J. Eur. Math. Soc. 26, 1713–1746 (2024) DOI 10.4171/JEMS/1443

Alina Bucur · Francesc Fité · Kiran S. Kedlaya

Effective Sato–Tate conjecture for abelian varieties and applications

Received March 16, 2020; revised September 25, 2023

Abstract. From the generalized Riemann hypothesis for motivic L-functions, we derive an effective version of the Sato–Tate conjecture for an abelian variety A defined over a number field k with connected Sato–Tate group. By *effective* we mean that we give an upper bound on the error term in the count predicted by the Sato–Tate measure that only depends on certain invariants of A. We discuss three applications of this conditional result. First, for an abelian variety defined over k, we consider a variant of Linnik's problem for abelian varieties that asks for an upper bound on the least norm of a prime whose normalized Frobenius trace lies in a given interval. Second, for an elliptic curve defined over k with complex multiplication, we determine (up to multiplication by a nonzero constant) the asymptotic number of primes whose Frobenius traces attain the integral part of the Hasse–Weil bound. Third, for a pair of abelian varieties A and A' defined over k with no common factors up to k-isogeny, we find an upper bound on the least norm of a prime at which the respective Frobenius traces of A and A' have opposite sign.

Keywords: effective Sato–Tate conjecture, abelian varieties over number fields, generalized Riemann hypothesis, Frobenius sign separation, variants of Linnik's problem.

1. Introduction

Let A be an abelian variety defined over a number field k of dimension $g \geq 1$. For a rational prime ℓ, we denote by

$$\varrho_{A,\ell}\colon G_k \to \mathrm{Aut}(V_\ell(A))$$

the ℓ-adic representation attached to A, obtained from the action of the absolute Galois group of k on the rational ℓ-adic Tate module $V_\ell(A) := T_\ell(A) \otimes \mathbb{Q}_\ell$. Let N denote the

Alina Bucur: Department of Mathematics, University of California San Diego, La Jolla, CA 92093, USA; alina@math.ucsd.edu

Francesc Fité: Departament de matemàtiques i informàtica, Universitat de Barcelona, 08007 Barcelona, Spain; ffite@ub.edu

Kiran S. Kedlaya: Department of Mathematics, University of California San Diego, La Jolla, CA 92093, USA; kedlaya@ucsd.edu

Mathematics Subject Classification 2020: 11G10 (primary); 11G05, 11R44 (secondary).

absolute norm of the conductor of A, which we will call the absolute conductor of A. For a nonzero prime ideal $\mathfrak{p}$ of the ring of integers of k not dividing $N\ell$, let $a_\mathfrak{p} := a_\mathfrak{p}(A)$ denote the trace of $\varrho_{A,\ell}(\mathrm{Frob}_\mathfrak{p})$, where $\mathrm{Frob}_\mathfrak{p}$ is a Frobenius element at $\mathfrak{p}$. The trace $a_\mathfrak{p}$ is an integer which does not depend on ℓ and, denoting by $\mathrm{Nm}(\mathfrak{p})$ the absolute norm of $\mathfrak{p}$, the Hasse–Weil bound asserts that the normalized trace

$$\overline{a}_\mathfrak{p} := \frac{a_\mathfrak{p}}{\sqrt{\mathrm{Nm}(\mathfrak{p})}}$$

lies in the interval $[-2g, 2g]$.

Attached to A there is a compact real Lie subgroup $\mathrm{ST}(A)$ of the unitary symplectic group $\mathrm{USp}(2g)$ that conjecturally governs the distribution of the normalized Frobenius traces. More precisely, the Sato–Tate conjecture predicts that the sequence $\{\overline{a}_\mathfrak{p}\}_\mathfrak{p}$, indexed by primes $\mathfrak{p}$ not dividing N ordered by norm, is equidistributed on the interval $[-2g, 2g]$ with respect to the pushforward via the trace map of the (normalized) Haar measure of the Sato–Tate group $\mathrm{ST}(A)$. We will denote this pushforward measure by μ.

Denote by δ_I the characteristic function of a subinterval I of $[-2g, 2g]$. Together with the prime number theorem, the Sato–Tate conjecture predicts that

$$\sum_{\mathrm{Nm}(\mathfrak{p})\le x} \delta_I(\overline{a}_\mathfrak{p}) \sim \mu(I)\,\mathrm{Li}(x) \quad \text{as } x \to \infty, \tag{1.1}$$

where $\mathrm{Li}(x) := \int_2^\infty dt/\log(t)$. Let $L(\chi, s)$ denote the (normalized) L-function attached to an irreducible character χ of $\mathrm{ST}(A)$. It is well known that (1.1) is implied by the conjectural nonvanishing and analyticity on the right half-plane $\Re(s) \ge 1$ of $L(\chi, s)$ for every nontrivial irreducible character χ. In this paper we derive an asymptotic upper bound on the error term implicit in (1.1) by further assuming the generalized Riemann hypothesis for the L-functions $L(\chi, s)$.

Our main result is a quantitative refinement of the Sato–Tate conjecture (see Theorem 3.8). In order to state it we need to introduce some notations. Let $\mathfrak{g}$ denote the complexified Lie algebra of $\mathrm{ST}(A)$, and write it as $\mathfrak{s} \times \mathfrak{a}$, where $\mathfrak{s}$ is semisimple and $\mathfrak{a}$ is abelian. Set

$$\varepsilon_\mathfrak{g} := \frac{1}{2(q+\varphi)}, \tag{1.2}$$

where φ is the size of the set of positive roots of $\mathfrak{s}$ and q is the rank of $\mathfrak{g}$, and define

$$\nu_\mathfrak{g}\colon \mathbb{R}_{>0} \to \mathbb{R}_{>0}, \quad \nu_\mathfrak{g}(z) = \max\left\{1, \frac{\log(z)^6}{z^{1/\varepsilon_\mathfrak{g}}}\right\} \tag{1.3}$$

For a subinterval I of $[-2g, 2g]$, let $|I|$ denote its length.

Theorem 1.1 (Effective Sato–Tate conjecture). *Let A be an abelian variety defined over the number field k of dimension $g \ge 1$, absolute conductor N, and such that* $\mathrm{ST}(A)$ *is connected. Suppose that the Mumford–Tate conjecture holds for A and that the generalized Riemann hypothesis holds for $L(\chi, s)$ for every irreducible character χ of* $\mathrm{ST}(A)$.

