TOROIDAL FAMILIES AND AVERAGES OF L-FUNCTIONS, I

ÉTIENNE FOUVRY, EMMANUEL KOWALSKI, AND PHILIPPE MICHEL

Abstract. We initiate the study of certain families of L-functions attached to characters of subgroups of higher-rank tori, and of their average at the central point. In particular, we evaluate the average of the values $\mathrm{L}(\frac{1}{2},\chi^a)\mathrm{L}(\frac{1}{2},\chi^b)$ for arbitrary integers a and b when χ varies of Dirichlet characters of a prime modulus.

Dedicated to Henryk Iwaniec, avec respect, avec gratitude, avec admiration

1. Introduction

The modern idea of "family of L-functions" in analytic number theory crystallized in part in the work of Iwaniec and Sarnak [14] on Landau–Siegel zeros (see also [15] for a general survey of L-functions). A notable illustration of this idea is then found in their paper [16] concerning the mollified first and second moments of central values of Dirichlet L-functions.

Our goal is to introduce a type of family related to Dirichlet characters, which seems very natural, but has not been considered to the best of our knowledge.

We call these families *toroidal families*, as they arise from families of Dirichlet (or automorphic) characters of *algebraic tori* by restricting to those characters that belong to some *algebraic* subgroup of its group of characters. In the simplest example, this means that we fix a non-zero vector $(a,b)\in\mathbf{Z}^2$, and for any modulus $q\geqslant 1$, we consider the group of pairs (χ_1,χ_2) of Dirichlet characters modulo q such that $\chi_1^a\chi_2^b=1$, and we study the properties as $q\to+\infty$ of the L-functions $\mathrm{L}(s,\chi_1)$ and $\mathrm{L}(s,\chi_2)$ when (χ_1,χ_2) satisfies this property. As we will see, even this first case leads to interesting phenomena.

Our main result in this first paper is the asymptotic computation of the average of the associated product L-values, in the case of prime moduli.

Theorem 1.1 (Second toroidal moment)**.** *Let a and b be non-zero integers. There exists an absolute and effective constant $c>0$ such that, defining*

$$\delta=\frac{c}{|a|+|b|},$$

we have, for any prime number q, the asymptotic formulas:

(1) *If $a+b=0$, then*

$$\frac{1}{q-1}\sum_{\chi\,(\mathrm{mod}\,q)}\mathrm{L}(\tfrac{1}{2},\chi^a)\mathrm{L}(\tfrac{1}{2},\chi^{-a})=\log q+2\mathrm{C}+\mathrm{O}(q^{-\delta})$$

where C is a real number independent of a.[1]

Date: Sat 18$^{\text{th}}$ Mar, 2023.

2010 *Mathematics Subject Classification.* 11M06,11T23,11A07.

Key words and phrases. L-functions, toroidal family, Dirichlet characters, exponential sums over finite fields, toric congruences, moments of L-functions.

[1] See (7.2) for the value of C.

(2) *If $a+b\neq 0$, then*

$$\frac{1}{q-1}\sum_{\chi\,(\mathrm{mod}\,q)}\mathrm{L}(\tfrac{1}{2},\chi^{a})\mathrm{L}(\tfrac{1}{2},\chi^{b})=\alpha(a,b)+\mathrm{O}(q^{-\delta}),$$

where

$$\alpha(a,b)=\begin{cases}\zeta(\frac{|a|+|b|}{2(a,b)}) & \text{if } ab<0,\\ 1 & \text{if } ab>0.\end{cases}$$

Here is an interpretation of this result. First, if we take $b=-a$ with $a\geqslant 2$ and suppose that we consider the primes $q\equiv 1\,(\mathrm{mod}\,a)$, then the toroidal average of Theorem 1.1 can also be expressed as the sum

$$\frac{1}{q-1}\sum_{\chi\,(\mathrm{mod}\,q)}|\mathrm{L}(\tfrac{1}{2},\chi^{a})|^{2}=\sum_{\varrho^{a}=1\,(\mathrm{mod}\,q)}\frac{1}{q-1}\sum_{\chi\,(\mathrm{mod}\,q)}\chi(\varrho)|\mathrm{L}(\tfrac{1}{2},\chi)|^{2}. \tag{1.1}$$

Our result is then an analogue of the estimation of the second moment of critical values of Dirichlet L-functions twisted by $\chi(\varrho)$, which is the building block of the mollified or amplified second moment (see, again, [16]). The crucial difference from earlier works is that in (1.1), the value of ϱ is *not* a fixed integer, except if $\varrho=1$ (which gives the main term) or if $\varrho=-1$, when a is even (see also Remark 7.1, where we explain that one can also bound the further variant where ϱ is a root of an irreducible polynomial congruence of degree $\geqslant 2$).

The fact that the second moment has the same order of magnitude as the full average

$$\frac{1}{q-1}\sum_{\chi\,(\mathrm{mod}\,q)}|\mathrm{L}(\tfrac{1}{2},\chi)|^{2}\sim\log q$$

(the special case $a=-b=1$ of Theorem 1.1, which was first proved by Paley [28, Th. B]) indicates a kind of "fair distribution" of the large values of $\mathrm{L}(\frac{1}{2},\chi)$ among the characters of the form χ^{a}, or in other words, among those that are trivial on a-th roots of unity modulo q.

Moreover, comparing with the formula for $a+b\neq 0$, we conclude that the sizes of the values $\mathrm{L}(\frac{1}{2},\chi^{a})$ and $\mathrm{L}(\frac{1}{2},\chi^{b})$ are "uncorrelated" if $a\neq b$. In this respect, the result is reminiscent of the well-known fact that

$$\sum_{n\leqslant x}d(n)d(n+1)$$

is of order of magnitude $x(\log x)^{2}$ for x large, whereas

$$\sum_{n\leqslant x}d(n)^{2}$$

is of order of magnitude $x(\log x)^{3}$.

Remark 1.2. (1) The first part of Theorem 1.1 implies (and is essentially equivalent) to the following: for a fixed integer $a\geqslant 1$, as q varies among prime numbers, we have

$$\frac{1}{q-1}\sum_{\substack{\chi\,(\mathrm{mod}\,q)\\ \chi=\eta^{a}\text{ for some }\eta}}|\mathrm{L}(\tfrac{1}{2},\chi)|^{2}=\frac{1}{(a,q-1)}\Big(\log q+2\mathrm{C}\Big)+\mathrm{O}(q^{-\delta}),$$

i.e., a second moment formula over the family of characters which are a-th powers (indeed, a character which is an a-th power is the a-th power of $(a,q-1)$ characters).

(2) We have not attempted to obtain the best possible error terms available with current technology; our goal here is primarily to grasp the interesting new aspects of non-trivial toroidal families.

(3) Suppose that $a+b\neq 0$. In a standard way, Theorem 1.1 establishes the existence of many characters modulo q such that

$$\mathrm{L}(\tfrac{1}{2},\chi^a)\mathrm{L}(\tfrac{1}{2},\chi^b)\neq 0$$

(a simultaneous non-vanishing result). Indeed, applying Cauchy's inequality twice and simply the upper bound

$$\frac{1}{q-1}\sum_{\chi\,(\mathrm{mod}\,q)}|\mathrm{L}(\tfrac{1}{2},\chi^a)|^4\ll\frac{1}{q-1}\sum_{\chi\,(\mathrm{mod}\,q)}|\mathrm{L}(\tfrac{1}{2},\chi)|^4\ll(\log q)^4$$

for q prime (also proved first by Paley [28, Th. IV]), we see that

$$|\{\chi\,(\mathrm{mod}\,q)\ \mid\ \mathrm{L}(\tfrac{1}{2},\chi^a)\mathrm{L}(\tfrac{1}{2},\chi^b)\neq 0\}|\gg\frac{q}{(\log q)^4}.$$

In fact, the power-saving error term in Theorem 1.1 means that it is straightforward to apply the mollification method (as in [16], see also [3]) and prove that

$$|\{\chi\,(\mathrm{mod}\,q)\ \mid\ \mathrm{L}(\tfrac{1}{2},\chi^a)\mathrm{L}(\tfrac{1}{2},\chi^b)\neq 0\}|\gg q$$

for q prime, where the implicit constant depends on (a,b) (see the work of Zacharias [36] for some related results).

(4) We note that although "families of L-functions" still appear most often in an informal manner based on concrete examples, there are significant attempts to provide a precise definition (e.g., those of Sarnak, Shin and Templier [34] and Kowalski [20]).

In an automorphic perspective, a natural framework seems to be the following. Let K be a number field with ring of integers $\mathscr{O}$ and adèle ring $\mathbf{A}_{\mathrm{K}}$. Let $\mathbf{T}_1$ and $\mathbf{T}_2$ be two tori defined over $\mathscr{O}$, and assume they are split for simplicity. Let $\varphi\colon\mathbf{T}_1\to\mathbf{T}_2$ be an algebraic morphism defined over $\mathscr{O}$. Then $\chi\mapsto\chi\circ\varphi$ gives a map φ^* from characters of $\mathbf{T}_2(\mathbf{A}_{\mathrm{K}})/\mathbf{T}_2(\mathrm{K})$ to those of $\mathbf{T}_1(\mathbf{A}_{\mathrm{K}})/\mathbf{T}_1(\mathrm{K})$, i.e., from automorphic characters of one torus to the other.

Let $\mathrm{X}_*(\mathbf{T}_1)$ be the abelian group of algebraic characters of $\mathbf{T}_1$. The L-group of $\mathbf{T}_1$ is the complex dual torus $\mathrm{Hom}(\mathrm{X}_*(\mathbf{T}_1),\mathbf{C}^\times)$ of $\mathbf{T}_1$, and similarly for $\mathbf{T}_2$ (because we assumed that the tori are split).

Let now $\varrho\colon{}^{\mathrm{L}}\mathbf{T}_1\to\mathbf{GL}_r$ be a faithful continuous representation. By the Langlands formalism, this defines for any automorphic character χ of $\mathbf{T}_1$ the associated L-function $\mathrm{L}(s,\chi,\varrho)$. Thus, varying χ over automorphic characters of $\mathbf{T}_2$, we obtain a family of L-functions $\mathrm{L}(s,\varphi^*\chi,\varrho)$. These are general forms of toroidal families.

To illustrate this construction, let $\mathrm{K}=\mathbf{Q}$, $\mathbf{T}_1=\mathbf{G}_m^2$, $\mathbf{T}_2=\mathbf{G}_m$ and $\varphi(x,y)=x^ay^b$. Then ${}^{\mathrm{L}}\mathbf{T}_1\simeq(\mathbf{C}^\times)^2$ and we can take the embedding $\varrho\colon{}^{\mathrm{L}}\mathbf{T}_1\to\mathbf{GL}_2$ of ${}^{\mathrm{L}}\mathbf{T}_1$ as diagonal matrices. For any $d\geqslant 1$, the automorphic characters of $\mathbf{G}_m^d$ over $\mathbf{Q}$ are simply d-tuples of (primitive) Dirichlet characters; the map φ^* then maps a primitive Dirichlet character χ modulo q to the pair of Dirichlet characters (χ^a,χ^b) (modulo primitivity issues), and we have

$$\mathrm{L}(s,\varphi^*\chi,\varrho)=\mathrm{L}(s,\chi^a)\mathrm{L}(s,\chi^b),$$

which recovers the type of L-functions in Theorem 1.1.

We note that the set of automorphic characters of a torus with bounded conductor is amenable to analytic investigations in (more than) the generality we have described, due to the remarkable recent work of Petrow [29]. Thus some questions about toroidal families should be accessible in this generality.

The outline of the remainder of this paper is the following: we will first discuss some natural questions and generalizations of toroidal averages, then consider separately some ingredients that should arise in their evaluation in all cases, namely properties of certain exponential sums and counting solutions of toric congruences, which (for later purposes) we discuss in a broader context than needed in the present paper (see Remark 5.2). We then give the proof of Theorem 1.1. After some elementary common considerations, this splits naturally in two cases, depending on whether $a + b = 0$ or not.

Notation. Given complex-valued functions f and g defined on a set X, we write $f \ll g$ if there exists a real number $\mathrm{A} \geqslant 0$ (called an "implicit constant") such that the inequality $|f(x)| \leqslant \mathrm{A}g(x)$ holds for all $x \in \mathrm{X}$. We write $f \asymp g$ if $f \ll g$ and $g \ll f$.

For any prime q, we denote by $\widehat{\mathbf{F}}_q^\times$ the group of Dirichlet characters modulo q (in other words, the group of characters of $(\mathbf{Z}/q\mathbf{Z})^\times$).

For a Dirichlet character χ modulo a prime number q, we denote by

$$\varepsilon(\chi) = \frac{1}{\sqrt{q}} \sum_{x \in \mathbf{F}_q^\times} \chi(x) e\Big(\frac{x}{q}\Big),$$

the normalized Gauss sums, so that $|\varepsilon(\chi)| = 1$ if χ is non-trivial. More generally, for an integer $k \geqslant 1$ and a k-tuple $\boldsymbol{\chi} = (\chi_1, \ldots, \chi_k)$ of characters modulo q, we put

$$\varepsilon(\boldsymbol{\chi}) = \prod_{j=1}^{k} \varepsilon(\chi_j).$$

Let K be a field. For an integer $k \geqslant 1$, we write $x \cdot y = x_1 y_1 + \cdots + x_k y_k$ for the standard bilinear form on K^k. For any integer $d \geqslant 1$, we denote by $\boldsymbol{\mu}_d(\mathrm{K})$ the group of d-th roots of unity which belong to K.

