NON-VANISHING OF DIRICHLET $L$-FUNCTIONS IN GALOIS ORBITS

RIZWANUR KHAN, DJORDJE MILIĆEVIĆ, AND HIEU T. NGO

Abstract. A well known result of Iwaniec and Sarnak states that for at least one third of the primitive Dirichlet characters to a large modulus $q$, the associated $L$-functions do not vanish at the central point. When $q$ is a large power of a fixed prime, we prove the same proportion already among the primitive characters of any given order. The set of primitive characters modulo $q$ of a given order can be described as an orbit under the action of the Galois group of the corresponding cyclotomic field. We also prove a positive proportion of nonvanishing within substantially shorter orbits generated by intermediate Galois groups as soon as they are larger than roughly the square-root of the prime-power conductor.

1. Introduction

Central values of $L$-functions are of fundamental importance in number theory. In particular, a host of results and conjectures, including the Birch and Swinnerton-Dyer Conjecture, the Riemann Hypothesis, and the Katz-Sarnak Density Conjecture, predict in various contexts that the central values of $L$-functions (or their derivatives, as appropriate for root number reasons) hold key arithmetic information and should vanish only when there are deep arithmetic reasons for them to do so and that this should be an exceptional occurrence in suitably generic families.

Introduced by Bohr and Landau \cite{BohLandau} in their study of zeroes of the Riemann zeta-function and notably used by Selberg \cite{Selberg} in the course of proving that a positive proportion of these zeroes lie on the critical line, the “mollifier” is the most versatile tool used in analytic number theory to prove the non-vanishing of central values of $L$-functions in families, often achieving a positive proportion result. What is by now a classical result using the mollifier is one of Iwaniec and Sarnak \cite{IwaniecSarnak}, concerning non-vanishing in the family of Dirichlet $L$-functions. They proved that for at least $(\frac{1}{3} - \epsilon)$ of the primitive Dirichlet characters modulo $q$, where $q$ is any integer sufficiently large in terms of $\epsilon$, the central value $L(\frac{1}{2}, \chi)$ is not zero. This is currently the best known result that can be proved with a “one-piece” mollifier. Earlier, Balasubramanian and Murty \cite{BalasMurty} had established a smaller positive proportion of non-vanishing, and recently, Bui \cite{Bui} proved, using a “two-piece” mollifier, that about 34\% of the central values in this family do not vanish as $q \to \infty$. When one restricts to the quadratic Dirichlet $L$-functions, Soundararajan \cite{SoundarV2} established that for at least $\frac{7}{8}$ of the fundamental discriminants $|d| \leq X$, the central value $L(\frac{1}{2}, (\frac{d}{\cdot}))$ is not zero, as $X \to \infty$. It is generally conjectured (see the discussion in \cite{SoundarV2}) that

\begin{equation}
L(\frac{1}{2}, \chi) \neq 0
\end{equation}
for any primitive character \( \chi \). Such a statement appears to be substantially beyond the reach of currently available technology.

Let \( \xi \) be a primitive \( \phi(q) \)th root of unity, where \( \phi \) is the Euler totient function. The Galois group \( G = \text{Gal}(\mathbb{Q}(\xi)/\mathbb{Q}) \) acts on the set of primitive Dirichlet characters modulo \( q \) as follows. For \( \sigma \in G \), we define \( \chi^\sigma \) to be the character given by \( \chi^\sigma(n) = \sigma(\chi(n)) \) for all integers \( n \). The Galois action partitions the set of characters into orbits, which are particularly natural from the arithmetic point of view, since the associated Dirichlet \( L \)-functions (by definition) share the same field of coefficients. One is led to wonder whether a positive proportion of non-vanishing can be proven for Dirichlet \( L \)-functions within each orbit. The aforementioned results [10] and [5] do not preclude the possibility that \( L(\frac{1}{2}, \chi) = 0 \) for all \( \chi \) in some orbit \( O \).

As a positive proportion statement toward (1.1) in this context, one would like to prove that for some \( c > 0 \), we have that

\[
\frac{1}{|O|} \sum_{\chi \in O, L(\frac{1}{2}, \chi) \neq 0} 1 \geq c - \epsilon
\]

for any \( \epsilon > 0 \), any orbit \( O \) of cardinality \( |O| > q^c \), and any integer \( q \) sufficiently large in terms of \( \epsilon \).

While of course \( c = 1 \) is expected in light of the conjecture (1.1), a positive proportion statement such as (1.2) with some \( c > 0 \) can be sufficient for some arithmetic applications, serve as support for the general philosophy of non-vanishing, and yet be more attainable with present methods. One might hope to match the best constant \( c = 0.34 \) currently available for the full set of primitive Dirichlet \( L \)-functions or the more classical proportion of Iwaniec and Sarnak. We establish the latter in the case that \( q \) is a large power of a prime.

**Theorem 1.1.** Let \( q = p^k \) for a fixed odd prime \( p \). For any \( \epsilon > 0 \) and \( k \) large enough in terms of \( \epsilon \) and \( p \), we have that (1.2) holds with \( c = \frac{1}{3} \).

When \( q = p^k \), the orbits under the Galois action can be described as follows. First note that a necessary and sufficient condition for the existence of a primitive character modulo \( p^k \) having order \( l \) is that \( l = p^{k-1}d \) for some \( d \mid (p - 1) \) (and \( d > 1 \) for \( k = 1 \)). The set of all primitive Dirichlet characters modulo \( q \) whose orders equal \( p^{k-1}d \) forms an orbit \( O \) of cardinality

\[
|O| = \phi(p^{k-1}d),
\]

and every orbit arises in this way. Thus \( O \) depends on \( d \) but we suppress this in the notation.

These claims about Galois orbits follow easily from the following well-known explicit description of primitive Dirichlet characters and of the action of the Galois group \( G \) on this set. Having fixed a primitive \( \phi(q)^{th} \) root of unity \( \xi \) and a primitive root \( g \) modulo \( p^k \), the set of Dirichlet characters \( \chi \) modulo \( p^k \) is in one-to-one correspondence with \( m \in \mathbb{Z}/\phi(p^k)\mathbb{Z} \) via the following identity [7, Chapter 5]:

\[
\chi_{[m]}(n) = \xi^{m \text{ind}_s(n)},
\]

where \( n \equiv g^{\text{ind}_s(n)} \mod p^k \), and the primitive characters \( \chi_{[m]} \) modulo \( p^k \) correspond precisely to the classes \( m \in \mathbb{Z}/\phi(p^k)\mathbb{Z} \) with \( p \nmid m \) (or such that \( m \neq 0 \mod \phi(p) \) for \( k = 1 \)). Moreover, the Galois automorphisms \( \sigma \) are in a one-to-one correspondence with primitive classes \( a \in (\mathbb{Z}/\phi(p^k)\mathbb{Z})^\times \) via \( \sigma_{[a]}(\xi) = \xi^a \), and, with respect to these isomorphisms, the action of \( G \) on the primitive Dirichlet characters is simply given by \( \sigma_{[a]}(\chi_{[m]}) = \chi_{[am]} \).

Although, from an analytic perspective, it may appear that we are dealing with a family only slightly thinner than the original unitary family of all primitive Dirichlet characters, in reality this
gives rise to a significant difficulty, which we describe below in the introduction. The key feature here is that the family is “thinning out” in a thoroughly arithmetic (rather than analytic) way. In fact, the same device that we use to overcome this basic difficulty subsequently allows us to prove a positive proportion of non-vanishing in substantially smaller “thin orbits” in Theorem 1.3 below.