Then for all subintervals I of $[-2g, 2g]$ of nonzero length, we have

$$\sum_{\mathrm{Nm}(\mathfrak{p})\le x} \delta_I(\overline{a}_{\mathfrak{p}}) = \mu(I)\,\mathrm{Li}(x) + O\left(\frac{x^{1-\varepsilon_{\mathfrak{g}}}\log(Nx)^{2\varepsilon_{\mathfrak{g}}}}{\log(x)^{1-4\varepsilon_{\mathfrak{g}}}}\right) \quad \text{for } x \ge x_0, \tag{1.4}$$

where the sum runs over primes not dividing N, the implied constant in the O-notation depends exclusively on k and g, and $x_0 = O(\nu_{\mathfrak{g}}(|I|)\log(2N)^2\log(\log(4N))^4)$.

The dependence on g in the implied constant of (1.4) can be traced through Propositions 3.1 and 3.2 and Lemmas 3.6 and 3.7; it is highly exponential. This theorem generalizes a result of Murty [25] concerning elliptic curves without complex multiplication (CM); see also [4, Theorem 3.1]. Its proof follows the strategy envisaged in [4, Section 5] and it occupies Section 3. A key ingredient is the construction of a multivariate Vinogradov function; this is a continuous periodic function, with rapidly decaying Fourier coefficients, and approximating the characteristic function of the preimage of I by the trace map in the parameter space of a Cartan subgroup H of $\mathrm{ST}(A)$. By identifying the quotient of this space by the action of the Weyl group with the set of conjugacy classes of $\mathrm{ST}(A)$, one can rewrite (a Weyl average of) the Vinogradov function as a combination of irreducible characters of $\mathrm{ST}(A)$. One can use purely Lie algebra-theoretic arguments (most notably Weyl's character dimension formula and a result due to Gupta [19, Theorem 3.8] on the boundedness of the inverse of the weight multiplicity matrix) to show that the coefficients in the character decomposition of the Vinogradov function also exhibit a rapid decay. The theorem can then be obtained by using an estimate of Murty (as presented in [4, (2.4)]) on truncated sums of an irreducible character χ over the prime ideals of k. The implied constant in the O-notation depends in principle on the exponents of the Cartan subgroup H. In order to bound these exponents purely in terms of g, we show that the Mumford–Tate conjecture implies that H is generated by the Hodge circles contained in it (see Theorem 3.5). This result may be of independent interest.

The conjectural background for Theorem 1.1 is presented in Section 2. We recall the Mumford–Tate conjecture and the related algebraic Sato–Tate conjecture, define the L-functions $L(\chi, s)$, and state the generalized Riemann hypothesis for them. In Section 4 we give three applications of Theorem 1.1. The first is what we call the *interval variant* of Linnik's problem for an abelian variety (see Corollary 4.1).

Corollary 1.2. *Assume the hypotheses and notations of Theorem 1.1. For every subinterval I of $[-2g, 2g]$ of nonzero length, there exists a prime $\mathfrak{p}$ not dividing N with norm*

$$\mathrm{Nm}(\mathfrak{p}) = O\big(\nu_{\mathfrak{g}}(\min\{|I|, \mu(I)\})\log(2N)^2\log(\log(4N))^4\big)$$

such that $\overline{a}_{\mathfrak{p}} \in I$.

The second application concerns what we call the *Frobenius sign separation problem* for a pair of abelian varieties (see Corollary 4.4).

Corollary 1.3. *Let A (resp. A') be an abelian variety defined over the number field k of dimension $g \ge 1$ (resp. $g' \ge 1$), absolute conductor N (resp. N'), and such that*

$\mathrm{ST}(A)$ *(resp.* $\mathrm{ST}(A')$*) is connected. Suppose that the Mumford–Tate conjecture holds for* A *(resp.* A'*) and that the generalized Riemann hypothesis holds for* $L(\chi, s)$ *(resp.* $L(\chi', s)$*) for every irreducible character* χ *of* $\mathrm{ST}(A)$ *(resp.* χ' *of* $\mathrm{ST}(A')$*). Suppose that* $\mathrm{ST}(A \times A') \simeq \mathrm{ST}(A) \times \mathrm{ST}(A')$*. Then there exists a prime* $\mathfrak{p}$ *not dividing* NN' *with norm*

$$\mathrm{Nm}(\mathfrak{p}) = O\big(\log(2NN')^2 \log(\log(4NN'))^6\big)$$

such that $a_{\mathfrak{p}}(A)$ *and* $a_{\mathfrak{p}}(A')$ *are nonzero and of opposite sign. Here, the implied constant in the* O*-notation depends exclusively on* k*,* g*, and* g'*.*

We also examine what our method says about the set of primes with "maximal Frobenius trace". Let $M_k(x)$ denote the set of primes $\mathfrak{p}$ not dividing N with norm up to x for which $a_{\mathfrak{p}} = \lfloor 2\sqrt{\mathrm{Nm}(\mathfrak{p})} \rfloor$. Vaguely formulated, a natural approach to compute (at least an asymptotic lower bound on) $M_k(x)$ is to compute the number of $\mathfrak{p}$ with norm up to x for which $\overline{a}_{\mathfrak{p}}$ lies in a sufficiently small neighborhood I_x of $2g$. However, for this idea to succeed, the neighborhood I_x should be sufficiently large in order for the "error term" in (1.4) to be still dominated by the "main term", which is now multiplied by the tiny quantity $\mu(I_x)$. In the case where A is an elliptic curve with CM it is possible to achieve this trade-off, yielding the following statement (see Proposition 4.9 and Corollary 4.10).

Corollary 1.4. *Let* A *be an elliptic curve defined over* k *with potential CM, that is, such that* $A_{\overline{\mathbb{Q}}}$ *has CM. Under the generalized Riemann hypothesis for the* L*-function attached to every power of the Hecke character of* A*, we have*

$$\#M_k(x) \asymp \frac{x^{3/4}}{\log(x)} \quad \text{as } x \to \infty.$$

This recovers a weaker version of a theorem of James and Pollack [21, Theorem 1], which asserts (unconditionally) that

$$\#M_k(x) \sim \frac{2}{3\pi} \frac{x^{3/4}}{\log(x)}.$$

A different result in a similar spirit, concerning numbers of points on diagonal curves, is due to Duke [11, Theorem 3.3].

Corollary 1.3 extends work of Bucur and Kedlaya [4, Theorem 4.3], who considered the case in which A and A' are elliptic curves without CM. Later Chen, Park, and Swaminathan [6, Theorem 1.3] reexamined this case, obtaining an upper bound of the form $O(\log(NN')^2)$ and relaxing the generalized Riemann hypothesis assumed in [4, Theorem 4.3]. Corollary 1.2 extends [6, Theorem 1.8], which again applies to elliptic curves without CM. It should be noted that the aforementioned results in [6] make explicit the constants involved in the respective upper bounds, a goal which we have not pursued in our work.