Let k be an integer. A *box* in $\mathbf{Z}^k$ is a product $\mathrm{B} = \mathrm{I}_1 \times \cdots \times \mathrm{I}_k$ where each $\mathrm{I}_j = \{a_j, \ldots, b_j\}$ is an interval in $\mathbf{Z}$ with $a_j \leqslant b_j$. By the size of the box, we mean simply its cardinality, and the boundary $\partial \mathrm{B}$ is the union of the sets

$$\mathrm{I}_1 \times \cdots \times \mathrm{I}_{j-1} \times \{a_j\} \times \mathrm{I}_{j+1} \times \cdots \times \mathrm{I}_k, \qquad \mathrm{I}_1 \times \cdots \times \mathrm{I}_{j-1} \times \{b_j\} \times \mathrm{I}_{j+1} \times \cdots \times \mathrm{I}_k$$

for $1 \leqslant j \leqslant k$.

Acknowledgments. This work was started during a visit of É.F. and Ph.M. to the Forschungssinstitut für Mathematik (FIM) at ETH Zürich, and continued during a visit of É.F. and E.K. at the Bernoulli Center for Fundamental Studies of EPF Lausanne. We all thank these institutions for the excellent and stimulating work conditions.

We thank K. Soundararajan for pointing out to us the related works of Pliego [31, 32] and of Khan–Milićević–Ngo [19] (see Section 2 below).

Ph. M. was partially supported by the SNF grant 200021_197045.

2. Prospective and related works

This section may be skipped in a first reading. We will list some obvious variants and potential generalizations of toroidal families, some of which we hope to study in later papers. See also the work of Blomer, Fouvry, Milićević, Kowalski, Michel and Sawin [3] for a relatively systematic discussion of problems (and results) about families of L-functions, many parts of which could be adapted to our context.

(1) A general form of toroidal averages is the following. Fix $k \geqslant 2$. Consider an integral matrix $\mathrm{A} = (a_{i,j})$ with k columns (and an arbitrary finite number of rows), and for any prime number q, define

$$\widehat{\mathrm{H}}_{\mathrm{A}}(q) = \{(\chi_1, \ldots, \chi_k) \pmod q \mid \prod_{j=1}^{k} \chi_j^{a_{i,j}} = 1 \text{ for all } i\},$$

a subgroup of the group of characters of $(\mathbf{F}_q^\times)^k$. We want to evaluate asymptotically

$$\sum_{\boldsymbol{\chi} \in \widehat{\mathrm{H}}_{\mathrm{A}}(q)} \prod_{j=1}^{k} \mathrm{L}(\tfrac{1}{2}, \chi_j)$$

for q prime. The case of Theorem 1.1 is essentially the case $k = 2$ with $\mathrm{A} = (-b, a)$. An analogue with $k = 3$ would be the study of a sum like

$$\sum_{\chi_1, \chi_2 \pmod q} \mathrm{L}(\tfrac{1}{2}, \chi_1^a) \mathrm{L}(\tfrac{1}{2}, \chi_2^b) \mathrm{L}(\tfrac{1}{2}, \chi_1^c \chi_2^d).$$

Besides the work in the present paper, we are only aware of one previous case which was considered by Nordentoft [25], and concerns the subgroup defined (for fixed $k \geqslant 2$) by $\chi_1 \cdots \chi_k = 1$.

(2) In addition to the characters in some subgroup $\widehat{\mathrm{H}}$ of tuples of Dirichlet characters modulo q, it is natural to consider those in a *coset* modulo such a subgroup. This means that in addition to the matrix A, we also consider for any prime q some characters $(\eta_1, \ldots, \eta_k)$ modulo q (arbitrarily chosen) and we look at the averages

$$\sum_{\boldsymbol{\chi} \in \widehat{\mathrm{H}}_{\mathrm{A}}(q)} \prod_{j=1}^{k} \mathrm{L}(\tfrac{1}{2}, \eta_j \chi_j).$$

An example of this type with $k = 3$ and $\mathrm{A} = (1, 1, 1)$ has been considered by Zacharias [36], and Nordentoft [25] also considers such twists in his special case.

(3) We focus here for simplicity on groups of characters which are dual groups of the groups of rational points on a torus which is split over $\mathbf{F}_q$. It is likely that non-split cases would also lead to interesting questions. See [24] and [26] for works in that direction.

(4) Similar questions can be raised for twisted L-functions of various kinds, for instance

$$\frac{1}{q-1} \sum_{\chi \pmod q} \mathrm{L}(\tfrac{1}{2}, f \otimes \chi^a) \mathrm{L}(\tfrac{1}{2}, g \otimes \chi^b) \tag{2.1}$$

for some modular forms f and g, or the mixed moment

$$\frac{1}{q-1}\sum_{\chi\,(\mathrm{mod}\,q)} \mathrm{L}(\tfrac{1}{2},\chi^a)\mathrm{L}(\tfrac{1}{2},f\otimes\chi^b). \tag{2.2}$$

There are currently very few pairs (a,b) for which these moments have been evaluated asymptotically; these are

$$(a,b)=(1,-1),\ (2,-2)$$

in the case of (2.1) (see [3, 21]), and

$$(a,b)=(1,-1),(2,-2),(1,1),\ (2,2)$$

in the case of (2.2) (see [6, 36]). Techniques inspired by [22] should make it possible to deal with more pairs (a,b) for the mixed moment (2.2) (see Remark 6.1).

One can also of course study averages with derivatives of central values, and one might look at values $\mathrm{L}(s,\chi)$ of the L-functions where $s\neq\frac{1}{2}$, or combine these variants.

(5) A natural archimedean analogue of our problem consists in studying

$$\int_{\mathrm{T}}^{2\mathrm{T}} \zeta(\tfrac{1}{2}+a_1 it)\cdots\zeta(\tfrac{1}{2}+a_k it)dt$$

for fixed $(a_1,\ldots,a_k)\in(\mathbf{R}^{\times})^k$. Some work has indeed been done by Pliego on this question in the case of three factors (see [31] and [32]). Some rather interesting phenomena related to diophantine approximation appear in his analysis.

(6) Yet another question would be to study the *distribution* of pairs

$$(\mathrm{L}(1,\chi^a),\mathrm{L}(1,\chi^b))\in\mathbf{C}^2,$$

when χ varies. As $q\to+\infty$, one will obtain a limiting distribution given by the law of the random vector

$$\Big(\prod_p(1-\mathrm{U}_p^a/p)^{-1},\prod_p(1-\mathrm{U}_p^b/p)^{-1}\Big),$$

where $(\mathrm{U}_p)_p$ is a sequence, indexed by primes, of independent random variables all equidistributed on the unit circle. Because of the dependency of the two components, it would be interesting to have, for instance, estimates for large values, large deviations, and related statistics, comparable to those known for $\mathrm{L}(1,\chi)$ (see the paper [23] of Lamzouri).

Similar properties hold for the values at any s with $\mathrm{Re}(s)>1/2$. At $s=1/2$, one can expect a Central Limit Theorem for

$$\frac{\log|\mathrm{L}(\frac{1}{2},\chi^a)\mathrm{L}(\frac{1}{2},\chi^b)|}{\sqrt{\log\log q}}$$

(but note that it remains a major open problem to obtain such a result even for $\log|\mathrm{L}(\frac{1}{2},\chi)|$; Radziwiłł and Soundararajan [33] have however announced a proof of a result of this type when conditioning on characters with $\mathrm{L}(\frac{1}{2},\chi)\neq 0$). Note that if we consider the pairs $(\log|\mathrm{L}(\frac{1}{2},\chi^a)|,\log|\mathrm{L}(\frac{1}{2},\chi^b)|)$, we can expect that after renormalization, the two components will become independent gaussians if $a\neq -b$ (formally, because

$$\mathbf{E}\Big(\sum_{p\leqslant \mathrm{X}}\frac{\mathrm{U}_p^a}{\sqrt{p}}\cdot\overline{\sum_{p\leqslant \mathrm{X}}\frac{\mathrm{U}_p^b}{\sqrt{p}}}\Big)=\sum_{p\leqslant \mathrm{X}}\frac{1}{p}\mathbf{E}(\mathrm{U}_p^{a+b})=0$$

by independance, which will imply that the limits should be uncorrelated gaussians).

(7) We have considered prime moduli q for simplicity, but more general moduli can also be investigated. Moreover, other families of Dirichlet characters can be used to build toroidal averages, e.g., we can average as $\mathrm{Q}\to+\infty$ over pairs (χ^a,χ^b) as χ ranges over all primitive Dirichlet characters with conductor $\leqslant \mathrm{Q}$. This may open the door to higher values of k, as happens with the moments of $\mathrm{L}(\frac{1}{2},\chi)$ (see for instance the work [5] of Conrey, Iwaniec and Soundararajan, and the recent work [4] of Chandee and Li).

The paper [19] of Khan, Milićević and Ngo can be understood as somewhat similar: roughly speaking, they consider (mollified) first and second moments of the type

$$\sum_{\chi\in\mathcal{O}}\mathrm{L}(\tfrac{1}{2},\chi),\qquad \sum_{\chi\in\mathcal{O}}|\mathrm{L}(\tfrac{1}{2},\chi)|^2$$

where $\mathcal{O}$ is a (suitably large) Galois orbit of characters modulo p^k for a fixed prime number p and $k\to+\infty$ (in the simplest case, this amounts to characters of a fixed order, large enough that the orbit is big enough for averaging). The key new arithmetic ingredient they use is related to p-adic diophantine approximation [19, § 2.4].

One can also look, from the point of view of changing moduli, at the toroidal family of primitive characters defined, e.g., by the equation $\chi^2=1$, recovering the well-known family of real characters of bounded conductor.

(8) We have worked with the base field $\mathbf{Q}$ for simplicity, but any other global field leads to similar questions. Function fields should, as usual, be somewhat more accessible, especially in the large finite field limit.

3. Exponential sums

The first subsection can be mostly skipped in a first reading, since the exponential sums relevant for the proof of Theorem 1.1 are treated from scratch in Section 3.2.

3.1. **General sums.** We consider a more general situation for future reference. Let $k\geqslant 1$ be an integer. For any prime number q and any subgroup H of $(\mathbf{F}_q^\times)^k$, we consider families of exponential sums of the type

$$u\mapsto \sum_{(x_j)\in u\mathrm{H}} e\Big(\frac{x_1+\cdots+x_k}{q}\Big)$$

for $u\in(\mathbf{F}_q^\times)^k$. Special cases of these sums include the hyper-kloosterman sums Kl_k, which correspond to the subgroup H defined by $x_1\cdots x_k=1$: we then have

$$\sum_{(x_j)\in u\mathrm{H}} e\Big(\frac{x_1+\cdots+x_k}{q}\Big)=\sum_{x_1\cdots x_k=u_1\cdots u_k} e\Big(\frac{x_1+\cdots+x_k}{q}\Big).$$

In fact, we observe that the condition $x \in u\mathrm{H}$ is equivalent to $x_j = u_j h_j$ for all j with $(h_j) \in \mathrm{H}$, so that for $u \in (\mathbf{F}_q^\times)^k$, we can also write

$$\sum_{(x_j)\in u\mathrm{H}} e\Big(\frac{x_1+\cdots+x_k}{q}\Big) = \sum_{h\in\mathrm{H}} e\Big(\frac{u_1h_1+\cdots+u_kh_k}{q}\Big).$$

This formula identifies our family of exponential sums as the restriction to $(\mathbf{F}_q^\times)^k$ of the discrete Fourier transform of the characteristic function of $\mathrm{H} \subset \mathbf{F}_q^k$. It is then most convenient to extend the definition to all $u \in \mathbf{F}_q^k$ by the last formula, as this allows us to use the standard formalism of Fourier analysis.

Finally, it is convenient to normalize these sums before giving them a name. There are various possibilities, and our choice is based on obtaining generically sums of size one.

More precisely, our interest lies in situations where q varies and H is for each q the group of $\mathbf{F}_q$-points of a fixed algebraic subgroup $\mathbf{H}$ of $\mathbf{G}_m^k$ over $\mathbf{Z}$. Concretely, this means that we fix an integer matrix $\mathrm{A} = (a_{i,j})$ with k columns (and a finite number of rows), and consider the subgroup

$$\mathrm{H}_\mathrm{A}(\mathbf{F}_q) = \{x = (x_j) \mid \prod_j x_j^{a_{i,j}} = 1 \text{ for all } i\} \tag{3.1}$$

for all prime numbers q. For instance, taking $\mathrm{A} = (1, \ldots, 1) \in \mathbf{Z}^k$, we recover the subgroup which gives rise to Kl_k.

We then define

$$\mathrm{T}_\mathrm{A}(u;q) = \frac{1}{q^{(k-\mathrm{rank(A)})/2}} \sum_{(x_j)\in \mathrm{H}_\mathrm{A}(\mathbf{F}_q)} e\Big(\frac{x\cdot u}{q}\Big)$$

for $u \in \mathbf{F}_q^k$ (note that $\dim(\mathbf{H}) = k - \mathrm{rank(A)}$; in the case of the hyper-Kloosterman sums, these sums are bounded for all $u \in (\mathbf{F}_q^\times)^k$ by Deligne's estimate).

