{\pi W^2}{4} - \frac{W}{2} + O(\sqrt{W}).$$
where \( \psi(t) = t - \lfloor t \rfloor - 1/2 \) is the sawtooth function. For the first integral, notice

\[
(3.2) \quad - \int_0^W tG'(t)dt = \int_0^W \frac{t^2}{\sqrt{W^2 - t^2}}dt = \frac{\pi W^2}{4}.
\]

For the second integral, we use the Fourier expansion (see e.g. [IK04, Equation (4.18)])

\[
\psi(x) = -\sum_{1 \leq n \leq N} (\pi n)^{-1} \sin(2\pi nx) + O((1 + ||x||N)^{-1}),
\]

where ||x|| is the distance of x to the nearest integer. It follows that

\[
(3.3) \quad \int_0^W \psi(t)G'(t)dt = 2 \sum_{1 \leq n \leq N} \int_0^W G(t) \cos(2\pi nt)dt + O\left(\int_0^W |G'(t)| \frac{dt}{1 + ||t||N}\right)
\]

after integrating by parts. Using a computer algebra package or table of integrals,

\[
(3.4) \quad \int_0^W G(t) \cos(2\pi nt)dt = \int_0^W \sqrt{W^2 - t^2} \cos(2\pi nt)dt = W \cdot \frac{J_1(2\pi nW)}{4n},
\]

where \( J_\nu(z) \) is the Bessel function of the first kind. Recall that, for \( z > 1 + |\nu|^2 \),

\[
J_\nu(z) = \left(\frac{1}{2}z\right)^\nu \sum_{m=0}^\infty \frac{(-1)^m(z^2/4)^m}{m!\Gamma(m + \nu + 1)} = \left(\frac{2}{\pi z}\right)^{1/2} \left(\cos(z - \frac{3}{2}\nu\pi - \frac{1}{4}\pi) + O\left(\frac{1 + |\nu|^2}{z}\right)\right).
\]

When summing (3.4) over \( 1 \leq n \leq N \), it follows that

\[
\begin{align*}
2 \sum_{1 \leq n \leq N} \int_0^W G(t) \cos(2\pi nt)dt &= \frac{W}{4} \cdot 2 \sum_{1 \leq n \leq N} \frac{J_1(2\pi nW)}{n} \\
&= \frac{\sqrt{W}}{2\pi} \sum_{1 \leq n \leq N} \left(\frac{\cos(2\pi nW - 3\pi/4)}{n^{3/2}} + O\left(\frac{1}{n^{5/2}W}\right)\right) \\
&\ll \sqrt{W}
\end{align*}
\]

uniformly over \( N \). Taking \( N \to \infty \), we conclude from (3.3) and (3.5) that

\[
(3.6) \quad \int_0^W \psi(t)G'(t)dt \ll \sqrt{W}.
\]

Combining (3.1), (3.2), and the above yields the result. \( \square \)

Next, in Lemma 3.2, we calculate some weighted average values of the Dirichlet convolution

\[
(1 * \chi)(n) = \sum_{d|n} \chi(d)
\]

for any quadratic Dirichlet character \( \chi \). Better estimates and certainly simpler proofs are available via Mellin inversion, but it is useful for us to explicitly express the error terms using the character sum quantity

\[
S_\chi(t) := \sum_{n \leq t} \chi(n).
\]
This feature will give us flexibility and streamline the proofs of Theorems 1.3 and 1.4 each of which use different bounds for character sums. One utilizes uniform bounds (conditional and unconditional) whereas the other applies average bounds.

**Lemma 3.2.** Let $\chi \pmod{D}$ be a quadratic Dirichlet character. For $x \geq 3$,

$$
\sum_{n \leq x} (1 \ast \chi)(n) \left(1 - \frac{n}{x}\right) = \frac{x}{2} L(1, \chi) + O(\mathcal{E}_0(x; \chi)),
$$

where

$$
\mathcal{E}_0(x; \chi) := \min_{1 \leq y \leq x} \left(\frac{y^2}{x} + |S_\chi(y)| + x \int_y^\infty \frac{|S_\chi(t)|}{t^2} dt\right)
$$

and $S_\chi(t) = \sum_{n \leq t} \chi(n)$. Moreover,

$$
\sum_{n \leq x} \frac{(1 \ast \chi)(n)}{n} \left(1 - \frac{n}{x}\right) = L(1, \chi)(\log x + \gamma - 1) + L'(1, \chi) + O(\mathcal{E}_1(x; \chi)),
$$

where

$$
\mathcal{E}_1(x; \chi) := \min_{1 \leq y \leq x} \left(\frac{y}{x} + \log x \int_y^\infty \frac{|S_\chi(t)| \log t}{t^2} dt\right).
$$

**Remark.** Bounds for $\mathcal{E}_0(x; \chi)$ and $\mathcal{E}_1(x; \chi)$ can be found in Sections 4 and 5.

**Proof.** The proofs of these facts are standard but we include the details for the sake of completeness. Equation (3.7) is a more precise version of [MV07, Section 4.3.1, Exercise 3] which originated from work of Mertens. Before we proceed, recall by partial summation that

$$
\sum_{d \leq y} \chi(d) \frac{A(x/d)}{d} = L(1, \chi) - \int_y^\infty \frac{S_\chi(t)}{t^2} dt,
$$

and

$$
\sum_{d \leq y} \chi(d) \log d \frac{A(x/d)}{d} = -L'(1, \chi) - \int_y^\infty \frac{S_\chi(t)(\log t - 1)}{t^2} dt.
$$

We will use these estimates in what follows. Let $A(u) = \sum_{n \leq u} (1 - \frac{n}{u})$. One can verify that

$$
A(u) = \frac{u}{2} - \frac{1}{2} \int_0^u \{t\} dt = \frac{u}{2} - \frac{1}{2} + O(u^{-1}).
$$

From the above integral formula for $A(u)$, it is straightforward to check that $A(u)$ is continuous and if $u > 1$ is not an integer then

$$
A'(u) = \frac{[u]([u] + 1)}{2u^2} \ll 1.
$$

In particular, $A(u)$ is increasing and absolutely continuous. Now, to calculate the sum in (3.7), we use Dirichlet’s hyperbola method with a parameter $1 \leq y \leq x$. Namely,

$$
\Sigma := \sum_{n \leq x} (1 \ast \chi)(n) \left(1 - \frac{n}{x}\right) = \sum_{d \leq y} \chi(d)A(x/d) + \sum_{y < d \leq x} \chi(d)A(x/d)
$$

$$
= \Sigma_1 + \Sigma_2.
$$
say. For $\Sigma_1$, we see by (3.10) that

$$\Sigma_1 = \frac{x}{2} \sum_{d \leq y} \frac{\chi(d)}{d} - \frac{1}{2} \sum_{d \leq y} \chi(d) + O(y^2/x).$$

For $\Sigma_2$, since $A$ is an absolutely continuous, decreasing, non-negative function and $A(1) = 0$, it follows by partial summation that

$$\Sigma_2 = \int_y^x A(x/t) dS_X(t) = A(1) S_X(x) + \int_y^x S_X(t) dA(x/t) = -\int_y^x S_X(t) A'(x/t) xt^{-2} dt.$$

Thus, by (3.11),

$$|\Sigma_2| \ll x \left( \int_y^x |S_X(t)| dt \right).$$

Combining the above estimate, (3.13), and (3.9) into (3.12) yields (3.7).

We prove (3.8) similarly. For $u \geq 1$, define

$$B(u) := \sum_{n \leq u} \frac{1}{n} \left( 1 - \frac{n}{u} \right) = \log u + \gamma - 1 + O(u^{-1}).$$

One can verify that $B(u)$ is a non-negative, increasing, and absolutely continuous function of $u$. Also, if $u > 1$ is not an integer then

$$B'(u) \ll \frac{1}{u}.$$

For some parameter $1 \leq y \leq x$,

$$\Sigma' := \sum_{n \leq x} \frac{(1 \ast \chi)(n)}{n} \left( 1 - \frac{n}{x} \right) = \sum_{d \leq y} \frac{\chi(d)}{d} B(x/d) + \sum_{y < d \leq x} \frac{\chi(d)}{d} B(x/d)$$

$$= \Sigma_1' + \Sigma_2',$$

say. To calculate $\Sigma_1'$, we apply (3.14) and deduce that

$$\Sigma_1' = (\log x + \gamma - 1) \sum_{d \leq y} \frac{\chi(d)}{d} - \sum_{d \leq y} \frac{\chi(d) \log d}{d} + O(y/x).$$

For $\Sigma_2'$, set $\tilde{S}_X(u) := \sum_{y < d \leq u} \frac{\chi(d)}{d}$. Since $B$ is an absolutely continuous, non-negative, increasing function and $B(1) = 0$, we similarly conclude that

$$\Sigma_2' = \int_y^x B(x/t) d\tilde{S}_X(t) = -x \int_y^x \tilde{S}_X(t) B'(x/t) xt^{-2} dt.$$

From (3.15), it follows that

$$|\Sigma_2'| \ll \int_y^x \frac{|\tilde{S}_X(t)|}{t} dt.$$

Substituting the identity

$$\tilde{S}_X(t) = \int_y^t \frac{1}{u} dS_X(u) = \frac{S_X(t)}{t} + \int_y^t \frac{S_X(u)}{u^2} du$$

10
into the previous estimate, we have that
\[ \Sigma'_2 \ll \int_y^x \frac{|S'\chi(t)|}{t^2} \, dt + \int_y^x \frac{1}{t} \int_y^t \frac{|S\chi(u)|}{u^2} \, du \, dt \]
\[ \ll \int_y^x \frac{|S\chi(t)|}{t^2} \, dt + \left( \int_y^x \frac{1}{t} \, dt \right) \left( \int_y^x \frac{|S\chi(u)|}{u^2} \, du \right) \]
\[ \ll \log x \int_y^x \frac{|S\chi(t)|}{t^2} \, dt. \]

Incorporating (3.17), (3.9), and the above into (3.16) establishes (3.8). □

4. Uniform bounds for quadratic characters

Here we collect known uniform bounds for character sums and values of the logarithmic derivatives of Dirichlet $L$-functions for quadratic characters. We apply the former to obtain estimates for the error terms arising in Lemma 3.2.

