Algebra & Number Theory

Volume 13
2019
No. 5

A unified and improved Chebotarev density theorem

Jesse Thorner and Asif Zaman

msp

msp

ALGEBRA AND NUMBER THEORY 13:5 (2019)
dx.doi.org/10.2140/ant.2019.13.1039

A unified and improved Chebotarev density theorem

Jesse Thorner and Asif Zaman

We establish an unconditional effective Chebotarev density theorem that improves uniformly over the well-known result of Lagarias and Odlyzko. As a consequence, we give a new asymptotic form of the Chebotarev density theorem that can count much smaller primes with arbitrary log-power savings, even in the case where a Landau–Siegel zero is present. Our main theorem also interpolates the strongest unconditional upper bound for the least prime ideal with a given Artin symbol as well as the Chebotarev analogue of the Brun–Titchmarsh theorem proved by the authors.

1. Introduction and statement of results

1A. ***Introduction.*** Let L/F be a Galois extension of number fields with Galois group G. For each prime ideal $\mathfrak{p}$ of F that is unramified in L, we use the Artin symbol $\left[\frac{L/F}{\mathfrak{p}}\right]$ to denote the conjugacy class of G consisting of the set of Frobenius automorphisms attached to the prime ideals $\mathfrak{P}$ of L which lie over $\mathfrak{p}$. For any conjugacy class $C \subseteq G$, define the function

$$\pi_C(x) = \pi_C(x, L/F) = \#\left\{\mathrm{N}_{F/\mathbb{Q}}\mathfrak{p} \le x : \mathfrak{p} \text{ unramified in } L, \left[\tfrac{L/F}{\mathfrak{p}}\right] = C\right\}, \tag{1-1}$$

where $\mathrm{N}_{F/\mathbb{Q}}$ is the absolute norm of $F/\mathbb{Q}$. The Chebotarev density theorem states that

$$\pi_C(x) \sim \frac{|C|}{|G|}\operatorname{Li}(x) \quad \text{as } x \to \infty.$$

It follows from work of V.K. Murty [1997, Section 4] that there exists an absolute, effective, and positive constant c_1 such that

$$\pi_C(x) = \frac{|C|}{|G|}(\operatorname{Li}(x) - \theta_1 \operatorname{Li}(x^{\beta_1}) + O(xe^{-c_1\sqrt{\frac{\log x}{n_L}}})), \quad \log x \gg \frac{(\log D_L)^2}{n_L} + n_L(\log n_L)^2, \tag{1-2}$$

which refines a well-known result of Lagarias and Odlyzko [1977, Theorem 1.2]. Here, D_L is the absolute discriminant of L, $n_L = [L : \mathbb{Q}]$ is the degree of L over $\mathbb{Q}$, β_1 is a possible Landau–Siegel zero of the Dedekind zeta function $\zeta_L(s)$ of L, and $\theta_1 = \theta_1(C) \in \{-1, 0, 1\}$ depends on C; in particular, $\theta_1(C) = 0$ if and only if β_1 does not exist. For comparison, Lagarias and Odlyzko [1977, Theorem 1.1] proved that the generalized Riemann hypothesis for $\zeta_L(s)$ implies the more uniform result

$$\pi_C(x) = \frac{|C|}{|G|}(\operatorname{Li}(x) + O(\sqrt{x}\log(D_L x^{n_L}))), \quad x \gg (\log D_L)^2(\log\log D_L)^4. \tag{1-3}$$

MSC2010: 11R44.
Keywords: distribution of primes, Chebotarev density theorem, effective, uniform, binary quadratic forms.

As of now, the best bound for β_1 is due to Stark [1974, Theorem 1', p. 148]; it implies that

$$1-\beta_1 \gg (n_L^{n_L}\log D_L + D_L^{1/n_L})^{-1}. \tag{1-4}$$

Therefore, in order to ensure that $\frac{|C|}{|G|}\operatorname{Li}(x)$ dominates all other terms in (1-2), one must take the range of x to be

$$\log x \gg n_L^{-1}(\log D_L)^2 + n_L(\log n_L)^2 + (1-\beta_1)^{-1} \tag{1-5}$$

and apply (1-4) if β_1 exists. Otherwise, one omits the last term in (1-5) if β_1 does not exist. Regardless, (1-5) is very prohibitive in many applications where uniformity in L/F is crucial. Thus it often helps in applications to have upper and lower bounds for $\pi_C(x)$ of order $\operatorname{Li}(x)$ in ranges of x which are more commensurate with (1-3). Lagarias, Montgomery, and Odlyzko [1979] made substantial progress on these problems; their work has been improved upon by Weiss [1983], the authors [Thorner and Zaman 2017; 2018], and Zaman [2017]. In particular, it follows from the joint work of the authors [Thorner and Zaman 2017; 2018] that there exist absolute, effective constants $A > 2$ and $B > 2$ such that if D_L is sufficiently large, then

$$\frac{1}{(D_L n_L^{n_L})^A}\frac{|C|}{|G|}\operatorname{Li}(x) \ll \pi_C(x) < (2+o(1))\frac{|C|}{|G|}\operatorname{Li}(x) \quad \text{for } x \geq (D_L n_L^{n_L})^B, \tag{1-6}$$

where the $o(1)$ term tends to zero as $(\log x)/\log(D_L n_L^{n_L})$ tends to infinity.[1]

To summarize the above discussion, suppose that we are in the worst case scenario with $\theta_1 = 1$ and β_1 is as bad as (1-4) permits. If one is willing to sacrifice an asymptotic equality for $\pi_C(x)$ in order to obtain estimates in noticeably better ranges than (1-5), then one might use (1-6). On the other hand, if one needs an asymptotic equality for $\pi_C(x)$, then one uses (1-2) in the prohibitive range (1-5).

1B. ***Results.*** Our main result, Theorem 1.4, is a new asymptotic equality for $\pi_C(x)$ which interpolates both of the aforementioned options while providing several new options. In other words, we prove a new asymptotic equality for $\pi_C(x)$ from which one may deduce both (1-2) and (1-6). First, we present a simplified version of the main result.

Theorem 1.1. *Let L/F be a Galois extension of number fields with Galois group G, and let $C \subseteq G$ be a conjugacy class. Let β_1 denote the Landau–Siegel zero of the Dedekind zeta function $\zeta_L(s)$, if it exists. There exist absolute and effective constants $c_2 > 0$ and $c_3 > 0$ such that if $L \neq \mathbb{Q}$ and $x \geq (D_L n_L^{n_L})^{c_2}$, then*

$$\pi_C(x) = \frac{|C|}{|G|}(\operatorname{Li}(x) - \theta_1 \operatorname{Li}(x^{\beta_1}))\left(1 + O\left(\exp\left[-\frac{c_3\log x}{\log(D_L n_L^{n_L})}\right] + \exp\left[-\frac{(c_3\log x)^{1/2}}{n_L^{1/2}}\right]\right)\right),$$

where $\theta_1 = \theta_1(C) \in \{-1, 0, 1\}$. In particular, $\theta_1 = 0$ precisely when β_1 does not exist.

[1] The term $n_L^{n_L}$ is usually negligible compared to a power of D_L. If not, one might appeal to [Zaman 2017, Theorem 1.3.1] which states that $\pi_C(x) \gg D_L^{-A}\frac{|C|}{|G|}\operatorname{Li}(x)$ for $x \geq D_L^B$.

The inequality

$$\exp\left[-\frac{(c_3\log x)^{1/2}}{n_L^{1/2}}\right] \gg \exp\left[-\frac{c_3\log x}{\log(D_L n_L^{n_L})}\right],$$

holds when $\log x \gg (\log D_L)^2/n_L + n_L(\log n_L)^2$, so we see that Theorem 1.1 recovers (1-2) and is therefore a uniform improvement over it. Also, it follows from the mean value theorem and (1-4) that

$$\operatorname{Li}(x) - \theta_1 \operatorname{Li}(x^{\beta_1}) \gg ((1-\beta_1)\log(D_L n_L^{n_L}))\operatorname{Li}(x) \gg \frac{\log(D_L n_L^{n_L})}{D_L^{1/n_L} + n_L^{n_L}\log D_L}\operatorname{Li}(x). \tag{1-7}$$

With this lower bound at our disposal, one can see that Theorem 1.1 recovers (1-6). Thus Theorem 1.1 unifies and improves both (1-2) and (1-6).

As noted above, if one wants $\frac{|C|}{|G|}\operatorname{Li}(x)$ to dominate all other terms in (1-2), then one must take x in the range (1-5). However, one can plainly see that

$$\frac{|C|}{|G|}(\operatorname{Li}(x) - \theta_1 \operatorname{Li}(x^{\beta_1})) \tag{1-8}$$

dominates all other terms in Theorem 1.1 for all x in the claimed range, provided that c_2 is suitably large compared to c_3. At first glance, it may seem awkward that we adjoin the contribution from β_1 to the "main term" when it is classically viewed as an error term. But without eliminating the existence of β_1, it is well known that in situations where $\theta_1 \neq 0$ and x is small, say $\log x \ll \log(D_L n_L^{n_L})$, the term $-\theta_1\frac{|C|}{|G|}\operatorname{Li}(x^{\beta_1})$ is more properly treated as a secondary term than an error term. When $\theta_1 = 1$ and β_1 is especially close to 1, this secondary term causes serious difficulties in the proof of Linnik's bound [1944] for the least prime in an arithmetic progression. Fortunately, it follows from (1-7) that regardless of whether β_1 exists, we have

$$\operatorname{Li}(x) \ll_L \operatorname{Li}(x) - \theta_1 \operatorname{Li}(x^{\beta_1}) < 2\operatorname{Li}(x). \tag{1-9}$$

Therefore, in the range of x where $-\frac{|C|}{|G|}\theta_1 \operatorname{Li}(x^{\beta_1})$ acts like a secondary term, (1-9) shows that Theorem 1.1 recovers upper and lower bounds of order $\operatorname{Li}(x)$ precisely because (1-8) dominates all other terms in Theorem 1.1. This perspective is implicit in Linnik's work. On the other hand, when x is sufficiently large in terms of L/F per (1-5), the contribution from β_1 can be safely absorbed into the O-term in Theorem 1.1. In light of these observations, we believe that viewing (1-8) as the "main term" in Theorem 1.1 helps to clarify the role of the contribution from β_1 when one transitions from small values of x to large values of x.

Upon considering the O-term in Theorem 1.1, we see that Theorem 1.1 noticeably improves the range of x in which we have an asymptotic equality for $\pi_C(x)$.

Corollary 1.2. *If* $\log x/\log(D_L n_L^{n_L}) \to \infty$, *then* $\pi_C(x) \sim \frac{|C|}{|G|}(\operatorname{Li}(x) - \theta_1 \operatorname{Li}(x^{\beta_1}))$.

Theorem 1.1 also produces a new asymptotic equality in which the error term saves an arbitrarily large power of $\log x$ in a much stronger range of x than (1-2).

Corollary 1.3. *Let* $A > 1$. *If* $\log x \gg_A (\log D_L)(\log\log D_L) + n_L(\log n_L)^2$, *then*

$$\pi_C(x) = \frac{|C|}{|G|}(\operatorname{Li}(x) - \theta_1 \operatorname{Li}(x^{\beta_1}))(1 + O_A((\log x)^{-A})). \tag{1-10}$$

In order to state the main result from which Theorem 1.1 follows, we introduce some additional notation. Let $H \subseteq G$ be an abelian subgroup of G such that $H \cap C$ is nonempty, and let $K = L^H$ be the fixed field of H. The characters χ in the dual group $\hat{H}$ are Hecke characters; we write the conductor of χ as $\mathfrak{f}_\chi$. Define

$$\mathcal{Q} = \mathcal{Q}(L/K) = \max_{\chi \in \hat{H}} \mathrm{N}_{K/\mathbb{Q}} \mathfrak{f}_\chi. \tag{1-11}$$

We write the L-function associated to such a Hecke character as $L(s, \chi, L/K)$. From work of Stark [1974], at most one real Hecke character $\chi_1 \in \hat{H}$ has an associated Hecke L-function $L(s, \chi_1, L/K)$ with a Landau–Siegel zero $\beta_1 = 1 - \lambda_1 / \log(D_K \mathcal{Q} n_K^{n_K})$, where $0 < \lambda_1 < \frac{1}{8}$.

Theorem 1.4. *Let L/F be a Galois extension of number fields with Galois group G, and let $C \subseteq G$ be a conjugacy class. Let $H \subseteq G$ be an abelian subgroup such that $C \cap H$ is nonempty, let K be the fixed field of H, and choose $g_C \in C \cap H$. If $x \geq (D_K \mathcal{Q} n_K^{n_K})^{c_2}$, then*

$$\pi_C(x) = \frac{|C|}{|G|}(\mathrm{Li}(x) - \theta_1 \mathrm{Li}(x^{\beta_1}))\left(1 + O\left(\exp\left[-\frac{c_3 \log x}{\log(D_K \mathcal{Q} n_K^{n_K})}\right] + \exp\left[-\frac{(c_3 \log x)^{1/2}}{n_K^{1/2}}\right]\right)\right),$$

where $\theta_1 = \chi_1(g_C)$ if β_1 exists and $\theta_1 = 0$ otherwise and $\mathcal{Q}$ is given by (1-11). The constants c_2 and c_3 are the same as in Theorem 1.1.

Remark 1.5. As a group-theoretic quantity, θ_1 depends on the choice of $g_C \in C \cap H$. However, if $\theta_1 \neq 0$, then the existence of β_1 implies that θ_1 is well defined.