Furthermore, we define

$$\mathrm{H}_\mathrm{A}^\perp(\mathbf{F}_q) = \{\boldsymbol{\chi} = (\chi_j) \in (\widehat{\mathbf{F}}_q^\times)^k \mid \boldsymbol{\chi}(x) = \prod_{j=1}^k \chi_j(x_j) = 1 \text{ for all } x = (x_j) \in \mathrm{H}_\mathrm{A}(\mathbf{F}_q)\}. \tag{3.2}$$

Proposition 3.1. *Let* A *be an integral matrix as above. For any prime number* q *and for any* $u \in (\mathbf{F}_q^\times)^k$*, we have*

$$\frac{1}{|\mathrm{H}_\mathrm{A}^\perp(\mathbf{F}_q)|} \sum_{\boldsymbol{\chi}\in \mathrm{H}_\mathrm{A}^\perp(\mathbf{F}_q)} \varepsilon(\boldsymbol{\chi})\overline{\boldsymbol{\chi}(u)} = \frac{1}{q^{\mathrm{rank(A)}/2}} \mathrm{T}_\mathrm{A}(u;q).$$

Proof. This is a direct consequence of the definitions, and of orthogonality, in the form of the relation

$$\frac{1}{|\mathrm{H}_\mathrm{A}^\perp(\mathbf{F}_q)|} \sum_{\boldsymbol{\chi}\in \mathrm{H}_\mathrm{A}^\perp(\mathbf{F}_q)} \boldsymbol{\chi}(xu^{-1}) = \begin{cases} 1 & \text{if } xu^{-1} \in \mathrm{H}_\mathrm{A}(\mathbf{F}_q), \\ 0 & \text{otherwise,} \end{cases}$$

which follows from the duality of finite abelian groups. □

We are interested in bounds for the exponential sums $\mathrm{T}_\mathrm{A}(u;q)$. The mean-square bound is elementary, but the best pointwise bounds lie obviously much deeper, and in fact we only state these in the case of a non-negative matrix, in the "stratified form" going back to Katz and Laumon (see also [7] for other applications of such stratification results).

Theorem 3.2. *Let* A *be an integral matrix as above.*

(1) *For any* q *prime, we have*

$$\frac{1}{q^k}\sum_{u\in\mathbf{F}_q^k}|\mathrm{T}_{\mathrm{A}}(u;q)|^2=\frac{|\mathrm{H}_{\mathrm{A}}(\mathbf{F}_q)|}{q^{k-\mathrm{rank}(\mathrm{A})}},$$

and this quantity converges as $q\to+\infty$ *to the number of connected components of* $\mathbf{H}_{\mathrm{A}}$*.*

(2) *Suppose that* $a_{i,j}\geqslant 0$ *for all* i*,* j *and that the gcd of all* $a_{i,j}$ *is one. There exists a dense open set* $\mathrm{U}\subset\mathbf{A}^k_{\mathbf{Z}}$ *such that*

$$|\mathrm{T}_{\mathrm{A}}(u;q)|\ll 1$$

for all primes q *and all* $u\in\mathrm{U}(\mathbf{F}_q)$*, where the implicit constant depends only on* rank(A)*.*

More precisely, there exist closed subschemes $\mathrm{X}_k\subset\cdots\subset\mathrm{X}_1\subset\mathbf{A}^k_{\mathbf{Z}}$*, defined by homogeneous equations, with* X_j *of relative dimension* $\leqslant k-j$*, such that for* q *prime and* $u\in(\mathbf{F}_q^\times)^k$ *with* $u\notin\mathrm{X}_j(\mathbf{F}_q)$*, then we have*

$$|\mathrm{T}_{\mathrm{A}}(u;q)|\ll q^{\max(0,(j-2)/2)}.$$

Proof. (1) is a direct consequence of the discrete Plancherel formula and the fact that the cardinality of $\mathrm{H}_{\mathrm{A}}(\mathbf{F}_q)$ is $\sim q^{\dim(\mathbf{H}_{\mathrm{A}})}$ as $q\to+\infty$ if $\mathbf{H}_{\mathrm{A}}$ is connected (an elementary fact in our setting where we consider split tori).

To prove (2), note that the second statement implies the first with U the complement of X_1. To establish the latter, we need to invoke algebraic geometry, in the form of the result of [8, Th. 1.2] of Fouvry and Katz. We apply this to the algebraic variety $\mathbf{H}_{\mathrm{A}}$; since A has non-negative coefficients, the equations defining it in $\mathbf{A}^k$ are polynomial equations (by opposition with Laurent polynomials), so that $\mathbf{H}_{\mathrm{A}}$ is a closed subscheme of $\mathbf{A}^k_{\mathbf{Z}}$. It is an algebraic torus, and is connected because of the gcd assumption,[2] so $\mathbf{H}_{\mathrm{A}}(\mathbf{C})$ is smooth and irreducible. Finally, it follows from [8, Th. 8.1] that the so-called A-number of $\mathbf{H}_{\mathrm{A}}$ is non-zero (because $\mathbf{H}_{\mathrm{A}}$ is a torus, so that $|\mathbf{H}_{\mathrm{A}}(\mathbf{F}_q)|$ is coprime to q for all q). Thus all assumptions of the result of Fouvry and Katz are satisfied. ☐

Remark 3.3. For future reference, we explain the origin of the exponential sums as trace functions in general. We work over the finite field $\mathbf{F}_q$ for a fixed prime q. Pick a prime number $\ell\neq q$ and identify $\overline{\mathbf{Q}}_\ell$ with $\mathbf{C}$ by some fixed isomorphism. Let ψ be the additive $\overline{\mathbf{Q}}_\ell$-valued character of $\mathbf{F}_q$ which is then identified with $x\mapsto e(x/q)$.

The injective map $i\colon\mathbf{H}_{\mathrm{A}}\to\mathbf{A}^k$ is a locally closed immersion (it is a closed immersion if all $a_{i,j}$ are non-negative). The formula defining $\mathrm{T}_{\mathrm{A}}(u;q)$ then shows that it coincides (up to a sign depending only on A) with the trace function over $\mathbf{F}_q$ of Deligne's ℓ-adic Fourier transform $\mathrm{F}_{\mathrm{A}}=\mathrm{FT}_\psi(\mathrm{S}_{\mathrm{A}})$, where $\mathrm{S}_{\mathrm{A}}=(i_!\overline{\mathbf{Q}}_\ell)[\dim(\mathbf{H}_{\mathrm{A}})](\dim(\mathbf{H}_{\mathrm{A}}))$.

Since i is an affine quasi-finite morphism, the object S_{A} is a perverse sheaf on $\mathbf{A}^k$ (see [1, Cor. 4.1.3]). According to the formalism of Deligne's Fourier transform, this implies that F_{A} is also a perverse sheaf (see [18, Cor. 2.1.5]);

Suppose now that $a_{i,j}\geqslant 0$ for all i and j, and that the gcd of the $a_{i,j}$ is one (as in the previous theorem). Then $\mathbf{H}_{\mathrm{A}}$ is (geometrically) connected, and the morphism i is a closed immersion, so the perverse sheaf S_{A} is in addition simple (see [1, Th. 4.3.1]) and pure of weight 0 (because $i_*=i_!$ and because of Deligne's Riemann Hypothesis, see [1, 5.1.14]), and

[2] In general, the number of connected components of $\mathbf{H}_{\mathrm{A}}$ is the greatest common divisor of all components of the matrix A.

so is F_{A} (see [18, Th. 2.2.1]); note that this means that we normalized the Fourier transform by a twist so that it preserves weights, in contrast with [18].

Dually, we may construct using the matrix A the subgroups $\mathrm{X}_{\mathrm{A}}(\mathbf{F}_q)$ of $(\widehat{\mathbf{F}}_q^\times)^k$ defined by the condition that $\boldsymbol{\chi} \in \mathrm{X}_{\mathrm{A}}(\mathbf{F}_q)$ if

$$\prod_{j=1}^{k} \chi_j^{a_{i,j}} = 1$$

for all i. We can then form the subgroups

$$\mathrm{X}_{\mathrm{A}}^{\perp}(\mathbf{F}_q) = \{x = (x_j) \in (\mathbf{F}_q^\times)^k \mid \boldsymbol{\chi}(x) = 1 \text{ for all } \boldsymbol{\chi} \in \mathrm{X}_{\mathrm{A}}(\mathbf{F}_q)\}.$$

Elementary properties of the duality of finite abelian groups shows that there exists a matrix B such that $\mathrm{X}_{\mathrm{A}}^{\perp}(\mathbf{F}_q) = \mathrm{H}_{\mathrm{B}}(\mathbf{F}_q)$ for all q, and then $\mathrm{H}_{\mathrm{B}}^{\perp}(\mathbf{F}_q) = \mathrm{X}_{\mathrm{A}}(\mathbf{F}_q)$ for all q. Thus we can pass back and forth between subgroups of $(\mathbf{F}_q^\times)^k$ and associated subgroups of $(\widehat{\mathbf{F}}_q^\times)^k$.

Example 3.4. The example of Nordentoft [25] arises from the vector $\mathrm{A} = (1, \ldots, 1)$ to construct $\mathrm{X}_{\mathrm{A}}(\mathbf{F}_q)$ which is the group of characters with $\chi_1 \cdots \chi_k = 1$. Then $\mathrm{X}_{\mathrm{A}}^{\perp}(\mathbf{F}_q)$ is the diagonal subgroup $\{(x, \ldots, x)\} \subset (\mathbf{F}_q^\times)^k$ and corresponds to a matrix B with $\mathrm{rank}(\mathrm{B}) = k-1$. The corresponding exponential sums are given by

$$\mathrm{T}_{\mathrm{B}}(u; q) = \frac{1}{\sqrt{q}} \sum_{x \in \mathbf{F}_q^\times} e\Big(\frac{x(u_1 + \cdots + u_k)}{q}\Big),$$

and are therefore elementary (Ramanujan sums modulo q). Similarly, in the (slightly different) setting of of Khan, Milićević and Ngo, only elementary exponential sums arise.

In practice (this is the case in particular in Theorem 1.1) we often start with an integer d and an integral matrix $\mathrm{B} = (b_{j,i})$ with d columns and k rows, and consider the subgroups of characters

$$\widehat{\mathrm{H}}(\mathbf{F}_q) \subset (\widehat{\mathbf{F}}_q^\times)^k$$

defined as the image of the map

$$\varphi_{\mathrm{B}} \colon (\chi_1, \ldots, \chi_d) \mapsto \Big(\prod_{i=1}^{d} \chi_i^{b_{i,1}}, \cdots, \prod_{i=1}^{d} \chi_i^{b_{i,k}}\Big).$$

It is an elementary fact that there exists a (fixed) integral matrix A such that

$$\widehat{\mathrm{H}}(\mathbf{F}_q) = \widehat{\mathrm{H}}_{\mathrm{A}}(\mathbf{F}_q)$$

for all q. Note that the map φ_{B} is not necessarily injective, so that summing over $(\chi_1, \ldots, \chi_d)$ is not equivalent to summing over the image of φ_{B}. However, since the number of pre-images of a character $\eta \in \varphi_{\mathrm{B}}((\widehat{\mathbf{F}}_q^\times)^d)$ is independent of η, this is only a cosmetic issue.

3.2. **The case of Theorem 1.1.** We take $k = 2$ and $\mathrm{A} = (a, b) \in \mathbf{Z}^2$, $ab \neq 0$. The group $\mathbf{H}_{\mathrm{A}}$ is defined by the equation $x^a y^b = 1$, and it has dimension 1; it is connected if and only if a and b are coprime.

For any prime number q, we define the exponential sums $\mathrm{T}_{a,b}(u, v; q)$ by

$$\mathrm{T}_{a,b}(u, v; q) = \frac{1}{\sqrt{q}} \sum_{\substack{x, y \in \mathbf{F}_q^\times \\ x^a y^b = 1}} e\Big(\frac{ux + vy}{q}\Big),$$

for u, $v \in \mathbf{F}_q$, and we define

$$\widetilde{\mathrm{T}}_{a,b}(u;q) = \frac{1}{\sqrt{q}} \sum_{\substack{x,y\in\mathbf{F}_q^\times \\ x^a y^b = u}} e\Big(\frac{x+y}{q}\Big),$$

for $u \in \mathbf{F}_q^\times$, so that $\mathrm{T}_{a,b}(u,v;q) = \widetilde{\mathrm{T}}_{a,b}(u^a v^b;q)$ for u and v in $\mathbf{F}_q^\times$.

The relation with Gauss sums can be expressed here by the formula

$$\widetilde{\mathrm{T}}_{a,b}(u;q) = \frac{\sqrt{q}}{q-1} \sum_{\chi \,(\mathrm{mod}\, q)} \varepsilon(\chi^a)\varepsilon(\chi^b)\bar{\chi}(u) \tag{3.3}$$

for all $u \in \mathbf{F}_q^\times$.

We note in passing that if $q \geqslant 3$, then we also have

$$\widetilde{\mathrm{T}}_{a,b}(u;q) + \widetilde{\mathrm{T}}_{a,b}(-u;q) = \widetilde{\mathrm{T}}_{2a,2b}(u^2;q) \tag{3.4}$$

for $u \in \mathbf{F}_q^\times$.

The algebraic structure of these exponential sums is different depending on whether $a+b = 0$ or not. In the former case, we have a simple result.