4.1. Character sums. We state the celebrated result of Burgess [Bur63] specialized to quadratic characters which was extended from cube-free moduli to all moduli by Heath-Brown [HB92, Lemma 2.4].

**Lemma 4.1.** Let \( \chi \pmod{D} \) be a quadratic Dirichlet character. For any \( \eta > 0, N \geq 1 \), and integer \( k \geq 3 \),
\[ \sum_{n \leq N} \chi(n) \ll_{\eta,k} D^{(k+1)/4k^2 + \eta N^{-1/k}}. \]

We also record the well known GRH-conditional estimate for character sums.

**Lemma 4.2.** Let \( \chi \pmod{D} \) be a non-principal Dirichlet character and suppose \( L(s, \chi) \) satisfies GRH. For any \( \eta > 0 \) and \( N \geq 1 \),
\[ \sum_{n \leq N} \chi(n) \ll_{\eta} D^\eta N^{1/2}. \]

4.2. Errors from Lemma [3.2]. The results of Section 4.1 allow us to obtain power-saving estimates for the error terms in Lemma 3.2 for small values of \( x \) relative to the conductor \( D \).

**Lemma 4.3.** Let \( \chi \pmod{D} \) be a quadratic Dirichlet character and let \( 0 < \epsilon < \frac{1}{20} \) be arbitrary. Let \( \mathcal{E}_0(x; \chi) \) and \( \mathcal{E}_1(x; \chi) \) be as in Lemma 3.2. If \( x \gg \epsilon D^{1/4+\epsilon} \) then
\[ \mathcal{E}_0(x; \chi) \ll_{\epsilon} x^{1-\epsilon^2/2} \quad \mathcal{E}_1(x; \chi) \ll_{\epsilon} x^{-\epsilon^2/2}. \]

If \( L(s, \chi) \) satisfies GRH and \( x \gg \epsilon D^\epsilon \) then
\[ \mathcal{E}_0(x; \chi) \ll_{\epsilon} x^{2/3+\epsilon} \quad \text{and} \quad \mathcal{E}_1(x; \chi) \ll_{\epsilon} x^{-1/3+\epsilon}. \]

**Proof.** First, we consider \( \mathcal{E}_0 = \mathcal{E}_0(x; \chi) \). Let \( 1 \leq y \leq x \) be a parameter yet to be chosen. From Lemma 4.1 and the definition of \( \mathcal{E}_0 \), we have that
\[ \mathcal{E}_0 \ll_{\eta,k} x^{-1} y^2 + D^{\frac{k+1}{12k^2} + \eta x^{-1/k}} y^{-1}. \]

Selecting
\[ y = D^{\frac{k+1}{12k^2} + \frac{3}{4} x^{-1/3}}, \quad \eta = \frac{1}{8k^2}, \quad k = \lceil 1/\epsilon \rceil \geq 20 \]
yields the estimate for $E_0$ in (1.1) since $x \gg_\varepsilon D^{1/4+\varepsilon}$. Now, assume GRH holds for $L(s, \chi)$ and $x \gg_\varepsilon D^\varepsilon$. Utilize Lemma 4.2 with $\eta = 3\varepsilon^2/2$ and select $y = D^{\varepsilon^2/2}x^{5/6}$. This implies that

$$E_0 \ll_\varepsilon \frac{y^2}{x} + \frac{D^{3\varepsilon^2/2}x^{3/2}}{y} \ll_\varepsilon D^\varepsilon x^{2/3} \ll_\varepsilon x^{2/3 + \varepsilon}$$

as desired. The arguments for $E_1 = E_1(x; \chi)$ are similar. Again, by Lemma 4.1

$$E_1 \ll_{k, \eta} \frac{y}{x} + \frac{D^{k+1}x^{-1/k}(\log x)^2}{y},$$

Selecting

$$y = D^{k+1}x^{1-1/2k}, \quad \eta = \frac{1}{8k^2}, \quad k = \lceil 1/\varepsilon \rceil \geq 20$$

yields the estimate for $E_1$ in (1.1) since $x \gg_\varepsilon D^{1/4+\varepsilon}$. If $L(s, \chi)$ satisfies GRH then we use Lemma 4.2 with $\eta = \varepsilon^2$ and select $y = D^\varepsilon x^{2/3}$. This implies that

$$E_1 \ll_\varepsilon \frac{y}{x} + D^{3\varepsilon^2/2}y^{-1/2}(\log x)^2 \ll_\varepsilon D^{\varepsilon^2/2}x^{-1/3}(\log x)^2 \ll_\varepsilon x^{-1/3 + \varepsilon}$$

as desired.

4.3. Logarithmic derivatives. We record two results related to the value of the logarithmic derivative of $L(s, \chi)$ at $s = 1$ for a quadratic Dirichlet character $\chi$.

**Lemma 4.4** (Heath-Brown). If $\chi (\mod D)$ is a quadratic character then, for $\eta > 0$,

$$-\frac{L'}{L}(1, \chi) \leq \left( \frac{1}{8} + \eta \right) \log D + O_\eta(1).$$

*Proof.* This follows from the proof of [HB92, Lemma 3.1] with some minor modifications to allow for any modulus $D$ (not just sufficiently large) and a slightly wider range of the quantity $\sigma$ therein, say $1 < \sigma < 1 + \varepsilon$. See [TZ17b, Proposition 2.6] for details. □

**Lemma 4.5.** If $\chi (\mod D)$ is a quadratic character and $L(s, \chi)$ satisfies GRH then for $\eta > 0$

$$-\frac{L'}{L}(1, \chi) \ll \log \log D \leq \eta \log D + O_\eta(1).$$

*Proof.* This is well known and can be deduced from the arguments in Lemma 5.5. □

5. Average bounds for quadratic characters

The purpose of this section is analogous to Section 4 except we focus on estimates averaging over a certain class of quadratic characters attached to discriminants. To be more specific, for $Q \geq 3$, recall

$$\mathcal{D}(Q) = \{ \text{discriminants } -D \text{ with } 3 \leq D \leq Q \}.$$ 

Here a discriminant is that of a primitive positive definite binary quadratic form. We emphasize that a discriminant $-D \in \mathcal{D}(Q)$ is not necessarily fundamental. The associated Kronecker symbol $\chi_{-D}(\cdot) = (-D)$ is itself a quadratic Dirichlet character. Note the character is primitive if and only if $-D$ is a fundamental discriminant. Our goal is to average certain quantities involving $\chi_{-D}$ over $-D \in \mathcal{D}(Q)$. 12
5.1. **Character sums.** We record a special case of Heath-Brown’s mean value theorem for primitive quadratic characters [HB95, Corollary 3].

**Lemma 5.1** (Heath-Brown). Let $N,Q \geq 1$ and let $a_1, \ldots, a_n$ be arbitrary complex numbers. Let $\mathcal{S}(Q)$ denote the set of all primitive quadratic characters of conductor at most $Q$. Then

$$
\sum_{\chi \in \mathcal{S}(Q)} \left| \sum_{n \leq N} a_n \chi(n) \right|^2 \ll \eta \left( (QN)^{1+\eta} + Q^\eta N^{2+\eta} \right) \max_{1 \leq n \leq N} |a_n|
$$

for any $\eta > 0$.

Using Lemma 5.1 we deduce an analogous mean value result for the quadratic characters attached to such discriminants.

**Lemma 5.2.** Let $N,Q \geq 1$ and let $\mathfrak{D}(Q)$ be defined by (1.11). Then

$$
\sum_{-D \in \mathfrak{D}(Q)} \left| \sum_{n \leq N} \chi_{-D}(n) \right|^2 \ll \eta \left( (QN)^{1+\eta} + Q^{1/2+\eta} N^{2+\eta} \right)
$$

for any $\eta > 0$.