1C. ***An application.*** While it is aesthetically appealing to be able to encapsulate the work in [Lagarias et al. 1979; Lagarias and Odlyzko 1977; Murty 1997; Thorner and Zaman 2017; 2018; Weiss 1983] with a single asymptotic equality, Theorem 1.4 can make progress in certain sieve-theoretic problems when one must compute the local densities. As an example, we prove a new result in the study of primes represented by binary quadratic forms. Let

$$f(u, v) = au^2 + buv + cv^2 \in \mathbb{Z}[u, v]$$

be a positive definite binary quadratic form of discriminant $D = b^2 - 4ac < 0$. We do not assume that D is fundamental. The group $\mathrm{SL}_2(\mathbb{Z})$ naturally acts on such forms by $(T \cdot f)(\mathbf{x}) = f(T\mathbf{x})$ for $T \in \mathrm{SL}_2(\mathbb{Z})$. The class number $h(D)$ is the number of such forms up to SL_2-equivalence. If f is primitive (that is, $(a, b, c) = 1$) then it is a classical consequence of the Chebotarev density theorem and class field theory that

$$\frac{1}{|\mathrm{stab}(f)|} \sum_{\substack{u,v \in \mathbb{Z} \\ au^2+buv+cv^2 \leq x}} \sum \mathbf{1}_{\mathbb{P}}(au^2 + buv + cv^2) \sim \frac{\mathrm{Li}(x)}{h(D)} \quad \text{as } x \to \infty, \tag{1-12}$$

where $\mathbf{1}_{\mathbb{P}}$ is the indicator function for the odd primes and

$$\mathrm{stab}(f) = \{T \in \mathrm{SL}_2(\mathbb{Z}) : T \cdot f = f\}.$$

Note $|\mathrm{stab}(f)| = 2$ unless $D = -3$ or -4 in which case it equals 6 and 4 respectively.

We consider the question of imposing restrictions on the integers u and v which comprise a solution to the equation $p = f(u, v)$. In the special case of $f(u, v) = u^2 + v^2$, Fouvry and Iwaniec [1997] proved that there are infinitely many primes p such that $p = u^2 + v^2$ and u is prime. Their proof, which relies on sieve methods, enables them to asymptotically count such primes.

One might ask whether their methods extend to all positive definite primitive $f(u, v)$ with strong uniformity in the discriminant D.[2] The answer is not clear to the authors. Nevertheless, Theorem 1.4 enables us to study the distribution of primes $p = f(u, v)$ with some control over the divisors of u and v while maintaining strong uniformity in D. We prove the following result in Section 7.

Theorem 1.6. *Let $D \le -3$ be an integer and let $f(u, v) = au^2 + buv + cv^2$ be a positive definite primitive integral binary quadratic form with discriminant $D = b^2 - 4ac$. Let P be any integer dividing the product of primes $p \le z$. For all $A \ge 1$, there exists a sufficiently small constant $\eta = \eta(A) > 0$ such that if $3 \le z \le x^{\eta/\log\log x}$ and $3 \le |D| \le x^{\eta/\log\log z}$, then*

$$\frac{1}{|\mathrm{stab}(f)|} \sum_{\substack{u,v\in\mathbb{Z} \\ au^2+buv+cv^2\le x \\ (uv,P)=1}}\sum \mathbf{1}_{\mathbb{P}}(au^2+buv+cv^2) = \delta_f(P)\frac{\mathrm{Li}(x)-\mathrm{Li}(x^{\beta_1})}{h(D)}\{1+O_A((\log z)^{-A})\}. \quad (1\text{-}13)$$

Here, β_1 is a real simple zero of the Dedekind zeta function $\zeta_{\mathbb{Q}(\sqrt{D})}(s)$ (if it exists),

$$\delta_f(P) = \prod_{p\mid P}\left(1 - \frac{2 - \mathbf{1}_{p\mid a}(p) - \mathbf{1}_{p\mid c}(p)}{p - \left(\frac{D}{p}\right)}\right), \quad (1\text{-}14)$$

$\left(\frac{D}{p}\right)$ is the Legendre symbol for $p \ne 2$, $\left(\frac{D}{2}\right)$ is defined by (7-6), and the term $\mathrm{Li}(x^{\beta_1})$ is omitted if β_1 does not exist.

Remark 1.7. The constant $\delta_f(P)$ is always nonnegative. It is possible that $\delta_f(P) = 0$ due to the local factor at $p = 2$ in the product but this occurs precisely when the form $f(u, v)$ does not represent any odd primes. Since $\mathbf{1}_{\mathbb{P}}$ is the indicator function for the *odd* primes, (1-13) trivially holds in this case. The details of this casework are verified in Section 7A1.

While it is natural to think of P as equal to the product of primes up to z, we immediately obtain from Theorem 1.6 the following corollary when P is a *fixed* divisor of the product of primes up to z and $z \to \infty$ arbitrarily slowly.

Corollary 1.8. *Keep the assumptions of Theorem 1.6. If the integer $P \ge 1$ is fixed, then*

$$\frac{1}{|\mathrm{stab}(f)|} \sum_{\substack{u,v\in\mathbb{Z} \\ au^2+buv+cv^2\le x \\ (uv,P)=1}}\sum \mathbf{1}_{\mathbb{P}}(au^2+buv+cv^2) \sim \delta_f(P)\frac{\mathrm{Li}(x)-\mathrm{Li}(x^{\beta_1})}{h(D)} \quad \text{as } \frac{\log x}{\log|D|} \to \infty.$$

[2] Added in proof, 17 June 2019: Lam, Schindler, and Xiao [2018] recently extended Fouvry and Iwaniec's result to all positive-definite primitive binary quadratic forms. However, their error terms do not possess uniformity in the discriminant.

In particular, there exists a prime $p \le |D|^{\alpha}$ *and* $u, v \in \mathbb{Z}$ *such that* $p = f(u, v)$, $p \nmid D$, *and* $(uv, P) = 1$, *where* $\alpha = \alpha(P) > 0$ *is a sufficiently large constant depending only on* P.

In order to prove Theorem 1.6 with strong uniformity in z and $|D|$, one needs asymptotic control over sums like (1-12) (see (7-4) below) when x is as small as a polynomial in the discriminant, regardless of whether $\zeta_{\mathbb{Q}(\sqrt{D})}(s)$ has a Landau–Siegel zero. This is precisely what Theorem 1.4 provides. For comparison, a slightly stronger version of (1-2) that follows from [Murty 1997] along with the effective bound $(1-\beta_1)^{-1} \ll |D|^{1/2} \log |D|$ can produce (1-13) with the inferior ranges

$$3 \le |D| \ll (\log x)^2/(\log\log x)^2 \quad \text{and} \quad 3 \le z \le \exp(c\sqrt{\log x})$$

where $c > 0$ is an absolute constant and $\varepsilon > 0$. As one can plainly see, Theorem 1.4 yields substantial gains over earlier versions of the Chebotarev density theorem. See Remark 7.3 for further discussion.

1D. ***Overview of the methods.*** We now give an overview of how the proof of Theorem 1.4 differs from the proofs in [Lagarias et al. 1979; Lagarias and Odlyzko 1977; Murty 1997; Thorner and Zaman 2017; 2018; Weiss 1983]. For convenience, we refer to

$$\frac{|C|}{|G|}(\mathrm{Li}(x) - \theta_1 \mathrm{Li}(x^{\beta_1}))$$

as the "main term" in Theorem 1.4 and all other terms as the "error term".

The key difference between the proof of (1-2) and the proof of Theorem 1.4 lies in the study of the nontrivial low-lying zeros of $\zeta_L(s)$. The standard zero-free region for $\zeta_L(s)$ indicates that the low-lying zeros of $\zeta_L(s)$ lie further away from the edge of the critical strip $\{s \in \mathbb{C} : 0 < \mathrm{Re}(s) < 1\}$ than zeros of large height. However, the treatments in [Lagarias and Odlyzko 1977; Murty 1997] handle the contribution from the all of the nontrivial zeros by assuming that the low-lying zeros (other than β_1, if it exists) lie just as close to the edge of the critical strip as zeros of large height. This unduly inflates the contribution from the low-lying zeros, leading to the poor field uniformity in (1-2) along with the poor dependence on the Landau–Siegel zero β_1 if it exists. Consequently, both the range of x and the quality the error term in (1-2) directly depend on the quality of zero-free region available for $\zeta_L(s)$.

In order to efficiently handle the contribution to $\pi_C(x)$ which arises from the low-lying zeros of $\zeta_L(s)$, we factor $\zeta_L(s)$ as a product of Hecke L-functions associated to the Hecke characters of the abelian extension L/K and apply a log-free zero density estimate and the zero repulsion phenomenon for these L-functions. As in Linnik's work on arithmetic progressions, one typically uses these tools to establish upper and lower bounds of $\pi_C(x)$ when x is small instead of asymptotic equalities [Thorner and Zaman 2017; 2018; Weiss 1983]. In order to facilitate the analysis involving the log-free zero density estimate, we weigh the contribution of each prime ideal counted by $\pi_C(x)$ with a weight whose Mellin transform has carefully chosen decay properties (Lemma 2.2). Similar variations are a critical component in the proofs of (1-6) in [Thorner and Zaman 2017; 2018; Weiss 1983].

By using a log-free zero density estimate and the zero repulsion phenomenon, we ensure that the main term in Theorem 1.4 *always* dominates the error term in Theorem 1.4 when x is at least a polynomial in

$D_K \mathcal{Q} n_K^{n_K}$, regardless of whether β_1 exists. As one can see from the ensuing analysis, the quality of the zero-free region dictates the quality of the error term but has no direct impact on the valid range of x. This "decoupling" feature contrasts with the proof of (1-2), where the quality of the zero-free region simultaneously determines both the quality of the error term and the range of x in which the main term dominates.

After we "decouple" the range of x from the influence of the zero-free region, we are finally prepared to separate the contribution of the low-lying zeros from the contribution of the zeros with large height using a dyadic decomposition. This leads to savings over (1-2) only because we have already ensured via the log-free zero density estimate and zero repulsion that the main term in Theorem 1.4 dominate the error term regardless of whether β_1 exists. An additional benefit of this argument is an expression for the error term in Theorem 1.4 as a straightforward single-variable optimization problem involving x and the zero-free region (Lemma 4.5 and (4-13)). This simplification allows us to easily determine the error term with complete uniformity in D_K, $[K : \mathbb{Q}]$, $\mathcal{Q}$, and x (Lemma 4.6).

The fact that Theorem 1.4 holds for *all* Galois extensions L/F is a fairly subtle matter. In the case where $F = \mathbb{Q}$ and $L/\mathbb{Q}$ is a cyclotomic extension, the Chebotarev density theorem reduces to the prime number theorem for arithmetic progressions. Stark's bound for β_1 (Theorem 3.3, a refinement of (1-4)) recovers a lower bound for $1 - \beta_1$ which is commensurate with the lower bound for $1 - \beta_1$ that follows from Dirichlet's analytic class number formula for cyclotomic extensions; this suffices for our purposes. In the cyclotomic setting, our proofs only need to quantify the zero repulsion from a Landau–Siegel zero with a strong zero-free region for low-lying zeros (Theorem A.1 with $t \leq 4$). However, if L/F is a Galois extension where the root discriminant of L is especially small, which can happen in infinite class field towers, then Stark's lower bound for $1 - \beta_1$ is quite small. In this case, the approach which worked well for cyclotomic extensions of $\mathbb{Q}$ appears insufficient to prove Theorem 1.1 for all x in our claimed range.

To address this problem, we use a log-free zero density estimate for Hecke L-functions that naturally incorporates the zero repulsion phenomenon. Roughly speaking, when β_1 is especially close to 1, the quality of the log-free zero density estimate improves by a factor of $1 - \beta_1$; this is stronger than the classical formulation of the zero repulsion phenomenon. Therefore, if $1 - \beta_1$ happens to be as small as Stark's lower bound allows, the quality of the log-free zero density estimate increases dramatically. This offsets the adverse effect of β_1 in the small root discriminant case. The idea of incorporating the zero repulsion phenomenon directly into the log-free zero density estimate goes back to Bombieri [1987] in the case of Dirichlet characters. For Hecke L-functions over number fields, this was first proved by Weiss (see Theorem 3.2 below). The details of this obstacle and why we genuinely need the particular log-free zero density estimate in Theorem 3.2 are contained in the Appendix, especially Remark A.3.

2. Setup and notation

Throughout the paper, let $c_1, c_2, c_3, \ldots$ be a sequence of absolute, effective, and positive constants. All implied constants in the inequalities $f \ll g$ and $f = O(g)$ are absolute and effective unless noted otherwise.

Recall F is a number field with ring of integers $\mathcal{O}_F$, absolute norm $\mathrm{N} = \mathrm{N}_{F/\mathbb{Q}}$, absolute discriminant $D_F = |\mathrm{disc}(F/\mathbb{Q})|$, and degree $n_F = [F : \mathbb{Q}]$. Integral ideals will be denoted by $\mathfrak{n}$ and prime ideals by $\mathfrak{p}$. Moreover, L/F is a Galois extension of number fields with Galois group $G = \mathrm{Gal}(L/F)$. For prime ideals $\mathfrak{p}$ of F unramified in L, the Artin symbol $\left[\frac{L/F}{\mathfrak{p}}\right]$ is the conjugacy class of Frobenius automorphisms of G associated to prime ideals $\mathfrak{P}$ of L lying above $\mathfrak{p}$.

2A. ***Prime counting functions.*** For a conjugacy class C of G and $x \geq 2$, let $\pi_C(x)$ be as in (1-1) and define

$$\psi_C(x) = \psi_C(x, L/F) = \frac{|C|}{|G|} \sum_{\psi} \overline{\psi}(C) \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} -\frac{L'}{L}(s, \psi, L/F) \frac{x^s}{s}\, ds, \tag{2-1}$$

where ψ runs over the irreducible Artin characters of $G = \mathrm{Gal}(L/F)$ and $L(s, \psi, L/F)$ is the Artin L-function of ψ. It follows from Mellin inversion [Lagarias et al. 1979, p.283] that

$$\psi_C(x) = \sum_{\mathrm{N}\mathfrak{n} \leq x} \Lambda_F(\mathfrak{n}) \mathbf{1}_C(\mathfrak{n}), \tag{2-2}$$

where

$$\Lambda_F(\mathfrak{n}) = \begin{cases} \log \mathrm{N}\mathfrak{p} & \text{if } \mathfrak{n} = \mathfrak{p}^j \text{ for some prime ideal } \mathfrak{p} \text{ and some integer } j \geq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{2-3}$$

Here, $0 \leq \mathbf{1}_C(\mathfrak{n}) \leq 1$ for all ideals $\mathfrak{n}$ and for prime ideals $\mathfrak{p}$ unramified in L and $j \geq 1$,

$$\mathbf{1}_C(\mathfrak{p}^j) = \begin{cases} 1 & \text{if } \left[\frac{L/F}{\mathfrak{p}}\right]^j \subseteq C, \\ 0 & \text{otherwise.} \end{cases} \tag{2-4}$$

The prime counting functions π_C and ψ_C are related via partial summation.