The problem of studying the non-vanishing of $L$-functions within Galois orbits is a natural one that has yielded some of the strongest known results in the subject. Let $q = p^k$ for the remainder of this paper. Let $f$ be a holomorphic newform of weight 2 and level coprime to $p$ which has rational Fourier coefficients (equivalently, $f$ is associated to an elliptic curve over $\mathbb{Q}$ of conductor coprime to $p$, by [21, Theorems 7.14, 7.15] and [24, 23, 4]). Let $L(s, f \times \chi)$ be the $L$-function of $f$ twisted by a primitive Dirichlet character $\chi$ modulo $q$, having central point $s = 1$ by a functional equation normalized so as to relate $L(s, f \times \chi)$ and $L(2 - s, f \times \overline{\chi})$. Rohrlich [17] showed that $L(1, f \times \chi)$ does not vanish as long as $k$ is large enough in terms of $p$ and $f$. To prove his result, Rohrlich appealed to an “algebraicity” theorem of Shimura [19, 20, Theorem 1], which implies that if $L(1, f \times \chi) = 0$ then $L(1, f \times \chi^\sigma) = 0$ for all $\sigma \in G$. By this, if the sum

$$\sum_{\chi \in \mathcal{O}} L(1, f \times \chi)$$

is nonzero, then every summand is nonzero. Rohrlich found an asymptotic for (1.5) for any orbit $\mathcal{O}$ when $k$ is large enough in terms of $p$ and $f$, and showed that the main term is indeed nonzero. Chinta [6] extended Rohrlich’s work to the case of prime moduli $q$ by considering instead the sum

$$\sum_{\chi \in \mathcal{O}} L(1, f \times \chi) M(f \times \chi),$$

where $M(f \times \chi)$ is a truncation of the formal Dirichlet series for $L(1, f \times \chi)^{-1}$. The effect of this mollifier is that each summand of (1.6) is “morally” close to 1, and this allows Chinta to show that the sum is nonzero provided $|\mathcal{O}| > q^{q^{\theta + \epsilon}}$ and $q$ is large enough in terms of $f$. In particular, this implies the non-vanishing of $L(1, f \times \chi)$ over big orbits when $q$ is a large enough prime in terms of $f$, a fact which does not follow from Rohrlich’s result.

All these results for twists of elliptic modular $L$-functions rely heavily on the algebraicity results of Shimura. Such a route is not available in the present context of central values of Dirichlet $L$-functions.

We establish Theorem 1.1 by evaluating the mollified moments

$$\sum_{\chi \in \mathcal{O}} L(\frac{1}{2}, \chi) M(\chi)$$

and

$$\sum_{\chi \in \mathcal{O}} |L(\frac{1}{2}, \chi)|^2 |M(\chi)|^2,$$

where

$$M(\chi) = \sum_{m \leq q^\theta} \frac{a_m \chi(m)}{m^2}$$

is a mollifier of length $q^\theta$, for some $\theta \geq 0$ and coefficients $a_m$ satisfying $a_m \ll m^\epsilon$ and $a_1 = 1$. We are able to obtain asymptotics with a power-saving error term for arbitrary mollifiers when $\theta$ is any
fixed constant satisfying \(0 \leq \theta < \frac{1}{2}\); see Theorem \(1.2\). As discussed in section 3.1, this allows us to deduce Theorem \(1.1\) with the proportion of non-vanishing \(c = \frac{1}{3}\) by taking \(\theta \to \frac{1}{2}\).

Although \(|L(\frac{1}{2}, \chi)|^2\) can be considered to be analogous to \(L(1, f \times \chi)\), our problem has some important differences from the one considered by Rohrlich and Chinta. Firstly, the additional factor \(|M(\chi)|^2\) in \((1.8)\) makes our problem more complex. This is readily seen when comparing with \((1.5)\), and to compare with \((1.6)\) we note that Chinta’s method only works for a specific mollifier while ours works for an arbitrary mollifier, and that \(M(f \times \chi)\) has length at most \(q^{1-\varepsilon}\) (see [6, pg 22]) while \(|M(\chi)|^2\) has length at most \(q^{1-\varepsilon}\). However one must keep in mind that Chinta’s result works for \(q\) prime, while ours does not.

The second difference is the method of proof. To evaluate \((1.5)\), Rohrlich had to consider the averages 
\[
\sum_{\chi \in \mathcal{O}} \chi(n),
\]
for \(n \leq q^{1+\varepsilon}\), while for \((1.8)\) we must consider the averages 
\[
\sum_{\chi \in \mathcal{O}} \chi(n_1m_1)(n_2m_2)
\]
in which \(n_1n_2 \leq q^{1+\varepsilon}\) and \(m_1, m_2 \leq q^\theta\). As we will see, \((1.10)\) is zero unless \(n^{p-1} \equiv 1 \mod p^{k-1}\). From this, Rohrlich could immediately conclude that \(n = 1\) or \(n > p^\frac{k-1}{q}\), thereby effectively isolating the contribution of the diagonal term \(n = 1\). In contrast, \((1.11)\) is zero unless 
\[
(n_1m_1)^{p-1} \equiv (n_2m_2)^{p-1} \mod p^{k-1}.
\]
Writing \(n_1m_1 \equiv \zeta n_2m_2 \mod p^k\), it is now much harder to isolate the contribution of the diagonal terms \(n_1m_1 = n_2m_2\) (that is, \(\zeta \equiv 1 \mod p^k\); a priori it is, for example, perfectly plausible that \(n_1m_1\) and \(n_2m_2\) could be fairly close to each other without actually being equal. To isolate the contribution of the diagonal terms \(n_1m_1 = n_2m_2\), we will appeal to the \(p\)-adic version of Roth’s theorem, Lemma 2.4 below. The upshot is that, keeping in mind that \(\zeta^{p^{-1}} \equiv 1 \mod p^k\), having two solutions to \((1.12)\) within the same class of \(\zeta \not\equiv \pm 1 \mod p^k\) too close to each other would ultimately yield too good of an approximation in the \(p\)-adic norm to a \((p - 1)\)th \(p\)-adic root of unity.

It is interesting that Roth’s theorem, a deep result from diophantine approximation, should be used to prove the non-vanishing of \(L\)-functions. This connection has been made before in different contexts by Rohrlich [10] and Greenberg [8]. Our paper offers another such example and it seems to be the first one involving a family of Dirichlet characters as well as the first one involving a genuinely second mollified moment.

We now give our evaluations of \((1.7)\) and \((1.8)\). This result could also be of independent interest; indeed, evaluations of moments of \(L\)-functions twisted by short sums such as our \(M(\chi)\) in \((1.9)\) are in many ways analogous to moments of \(L\)-functions in the \(t\)-aspect weighted against short Dirichlet polynomials and are a mainstay of the analytic theory of \(L\)-functions since they can also be used, for example, for amplification or the method of resonators. Before stating our result, we note that the value \(\chi(-1)\) is the same for every character in any given orbit \(\mathcal{O}\), as it is a rational number.

\textbf{Theorem 1.2.} Let \(q = p^k\) for a fixed odd prime \(p\) and let \(\mathcal{O}\) be any Galois orbit of primitive Dirichlet characters mod \(q\). Let \(\iota \in \{0, 1\}\) be such that \(\chi(-1) = (-1)^\iota\) for any \(\chi \in \mathcal{O}\).
(1) For $0 \leq \theta < 1$ in $[1.9]$, we have that
\[
\frac{1}{|\mathcal{O}|} \sum_{\chi \in \mathcal{O}} L\left(\frac{1}{2}, \chi \right) M(\chi) = 1 + O\left(q^{-\frac{1}{4} + \epsilon} + q^{-\frac{1}{2} + \frac{\theta}{2} + \epsilon}\right)
\]
for any $\epsilon > 0$, where the implied constant depends on $\epsilon$ and $p$.