**Proof.** If $-D \in \mathfrak{D}(Q)$ then $-D = \Delta k^2$ where $\Delta \in \mathbb{Z}$ is the discriminant of some imaginary quadratic field and $k \geq 1$ is some integer. Consequently, the Kronecker symbol $\chi_{-D}(\cdot) = \left( \frac{-D}{\cdot} \right)$ is induced by the primitive quadratic character $\chi_{\Delta}(\cdot) = \left( \frac{\Delta}{\cdot} \right)$. Moreover,

$$
\chi_{-D}(n) = \chi_{k^2}(n) \chi_{\Delta}(n).
$$

For the details on these facts, see [Cox13, §7]. Therefore, by Lemma 5.1

$$
\sum_{-D \in \mathfrak{D}(Q)} \left| \sum_{n \leq N} \chi_{-D}(n) \right|^2 \leq \sum_{1 \leq k \leq \sqrt{Q}} \sum_{1 \leq |\Delta| \leq Q/k^2} \left| \sum_{n \leq N} \chi_{k^2}(n) \chi_{\Delta}(n) \right|^2
$$

\[ \ll \eta \sum_{1 \leq k \leq \sqrt{Q}} \left( (QN)^{1+\eta} k^{-2-2n} + Q^\eta k^{-2\eta} N^{2+\eta} \right) \]

\[ \ll \eta \left( (QN)^{1+\eta} + Q^{1/2+\eta} N^{2+\eta} \right) \]

as desired. \qed

5.2. **Error terms from Lemma 3.2.** Again, the results in Section 5.1 leads to estimates for the error terms in Lemma 3.2.

**Lemma 5.3.** Let $X \geq 1, Q \geq 3$ and let $\mathfrak{D}(Q)$ be defined by (1.11). For $x \geq 1$ and any quadratic character $\chi$, define $E_0(x; \chi)$ and $E_1(x; \chi)$ as in Lemma 3.2. Then

$$
\sum_{-D \in \mathfrak{D}(Q)} \max_{X \leq x \leq 2X} E_0(x; \chi_{-D}) \ll \eta Q^{1+\eta} X^{3/4+\eta} + Q^{3/4+\eta} X^{1+\eta}
$$

(5.1)

$$
\sum_{-D \in \mathfrak{D}(Q)} \max_{X \leq x \leq 2X} E_1(x; \chi_{-D}) \ll \eta Q^{1+\eta} X^{-1/3+\eta} + Q^{3/4+\eta} X^{\eta}
$$

for $\eta > 0$. 13
Proof. For $X \leq x \leq 2X$, let $1 \leq y \leq X$ be an unspecified parameter, depending only on $X$. Set $y_0 := yQ^{1/2+\eta}$. By Polya-Vinogradov, $|S_\chi(t)| \ll Q^{1/2} \log Q$ for any character $\chi \pmod{q}$ with $q \leq Q$. Therefore, for such $\chi$,

$$E_0(x; \chi) \ll \frac{y^2}{x} + |S_\chi(y)| + x \int_y^\infty \frac{|S_\chi(t)|}{t^2} dt \ll \frac{y^2}{x} + |S_\chi(y)| + x \int_y^{y_0} \frac{|S_\chi(t)|}{t^2} dt + \frac{x}{y_0}.$$ 

As $y$ and $y_0$ depend only on $X$ and $Q$, it follows that

$$\max_{X \leq x \leq 2X} E_0(x; \chi) \ll \frac{y^2}{X} + |S_\chi(y)| + X \int_y^{y_0} \frac{|S_\chi(t)|}{t^2} dt + \frac{X}{y}.$$ 

Summing the above expression over $\chi = \chi_{-D}$ with $-D \in \mathcal{D}(Q)$, applying Cauchy-Schwarz, and invoking the previous lemma, we see that

$$\sum_{-D \in \mathcal{D}(Q)} \max_{X \leq x \leq 2X} E_0(x; \chi_{-D}) \ll \frac{Qy^2}{X} + Q^{1+\eta}y^{1/2+\eta} + Q^{3/4+\eta}y^{1+\eta} + \frac{XQ}{y}$$

$$+ XQ^{1/2} \int_y^{y_0} \frac{(Qt)^{1/2+\eta} + Q^{1/4+\eta/2}t^{1+\eta/2}}{t^2} dt \ll \frac{Qy^2}{X} + Q^{1+\eta}y^{1/2+\eta} + Q^{3/4+\eta}y^{1+\eta} + \frac{XQ}{y}$$

$$+ XQ^{1+\eta}y^{-1/2+\eta} + XQ^{3/4+\eta/2}y_0^{\eta/2} \ll \frac{Qy^2}{X} + \frac{XQ}{y} + XQ^{1+\eta}y^{-1/2+\eta} + X^{1+\eta}Q^{3/4+\eta}.$$ 

In the last step, we used the definition of $y_0$ and the fact that $y \leq X$. Selecting $y = X^{4/5}$ implies the desired result after rescaling $\eta$ if necessary. We follow the same procedure for the average of $E_1$. First, we deduce that

$$\max_{X \leq x \leq 2X} E_1(x; \chi_{-D}) \ll \frac{y}{X} + \log X \int_y^{y_0} |S_\chi(t)| \log t \frac{dt}{t^2} + \frac{\log(Qy)}{y}$$

Again, summing over $\chi = \chi_{-D}$ with $-D \in \mathcal{D}(Q)$, we similarly conclude that

$$\sum_{-D \in \mathcal{D}(Q)} \max_{X \leq x \leq 2X} E_1(x; \chi_{-D}) \ll \frac{Qy}{X} + \frac{Q\log(Qy)}{y} + Q^{1/2} \log X \int_y^{y_0} (Qt)^{1/2+\eta} + Q^{1/4+\eta/2}t^{1+\eta/2} dt$$

$$\ll \frac{Qy}{X} + \frac{Q\log(Qy)}{y} + (Q^{1+\eta}y^{-1/2+\eta} + Q^{3/4+\eta/2}y_0^{\eta/2}) \log X$$

$$\ll \frac{Qy}{X} + \frac{Q^{1+\eta}}{y^{1-\eta}} + X^\eta Q^{1+\eta}y^{-1/2+\eta} + X^\eta Q^{3/4+\eta}.$$ 

Selecting $y = X^{2/3}$ yields the desired result. \qed
Lemma 5.4. Let $0 < \varepsilon < 1/8$, $X \geq 1$ and $Q \geq 3$. Let $c = c(\varepsilon) > 0$ and $C = C(\varepsilon) \geq 1$ be arbitrary constants. For all except at most $O(\varepsilon Q^{1-\varepsilon/10})$ discriminants $-D \in \mathcal{D}(Q)$, \n
\begin{equation}
\mathcal{E}_0(x; -D) \leq x^{7/8+\varepsilon}, \quad \mathcal{E}_1(x; -D) \leq x^{-1/8+\varepsilon}, \n
\end{equation}

uniformly for $cD^\varepsilon \leq x \leq CD^{2+\varepsilon}$. Here $\mathcal{E}_0$ and $\mathcal{E}_1$ are defined as in Lemma 3.2.

Proof. Without loss, we need only consider discriminants $-D \in \mathcal{D}(Q)$ satisfying $D \geq Q^{1-\varepsilon}$ since the remainder are a collection of negligible size $O(Q^{1-\varepsilon})$. For $X \geq 1$, define

$$\mathcal{E}(X, Q, \varepsilon) := \{-D \in \mathcal{D}(Q) : \text{there exists } X \leq x \leq 2X \text{ violating (5.2) for } \chi_{-D}\}.$$ 

By Lemma 5.3 with $\eta = \varepsilon/8$, it follows that

$$\mathcal{E}(X, Q, \varepsilon) \ll_{\varepsilon} Q^{1+\varepsilon/8}X^{-11/40-7\varepsilon/8} + Q^{3/4+\varepsilon/8}X^{1/8-7\varepsilon/8}.$$ 

Dyadically summing this estimate over $X$ between $cQ^{\varepsilon(1-\varepsilon)}$ and $CQ^{2+\varepsilon}$, we see that the total number of discriminants $-D \in \mathcal{D}(Q)$ satisfying $D \geq Q^{1-\varepsilon}$ and violating (5.2) anywhere in the range $cD^\varepsilon \leq x \leq CD^{2+\varepsilon}$ is bounded by

$$\ll_{\varepsilon} \left(Q^{1+\varepsilon/8}X^{-11/40-7\varepsilon/8} + Q^{-2+\varepsilon/8}\right) \log Q \ll_{\varepsilon} Q^{1-\varepsilon/10} \log Q \ll_{\varepsilon} Q^{1-\varepsilon/10}$$ 

as $\varepsilon < 1/8$. \hfill $\square$

5.3. Logarithmic derivatives. For $Q \geq 3$, define \n
$$\mathcal{D}^*(Q) = \{\text{fundamental discriminants } \Delta \text{ with } 3 \leq |\Delta| \leq Q\}.$$ 

We show that, aside from a sparse set of fundamental discriminants in $\mathcal{D}^*(Q)$, the logarithmic derivative of $L(s, \chi_\Delta)$ at $s = 1$ satisfies a GRH-quality bound. The key inputs are the explicit formula and Jutila's zero density estimate for primitive quadratic characters.