Lemma 2.1. *For $x \geq 2$,*

$$\pi_C(x) = \frac{\psi_C(x)}{\log x} + \int_{\sqrt{x}}^{x} \frac{\psi_C(t)}{t(\log t)^2}\, dt + O\left(\log D_L + \frac{n_F x^{1/2}}{\log x}\right).$$

Proof. Note the norm of the product of ramified prime ideals divides D_L and the number of prime ideals $\mathfrak{p}$ with norm equal to a given rational prime p is at most n_F. Thus,

$$\pi_C(x) = \sum_{\sqrt{x} < \mathrm{N}\mathfrak{p} \leq x} \mathbf{1}_C(\mathfrak{p}) + O\left(\frac{n_F x^{1/2}}{\log x} + \log D_L\right).$$

Define $\theta_C(x) = \sum_{\mathrm{N}\mathfrak{p} \leq x} \mathbf{1}_C(\mathfrak{p}) \log \mathrm{N}\mathfrak{p}$. It follows by partial summation as well as the previous observations that

$$\sum_{\sqrt{x} < \mathrm{N}\mathfrak{p} \leq x} \mathbf{1}_C(\mathfrak{p}) = \int_{\sqrt{x}}^{x} \frac{\theta_C(t)}{t(\log t)^2}\, dt + \frac{\theta_C(x)}{\log x}$$

Finally, one can verify that $|\theta_C(x) - \psi_C(x)| \ll n_F x^{1/2}$ by trivially estimating the number of prime ideal powers with norm at most x. Collecting all of these estimates yields the lemma. □

2B. *Choice of weight.* We now define a weight function which will be used to count prime ideals with norm between $\sqrt{x}$ and x.

Lemma 2.2. *Choose $x \geq 3$, $\varepsilon \in \left(0, \frac{1}{4}\right)$, and a positive integer $\ell \geq 1$. Define $A = \varepsilon/(2\ell \log x)$. There exists a continuous function $f(t) = f(t; x, \ell, \varepsilon)$ of a real variable t such that*:

(i) *$0 \leq f(t) \leq 1$ for all $t \in \mathbb{R}$, and $f(t) \equiv 1$ for $\frac{1}{2} \leq t \leq 1$.*

(ii) *The support of f is contained in the interval $\left[\frac{1}{2} - \frac{\varepsilon}{\log x}, 1 + \frac{\varepsilon}{\log x}\right]$.*

(iii) *Its Laplace transform $F(z) = \int_{\mathbb{R}} f(t)e^{-zt}\,dt$ is entire and is given by*

$$F(z) = e^{-(1+2\ell A)z} \cdot \left(\frac{1 - e^{(1/2+2\ell A)z}}{-z}\right)\left(\frac{1 - e^{2Az}}{-2Az}\right)^{\ell}. \tag{2-5}$$

(iv) *Let $s = \sigma + it$, $\sigma > 0$, $t \in \mathbb{R}$ and α be any real number satisfying $0 \leq \alpha \leq \ell$. Then*

$$|F(-s \log x)| \leq \frac{e^{\sigma\varepsilon} x^{\sigma}}{|s| \log x} \cdot (1 + x^{-\sigma/2}) \cdot \left(\frac{2\ell}{\varepsilon |s|}\right)^{\alpha}.$$

Moreover, $|F(-s \log x)| \leq e^{\sigma\varepsilon} x^{\sigma}$ and $\frac{1}{2} < F(0) < \frac{3}{4}$.

(v) *If $\frac{3}{4} < \sigma \leq 1$ and $x \geq 10$, then*

$$F(-\log x) \pm F(-\sigma \log x) = \left(\frac{x}{\log x} \pm \frac{x^{\sigma}}{\sigma \log x}\right)\{1 + O(\varepsilon)\} + O\left(\frac{x^{1/2}}{\log x}\right). \tag{2-6}$$

(vi) *Let $s = -\frac{1}{2} + it$ with $t \in \mathbb{R}$. Then*

$$|F(-s \log x)| \leq \frac{5x^{-1/4}}{\log x}\left(\frac{2\ell}{\varepsilon}\right)^{\ell} \left(\tfrac{1}{4} + t^2\right)^{-\ell/2}.$$

Proof. These are the contents of [Thorner and Zaman 2018, Lemma 2.2] except for (2-6), which we now prove. Let $\frac{3}{4} < \sigma \leq 1$. From (iii), we observe that

$$F(-\sigma \log x) = \frac{x^{\sigma}}{\sigma \log x}\left(\frac{e^{\varepsilon\sigma/\ell} - 1}{\varepsilon\sigma/\ell}\right)^{\ell} + O\left(\frac{x^{\sigma/2}}{\sigma \log x}\right). \tag{2-7}$$

The two cases of $F(-\log x) \pm F(-\sigma \log x)$ are proved differently; we first handle the $+$ case. It follows from (2-7) that

$$F(-\log x) + F(-\sigma \log x) = \frac{x}{\log x}\left(\frac{e^{\varepsilon/\ell} - 1}{\varepsilon/\ell}\right)^{\ell} + \frac{x^{\sigma}}{\sigma \log x}\left(\frac{e^{\varepsilon\sigma/\ell} - 1}{\varepsilon\sigma/\ell}\right)^{\ell} + O\left(\frac{x^{\sigma/2}}{\sigma \log x}\right).$$

The desired asymptotic for $F(-\log x) + F(-\sigma \log x)$ now follows from the Taylor series expansion

$$\left(\frac{e^{\varepsilon\sigma/\ell} - 1}{\varepsilon\sigma/\ell}\right)^{\ell} = 1 + O(\sigma\varepsilon),$$

which is valid for $0 < \sigma \leq 1$.

For the case of $F(-\log x) - F(-\sigma \log x)$, we first observe that (2-7) implies

$$(\log x)(F(-\log x) - F(-\sigma \log x)) = x\left(\frac{e^{\varepsilon/\ell} - 1}{\varepsilon/\ell}\right)^{\ell} - \frac{x^{\sigma}}{\sigma}\left(\frac{e^{\varepsilon\sigma/\ell} - 1}{\varepsilon\sigma/\ell}\right)^{\ell} + O(x^{1/2}). \tag{2-8}$$

Set

$$a = \frac{e^{\varepsilon/\ell} - 1}{\varepsilon/\ell}, \quad b = \frac{e^{\sigma\varepsilon/\ell} - 1}{\sigma\varepsilon/\ell}$$

so that $a > b \geq 1$. With this convention, we rewrite (2-8) as

$$(\log x)(F(-\log x) - F(-\sigma \log x)) = xa^{\ell} - \frac{x^{\sigma}}{\sigma} b^{\ell} + O(x^{1/2}). \tag{2-9}$$

Since $a > b \geq 1$, it follows from the bound $a^{\ell} - b^{\ell} \ll (a-b) \cdot \ell a^{\ell}$ that

$$xa^{\ell} - \frac{x^{\sigma}}{\sigma} b^{\ell} = \left(x - \frac{x^{\sigma}}{\sigma}\right) a^{\ell} + \frac{x^{\sigma}}{\sigma}(a^{\ell} - b^{\ell}) = \left(x - \frac{x^{\sigma}}{\sigma}\right) a^{\ell} + O\left(\frac{x^{\sigma}}{\sigma}(a-b)\ell a^{\ell}\right). \tag{2-10}$$

Since $\frac{3}{4} < \sigma \leq 1$, it follows from taking Taylor series expansions that $a^{\ell} = 1 + O(\varepsilon)$ and

$$a - b = \sum_{n=1}^{\infty} \frac{(1-\sigma^n)(\varepsilon/\ell)^n}{(n+1)!} \leq \sum_{n=1}^{\infty} \frac{n(1-\sigma)(\varepsilon/\ell)^n}{(n+1)!} \ll (1-\sigma)\frac{\varepsilon}{\ell}.$$

We apply these two Taylor expansions to (2-9) and (2-10) to obtain

$$(\log x)(F(-\log x) - F(-\sigma \log x)) = \left(x - \frac{x^{\sigma}}{\sigma}\right)(1 + O(\varepsilon)) + O\left(\frac{x^{\sigma}}{\sigma}(1-\sigma)\varepsilon\right) + O(x^{1/2}). \tag{2-11}$$

Finally, we observe that since $\sigma^{-2} x^{\sigma} \leq x$ for $\sigma > \frac{3}{4}$ and $x \geq 10$, we have that

$$\frac{x^{\sigma}}{\sigma}(1-\sigma) = \sigma\left(\frac{x^{\sigma}}{\sigma^2} - \frac{x^{\sigma}}{\sigma}\right) \leq \sigma\left(x - \frac{x^{\sigma}}{\sigma}\right).$$

We apply this observation to (2-11) to obtain

$$(\log x)(F(-\log x) - F(-\sigma \log x)) = \left(x - \frac{x^{\sigma}}{\sigma}\right)(1 + O(\varepsilon)) + O(x^{1/2}). \tag{2-12}$$

The desired result follows by dividing both sides of (2-12) by $\log x$. □

Let $\ell \geq 2$ be an integer, $x \geq 3$, and $\varepsilon \in \left(0, \frac{1}{4}\right)$. Define

$$\tilde{\psi}_C(x; f) = \tilde{\psi}_C(x, L/F; f) = \sum_{\mathfrak{n}} \Lambda_F(\mathfrak{n}) \mathbf{1}_C(\mathfrak{n}) f\left(\frac{\log \mathrm{N}\mathfrak{n}}{\log x}\right), \tag{2-13}$$

where $f = f(\cdot; x, \ell, \varepsilon)$ is given by Lemma 2.2. To understand ψ_C, it suffices to study the smooth variant $\tilde{\psi}_C$.

Lemma 2.3. *Let $\ell \geq 2$ be an integer, $x \geq 3$, and $\varepsilon \in \left(0, \frac{1}{4}\right)$. Then*

$$\psi_C(x) \leq \tilde{\psi}_C(x; f) + O(n_F x^{1/2}) \leq \psi_C(xe^{\varepsilon}).$$

Moreover, $\tilde{\psi}_C(x;\, f) = \psi_C(x) + O(n_F x^{1/2} + \varepsilon x)$.

Proof. By Lemma 2.2(i,ii) and definitions (2-2) and (2-13), we observe that

$$\sum_{\sqrt{x}\le \mathrm{N}\mathfrak{n}\le x} \Lambda_F(\mathfrak{n})\mathbf{1}_C(\mathfrak{n}) \le \tilde{\psi}_C(x;\, f) \le \psi_C(xe^{\varepsilon}).$$

The lemma now follows from (2-2) and the trivial estimate

$$\sum_{z\le \mathrm{N}\mathfrak{n}\le y} \Lambda_F(\mathfrak{n})\mathbf{1}_C(\mathfrak{n}) \le n_F \sum_{z\le n\le y} \Lambda_{\mathbb{Q}}(n) \ll n_F(y-z) \quad \text{for } 2\le z\le y. \qquad \square$$

2C. ***Dedekind zeta functions and Hecke L-functions.*** Now, assume L/K is an abelian extension of number fields. The Dedekind zeta function $\zeta_L(s)$ satisfies

$$\zeta_L(s) = \prod_{\chi} L(s,\, \chi,\, L/K), \tag{2-14}$$

where χ runs over the irreducible 1-dimensional Artin characters of $\mathrm{Gal}(L/K)$. By class field theory, each Artin L-function $L(s,\, \chi,\, L/K)$ is equal to a Hecke L-function $L(s,\, \chi,\, K)$, where (abusing notation) χ is a certain primitive Hecke character of K. For simplicity, write $L(s,\, \chi)$ in place of $L(s,\, \chi,\, L/K)$ or $L(s,\, \chi,\, K)$. Let the integral $\mathfrak{f}_\chi \subseteq \mathcal{O}_K$ denote the conductor associated to χ. For each χ, there exist nonnegative integers $a(\chi)$ and $b(\chi)$ satisfying $a(\chi) + b(\chi) = n_K$ such that if we define

$$\gamma(s,\, \chi) = \left[\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\right]^{a(\chi)}\left[\pi^{-(s+1)/2}\Gamma\left(\frac{s+1}{2}\right)\right]^{b(\chi)}$$

and

$$\delta(\chi) = \begin{cases} 1 & \text{if } \chi \text{ is trivial,} \\ 0 & \text{otherwise,} \end{cases}$$

then $\xi(s,\, \chi) := [s(1-s)]^{\delta(\chi)}(D_K \mathrm{N}\mathfrak{f}_\chi)^{s/2}\gamma(s,\, \chi)L(s,\, \chi)$ satisfies the functional equation

$$\xi(s,\, \chi) = \varepsilon(\chi)\xi(1-s,\, \overline{\chi}), \tag{2-15}$$

where $\varepsilon(\chi)$ is a complex number with unit modulus. Furthermore, $\xi(s,\, \chi)$ is an entire function of order 1 which does not vanish at $s = 0$. Note $L(s,\, \chi)$ has a simple pole at $s = 1$ if and only if χ is trivial. The nontrivial zeros ρ of $L(s,\, \chi)$ (which are the zeros of $\xi(s,\, \chi)$) satisfy $0 < \mathrm{Re}(\rho) < 1$, and the trivial zeros ω of $L(s,\, \chi)$ (which offset the poles of $\gamma(s,\, \chi)$) are at the nonnegative integers, each with order at most n_K.