(2) For $0 \leq \theta < \frac{1}{2}$ in $[1.9]$, we have that
\[
\frac{1}{|\mathcal{O}|} \sum_{\chi \in \mathcal{O}} |L\left(\frac{1}{2}, \chi \right)|^2 |M(\chi)|^2
= \frac{p-1}{p} \sum_{\substack{m_1, m_2 \leq q^\theta \\
(m_1, m_2, q)=1}} a_{m_1 \pi_m m_2} \left( \log \left( \frac{q(m_1, m_2)^2}{\pi m_1 m_2} \right) + C \right) + O\left(q^{-\frac{3}{4} + \epsilon} + q^{-\frac{3}{2} + \theta + \epsilon}\right)
\]
for any $\epsilon > 0$, where the implied constant depends on $\epsilon$ and $p$, $(m_1, m_2)$ denotes the greatest common divisor of $m_1$ and $m_2$, $[m_1, m_2]$ denotes the least common multiple of $m_1$ and $m_2$, and $\gamma = 0.57721 \cdots$ is the Euler constant.

Taking $\theta = 0$ in Theorem $[1.2]$ yields in particular the first and second moments of $L\left(\frac{1}{2}, \chi \right)$ over Galois orbits. Already this evaluation is new and rests on our nonarchimedean Diophantine argument; see the discussion after Lemma 2.4. The proof of Theorem 1.2 forms the technical heart of the paper and is presented in sections 3.2 and 3.3.

We remark that the implicit constants in all our main results are ineffective due to their dependence on $p$-adic Roth’s theorem. However, the mollified first moment in Theorem $[1.2]$ (1) can also be evaluated without recourse to $p$-adic Roth’s theorem, at the expense of the error term $O(q^{-\frac{1}{4} + \epsilon})$ being replaced by the weaker but effective error term $O(q^{-\frac{1}{m_1 m_2} + \epsilon})$. This will be shown in the course of the proof of Theorem $[1.2]$ (1).

Finally, we address the refined question of non-vanishing in smaller sub-families within the Galois orbits of primitive characters modulo $q$. A rather natural sub-family emerges when considering orbits of primitive characters under various subgroups $H$ of the Galois group $G = \text{Gal}(\mathbb{Q}(\xi)/\mathbb{Q})$. These subgroups form a partially ordered set (corresponding by Galois theory to the tree of intermediary field extensions $\mathbb{Q} \leq K \leq \mathbb{Q}(\xi)$), and, as the subgroup $H$ varies from $G$ through its various subgroups to the identity, the corresponding orbits of a fixed primitive character modulo $q$ can be seen as interpolating (or shrinking) between its full Galois orbit, considered in Theorem 1.1 and the individual character.

We describe these “thin orbits” explicitly in cases of our interest. For every $0 \leq \kappa \leq k - 1$, denote $K_{k-1-\kappa} = \mathbb{Q}(p^\kappa)$. (Note that the field $K_{\ell}$ is independent of $k$.) Since $[K_{k-1} : K_0] = \phi(p^{k-1}) \asymp_p \phi(q)$ and we are primarily concerned with the case of fixed $p$ and large $k$, we focus here on the tower of these intermediate fields
\[
\mathbb{Q}(\xi) = K_{k-1} \supseteq \cdots \supseteq K_0 \supseteq \mathbb{Q}.
\]
Recall that $\mathcal{O}$ denotes an orbit induced by the action of $\text{Gal}(\mathbb{Q}(\xi)/\mathbb{Q})$. Thus the intermediate Galois group $\text{Gal}(\mathbb{Q}(\xi)/K_{k-1-\kappa})$ acts on $\mathcal{O}$; we call an orbit induced by this action a thin Galois orbit, and we write $\mathcal{O}_\kappa$ for any one of these thin orbits. Note that already the thin orbits $\mathcal{O}_k$ refine the full Galois orbits $\mathcal{O}$, with thin orbits $\mathcal{O}_\kappa$ for smaller $\kappa$ being progressively smaller (so that we may
think of the parameter $κ$ essentially as an indicator or the logarithmic size of the corresponding thin orbits), all the way to the extreme case of $κ = 0$, which corresponds to the single primitive characters.

It is apparent that $σ ∈ \text{Gal}(ℚ(ξ)/ℚ)$ satisfies $σ ∈ \text{Gal}(ℚ(ξ)/K_{k−1−κ})$ if and only if

$$σ(ξ) = ξ^a \quad \text{for some } a ≡ 1 \mod p^{k−1−κ}(p−1);$$

indeed, on setting $σ(ξ) = ξ^a$, the condition that $σ ∈ \text{Gal}(ℚ(ξ)/K_{k−1−κ})$ is tantamount to $σ(ξ^p) = ξ^p$, which amounts to $φ(p^k)|p^κ(a−1)$. We thus see, using (1.4), that two characters $χ_1, χ_2$ are in the same thin Galois orbit $O_κ$ if and only if $χ_1|\overline{χ_2}$ has order dividing $p^κ$ (informally speaking, if $χ_1$ and $χ_2$ differ by an “algebraically simpler” character, one that is “shallower” in the sense of the “depth aspect” of modulus $q = p^k$ with large $k$). In particular, all thin orbits $O_κ$ for the same $κ$ are of equal size given by (keeping in mind that $(a, φ(q)) = 1)$

$$|O_κ| = \begin{cases} p^κ, & 0 ≤ κ < k−1, \\ φ(p^{k−1}), & κ = k−1. \end{cases}$$

Another, more “analytic”, way to think about the thin orbits $O_κ$ is provided by the explicit characterization of the “principal part” of the dual of the group $(ℤ/qℤ)^κ$ for a high prime power $q = p^k$, which is essentially due to Postnikov [13]. Let $X_k = (ℤ/p^k ℤ)^κ$, and let $π_a$ denote the power map $π_a : X_k → X_k, [x] → [x^a]$. Corresponding to the decomposition $X_k = X_{k0} × X_{k1}$, where $X_{k0} = \text{im } π_{p^{k−1}} = \text{ker } π_{p−1}, |X_{k0}| = p−1$, and $X_{k1} = \text{im } π_{p−1} = \text{ker } π_{p^{k−1}} = \{[x] ∈ X_k : x ≡ 1 \mod p\}$, we have the canonical decomposition of dual groups $\hat{X}_k ≅ \hat{X}_{k0} × \hat{X}_{k1}$. Let $log_p$ denote the $p$-adic logarithm and $ψ(x)$ denote the “standard” additive character $ψ : ℤ_p → ℂ^×$ such that its kernel is exactly $ℤ_p$ and that $ψ(x) = e^{2πix}$ for $x ∈ ℤ[1/p] ⊆ ℚ_p ∩ ℜ$. According to Postnikov’s lemma (for $p > 2$; see also [13 Lemma 13]), every character $χ^{(1)} ∈ \hat{X}_{k1}$ is of the form

$$χ^{(1)}(1+pt) = a_0(p^{k−1}) = ψ\left(a_0 log_p(1+pt)\right)$$

for some $a = a_0p^{−k} ∈ p^{−k}ℤ_p/p^{−1}ℤ_p$, with primitive characters corresponding to $a ∈ p^{−k}ℤ_p^×/p^{−1}ℤ_p$.

The isomorphism of $p^{−k}ℤ_p/p^{−1}ℤ_p → \hat{X}_{k1}$ given by $a ↦ χ^{(1)}_a$ induces a metric on $\hat{X}_{k1}$ via $d(χ^{(1)}_a, χ^{(1)}_b) = |a − b|_p/p = \text{cond}(χ^{(1)}_a, χ^{(1)}_b)$ for $χ^{(1)}_a ≠ χ^{(1)}_b$. Relative to the above decomposition of $\hat{X}_k$ and this metric on $\hat{X}_{k1}$, the thin orbit $O_κ$, containing a character $χ = χ^{(0)}χ^{(1)}$ is, for $κ > 0$, precisely the set $\{χ^{(0)}\} × B[χ^{(1)}, p^{κ}]$; here $B[χ^{(1)}, p^{κ}]$ denotes the closed ball in $\hat{X}_{k1}$ with center $χ^{(1)}$ and radius $p^{κ}$ with respect to the above-defined metric. In particular, all characters in a thin orbit $O_κ$ share the same $X_{k0}$-component and their $X_{k1}$-components are all close to each other, with the corresponding neighborhood around a fixed character $χ$ shrinking as $κ$ decreases.