Lemma 5.5. Let $Q \geq 3$ and $\varepsilon > 0$ be arbitrary. For all except at most $O(\varepsilon Q^{3/4+\varepsilon})$ fundamental discriminants $\Delta \in \mathcal{D}^*(Q)$, \n
\begin{equation}
-L'(1, \chi_\Delta) \ll_{\varepsilon} \log \log |\Delta|.
\end{equation}

Proof. We modify the arguments leading to [MJK13, Theorem 3]. Define $\mathcal{D}^*_\varepsilon(Q)$ to be the set of fundamental discriminants $\Delta \in \mathcal{D}^*(Q)$ such that $|\Delta| \leq Q$ and whose $L$-function $L(s, \chi_\Delta)$ is zero-free in the rectangle

\begin{equation}
\frac{1}{2} < \Re{s} < 1 \quad |\Im{s}| \leq |\Delta|^\varepsilon.
\end{equation}

First, we estimate $-\frac{L'}{L}(s, \chi_\Delta)$ for $\Delta \in \mathcal{D}^*_\varepsilon(Q)$. For simplicity, write $\chi = \chi_\Delta$. From the explicit formula in the form given by [IMS09, p. 261], one can verify that

\begin{equation}
-L'(1, \chi) = \frac{1}{y-1} \sum_{m<y} \left(\frac{y}{m} - 1\right) \Lambda(m) \chi(m) - \frac{1}{y-1} \sum_{\rho} \frac{y^\rho - 1}{\rho(1 - \rho)} + O\left(\frac{\log y}{y}\right)
\end{equation}

for $y \geq 2$, where the sum is taken over all non-trivial zeros $\rho = \beta + i\gamma$ of $L(s, \chi)$. From [IMS09, Equation (5.4.6)] and the prime number theorem, it follows for $T \geq 1$ that

$$-\frac{L'}{L}(1, \chi) = -\frac{1}{y-1} \sum_{\rho \geq y} \frac{y^\rho - 1}{\rho(1 - \rho)} + O\left(\log y + \log\left(\frac{|\Delta| T}{T} + \frac{\log^2 |\Delta|}{y}\right)\right).$$
Set $T = |\Delta|^\varepsilon$. By the symmetry of the functional equation for real characters $\chi$ and the fact that $L(s, \chi)$ has no zeros in $(5.4)$, it follows that every zero appearing in the sum over $\rho$ satisfies $\Re\{\rho\} = \frac{1}{2}$. Thus, trivially bounding the remaining zeros, we deduce that

$$-\frac{L'}{L}(1, \chi) \ll \log y + y^{-1/2} \sum_{\substack{\rho = \frac{1}{2} + i\gamma \\ |\gamma| \leq |\Delta|^\varepsilon}} \frac{1}{1 + |\gamma|^2} + \frac{\log |\Delta|}{|\Delta|^\varepsilon} + \frac{\log^2 |\Delta|}{y}$$

for $y \geq 2$. Setting $y = (\log |\Delta|)^2 + 2$ implies $(5.3)$ holds for all $\Delta \in \mathcal{D}^*_\varepsilon(Q)$.

It remains to show that the number of discriminants $\Delta \in \mathcal{D}^*(Q) \setminus \mathcal{D}^*_\varepsilon(Q)$ is small. Jutila’s zero density estimate \cite[Theorem 2]{Jut75} implies that

$$\sum_{\Delta \in \mathcal{D}^*(Q)} N(\sigma, T, \chi_\Delta) \ll_{\varepsilon} (QT)^{\frac{7-6\sigma}{6-4\sigma}} + \frac{\varepsilon}{4},$$

where $N(\sigma, T, \chi)$ is the number of zeros $\rho = \beta + i\gamma$ of $L(s, \chi)$ with $\sigma < \beta < 1$ and $|\gamma| \leq T$. Setting $\sigma = 1/2$ and $T = Q^\varepsilon$, we see that the number of fundamental discriminants $\Delta \in \mathcal{D}^*(Q)$ whose $L$-function $L(s, \chi_\Delta)$ has a zero in the rectangle $(5.4)$ is at most $O_{\varepsilon}(Q^{3/4+\varepsilon})$. Hence, $|\mathcal{D}^*(Q) \setminus \mathcal{D}^*_\varepsilon(Q)| \ll_{\varepsilon} Q^{3/4+\varepsilon}$ as required. \hfill $\square$

Lemma 5.5 implies the same type of result for the set of all discriminants $\mathcal{D}(Q)$.
6. Congruence sum decomposition

Let \( f(u, v) = au^2 + buv + cv^2 \) be a form with discriminant \(-D\). For this section, we will not require \( f \) to be primitive. For any integer \( n \geq 0 \), define
\[
(6.1) \quad r_f(n) := \{|(u, v) \in \mathbb{Z}^2 : n = f(u, v)\}.
\]
Moreover, for \( x \geq 1 \) and positive integers \( \ell \) and \( d \), define
\[
(6.2) \quad A = A(x, f) := \{(u, v) \in \mathbb{Z}^2 : f(u, v) \leq x\},
A_\ell = A_\ell(x, f) := \{(u, v) \in A : f(u, v) \equiv 0 \pmod{\ell}\},
A_\ell(d) = A_\ell(x, f; d) := \{(u, v) \in A : (v, \ell) = d\}.
\]

We will suppress the dependence on \( x \) and \( f \) whenever it is clear from context. This will be the case for almost the entirety of the paper. Observe that
\[
(6.3) \quad |A| = \sum_{n \leq x} r_f(n) \quad \text{and} \quad |A_\ell| = \sum_{n \leq x} r_f(n) = \sum_{d | \ell} |A_\ell(d)|.
\]

Note the last identity holds since \( A_\ell \) is a disjoint union of the sets \( A_\ell(d) \) over \( d | \ell \). To calculate \( |A_\ell(d)| \), and subsequently \( |A_\ell| \), we will need to decompose it into sums similar to \( |A_\ell(1)| \) and estimate them with uniformity over all parameters. To this end, we introduce some additional notation. For any integer \( \ell \geq 1 \) and \( m \in \mathbb{Z}/\ell\mathbb{Z} \), define
\[
(6.4) \quad B_\ell = B_\ell(x, f) := \{(u, v) \in A : (v, \ell) = 1, f(u, v) \equiv 0 \pmod{\ell}\},
B_\ell(m) = B_\ell(x, f; m) := \{(u, v) \in A : (v, \ell) = 1, u \equiv mv \pmod{\ell}\}.
\]

Note that \( B_\ell \) is exactly the same as \( A_\ell(1) \), but we distinguish it for the sake of clarity. The crucial property of the sets \( B_\ell \) and \( B_\ell(m) \) is summarized in the following lemma.

**Lemma 6.1.** Let \( f(u, v) = au^2 + buv + cv^2 \) be a positive definite binary integral quadratic form of discriminant \(-D\) and let \( \ell \geq 1 \) be a squarefree integer. Define
\[
M(\ell) = M_f(\ell) := \{m \in \mathbb{Z}/\ell\mathbb{Z} : am^2 + bm + c \equiv 0 \pmod{\ell}\}.
\]

Then
\[
(6.5) \quad |B_\ell| = \sum_{m \in M(\ell)} |B_\ell(m)|.
\]

Furthermore, \( M(\ell) = M_f(\ell) := |M(\ell)| \) is a non-negative multiplicative function satisfying
\[
(6.6) \quad M(p) = \begin{cases} 1 + \chi(p) & \text{if } p \nmid a, \\ \chi(p) & \text{if } p \mid a \text{ and } p \nmid (a, b, c), \\ p & \text{if } p \mid (a, b, c), \end{cases}
\]
for all primes \( p \). Here \( \chi = \chi_{-D} \) is the corresponding Kronecker symbol.

**Proof.** Let \((u, v) \in B_\ell \). As \((v, \ell) = 1\), select \( m \in \mathbb{Z}/\ell\mathbb{Z} \) such that \( u \equiv mv \pmod{\ell} \). Thus,
\[
f(u, v) \equiv 0 \pmod{\ell} \iff (am^2 + bm + c)v^2 \equiv 0 \pmod{\ell} \iff m \in M(\ell).
\]
This implies \( B_\ell \) is a union of \( B_\ell(m) \) over \( m \in M(\ell) \). One can verify from (6.4) that \( m_1 \neq m_2 \pmod{\ell} \) implies \( B_\ell(m_1) \cap B_\ell(m_2) = \emptyset \). Thus, the union is in fact disjoint yielding (6.5).
Next, we count $M(\ell) = |\mathcal{M}(\ell)|$. The function $M(\ell)$ is multiplicative by the Chinese Remainder Theorem. Let $p$ be an odd prime. If $p \nmid a$ then $M(p) = 1 + \chi(p)$ by the definition of $\chi$. If $p \mid a$ then for $m \in \mathcal{M}(\ell)$

\begin{equation}
0 \equiv am^2 + bm + c \equiv bm + c \pmod{p}.
\end{equation}

Note in this scenario $\chi(p) = 0$ or 1 only. We consider cases.

- If $p \nmid b$ then $m \equiv -b^{-1}c \pmod{p}$ is the only solution to (6.7). Thus, $M(p) = 1 = (\frac{b^2 - 4ac}{p}) = \chi(p)$. If $p = 2$ then note $b^2 - 4ac \equiv 1 \pmod{8}$, so $\chi(2) = 1$ indeed.
- If $p \mid b$ then condition (6.7) becomes $c \equiv 0 \pmod{p}$. We further subdivide the cases.
  - If $p \nmid c$ then no value of $m$ satisfies (6.7) implying $M(p) = 0 = (\frac{b^2 - 4ac}{p}) = \chi(p)$.
  - If $p \mid c$ then $p \mid (a, b, c)$ in this subcase. Hence, all $m \in \mathbb{Z}/p\mathbb{Z}$ vacuously satisfy (6.7) so $M(p) = p$.