The Dedekind zeta function $\zeta_L(s)$ possesses the same qualities (by considering the case $K = L$ and χ trivial). Namely, its completed L-function is

$$\xi_L(s) = [s(1-s)]D_L^{s/2}\left[\pi^{-s/2}\Gamma\left(\tfrac{s}{2}\right)\right]^{a_L}\left[(2\pi)^{-s}\Gamma\left(\tfrac{s+1}{2}\right)\right]^{b_L}\zeta_L(s) \tag{2-16}$$

for certain integers $a_L, b_L \geq 0$ satisfying $a_L + b_L = [L : \mathbb{Q}]$. The trivial zeros ω of $\zeta_L(s)$ are at the nonnegative integers with orders

$$\operatorname*{ord}_{s=\omega} \zeta_L(s) = \begin{cases} a_L & \omega = -2, -4, \dots, \\ b_L & \omega = -1, -3, \dots, \\ a_L - 1 & \omega = 0. \end{cases} \tag{2-17}$$

Moreover, the conductor-discriminant formula states that

$$\log D_L = \sum_{\chi} \log(D_K \mathrm{N}\mathfrak{f}_\chi). \tag{2-18}$$

From (1-11) with $\mathcal{Q} = \mathcal{Q}(L/K)$, it follows that

$$\log D_L \leq [L : K] \log(D_K \mathcal{Q}). \tag{2-19}$$

From this we deduce a somewhat crude bound for $\log D_L$ in terms of D_K, $\mathcal{Q}$, and n_K.

Lemma 2.4. *If L/K is abelian, then* $\log D_L \ll (D_K \mathcal{Q} n_K^{n_K})^2$.

Proof. By class field theory, L is contained in some ray class field L' of K whose Artin conductor has norm at most $\mathcal{Q}$. From [Weiss 1983, Lemma 1.16], it follows that $[L : K] \leq [L' : K] \leq D_K \mathcal{Q} e^{O(n_K)}$. The result now follows from (2-19). □

We also record a few standard estimates for Hecke L-functions.

Lemma 2.5 [Lagarias and Odlyzko 1977, Lemma 5.4]. *If $t \in \mathbb{R}$ and χ is a Hecke character of K, then*

$$\#\{\rho = \beta + i\gamma : L(\rho, \chi) = 0, 0 < \beta < 1, |\gamma - t| \leq 1\} \ll \log(D_K \mathrm{N}\mathfrak{f}_\chi) + n_K \log(|t| + 3),$$

where the zeros ρ are counted with multiplicity.

Lemma 2.6 [Lagarias and Odlyzko 1977, Lemma 5.6]. *Let χ be a Hecke character of K. Then*

$$-\frac{L'}{L}(s, \chi) \ll \log(D_K \mathrm{N}\mathfrak{f}_\chi) + n_K \log(|\mathrm{Im}(s)| + 3)$$

uniformly for $\mathrm{Re}(s) = -\frac{1}{2}$.

3. The distribution of zeros

For Sections 3 and 4, we will assume that the extension L/K is abelian. For notational simplicity, define

$$Q = Q(L/K) := D_K \mathcal{Q} n_K^{n_K}, \tag{3-1}$$

where $\mathcal{Q} = \mathcal{Q}(L/K)$ is given by (1-11). Any sum $\sum_\chi$ or product $\prod_\chi$ is over the primitive Hecke characters χ associated with L/K per the factorization in (2-14). Here we list three key results regarding the distribution of zeros of Hecke L-functions.

Theorem 3.1 (zero-free region). *There exists $c_4 > 0$ such that the Dedekind zeta function*

$$\zeta_L(s) = \prod_{\chi} L(s, \chi, L/K)$$

has at most one zero in the region $\mathrm{Re}(s) > 1 - \Delta(|\mathrm{Im}(s)| + 3)$, *where the function* Δ *satisfies*

$$\Delta(t) \geq \frac{c_4}{\log(Qt^{n_K})} \quad \text{for } t \geq 3. \tag{3-2}$$

If such an exceptional zero β_1 exists then it is real, simple, and attached to the L-function of a real Hecke character χ_1.

Proof. This is well known; see, for example, [Weiss 1983, Theorem 1.9]. □

We also refer to the exceptional zero β_1 as a Landau–Siegel zero. Now, for $0 \leq \sigma \leq 1$, $T \geq 1$ and any Hecke character χ, define

$$N(\sigma, T, \chi) = \#\{\rho = \beta + i\gamma : L(\rho, \chi) = 0, \sigma < \beta < 1, |\gamma| \leq T\}, \tag{3-3}$$

where the zeros ρ are counted with multiplicity.

Theorem 3.2 (log-free zero density estimate). *There exists an integer $c_5 \geq 1$ such that*

$$\sum_{\chi} N(\sigma, T, \chi) \ll B_1 (QT^{n_K})^{c_5(1-\sigma)} \tag{3-4}$$

uniformly for any $0 < \sigma < 1$ and $T \geq 1$, where

$$B_1 = B_1(T) = \min\{1, (1 - \beta_1)\log(QT^{n_K})\}. \tag{3-5}$$

Proof. Let $\varepsilon_0 > 0$ be a sufficiently small absolute and effective constant. It follows from [Thorner and Zaman 2017, Theorem 3.2] or its variant [Thorner and Zaman 2018, Theorem 4.5] that if $1 - \varepsilon_0 < \sigma < 1$ and $T \geq 1$, then

$$\sum_{\chi} N(\sigma, T, \chi) \ll (QT^{n_K})^{c_5(1-\sigma)}$$

regardless of whether β_1 exists. Weiss [1983, Theorem 4.3] proved that if β_1 exists, then for $1 - \varepsilon_0 < \sigma < 1$ and $T \geq 1$,

$$\sum_{\chi} N(\sigma, T, \chi) \ll (1 - \beta_1)\log(QT^{n_K})(QT^{n_K})^{c_5(1-\sigma)}.$$

Thus for $T \geq 1$, (3-4) holds with B_1 given by (3-5) in the range $1 - \varepsilon_0 < \sigma < 1$. By enlarging c_5 if necessary and using Stark's bound from Theorem 3.3, one can extend (3-4) to the remaining interval $0 < \sigma < 1 - \varepsilon_0$ by employing the trivial bound that follows from Lemma 2.5. □

Theorems 3.1 and 3.2 comprise the three principles used to prove Linnik's theorem on the least prime in an arithmetic progression: a zero-free region, a log-free zero density estimate, and a quantitative form of the zero repulsion phenomenon. Theorem 3.2 combines the second and third principles by following

the ideas of Bombieri [1987], and this is crucial to our arguments for certain choices of Galois extensions (see the Appendix).

We record an effective lower bound for the size of $1-\beta_1$ which follows from [Stark 1974, Theorem 1', p.148].

Theorem 3.3 (Stark's bound). *Let $\beta_1 = 1 - \lambda_1/\log Q$ be a real zero of a real Hecke character χ of the abelian extension L/K. Then $\lambda_1 \gg Q^{-2}$.*

Proof. This follows readily from (1-4) for $1-\beta$ when β is the real zero of a Dedekind zeta function. If χ is trivial then consider the Dedekind zeta function $\zeta_K(s)$. If χ is quadratic then consider the Dedekind zeta function $\zeta_K(s)L(s, \chi, L/K)$ corresponding to the quadratic extension of K defined by χ. □

As we shall see, these three theorems yield a unified Chebotarev density theorem which produces an asymptotic count for primes even in the presence of a Landau–Siegel zero.

4. Weighted counts of primes in abelian extensions

4A. ***Main technical result.*** The proof of Theorem 1.4 rests on the analysis on the weighted prime counting function $\tilde{\psi}_C(x; f) = \tilde{\psi}_C(x, L/K; f)$ given by (2-13), where f is given by Lemma 2.2 and L/K is abelian. The goal of this section is to prove the following proposition.

Proposition 4.1. *Assume L/K is abelian with Galois group G. Let $C \subseteq G$ be a conjugacy class of G. Let $f = f(\cdot; x, \ell, \varepsilon)$ be defined as in Lemma 2.2 with*

$$\varepsilon = 8\ell x^{-1/8\ell}, \quad \ell = 4c_5 n_K. \tag{4-1}$$

If $2 \le Q \le x^{1/(36c_5)}$ and $\varepsilon < \frac{1}{4}$, then

$$\frac{|G|}{|C|}\tilde{\psi}_C(x; f) = \left(x - \chi_1(C)\frac{x^{\beta_1}}{\beta_1}\right)\left(1 + O\left(e^{-\frac{c_4}{2}\frac{\log x}{\log Q}} + e^{-\sqrt{c_4(\log x)/4n_K}}\right)\right). \tag{4-2}$$

Remark 4.2. The constants c_4 and c_5 are defined in Theorems 3.1 and 3.2 respectively.

While f and its parameters are chosen in Proposition 4.1, we will assume throughout this section that $\varepsilon \in \left(0, \frac{1}{4}\right)$ and $\ell \ge 2$ are arbitrary, unless otherwise specified. The arguments leading to Proposition 4.1 are divided into natural steps: shifting a contour, estimating the arising zeros with the log-free zero density estimate, and optimizing the error term with a classical zero-free region.

4B. ***Shifting the contour.***

Lemma 4.3. *If $x \ge 3$, then*

$$\frac{|G|}{|C|}\frac{\tilde{\psi}_C(x; f)}{\log x} = F(-\log x) - \chi_1(C)F(-\beta_1 \log x) - \sum_{\chi}\overline{\chi}(C)\sideset{}{^\star}\sum_{\rho_\chi} F(-\rho_\chi \log x) + O\left(\frac{(2\ell/\varepsilon)^\ell \log D_L}{x^{1/4}\log x} + \frac{n_L}{\log x}\right),$$

where the sum $\sum^\star$ *is over all nontrivial zeros* $\rho_\chi \neq \beta_1$ *of* $L(s,\chi)$*, counted with multiplicity. Here the term* $F(-\beta_1 \log x)$ *may be omitted if the exceptional zero* β_1 *does not exist.*

Proof. By (2-1), (2-13), Lemma 2.2 and a standard Mellin inversion calculation,

$$\frac{|G|}{|C|}\tilde{\psi}_C(x;f) = \sum_\chi \overline{\chi}(C) I_\chi, \quad \text{where } I_\chi = \frac{\log x}{2\pi i}\int_{2-i\infty}^{2+i\infty} -\frac{L'}{L}(s,\chi)F(-s\log x)\,ds. \tag{4-3}$$

For each Hecke character χ, shift the contour I_χ to the line $\mathrm{Re}(s) = -\frac{1}{2}$. Note F is entire by Lemma 2.2(iii), so we need only consider the zeros and poles of $L(s,\chi)$. We pick up the simple pole at $s=1$ of $L(s,\chi)$ when χ is trivial and the trivial zero at $s=0$ of $L(s,\chi)$ of order at most n_K. Moreover, we also pick up all of the nontrivial zeros ρ_χ of $L(s,\chi)$. For the remaining contour along $\mathrm{Re}(s) = -\frac{1}{2}$, we apply Lemma 2.6, Minkowski's estimate $n_K \ll \log D_K$, and Lemma 2.2(vi) to deduce that

$$-\frac{\log x}{2\pi i}\int_{-1/2-i\infty}^{-1/2+i\infty} \frac{L'}{L}(s,\chi,L/K)F(-s\log x)\,ds \ll \frac{(2\ell/\varepsilon)^\ell \log(D_K \mathrm{N}\mathfrak{f}_\chi)}{x^{1/4}}.$$

Combining all of these observations yields

$$(\log x)^{-1} I_\chi = \delta(\chi)F(-\log x) - \sum_{\rho_\chi} F(-\rho_\chi \log x) + O\bigg(F(0)n_K + \frac{(2\ell/\varepsilon)^\ell \log(D_K \mathrm{N}\mathfrak{f}_\chi)}{x^{1/4}\log x}\bigg). \tag{4-4}$$

Here, ρ_χ runs over all nontrivial zeros of $L(s,\chi)$, including β_1 if it exists. Substituting (4-4) into (4-3) and dividing through by $\log x$, we obtain the desired result but with an error term of

$$O\bigg(\frac{|F(0)|n_K}{\log x}\sum_\chi |\overline{\chi}(C)| + \frac{(2\ell/\varepsilon)^\ell}{x^{1/4}\log x}\sum_\chi |\overline{\chi}(C)|\log(D_K \mathrm{N}\mathfrak{f}_\chi)\bigg).$$

As L/K is abelian, the characters χ are 1-dimensional so $|\overline{\chi}(C)| = 1$. Thus, applying the conductor-discriminant formula (2-18), the observation $n_K \sum_\chi 1 = [L:K]n_K = n_L$, and Lemma 2.2(iv), we obtain the desired error term. □

4C. ***Estimating the zeros.*** Now we estimate the sum over nontrivial zeros ρ in Lemma 4.3, beginning with those ρ of small modulus.

Lemma 4.4. *If* $x \geq 3$*, then*

$$\sum_\chi \sum_{\substack{\rho_\chi \\ |\rho_\chi| \leq 1/4}} |F(-\rho_\chi \log x)| \ll x^{1/4}\log D_L.$$

Proof. From Lemmas 2.2(iv) and 2.5,

$$\sum_\chi \sum_{\substack{\rho_\chi \\ |\rho_\chi| \leq 1/4}} |F(-\rho_\chi \log x)| \ll \sum_\chi \sum_{\substack{\rho_\chi \\ |\rho_\chi| \leq 1/4}} x^{1/4} \ll x^{1/4}\sum_\chi (\log(D_K \mathrm{N}\mathfrak{f}_\chi) + n_K).$$

The result now follows from Minkowski's estimate $n_K \ll \log D_K$ and (2-18). □

Next, we use the log-free zero density estimate to analyze the remaining contribution.