The framework of the generalized Sato–Tate conjecture includes many additional questions about distinguishing L-functions, a number of which have been considered previously. For instance, Goldfeld and Hoffstein [17] established an upper bound on the

first distinguishing coefficient for a pair of holomorphic Hecke newforms, by an argument similar to ours but with a milder analytic hypothesis (the Riemann hypothesis for the Rankin–Selberg convolutions of the two forms with themselves and each other). Sengupta [26] carried out the analogous analysis with the Fourier coefficients replaced by normalized Hecke eigenvalues (this only makes a difference when the weights are distinct).

There is an alternative approach to the above kind of questions, which is based on the use of effective forms of Chebotaryov's density theorem conditional to Riemann hypothesis for Artin L-functions. This approach was introduced by Serre [29], who gave an upper bound on the smallest prime at which two nonisogenous elliptic curves have different Frobenius traces. The analogue of Serre's argument for modular forms was given by Ram Murty [24] and subsequently extended to Siegel modular forms by Ghitza [15] for Fourier coefficients and Ghitza and Sayer [16] for Hecke eigenvalues. Building on Serre's method, several recent works have explored the asymptotic number of zero Frobenius traces for abelian varieties which are either generic (see [9]) or isogenous to a product of elliptic curves (see [7, 8]).

Notation and terminology

Throughout this article, k is a fixed number field and g and g' are fixed positive integers. For an ordered set $(X, \leq)$ and functions $f, h\colon X \to \mathbb{R}$ we write $f(x) = O(h(x))$ to denote that there exist a real number $K > 0$ and an element $x_0 \in X$ such that $|f(x)| \leq Kh(x)$ for every $x \geq x_0$. We will generally specify the element x_0 in the statements of theorems, but we will usually obviate it in their proofs, where it can be inferred from the context. We refer to K as the *implied constant* in the O-notation. As we did in this introduction, whenever using the O-notation in a statement concerning an arbitrary abelian variety A of dimension g defined over the number field k, the corresponding implied constant is computable exclusively in terms of g and k (in fact the dependence on k is just on the absolute discriminant $|\mathrm{disc}_{k/\mathbb{Q}}|$ and the degree $[k : \mathbb{Q}]$). For statements concerning a pair of arbitrary abelian varieties A and A' of respective dimensions g and g' defined over k, the implied constant in the O-notation is computable exclusively in terms of g, g', and k. Section 4.3 is the only exception to the previous convention and to emphasize the dependence on N of the implied constants in the asymptotic bounds therein, we use the notations O_N and $\asymp_N$. We write $f \asymp g$ if $f = O(g)$ and $g = O(f)$. By a prime of k, we refer to a nonzero prime ideal of the ring of integers of k. Additional notation introduced later in the paper is summarized in Table 1.

2. Conjectural framework

Throughout this section, A will denote an abelian variety of dimension g defined over the number field k, and of absolute conductor $N := N_A$. We will define its Sato–Tate group, introduce the motivic L-functions attached to it, and present the conjectural framework on which Section 3 is sustained.

2.1. Sato–Tate groups

Following [32, Chapter 8] (see also [12, Section 2]), one defines the Sato–Tate group of A, denoted $\mathrm{ST}(A)$, in the following manner. Let G_ℓ^{Zar} denote the Zariski closure of the image of the ℓ-adic representation $\varrho_{A,\ell}$, which we may naturally see as lying in $\mathrm{GSp}_{2g}(\mathbb{Q}_\ell)$. Denote by $G_\ell^{1,\mathrm{Zar}}$ the intersection of G_ℓ^{Zar} with $\mathrm{Sp}_{2g}/\mathbb{Q}_\ell$. Fix an isomorphism $\iota\colon \overline{\mathbb{Q}}_\ell \simeq \mathbb{C}$ and let $G_{\ell,\iota}^{1,\mathrm{Zar}}$ denote the base change $G_\ell^{1,\mathrm{Zar}} \times_{\mathbb{Q}_\ell,\iota} \mathbb{C}$. The Sato–Tate group $\mathrm{ST}(A)$ is defined to be a maximal compact subgroup of the group of $\mathbb{C}$-points of $G_{\ell,\iota}^{1,\mathrm{Zar}}$. In the present paper, to avoid the a priori dependence on ℓ and ι of the definition of $\mathrm{ST}(A)$, we formulate the following conjecture.

Conjecture 2.1 (Algebraic Sato–Tate conjecture; [1]). *There exists an algebraic subgroup* $\mathrm{AST}(A)$ *of* $\mathrm{Sp}_{2g}/\mathbb{Q}$*, called the* algebraic Sato–Tate group*, such that* $G_\ell^{1,\mathrm{Zar}} \simeq \mathrm{AST}(A) \times_{\mathbb{Q}} \mathbb{Q}_\ell$ *for every prime* ℓ*.*

The Sato–Tate group $\mathrm{ST}(A)$ is then a maximal compact subgroup of $\mathrm{AST}(A) \times_{\mathbb{Q}} \mathbb{C}$. It should be noted that, following [31], Banaszak and Kedlaya [1] have given an alternative definition of $\mathrm{ST}(A)$ that also avoids the dependence on ℓ and ι. However, this is rendered mostly unnecessary by Theorem 2.3 below.

The algebraic Sato–Tate group is related to the Mumford–Tate group and the Hodge group. Fix an embedding $k \hookrightarrow \mathbb{C}$. The *Mumford–Tate group* $\mathrm{MT}(A)$ is the smallest algebraic subgroup G of $\mathrm{GL}(H_1(A_{\mathbb{C}}, \mathbb{Q}))$ over $\mathbb{Q}$ such that $G(\mathbb{R})$ contains $h(\mathbb{C}^\times)$, where

$$h\colon \mathbb{C} \to \mathrm{End}_{\mathbb{R}}(H_1(A_{\mathbb{C}}, \mathbb{R}))$$

is the complex structure on the $2g$-dimensional real vector space $H_1(A_{\mathbb{C}}, \mathbb{R})$ obtained by identifying it with the tangent space of A at the identity. The *Hodge group* $\mathrm{Hg}(A)$ is the intersection of $\mathrm{MT}(A)$ with $\mathrm{Sp}_{2g}/\mathbb{Q}$. Let $G_\ell^{\mathrm{Zar},0}$ (resp. $G_\ell^{1,\mathrm{Zar},0}$) denote the identity component of G_ℓ^{Zar} (resp. $G_\ell^{1,\mathrm{Zar}}$).