Comparing these cases, we see $M(p)$ indeed satisfies (6.6) for all odd primes $p$. For $p = 2$, one can verify by a tedious case analysis that $M(2)$ also satisfies (6.6). \hfill \Box

In light of Lemma 6.1, the main goal of this section is to determine the size of $|\mathcal{B}_\ell(m)|$ for any $m \in \mathbb{Z}/\ell\mathbb{Z}$. For convenience, set

\begin{equation}
V = V(x, f) := \sqrt{\frac{4ax}{\delta}}.
\end{equation}

This notation will be used throughout the paper. While we are more interested when $V \geq 1$, we only assume $x \geq 1$ in all of our arguments so it is possible that $0 < V < 1$. Recall $\varphi$ denotes the Euler totient function and $\tau$ is the divisor function.

\textbf{Lemma 6.2.} Let $f(u, v) = au^2 + buv + cv^2$ be a positive definite binary integral quadratic form of discriminant $-D$. Let $\ell \geq 1$ be a squarefree integer and $m \in \mathbb{Z}/\ell\mathbb{Z}$. For $x \geq 1$,

\begin{equation}
|\mathcal{B}_\ell(m)| = \frac{\varphi(\ell)}{\ell^2} \cdot \frac{\pi \sqrt{-D}}{2a} V^2 + O\left( V + \ell^{1/2} \tau(\ell) \frac{\sqrt{-D}}{a} V^{1/2} + \delta(\ell) \right),
\end{equation}

where $\mathcal{B}_\ell(m) = \mathcal{B}_\ell(x, f; m)$ is defined by (6.4), $V = V(x, f)$ is defined by (6.8), and $\delta(\ell)$ is the indicator function for $\ell = 1$.

\textbf{Remark.} We emphasize that the righthand side of (6.9) is independent $m \in \mathbb{Z}/\ell\mathbb{Z}$.

\textbf{Proof.} Counting the numbers of pairs $(u, v) \in \mathbb{Z}^2$ satisfying $f(u, v) \leq x$ amounts to verifying the inequality

$$(2au + bv)^2 + Dv^2 \leq 4ax.$$

Fixing $v$, any $u$ satisfying the above inequality lies in the range

\begin{equation}
\frac{-bv - \sqrt{4ax - Dv^2}}{2a} \leq u \leq \frac{-bv + \sqrt{4ax - Dv^2}}{2a}.
\end{equation}

Without loss, we may assume $m$ is an integer lying in $\{0, 1, \ldots, \ell - 1\}$. Restricting to $u = mv + j\ell$, we see that the integer $j$ must lie in the range

\begin{equation}
\frac{-(b + 2am)v - \sqrt{4ax - Dv^2}}{2a\ell} \leq j \leq \frac{-(b + 2am)v + \sqrt{4ax - Dv^2}}{2a\ell}.
\end{equation}

For each fixed $v$ and solution $u \equiv mv \pmod{\ell}$, the total number of such integers $j$ is therefore

$$\frac{1}{\ell} F(v) + O(1),$$

where
where

\[(6.10)\quad F(v) = \frac{1}{a} \sqrt{4av^2 - D} = \frac{\sqrt{D}}{a} \sqrt{V^2 - v^2}.\]

Now, summing over all integers \(v\) satisfying \(|v| \leq V\) and \((v, \ell) = 1\), we deduce that

\[
|B_\ell(m)| = \frac{1}{\ell} \sum_{|v| \leq V, (v, \ell) = 1} F(v) + O(\sum_{|v| \leq V})\]

The term \(v = 0\) contributes to the above sums if and only if \(\ell = 1\). Let \(\delta(\ell)\) be the indicator function for \(\ell = 1\). We separate the term \(v = 0\), if necessary, in the sums above and note \(F(v)\) is even to see that

\[(6.11)\quad |B_\ell(m)| = \frac{2}{\ell} \sum_{1 \leq v \leq V, (v, \ell) = 1} F(v) + \frac{\delta(\ell) \sqrt{D}}{\ell a} V + O(V + \delta(\ell)).\]

We remove the condition \((v, \ell) = 1\) via Mobius inversion and deduce that

\[(6.12)\quad \sum_{1 \leq v \leq V, (v, \ell) = 1} F(v) = \sum_{d \mid \ell} \mu(d) \sum_{1 \leq w \leq V/d} F(dw).\]

Set \(W_d := V/d\) and notice \(F(dw) = \frac{d \sqrt{D}}{a} \sqrt{w^2 - w^2}\). By Lemma 3.1 we see that

\[
\sum_{1 \leq w \leq V/d} F(dw) = \frac{d \sqrt{D}}{a} \left(\frac{\pi V^2}{4d^2} - \frac{V}{2d} + O(\sqrt{V/d})\right).
\]

Since \(\sum_{d \mid \ell} \frac{\mu(d)}{d} = \varphi(\ell)/\ell\), we deduce that

\[(6.13)\quad \sum_{1 \leq v \leq V, (v, \ell) = 1} F(v) = \frac{\varphi(\ell)}{\ell} \frac{\pi \sqrt{D}}{4ad} V^2 - \delta(\ell) \frac{\sqrt{D}}{2a} V + O\left(V^{1/2} \tau(\ell) \frac{\sqrt{D}}{a} V^{1/2}\right).
\]

Combining (6.11) and (6.13) yields (6.9). Note that the terms involving \(\delta(\ell)\) cancel. \(\square\)

We conclude this section by calculating \(|B_\ell|\).

**Lemma 6.3.** Let \(f(u, v) = au^2 + buv + cv^2\) be a positive definite binary integral quadratic form of discriminant \(-D\). If \(\ell \geq 1\) is a squarefree integer then

\[
|B_\ell| = M(\ell) \left(\frac{\varphi(\ell)}{\ell^2} \frac{\pi \sqrt{D}}{2a} V^2 + O \left(V + \ell^{1/2} \tau(\ell) \frac{\sqrt{D}}{a} V^{1/2} + \delta(\ell)\right)\right),
\]

where \(B_\ell = B_\ell(x, f)\) is defined by (6.4), \(V = V(x, f)\) is defined by (6.8), \(\delta(\ell)\) is the indicator function for \(\ell = 1\), and \(M(\ell) = M_\ell(\ell)\) is a multiplicative function defined by (6.6).

**Proof.** This is an immediate consequence of Lemmas 6.1 and 6.2 since the latter lemma’s estimates are uniform over all \(m \in \mathbb{Z}/\ell \mathbb{Z}\). \(\square\)
7. Local densities

We may now assemble our tools to establish the key technical proposition. Namely, we estimate the congruence sums given by (6.3) and calculate the local densities.

**Proposition 7.1.** Let $f$ be a primitive positive definite binary quadratic form with discriminant $-D$. If $\ell \geq 1$ is a squarefree integer then for any $\varepsilon > 0$ and $x \geq 1$,

$$|\mathcal{A}_\ell| = \sum_{n \leq x} r_f(n) = g(\ell)\frac{\pi \sqrt{-D}}{2a}V^2 + O\left(\tau_3(\ell)V + \ell^{1/2}\tau(\ell)\tau_3(\ell)\frac{\sqrt{-D}}{a}V^{1/2} + 1\right),$$

where $V = \sqrt{4ax/D}$ and $g$ is a multiplicative function satisfying

$$g(p) = \frac{1}{p}\left(1 + \chi(p) - \frac{\chi(p)}{p}\right)$$

for all primes $p$.

Here $\chi = \chi_{-D}$ is the corresponding Kronecker symbol.

**Proof.** Let $d \mid \ell$ and let $\mathcal{A}_\ell(d) = \mathcal{A}_\ell(x, f; d)$ be defined by (6.2). From observation (6.3), it suffices to calculate $|\mathcal{A}_\ell(d)|$. First, we introduce some notation. For any integer $r \geq 1$, set

$$f_r(u, w) := f(u, rw) = au^2 + bruw + cr^2w^2.$$  

Notice that its discriminant is $-r^2D$. Therefore, it follows for any $\alpha > 0$ that

$$V(\alpha^2 x, f_r) = \frac{\alpha V}{r}$$

and

$$\chi_{-r^2D}(n) = \begin{cases} \chi(n) & \text{if } (n, r) = 1 \\ 0 & \text{otherwise}, \end{cases}$$

where $V = V(x, f)$ and $\chi = \chi_{-D}$ as usual.