From the point of view of harmonic analysis, we can now see the basic difficulty of restricting to a smaller family of characters in a single Galois orbit (or a thin Galois orbit) as follows. Characters in $\hat{X}_{k0}$ have orders which are divisors of $p−1$, while those in $\hat{X}_{k1}$ have orders which are powers of $p$. Thus, keeping in mind the description of Galois orbits as in (1.3), restricting from the full family of all primitive characters modulo $q$ down to a single Galois orbit is tantamount to restricting to primitive characters with their $X_{k0}$-components in a specified, fixed subset of $\hat{X}_{k0}$. This shrinking to a smaller subset of $\hat{X}_{k0}$ is (naturally) reflected on the dual side in the survival of further terms in Lemma 2.3, namely, of those $n$ congruent modulo a high power of $p$ to one of the $(p−1)^{th}$ roots of unity (that is, exactly to elements of the dual group $\hat{X}_{k0}$, as in (2.6). With the further shrinking of the family down to thin Galois orbits, we are isolating an individual component $χ^{(0)} ∈ \hat{X}_{k0}$,
which on the dual side is again reflected by the initial survival of the \((p - 1)\)th roots of unity in Lemma 3.1, followed by isolating the components \(\chi^{(1)}\) in smaller neighborhoods within \(\hat{X}_{k,1}\), which corresponds to the survival of further terms in more permissive congruence classes containing these roots of unity as in (3.15).

Our techniques, which ultimately rely on the impossibility of overly good \(p\)-adic approximations to algebraic integers, are very well suited to the study of thin orbits of primitive characters to prime power moduli and give the following refinement of Theorem 1.1.

\[ \text{Theorem 1.3.} \quad \text{Let } q = p^k \text{ for a fixed odd prime } p. \text{ For any } \epsilon > 0 \text{ and } k \text{ large enough in terms of } \epsilon \text{ and } p, \text{ and for any } \kappa > k/2, \text{ we have that (1.2) holds also when } O \text{ is replaced by any } \text{“thin orbit” } O_\kappa, \text{ with} \]

\[ c = c_\kappa = \frac{\kappa/k - 1/2}{\kappa/k + 1/2}. \]

To keep the article light and readable, we present our arguments in the context of Theorem 1.1 first and then indicate the adjustments needed for the proof of Theorem 1.3 in section 3.4, emphasizing the parallels between the two cases. In particular, we derive the non-vanishing statement of Theorem 1.3 by using the mollified first and second moments of the central values of Dirichlet \(L\)-functions over thin Galois orbits. These moments are evaluated in Theorem 3.2, whose statement is also found in section 3.4.

2. Preliminaries

**Notation.** Throughout the paper, \(\epsilon > 0\) denotes a parameter which may be chosen to be as small as we like, but need not have the same value from one occurrence to another. The letter \(p\) denotes an odd prime and \(q = p^k\). We use \(\mu_{p-1}\) to denote the set of \((p-1)\)th roots of unity in the algebraic closure \(\overline{\mathbb{Q}}\) of \(\mathbb{Q}\). All implicit constants may depend on \(\epsilon\), \(p\) and the parameter \(\theta\) introduced in (1.9), but not on \(k\).

2.1. Approximate functional equations. We have the following standard approximate functional equations.

**Lemma 2.1.** For a primitive Dirichlet character \(\chi\) modulo \(q\), let \(\iota \in \{0, 1\}\) be such that \(\chi(-1) = (-1)^\iota\), and let

\[ U(x) = \frac{1}{2\pi i} \int_{(2)} \frac{\Gamma\left(\frac{s + \iota}{2} + \frac{1}{2}\right)}{\Gamma\left(\frac{s + 1}{2}\right)} (\pi x)^{-s} \frac{ds}{s}, \]

\[ V(x) = \frac{1}{2\pi i} \int_{(2)} \frac{\Gamma\left(\frac{s + \iota}{2} + \frac{1}{2}\right)^2}{\Gamma\left(\frac{s + 1}{2}\right)^2} (\pi x)^{-s} \frac{ds}{s}. \]

We have that

\[ U(x) \ll_{\iota} x^{-\epsilon}, \quad V(x) \ll_{\iota} x^{-\epsilon} \]

for any \(x, \epsilon > 0\). For any \(\lambda > 0\), we have that

\[ L\left(\frac{1}{2}, \chi\right) = \sum_{n \geq 1} \frac{\chi(n)}{n^{\frac{1}{2}}} U\left(\frac{n}{q^{1+x}}\right) + O(q^{-100}), \]

\[ |L\left(\frac{1}{2}, \chi\right)|^2 = 2 \sum_{n_1, n_2 \geq 1} \frac{\chi(n_1)\chi(n_2)}{(n_1 n_2)^{\frac{1}{2}}} V\left(\frac{n_1 n_2}{q}\right), \]
where the implied constant in (2.3) depends on $\lambda$.

Proof. The first equation (2.3) is established by the functional equation of $L(s, \chi)$, which may be found in [9, Theorem 4.15], together with [9, Theorem 5.3], in which we take $G(u) = 1$ and $X = q^{\frac{1}{2} + \lambda}$. With this choice of $X$, the second sum in [9 (5.12)] may be bounded by $q^{-100}$. The second equation is established by applying [9, Theorem 5.3] to the product $L(s, \chi)L(s, \chi)$, rather than to each factor individually, with $G(u) = 1$ and $X = 1$. The estimates (2.2) may be found in [9, Proposition 5.4]. □

The sums in (2.3) and (2.4) are essentially restricted to $n < q^1 + \lambda + \epsilon$ and $n_1 n_2 < q^{1 + \epsilon}$, by (2.2).

2.2. Character averages. In this section, we record the orthogonality relations provided by averaging over the family of Dirichlet characters in a Galois orbit; see Lemma 2.3 below. We start with a familiar auxiliary result.

Lemma 2.2. Let $m$ and $k \geq 1$ be integers. If $p^k \mid (m^p - 1)$, then $p^{k-1} \mid (m - 1)$.

Proof. The claim is trivially true for $m = 1$, so assume $m > 1$. Since $m \equiv m^p \equiv 1 \pmod{p}$, we may write $m = 1 + p^r t$ for some $r \geq 1$ and some integer $t$ with $p \nmid t$. Then we have

$$m^p - 1 = (1 + p^r t)^p - 1 \equiv p^{r+1} t \pmod{p^{r+2}},$$

so that $p^{r+1}$ is the highest power of $p$ that divides $m^p - 1$. In particular, since $p^k \mid (m^p - 1)$, we get $r + 1 \geq k$, and so $r \geq k - 1$ and $p^{k-1} \mid (m - 1)$.