Proposition 3.5. *Suppose that $a + b = 0$. For any $u \in \mathbf{F}_q^\times$, we have*

$$\widetilde{\mathrm{T}}_{a,b}(u^a;q) = \begin{cases} \sqrt{q} - \frac{(a,q-1)}{\sqrt{q}} & \text{if } u^a = 1 \\ -\frac{(a,q-1)}{\sqrt{q}} & \text{if } u^a \neq 1. \end{cases}$$

Proof. It is simpler here to use the formula (3.3) and the fact that $\varepsilon(\chi^a)\varepsilon(\chi^{-a})$ is equal to 1 if $\chi^a \neq 1$ and to $1/q$ if $\chi^a = 1$. The result then follows easily. ☐

Remark 3.6. From the point of view of trace functions, this case is actually rather delicate, since one should see $\widetilde{\mathrm{T}}_{a,-a}(u;q)$ as the trace function of a perverse object, and not of a single sheaf, and this object is neither simple nor pure (see [17, Lemma 8.4.8] for what amounts to the case $a = -1$).

The case $a + b \neq 0$ gives rise to more interesting exponential sums. We can view them as trace functions again, but the main property that we will use also has a more elementary proof.

Proposition 3.7. *Suppose that $a + b \neq 0$. We have*

$$|\widetilde{\mathrm{T}}_{a,b}(u;q)| \leqslant |a| + |b|$$

for all primes $q \geqslant \max(|a|,|b|)^2$ and all $u \in \mathbf{F}_q^\times$. Additionally, the bound above can be improved to $\max(|a|,|b|)$ if a and b have different signs.

Proof. Let $\delta = (a,b)$ and pick integers α, β such that

$$\beta a + \alpha b = \delta.$$

It is then straightforward to check that the map

$$\mathbf{F}_q^\times \times \boldsymbol{\mu}_\delta(\mathbf{F}_q) \to \{(x,y) \in \mathbf{F}_q^\times \mid x^a y^b = 1\}$$

defined by

$$(t,\varrho) \mapsto (\varrho^\beta t^{b/\delta}, \varrho^\alpha t^{-a/\delta})$$

is a bijection with inverse

$$(x,y)\mapsto (x^{\alpha}y^{-\beta},x^{a/\delta}y^{b/\delta}).$$

Let now $u\in \mathbf{F}_q^{\times}$. If u is not of the form $x_0^ay_0^b$ with $(x_0,y_0)\in(\mathbf{F}_q^{\times})^2$, then $\widetilde{\mathrm{T}}_{a,b}(u;q)=0$. Otherwise, fix such representation $u=x_0^ay_0^b$. Then the above allows us to write

$$\widetilde{\mathrm{T}}_{a,b}(u;q)=\sum_{\varrho^{\delta}=1}\frac{1}{\sqrt{q}}\sum_{t\in\mathbf{F}_q^{\times}}e\Big(\frac{x_0\varrho^{\beta}t^{b/\delta}+y_0\varrho^{\alpha}t^{-a/\delta}}{q}\Big).$$

For each $\varrho\in\mathbf{F}_q^{\times}$ such that $\varrho^{\delta}=1$, the inner sum over t is a standard Weil sum with a rational function. The corresponding bounds are as follows, where we assume that $|a|/\delta>q$ and $|b|/\delta>q$.

Case 1. If a and b have the same sign, then the rational function has poles at 0 and ∞, and the Weil bound is

$$|\widetilde{\mathrm{T}}_{a,b}(u;q)|\leqslant\frac{|a|+|b|}{\delta}.$$

In fact, the sum is "pure" in that case, i.e., it is the sum of as many complex numbers of modulus 1 as the right-hand upper-bound.

Case 2. If a is negative and b is positive, then the function $t\mapsto\varrho^{\beta}t^{b/\delta}+\varrho^{\alpha}t^{-a/\delta}$ is a *polynomial* of degree $\max(|a|/\delta,b/\delta)$ (the assumption $a+b\neq0$ ensures there is no cancellation), so the Weil bound is

$$\Big|\frac{1}{\sqrt{q}}\sum_{t\in\mathbf{F}_q}e\Big(\frac{\varrho^{\beta}t^{b/\delta}+\varrho^{\alpha}t^{-a/\delta}}{q}\Big)\Big|\leqslant\frac{\max(-a,b)}{\delta}-1,$$

(and this sum is "pure") where the sum ranges over all $t\in\mathbf{F}_q$. Thus, we get a bound

$$|\widetilde{\mathrm{T}}_{a,b}(u;q)|\leqslant\max(|a|-1,b-1)+\frac{\delta}{q^{1/2}}\leqslant\max(|a|,|b|)$$

provided $q>\max(|a|,|b|)^2$.

Case 3. If a is positive and b is negative, then the symmetry $\widetilde{\mathrm{T}}_{-a,-b}(u;q)=\widetilde{\mathrm{T}}_{a,b}(u^{-1};q)$, reduces the question to Case 2. ☐

Remark 3.8. (1) This result is also proved by Pierce [30, § 3.1], in the case $(a,b)=1$.

(2) Elaborating this argument using the formalism of ℓ-adic sheaves and trace functions, one can show that for any prime number $q>\max(|a|,|b|)^2$, there exists a lisse sheaf $\mathscr{T}_{a,b}$ on $\mathbf{G}_m$ over $\mathbf{F}_q$ such that:

- The trace function of $\mathscr{T}_{a,b}$ coincides with $\widetilde{\mathrm{T}}(u;q)$;
- We have

  $$\mathrm{rank}(\mathscr{T}_{a,b})=\begin{cases}|a|+|b| & \text{if } ab\geqslant1\\ \max(|a|,|b|) & \text{if } ab\leqslant-1,\end{cases}$$

  and moreover the conductor (in the sense of [9]) of $\mathscr{T}_{a,b}$ is $\ll1$ for all q;
- The sheaf $\mathscr{T}_{a,b}$ is mixed of weights $\leqslant0$, so that

  $$|\widetilde{\mathrm{T}}(u;q)|\leqslant|a|+|b|$$

  for all $u\in\mathbf{F}_q^{\times}$;
- If a and b have the same sign, then the sheaf $\mathscr{T}_{a,b}$ is pure of weight 0. It is geometrically irreducible if and only if $(a,b)=1$.

This means in particular that we can apply to these exponential sums results such as those in [9, 10].

4. Counting solutions of toric congruences

4.1. A general problem. We consider first the general setting, as in Section 3.1, from which we borrow some notation.

Fix an integer $k \geqslant 1$. The basic question is now, given a prime q and a subgroup $\mathrm{H} \subset (\mathbf{F}_q^\times)^k$, to estimate the number of elements of some non-empty box $\mathrm{B} = \mathrm{I}_1 \times \cdots \times \mathrm{I}_k$ in $\mathbf{Z}^k$, where I_j is an interval, whose reduction modulo q lies in H.

More generally, it is natural to count solutions in cosets, since this provides some potential flexibility in arguing by induction. Thus, given an integral matrix A with k columns (and an arbitrary finite number of rows) and the subgroups $\mathrm{H}_\mathrm{A}(\mathbf{F}_q)$ defined in (3.1), we define

$$\mathrm{M}_\mathrm{A}(u, \mathrm{B}; q) = |\{x \in \mathrm{B} \mid x \,(\mathrm{mod}\, q) \in u\mathrm{H}_\mathrm{A}(\mathbf{F}_q)\}|$$

for $u \in (\mathbf{F}_q^\times)^k$, and we write simply $\mathrm{M}_\mathrm{A}(\mathrm{B}; q) = \mathrm{M}_\mathrm{A}(1, \mathrm{B}; q)$.

Concretely, this amounts to counting the solutions to a system of congruences of the form

$$x_1^{a_{i,1}} \cdots x_k^{a_{i,k}} \equiv y_i \,(\mathrm{mod}\, q) \tag{4.1}$$

for all i, with $(x_j) \in \mathrm{B}$ and $y_i \in \mathbf{F}_q^\times$. We call such systems "toric congruences".

Here also, some normalization plays an important role in applications to toroidal families, and we define for this purpose

$$\mathrm{N}_\mathrm{A}(u, \mathrm{B}; q) = \frac{\mathrm{M}_\mathrm{A}(u, \mathrm{B}; q)}{\sqrt{|\mathrm{B}|}}, \qquad \mathrm{N}_\mathrm{A}(\mathrm{B}; q) = \frac{\mathrm{M}_\mathrm{A}(1, \mathrm{B}; q)}{\sqrt{|\mathrm{B}|}}. \tag{4.2}$$

One can expect the problem of estimating these quantities to be quite difficult when the intervals I_j have small size, whereas the Riemann Hypothesis over finite fields can be applied to get an asymptotic formula once they are big enough. Here is a sample result of the latter kind.

Proposition 4.1. *Assume that* A *is non-negative and that the gcd of the coefficients of* A *is* 1*. Assume moreover that* $\mathbf{H}_\mathrm{A}(\mathbf{C})$ *is not contained in any of the sets defined by* $x_i = x_j$ *for some integers* $i \neq j$*, or* $x_j = 1$ *for some integer* j*.*

For any $\varepsilon > 0$*, for any prime number* q *and for any box of the type* $\mathrm{B} = \{1, \ldots, x\}^k$ *with* $1 \leqslant x < q$*, we have*

$$\mathrm{M}_\mathrm{A}(\mathrm{B}; q) = \frac{|\mathrm{B}|}{q^{\mathrm{rank}(\mathrm{A})}} + \mathrm{O}\Big(q^{(k-\mathrm{rank}(\mathrm{A}))/2+\varepsilon}\Big(1 + \frac{1}{q^{1/2}}\Big(\frac{|\mathrm{B}|}{q^{1/2}}\Big)^{k-\mathrm{rank}(\mathrm{A})}\Big)\Big)$$

where the implicit constant depends only on A *and* ε*.*

Although the result is valid uniformly, it is of course only non-trivial when $|\mathrm{B}|$ is large enough.

Proof. This is a direct application of the general result [8, Cor. 1.5] of Fouvry and Katz. We check the assumptions of this result:

(1) since the matrix A is non-negative, the algebraic torus $\mathbf{H}_\mathrm{A}$ is a closed subvariety of $\mathbf{A}^k$; since the gcd of the coefficients is one, the variety $\mathbf{H}_\mathrm{A}(\mathbf{C})$ is connected;

(2) by [8, Th. 8.1], the so-called "A-number" of $\mathbf{H}_\mathrm{A}$ over $\mathbf{F}_q$ is non-zero for all primes q;

(3) finally, $\mathbf{H}_{\mathrm{A}}(\mathbf{C})$ is not contained in any affine hyperplane in $\mathbf{C}^k$; indeed, if this were so, it would follow that the $k+1$ characters

$$x \mapsto x_j, \quad 1 \leqslant j \leqslant k, \qquad x \mapsto 1$$

of $\mathbf{H}_{\mathrm{A}}(\mathbf{C})$ are linearly dependent. Since the characters of any group are linearly independent, this is only possible if two at least of these characters coincide, which contradicts our assumption. ☐

We will use another bound in the proof of Theorem 1.1, which is based on the geometry of numbers and applies to linear toric congruences, i.e., congruences of the form $x_i \equiv \varrho x_j \pmod q$ for some $\varrho \in \mathbf{F}_q^\times$. This corresponds to

$$\mathrm{A} = (0, \ldots, 0, 1, 0, \ldots, 0, -1, 0, \ldots, 0),$$

with a 1 at the i-th position and a -1 at the j-th position.

Proposition 4.2. *Let $k \geqslant 1$ be an integer. Let q be a prime number. Fix integers i, j with $1 \leqslant i \neq j \leqslant k$.*

Let $u \in (\mathbf{F}_q^\times)^k$. Let λ_1 be the minimum of the lattice

$$\Lambda = \{y \in \mathbf{Z}^k \mid u_i y_i \equiv u_j y_j \pmod q\} \subset \mathbf{Z}^k.$$

For any box $\mathrm{B} \subset \mathbf{Z}^k$, we have

$$|\{x \in \mathrm{B} \mid u_i x_i \equiv u_j x_j \pmod q\}| \ll \frac{|\mathrm{B}|}{q} + \Big(\frac{\Delta}{\lambda_1} + 1\Big)^{k-1},$$

where Δ is the size of $\partial \mathrm{B}$ and the implicit constant depends only on k.

Proof. Write

$$\mathrm{B} = \prod_{j=1}^{k} \{a_j, a_j + 1, \ldots, b_j\} \subset \mathbf{Z}^k,$$

and let

$$\mathrm{B}^0 = \prod_{j=1}^{k} [a_j, b_j] \subset \mathbf{R}^k.$$

We are therefore computing the size of $\Lambda \cap \mathrm{B}^0$. The set B^0 is convex and bounded, and its volume is $\ll |\mathrm{B}|$. The result thus follows from the estimate

$$|\Lambda \cap \mathrm{B}^0| \ll \frac{\mathrm{Vol}(\mathrm{B}^0)}{q} + \Big(\frac{\Delta}{\lambda_1} + 1\Big)^{k-1},$$

where the implicit constant depends only on k, which is a case of the so-called Lipschitz Principle in the geometry of numbers. Among results of this type which imply this estimate, we appeal to one of Widmer [35, Cor. 5.3] (which is much more general).

More precisely, the determinant of Λ is q under our assumptions on u. Denoting by Ω the orthogonality defect of Λ (see [35, p. 4805]), and applying loc. cit., we get

$$\Big| |\Lambda \cap \mathrm{B}^0| - \frac{\mathrm{Vol}(\mathrm{B}^0)}{q} \Big| \leqslant 3^k \, (2k) \, \Big(\frac{\sqrt{k}\Omega\Delta}{\lambda_1} + 1\Big)^{k-1},$$

since the boundary of the box B^0 is the union of $2k$ faces, each of which is the image of a Lipschitz map $[0,1]^{k-1} \to \partial \mathrm{B}^0$ with Lipschitz constant bounded by the constant Δ. The bound $\Omega \ll 1$ holds by [35, Lemma 4.4], where the implicit constant depends only on k, and the proposition follows. □

Finally, we conclude with some speculations, maybe rather naive, but which we hope can provide definite targets for investigations of general toric congruences.