Now, write $\ell = dk$ so $(d, k) = 1$ as $\ell$ is squarefree. We wish to character each point $(u, v) \in \mathcal{A}_\ell(x, f; d)$. Since $(v, \ell) = d$, it follows by the Chinese Remainder Theorem that

$$f(u, v) \equiv 0 \pmod{\ell} \iff au^2 \equiv 0 \pmod{d} \quad \text{and} \quad f(u, v) \equiv 0 \pmod{k}$$

$$\iff u \equiv 0 \pmod{\frac{d}{(a, d)}} \quad \text{and} \quad f(u, v) \equiv 0 \pmod{k}.$$  

Write $u = \frac{d}{(a, d)}s$ and $v = dt$ for integers $s$ and $t$. Note $(t, k) = 1$ as $(v, \ell) = d$ and $\ell$ is squarefree. Then one can verify that

$$f(u, v) = \frac{d^2}{(a, d)^2} \cdot (as^2 + b(a, d)st + c(a, d)t^2) = \frac{d^2}{(a, d)^2} \cdot f(a, d)(s, t).$$

From this change of variables, (7.4), and (7.5), we see that

$$f(u, v) \equiv 0 \pmod{\ell} \iff f_{(a, d)}(s, t) \equiv 0 \pmod{k}$$

$$f(u, v) \leq x \iff f_{(a, d)}(s, t) \leq \frac{(a, d)^2}{d^2}x.$$  

Note by (7.5) that the congruence conditions are equivalent as $(d, k) = 1$ and $\ell = dk$. Since $(t, k) = 1$ necessarily, we have therefore established that

$$|\mathcal{A}_\ell(x, f; d)| = \left|\mathcal{B}_k\left(\frac{(a, d)^2x}{d^2}, f(a, d)\right)\right|.$$
Summing this identity over $d \mid \ell$, we apply observation (6.3) and Lemma 6.3 to deduce that

$$|A_{\ell}(x, f)| = \sum_{d=dk} B_k \left( \frac{(a, d)^2 x}{d^2}, f_{(a, d)} \right)$$

(7.8)

$$= \pi \sqrt{D} 2a V^2 \cdot \sum_{d=dk} M_{f(a, d)}(k) \frac{\varphi(k)}{k^2} \frac{(a, d)}{d^2} + O \left( V \cdot \sum_{d=dk} \frac{M_{f(a, d)}(k)}{d} \right)$$

$$+ \sqrt{D} a V^{1/2} \sum_{d=dk} M_{f(a, d)}(k) k^{1/2} \tau(k) d^{1/2} + \sum_{d=dk} M_{f(a, d)}(k) \delta(k) \right).$$

We wish to simplify the remaining sums and error term. Let $r \mid \ell$. As $f$ is primitive,

$$M_f(p) = \begin{cases} 1 + \chi(p) & \text{if } p \nmid a, \\ \chi(p) & \text{if } p \mid a, \end{cases}$$

by (6.6). To compute $M_f$, observe by the primitivity of $f$ that a prime $p$ divides $(a, br, cr^2)$ if and only if $p$ divides $(a, r)$. Moreover, if $p \mid r$ then $\chi_{r^2D}(p) = (\frac{r}{Dp}) = (\frac{r}{p}) (\frac{D}{p}) = 0$ and, similarly, if $p \nmid r$ then $\chi_{r^2D}(p) = \chi(p)$. Combining these observations with (6.6) and (7.3), we see that

$$M_{f(a, d)}(p) = \begin{cases} 1 + \chi(p) & \text{if } p \nmid a, \\ \chi(p) & \text{if } p \mid a \text{ and } p \nmid (a, d), \\ p & \text{if } p \mid (a, d). \end{cases}$$

In particular, as $(d, k) = 1$, it follows that $M_{f(a, d)}(k) = M_f(k)$. Hence,

$$\sum_{d=dk} M_{f(a, d)}(k) \frac{\varphi(k)}{k^2} \frac{(a, d)}{d^2}$$

$$= \sum_{d=dk} M_f(k) \cdot \frac{\varphi(k)}{k^2} \cdot \frac{(a, d)}{d^2}$$

(7.9)

$$= \prod_{p \mid \ell, p \nmid a} \left( (1 + \chi(p)) \left( \frac{1}{p} - \frac{1}{p^2} \right) + \frac{1}{p} \right) \times \prod_{p \mid (\ell, a)} \left( \chi(p) \left( \frac{1}{p} - \frac{1}{p^2} \right) + \frac{1}{p} \right)$$

$$= \prod_{p \mid \ell} \left( 1 + \chi(p) \right) \frac{1}{p} - \chi(p) \frac{1}{p^2} = g(\ell).$$

Similarly,

$$\sum_{d=dk} \frac{M_{f(a, d)}(k)}{d} = \prod_{p \mid \ell, p \nmid a} \left( 1 + \chi(p) + \frac{1}{p} \right) \times \prod_{p \mid (\ell, a)} \left( \chi(p) + \frac{1}{p} \right) \ll \tau_3(\ell),$$

(7.10)

which implies that

$$\sum_{d=dk} M_{f(a, d)}(k) k^{1/2} \tau(k) d^{-1/2} \ll \ell^{1/2} \tau(\ell) \sum_{d=dk} \frac{M_{f(a, d)}(k)}{d} \ll \ell^{1/2} \tau(\ell) \tau_3(\ell).$$

(7.11)
Combining the observation that
\[ \sum_{\ell=dk} M_{f(a,d)}(k)\delta(k) = M_{f(a,d)}(1) = 1 \]
with (7.8), (7.9), (7.10), and (7.11) yields the desired result. \(\square\)

To obtain a better intuition for the quality of Proposition 7.1, we present the special case when \(\ell = 1\) as a corollary below. We do not claim that this corollary is new, but we have not seen it stated in the literature and thought it may be of independent interest.

**Corollary 7.2.** Let \(f(u,v) = au^2 + buv + cv^2\) be a primitive positive definite binary quadratic form with discriminant \(-D\). For \(x \geq 1\),
\[ \sum_{n \leq x} r_f(n) = \frac{2\pi x}{\sqrt{D}} + O\left(\frac{(ax)^{1/2}}{D^{1/2}} + \frac{(Dx)^{1/4}}{a^{3/4}} + 1\right). \]

**Remarks.** Suppose \(f\) is reduced so \(|b| \leq a \leq c\). It is well known (see e.g. [BG06, Lemma 3.1]) that
\[ \sum_{n \leq x} r_f(n) = \frac{2\pi x}{\sqrt{D}} + O\left(\frac{x^{1/2}}{a^{1/2}} + 1\right). \]
This estimate, like Corollary 7.2, gives the asymptotic \(\sim \frac{2\pi x}{\sqrt{D}}\) as long as \(x/c \to \infty\), but the error term in Corollary 7.2 is stronger than the above whenever \(x \geq c\). The source of this improvement is a standard analysis of the sawtooth function in Lemma 3.1 and its subsequent application in Lemma 6.2. As discussed in Section 1.2, the condition \(x \geq c\) is the ‘non-trivial’ range for counting the lattice points inside the ellipse \(f(u,v) \leq x\) whenever \(f\) is reduced.

### 8. Application of Selberg’s sieve

We now apply Selberg’s sieve to give an upper bound for the number of primes in a short interval represented by a reduced positive definite primitive integral binary quadratic form. We leave the calculation of the main term’s implied constant unfinished as the final arguments vary slightly for Theorems 1.3 and 1.4.

**Proposition 8.1.** Let \(f(u,v) = au^2 + buv + cv^2\) be a reduced positive definite integral binary quadratic form with discriminant \(-D\). Let \((ax)^{1/2} \leq y \leq x\). Set
\[ z = \left(\frac{a}{Dx}\right)^{1/4} y^{1/2} (\log y)^{-7} + 1. \]
If \(x \geq D/a\) then
\[ \pi_f(x) - \pi_f(x-y) < \left\{ \frac{\log y}{\mathcal{J}} + O((\log y)^{-1}) \right\} \frac{\delta_{f,y}}{h(-D) \log y}, \]
where
\[ \mathcal{J} = \begin{cases} \frac{1}{L(1,\chi)} \sum_{\ell<z} g(\ell) & \text{if } L(1,\chi) \geq (\log y)^{-2}, \\ (\log y)^2 & \text{otherwise.} \end{cases} \]
Here \(\chi = \chi_{-D}\) is the corresponding Kronecker symbol and \(g\) is the completely multiplicative function defined by (7.2).
Proof. As \( f \) is reduced, we have that \(|b| \leq a \leq c\) and moreover \( c \gtrsim D/a \geq \sqrt{D} \geq a\). We will frequently apply these properties while only mentioning that \( f \) is reduced.