Lemma 2.3. Let $q = p^k$ for an odd prime $p$ and let $O$ be any Galois orbit of primitive Dirichlet characters mod $q$. For any integer $n$, we have that

$$\sum_{\chi \in O} \chi(n) = 0 \quad (2.5)$$

unless

$$n^{p-1} \equiv 1 \pmod{p^{k-1}}. \quad (2.6)$$

Proof. Suppose that $O$ has cardinality given by (1.3), for some $d \mid (p - 1)$. We first show that (2.5) holds unless

$$n^{p(p-1)} \equiv 1 \pmod{q}. \quad (2.7)$$

It was shown in [6, pg 17] that

$$\frac{1}{|O|} \sum_{\chi \in O} \chi(n) = \frac{\mu(\text{ord}(n^{\frac{p-1}{d}}))}{\phi(\text{ord}(n^{\frac{p-1}{d}}))}, \quad (2.8)$$

where $\text{ord}(n^{\frac{p-1}{d}})$ denotes the multiplicative order of $n^{\frac{p-1}{d}}$ in the group $(\mathbb{Z}/q\mathbb{Z})^\times$. Since $|(\mathbb{Z}/q\mathbb{Z})^\times| = p^{k-1}(p - 1)$, if (2.7) is not satisfied, then $p^2$ divides $\text{ord}(n^{\frac{p-1}{d}})$, and so (2.8) is zero.

Now if (2.7) holds, then Lemma 2.2 implies, by taking $m = n^{p-1}$, that (2.6) holds. □

Note that the condition (2.6), which is all we will need from our orthogonality relations, requires less information than what is provided by the full sum over all $\chi \in O$ or the explicit evaluation (2.8). This will be transparent in Lemma 3.1 which features a thinner average and for which we provide an independent proof.
2.3. Roots of unity. The averaging over the family of primitive Dirichlet characters in a single Galois orbit or a single thin Galois orbit leads one to consider the congruence in $x \in \mathbb{Z}$:

$$x^{p-1} \equiv 1 \mod p^a,$$

where $a$ is a positive integer; see Lemmas 2.3 and 3.1. This congruence can be regarded as a condition on $x \in \overline{\mathbb{Q}}_p$, as we now explain.

Fix an embedding $\tau : \overline{\mathbb{Q}} \to \overline{\mathbb{Q}}_p$, which realizes $\overline{\mathbb{Q}}$ as a subset of $\overline{\mathbb{Q}}_p$ and which restricts the standard absolute value $|\cdot |_p$ from $\overline{\mathbb{Q}}_p$ to $\overline{\mathbb{Q}}$. Then the set $\mu_{p-1} \subset \overline{\mathbb{Q}}$ of $(p-1)^{th}$ roots of unity can be viewed as a subset of $\overline{\mathbb{Q}} \cap \mathbb{Z}_p \subset \overline{\mathbb{Q}}_p$. By Hensel’s lemma [2, Chapter II.2, Proposition 2], the equation $y^{p-1} = 1 = 0$ in $y \in \overline{\mathbb{Q}}_p$ has as solutions the set $\mu_{p-1}$ and one can write $\mu_{p-1} = \{\zeta_j : 1 \leq j \leq p-1\}$ so that $\zeta_j \equiv j \mod p$ or equivalently $\zeta_j - j \in p\mathbb{Z}_p$; in particular, $\zeta_1 = 1$.

A useful observation is that if $x \in \mathbb{Z}$ satisfies

$$x^{p-1} \equiv 1 \mod p^a$$

for some positive integer $a$, then

$$x \equiv \zeta_j \mod p^a$$

for exactly one $1 \leq j \leq p-1$. This is clear on noting that the hypothesis is equivalent to $\prod_{j=1}^{p-1}|x - \zeta_j|_p \leq p^{-a}$ and that two distinct values $x - \zeta_i$ and $x - \zeta_j$ can not both have absolute value less than one.

2.4. Roth’s theorem. The following lemma is a special case of a result of Rohrlich [16, Proposition 1, pg 401]. It is a consequence of the $p$-adic version of Roth’s theorem, due to Ridout [15].

Lemma 2.4. Let $p$ be a fixed odd prime, let $\beta \in \mathbb{Z}_p$ and let $0 < \delta < \frac{1}{2}$. Suppose that $\beta$ is algebraic over $\mathbb{Q}$ of degree at least 2. Then, for sufficiently large $k \geq k_0(\beta, \delta)$, there are no nonzero integers $a$ and $b$ which satisfy

$$|a - b\beta|_p \leq p^{-k+1}$$

and

$$|a|, |b| < (p^k)^{\frac{3}{2} - \delta}.$$ 

Proof. In [16, Proposition 1, pg 401], take $\alpha_p = 1$ and $\beta_p = \beta$. Note that condition (ii) directly above the proposition is satisfied because $\beta \notin \mathbb{Q}$. Note that $k_0$ is not computable.

When $(b, p) = 1$, this result says that an approximation of the algebraic number $\beta$ by $\frac{a}{b}$ to an error within $(\max\{|a|, |b|\})^{-2-\delta}$ is too good to exist. Note that there are infinitely many approximations of any $\beta \in \mathbb{Z}_p$ up to within $\ll (\max\{|a|, |b|\})^{-2}$ by a $p$-adic incarnation of Dirichlet’s Approximation Theorem (see also [2] for an analogue of the Farey dissection in this context), and it is a hallmark of Roth’s theorem that the exponent in Lemma 2.4 is essentially the best possible. For an interesting investigation into existence of algebraic numbers exhibiting approximability by rationals within the transition range between the exponents $-2$ and $-2 - \delta$, see [11].

We stress that the $p$-adic analogues of previous partial results toward Roth’s Theorem, such as the theorems of Liouville, Thue, Siegel, and Dyson, do not (except for a few small primes $p$) suffice to obtain an asymptotic in the situation of Theorem 1.2 (2) with our methods. This is so because, for example when estimating (3.12), in order to obtain (3.13) with a $o(1)$-upper bound on the right-hand side for large $q$, we need to be able to take $Q$ (essentially the allowable length of intervals in Lemma 2.5 below) to be at least $(p^k)^{\frac{3}{2} + \delta}$, whereas all results prior to the actual $p$-adic Roth’s theorem furnish only $(p^k)^{o(1)}$. 
We also have, by the assumption on the lengths of $A$ and $B$, it follows from the strong triangle inequality that

$$\zeta \in \mathbf{\mu}_{p-1}(1 \leq j \leq p-1)$$

By the discussion in section 2.3, the roots of unity $\zeta_j \in \mathbf{\mu}_{p-1}(1 \leq j \leq p-1)$ are distinct modulo $p$. Hence, using the assumption $(ab, p) = 1$, we have that $a - b\zeta_i$ and $a - b\zeta_j$ are distinct modulo $p$ whenever $1 \leq i < j \leq p-1$. Thus, by (2.11) and (2.9), for some $\zeta_j \in \mathbf{\mu}_{p-1}\{\pm 1\}$, we have

$$|a - b\zeta_j|_p \leq p^{-k+1}.$$ 

Now suppose for a contradiction that at least $p-2$ pairs $(a, b)$ satisfy the conditions of the lemma. Then by the argument above and Dirichlet’s Box Principle, we have for some $\zeta_j' \in \mathbf{\mu}_{p-1}\{\pm 1\}$, at least two distinct pairs $(a_1, b_1), (a_2, b_2) \in A_k \times B_k$ satisfying

$$|a_i - b_i\zeta_j'|_p \leq p^{-k+1} \quad (i = 1, 2).$$

It follows from the strong triangle inequality that

$$|(a_2 - a_1) - (b_2 - b_1)\zeta_j'|_p \leq \max\{|a_1 - b_1\zeta_j'|_p, |a_2 - b_2\zeta_j'|_p\} \leq p^{-k+1}.$$ 

We also have, by the assumption on the lengths of $A_k$ and $B_k$, that

$$|a_2 - a_1|, |b_2 - b_1| \leq (p^k)^{\frac{1}{2} - \delta}.$$ 

Thus by Lemma 2.4, in which we take $\beta = \zeta_j'$, we deduce that for $k > k_1$ sufficiently large, where $k_1$ is not computable, we must have $a_2 - a_1 = b_2 - b_1 = 0$. This is the desired contradiction. \hfill \Box

3. PROOFS OF THE MAIN RESULTS

3.1. Mollifiers and nonvanishing. In this section, we explain how Theorem 1.2 follows from Theorem 1.1. Theorem 1.2 itself will be proved in sections 3.2 and 3.3.