We consider a box B where the sides I_j are of the form $\mathrm{I}_j = \{1, \ldots, \mathrm{A}_j\}$ for some positive integers $\mathrm{A}_j < q$. There are some "systematic" sources of solutions of the toric congruences (4.1). These are provided by the non-trivial morphisms

$$\varphi\colon \mathbf{F}_q^\times \to (\mathbf{F}_q^\times)^k$$

of the form

$$\varphi(x) = (x^{b_1}, \ldots, x^{b_k})$$

with $b_i \geqslant 0$, not all zero, whose image lies in H_A. This means that (b_i) is a non-zero and non-negative solution of the system of linear equations

$$b_1 a_{i,1} + \cdots + b_k a_{i,k} = 0$$

for all i. These homomorphisms correspond to algebraic group morphisms $\mathbf{G}_m \to \mathbf{H}_\mathrm{A}$, and only depend on the matrix A.

Given such a morphism φ, we have $\varphi(x) \in \mathrm{B} \cap \mathrm{H}_\mathrm{A}(\mathbf{F}_q)$ provided

$$1 \leqslant x \leqslant \mathrm{A}_j^{1/b_j} \qquad \text{for all } j \text{ such that } b_j \geqslant 1,$$

where we use in an essential way the condition $b_j \geqslant 0$.

The kernel of the morphism φ if of size $\ll 1$ for all q (it is the group of $\gcd(b_1, \ldots, b_k)$-th roots of unity in $\mathbf{F}_q^\times$). The morphism φ therefore provides

$$\gg \min_{b_j \geqslant 1} (\mathrm{A}_j^{1/b_j})$$

solutions to the system of congruences (4.1) with right-hand side 1.

We now assume that $\mathrm{A}_j \geqslant q^{\alpha_j}$ for some $\alpha_j > 0$, so that the systematic solutions associated to a given φ are increasingly numerous. We then let $\mathrm{M}_\mathrm{A}^{\mathrm{syst}}(\mathrm{B}; q)$ denote the number of elements of $\mathrm{B} \cap \mathrm{H}_\mathrm{A}(\mathbf{F}_q)$ which belong to the image of a positive morphism (some solutions of the congruences might of course lie in the image of more than one morphism). One may then speculate that

$$\mathrm{M}_\mathrm{A}(\mathrm{B}; q) = \mathrm{M}_\mathrm{A}^{\mathrm{syst}}(\mathrm{B}; q) + \mathrm{O}\Big(q^\varepsilon \Big(1 + \frac{|\mathrm{B}|}{q^{k - \mathrm{rank}(\mathrm{A})}}\Big)\Big)$$

for any $\varepsilon > 0$ (the quantity $|\mathrm{B}|/q^{k-\mathrm{rank}(\mathrm{A})}$ on the right-hand side is of probabilistic nature: the probability for a uniformly random element of $(\mathbf{F}_q^\times)^k$ to belong to $\mathrm{H}_\mathrm{A}(\mathbf{F}_q)$ is about $q^{-\dim(\mathbf{H}_\mathrm{A})} = q^{-(k-\mathrm{rank}(\mathrm{A}))}$).

Remark 4.3. (1) There is a vague analogy here with the Pila–Wilkie counting theorem and the distinction between algebraic and transcendental parts of definable sets in o-minimal structures (see, e.g., the recent survey of Bhardwaj and van den Dries [2] for a discussion of the Pila–Wilkie theorem and its applications).

(2) If the matrix A is non-negative, so that $a_{i,j} \geqslant 0$, then we have $\mathrm{M}_\mathrm{A}^{\mathrm{syst}}(\mathrm{B}; q) = 0$.

4.2. **The case of Theorem 1.1.** We now discuss the example which is relevant for Theorem 1.1.

We have now $k = 2$ and $\mathrm{A} = (a, b) \in \mathbf{Z}^2$ with $ab \neq 0$. For q prime, the subgroup $\mathrm{H}_{\mathrm{A}}(\mathbf{F}_q)$ is the subgroup of $(\mathbf{F}_q^\times)^2$ given by the equation $x^a y^b = 1$.

A trivial bound is

$$\mathrm{M}_{a,b}(u, \mathrm{I}\times\mathrm{J}; q) \ll |\mathrm{I}|\Big(\frac{|\mathrm{J}|}{q}+1\Big), \qquad \mathrm{M}_{a,b}(u, \mathrm{I}\times\mathrm{J}; q) \ll |\mathrm{J}|\Big(\frac{|\mathrm{I}|}{q}+1\Big),$$

(indeed, for the first, note that for each $m \in \mathrm{I}$, the possible n such that $m^a n^b \equiv 1 \pmod q$ lie in the intersection of I with at most b arithmetic progressions modulo q). In particular

$$\mathrm{N}_{a,b}(u, \mathrm{I}\times\mathrm{J}; q) \ll \frac{\sqrt{|\mathrm{I}|}}{\sqrt{|\mathrm{J}|}}\Big(\frac{|\mathrm{J}|}{q}+1\Big). \tag{4.3}$$

The following is rather deeper, and is related to the methods used to prove the Burgess bound.

Theorem 4.4 (Pierce)**.** *Suppose that $b \neq 0$ and that a/b is not a negative integer. Let $k \geqslant 1$ be an integer. If* M *and* N *are positive integers $< q$ such that*

$$\mathrm{M} \leqslant \tfrac{1}{2} q^{\frac{k+1}{2k}}, \ \mathrm{N} \leqslant \tfrac{q}{4},$$

then we have

$$\mathrm{M}_{a,b}([1,\mathrm{M}]\times[1,\mathrm{N}]; q) \ll \mathrm{M}^{\frac{k}{k+1}} \mathrm{N}^{\frac{1}{2k}} (\log q)^{1/(2k)}$$

for all primes q, where the implicit constant depends only on (a, b, k).

Proof. This is a special case of [30, Th. 4] for q prime, noting that Pierce's equations are written in the form $x^a = y^b$, hence our condition on (a, b) is the opposite of that in loc. cit.; the other condition $(b, q) = 1$ in loc. cit. is true for all large enough q. □

On the other hand, Proposition 4.2 implies the following result.

Proposition 4.5. *Let q be a prime number. For any $\alpha \in \mathbf{F}_q^\times$, and any box $\mathrm{B} = \mathrm{I}\times\mathrm{J} \subset \mathbf{Z}^2$, we have*

$$\mathrm{M}_{1,-1}((\alpha, 1), \mathrm{I}\times\mathrm{J}; q) \ll \frac{|\mathrm{B}|}{q} + \frac{|\mathrm{I}|+|\mathrm{J}|}{\lambda} + 1$$

where λ is the minimum of the lattice

$$\Lambda = \{(m,n) \in \mathbf{Z}^2 \mid m \equiv \alpha n \pmod q\}.$$

Remark 4.6. (1) The motivation in the paper of Pierce is the study of class groups of imaginary quadratic fields. In this respect, we may also mention another work of Heath–Brown and Pierce [11, § 4], where toric congruences also appear and are studied in part through lattice-point counting methods.

(2) Special cases of these toric congruences appeared in earlier work of Heath–Brown [12] concerning squarefree numbers in arithmetic progressions, and more recently in related work of Nunes [27].

5. Beginning of the proof

We begin the proof of Theorem 1.1 with some general preliminaries. For any Dirichlet character χ modulo q, we denote $t(\chi) = (1-\chi(-1))/2$.

Let $(a,b) \in \mathbf{Z}^2$ be integers with $ab \neq 0$. Let q be a prime number. We intend to compute

$$\mathrm{M}_{a,b}(q) = \frac{1}{q-1} \sum_{\chi \,(\mathrm{mod}\, q)} \mathrm{L}(\tfrac{1}{2}, \chi^a)\mathrm{L}(\tfrac{1}{2}, \chi^b).$$

For any Dirichlet character χ modulo q, we have

$$\mathrm{L}(s, \chi^a)\mathrm{L}(s, \chi^b) = \sum_{m,n \geqslant 1} \frac{\chi(m^a n^b)}{(mn)^s}, \qquad \mathrm{Re}(s) > 1.$$

Suppose that $\chi^a \neq 1$ and $\chi^b \neq 1$, so that the L-functions $\mathrm{L}(s,\chi^a)$ and $\mathrm{L}(s,\chi^b)$ have functional equations with signs

$$i^{-t(\chi^a)}\varepsilon(\chi^a), \qquad i^{-t(\chi^b)}\varepsilon(\chi^b),$$

respectively. For any positive real numbers X and Y such that $\mathrm{XY} = q^2$, the approximate functional equation for the product L-function gives the formula

$$\mathrm{L}(\tfrac{1}{2}, \chi^a)\mathrm{L}(\tfrac{1}{2}, \chi^b) = \sum_{m,n \geqslant 1} \frac{\chi(m^a n^b)}{\sqrt{mn}} \mathrm{V}_{a,b,\chi}\Big(\frac{mn}{\mathrm{X}}\Big) + \frac{\varepsilon(\chi^a)\varepsilon(\chi^b)}{i^{t(\chi^a)+t(\chi^b)}} \sum_{m,n \geqslant 1} \frac{\bar{\chi}(m^a n^b)}{\sqrt{mn}} \mathrm{V}_{a,b,\chi}\Big(\frac{mn}{\mathrm{Y}}\Big) \tag{5.1}$$

where the function $\mathrm{V}_{a,b,\chi}$ has rapid decay at infinity and only depends on $(\chi^a(-1), \chi^b(-1))$.

To be precise, we apply [13, Th. 5.3] to the L-function $\mathrm{L}(s,\chi^a)\mathrm{L}(s,\chi^b)$ with conductor q^2 and gamma factor

$$\gamma_{a,b}(s) = \pi^{-s}\Gamma\Big(\frac{s+t(\chi^a)}{2}\Big)\Gamma\Big(\frac{s+t(\chi^b)}{2}\Big),$$

and with the parameter X in loc. cit. replaced by $\mathrm{X}/\sqrt{q}$. Furthermore, we choose a test function G such that

$$\begin{cases} \mathrm{G} \text{ is holomorphic and bounded in the strip } -4 < \mathrm{Re}(u) < 4, \\ \mathrm{G} \text{ is even}, \\ \mathrm{G}(0) = 1. \end{cases} \tag{5.2}$$

Then the formula (5.1) holds with

$$\mathrm{V}_{a,b,\chi}(y) = \frac{1}{2i\pi} \int_{(3)} y^{-u} \mathrm{G}(u) \frac{\gamma_{a,b}(\frac{1}{2}+u)}{\gamma_{a,b}(\frac{1}{2})} \frac{du}{u}.$$

Note that we can define $\mathrm{V}_{a,b,\chi}$ by the same formula even if either χ^a or χ^b is trivial. For any $\mathrm{A} > 0$, the following bounds

$$\mathrm{V}_{a,b,\chi}(y) = 1 + \mathrm{O}(y^{\mathrm{A}}) \tag{5.3}$$

$$\mathrm{V}_{a,b,\chi}(y) \ll y^{-\mathrm{A}} \tag{5.4}$$

hold for $y > 0$ (see [13, Prop. 5.4]).

For *any* character χ, including the case where χ^a or χ^b is trivial, we have

$$\mathrm{L}(\tfrac{1}{2}, \chi^a)\mathrm{L}(\tfrac{1}{2}, \chi^b) \ll q^{1/2}$$

(by the convexity bound, see e.g. [13, Th. 5.23]). Moreover, for any $\varepsilon > 0$ and $\mathrm{X}, \mathrm{Y} \geqslant 1$, we have

$$\sum_{m,n\geqslant 1} \frac{\chi(m^a n^b)}{\sqrt{mn}} \mathrm{V}_{a,b,\chi}\Big(\frac{mn}{\mathrm{X}}\Big) \ll \mathrm{X}^{1/2+\varepsilon}$$

$$\varepsilon(\chi^a)\varepsilon(\chi^b) \sum_{m,n\geqslant 1} \frac{\bar{\chi}(m^a n^b)}{\sqrt{mn}} \mathrm{V}_{a,b,\chi}\Big(\frac{mn}{\mathrm{Y}}\Big) \ll \mathrm{Y}^{1/2+\varepsilon}$$

(since $|\varepsilon(\chi)| \leqslant 1$ for all characters). Thus, bringing back the contributions of the characters where either χ^a or χ^b is trivial, we have

$$\begin{aligned}(5.5)\quad \mathrm{M}_{a,b}(q) = \frac{1}{q-1} \sum_{\chi \,(\mathrm{mod}\, q)} \Big\{ \sum_{m,n\geqslant 1} \frac{\chi(m^a n^b)}{\sqrt{mn}} \mathrm{V}_{a,b,\chi}\Big(\frac{mn}{\mathrm{X}}\Big) \\ + \frac{\varepsilon(\chi^a)\varepsilon(\chi^b)}{i^{t(\chi^a)+t(\chi^b)}} \sum_{m,n\geqslant 1} \frac{\bar{\chi}(m^a n^b)}{\sqrt{mn}} \mathrm{V}_{a,b,\chi}\Big(\frac{mn}{\mathrm{Y}}\Big)\Big\} \\ + \mathrm{O}(q^{-1/2} + q^{-1}(\mathrm{X}^{1/2+\varepsilon} + \mathrm{Y}^{1/2+\varepsilon})),\end{aligned}$$

where the implicit constant depends on a, b and ε.