Our argument is divided according to the size of \( L(1, \chi) \). First, assume \( L(1, \chi) < (\log y)^{-2} \). Let \( w_{-D} \) be the number of roots of unity contained in \( \mathbb{Q}(\sqrt{-D}) \). Trivially, by Corollary 7.2

\[
\frac{w_{-D}}{\delta_f} (\pi_f(x) - \pi_f(x - y)) \leq \sum_{x-y < n \leq x} r_f(n) = \frac{2\pi y}{\sqrt{D}} + O\left(\frac{(ax)^{1/2}}{D^{1/2}}\right)
\]

because \( x \geq D/a \) and \( f \) is reduced. Thus, by the class number formula

\[
(8.3) \quad h(-D) = \frac{w_{-D} \sqrt{D}}{2\pi} L(1, \chi)
\]

and our assumption on \( L(1, \chi) \),

\[
\pi_f(x) - \pi_f(x - y) \leq L(1, \chi) \left\{ 1 + O\left(\frac{(ax)^{1/2}}{y}\right) \right\} \frac{\delta_f y}{h(-D)} \ll \frac{y}{h(-D) (\log y)^2}.
\]

In the last step, we used that \( y \geq (ax)^{1/2} \). This establishes \([8.2]\) when \( L(1, \chi) < (\log y)^{-2} \). Therefore, we may henceforth assume

\[
(8.4) \quad L(1, \chi) \geq (\log y)^{-2}.
\]

Defining \( P = P(z) = \prod_{p \leq z} p \), it follows that

\[
(8.5) \quad \frac{w_{-D}}{\delta_f} (\pi_f(x) - \pi_f(x - y)) \leq \sum_{x-y < n \leq x} r_f(n) + \frac{w_{-D}}{\delta_f} \pi(z),
\]

where \( \pi(z) \) is the number of primes up to \( z \). We proceed to estimate the sieved sum. Using Proposition 7.1 and Selberg’s upper bound sieve [FI10, Theorem 7.1] with level of distribution \( z^2 \), we see that

\[
(8.6) \quad \sum_{x-y < n \leq x} r_f(n) < \frac{2\pi y}{\sqrt{D} J} + \sum_{\ell \mid P \atop \ell < z^2} r_\ell \lambda_\ell,
\]

where

\[
J = \sum_{\ell \mid P \atop \ell < z} h(\ell), \quad h(\ell) = \prod_{p \mid \ell} \frac{g(p)}{1 - g(p)}, \quad |\lambda_\ell| \leq \tau_3(\ell),
\]

and

\[
r_\ell \ll \tau_3(\ell) V + \ell^{1/2} \tau(\ell) \tau_3(\ell) \frac{\sqrt{D}}{a} V^{1/2}.
\]

Here, as usual, \( V = \sqrt{4ax/D} \). Note 1 \( \leq V \leq x^{1/2} \) as \( x \geq D/a \) and \( a \leq \sqrt{D} \). For the quantity \( J \) in the main term, we treat \( g \) as a completely multiplicative function and note that

\[
(8.7) \quad J \geq \sum_{\ell \mid z} g(\ell) = L(1, \chi) J
\]

23
by \((8.4)\). The remainder term in \((8.6)\) is bounded in a straightforward manner. Using standard estimates for the \(k\)-divisor function \(\tau_k(\ell)\) (see, e.g., \([\text{LT17}]\)) and the prime number theorem, one can verify that

\[
\sum_{\ell < z^2} \tau_3(\ell) + \sqrt{D} V^{1/2} \sum_{\ell < z^2} \ell^{1/2} \tau_3(\ell) \\
\ll z^2 (\log z)^8 \cdot V + z^3 (\log z)^{17} \cdot \frac{\sqrt{D}}{a} V^{1/2} \\
\ll ay D (\log y)^{-6} + \frac{y^{3/2}x^{-1/2}}{\sqrt{D}} (\log y)^{-4}.
\]

In the last step, we used that \(V = \sqrt{4ax/D}\) and, by \((8.1)\), \(z = \left(\frac{a}{Dx}\right)^{1/4} y^{1/2} (\log y)^{-7} + 1 \leq y\). Since \(y \leq x\) and \(a \leq \sqrt{D}\), we see that the above is

\[
\ll \frac{y}{\sqrt{D} (\log y)^4}.
\]

Thus, applying the class number formula \((8.3)\) and the well-known estimate \(L(1, \chi) \ll \log D \ll \log y\), we conclude that

\[
\frac{w-D}{\delta_f} \pi(z) + \sum_{\ell \mid p, \ell < z^2} r_\ell \lambda_\ell \ll \frac{y}{h(-D)(\log y)^3}.
\]

(8.8)

Combining \((8.5)\), \((8.6)\), \((8.7)\), and \((8.8)\) completes the proof of the proposition with a final application of the class number formula.

Evidently, from Proposition \(8.1\), we will require a lower bound for the sum of local densities. We execute the first steps here.

**Lemma 8.2.** Let \(g\) be the completely multiplicative function defined by \((7.2)\). For \(z \geq 1\),

\[
\sum_{\ell < z} g(\ell) \geq L(1, \chi) \log z + L'(1, \chi) + O\left( L(1, \chi) + \mathcal{E}_1(z; \chi) + z^{-1} \mathcal{E}_0(z; \chi) \right).
\]

Here \(\mathcal{E}_1(z; \chi)\) and \(\mathcal{E}_0(z; \chi)\) are defined as in Lemma \(3.2\).

**Proof.** Define

\[
G(s) := \sum_{n=1}^{\infty} g(n)n^{-s} = \prod_p \left( 1 - g(p)p^{-s} \right)^{-1},
\]

which absolutely converges for \(\Re\{s\} > 0\) since \(|g(p)| \leq 2/p\). One can verify that for \(\Re\{s\} > 0\)

\[
G(s) = \zeta(s+1)L(s+1, \chi)\tilde{G}(s),
\]

where \(\zeta(s)\) is the Riemann zeta function, \(L(s, \chi)\) is the Dirichlet \(L\)-function attached to the quadratic character \(\chi = \chi_{-D}\), and

\[
\tilde{G}(s) := \prod_p \left( 1 - \frac{\chi(p)p^{-s-2} - \chi(p)p^{-2s-2}}{1 - (1 + \chi(p))p^{-s-1} + \chi(p)p^{-s-2}} \right).
\]

24
It is straightforward to check that \( \tilde{G}(s) \) is absolutely convergent for \( \Re\{s\} > -1 \) and, in particular, \( \tilde{G}(0) = 1 \). Expanding the Euler product for \( \tilde{G}(s) \) and writing \( \tilde{G}(s) = \sum_n \tilde{g}(n)n^{-s} \) for some multiplicative function \( \tilde{g} \), one can see that
\[
\tilde{g}(n) \ll n^{-2}.
\]

As \( \tilde{G}(0) = 1 \), it follows that
\[
\sum_{n \leq N} \tilde{g}(n) = 1 + O(N^{-1}).
\]

Therefore, from (8.9),
\[
\sum_{\ell < z} g(\ell) = \sum_{\ell < \ell = mn} (1 * \chi)(m) \tilde{g}(n) = \sum_{m < z} (1 * \chi)(m) + O\left(\frac{z^{-1} \sum_{m < z} (1 * \chi)(m)}{z}\right).
\]

The desired result now follows from Lemma 3.2. \( \square \)

9. Representation of primes

We may finally prove Theorems 1.3 and 1.4. In both cases, we will need to apply Proposition 8.1 from which one can see that it suffices to provide an appropriate lower bound for \( J \) when \( L(1, \chi) \geq (\log y)^{-2} \). By Lemma 8.2, it follows that
\[
(9.1) \quad J \geq \log z + \frac{L'}{L}(1, \chi) + O\left(1 + \mathcal{E}_1(z; \chi)(\log y)^2 + \mathcal{E}_0(z; \chi)(\log y)^2\right)
\]

provided \( L(1, \chi) \geq (\log y)^{-2} \), \( (ax)^{1/2} \leq y \leq x \), and \( z \) is given by (8.1). Recall that \( \mathcal{E}_0 \) and \( \mathcal{E}_1 \) are given by Lemma 3.2 with estimates exhibited in Sections 4 and 5. The proofs for both theorems will employ (9.1).

Before we proceed, we wish to emphasize that \( f(u, v) = au^2 + buv + cv^2 \) is assumed to be a reduced positive definite binary integral quadratic form of discriminant \(-D\). Thus, \( |b| \leq a \leq c \) and \( a \leq \sqrt{D} \).

9.1. Proof of Theorem 1.3. Recall we are assuming that
\[
\left(\frac{D^{1+4\phi}}{a}\right)^{1/2+\epsilon} x^{1/2+\epsilon} \leq y \leq x,
\]
where \( \phi \) is given by (1.8). As \( a \leq \sqrt{D} \), this implies that \( y \geq (ax)^{1/2} \). Furthermore,
\[
z = \left(\frac{a}{Dx}\right)^{1/4} y^{1/4} (\log y)^{-7} + 1 \gg \varepsilon \, D^{\phi+\epsilon/2}, \quad \text{and} \quad \log z \asymp \log y \asymp \log x.
\]

Therefore, applying Lemma 4.4 (or Lemma 4.5 when assuming GRH) and Lemma 4.3 to (9.1), it follows that
\[
J \geq \log z - \left(\phi + \frac{\varepsilon}{4}\right) \log D + O_\varepsilon(1 + z^{-\varepsilon^2} \log^2 y)
\]
\[
\geq \frac{1}{2} \log y - \frac{1}{4} \log x - \left(\frac{1}{4} + \frac{\phi}{2} + \frac{\varepsilon}{4}\right) \log D + \frac{1}{4} \log a + O_\varepsilon(\log \log y)
\]
\[
= \frac{1 - \theta}{2} \log y + O_\varepsilon(\log \log y),
\]

\[
\log \log n + O_\varepsilon(\log \log n),
\]
Thus, incorporating (9.2) and (9.3) into (9.1), we again obtain (9.6) whenever (9.7) holds and

\[ L(1, \chi) = (\log y)^{-2}. \]