The starting point of the mollifier method is the observation that

$$\frac{1}{|\mathcal{O}|} \sum_{\chi \in \mathcal{O}, L(\frac{1}{2}, \chi) \neq 0} 1 \geq \frac{1}{|\mathcal{O}|} \sum_{\chi \in \mathcal{O}, L(\frac{1}{2}, \chi)M(\chi) \neq 0} 1 \geq \frac{1}{|\mathcal{O}|} \sum_{\chi \in \mathcal{O}} \left| \frac{L(\frac{1}{2}, \chi)M(\chi)}{L(\frac{1}{2}, \chi)|M(\chi)|} \right|^2.$$ 

The second inequality above is the Cauchy-Schwarz inequality. The next step is to insert the main terms of Theorem 1.2 into the ratio on the right hand side of (3.1). The final step is to choose the coefficients $a_m$ so that the ratio is maximized. We note that the main terms of our mollified
moments are identical to those in Iwaniec and Sarnak's problem (see [10] (5.5)) because in both problems the main terms arise from the diagonal contributions. That is, if \( \chi(-1) = (-1)^t \) for all \( \chi \in \mathcal{O} \), we have for any \( 0 \leq \theta < \frac{1}{2} \) and some \( \delta > 0 \) depending on \( p \) and \( \theta \) that
\[
\frac{1}{|\mathcal{O}|} \sum_{\chi \in \mathcal{O}} L(\frac{1}{2}, \chi) M(\chi) - \frac{2}{\varphi^*(q)} \sum_{\chi \mod q}^* L(\frac{1}{2}, \chi) M(\chi) \ll q^{-\delta},
\]
\[
\frac{1}{|\mathcal{O}|} \sum_{\chi \in \mathcal{O}} |L(\frac{1}{2}, \chi)|^2 |M(\chi)|^2 - \frac{2}{\varphi^*(q)} \sum_{\chi \mod q}^* |L(\frac{1}{2}, \chi)|^2 |M(\chi)|^2 \ll q^{-\delta},
\]
where \( \sum^* \) means that summation is restricted to the primitive characters and \( \varphi^*(q) = \sum_{\chi \mod q}^* 1 \).

Note that the exchange of summation above is valid because although the function \( U \) \((3.2)\) is \( \Re \) large, using also the divisor bound, we have that (3.2) equals to one of \( \zeta \) unless \((3.2)\) is \( \frac{\theta}{1+\theta} + o(1) \) as \( k \to \infty \). Taking \( \theta \to \frac{1}{2} \), Theorem 1.1 thus follows from Theorem 1.2.

3.2. Evaluation of the first mollified moment. In this section, we prove Theorem 1.2 (1).

In Lemma 2.1 let \( \lambda > 0 \) be such that \( \lambda + \theta < 1 \). We have that
\[
\frac{1}{|\mathcal{O}|} \sum_{\chi \in \mathcal{O}} L(\frac{1}{2}, \chi) M(\chi) = \frac{1}{|\mathcal{O}|} \sum_{m \leq q^\theta, n \geq 1} \frac{a_m}{(nm)^{\frac{1}{2}}} U\left(\frac{n}{q^{1+\lambda}}\right) \frac{1}{|\mathcal{O}|} \sum_{\chi \in \mathcal{O}} \chi(nm) + O(q^{-99}).
\]

Note that the exchange of summation above is valid because although the function \( U(x) \) depends on \( \chi(-1) \), this value is the same for every character in \( \mathcal{O} \). Now, by Lemma 2.3 we have that
\[
\frac{1}{|\mathcal{O}|} \sum_{\chi \in \mathcal{O}} \chi(nm) = 0
\]
unless \( (nm)^{p-1} \equiv 1 \mod p^{k-1} \). It follows from the observation in section 2.3 that \( nm \) is congruent to one of \( \zeta_1, \ldots, \zeta_{p-1} \in \mu_{p-1} \) modulo \( p^{k-1} \mathbb{Z} \). Separating the term \( nm = 1 \) for the main term in Theorem 1.2 (1) and using also the divisor bound, we have that (3.2) equals
\[
U\left(\frac{1}{q^{1+\lambda}}\right) + O\left(q^\theta \sum_{1 \leq j \leq p-1} \sum_{1 \leq s \leq q \mod p^{k-1}} s^{-\frac{1}{2}}\right) + O\left(q^\theta \sum_{1 \leq j \leq p-1} \sum_{1 \leq s \leq q \mod p^{k-1}} s^{-\frac{1}{2}}\right).
\]

By shifting the line of integration in (2.1) to \( \Re(s) = -\frac{1}{2} + \epsilon \), we see that the main term above equals
\[
1 + O(q^{-\frac{1}{2} + \epsilon}).
\]

The summands in the first error term are particularly sensitive to the size of \( s \). From the condition that \( s^{p-1} \equiv 1 \mod p^{k-1} \) and \( s > 1 \), we immediately have that \( s > p^\frac{k-1}{2} \), and the first error term is seen to be
\[
O\left(q^{-\frac{1}{2} + \frac{1}{2} + \epsilon}\right).
\]

without recourse to \( p \)-adic Roth's theorem. However, we can improve this estimate by appealing to Lemma 2.4. The terms with \( s \equiv \pm 1 \mod p^{k-1} \), \( s > 1 \), contribute \( O(p^{-(k-1)/2}) \). As for the terms corresponding to \( s \equiv \zeta_j \neq \pm 1 \mod p^{k-1} \), Lemma 2.4 with \( \beta = \zeta_j \) guarantees that, for sufficiently large \( k \), there are no values of \( |s| < (p^k)^\frac{1}{2} - \delta \) with \( |s - \zeta_j - 1| \leq p^{-k+1} \), so that all these terms must in fact satisfy \( s \geq (p^k)^{\frac{1}{2} - \delta} \), and in total the first error term is
\[
O\left(q^{-\frac{1}{2} + \frac{1}{2} + \delta + \epsilon}\right).
\]
Finally, writing \( s = \zeta_j + p^{k-1} \), we have that the second error term above is bounded by
\[
q^\epsilon \sum_{1 \leq r < q^{k+\theta + \epsilon}} (qr)^{-1/2} \ll q^{\frac{k+\theta - 1}{2} + \epsilon}.
\]
Taking the positive \( \lambda \) and \( \delta \) to be as small as we like and adjusting the implied constants completes the proof.