Lemma 4.5. *Keep the assumptions and notation of Lemma 4.3. Select ε and ℓ as in (4-1) and assume $\varepsilon < \frac{1}{4}$. For $2 \le Q \le x^{1/(8c_5)}$,*

$$\log x \sum_{\chi} \sideset{}{^\star}\sum_{\substack{\rho_\chi \\ |\rho_\chi| \ge 1/4}} |F(-\rho_\chi \log x)| \ll \nu_1 x e^{-\eta(x)/2}, \tag{4-5}$$

where

$$\nu_1 = \begin{cases} (1-\beta_1)\log Q & \text{if } \beta_1 \text{ exists}, \\ 1 & \text{otherwise}, \end{cases} \tag{4-6}$$

and η is given by

$$\eta(x) = \inf_{t \ge 3} [\Delta(t) \log x + \log t]. \tag{4-7}$$

Proof. We dyadically estimate the zeros. For $j \ge 1$, set $T_0 = 0$ and $T_j = 2^{j-1}$ for $j \ge 1$. Consider the sum

$$Z_j := \frac{\log x}{x} \sum_{\chi} \sum_{\substack{\rho_\chi = \beta_\chi + i\gamma_\chi \\ T_{j-1} \le |\gamma_\chi| \le T_j \\ |\rho_\chi| \ge 1/4}} |F(-\rho_\chi \log x)| \tag{4-8}$$

for $j \ge 1$. First, we estimate the contribution of each zero $\rho = \rho_\chi$ appearing in Z_j. Let $\rho = \beta + i\gamma$ satisfy $T_{j-1} \le |\gamma| \le T_j$ and $|\rho| \ge \frac{1}{4}$, so $|\rho| \ge \max\{T_{j-1}, \frac{1}{4}\} \ge T_j/4$ and $|\rho| \gg |\gamma| + 3$. Thus, Lemma 2.2(iv) with $\alpha = \ell(1-\beta)$ and our choice of ε imply that

$$\frac{\log x}{x} |F(-\rho \log x)| \ll \frac{x^{\beta-1}}{|\rho|} \left(\frac{2\ell}{\varepsilon |\rho|} \right)^{\ell(1-\beta)} \ll T_j^{-1/2} (|\gamma| + 3)^{-1/2} \cdot x^{-(1-\beta)/2} \cdot (x^{3/8} T_j^{\ell})^{-(1-\beta)}.$$

Since $Q \le x^{1/(8c_5)}$ and $\ell = 4c_5 n_K$, it follows that

$$\frac{\log x}{x} |F(-\rho \log x)| \ll T_j^{-1/2} \cdot (|\gamma| + 3)^{-1/2} x^{-(1-\beta)/2} (Q T_j^{n_K})^{-2c_5(1-\beta)}. \tag{4-9}$$

From Theorem 3.1 and (4-7), we deduce

$$(|\gamma| + 3)^{-1/2} x^{-(1-\beta)/2} \le (|\gamma| + 3)^{-1/2} x^{-\Delta(|\gamma|+3)/2} \le e^{-\eta(x)/2}.$$

Note the right-hand side is uniform over all nontrivial zeros ρ appearing in (4-5). Combining (4-9) and the above inequality with (4-8), we deduce that

$$Z_j \ll e^{-\eta(x)/2} T_j^{-1/2} \sum_{\chi} \sum_{\substack{\rho_\chi = \beta_\chi + i\gamma_\chi \\ T_{j-1} \le |\gamma_\chi| \le T_j}} (Q T_j^{n_K})^{-2c_5(1-\beta)}.$$

Defining $N(\sigma, T) = \sum_\chi N(\sigma, T, \chi)$, we use partial summation and Theorem 3.2 to see that

$$\begin{aligned} e^{\eta(x)/2} T_j^{1/2} Z_j &\ll \int_0^1 (QT_j^{n_K})^{-2c_5\alpha}\, dN(1-\alpha, T_j) \\ &\ll \left[(QT_j^{n_K})^{-2c_5} N(0, T_j) + \log(QT_j^{n_K}) \int_0^1 (QT_j^{n_K})^{-2c_5\alpha} N(1-\alpha, T_j)\, d\alpha \right] \\ &\ll B_1(T_j) \left[(QT_j^{n_K})^{-c_5} + \log(QT_j^{n_K}) \int_0^1 (QT_j^{n_K})^{-c_5\alpha}\, d\alpha \right] \\ &\ll B_1(T_j). \end{aligned}$$

If a Landau–Siegel zero does not exist then $B_1(T_j) = 1 = \nu_1$. Otherwise, if a Landau–Siegel zero exists then one can verify by (3-5) and a direct calculation that

$$B_1(T_j) T_j^{-1/4} \le (1-\beta_1) \cdot \sup_{t \ge 1} [\log(Qt^{n_K}) t^{-1/4}] \ll (1-\beta_1) \log Q = \nu_1.$$

The supremum occurs at $t \ll 1$ since $n_K \le \log Q$. Therefore,

$$\sum_{j \ge 1} Z_j \ll e^{-\eta(x)/2} \sum_{j \ge 1} \frac{B_1(T_j)}{T_j^{1/4}} \cdot \frac{1}{T_j^{1/4}} \ll \nu_1 e^{-\eta(x)/2} \sum_{j \ge 1} 2^{-j/4} \ll \nu_1 e^{-\eta(x)/2},$$

which yields the lemma by definition (4-8). □

4D. ***Error term with a classical zero-free region.*** The quality of the error term in Lemma 4.5, and hence in Proposition 4.1, is reduced to computing $\eta(x)$. This is a single-variable optimization problem.

Lemma 4.6. *Let η be defined by (4-7). If $x \ge 2$ then $e^{-\eta(x)} \le e^{-c_4 \frac{\log x}{\log Q}} + e^{-\sqrt{c_4 (\log x)/n_K}}$.*

Proof. It follows from Theorem 3.1, (4-7), and a change of variables $t = e^u$ that

$$\eta(x) \ge \inf_{u \ge 0} \phi_x(u) \quad \text{where } \phi_x(u) = \frac{c_4 \log x}{\log Q + n_K u} + u.$$

Note that $\phi_x(u) \to \infty$ as $u \to \infty$. By standard calculus arguments, one can verify that

$$\eta(x) \ge \begin{cases} \frac{c_4 \log x}{\log Q} & \text{if } 2 \le x \le \exp\left(\frac{(\log Q)^2}{c_4 n_K}\right), \\ \sqrt{\frac{c_4 \log x}{n_K}} & \text{if } x \ge \exp\left(\frac{(\log Q)^2}{c_4 n_K}\right). \end{cases} \tag{4-10}$$

This proves the lemma. □

4E. ***Proof of Proposition 4.1.*** Choose ε and ℓ as in (4-1) and continue to assume $\varepsilon < \frac{1}{4}$. By Lemmas 4.3–4.5, it follows for $2 \le Q \le x^{1/(36c_5)}$ that

$$\frac{|G|}{|C|} \tilde{\psi}_C(x; f) = (\log x)[F(-\log x) - \chi_1(C) F(-\beta_1 \log x)] + O(\nu_1 x e^{-\eta(x)/2} + \mathcal{E}(x)),$$

where $\mathcal{E}(x) = x^{-1/4} (2\ell/\varepsilon)^\ell \log D_L + n_L + x^{1/4} (\log x)(\log D_L)$. From (4-1) and Minkowski's estimate $n_L \ll \log D_L$, we see that $\mathcal{E}(x) \ll x^{1/4} (\log D_L)(\log x)$. From Lemma 2.4, $\log D_L \ll Q^2 \ll x^{1/10}$ since

$x \geq Q^{36c_5}$ and $c_5 \geq 1$. Hence, $\mathcal{E}(x) \ll x^{1/2}$. Using Lemma 2.2(v), (4-1), and noting $\beta_1 > \frac{1}{2}$, we deduce that

$$\frac{|G|}{|C|}\tilde{\psi}_C(x; f) = \left(x - \chi_1(C)\frac{x^{\beta_1}}{\beta_1}\right)(1 + O(n_K x^{-1/(32c_5 n_K)})) + O(\nu_1 x e^{-\eta(x)/2} + x^{1/2}) \tag{4-11}$$

for $2 \leq Q \leq x^{1/36c_5}$. Now, we claim that

$$x - \chi_1(C)\frac{x^{\beta_1}}{\beta_1} \gg \nu_1 x \gg x^{3/4}. \tag{4-12}$$

If β_1 does not exist, then $\nu_1 = 1$ and (4-12) is immediate. If β_1 exists and $(1-\beta_1)\log x < 1$, then since $x \geq Q^{36c_5}$ and $e^{-t} \geq 1 - t$ for $0 < t < 1$, we have

$$x - \chi_1(C)\frac{x^{\beta_1}}{\beta_1} \geq x\left(1 - \frac{x^{-(1-\beta_1)}}{\beta_1}\right) \geq (1-\beta_1)x\log\left(\frac{x}{e}\right) \gg (1-\beta_1)x\log Q = \nu_1 x.$$

Otherwise, β_1 exists and $(1-\beta_1)\log x \geq 1$ so $\beta_1 > \frac{1}{2}$ implies that

$$x - \chi_1(C)\frac{x^{\beta_1}}{\beta_1} \geq x\left(1 - \frac{x^{-(1-\beta_1)}}{\beta_1}\right) \geq x(1 - 2e^{-1}) \gg x \gg \nu_1 x,$$

Thus, the claim (4-12) follows upon noting that $\nu_1 \gg Q^{-2} \gg x^{-1/4}$ by Stark's bound Theorem 3.3 and the condition $x \geq Q^{36c_5}$. Combining (4-12) with (4-11), it follows that

$$\frac{|G|}{|C|}\tilde{\psi}_C(x; f) = \left(x - \chi_1(C)\frac{x^{\beta_1}}{\beta_1}\right)(1 + O(e^{-\eta(x)/2} + n_K x^{-1/(32c_5 n_K)})). \tag{4-13}$$

Finally, we apply Lemma 4.6 and note $n_K x^{-1/(32c_5 n_K)} \ll x^{-1/(300c_5 n_K)} \ll e^{-\sqrt{c_4(\log x)/(4n_K)}}$ for $x \geq Q^{36c_5}$. This completes the proof of Proposition 4.1. □

5. Proof of Theorems 1.1 and 1.4

5A. ***Abelian extensions.*** First, we prove Theorem 1.4 in the case of abelian extensions.

Theorem 5.1. *Assume L/K is abelian with Galois group G. Let $C \subseteq G$ be a conjugacy class. Define Q by (3-1). for $2 \leq Q \leq x^{1/c_2}$,*

$$\pi_C(x, L/K) = \frac{|C|}{|G|}(\mathrm{Li}(x) - \chi_1(C)\,\mathrm{Li}(x^{\beta_1}))(1 + O(e^{-\frac{c_4}{4}\frac{\log x}{\log Q}} + e^{-\sqrt{c_4(\log x)/8n_K}})). \tag{5-1}$$

Here β_1 is a putative exceptional zero with associated real Hecke character χ_1 of L/K.

Proof. Write $g(x) = x - \chi_1(C)x^{\beta_1}/\beta_1$. Select ε as in (4-1). Note the assumption $2 \leq Q \leq x^{1/c_2}$ guarantees $\varepsilon < \frac{1}{4}$ provided c_2 is sufficiently large. From Proposition 4.1 and Lemma 2.3, it follows that

$$\psi_C(x) \leq \frac{|C|}{|G|}g(x)(1 + O(e^{-\frac{c_4}{2}\frac{\log x}{\log Q}} + e^{-\sqrt{c_4(\log x)/4n_K}})) \quad \text{for } x \geq Q^{36c_5}. \tag{5-2}$$

On the other hand, writing $y = xe^{\varepsilon}$, Proposition 4.1 and Lemma 2.3 also imply

$$\psi_C(y) \geq \frac{|C|}{|G|} g(ye^{-\varepsilon})(1 + O(e^{-\frac{c_4}{2}\frac{\log y}{\log Q}} + e^{-\sqrt{c_4(\log y)/4n_K}}))$$

for $y \geq 2Q^{36c_5}$. By (4-12) and elementary arguments,

$$|g(ye^{-\varepsilon}) - g(y)e^{-\varepsilon}| \leq \frac{y^{\beta_1}}{\beta_1}(e^{-\varepsilon\beta_1} - e^{-\varepsilon}) \ll y\varepsilon(1-\beta_1) \ll \varepsilon g(y).$$

In particular, $g(ye^{-\varepsilon}) = g(y)(1 + O(\varepsilon))$. From our choice of ε in (4-1) and the condition $y \geq 2Q^{36c_5}$, one can see that $\varepsilon \ll n_K y^{-1/32c_5 n_K} \ll y^{-1/300c_5 n_K} \ll e^{-\sqrt{c_4(\log y)/4n_K}}$ so

$$\psi_C(y) \geq \frac{|C|}{|G|} g(y)(1 + O(e^{-\frac{c_4}{2}\frac{\log y}{\log Q}} + e^{-\sqrt{c_4(\log y)/4n_K}})) \quad \text{for } y \geq 2Q^{36c_5}.$$

Comparing the above with (5-2), we conclude that

$$\psi_C(x) = \frac{|C|}{|G|} g(x)(1 + O(e^{-\frac{c_4}{2}\frac{\log x}{\log Q}} + e^{-\sqrt{c_4(\log x)/4n_K}}))$$

for $x \geq Q^{40c_5}$. By partial summation (Lemma 2.1) and the observation that, for $\frac{1}{2} < \sigma \leq 1$,

$$\frac{x^\sigma}{\sigma \log x} + \int_{\sqrt{x}}^{x} \frac{t^{\sigma-1}}{\sigma(\log t)^2}\, dt = \int_{x^{\sigma/2}}^{x^\sigma} \frac{1}{\log t}\, dt = \mathrm{Li}(x^\sigma) + O\left(\frac{x^{1/2}}{\log x}\right), \tag{5-3}$$

it follows for $x \geq Q^{40c_5}$ that

$$\frac{|G|}{|C|}\pi_C(x) = (\mathrm{Li}(x) - \chi_1(C)\,\mathrm{Li}(x^{\beta_1}))(1 + O(e^{-\frac{c_4}{4}\frac{\log x}{\log Q}} + e^{-\sqrt{c_4(\log x)/8n_K}})) + \mathcal{E}_0(x),$$

where $\mathcal{E}_0(x) = \log D_L + n_K x^{1/2}/\log x$. By Lemma 2.4 and the observation that $n_K \ll \log x$, one can verify that $\mathcal{E}_0(x) \ll x^{1/2}$ for $x \geq Q^{40c_5}$. Hence, by (4-12), $\mathcal{E}_0(x)$ can be absorbed into the error term of Section 5A. As c_2 is sufficiently large, this completes the proof of Theorem 5.1. □

5B. ***Proof of Theorem 1.4.*** Now we finish the proof of Theorem 1.4 for any Galois extension L/F with any Galois group G. Using well-known arguments from class field theory, we reduce to the case of abelian extensions.