Conjecture 2.2 (Mumford–Tate conjecture). *There is an isomorphism*

$$G_\ell^{\mathrm{Zar},0} \simeq \mathrm{MT}(A) \times_{\mathbb{Q}} \mathbb{Q}_\ell.$$

Equivalently, we have $G_\ell^{1,\mathrm{Zar},0} \simeq \mathrm{Hg}(A) \times_{\mathbb{Q}} \mathbb{Q}_\ell$*.*

The identity component of $\mathrm{AST}(A)$ should thus be the Hodge group $\mathrm{Hg}(A)$. It follows from the definition that $\mathrm{ST}(A)$ has a faithful unitary symplectic representation

$$\varrho\colon \mathrm{ST}(A) \to \mathrm{GL}(V),$$

where V is a $2g$-dimensional $\mathbb{C}$-vector space. Via this representation, we regard $\mathrm{ST}(A)$ as a compact real Lie subgroup of $\mathrm{USp}(2g)$.

The following result has recently been established by Cantoral-Farfán–Commelin [5].

Theorem 2.3 (Cantoral-Farfán–Commelin). *If the Mumford–Tate conjecture holds for* A*, then the algebraic Sato–Tate conjecture also holds for* A*.*

2.2. *Motivic L-functions*

As described in [32, Section 8.3.3], to each prime $\mathfrak{p}$ of k not dividing N one can attach an element $y_{\mathfrak{p}}$ in the set of conjugacy classes Y of $\mathrm{ST}(A)$ with the property that

$$\det\bigl(1-\varrho_{A,\ell}(\mathrm{Frob}_{\mathfrak{p}})\,\mathrm{Nm}(\mathfrak{p})^{-1/2}T\bigr)=\det(1-\varrho(y_{\mathfrak{p}})T),$$

where $\mathrm{Frob}_{\mathfrak{p}}$ denotes a Frobenius element at $\mathfrak{p}$. More generally, via Weyl's unitarian trick, any complex representation

$$\sigma\colon \mathrm{ST}(A)\to \mathrm{GL}(V_\chi),$$

say of character χ and degree d_χ, gives rise to an ℓ-adic representation

$$\sigma_{A,\ell}\colon G_k\to \mathrm{Aut}(V_{\chi,\ell}),$$

where $V_{\chi,\ell}$ is a $\bar{\mathbb{Q}}_\ell$-vector space of dimension d_χ, such that for each prime $\mathfrak{p}$ of k not dividing N one has

$$\det\bigl(1-\sigma_{A,\ell}(\mathrm{Frob}_{\mathfrak{p}})\,\mathrm{Nm}(\mathfrak{p})^{-w_\chi/2}T\bigr)=\det(1-\sigma(y_{\mathfrak{p}})T),$$

where w_χ denotes the motivic weight of χ. For a prime $\mathfrak{p}$ of k, define

$$L_{\mathfrak{p}}(\chi,T):=\det\bigl(1-\sigma_{A,\ell}(\mathrm{Frob}_{\mathfrak{p}})\,\mathrm{Nm}(\mathfrak{p})^{-w_\chi/2}T\;\big|\;V_{\chi,\ell}^{I_{\mathfrak{p}}}\bigr),$$

where $I_{\mathfrak{p}}$ denotes the inertia subgroup of the decomposition group $G_{\mathfrak{p}}$ at $\mathfrak{p}$. The polynomials $L_{\mathfrak{p}}(\chi,T)$ do not depend on ℓ, and have degree $d_\chi(\mathfrak{p})\le d_\chi$. Moreover, writing $\alpha_{\mathfrak{p},j}$ for $j=1,\dots,d_\chi(\mathfrak{p})$ to denote the reciprocal roots of $L_{\mathfrak{p}}(\chi,T)$, we have

$$|\alpha_{\mathfrak{p},j}|\le 1.$$

In fact, for a prime $\mathfrak{p}$ not dividing N, we have $d_\chi(\mathfrak{p})=d_\chi$ and $|\alpha_{\mathfrak{p},j}|=1$. Therefore, the Euler product

$$L(\chi,s):=\prod_{\mathfrak{p}}L_{\mathfrak{p}}(\chi,\mathrm{Nm}(\mathfrak{p})^{-s})^{-1}$$

is absolutely convergent for $\Re(s)>1$. We will make strong assumptions on the analytic behavior of the above Euler product. Before doing so, following [27, Section 4.1], define the positive integer

$$B_\chi:=|\mathrm{disc}_{k/\mathbb{Q}}|^{d_\chi}\cdot N_\chi,$$

where N_χ is the absolute conductor attached to the ℓ-adic representation $\sigma_{A,\ell}$. For $j=1,\dots,d_\chi$, let $0\le\kappa_{\chi,j}\le 1+w_\chi/2$ be the local parameters at infinity (they are semi-integers that can be explicitly computed from the discussion in [27, Section 3]). Define the completed L-function

$$\Lambda(\chi,s):=B_\chi^{s/2}L(\chi,s)\Gamma(\chi,s),\quad\text{where}\quad\Gamma(\chi,s):=\pi^{d_\chi s/2}\prod_{j=1}^{d_\chi}\Gamma\left(\frac{s+\kappa_{\chi,j}}{2}\right).\tag{2.1}$$

Let $\delta(\chi)$ be the multiplicity of the trivial representation in the character χ of $\mathrm{ST}(A)$.

Conjecture 2.4 (Generalized Riemann hypothesis). *For every irreducible character χ of* $\mathrm{ST}(A)$*, the following holds:*

(i) *The function $s^{\delta(\chi)}(s-1)^{\delta(\chi)}\Lambda(\chi,s)$ extends to an analytic function on $\mathbb{C}$ of order* 1 *which does not vanish at $s = 0, 1$.*

(ii) *There exists $\epsilon \in \mathbb{C}^\times$ with $|\epsilon| = 1$ such that for all $s \in \mathbb{C}$ we have*

$$\Lambda(\chi, s) = \epsilon\Lambda(\overline{\chi}, 1-s),$$

where $\overline{\chi}$ is the character of the contragredient representation of σ.

(iii) *The zeros ρ of $\Lambda(\chi,s)$ (equivalently, the zeros ρ of $L(\chi,s)$ with $0 < \Re(\rho) < 1$) all have $\Re(\rho) = 1/2$.*

The following estimate of Murty [25, Proposition 4.1] will be crucial in Section 3. We will need the formulation with the level of generality of [4, (2.3)].