We will assume that $q \geqslant 3$ and focus on the contribution, denoted $\mathrm{M}_{a,b}^{\mathrm{even}}(q)$, of the even characters to the sum. Note that we then have $t(\chi^a) = t(\chi^b) = 0$, and $\mathrm{V}_{a,b,\chi}$ is independent of χ, and is denoted simply V. We have therefore

$$\mathrm{M}_{a,b}^{\mathrm{even}}(q) = \frac{1}{q-1} \sum_{\substack{\chi \,(\mathrm{mod}\, q) \\ \chi \text{ even}}} \Big\{ \sum_{m,n\geqslant 1} \frac{\chi(m^a n^b)}{\sqrt{mn}} \mathrm{V}\Big(\frac{mn}{\mathrm{X}}\Big) + \varepsilon(\chi^a)\varepsilon(\chi^b) \sum_{m,n\geqslant 1} \frac{\bar{\chi}(m^a n^b)}{\sqrt{mn}} \mathrm{V}\Big(\frac{mn}{\mathrm{Y}}\Big)\Big\}.$$

Remark 5.1. Note that the sum over even characters can also be interpreted as the restriction to the toroidal family which is the image of the morphism $\chi \mapsto \chi^2$.

By orthogonality of Dirichlet characters modulo q, combined with (3.3) and (3.4), we obtain

$$\begin{aligned}(5.6)\quad \mathrm{M}_{a,b}^{\mathrm{even}}(q) = \frac{1}{q-1} \sum_{\chi \,(\mathrm{mod}\, q)} \frac{1}{2}(1+\chi(-1)) \\ \Big\{ \sum_{m,n\geqslant 1} \frac{\chi(m^a n^b)}{\sqrt{mn}} \mathrm{V}\Big(\frac{mn}{\mathrm{X}}\Big) + \varepsilon(\chi^a)\varepsilon(\chi^b) \sum_{m,n\geqslant 1} \frac{\bar{\chi}(m^a n^b)}{\sqrt{mn}} \mathrm{V}\Big(\frac{mn}{\mathrm{Y}}\Big)\Big\} \\ = \mathrm{N}_{a,b}(\mathrm{X}) + \mathrm{P}_{a,b}(\mathrm{Y}),\end{aligned}$$

where

$$\mathrm{N}_{a,b}(\mathrm{X}) = \frac{1}{2} \sum_{\substack{m,n\geqslant 1 \\ (m^a n^b)^2 \equiv 1 \,(\mathrm{mod}\, q)}} \frac{1}{\sqrt{mn}} \mathrm{V}\Big(\frac{mn}{\mathrm{X}}\Big)$$

and

$$\begin{aligned}\mathrm{P}_{a,b}(\mathrm{Y}) &= \frac{1}{2\sqrt{q}} \sum_{\substack{m,n\geqslant 1\\(mn,q)=1}} \frac{1}{\sqrt{mn}}(\widetilde{\mathrm{T}}_{a,b}(m^an^b;q)+\widetilde{\mathrm{T}}_{a,b}(-m^an^b;q))\mathrm{V}\Big(\frac{mn}{\mathrm{Y}}\Big)\\ &= \frac{1}{2\sqrt{q}} \sum_{m,n\geqslant 1} \frac{1}{\sqrt{mn}}\widetilde{\mathrm{T}}_{2a,2b}(m^{2a}n^{2b};q)\mathrm{V}\Big(\frac{mn}{\mathrm{Y}}\Big).\end{aligned}$$

We see that the toric congruences of Section 4.2 appear in $\mathrm{N}_{a,b}(\mathrm{X})$, and that the exponential sums of Section 3.2 appear in $\mathrm{P}_{a,b}(\mathrm{Y})$. The treatment of these terms now varies depending on whether $a+b=0$ or not, the key difference being how to handle the first term, and in particular how to extract the main term.

Before proceeding, we observe that

$$\mathrm{M}_{a,b}(q) = \overline{\mathrm{M}_{-a,-b}(q)},$$

which allows us to assume from now on that a is positive.

Remark 5.2. More generally, let A be an integral matrix A with k columns as in Sections 3 and 4, and define the subroups $\mathbf{H}_{\mathrm{A}}^{\perp}(\mathbf{F}_q)$ as in (3.2). Then the toroidal average

$$\frac{1}{|\mathrm{H}_{\mathrm{A}}^{\perp}(\mathbf{F}_q)|} \sum_{\boldsymbol{\chi}\in \mathrm{H}_{\mathrm{A}}^{\perp}(\mathbf{F}_q)} \prod_{j=1}^{k} \mathrm{L}(\tfrac{1}{2},\chi_j)$$

associated to A will essentially be a combination of expressions of the type

$$\sum_{\substack{m_1,\ldots,m_k\geqslant 1\\(m_1,\ldots,m_k)\,(\mathrm{mod}\,q)\in \mathrm{H}_{\mathrm{A}}(\mathbf{F}_q)}} \frac{1}{\sqrt{m_1\cdots m_k}}\mathrm{V}_{\boldsymbol{\chi}}\Big(\frac{m_1\cdots m_k}{\mathrm{X}}\Big)$$

and

$$q^{\mathrm{rank}(\mathrm{A})/2} \sum_{m_1,\ldots,m_k\geqslant 1} \frac{\mathrm{T}_{\mathrm{A}}(m_1,\ldots,m_k;q)}{\sqrt{m_1\cdots m_k}}\mathrm{V}_{\boldsymbol{\chi}}\Big(\frac{m_1\cdots m_k}{\mathrm{Y}}\Big),$$

for suitable test functions $\mathrm{V}_{\boldsymbol{\chi}}$, with $\mathrm{XY}=q^k$ (depending on the parity of the characters, as usual).

6. Proof of Theorem 1.1 for $a+b\neq 0$

6.1. **The main term when a, b are positive.** We assume here that a and b are positive. The toric congruence

$$(m^an^b)^2 \equiv 1\,(\mathrm{mod}\,q) \tag{6.1}$$

has the solution $(m,n)=(1,1)$, which contributes

$$\frac{1}{2}\mathrm{V}\Big(\frac{1}{\mathrm{X}}\Big) = \frac{1}{2}+\mathrm{O}(\mathrm{X}^{-1/2})$$

to $\mathrm{N}_{a,b}(\mathrm{X})$ (see (5.3)). For any other solution (m,n) of the congruence with m, n positive, we have $mn>1$, hence

$$\max(m,n)\geqslant (q-1)^{\frac{1}{a+b}}.$$

Fix some $\delta > 0$ to be chosen later. Using a dyadic partition of unity and the rapid decay of V (see (5.4)), it follows that

$$\sum_{\substack{mn>1 \\ (m^a n^b)^2 \equiv 1 \pmod q}} \frac{1}{\sqrt{mn}} \mathrm{V}\Big(\frac{mn}{\mathrm{X}}\Big) \ll (\log q)^2 \max_{\mathrm{I},\mathrm{J}} \mathrm{N}_{2a,2b}(\mathrm{I} \times \mathrm{J}; q) + \mathrm{O}(\mathrm{X}^{-1}) \tag{6.2}$$

(with notation as in (4.2)), where (I, J) runs over pairs of intervals of integers $n \leqslant \mathrm{X}$ such that

$$\max(|\mathrm{I}|, |\mathrm{J}|) \geqslant \tfrac{1}{2} q^{1/(a+b)}, \qquad |\mathrm{I}|\,|\mathrm{J}| \ll \mathrm{X}^{1+\delta}.$$

Fix one such pair (I, J). We distinguish two cases.

Case 1. If $\max(|\mathrm{I}|, |\mathrm{J}|) \geqslant q/4$, we can use the bound (4.3) to get

$$\mathrm{N}_{2a,2b}(\mathrm{I} \times \mathrm{J}; q) \ll \frac{(|\mathrm{I}||\mathrm{J}|)^{1/2}}{q} \ll \frac{\mathrm{X}^{1/2+\delta/2}}{q}. \tag{6.3}$$

Case 2. If $\max(|\mathrm{I}|, |\mathrm{J}|) < q/4$, then we will appeal to Theorem 4.4, with $k = 2$. Since $a/b > 0$, we may assume that $|\mathrm{J}| \geqslant |\mathrm{I}|$, by switching the coordinates if needed. We then have $\sqrt{|\mathrm{I}||\mathrm{J}|} \leqslant |\mathrm{J}|$.

Under the assumption that

$$|\mathrm{I}| \leqslant \tfrac{1}{2} q^{3/4}, \tag{6.4}$$

Theorem 4.4 implies the estimate

$$\mathrm{N}_{2a,2b}(\mathrm{I} \times \mathrm{J}; q) \ll (\log q)|\mathrm{I}|^{2/3-1/2}|\mathrm{J}|^{1/4-1/2} = (\log q)|\mathrm{I}|^{1/6}|\mathrm{J}|^{-1/4}.$$

Since $|\mathrm{I}||\mathrm{J}| \leqslant |\mathrm{J}|^2$, we have then

$$|\mathrm{I}|^{1/6}|\mathrm{J}|^{-1/4} = (|\mathrm{I}||\mathrm{J}|)^{1/6}|\mathrm{J}|^{-5/12} \leqslant (|\mathrm{I}||\mathrm{J}|)^{1/6-5/24} = (|\mathrm{I}||\mathrm{J}|)^{-1/24} \leqslant \max(|\mathrm{I}|, |\mathrm{J}|)^{-1/24},$$

and hence

$$\mathrm{N}_{2a,2b}(\mathrm{I} \times \mathrm{J}; q) \ll (\log q) \max(|\mathrm{I}|, |\mathrm{J}|)^{-1/24} \ll (\log q) q^{1/(24(a+b))}. \tag{6.5}$$

From the combination of (6.2), (6.3) and (6.5), we obtain

$$\sum_{\substack{mn>1 \\ (m^a n^b)^2 \equiv 1 \pmod q}} \frac{1}{\sqrt{mn}} \mathrm{V}\Big(\frac{mn}{\mathrm{X}}\Big) \ll (\log q)^2 \Big(q^{-1}\mathrm{X}^{(1+\delta)/2} + (\log q) q^{-1/(24(a+b))}\Big) + \mathrm{X}^{-1},$$

if the condition (6.4) is satisfied for all relevant intervals I and J. Since these satisfy

$$|\mathrm{I}| \leqslant (|\mathrm{I}||\mathrm{J}|)^{1/2} \ll \mathrm{X}^{(1+\delta)/2}$$

and since $1 - \delta < 1/(1 + \delta)$ if $0 < \delta < 1$, the condition (6.4) is indeed satisfied, provided

$$\mathrm{X} \ll q^{3(1-\delta)/2}.$$

This last estimate also implies that $\mathrm{X}^{(1+\delta)/2} q^{-1} \ll q^{-(1-3\delta)/4}$, and we conclude that if δ is fixed and small enough, and if $\mathrm{X} \ll q^{3(1-\delta)/2}$, then we have

$$\mathrm{N}_{a,b}(\mathrm{X}) = \frac{1}{2} + \mathrm{O}(q^{-1/(24(a+b))+\varepsilon} + \mathrm{X}^{-1}) \tag{6.6}$$

for any $\varepsilon > 0$, where the implicit constant depends on a, b and ε.

6.2. **The main term when a is positive and b is negative.** We assume here that $a \geqslant 1$ and $-b \geqslant 1$, and recall that $a+b \neq 0$. We denote by $d \geqslant 1$ the gcd of a and b.

In this case the congruence (6.1) becomes

$$m^{2a} \equiv n^{-2b} \,(\mathrm{mod}\, q).$$

It admits "systematic" (positive) solutions of the form

$$(m, n) = (r^{-b/d}, r^{a/d})$$

for $r \geqslant 1$. The contribution $\mathrm{N}^{\mathrm{syst}}_{a,b}(\mathrm{X})$ of solutions of this form to $\mathrm{N}_{a,b}(\mathrm{X})$ is equal to

$$\mathrm{N}^{\mathrm{syst}}_{a,b}(\mathrm{X}) = \frac{1}{2} \sum_{r \geqslant 1} \frac{1}{r^{(a-b)/2d}} \mathrm{V}\Big(\frac{r^{(a-b)/d}}{\mathrm{X}}\Big) = \frac{1}{2} \zeta\Big(\frac{a-b}{2d}\Big) + \mathrm{O}(\mathrm{X}^{-1/2+\varepsilon}), \tag{6.7}$$

for any $\varepsilon > 0$, where we note that the fact that $a+b \neq 0$ implies that $(a-b)/(2d) > 1$.

We denote by $\mathrm{N}^{\mathrm{err}}_{a,b}(\mathrm{X})$ the contribution of the positive solutions (m, n) to the congruence which are not of the form above. Let (m, n) be one of these. Then the integer

$$m^{2a} - n^{-2b} = (m^a - n^{-b})(m^a + n^{-b})$$

is non-zero and divisible by q, so that we obtain the lower bound

$$\max(m, n) \geqslant (\tfrac{1}{2} q)^{1/\max(a,-b)}. \tag{6.8}$$

Fix again $\delta > 0$ small enough. By a dyadic partition of unity and the rapid decay (5.4) of V (applied for some fixed A large enough, depending on δ), we have

$$\mathrm{N}^{\mathrm{err}}_{a,b}(\mathrm{X}) \ll (\log q)^2 \max_{\mathrm{I},\mathrm{J}} \mathrm{N}_{2a,2b}(\mathrm{I} \times \mathrm{J}; q) + \mathrm{X}^{-1} \tag{6.9}$$

(with notation as in (4.2)), where (I, J) runs over pairs of intervals of integers $n \leqslant \mathrm{X}$ such that

$$\max(|\mathrm{I}|, |\mathrm{J}|) \gg q^{1/\max(a,-b)}, \qquad |\mathrm{I}|\,|\mathrm{J}| \ll \mathrm{X}^{1+\delta},$$

where the implicit constants depend only on a and b.