9.2. Proof of Theorem 1.4. Recall \( \mathcal{D}(Q) \) is given by (1.11). Let \( c_1(\varepsilon) > 0 \) be a sufficiently small constant and \( C_1(\varepsilon), C_2(\varepsilon) \geq 1 \) be sufficiently large constants, all of which depend only on \( \varepsilon \). For \( Q \geq 3 \), let \( \mathcal{D}_\varepsilon(Q) \) be the subset of discriminants \( -D \in \mathcal{D}(Q) \) such that

\[ (\log y)^{-1} \]

Thus, it suffices to show for every discriminant \( -D \in \mathcal{D}_\varepsilon(Q) \) and reduced positive definite binary quadratic form \( f \) of discriminant \( -D \) that

\[ \pi_f(x) - \pi_f(x - y) < \frac{2}{1 - \theta'} \delta_f y \log y \{ 1 + O_\varepsilon \left( \frac{\log \log y}{\log y} \right) \} \]

provided \( \left( \frac{Dx}{a} \right)^{1/2+\varepsilon} \leq y \leq x \). Here \( \theta' \) is defined as in Theorem 1.3 with \( \phi = 0 \). First, assume

\[ (\log y)^{-1} \]

Arguing as in Section 9.1, it follows that \( y \geq (ax)^{1/2} \),

\[ z = \left( \frac{a}{Dx} \right)^{1/4} y^{1/2} (\log y)^{-7} + 1 \geq \varepsilon D^{1/4+\varepsilon/2}, \quad \text{and} \quad \log z \preceq \log y \preceq \log x. \]

Thus, incorporating (9.2) and (4.1) from Lemma 4.3 into (9.4), it similarly follows that

\[ \mathcal{J} \geq \frac{1 - \theta'}{2} \log y + O_\varepsilon (\log \log y) \]

whenever \( L(1, \chi) \geq (\log y)^{-2} \). Therefore, by Proposition 8.1, this establishes (9.4) provided (9.5) holds and \( -D \in \mathcal{D}_\varepsilon(Q) \). It remains to consider the case when

\[ (\log y)^{-1} \]

Note we continue to assume \( -D \in \mathcal{D}_\varepsilon(Q) \). As before, it follows that \( y \geq (ax)^{1/2} \),

\[ z = \left( \frac{a}{Dx} \right)^{1/4} y^{1/2} (\log y)^{-7} + 1 \geq \varepsilon D^{1/2+\varepsilon}, \quad \text{and} \quad \log z \preceq \log y \preceq \log x. \]

Thus, incorporating (9.2) and (9.3) into (9.1), we again obtain (9.6) whenever \( L(1, \chi) \geq (\log y)^{-2} \). By Proposition 8.1, this establishes (9.4) provided (9.7) holds and \( -D \in \mathcal{D}_\varepsilon(Q) \). This completes the proof in all cases. \( \square \)
10. Representation of small integers with few prime factors

Proof of Theorem 10.3. We apply the beta sieve to the sequence $A = A(x, f)$ given by (6.2). Let $g$ be the local density function defined in Proposition 7.1, so $g(p) \leq 2/p$. Thus, the sequence $A$ is of dimension at most $\kappa = 2$ and has sifting limit $\beta = \beta(\kappa) < 4.85$ according to [FI10, Section 11.19]. Let $x \geq D/a$ and select $z := V^{10/49}$.

where $V = \sqrt{4ax/D} \geq 2$. Select the level of distribution to be $R = z^{485/100} > z^\beta$. Thus, by [FI10, Theorem 11.13] and Proposition 7.1, it follows that

\[
\sum_{n \leq x \atop (n, P(z)) = 1} r_f(n) \gg \frac{x}{\sqrt{D} (\log x)^2} + O\left( \sum_{\ell \mid P(z) \atop \ell < R} |r_{\ell}| \right),
\]

where

\[ |r_{\ell}| \ll \ell^\eta V + \ell^{1/2+\eta} \sqrt{\frac{D}{a}} V^{1/2} + 1 \]

for fixed $\eta > 0$ sufficiently small. Since $f$ is reduced and $R = z^{485/100} = V^{97/98}$, we see that

\[
\sum_{\ell \mid P(z) \atop \ell < R} |r_{\ell}| \ll R^{1+\eta} V + \frac{\sqrt{D}}{a} R^{3/2+\eta} V^{1/2} + R^{1-\eta} \ll \frac{\sqrt{D}}{a} V^{2-\frac{3}{196}+\eta} \ll \frac{x^{1-\frac{3}{196}+\eta}}{\sqrt{D}}.
\]

Thus, as $\eta > 0$ is sufficiently small,

\[
\sum_{n \leq x \atop (n, P(z)) = 1} r_f(n) \gg \frac{x}{\sqrt{D} (\log x)^2}
\]

for $x \geq D/a$. For an integer $k \geq 10$, observe that $z \geq x^{1/k}$ if and only if

\[
\left( \frac{ax}{D} \right)^{\frac{5k}{49}} \geq x \iff x \geq \left( \frac{D}{a} \right)^{1+\frac{49}{5k-49}}.
\]

This completes the proof. \qed

\section*{References}

[BG06] Valentin Blomer and Andrew Granville. Estimates for representation numbers of quadratic forms. \textit{Duke Math. J.}, 135(2):261–302, 2006.

[Bur63] D. A. Burgess. On character sums and $L$-series. II. \textit{Proc. London Math. Soc. (3)}, 13:524–536, 1963.

[Cox13] David A. Cox. \textit{Primes of the form $x^2 + ny^2$}. Pure and Applied Mathematics (Hoboken). John Wiley & Sons, Inc., Hoboken, NJ, second edition, 2013. Fermat, class field theory, and complex multiplication.

[Deb16] Korneel Debaene. Explicit counting of ideals and a brun-titchmarsh inequality for the chebotarev density theorem. \textit{arXiv preprint arXiv:1611.10103}, 2016.

[Dit13a] Jakob Ditchen. On the average distribution of primes represented by binary quadratic forms. \textit{arXiv preprint arXiv:1312.1502}, 2013.

[Dit13b] Jakob Johann Ditchen. \textit{Primes of the shape $x^2 + ny^2$. The distribution on average and prime number races}. PhD thesis, ETH Zürich, 2013.

[FI10] John Friedlander and Henryk Iwaniec. \textit{Opera de cribro}, volume 57 of American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, RI, 2010.

[Fog62] E. Fogels. On the zeros of Hecke’s $L$-functions. I, II. \textit{Acta Arith.}, 7:87–106, 131–147, 1961/1962.
D. R. Heath-Brown. Zero-free regions for Dirichlet $L$-functions, and the least prime in an arithmetic progression. *Proc. London Math. Soc. (3)*, 64(2):265–338, 1992.

D. R. Heath-Brown. A mean value estimate for real character sums. *Acta Arith.*, 72(3):235–275, 1995.

Henryk Iwaniec and Emmanuel Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

Yasutaka Ihara, V. Kumar Murty, and Mahoro Shimura. On the logarithmic derivatives of Dirichlet $L$-functions at $s = 1$. *Acta Arith.*, 137(3):253–276, 2009.

Matti Jutila. On mean values of Dirichlet polynomials with real characters. *Acta Arith.*, 27:191–198, 1975. Collection of articles in memory of Jurii Vladimirovič Linnik.

E. Kowalski and P. Michel. Zeros of families of automorphic $L$-functions close to 1. *Pacific J. Math.*, 207(2):411–431, 2002.

J. C. Lagarias, H. L. Montgomery, and A. M. Odlyzko. A bound for the least prime ideal in the Chebotarev density theorem. *Invent. Math.*, 54(3):271–296, 1979.

J. C. Lagarias and A. M. Odlyzko. Effective versions of the Chebotarev density theorem. pages 409–464, 1977.

Florian Luca and László Tóth. The $r$th moment of the divisor function: an elementary approach. *Journal of Integer Sequences*, 20:17.7.4, 2017.

M. Mourtada and V. Kumar Murty. Omega theorems for $L'(1,\chi_D)$. *Int. J. Number Theory*, 9(3):561–581, 2013.

H. L. Montgomery and R. C. Vaughan. The large sieve. *Mathematika*, 20:119–134, 1973.

Hugh L. Montgomery and Robert C. Vaughan. *Multiplicative number theory. I. Classical theory*, volume 97 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2007.

N. Tschebotareff. Die Bestimmung der Dichtigkeit einer Menge von Primzahlen, welche zu einer gegebenen Substitutionsklasse gehören. *Math. Ann.*, 95(1):191–228, 1926.

Jesse Thorner and Asif Zaman. A Chebotarev Variant of the Brun-Titchmarsh Theorem and Bounds for the Lang-Trotter conjectures. *Int. Math. Res. Not.*, 2017.

Jesse Thorner and Asif Zaman. An explicit bound for the least prime ideal in the Chebotarev density theorem. *Algebra Number Theory*, 11(5):1135–1197, 2017.

Alfred Weiss. The least prime ideal. *J. Reine Angew. Math.*, 338:56–94, 1983.

Asif Zaman, Department of Mathematics, Stanford University, 450 Serra Mall, Building 380, Stanford, CA, USA, 94305

*E-mail address*: aazaman@stanford.edu