Proposition 2.5 (Murty's estimate). *Assume that Conjecture 2.4 holds for the irreducible character χ of* $\mathrm{ST}(A)$*. Then*

$$\sum_{\mathrm{Nm}(\mathfrak{p})\le x} \chi(y_\mathfrak{p})\log(\mathrm{Nm}(\mathfrak{p})) = \delta(\chi)x + O\big(d_\chi\sqrt{x}\log(x)\log(N(x+w_\chi))\big) \quad \text{for } x \ge 2. \tag{2.2}$$

By applying Abel's summation trick, the above gives

$$\sum_{\mathrm{Nm}(\mathfrak{p})\le x} \chi(y_\mathfrak{p}) = \delta(\chi)\,\mathrm{Li}(x) + O\big(d_\chi\sqrt{x}\log(N(x+w_\chi))\big) \quad \text{for } x \ge 2. \tag{2.3}$$

Remark 2.6. In (2.2) and thereafter, we make the convention that all sums involving the classes $y_\mathfrak{p}$ run over primes $\mathfrak{p}$ not dividing N. A similar convention applies for sums involving the normalized Frobenius traces $\overline{a}_\mathfrak{p} = \mathrm{Trace}(y_\mathfrak{p})$.

Remark 2.7. We alert the reader to a small discrepancy between (2.3) and [4, (2.4)]: in the latter, the error term stated is $O(d_\chi\sqrt{x}\log(N(x+d_\chi)))$. We make this precise here, although we note that it has no effect on the subsequent results of [4]. Indeed, in many cases (as those of interest in [4] involving elliptic curves without CM) the weight w_χ is bounded by the dimension d_χ.

Remark 2.8. The proof of Proposition 2.5 uses the bound

$$\log(B_\chi) = O(d_\chi\log(N)) \quad \text{for every character } \chi \text{ of } \mathrm{ST}(A). \tag{2.4}$$

In order to show (2.4), let us recall the definition of N_χ as a product

$$N_\chi := \prod_\mathfrak{p} \mathrm{Nm}(\mathfrak{p})^{f_\chi(\mathfrak{p})}$$

over primes of k, where $f_\chi(\mathfrak{p})$ is the exponent conductor at $\mathfrak{p}$; this is a nonnegative integer whose definition can be found in [27, Section 2], for example. If A has good reduction at $\mathfrak{p}$, then $f_\chi(\mathfrak{p})$ is zero and so the product is finite. Let $T_{\chi,\ell}$ denote a $\mathbb{Z}_\ell$-lattice in $V_{\chi,\ell}$

stable by the action of $G_{\mathfrak{p}}$. By Grothendieck [18, Section 4], the exponent conductor can be written as

$$f_\chi(\mathfrak{p}) = \varepsilon_\chi(\mathfrak{p}) + \delta_\chi(\mathfrak{p}),$$

where $\varepsilon_\chi(\mathfrak{p}) = d_\chi - \dim(V_{\chi,\ell}^{I_{\mathfrak{p}}})$ and $\delta_\chi(\mathfrak{p})$ is the Swan conductor of $V_\chi[\ell] := T_{\chi,\ell}/\ell T_{\chi,\ell}$ for every ℓ coprime to $\mathfrak{p}$. Since the kernel of the action

$$\overline{\sigma}_{A,\ell}\colon G_{\mathfrak{p}} \to \mathrm{Aut}(V_\chi[\ell])$$

on this quotient is contained in the kernel of the action of $G_{\mathfrak{p}}$ on $T_\ell(A)/\ell T_\ell(A)$, we see that $\overline{\sigma}_{A,\ell}$ factors through a finite group $G_{\chi,\mathfrak{p}}$ whose order is $O(1)$. Consider the normal filtration of ramification groups

$$G_{\chi,\mathfrak{p}} \supseteq G_0 \supseteq G_1 \supseteq \cdots$$

of $G_{\chi,\mathfrak{p}}$. Let us simply write V (resp. V_i) for $V_\chi[\ell]$ (resp. $V_\chi[\ell]^{G_i}$). By [3, Proposition 5.4], we have

$$f_\chi(\mathfrak{p}) = \dim(V/V_0) + \big(a + h(G_1) + 1/(p-1)\big)e\dim(V/V_1),$$

where e is the ramification index of $\mathfrak{p}$ over $\mathbb{Q}$, $p^{h(G_1)}$ is the exponent of the p-group G_1, and p^a is the maximal dimension among absolutely simple components of V/V_1 as a G_1-module. Since $\#G_1$ is $O(1)$, so are $h(G_1)$ and a, because the dimension of an irreducible representation of a group is bounded by the order of the group. We deduce that

$$f_\chi(\mathfrak{p}) = O(d_\chi),$$

from which (2.4) is immediate.

3. Effective Sato–Tate Conjecture

In this section we derive, from the conjectural framework described in Section 2, an effective version of the Sato–Tate conjecture for an arbitrary abelian variety A of dimension g defined over the number field k (see Theorem 3.8). Let I be a subinterval of $[-2g, 2g]$. By *effective* we mean that we provide an upper bound on the error term in the count of primes with normalized Frobenius trace lying in I relative to the prediction made by the Sato–Tate measure.

The proof is based on the strategy hinted at in [4, Section 5]. The first step is the construction of a multivariate Vinogradov function aproximating the characteristic function of the preimage of I by the trace map. This is a continuous periodic function with rapidly decaying Fourier coefficients that generalizes the classical Vinogradov function [34, Lemma 12]. This construction is accomplished in Section 3.2.

The core of the proof consists in rewriting the Vinogradov function in terms of the irreducible characters of $\mathrm{ST}(A)$ and applying Murty's estimate (Proposition 2.5) to each of its irreducible constituents. This is the content of Section 3.4.

In order to control the size of the coefficients of the character decomposition, we use a result of Gupta [19, Theorem 3.8] bounding the size and number of nonzero entries of the inverse of the weight multiplicity matrix. Gupta's result and other background material on representations of Lie groups is recalled in Section 3.1.

A first analysis does not yield the independence of the implied constant in the O-notation from the Lie algebra of $\mathrm{ST}(A)$. This independence is shown to follow from the density of the subgroup generated by the Cartan Hodge circles in the Cartan subgroup. In a result which may be of independent interest (see Theorem 3.5), this density is shown to follow from the Mumford–Tate conjecture in Section 3.3.