Fix one such pair (I, J). Let

$$\mathrm{K}_+ = \max(|\mathrm{I}|, |\mathrm{J}|), \qquad \mathrm{K}_- = \min(|\mathrm{I}|, |\mathrm{J}|).$$

The bound (4.3) gives the estimate

$$\mathrm{N}_{2a,2b}(\mathrm{I} \times \mathrm{J}; q) \ll \frac{\mathrm{K}_-(\mathrm{K}_+/q + 1)}{(\mathrm{K}_+\mathrm{K}_-)^{1/2}} \ll \frac{\mathrm{X}^{1/2+\delta/2}}{q} + \Big(\frac{\mathrm{K}_-}{\mathrm{K}_+}\Big)^{1/2}.$$

We now distinguish two cases again, depending on some parameter $\eta > 0$ to be chosen later, depending on a and b.

Case 1. If

$$\frac{\mathrm{K}_+}{\mathrm{K}_-} \geqslant q^{\eta},$$

we deduce that

$$\mathrm{N}_{2a,2b}(\mathrm{I} \times \mathrm{J}; q) \ll \frac{\mathrm{X}^{(1+\delta)/2}}{q} + q^{-\eta/2}. \tag{6.10}$$

Case 2. Suppose that $K_+/K_- < q^\eta$. The condition $a+b \neq 0$ implies that at least one of a/b or b/a is not an integer, hence is not a negative integer. Under the condition that

$$K_+ \leqslant \tfrac{1}{2} q^{3/4},$$

we can apply Theorem 4.4 with $k=2$ in either case with $(M,N) = (K_+, K_-)$, and deduce that

$$N_{2a,2b}(I \times J; q) \ll (\log q) K_+^{1/6} K_-^{-1/4} = (\log q)(K_+K_-)^{1/6} K_-^{-5/12},$$

and (since $K_- \geqslant q^{-\eta}\sqrt{K_-K_+}$) further

$$\begin{aligned} N_{2a,2b}(I \times J; q) &\ll (\log q)(K_+K_-)^{-1/24} q^{5\eta/12} = (\log q)(|I||J|)^{-1/24} q^{5\eta/12} \\ &\ll (\log q) q^{-1/(24\max(a,-b)) + 5\eta/12}. \end{aligned} \tag{6.11}$$

The bounds (6.10) and (6.11) lead by (6.9) to the estimate

$$N^{\mathrm{err}}_{a,b}(X) \ll (\log q)^2 \Big(q^{-1} X^{(1+\delta)/2} + q^{-\eta/2} + (\log q) q^{-1/(24\max(a,-b))+5\eta/12} \Big) + X^{-1},$$

provided the condition $K_+ \leqslant \frac{1}{2} q^{3/4}$ is satisfied for all relevant intervals I and J. Since

$$K_+^2 q^{-\eta} \leqslant K_+K_- = |I||J| \ll X^{1+\delta},$$

and $1-\delta < 1/(1+\delta)$ if $0 < \delta < 1$, this condition will be true as soon as $X \ll q^{(3/2-\eta)(1-\delta)}$.

Since this last condition on X implies that

$$q^{-1} X^{(1+\delta)/2} \ll q^{-1/4-\eta/2},$$

we conclude, by combining this with (6.7), that if η and δ are fixed and small enough, then

$$N_{a,b}(X) = \frac{1}{2}\zeta\Big(\frac{a-b}{2d}\Big) + O(X^{-1/2+\varepsilon} + q^{-1/(24\max(a,-b))+5\eta/12+\varepsilon} + q^{-\eta/2}) \tag{6.12}$$

for any $\varepsilon > 0$, provided $X \ll q^{(3/2-\eta)(1-\delta)}$.

6.3. **Conclusion of the proof.** Since $a+b \neq 0$, we can apply Proposition 3.7 to estimate the exponential sums $\widetilde{T}_{2a,2b}(m^{2a}n^{2b}; q)$ in $P_{a,b}(Y)$. Using the decay of V (see (5.4)), we obtain

$$P_{a,b}(Y) \ll q^{-1/2} Y^{1/2+\varepsilon} \tag{6.13}$$

for any $\varepsilon > 0$, where the implicit constant depends on a, b and ε.

Remark 6.1. This bound is trivial in the "balanced" case, i.e., when $Y = q$. It should however be possible with current methods to improve it by a factor $q^{-\eta}$ for some $\eta > 0$ if Y has size about q, but this would require much more elaborate arguments, in the spirit of the work of Kowalski, Michel and Sawin [22, Thm. 1.3].

Combining (6.13) with (6.6) or (6.12), we obtain in all cases (recalling that $a \geqslant 1$) the estimate

$$M^{\mathrm{even}}_{a,b}(q) = \tfrac{1}{2}\alpha(a,b) + O(q^{-1/(24(|a|+|b|))+5\eta/12+\varepsilon} + X^{-1/2+\varepsilon} + q^{-\eta/2} + q^{-1/2}Y^{1/2+\varepsilon})$$

where

$$\alpha(a,b) = \begin{cases} 1 & \text{if } ab \geqslant 1, \\ \zeta(\frac{a-b}{2(a,b)}) & \text{if } ab \leqslant -1, \end{cases}$$

provided $X \ll q^{(3/2-\eta)(1-\delta)}$ for some fixed small enough $\delta > 0$ and $\eta > 0$.

We pick (say) $\delta = 1/4$ and $\eta = \frac{c_0}{|a|+|b|}$ for c_0 fixed and small enough, and define

$$X = q^{(3/2-\eta)(1-\delta)} = q^{9/8-3\eta/4},$$

so that $Y = q^{7/8+3\eta/4}$. Inspecting the error terms above, we see that if c_0 is small enough ($c_0 < 1/20$ suffices), then there exists $c > 0$, depending only on c_0, such that the asymptotic formula

$$M^{\text{even}}_{a,b}(q) = \tfrac{1}{2}\alpha(a,b) + O(q^{-c/(|a|+|b|))+\varepsilon})$$

holds for any $\varepsilon > 0$, where the implicit constant depends only on a, b and ε.

We now consider briefly the contribution from the odd characters. This is

$$M^{\text{odd}}_{a,b}(q) = \frac{1}{q-1}\sum_{\substack{\chi\ (\mathrm{mod}\ q)\\ \chi\ \text{odd}}}\Big\{\sum_{m,n\geqslant 1}\frac{\chi(m^a n^b)}{\sqrt{mn}}W\Big(\frac{mn}{X}\Big) + \frac{\varepsilon(\chi^a)\varepsilon(\chi^b)}{i^\beta}\sum_{m,n\geqslant 1}\frac{\bar{\chi}(m^a n^b)}{\sqrt{mn}}W\Big(\frac{mn}{Y}\Big)\Big\}$$

for some integer β independent of χ (depending on the parity of a and b) and some function W that also depends only on the parity of (a,b). This leads to the variant

$$M^{\text{odd}}_{a,b}(q) = N'_{a,b}(X) + i^{-\beta}P'_{a,b}(Y)$$

of (5.6), where

$$N'_{a,b}(X) = \frac{1}{2}\sum_{\substack{m,n\geqslant 1\\ m^a n^b\equiv 1\ (\mathrm{mod}\ q)}}\frac{1}{\sqrt{mn}}W\Big(\frac{mn}{X}\Big) - \frac{1}{2}\sum_{\substack{m,n\geqslant 1\\ m^a n^b\equiv -1\ (\mathrm{mod}\ q)}}\frac{1}{\sqrt{mn}}W\Big(\frac{mn}{X}\Big)$$

and

$$P'_{a,b}(Y) = \frac{1}{2\sqrt{q}}\sum_{\substack{m,n\geqslant 1\\ (mn,q)=1}}\frac{1}{\sqrt{mn}}(\widetilde{T}_{a,b}(m^a n^b;q) - \widetilde{T}_{a,b}(-m^a n^b;q))W\Big(\frac{mn}{Y}\Big).$$

Proceeding exactly as before (with the same choices of X and Y), we obtain

$$M^{\text{odd}}_{a,b}(q) = \tfrac{1}{2}\alpha(a,b) + O(q^{-c/(|a|+|b|)+\varepsilon}).$$

Finally, combining these with (5.5), we conclude the proof of Theorem 1.1 in the case where $a + b \neq 0$.

7. Proof of Theorem 1.1 for $a + b = 0$

We will write simply $M_a(q) = M_{a,-a}(q)$. As in the previous section, we treat first in details the average $M^{\text{even}}_a(q)$ over even characters. Starting from (5.6) with $X = Y = q$ and using the fact that $\varepsilon(\chi^a)\varepsilon(\chi^{-a}) = 1$ for χ even and non-trivial, we get

$$M^{\text{even}}_a(q) = \sum_{m^{2a}\equiv n^{2a}\ (\mathrm{mod}\ q)}\frac{1}{\sqrt{mn}}V\Big(\frac{mn}{q}\Big),$$

where $V = V_{0,0,\chi}$ is independent of χ even.

The equation $m^{2a} = n^{2a}$ has solutions in positive integers given by $m = n \geqslant 1$, and no other integral solution. The contribution $M^{\text{syst}}_a(X)$ of these solutions to $M^{\text{even}}_a(q)$ is equal to

$$M^{\text{syst}}_a(X) = \sum_{\substack{m\geqslant 1\\ (m,q)=1}}\frac{1}{m}V\left(\frac{m^2}{q}\right) = \frac{1}{2}\log q + C + O(q^{-1/2}), \tag{7.1}$$

for some constant C, independent of a and b (see [16, Lemma 4.1], which provides the value

$$\mathrm{C} = 2\gamma + \Gamma'(1/4)/\Gamma(1/4) = \gamma - \frac{\pi}{2} - 3\log(2) = -3.073\ldots, \tag{7.2}$$

where γ is Euler's constant).[3]

We denote by $\mathrm{M}_a^{\mathrm{err}}(q)$ the contribution of the positive solutions (m, n) to the congruence which are not of the form above. Let (m, n) be one of these. Then the integer

$$m^{2a} - n^{2a} = (m^a - n^a)(m^a + n^a)$$

is non-zero and divisible by q, and we obtain the lower bound

$$\max(m, n) \geqslant (\tfrac{1}{2}q)^{1/a}$$

(as in (6.8)).

We split the sum further as follows:

$$\mathrm{M}_a^{\mathrm{err}}(q) = \sum_{\substack{\varrho\in\mathbf{F}_q^\times \\ \varrho^{2a}=1}} \mathrm{M}_a^{\mathrm{err}}(q, \varrho),$$

where

$$\mathrm{M}_a^{\mathrm{err}}(q, \varrho) = \sum_{\substack{m\equiv\varrho n \,(\mathrm{mod}\, q) \\ m^{2a}\neq n^{2a}}} \frac{1}{\sqrt{mn}}\mathrm{V}\Big(\frac{mn}{q}\Big).$$

We define

$$\mathrm{M}^*_{1,-1}((\varrho, 1), \mathrm{I}\times\mathrm{J}; q) = \mathrm{M}_{1,-1}((\varrho, 1), \mathrm{I}\times\mathrm{J}; q)$$

if $\varrho^2 \neq 1$, whereas if $\varrho^2 = 1$, then we define

$$\mathrm{M}^*_{1,-1}((\varrho, 1), \mathrm{I}\times\mathrm{J}; q) = |\{(m, n) \in \mathrm{I}\times\mathrm{J} \mid n \equiv \varrho m \,(\mathrm{mod}\, q) \text{ and } n^2 \neq m^2\}|.$$

Let $\delta > 0$ be a fixed parameter to be chosen later. By a dyadic partition of unity and the rapid decay of V (5.4), we have

$$\mathrm{M}_a^{\mathrm{err}}(q, \varrho) \ll (\log q)^2 \max_{\mathrm{I},\mathrm{J}} \frac{\mathrm{M}^*_{1,-1}((\varrho, 1), \mathrm{I}\times\mathrm{J}; q)}{(|\mathrm{I}||\mathrm{J}|)^{1/2}} + q^{-1} \tag{7.3}$$

where (I, J) runs over pairs of intervals of integers $n \leqslant q$ such that

$$\max(|\mathrm{I}|, |\mathrm{J}|) \gg q^{1/a}, \qquad |\mathrm{I}|\,|\mathrm{J}| \ll q^{1+\delta},$$

the implicit constants depending only on a and b.

Fix one such pair (I, J). If $\mathrm{M}^*_{1,-1}((\varrho, 1), \mathrm{I}\times\mathrm{J}; q)$ is zero, then there is nothing to do. Otherwise, let

$$\mathrm{K}_+ = \max(|\mathrm{I}|, |\mathrm{J}|), \qquad \mathrm{K}_- = \min(|\mathrm{I}|, |\mathrm{J}|).$$

Let $\eta > 0$ be a real number to be fixed later.

Case 1. Suppose that

$$\frac{\mathrm{K}_+}{\mathrm{K}_-} \geqslant q^\eta. \tag{7.4}$$

[3] Actually, [16, Lemma 4.1] is established for a test function G satisfying conditions slightly different from (5.2) (see [16, p. 942]), but the extension to the equality (7.1) is straightforward.