Lemma 5.2 (Murty, Murty and Saradha). *Let L/F be a Galois extension of number fields with Galois group G, and let $C \subseteq G$ be a conjugacy class. Let H be a subgroup of G such that $C \cap H$ is nonempty, and let K be the fixed field of L by H. Let $g \in C \cap H$, and let $C_H(g)$ denote the conjugacy class of H which contains g. If $x \geq 2$, then*

$$\left|\pi_C(x, L/F) - \frac{|C|}{|G|}\frac{|H|}{|C_H|}\pi_{C_H}(x, L/K)\right| \leq \frac{|C|}{|G|}\left(n_L x^{1/2} + \frac{2}{\log 2}\log D_L\right).$$

Proof. This is carried out during the proof of [Murty et al. 1988, Proposition 3.9]. □

Now, we apply Lemma 5.2 and subsequently Theorem 5.1 to $\pi_{C_H}(x, L/K)$ of the abelian extension L/K. Consequently, for $2 \le Q \le x^{1/c_2}$,

$$\frac{|G|}{|C|}\pi_C(x, L/F) = (\mathrm{Li}(x) - \chi_1(C)\,\mathrm{Li}(x^{\beta_1}))(1 + O(e^{-\frac{c_4}{4}\frac{\log x}{\log Q}} + e^{-\sqrt{c_4(\log x)/8n_K}})) + O(n_L x^{1/2} + \log D_L), \tag{5-4}$$

where $Q = Q(L/K)$ is defined by (3-1). Since we may assume $c_2 \ge 20$, it follows from Lemma 2.4 and Minkowski's estimate $n_L \ll \log D_L$ that $n_L x^{1/2} + \log D_L \ll x^{5/8}$ for $x \ge Q^{c_2}$. From (4-12), this estimate may be absorbed into the first error term of (5-4) since $x^{5/8-3/4} = x^{-1/8} \ll e^{-\sqrt{c_4(\log x)/8n_K}}$. This completes the proof of Theorem 1.4. □

Theorem 1.4 implies Theorem 1.1. Fix $g \in C$, let H in Theorem 1.4 be the cyclic group generated by g, and let K be the fixed field of H. Clearly $n_K \le n_L$, and the centered equation immediately below [Thorner and Zaman 2017, Equation 1-7] states $D_L^{1/|H|} \le D_K \mathcal{Q} \le D_L^{1/\varphi(|H|)}$. Theorem 1.1 now follows. □

6. Reduced composition of beta-sieves

Before proceeding to the proof of Theorem 1.6, we require some sieve machinery that follows from standard results. The setup and discussion here closely follow [Friedlander and Iwaniec 2010, Sections 5.9 and 6.3–6.5]. Let $\Lambda = (\lambda'_d)$ and $\Lambda' = (\lambda''_d)$ be beta sieve weights with the same sifting level z and same level of distribution R. That is, λ'_d and λ''_d satisfy

$$\lambda'_1 = \lambda''_1 = 1, \quad |\lambda'_d| \le 1, \quad |\lambda''_d| \le 1,$$

and are supported on squarefree numbers $d < R$ consisting of prime factors $\le z$. Let

$$s = \frac{\log R}{\log z}$$

be the sifting variable for both sieves. Let g' and g'' be multiplicative functions satisfying

$$0 \le g'(p) < 1, \quad 0 \le g''(p) < 1, \quad g'(p) + g''(p) < 1 \quad \text{for all primes } p. \tag{6-1}$$

Assume there exists $K > 1$ and $\kappa > 0$ such that, for all $2 \le w \le z$, we have

$$\begin{aligned} \prod_{w \le p < z} \left(1 - \frac{g'(p)}{1 - g'(p) - g''(p)}\right)^{-1} &\le K\left(\frac{\log z}{\log w}\right)^{\kappa}, \\ \prod_{w \le p < z} \left(1 - \frac{g''(p)}{1 - g'(p) - g''(p)}\right)^{-1} &\le K\left(\frac{\log z}{\log w}\right)^{\kappa}. \end{aligned} \tag{6-2}$$

The goal of this section is to estimate the reduced composition given by

$$G := \sum_{(d_1,d_2)=1} \sum \lambda'_{d_1} \lambda''_{d_2} g'(d_1) g''(d_2). \tag{6-3}$$

This expression can arise as the main term when two different sieves are applied to two different sequences that are linearly independent. Keeping this setup, the remainder of this section will be dedicated to the proof of the following theorem.

Theorem 6.1. *Assume $s > 9\kappa + 1 + 10\log K$, (6-1) holds, and (6-2) holds. If Λ and Λ' are upper bound beta sieves, then*

$$\sum_{(d_1,d_2)=1}\sum \lambda'_{d_1}\lambda''_{d_2}g'(d_1)g''(d_2) \le \prod_p (1-g'(p)-g''(p))\{1+e^{9-s}K^{10}\}^2.$$

If Λ is a lower bound beta sieve and Λ' is an upper bound beta sieve, then

$$\sum_{(d_1,d_2)=1}\sum \lambda'_{d_1}\lambda''_{d_2}g'(d_1)g''(d_2) \ge \prod_p (1-g'(p)-g''(p))\{1-e^{9-s}K^{10}\}.$$

Assume Λ is a lower bound beta sieve and Λ' is an upper bound beta sieve. The other case is entirely analogous. Thus, if $\theta' = 1 * \lambda'$ and $\theta'' = 1 * \lambda''$ then

$$\theta'_1 = \theta''_1 = 1 \quad \text{and} \quad \theta'_n \le 0 \le \theta''_n \quad \text{for } n \ge 2. \tag{6-4}$$

First, we apply [Friedlander and Iwaniec 2010, Lemma 5.6] to (6-3) and, keeping with their notation, we see that

$$G = \sum_{(b_1,b_2)=1}\sum \theta'_{b_1}\theta''_{b_2}g'(b_1)g''(b_2) \prod_{p\nmid b_1b_2}(1-g'(p)-g''(p)). \tag{6-5}$$

Define $\tilde{h}'$, $\tilde{h}''$ and $\tilde{g}'$, $\tilde{g}''$ to be multiplicative functions supported on squarefree numbers with

$$\tilde{h}'(p) = \frac{g'(p)}{1-g'(p)-g''(p)}, \quad \tilde{h}''(p) = \frac{g''(p)}{1-g'(p)-g''(p)}, \quad \tilde{g}'(p) = \frac{g'(p)}{1-g''(p)}, \quad \tilde{g}''(p) = \frac{g''(p)}{1-g'(p)}.$$

Thus we obtain the usual relations

$$\tilde{h}'(p) = \frac{\tilde{g}'(p)}{1-\tilde{g}'(p)} \quad \text{and} \quad \tilde{h}''(p) = \frac{\tilde{g}''(p)}{1-\tilde{g}''(p)}. \tag{6-6}$$

Note $\tilde{h}'(p), \tilde{h}''(p) \ge 0$ and $0 \le \tilde{g}'(p), \tilde{g}''(p) < 1$ by (6-1). Inserting these definitions into (6-5), we observe that

$$G = \left(\prod_p (1-g'(p)-g''(p))\right)\sum_{(b_1,b_2)=1}\sum \theta'_{b_1}\theta''_{b_2}\tilde{h}'(b_1)\tilde{h}''(b_2).$$

If $(b_1, b_2) \ne 1$ then the expression $\theta'_{b_1}\theta''_{b_2}\tilde{h}'(b_1)\tilde{h}''(b_2)$ is nonpositive by (6-4), so we may introduce all of these terms at the cost of a lower bound for G. Thus

$$G \ge \left(\prod_p (1-g'(p)-g''(p))\right)\left(\sum_{b_1}\theta'_{b_1}\tilde{h}'(b_1)\right)\left(\sum_{b_2}\theta''_{b_2}\tilde{h}''(b_2)\right). \tag{6-7}$$

The two sums in (6-7) are prepared for standard beta-sieve analysis.

Lemma 6.2. *If Λ' is a lower bound beta-sieve with $\beta = 9\kappa + 1$ and $s \geq \beta$ then*

$$\sum_b \theta'_b \tilde{h}'(b) \geq 1 - e^{9\kappa - s} K^{10}.$$

If Λ'' is an upper bound beta-sieve with $\beta = 9\kappa + 1$ and $s \geq \beta$ then

$$\sum_b \theta''_b \tilde{h}''(b) \leq 1 + e^{9\kappa - s} K^{10}.$$

Proof. This statement is essentially the fundamental lemma [Friedlander and Iwaniec 2010, Lemma 6.8]. To make the comparison clear with [loc. cit., Sections 6.3–6.5], one begins with [loc. cit., Equation 6.40] with their D, h, g replaced by our $R, \tilde{h}', \tilde{g}'$ (or $R, \tilde{h}'', \tilde{g}''$, respectively). Per the definition of $V(z)$ on [loc. cit., p. 56], it follows that

$$V(z) = \prod_{p<z} (1 - \tilde{g}'(p)).$$

Thus the assumption [loc. cit., Equation 5.38] corresponds to our (6-2). Next, one defines V_n just as in the equation at the top of [loc. cit., p. 63]; in doing so, we obtain [loc. cit., Equations 6.43 and 6.44]. Finally, using the same truncation parameters, the analysis of [loc. cit., Section 6.5] leading up to [loc. cit., Lemma 6.8] yields our result. □

Now, we apply Lemma 6.2 to the sum over b_1 (the lower bound sieve Λ') in (6-7). Note that the assumption $s > 9\kappa + 1 + 10 \log K$ implies that this sum over b_1 is positive. By the positivity of $\tilde{h}$ and (6-4), we may trivially estimate the sum over b_2 in (6-7) by

$$\sum_{b_2} \tilde{h}''(b_2)\, \theta''_{b_2} \geq \tilde{h}''(1)\, \theta''_1 = 1.$$

This proves the lower bound in Theorem 6.1. For the upper bound, we follow the same arguments and apply Lemma 6.2 twice (once to each sieve) in these final steps. □

7. Restricted primes represented by binary quadratic forms

We recall the setup in Section 1C. Let

$$f(u, v) = au^2 + buv + cv^2 \in \mathbb{Z}[u, v]$$

be a positive definite binary quadratic form of discriminant $D = b^2 - 4ac < 0$, not necessarily fundamental. The group $\mathrm{SL}_2(\mathbb{Z})$ naturally acts on such forms by $(T \cdot f)(\mathbf{x}) = f(T\mathbf{x})$ for $T \in \mathrm{SL}_2(\mathbb{Z})$. The class number $h(D)$ is the number of such forms up to SL_2-equivalence. We assume that f is primitive (that is, $(a, b, c) = 1$), and we define

$$\mathrm{stab}(f) = \{T \in \mathrm{SL}_2(\mathbb{Z}) : T \cdot f = f\}.$$

Note $|\mathrm{stab}(f)| = 2$ unless $D = -3$ or -4 in which case it equals 6 and 4 respectively.

7A. ***Proof of Theorem 1.6.*** Recall by assumption that $3 \le z \le x^{\eta/\log\log x}$ where $\eta = \eta(A) > 0$ is sufficiently small. Further, P is any integer dividing the product of primes $\le z$. Let $1 \le R \le x^{1/10}$ be a parameter yet to be specified. Let $\Lambda' = (\lambda'_d)$ and $\Lambda'' = (\lambda''_d)$ be sieve weights supported on squarefree integers $d \mid P$ satisfying

$$\lambda'_1 = \lambda''_1 = 1, \quad |\lambda'_d| \le 1, \quad |\lambda''_d| \le 1 \quad \text{for } d \ge 1, \quad \lambda'_d = \lambda''_d = 0 \quad \text{for } d \ge R. \tag{7-1}$$

We approximate the condition $(uv, P) = 1$ in (1-13) by considering the sieved sum

$$S(x) = S(x; \Lambda', \Lambda'') := \frac{1}{|\mathrm{stab}(f)|} \sum_{\substack{u,v\in\mathbb{Z}\\ f(u,v)\le x}} \sum \mathbf{1}_{\mathbb{P}}(f(u,v)) \Big(\sum_{d_1 \mid u} \lambda'_{d_1}\Big)\Big(\sum_{d_2 \mid v} \lambda''_{d_2}\Big). \tag{7-2}$$

By swapping the order of summation,

$$S(x) = \sum_{\substack{d_1,d_2\\ (d_1,d_2)=1}} \sum \lambda'_{d_1}\lambda''_{d_2} A_{d_1,d_2}(x), \tag{7-3}$$

where

$$A_{d_1,d_2}(x) = \frac{1}{|\mathrm{stab}(f)|} \sum_{\substack{f(u,v)\le x\\ d_1 \mid u, d_2 \mid v}} \sum \mathbf{1}_{\mathbb{P}}(f(u,v)). \tag{7-4}$$

Before computing the congruence sums $A_{d_1,d_2}(x)$, we introduce the local densities g' and g''. These are multiplicative functions defined by

$$g'(p) = \begin{cases} \big(p - \big(\frac{D}{p}\big)\big)^{-1} & \text{if } p \mid P \text{ and } p \nmid c,\\ 0 & \text{otherwise,}\end{cases} \quad \text{and} \quad g''(p) = \begin{cases} \big(p - \big(\frac{D}{p}\big)\big)^{-1} & \text{if } p \mid P \text{ and } p \nmid a,\\ 0 & \text{otherwise.}\end{cases} \tag{7-5}$$

Here $\big(\frac{D}{p}\big)$ is the usual Legendre symbol for $p \ne 2$ and

$$\left(\frac{D}{2}\right) = \begin{cases} 0 & \text{if } 2 \mid D,\\ 1 & \text{if } D \equiv 1 \pmod 8,\\ -1 & \text{if } D \equiv 5 \pmod 8.\end{cases} \tag{7-6}$$

Our main result on the Chebotarev density theorem, Theorem 1.4, yields the following key lemma whose proof is postponed to Section 7B.