3.1. Lie group theory background

Let $\mathfrak{s}$ be a finite-dimensional complex semisimple Lie algebra with Cartan subalgebra $\mathfrak{h}$ of rank h. Let $\Phi \subseteq \mathfrak{h}^*$ be a root system for $\mathfrak{s}$, $\mathfrak{h}_0^*$ be the real vector subspace generated by Φ, and $\mathcal{R} \subseteq \mathfrak{h}_0^*$ denote the lattice of integral weights of $\mathfrak{s}$.

Fix a base S for the root system Φ. The choice of S determines a Weyl chamber in $\mathfrak{h}_0^*$ and a partition $\Phi = \Phi^+ \cup \Phi^-$, where Φ^+ (resp. Φ^-) denotes the set of positive (resp. negative) roots of $\mathfrak{s}$. Let $\mathcal{C}$ denote the set of dominant weights, that is, the intersection of the set of integral weights $\mathcal{R}$ with this Weyl chamber. The choice of a basis $\{\omega_j\}_j$ of fundamental weights determines an isomorphism $\mathcal{C} \simeq \mathbb{Z}^h_{\geq 0}$.

For $\lambda, \mu \in \mathcal{C}$, the multiplicity m_λ^μ of μ in λ is defined to be the dimension of the space

$$\Gamma_\lambda^\mu = \{v \in \Gamma_\lambda;\ b(v) = \mu(b)v,\ \forall b \in \mathfrak{h}\},$$

where Γ_λ is the irreducible representation of $\mathfrak{s}$ of highest weight λ. Write

$$\rho := \frac{1}{2} \sum_{\alpha \in \Phi^+} \alpha$$

for the Weyl vector and $\mathcal{W}$ for the Weyl group of $\mathfrak{s}$. The multiplicity of μ in λ can be computed via Kostant's multiplicity formula

$$m_\lambda^\mu = \sum_{w \in \mathcal{W}} \epsilon(w) p(w(\lambda + \rho) - (\mu + \rho)), \tag{3.1}$$

where $\epsilon(w)$ is the sign of w, and $p(v)$ is defined by the identity

$$\sum_{v \in \mathcal{R}} p(v) e^v := \prod_{\alpha \in \Phi^+} (1 - e^\alpha)^{-1},$$

where we make a formal use of the exponential notation e^α (see [14, Proposition 25.21]). The natural number $p(v)$ is thus the number of ways to write the weight v as a sum of positive roots with nonnegative coefficients.

Write $\mu \preceq \lambda$ if and only if $\lambda - \mu$ is a sum of positive roots with nonnegative coefficients. The lattice $\mathcal{R} \subseteq \mathfrak{h}_0^*$ is then partially ordered with respect to the relation $\preceq$. Relative to this ordering of $\mathcal{C}$, the matrix $(m_\lambda^\mu)_{\lambda,\mu}$ of weight multiplicities is lower triangular. Let $(d_\lambda^\mu)_{\lambda,\mu}$ denote the inverse of $(m_\lambda^\mu)_{\lambda,\mu}$.

Gupta has obtained a formula[1] in the spirit of Kostant's multiplicity formula for the entries of the inverse matrix $(d_\lambda^\mu)_{\lambda,\mu}$. More precisely, by [19, Theorem 3.8], we have $d_\lambda^\mu = a_\lambda^\mu t_\lambda^{-1}$, where t_λ is the size of the stabilizer of λ in $\mathcal{W}$ and

$$a_\mu^\lambda := \sum_{w\in\mathcal{W}} \epsilon(w) f(w(\lambda+\rho)-\mu).$$

Here, for each $v \in \mathcal{R}$, the integer $f(v)$ is defined by

$$\sum_{v\in\mathcal{R}} f(v)e^v := e^\rho \prod_{\alpha\in\Phi^+} (1-e^{-\alpha}). \tag{3.2}$$

Let φ denote the size of the set Φ^+ of positive roots.

Proposition 3.1. *The sum of the absolute values of the elements in each row (resp. column) of $(d_\lambda^\mu)_{\lambda,\mu}$ is bounded by $\#\mathcal{W}\cdot 2^\varphi$. In particular, $d_\lambda^\mu = O(1)$ and the number of nonzero entries in each row (resp. column) of $(d_\lambda^\mu)_{\lambda,\mu}$ is $O(1)$.*

Proof. The proof follows from the aforementioned result by Gupta. Indeed, the sum of the absolute values of the entries in each row (resp. column) of $(d_\lambda^\mu)_{\lambda,\mu}$ is bounded by $\#\mathcal{W}$ times the norm

$$\sum_{v\in\mathcal{R}} |f(v)|.$$

But this number is bounded by 2^φ, as one observes from (3.2). Now the other two statements are implied by the fact that φ, $\#\mathcal{W}$ can be bounded in terms of g, as follows from the general classification of complex semisimple Lie algebras, and thus are $O(1)$. ∎

For $\lambda \in \mathcal{C}$, write λ as a nonnegative integral linear combination $\sum_{j=1}^s m_j\omega_j$ of the fundamental weights and define

$$\|\lambda\|_{\mathrm{fund}} := \max_j m_j.$$

Proposition 3.2. *The previous definition has the following properties:*

(i) $\dim(\Gamma_\lambda) = O(\|\lambda\|_{\mathrm{fund}}^\varphi)$ *for every* $\lambda \in \mathcal{C}$.

(ii) $\dim(\Gamma_{\lambda'}) = O(\dim(\Gamma_\lambda))$ *for all* $\lambda, \lambda' \in \mathcal{C}$ *with* $\lambda' \preceq \lambda$.

(iii) *For every* $\lambda \in \mathcal{C}$, *the motivic weight of the* ℓ*-adic representation* $(\Gamma_\lambda)_{A,\ell}$ *attached to* Γ_λ *as in Section 2 is* $O(\|\lambda\|_{\mathrm{fund}})$.