3.3. Evaluation of the second mollified moment. In this section, we prove Theorem 1.2 (2).

By Lemma 2.1, we have that
\[
\frac{1}{|O|} \sum_{\chi \in O} |L(\frac{1}{2}, \chi)|^2 |M(\chi)|^2
\]
\[
= 2 \sum_{m_1, m_2 \leq q^\delta} \sum_{n_1, n_2 \geq 1} \frac{a_{m_1} \overline{a}_{m_2}}{(n_1 n_2 m_1 m_2)^{1/2}} V \left( \frac{n_1 n_2}{q} \right) \frac{1}{|O|} \sum_{\chi \in O} \chi(n_1 m_1) \overline{\chi}(n_2 m_2).
\]

We write this as a sum of diagonal terms and off-diagonal terms,
\[
\sum_{n_1 m_1 = n_2 m_2} + \sum_{n_1 m_1 \neq n_2 m_2}.
\]

The diagonal. We first consider the diagonal terms. The equality \( n_1 m_1 = n_2 m_2 \) is the same as requiring \( n_1 = rm_2/(m_1, m_2) \) and \( n_2 = rm_1/(m_1, m_2) \) for some \( r \in \mathbb{N} \). Thus the diagonal contribution is
\[
\sum_{n_1 m_1 = n_2 m_2} = 2 \sum_{m_1, m_2 \leq q^\delta} \frac{a_{m_1} \overline{a}_{m_2}}{(m_1, m_2)} \sum_{r \geq 1} \frac{1}{r} V \left( \frac{r^2 m_1 m_2}{q(m_1, m_2)^2} \right).
\]
The innermost sum above can be evaluated by the calculation in [10] Lemma 4.1 and equals
\[
\frac{p-1}{2p} \left( \log \left( \frac{q(m_1, m_2)^2}{\pi m_1 m_2} \right) + C \right) + O \left( \left( \frac{q(m_1, m_2)^2}{m_1 m_2} \right)^{-1/2 + \epsilon} \right),
\]
where \( C \) is as in the statement of Theorem 1.2 (2) The main term above gives the main term of Theorem 1.2 (2) and the total error is less than
\[
q^\epsilon \sum_{m_1, m_2 \leq q^\delta} \frac{1}{(m_1, m_2)} \left( \frac{q(m_1, m_2)^2}{m_1 m_2} \right)^{-1} \ll q^{-\frac{1}{2} + \delta} \sum_{m_1, m_2 \leq q^\delta} \frac{1}{(m_1 m_2)^{1/2}} \ll q^{-\frac{1}{2} + \theta + \epsilon}.
\]

The off-diagonal. We now turn to the off-diagonal terms; we show that these are bounded by the error terms in Theorem 1.2 (2). By Lemma 2.3 and (2), we have that
\[
\sum_{n_1 m_1 \neq n_2 m_2} \ll q^\epsilon \sum_{m_1, m_2 \leq q^\delta} \sum_{n_1 m_1 \neq n_2 m_2} \frac{1}{(n_1 n_2 m_1 m_2)^{1/2}}.
\]
Write \( a = n_1 m_1 \) and \( b = n_2 m_2 \) and split the resulting sum above into dyadic intervals \( A \leq a < 2A \) and \( B \leq b < 2B \). To isolate the diagonal contribution and complete the proof of Theorem 1.2 (2),
it is sufficient to show that
\begin{equation}
\frac{q^e}{(AB)^{1/2}} \sum_{\substack{A \leq a < 2A \\ AB \leq b < 2B \\ (ab,p) = 1 \\ a^{p-1} \equiv b^{p-1} \mod p^{k-1} \\ a \neq b}} 1
\end{equation}
is less than the announced error term for
\begin{equation}
1 \leq AB \leq q^{1+2\theta+\epsilon}.
\end{equation}

This is easy to do in the “unbalanced range”,
\begin{equation}
\frac{A}{B} \notin (q^{\frac{1}{2}}, q^{\frac{1}{2}}),
\end{equation}
which we consider first. Indeed, assume by symmetry that $A/B \geq q^{\frac{1}{2}}$. For each of the $B$ choices for $b$, there are $O(A/q + 1)$ choices for $a$; hence there are $O(AB/q + B)$ choices for $a$ and $b$. Therefore (3.5) is bounded by
\begin{equation}
\frac{q^e}{(AB)^{1/2}} (\frac{AB}{q} + B) \ll q^{-\frac{1}{4}+\theta+\epsilon} + q^{1/2} \leq q^{-\frac{1}{4}+\theta+\epsilon} + q^{-\frac{1}{4}+\epsilon},
\end{equation}
as desired.

From now on, we assume that
\begin{equation}
q^{\frac{1}{4}} < \frac{A}{B} < q^{\frac{1}{2}}.
\end{equation}
The sum (3.5) can contain terms for which $a \equiv \pm b \mod p^{k-1}$ only if $2A > p^{k-1}$ or $2B > p^{k-1}$. In that case, by the same argument as in the “unbalanced range”, the contribution of terms with $a \equiv \pm b \mod p^{k-1}$ to (3.5) is bounded by
\begin{equation}
\ll \frac{q^e}{(AB)^{1/2}} \frac{AB}{q} \ll q^{-\frac{1}{4}+\theta+\epsilon}.
\end{equation}
We are thus reduced to bounding the sum
\begin{equation}
\frac{q^e}{(AB)^{1/2}} \sum_{\substack{A \leq a < 2A \\ AB \leq b < 2B \\ (ab,p) = 1 \\ a^{p-1} \equiv b^{p-1} \mod p^{k-1} \\ a \neq \pm b \mod p^{k-1}}} 1,
\end{equation}
assuming (3.6) and (3.8). This is the key sum to which Lemmas 2.4 and 2.5 will be crucially applied. Letting $0 < \delta < \frac{1}{p}$ and $Q = q^{\frac{1}{2}-2\delta}$, we analyze (3.9) according to the sizes of $A$ and $B$ as follows.

**Case 1.** Suppose that $2A < q^{\frac{1}{4}-\delta}$ and $2B < q^{\frac{1}{4}-\delta}$. Every term in (3.9) would have to satisfy $a \equiv b\zeta_j \mod p^{k-1}$ for some $\zeta_j \in \mu_{p-1} \setminus \{\pm 1\}$, and hence $|a - b\zeta_j| \leq p^{k-1}$ with $|a|, |b| < (p^k)^{\frac{1}{2}-\delta}$. For sufficiently large $k$, however, by Lemma 2.4 there are no such nonzero integers $a$ and $b$, and the corresponding sum (3.9) is actually empty.
Case 2. Suppose that $2A \geq q^{\frac{3}{2} - \delta}$ and $2B < q^{\frac{1}{2} - \delta}$. (The case $2A < q^{\frac{3}{2} - \delta}$ and $2B \geq q^{\frac{1}{2} - \delta}$ is treated similarly.) Dividing the dyadic interval $A \leq a < 2A$ into smaller pieces of length $Q$, we may bound (3.9) by

\[
\frac{q^\epsilon}{(AB)^{\frac{3}{2}}} \sum_{1 \leq u \leq \frac{A}{4}} \sum_{B \leq b < 2B} \frac{1}{A+(u-1)Q \leq a < a+uQ} \frac{1}{B \leq b < 2B} \frac{1}{(ab,p) = 1} \frac{a^{p-1} \equiv b^{p-1} \mod{p^{k-1}}}{a \not\equiv \pm b \mod{p^{k-1}}}.
\]

By Lemma 2.5, the innermost sum of (3.10) is, for sufficiently large $k$, bounded by a constant (depending on $p$). Using this fact and (3.8), we see that (3.10) is bounded by

\[
\frac{q^\epsilon}{(AB)^{\frac{3}{2}}} A \leq q^{\frac{3}{4} + 2\delta + \epsilon}.
\]

This falls into the error term of Theorem 1.2 (2) as we may take $\delta$ to be as small as we like.

Case 3. Suppose that $2A \geq q^{\frac{3}{2} - \delta}$ and $2B \geq q^{\frac{1}{2} - \delta}$. Dividing the dyadic intervals $A \leq a < 2A$ and $B \leq b < 2B$ into smaller pieces of length $Q$, we may rewrite (3.9) as

\[
\frac{q^\epsilon}{(AB)^{\frac{3}{2}}} \sum_{1 \leq u \leq \frac{A}{4}} \sum_{1 \leq v \leq \frac{B}{4}} \frac{1}{A+(u-1)Q \leq a < a+uQ} \frac{1}{B+(v-1)Q \leq b < b+vQ} \frac{1}{(ab,p) = 1} \frac{a^{p-1} \equiv b^{p-1} \mod{p^{k-1}}}{a \not\equiv \pm b \mod{p^{k-1}}}.
\]

By Lemma 2.5, the innermost sum of (3.10) is, for sufficiently large $k$, bounded by a constant (depending on $p$). Using this fact and (3.6), we see that (3.12) is bounded by

\[
\frac{q^\epsilon}{(AB)^{\frac{3}{2}}} AB \leq q^{\frac{1}{4} + \theta + 4\delta + \epsilon}.
\]

This falls into the error term of Theorem 1.2 (2) as we may take $\delta$ to be as small as we like.