Lemma 7.1. *Let $\gamma > 0$ and $\vartheta > 0$ be a sufficiently small absolute constants, and let d_1, d_2 be relatively prime integers dividing P. If $|d_1 d_2 D| \le x^{\gamma}$ then*

$$A_{d_1,d_2}(x) = g'(d_1)g''(d_2)\frac{\mathrm{Li}(x) - \mathrm{Li}(x^{\beta_1})}{h(D)}\{1 + O(\varepsilon_{d_1d_2}(x))\} + O(\sqrt{x}\log x), \tag{7-7}$$

where β_1 is a simple real zero of the Dedekind zeta function $\zeta_{\mathbb{Q}(\sqrt{D})}(s)$ (if it exists) and

$$\varepsilon_d(x) = \varepsilon_d(x; D) = \exp\left[-\vartheta\frac{\log x}{\log|dD|}\right] + \exp[-(\vartheta\log x)^{1/2}] \quad \textit{for } d \ge 1. \tag{7-8}$$

Remark 7.2. For the remainder of the proof of Theorem 1.6, the constant ϑ may be allowed to vary from line-to-line. This will occurs finitely many times, so this is no cause for concern.

Remark 7.3. For the sieve to succeed, one crucially requires an asymptotic equality for $A_{d_1,d_2}(x)$ as in (7-7) with small remainder terms. Proceeding via the Chebotarev density theorem, one might use a stronger version of (1-2) in [Murty 1997] to obtain the asymptotic

$$A_{d_1,d_2}(x) = \frac{g'(d_1)g''(d_2)}{h(D)}(\mathrm{Li}(x) + O(xe^{-c_1\sqrt{\log x}})), \quad \text{for } \log x \gg (\log|d_1d_2D|)^2 + \frac{1}{1-\beta_1}. \tag{7-9}$$

Currently, $(1-\beta_1)^{-1} \ll |D|^{1/2}\log|D|$ is the best unconditional effective bound for β_1. Thus x must be quite large with respect to $|D|$, d_1, and d_2; this adversely impacts the permissible ranges of $|D|$ and z in Theorem 1.6. To improve the range of x, one might instead appeal to variants of (1-6) found in [Thorner and Zaman 2017; 2018; Weiss 1983] but this only yields lower and upper bounds for $A_{d_1,d_2}(x)$, rendering the sieve powerless. Fortunately, Theorem 1.4 addresses all of these obstacles simultaneously. Regardless of whether β_1 exists, it maintains an asymptotic with an improved range of x that is polynomial in $|D|$, d_1, and d_2 while keeping satisfactory control on the error terms. This allows us to strengthen the uniformity of both z and $|D|$ in Theorem 1.6 beyond what earlier versions of the Chebotarev density theorem permit.

Now, set the level of distribution to be

$$R := z^{\frac{1}{\sqrt{\eta}}\log\log z}. \tag{7-10}$$

Recall the constant $\eta = \eta(A) > 0$ should be thought of as very small. Since $z \le x^{\eta/\log\log x}$ and $|D| \le x^{\eta/\log\log z}$ by assumption, we have that $R \le x^{1/10}$ and also $|d_1d_2D| \le x^{4\sqrt{\eta}}$ for any integers $d_1, d_2 < R$. Thus, by Lemma 7.1 and (7-1), it follows that

$$S(x) = (\mathcal{G} + O(\mathcal{R}))\frac{\mathrm{Li}(x) - \mathrm{Li}(x^{\beta_1})}{h(D)} + O(x^{3/4}), \tag{7-11}$$

where

$$\mathcal{G} = \sum_{\substack{d_1,d_2\\(d_1,d_2)=1}}\sum \lambda'_{d_1}\lambda''_{d_2}g'(d_1)g''(d_2), \quad \mathcal{R} = \sum_{\substack{d<R^2\\d\mid P}}\frac{\tau(d)}{\varphi(d)}\varepsilon_d(x).$$

Here τ is the divisor function and φ is Euler phi function. We obtained $\mathcal{R}$ by observing that

$$\sum_{\substack{d_1,d_2<R\\d_1,d_2\mid P\\(d_1,d_2)=1}}\sum .|\lambda'_{d_1}\lambda''_{d_2}|g'(d_1)g''(d_2)\varepsilon_{d_1d_2}(x) \le \sum_{\substack{d<R^2\\d\mid P}}\varepsilon_d(x)\sum_{\substack{d_1d_2=d\\(d_1,d_2)=1}} g'(d_1)g''(d_2) \le \mathcal{R}$$

since $g'(d_1)g''(d_2) \le 1/(\varphi(d_1)\varphi(d_2)) = 1/\varphi(d)$ from (7-5) and $\sum_{\substack{d_1d_2=d\\(d_1,d_2)=1}} 1 \le \tau(d)$. Now, we proceed to calculate the main term $\mathcal{G}$ and remainder terms $\mathcal{R}$.

7A1. *Main term $\mathcal{G}$.* For the main term $\mathcal{G}$, suppose we have chosen a lower bound sieve for the sum in (1-13); namely, suppose Λ is a lower bound beta sieve and Λ' is an upper bound beta sieve, each with level of distribution R. Our aim is to apply the Fundamental Lemma in the form of Theorem 6.1. One can see that g' and g'' each satisfy (6-2) with $\kappa = 1$ and K absolutely bounded. Moreover, our choice of sieve has a sufficiently large sifting variable $s = \log R/\log z \gg \eta^{-1}$ because $\eta > 0$ is sufficiently small.

We claim that we may assume

$$g'(p) + g''(p) < 1 \quad \text{for all primes } p$$

and hence g' and g'' also satisfy (6-1). From (7-5), the only concern occurs when $p = 2$ and $2 \mid P$. We prove the claim by checking cases and verifying that $g'(2) + g''(2) \geq 1$ only if Theorem 1.6 is trivially true.

- Suppose $D \equiv 5 \pmod 8$. By (7-5), we have $g'(2) + g''(2) \leq \frac{1}{3} + \frac{1}{3} < 1$.
- Suppose $D \equiv 1 \pmod 8$ so $b \equiv 1 \pmod 2$ and $ac \equiv 0 \pmod 2$. If $a+b+c \equiv 0 \pmod 2$ then the sum in (1-13) is necessarily empty because $\mathbf{1}_{\mathbb{P}}$ only detects odd primes. In this case, a and c have opposite parity so $g'(2) + g''(2) = 1$. Hence, $\delta_f(P) = 0$ by (1-14) and Theorem 1.6 is therefore trivially true. Otherwise, if $a + b + c \equiv 1 \pmod 2$ then a and c have the same parity. As $ac \equiv 0 \pmod 2$, it must be that $a \equiv c \equiv 0 \pmod 2$ implying $g'(2) + g''(2) = 0 < 1$ by definition (7-5).
- Suppose $2 \mid D$ so $b \equiv 0 \pmod 2$. If one of a or c is even then $g'(2) + g''(2) \leq \frac{1}{2} < 1$. Otherwise, if both a and c are odd then $g'(2) + g''(2) = 1$ and $a + b + c \equiv 0 \pmod 2$. This implies $\delta_f(P) = 0$ and also the sum in (1-13) is necessarily empty so Theorem 1.6 is trivially true.

This proves the claim. Therefore, by Theorem 6.1 and (7-10), it follows that

$$\mathcal{G} \geq \delta_f(P)\{1 + O_A((\log z)^{-A})\} \tag{7-12}$$

since $\eta = \eta(A)$ is sufficiently small. If Λ and Λ' are both upper bound beta sieves with level of distribution $x^{1/10}$ then one similarly obtains the reverse inequality.

7A2. *Remainder terms $\mathcal{R}$.* We estimate $\mathcal{R}$ dyadically. By the Cauchy–Schwarz inequality and standard estimates for τ and φ, we see for $0 \leq N \leq \lceil 2 \log R/\log z \rceil$ that

$$\begin{aligned}\sum_{\substack{z^N \leq d < z^{N+1} \\ d \mid P}} \frac{\tau(d)}{\varphi(d)} \varepsilon_d(x) &\ll \varepsilon_{z^{N+1}}(x) \Bigg(\sum_{\substack{z^N \leq d < z^{N+1} \\ p \mid d \Rightarrow p \leq z}} \frac{1}{d} \Bigg)^{1/2} \Bigg(\sum_{z^N \leq d < z^{N+1}} \frac{\tau(d)^2 d}{\varphi(d)^2} \Bigg)^{1/2} \\ &\ll \varepsilon_{z^{N+1}}(x)((N+1)\log z)^{3/2} \Bigg(\sum_{\substack{z^N \leq d < z^{N+1} \\ p \mid d \Rightarrow p \leq z}} \frac{1}{d} \Bigg)^{1/2}.\end{aligned}$$

By (7-10), one has that $R^{\eta'/\log\log R} \leq z \leq R$ where $\eta' > 0$ is sufficiently small depending only on η. In other words, $\log R/\log z \ll \log\log z$. Thus, we may apply Hildebrand's estimate [1986, Theorem 1] for

z-smooth numbers via partial summation to conclude from (7-8) that the above is

$$\ll (e^{-\vartheta \frac{\log x}{(N+1)\log z}} + e^{-\vartheta \frac{\log x}{\log |D|}} + e^{-\vartheta\sqrt{\log x}})\rho(N)(N+1)^2 \log^2 z,$$

where ρ is the Dickman–de Bruijn function. Recall we allow the constant $\vartheta > 0$ to change from line-to-line and be replaced by a smaller value if necessary. Summing this estimate over $0 \le N \le \lceil 2\log R/\log z\rceil$ and using the crude estimate $\rho(N) \ll N^{-N}$ for $N \ge 1$, we deduce that

$$\begin{aligned}\mathcal{R} &\ll (\max_{N\ge 1} e^{-\frac{c\log x}{N\log z}} N^{-N+2})\log^2 z + (e^{-\vartheta\frac{\log x}{\log z}} + e^{-\vartheta\frac{\log x}{\log |D|}} + e^{-\vartheta\sqrt{\log x}})\log^2 z \\ &\ll (e^{-\vartheta\sqrt{\frac{\log x \log\log x}{\log z}}} + e^{-\vartheta\frac{\log x}{\log z}} + e^{-\vartheta\frac{\log x}{\log |D|}} + e^{-\vartheta\sqrt{\log x}})\log^2 z.\end{aligned}$$

Since $|D| \le x^{\eta/\log\log z}$ and $z \le x^{\eta/\log\log x}$ with $\eta = \eta(A) > 0$ sufficiently small, we have that

$$\mathcal{R} \ll_A (\log z)^{-A}. \tag{7-13}$$

7A3. *Concluding the proof.* Inserting (7-12) and (7-13) into (7-11) along with the fact that $\delta_f(P) \gg (\log z)^{-2}$ from Mertens' estimate, we conclude that

$$\sum_{\substack{u,v\in\mathbb{Z}\\ au^2+buv+cv^2\le x\\ (uv,P)=1}} \frac{\mathbf{1}_{\mathbb{P}}(au^2+buv+cv^2)}{|\mathrm{stab}(f)|} \ge \delta_f(P)\frac{\mathrm{Li}(x)-\mathrm{Li}(x^{\beta_1})}{h(D)}\{1+O_A((\log z)^{-A})\} + O(x^{3/4}).$$

By using an upper bound sieve instead (as mentioned at the end of Section 7A1), one also obtains the reverse inequality. Thus, it remains to show the secondary error term $O(x^{3/4})$ may be absorbed into the primary error term. If $\delta_f(P) = 0$ then the arguments in Section 7A1 imply Theorem 1.6 trivially true so we may assume $\delta_f(P) > 0$. By the effective lower bound that $1-\beta_1 \gg_\varepsilon |D|^{-1/2-\varepsilon}$, the fact that $h(D) \ll_\varepsilon |D|^{1/2+\varepsilon}$, and the assumption that $|D| \le x^{\eta/\log\log z}$, we see

$$\frac{\mathrm{Li}(x)-\mathrm{Li}(x^{\beta_1})}{h(D)} \gg x^{4/5}.$$

As $\delta_f(P) \gg (\log z)^{-2}$, this implies the claim and hence proves Theorem 1.6. □

7B. ***Proof of Lemma 7.1.*** The pair (d_1, d_2) induces another form f_{d_1,d_2} given by

$$f_{d_1,d_2}(s,t) := f(d_1 s, d_2 t).$$

Note its discriminant is $D(d_1d_2)^2$. With this definition, it follows that

$$A_{d_1,d_2}(x) = \frac{1}{|\mathrm{stab}(f)|}\sum_{p\le x} \#\{(s,t)\in\mathbb{Z}^2 : p = f_{d_1,d_2}(s,t)\}. \tag{7-14}$$

Observe

$$A_{d_1,d_2}(x) \ll 1 \quad \text{if } (d_1,c)\ne 1 \text{ or } (d_2,a)\ne 1$$

since, in this case, f_{d_1,d_2} is not primitive and hence represents an absolutely bounded number of primes. This trivially establishes Lemma 7.1 in this case. To evaluate $A_{d_1,d_2}(x)$ for all other d_1 and d_2, we use class field theory.

Lemma 7.4. *Let $\mathcal{O}_K$ be the ring of integers of $K=\mathbb{Q}(\sqrt{D})$. For $d\ge 1$, let $\mathcal{O}_d$ be the order of discriminant $-Dd^2$ in K and let L_d be the ring class field of $\mathcal{O}_d$. If F is a primitive binary quadratic form of discriminant $-Dd^2$ then*

$$|\mathcal{O}_d^\times| = |\mathrm{stab}(F)|.$$

Moreover, if C_F is the conjugacy class corresponding to F in the Galois group of L_d/K then

$$\#\{(s,t)\in\mathbb{Z}^2 : p=F(s,t)\} = |\mathcal{O}_d^\times|\cdot\#\Big\{\mathfrak{p}\subseteq\mathcal{O}_K : \mathrm{N}\mathfrak{p}=p, \Big[\frac{L_d/K}{\mathfrak{p}}\Big]=C_F\Big\} \quad \text{for } p\nmid Dd.$$

Here $\left[\frac{L_d/K}{\mathfrak{p}}\right]$ is the Artin symbol of $\mathfrak{p}$ and $\mathrm{N}=\mathrm{N}_{K/\mathbb{Q}}$ is the absolute norm of $K/\mathbb{Q}$.