Proof. For (i), recall Weyl's dimension formula [30, Corollary 1 to Theorem 4, Chapter VII], which states

$$\dim(\Gamma_\lambda) = \prod_{\alpha\in\Phi^+} \frac{(\lambda+\rho,\alpha)}{(\rho,\alpha)},$$

[1] In fact, Gupta's result is of a more general nature: it applies to a q-analog of d_λ^μ. The version of interest to us is obtained by specialization.

where $(\cdot,\cdot)$ denotes a $\mathcal{W}$-invariant positive definite form on the real vector space $\mathfrak{h}_0^*$ spanned by the base S. This trivially implies

$$\dim(\Gamma_\lambda) \asymp \prod_{\alpha\in\Phi^+} (\lambda,\alpha).$$

It remains to show that $(\lambda,\alpha) = O(\|\lambda\|_{\mathrm{fund}})$ for every $\alpha \in \Phi^+$. Let α_j, for $j = 1,\dots,h$, be the constituents of the base S, the so-called *simple roots*. The desired result follows from the following relation linking simple roots and fundamental weights:

$$2\frac{(\omega_l,\alpha_j)}{(\alpha_j,\alpha_j)} = \delta_{lj}. \tag{3.3}$$

As for (ii), suppose that the expression of $\lambda \in \mathcal{C}$ (resp. $\lambda' \in \mathcal{C}$) as a nonnegative linear combination of the simple roots is $\sum_{j=1}^h r_j\alpha_j$ (resp. $\sum_{j=1}^h r'_j\alpha_j$). Note that $\lambda' \preceq \lambda$ implies that $r'_j \le r_j$. Therefore

$$\dim(\Gamma_{\lambda'}) = O\Big(\prod_{\alpha\in\Phi^+}(\lambda',\alpha)\Big) = O\Big(\prod_{\alpha\in\Phi^+}(\lambda,\alpha)\Big) = O(\dim(\Gamma_\lambda)).$$

Part (iii) is a consequence of the weight decomposition of Γ_λ. ∎

3.2. A multivariate Vinogradov function

The main result of this section is Proposition 3.4, which is a generalization of [34, Lemma 12]. Let $q \ge 1$ be a positive integer. We will write $\boldsymbol{\theta}$ to denote the q-tuple $(\theta_1,\dots,\theta_q) \in \mathbb{R}^q$ (a similar convention applies to $\boldsymbol{z}$, $\boldsymbol{\delta}$, etc.). We also write $\boldsymbol{m}$ to denote $(m_1,\dots,m_q) \in \mathbb{Z}^q$. We will say that a function $h\colon \mathbb{R}^q \to \mathbb{R}$ is *periodic* of period 1 if it is so in each variable.

For $\boldsymbol{\delta} = (\delta_1,\dots,\delta_q) \in [0,1)^q$, denote by $R(\boldsymbol{\delta})$ the parallelepiped $\prod_{j=1}^q[-\delta_j,\delta_j]$. Define also the multiplier

$$\nu(\boldsymbol{m},\boldsymbol{\delta}) := \prod_{j=1}^q \nu(m_j,\delta_j), \quad \text{where} \quad \nu(m_j,\delta_j) := \begin{cases} 1, & m_j = 0,\\ \dfrac{\sin(2\pi m_j\delta_j)}{2\pi m_j\delta_j}, & m_j \ne 0.\end{cases}$$

Lemma 3.3. *Suppose that $h\colon \mathbb{R}^q \to \mathbb{R}$ admits a Fourier series expansion*

$$h(\boldsymbol{\theta}) = \sum_{\boldsymbol{m}\in\mathbb{Z}^q} c_{\boldsymbol{m}}(h)e^{2\pi i(\boldsymbol{m}\cdot\boldsymbol{\theta})}, \quad \textit{where} \quad c_{\boldsymbol{m}}(h) := \int_{[0,1]^q} h(\boldsymbol{\theta})e^{-2\pi i(\boldsymbol{m}\cdot\boldsymbol{\theta})}\,d\boldsymbol{\theta}.$$

For $\boldsymbol{\delta} \in [0,1)^q$, define

$$f(\boldsymbol{\theta}) := \Big(\prod_{j=1}^q \frac{1}{2\delta_j}\Big)\int_{R(\boldsymbol{\delta})} h(\boldsymbol{\theta}+\boldsymbol{z})\,d\boldsymbol{z}. \tag{3.4}$$

Then

$$c_{\boldsymbol{m}}(f) = c_{\boldsymbol{m}}(h)\nu(\boldsymbol{m},\boldsymbol{\delta}). \tag{3.5}$$

Proof. The proof follows the same lines as Vinogradov's one-dimensional version. We have

$$\begin{aligned} c_{\boldsymbol{m}}(f) &= \int_{[0,1]^q} f(\boldsymbol{\theta}) e^{-2\pi i(\boldsymbol{m}\cdot\boldsymbol{\theta})}\, d\boldsymbol{\theta} \\ &= \left(\prod_{j=1}^{q} \frac{1}{2\delta_j}\right) \int_{[0,1]^q} \int_{R(\boldsymbol{\delta})} h(\boldsymbol{\theta}+\boldsymbol{z}) e^{-2\pi i(\boldsymbol{m}\cdot\boldsymbol{\theta})}\, d\boldsymbol{z}\, d\boldsymbol{\theta} \\ &= \left(\prod_{j=1}^{q} \frac{1}{2\delta_j}\right) \int_{R(\boldsymbol{\delta})} \int_{[0,1]^q} h(\boldsymbol{\theta}+\boldsymbol{z}) e^{-2\pi i(\boldsymbol{m}\cdot\boldsymbol{\theta})}\, d\boldsymbol{\theta}\, d\boldsymbol{z}. \end{aligned}$$

Setting $\boldsymbol{t} = \boldsymbol{\theta} + \boldsymbol{z}$ so that $\boldsymbol{\theta} = \boldsymbol{t} - \boldsymbol{z}$ and $d\boldsymbol{\theta} = d\boldsymbol{t}$ in the above equation, we obtain

$$\begin{aligned} c_{\boldsymbol{m}}(f) &= \int_{[0,1]^q} h(\boldsymbol{t}) e^{-2\pi i(\boldsymbol{m}\cdot\boldsymbol{t})}\, d\boldsymbol{t} \cdot \prod_{j=1}^{q} \frac{1}{2\delta_j} \int_{-\delta_j}^{\delta_j} e^{2\pi i m_j z_j}\, dz_j \\ &= c_{\boldsymbol{m}}(h) \prod_{j=1}^{q} \frac{1}{2\delta_j} \int_{-\delta_j}^{\delta_j} e^{2\pi i m_j z_j}\, dz_j. \end{aligned}$$

For $m_j = 0$ the corresponding term in the product is

$$\frac{1}{2\delta_j} \int_{-\delta_j}^{\delta_j} 1\, dz_j = 1 = v(0, \delta_j).$$

For $m_j \neq 0$ the corresponding term becomes

$$\frac{1}{2\delta_j} \int_{-\delta_j}^{\delta_j} e^{2\pi i m_j z_j}\, dz_j = \frac{1}{2\delta_j} \cdot \frac{e^{2\pi i m_j \delta_j} - e^{-2\pi i m_j \delta_j}}{2\pi i m_j} = \frac{\sin(2\pi m_j \delta_j)}{2\pi m_j \delta_j} = v(m_j, \delta_j).$$

The desired formula follows. ∎

For $1 \le j \le q$, let $\pi_j \colon [0,1]^q \to [0,1]^{q-1}$ be the map that sends $\boldsymbol{\theta} \in [0,1]^q$ to the $(q-1)$-tuple obtained from $\boldsymbol{\theta}$ by suppressing its j-th component. For $\boldsymbol{\vartheta} \in [0,1]^{q-1}$, define $X_j(\boldsymbol{\vartheta}) = \pi_j^{-1}(\boldsymbol{\vartheta})$.