The bound (4.3) then gives the estimate

$$\text{(7.5)}\quad \frac{\mathrm{M}^*_{1,-1}((\varrho,1),\mathrm{I}\times\mathrm{J};q)}{(|\mathrm{I}||\mathrm{J}|)^{1/2}} \ll \frac{\mathrm{K}_-(\mathrm{K}_+/q+1)}{(\mathrm{K}_+\mathrm{K}_-)^{1/2}}$$

$$\ll \frac{(|\mathrm{I}||\mathrm{J}|)^{1/2}}{q} + \Big(\frac{\mathrm{K}_-}{\mathrm{K}_+}\Big)^{1/2} \ll q^{-1/2+\delta/2} + q^{-\eta/2}.$$

Case 2. Suppose that (7.4) is not satisfied. Then $\mathrm{K}_+ < q^\eta \mathrm{K}_-$, and

$$|\mathrm{I}| \leqslant q^{(1+\delta)/2+\eta}, \qquad |\mathrm{J}| \leqslant q^{(1+\delta)/2+\eta},$$

in addition to the previous lower bound on $\max(|\mathrm{I}|,|\mathrm{J}|)$.

Let $\Lambda_\varrho \subset \mathbf{Z}^2$ be the lattice determined by the condition

$$m \equiv \varrho n \,(\mathrm{mod}\, q),$$

and let $\lambda_\varrho \geqslant 1$ be its minimum. By Proposition 4.5, we have

$$\mathrm{M}^*_{1,-1}((\varrho,1),\mathrm{I}\times\mathrm{J};q) \leqslant \mathrm{M}_{1,-1}((\varrho,1),\mathrm{I}\times\mathrm{J};q) \ll \frac{|\mathrm{I}||\mathrm{J}|}{q} + \frac{|\mathrm{I}|+|\mathrm{J}|}{\lambda_\varrho} + 1,$$

hence

$$\frac{\mathrm{M}^*_{1,-1}((\varrho,1),\mathrm{I}\times\mathrm{J};q)}{(|\mathrm{I}||\mathrm{J}|)^{1/2}} \ll \frac{(|\mathrm{I}||\mathrm{J}|)^{1/2}}{q} + \frac{|\mathrm{I}|+|\mathrm{J}|}{(|\mathrm{I}||\mathrm{J}|)^{1/2}\lambda_\varrho} + \frac{1}{(|\mathrm{I}||\mathrm{J}|)^{1/2}}.$$

The first term is $\ll q^{-1/2+\delta/2}$ and the third is $\ll 1$. For the second term, we already observed that since $\varrho^2 \neq 1$, any solution of the congruence $m \equiv \varrho n$ satisfies $\max(|m|,|n|) \gg q^{1/a}$, and therefore $\lambda_\varrho \gg q^{1/a}$. Thus

$$\frac{|\mathrm{I}|+|\mathrm{J}|}{(|\mathrm{I}||\mathrm{J}|)^{1/2}\lambda_\varrho} \leqslant 2\frac{\mathrm{K}_+}{\mathrm{K}_-\lambda_\varrho} \ll q^{\eta-1/a},$$

and consequently

$$\text{(7.6)}\qquad \frac{\mathrm{M}^*_{1,-1}((\varrho,1),\mathrm{I}\times\mathrm{J};q)}{(|\mathrm{I}||\mathrm{J}|)^{1/2}} \ll q^{-1/2+\delta/2} + q^{\eta-1/a}.$$

From (7.3) and the bounds (7.5) and (7.6), we deduce

$$\mathrm{M}^{\mathrm{err}}_a(q,\varrho) \ll (\log q)^2(q^{-\eta/2} + q^{-1/2+\delta/2} + q^{\eta-1/a}).$$

Using (7.1), we conclude that

$$\mathrm{M}^{\mathrm{even}}_a(q) = \frac{1}{2}\log q + \mathrm{C} + \mathrm{O}\Big(q^{-1/2} + (\log q)^2(q^{-\eta/2} + q^{-1/2+\delta/2} + q^{\eta-1/a})\Big).$$

Taking, e.g., $\delta = 1/4$ and $\eta = 1/(2a)$, we conclude that

$$\mathrm{M}^{\mathrm{even}}_a(q) = \frac{1}{2}\log q + \mathrm{C} + \mathrm{O}(q^{-c/a})$$

for some absolute constant $c > 0$.

Finally, one finds by looking at the functional equation that the contribution of odd characters is

$$\mathrm{M}^{\mathrm{odd}}_a(q) = \sum_{m^a \equiv n^a \,(\mathrm{mod}\, q)} \frac{1}{\sqrt{mn}} \mathrm{V}\Big(\frac{mn}{q}\Big) - \sum_{m^a \equiv -n^a \,(\mathrm{mod}\, q)} \frac{1}{\sqrt{mn}} \mathrm{V}\Big(\frac{mn}{q}\Big),$$

and hence the same analysis as above gives the asymptotic formula

$$\mathrm{M}_a^{\mathrm{odd}}(q) = \frac{1}{2}\log q + \mathrm{C} + \mathrm{O}(q^{-c/a}),$$

so that Theorem 1.1 follows by adding these two terms.

Remark 7.1. A similar argument leads to the following result.

Proposition 7.2. *Let $f \in \mathbf{Z}[\mathrm{X}]$ be a polynomial of degree $d \geqslant 2$, irreducible over $\mathbf{Q}$. There exists $\delta > 0$, depending only on d, such that the estimate*

$$\frac{1}{q-1}\sum_{\chi \,(\mathrm{mod}\, q)} \chi(\varrho)|\mathrm{L}(\tfrac{1}{2},\chi)|^2 \ll q^{-\delta}$$

holds for all primes q such that f has at least one root modulo q and ϱ is an arbitrarily chosen root of $f\,(\mathrm{mod}\, q)$.

More precisely, the application of the approximate functional equation shows that this follows from the bound

$$\sum_{\substack{m,n\geqslant 1\\ (\varrho m)^2\equiv n^2 \,(\mathrm{mod}\, q)}} \frac{1}{\sqrt{mn}}\mathrm{V}\Big(\frac{mn}{q}\Big) \ll q^{-\delta},$$

which is estimated exactly like $\mathrm{M}_a^{\mathrm{err}}(q,\varrho)$ above. The point is once again that the minimum of the lattices

$$\{(m,n)\in\mathbf{Z}^2 \mid \varrho m \equiv n\,(\mathrm{mod}\, q)\}, \quad \{(m,n)\in\mathbf{Z}^2 \mid -\varrho m \equiv n\,(\mathrm{mod}\, q)\}$$

are large, namely of size $\gg q^{1/d}$.

It would be interesting to study "root twists" of this kind for other families, e.g., to bound

$$\sum_{\chi \,(\mathrm{mod}\, q)} \chi(\varrho)|\mathrm{L}(\tfrac{1}{2},\chi)|^4, \qquad \sum_{f\in \mathrm{S}_k(q)} \lambda_f(\varrho)|\mathrm{L}(\tfrac{1}{2},f)|^2,$$

where $\mathrm{S}_k(q)$ is the family of Hecke eigenforms of conductor q with Hecke eigenvalues $\lambda_f(n)$.

References

[1] A. Beilinson, J. Bernstein, P. Deligne and O. Gabber: *Perverse sheaves*, Astérisque 100, S.M.F, 2nd edition, 2018.

[2] N. Bhardwaj and L. van den Dries: *On the Pila-Wilkie theorem*, Expo. Math. 40 (2022), 495–542.

[3] V. Blomer, D. Milićević, É. Fouvry, E. Kowalski, Ph. Michel and W. Sawin: *The second moment theory of families of* L*-functions*, Memoirs of the AMS 1394, 2023.

[4] V. Chandee and X. Li: *The eighth moment of Dirichlet* L*-functions*, Adv. Math. 259 (2014), 339–375.

[5] B. Conrey, H. Iwaniec and K. Soundararajan: *The sixth power moment of Dirichlet* L*-Functions*, GAFA 22 (2012), 1257–1288.

[6] S. Das, R. Khan: *Simultaneous nonvanishing of Dirichlet* L*-functions and twists of Hecke-Maass* L*-functions*, J. Ramanujan Math. Soc. 30 (2015), 237–250.

[7] É. Fouvry: *Consequences of a result of N. Katz and G. Laumon concerning trigonometric sums*, Israel J. Math. 120 (2000), part A, 81–96.

[8] É. Fouvry and N.M. Katz: *A general stratification theorem for exponential sums, and applications*, J. reine angew. Math. 540 (2001), 115–166.

[9] É. Fouvry, E. Kowalski, Ph. Michel: *Algebraic twists of modular forms and Hecke orbits*, Geom. Funct. Anal. 25 (2015), no. 2, 580–657.

[10] É. Fouvry, E. Kowalski, Ph. Michel: *Algebraic trace weights over the primes*, Duke Math. J. 163 (2014), no. 9, 1683–1736.
[11] D.R. Heath-Brown and L. Pierce: *Averages and moments associated to class numbers of imaginary quadratic fields*, Compositio Math. 153 (2017), 2287–2309.
[12] D.R. Heath-Brown: *The least square-free number in an arithmetic progression*, J. reine angew. Math. 332 (1982), 204–220.
[13] H. Iwaniec and E. Kowalski: *Analytic number theory*, Colloquium Publ. 53, AMS, 2004.
[14] H. Iwaniec and P. Sarnak: *The non-vanishing of central values of automorphic* L*-functions and Landau-Siegel zeros*, Israel J. Math. 120 (2000), part A, 155–177.
[15] H. Iwaniec and P. Sarnak: *Perspectives on the analytic theory of L-functions*, GAFA 2000 (Tel Aviv, 1999), Geom. Funct. Anal. 2000, Special Volume, Part II, 705–741.
[16] H. Iwaniec and P. Sarnak: *Dirichlet* L*-functions at the central point*, in "Number theory in progress", Vol. 2, 941–952, de Gruyter, 1999.
[17] N.M. Katz: *Exponential sums and differential equations*, Annals of Math. Studies 124, Princeton Univ. Press, 1990.
[18] N.M. Katz and G. Laumon: *Transformation de Fourier et majoration de sommes exponentielles*, Publ. Math. IHÉS 62 (1985), 145–202.
[19] R. Khan, D. Milićević and H.T. Ngo: *Non-vanishing of Dirichlet* L*-functions in Galois orbits*, I.M.R.N (2016), 6955–6978.
[20] E. Kowalski: *Families of cusp forms*, Publ. Math. Besançon Algèbre Théorie Nr. (2013), 5–40.
[21] E. Kowalski, Ph. Michel and W. Sawin: *Bilinear forms with Kloosterman sums and applications*, Ann. of Math. 186 (2017), 413–500.
[22] E. Kowalski, Ph. Michel and W. Sawin: *Stratification and averaging for exponential sums: bilinear forms with generalized Kloosterman sums*, Ann. Sc. Norm. Super. Pisa 21 (2020), 1453–1530.
[23] Y. Lamzouri: *Distribution of values of* L*-functions at the edge of the critical strip*, Proceedings L.M.S 100 (2010), 835–863.
[24] Ph. Michel, A. Venkatesh: *Heegner points and non-vanishing of Rankin-Selberg* L*-functions*, Analytic number theory, Clay Math. Proc. Vol. 7 (2007), 169–183.
[25] A.B. Nordentoft: *Wide moments of* L*-functions II: Dirichlet* L*-functions*, Quarterly J. Math. 2022, haac026, https://doi.org/10.1093/qmath/haac026
[26] A.B. Nordentoft: *Wide moments of* L*-functions I: Twists by class group characters of imaginary quadratic fields*, preprint (2021), https://arxiv.org/abs/2105.09130.
[27] R. Nunes: *On the least squarefree number in an arithmetic progression*, Mathematika 63 (2017), 483–498.
[28] R.E.A.C. Paley: *On the k-Analogues of some Theorems in the Theory of the Riemann ζ-Function*, Proc. London Math. Soc. (2) 32 (1931), 273–311.
[29] I. Petrow: *The Weyl law for algebraic tori*, preprint (2018), arXiv:1808.09991.
[30] L. Pierce: *The* 3*-part of class numbers of quadratic fields*, Journal London Math. Soc. 71 (2005), 579–598.
[31] J. Pliego: *Mixed moments of the Riemann zeta function*, preprint (2022), arXiv:2210.15321.
[32] J. Pliego: *Twisted mixed moments of the Riemann zeta function*, preprint (2022), arXiv:2211.11450.
[33] M. Radziwiłł and K. Soundararajan: *Value distribution of* L*-functions*, Oberwolfach report, 40/2017.
[34] P. Sarnak, S.W. Shin and N. Templier: *Families of* L*-functions and their symmetry*, in "Families of automorphic forms and the trace formula", 531–578, Simons Symp., Springer, 2016.
[35] M. Widmer: *Counting primitive points of bounded height*, Transactions A.M.S 362 (2010), 4793–4829.
[36] R. Zacharias: *Simultaneous non-vanishing for Dirichlet* L*-functions*, Ann. Inst. Fourier (Grenoble) 69 (2019), 1459–1524.

Université Paris–Saclay, CNRS, Laboratoire de Mathématiques d'Orsay, 91405 Orsay, France
Email address: etienne.fouvry@universite-paris-saclay.fr

ETH Zürich – D-MATH, Rämistrasse 101, CH-8092 Zürich, Switzerland
Email address: kowalski@math.ethz.ch

EPFL/SB/IMB/TAN, Station 8, CH-1015 Lausanne, Switzerland
Email address: philippe.michel@epfl.ch