The proof of Theorem 1.2 (2) is now complete. □

3.4. The case of thin orbits. In this section, we prove Theorem 1.3, including evaluating the first and second mollified moments over thin Galois orbits in Theorem 3.2.

As is customary in analytic number theory (and as is already the case with full Galois orbits of primitive characters), the principal change introduced by the shrinking family of characters in the orbit $O_\kappa$ is that more terms survive averaging over the family. We quantify this effect with the following modification of Lemma 2.3 on orthogonality relations.

Lemma 3.1. Let $q = p^k$ for an odd prime $p$, let $0 < \kappa \leq k - 1$, and let $O_\kappa$ be a thin Galois orbit of primitive Dirichlet characters mod $q$. For any integer $n$, we have that

\[
\sum_{\chi \in O_\kappa} \chi(n) = 0
\]

unless

\[
n^{p-1} \equiv 1 \mod{p^{\tilde{\kappa} + 1}},
\]

where $\tilde{\kappa} = \min(\kappa, k - 2)$. 


Theorem 3.2. Let \( q = p^k \) for a fixed odd prime \( p \), let \( k/2 < \kappa < k-1 \), and let \( O_\kappa \) be a thin Galois orbit of primitive Dirichlet characters mod \( q \). Then:

1. For \( 0 \leq \theta < 2\left(\frac{k}{k} - \frac{1}{2}\right) \) in (1.9), we have
   \[
   \frac{1}{|O_\kappa|} \sum_{\chi \in O_\kappa} L\left(\frac{1}{2}, \chi\right) M(\chi) = 1 + O\left(q^{-\frac{k}{2}+\epsilon} + q^{\frac{k}{2}+\theta - \frac{k}{2} + \epsilon}\right).
   \]

2. For \( 0 \leq \theta < \frac{k}{k} - \frac{1}{2} \) in (1.9), we have
   \[
   \frac{1}{|O_\kappa|} \sum_{\chi \in O_\kappa} |L\left(\frac{1}{2}, \chi\right)|^2 |M(\chi)|^2 = \frac{p-1}{p} \sum_{\substack{m,m_2 \leq q^\epsilon \\text{ for } (m_1,m_2) = 1}}^{m_1 \mid pm_2} \alpha_{m_1,m_2} \left( \log \left( \frac{q(m_1,m_2)^2}{\pi m_1 m_2} \right) + C \right) + O\left(q^{-\frac{k}{2}+\epsilon} + q^{\frac{k}{2}+\theta - \frac{k}{2} + \epsilon}\right),
   \]
   with notations as in Theorem 1.2 (2).

Proof. The proof follows the proof of Theorem 1.2 with Lemma 3.1 as the orthogonality relation in place of Lemma 2.3.

For the first mollified moment, we start by inserting the approximate functional equation as in (3.2). By Lemma 3.1, the character average isolates the main term, which comes from \( mn = 1 \) and is identical as before, and summands with \( mn = \zeta_j \) mod \( p^k \), \( s > 1 \), which we split into two terms, corresponding to the ranges \( 1 < s \leq p^k \) and \( p^k < s < q^{1+\lambda+\theta+\epsilon} \). By Lemma 2.4 for sufficiently large \( k \), all summands in the first sum satisfy \( s \geq (p^k)^{\frac{k}{2} - \delta} \), and in total the first error term is \( O\left((p^k)^{-\frac{k}{2}+\delta}\right) \). Writing \( s = \zeta_j + p^s r \), the second error term is similarly bounded by

\[
q^\epsilon \sum_{1 \leq r < q^{1+\lambda+\theta+\epsilon - k/k}} (p^s r)^{-\frac{1}{2}} \ll q^{1+\lambda+\theta+\epsilon - k/k}.\]
completing the proof of the first part of our theorem.

As for the second mollified moment, we start by expanding using the same functional equation (3.4). The main term arises from the diagonal terms, when \( n_1 m_1 = n_2 m_2 \), and is the same as above, while in the off-diagonal terms we are left with Lemma 3.1 with the sum as in (3.5), now subject to \( a^{p^{-1}} \equiv b^{p^{-1}} \mod p^\kappa \). In the “unbalanced range”, which we now take to be \( A/B \notin (p^{-\frac{k}{2}}, p^{\frac{\kappa}{2}}) \), the sum is bounded as in (3.7) by

\[
\frac{q^\theta}{(AB)^{\frac{\kappa}{2}}} \left( \frac{AB}{p^\kappa} + B \right) \ll q^{\frac{1}{2}+\theta-\frac{\kappa}{2}+\epsilon} + q^{-\frac{1}{2}+\epsilon}.
\]

Assuming from now on \( p^{-\frac{k}{2}} < A/B < p^{\frac{\kappa}{2}} \), the terms for which \( a \equiv \pm b \mod p^\kappa \) can appear only if \( 2A > p^\kappa \) or \( 2B > p^\kappa \), and the total contribution from these terms is bounded by

\[
\frac{q^\theta}{(AB)^{\frac{\kappa}{2}}} \frac{AB}{p^\kappa} \ll q^{\frac{1}{2}+\theta-\frac{\kappa}{2}+\epsilon}.
\]

It remains to consider the terms for which \( a^{p^{-1}} \equiv b^{p^{-1}} \mod p^\kappa \) but \( a \not\equiv \pm b \mod p^\kappa \) as in (3.9); we set \( Q = (p^\kappa)^{\frac{1}{2}-\delta} \) and consider three cases. If \( 2A < (p^\kappa)^{\frac{1}{2}-\delta} \) and \( 2B < (p^\kappa)^{\frac{1}{2}-\delta} \), then the sum is empty for sufficiently large \( k \) by Lemma 2.4. If \( 2A \geq (p^\kappa)^{\frac{1}{2}-\delta} \) and \( 2B < (p^\kappa)^{\frac{1}{2}-\delta} \), then by splitting the \( a \)-sum into intervals of length \( Q \) the sum is bounded as in (3.11) by

\[
\frac{q^\theta}{(AB)^{\frac{\kappa}{2}}} \frac{A}{Q} \ll q^{-\frac{1}{2}+\epsilon+2\delta+\epsilon},
\]

and, finally, if both \( 2A \geq (p^\kappa)^{\frac{1}{2}-\delta} \) and \( 2B \geq (p^\kappa)^{\frac{1}{2}-\delta} \), then the sum is bounded as in (3.13) by

\[
\frac{q^\theta}{(AB)^{\frac{\kappa}{2}}} \frac{AB}{Q^2} \ll q^{\frac{1}{2}+\theta-\frac{\kappa}{2}+4\delta+\epsilon},
\]

which completes the proof as we may take \( \delta \) as small as we wish.

Theorem 1.3 now follows from Theorem 3.2 in the same way in which Theorem 1.1 is deduced from Theorem 1.2 in section 3.1 by taking \( \theta \to \frac{k}{2} - \frac{1}{2} \).

**Acknowledgements.** Part of this work was done during the excellent National Science Foundation-supported 29th Automorphic Forms Workshop held at the University of Michigan, Ann Arbor; we wish to express our thanks to the organizers of the conference. The second author would like to thank the University of Melbourne, where he completed part of the work, for their support and excellent research atmosphere and facilities. The third author is grateful to Jeffrey C. Lagarias for his constant encouragement. Finally, we would like to thank the referee for a careful reading of our paper and for the constructive comments.

**References**