Proof. These are straightforward consequences of the theory for positive definite binary quadratic forms, so we only sketch the details. Standard references include for example [Cassels 1978; Cox 1989]. First, one can verify that $\mathcal{O}_d^\times=\{\pm1\}$ unless $\mathcal{O}_d$ is the ring of integers for $\mathbb{Q}(i)$ or $\mathbb{Q}(\sqrt{-3})$. Similarly, the SL_2-automorphism group of F is $\left\{\pm\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)\right\}$ unless F is properly equivalent to either x^2+y^2 or x^2+xy+y^2. These are respectively the unique reduced forms of discriminant -4 or -3. These remaining two cases can be checked by direct calculation.

The second claim follows from the first claim and the one-to-one correspondence between inequivalent representations of a prime p by F and degree 1 prime ideals $\mathfrak{p}\subseteq\mathcal{O}_K$ in the class C_F. For more details, see [Cox 1989, Theorem 7.7]. □

Now, assuming $(d_1,c)=(d_2,a)=1$, we return to computing $A_{d_1,d_2}(x)$. It follows that f_{d_1,d_2} is primitive so by Lemma 7.4 with $F=f_{d_1,d_2}$ and $d=d_1d_2$,we deduce that

$$A_{d_1,d_2}(x) = \frac{1}{|\mathrm{stab}(f)|}\sideset{}{^\dagger}\sum_{\substack{\mathrm{N}\mathfrak{p}\le x\\ \deg(\mathfrak{p})=1}} |\mathcal{O}_{d_1d_2}^\times| + O\Big(\sum_{p\mid Dd_1d_2} 1\Big), \tag{7-15}$$

where $\sum^\dagger$ runs over prime ideals $\mathfrak{p}$ in $\mathcal{O}_K$ unramified in $L_{d_1d_2}$ satisfying $[(L_{d_1d_2}/K)/\mathfrak{p}]=C_{f_{d_1,d_2}}$. Note, for the primes $p\mid Dd_1d_2$ in (7-15), we have used that each prime p is represented by f with absolutely bounded multiplicity. We may add the remaining degree 2 prime ideals $\mathfrak{p}$ to the †-marked sum with error at most $O(|\mathcal{O}_{d_1d_2}^\times|\sqrt{x}\log x)=O(\sqrt{x}\log x)$. Further, we have

$$\sum_{p\mid Dd_1d_2} 1 \ll \log|Dd_1d_2| \ll \log x$$

since $|d_1d_2D|\le x^\gamma$. Collecting these observations, it follows that

$$A_{d_1,d_2}(x) = \frac{|\mathcal{O}_{d_1d_2}^\times|}{|\mathrm{stab}(f)|}\sideset{}{^\dagger}\sum_{\mathrm{N}\mathfrak{p}\le x} 1 + O(\sqrt{x}\log x). \tag{7-16}$$

We invoke Theorem 1.4 to compute the sum in (7-16), thus

$$\sum_{\mathrm{N}\mathfrak{p}\le x}^{\dagger} 1 = \frac{\mathrm{Li}(x) - \theta_1 \mathrm{Li}(x^{\beta_1})}{h(D(d_1d_2)^2)}\{1 + O(\varepsilon_{d_1d_2}(x))\} \quad \text{for } |d_1d_2D| \le x^{\gamma}, \tag{7-17}$$

where $\varepsilon_{d_1d_2}(x)$ is defined by (7-8) and $\gamma > 0$ is fixed and sufficiently small. We make two simplifications for (7-17). First, we claim that $\theta_1 = 1$ if the exceptional zero β_1 exists. By a theorem of Heilbronn [1972] generalized by Stark [1974, Theorem 3], since β_1 is a real simple zero of $\zeta_{L_{d_1d_2}}(s)$ and $L_{d_1d_2}$ is Galois over $\mathbb{Q}$ with K being its only quadratic subfield, it follows that $\zeta_K(\beta_1) = 0$. Hence, the exceptional Hecke character χ_1 of K from Theorem 1.4 is trivial implying $\theta_1 = 1$. Second, we have for $d \ge 1$ that

$$h(Dd^2) = \frac{h(D)}{[\mathcal{O}^{\times} : \mathcal{O}_d^{\times}]} d \prod_{p \mid d} \left(1 - \left(\frac{D}{p}\right)\frac{1}{p}\right). \tag{7-18}$$

For a proof, see for example [Cox 1989, Theorem 7.4 and Corollary 7.28].

Finally, with these observations, Lemma 7.1 follows by inserting (7-17) and (7-18) into (7-16) and noting that $[\mathcal{O}_1^{\times} : \mathcal{O}_d^{\times}] \cdot |\mathcal{O}_d^{\times}| = |\mathcal{O}_1^{\times}| = |\mathrm{stab}(f)|$ from Lemma 7.4. □

Appendix: Error term with an exceptional zero

Theorem 3.2 states that if $T \ge 1$, then

$$\sum_{\chi} N(\sigma, T, \chi) \ll B_1 (QT^{n_K})^{c_5(1-\sigma)}, \quad B_1 = \min\{1, (1-\beta_1)\log(QT^{n_K})\}. \tag{A-1}$$

This clearly implies that regardless of whether β_1 exists, we have

$$\sum_{\chi} N(\sigma, T, \chi) \ll (QT^{n_K})^{c_5(1-\sigma)}. \tag{A-2}$$

If β_1 exists, Theorem 3.2 produces the following strong zero-free region:

Theorem A.1 (zero repulsion). *Suppose the exceptional zero β_1 of Theorem 3.1 exists. There exists $c_6 > 0$ such that if Δ is given in Theorem 3.1, then*

$$\Delta(t) \ge \min\left\{\frac{1}{2}, \frac{c_6 \log([(1-\beta_1)\log(Qt^{n_K})]^{-1})}{\log(Qt^{n_K})}\right\}.$$

Let $q \ge 1$ be an integer. In the context of arithmetic progressions, in which case $L = \mathbb{Q}(e^{2\pi i/q})$ and $F = K = \mathbb{Q}$, it is preferable to use (A-2) and Theorem A.1 instead of (A-1), as one can typically obtain numerically superior results with the former. However, in the context of arithmetic progressions, one has the benefit of working with characters of an extension which is abelian over $\mathbb{Q}$, in which case Theorem 3.3 gives an adequate upper bound for β_1 (should it exist). However, for abelian extensions L/K where the root discriminant of K is rather small, Theorem 3.3 gives an upper bound for β_1 which is not commensurate with the corresponding result for cyclotomic extensions of $\mathbb{Q}$. In fact, this weak upper bound leads us to actually require a version of the log-free zero density estimate that improves as

β_1 approaches 1 to handle the case when K has a small root discriminant. This is why we use (A-1) in our proofs instead of using (A-2) and Theorem A.1 separately.

For comparison with Lemma 4.6, we quantify the effect of (A-2) and Theorem A.1 on the error term in Lemma 4.5 and subsequently (4-13) in the proof of Proposition 4.1. Since the calculations are tedious, we omit the proof.

Lemma A.2. *Let η be defined by (4-7). Suppose the exceptional zero $\beta_1 = 1 - \lambda_1/\log Q$ of Theorem 3.1 exists. There exists absolute constants $c_7, c_8, c_9 > 0$ such that if $\lambda_1 \le c_7$ and $Q \le x^{1/c_9}$,*

$$e^{-\eta(x)} \ll x^{-1/2} + \lambda_1^{10}(e^{-\frac{c_6 \log x}{2\log Q}} + e^{-c_8\sqrt{(\log x)/n_K}}) \qquad \text{if } \lambda_1 \ge Q^{-20/n_K}, \tag{A-3}$$

$$e^{-\eta(x)} \ll x^{-1/2} + e^{-10\sqrt{\log(1/\lambda_1)}}(e^{-\frac{c_6 \log x}{2\log Q}} + e^{-c_8\sqrt{(\log x)/n_K}}) \qquad \text{if } \lambda_1 < Q^{-20/n_K}. \tag{A-4}$$

Remark A.3. Recall the definition of ν_1 in (4-6). From (4-11) and (4-12), one can see it is critical to prove an estimate at least as strong as

$$\nu_1 x e^{-\eta(x)} = o(\lambda_1 x). \tag{A-5}$$

Notice that the density estimate in (A-1) decays linearly with respect to $1 - \beta_1$ (that is, $\nu_1 = \lambda_1$), so we easily obtain (A-5). Suppose we instead use (A-2), which is tantamount to the trivial estimate $\nu_1 \le 1$ when β_1 exists. From (A-3), one obtains (A-5) when $\lambda_1 \ge Q^{-20/n_K}$. Otherwise, from (A-4), if $\lambda_1 < Q^{-20/n_K}$ then we can at best show $xe^{-\eta(x)} = o(e^{-10\sqrt{\log(1/\lambda_1)}}x)$. The situation $\lambda_1 < Q^{-20/n_K}$ is not uniformly excluded by Stark's bound (1-4). For example, when the root discriminant D_K^{1/n_K} is bounded and the extension L/K is unramified (that is, $\mathcal{Q} = 1$), then

$$Q^{100/n_K} = (D_K \mathcal{Q})^{100/n_K} n_K^{100} \ll n_K^{100}$$

and Stark's bound (1-4) implies $\lambda_1^{-1} \ll n_K^{n_K} \log D_K$ so it may very well be the case that $\lambda_1^{-1} \gg n_K^{100} \gg Q^{100/n_K}$. This situation with a bounded root discriminant is entirely possible as Minkowski's unconditional estimate $n_K \ll \log D_K$ is tight when varying over all number fields K. Infinite class field towers are well known sources of this scenario. Thus, we cannot see how to unconditionally obtain the desired linear decay demanded by (A-5) with only (A-2) and Theorem A.1.

Acknowledgements

We thank Kannan Soundararajan for helpful discussions and the anonymous referee for providing very thorough comments on our initial submission. Jesse Thorner is partially supported by a NSF Mathematical Sciences Postdoctoral Fellowship, and Asif Zaman is partially supported by a NSERC Postdoctoral Fellowship.

References

[Bombieri 1987] E. Bombieri, "Le grand crible dans la théorie analytique des nombres", *Astérisque* 18 (1987), 103. MR Zbl

[Cassels 1978] J. W. S. Cassels, *Rational quadratic forms*, London Mathematical Society Monographs **13**, Academic Press, London, 1978. MR Zbl

[Cox 1989] D. A. Cox, *Primes of the form $x^2 + ny^2$*, Wiley, New York, 1989. MR Zbl

[Fouvry and Iwaniec 1997] E. Fouvry and H. Iwaniec, "Gaussian primes", *Acta Arith.* **79**:3 (1997), 249–287. MR Zbl

[Friedlander and Iwaniec 2010] J. Friedlander and H. Iwaniec, *Opera de cribro*, American Mathematical Society Colloquium Publications **57**, American Mathematical Society, Providence, RI, 2010. MR Zbl

[Heilbronn 1972] H. Heilbronn, "On real simple zeros of Dedekind ζ-functions", pp. 108–110 in *Proceedings of the Number Theory Conference* (Univ. Colorado, Boulder, Colo., 1972), Univ. Colorado, Boulder, Colo., 1972. MR Zbl

[Hildebrand 1986] A. Hildebrand, "On the number of positive integers $\leq x$ and free of prime factors $> y$", *J. Number Theory* **22**:3 (1986), 289–307. MR Zbl

[Lagarias and Odlyzko 1977] J. C. Lagarias and A. M. Odlyzko, "Effective versions of the Chebotarev density theorem", pp. 409–464 in *Algebraic number fields: L-functions and Galois properties* (Durham, UK, 1975), edited by A. Fröhlich, Academic Press, London, 1977. MR Zbl

[Lagarias et al. 1979] 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 (1979), 271–296. MR Zbl

[Lam et al. 2018] C. Lam, D. Schindler, and S. Xiao, "On prime values of binary quadratic forms with a thin variable", preprint, 2018. arXiv

[Linnik 1944] U. V. Linnik, "On the least prime in an arithmetic progression, I: The basic theorem", *Rec. Math. [Mat. Sbornik] N.S.* **15(57)** (1944), 139–178. MR Zbl

[Murty 1997] V. K. Murty, "Modular forms and the Chebotarev density theorem, II", pp. 287–308 in *Analytic number theory* (Kyoto, 1996), edited by Y. Motohashi, London Math. Soc. Lecture Note Ser. **247**, Cambridge Univ. Press, 1997. MR Zbl

[Murty et al. 1988] M. R. Murty, V. K. Murty, and N. Saradha, "Modular forms and the Chebotarev density theorem", *Amer. J. Math.* **110**:2 (1988), 253–281. MR Zbl

[Stark 1974] H. M. Stark, "Some effective cases of the Brauer–Siegel theorem", *Invent. Math.* **23** (1974), 135–152. MR Zbl

[Thorner and Zaman 2017] J. Thorner and A. Zaman, "An explicit bound for the least prime ideal in the Chebotarev density theorem", *Algebra Number Theory* **11**:5 (2017), 1135–1197. MR Zbl

[Thorner and Zaman 2018] J. Thorner and A. Zaman, "A Chebotarev variant of the Brun–Titchmarsh theorem and bounds for the Lang–Trotter conjectures", *Int. Math. Res. Not.* **2018**:16 (2018), 4991–5027. MR Zbl

[Weiss 1983] A. Weiss, "The least prime ideal", *J. Reine Angew. Math.* **338** (1983), 56–94. MR Zbl

[Zaman 2017] A. A. Zaman, *Analytic estimates for the Chebotarev density theorem and their applications*, Ph.D. thesis, University of Toronto, 2017, Available at https://search.proquest.com/docview/1993455319. MR

Communicated by Roger Heath-Brown
Received 2018-03-22 Revised 2018-11-29 Accepted 2019-01-30

jthorner@stanford.edu *Department of Mathematics, Stanford University, Stanford, CA, United States*

aazaman@stanford.edu *Department of Mathematics, Stanford University, Stanford, CA, United States*

Algebra & Number Theory

msp.org/ant

See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2019 is US $/year for the electronic version, and $/year (+$, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

ANT peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

Algebra & Number Theory

Volume 13 No. 5 2019