Proposition 3.4. *Let $T \colon \mathbb{R}^q \to \mathbb{R}$ be a differentiable function satisfying the following hypotheses:*

(1) *It is periodic of period* 1.

(2) *There exists a real number $K > 0$ such that $|\nabla T(\boldsymbol{\theta})| \le K$ for every $\boldsymbol{\theta} \in \mathbb{R}^q$.*

(3) *There exists a positive integer $C > 0$ such that, for every $\gamma \in \mathbb{R}$, $1 \le j \le q$, and $\boldsymbol{\vartheta} \in [0,1]^{q-1}$, we have*

$$\#(T^{-1}(\gamma) \cap X_j(\boldsymbol{\vartheta})) \le C.$$

Let α, β, Δ be real numbers satisfying

$$\Delta > 0, \quad 2\Delta \le \beta - \alpha. \tag{3.6}$$

Let I denote the open interval (α, β). By (3.6) we can define the disjoint sets

$$R_1 := R_1(\Delta, \alpha, \beta) := T^{-1}((\alpha + \Delta, \beta - \Delta)) \cap [0,1]^q,$$
$$R_0 := R_0(\Delta, \alpha, \beta) := T^{-1}(\mathbb{R} \setminus [\alpha - \Delta, \beta + \Delta]) \cap [0,1]^q.$$

Then for every positive integer $r \geq 1$, there exists a continuous function $D := D_{\Delta,I} : \mathbb{R}^q \to \mathbb{R}$ periodic of period 1 satisfying the following properties:

(i) *For $\boldsymbol{\theta} \in R_1$, we have $D(\boldsymbol{\theta}) = 1$.*

(ii) *For $\boldsymbol{\theta} \in R_0$, we have $D(\boldsymbol{\theta}) = 0$.*

(iii) *$D(\boldsymbol{\theta})$ has a Fourier series expansion*

$$D(\boldsymbol{\theta}) = \sum_{\boldsymbol{m} \in \mathbb{Z}^q} c_{\boldsymbol{m}} e^{2\pi i (\boldsymbol{m} \cdot \boldsymbol{\theta})},$$

where $c_{\mathbf{0}} = \int_{T^{-1}((\alpha,\beta)) \cap [0,1]^q} d\boldsymbol{\theta}$ and for all $\boldsymbol{m} \neq 0$ we have

$$|c_{\boldsymbol{m}}| \leq \min\left\{|c_{\mathbf{0}}|, \left\{\frac{C}{\pi \max_j |m_j|} \prod_{j=1, m_j \neq 0}^{q} \min\left\{1, \left(\frac{rK\sqrt{q}}{2\pi |m_j| \Delta}\right)^{\rho}\right\}\right\}_{\rho=0,\ldots,r}\right\}.$$

Proof. Start by defining the function ψ_0 periodic of period 1 as

$$\psi_0(\boldsymbol{\theta}) := \begin{cases} 1 & \text{if } \boldsymbol{\theta} \in T^{-1}((\alpha, \beta)), \\ 0 & \text{if } \boldsymbol{\theta} \in T^{-1}(\mathbb{R} \setminus [\alpha, \beta]), \\ 1/2 & \text{if } \boldsymbol{\theta} \in T^{-1}(\alpha) \cup T^{-1}(\beta). \end{cases}$$

Then clearly

$$c_{\mathbf{0}}(\psi_0) = \int_{T^{-1}((\alpha,\beta)) \cap [0,1]^q} d\boldsymbol{\theta},$$

and for $\boldsymbol{m} \neq 0$ we find the bound

$$|c_{\boldsymbol{m}}(\psi_0)| = \left| \int_{T^{-1}((\alpha,\beta)) \cap [0,1]^q} e^{-2\pi i (\boldsymbol{m} \cdot \boldsymbol{\theta})} \, d\boldsymbol{\theta} \right| \leq |c_{\mathbf{0}}(\psi_0)|. \tag{3.7}$$

We next derive an alternative upper bound for $c_{\boldsymbol{m}}(\psi_0)$. Let m denote $\max_l |m_l|$ and let j be such that $m = |m_j|$. Then by Fubini's theorem we have

$$c_{\boldsymbol{m}}(\psi_0) = \int_{T^{-1}((\alpha,\beta)) \cap [0,1]^{q-1}} \left(\int_{T^{-1}((\alpha,\beta)) \cap X_j(\pi_j(\boldsymbol{\theta}))} e^{-2\pi i m \theta_j} \, d\theta_j \right) e^{-2\pi i \pi_j(\boldsymbol{m}) \cdot \pi_j(\boldsymbol{\theta})} \, d\pi_j(\boldsymbol{\theta}).$$

By condition (3) we find that $T^{-1}((\alpha, \beta)) \cap X_j(\pi_j(\boldsymbol{\theta}))$ is a union of at most C intervals. It follows that

$$|c_{\boldsymbol{m}}(\psi_0)| \leq 2C \frac{1}{2\pi m} = \frac{C}{\pi m}. \tag{3.8}$$

Fix $\delta > 0$ such that $r\sqrt{q}K\delta = \Delta$ and set $\boldsymbol{\delta} := (\delta, \dots, \delta)$. By averaging over the region $R(\boldsymbol{\delta})$ as in (3.4), we recursively define the function ψ_ρ, for $1 \le \rho \le r$, as

$$\psi_\rho(\boldsymbol{\theta}) = \frac{1}{(2\delta)^q} \int_{R(\boldsymbol{\delta})} \psi_{\rho-1}(\boldsymbol{\theta} + \boldsymbol{z})\, d\boldsymbol{z}. \tag{3.9}$$

We will prove inductively that:

(a) $\psi_\rho(\boldsymbol{\theta}) \in \mathbb{R}$ for all $\boldsymbol{\theta}$.

(b) $0 \le \psi_\rho(\boldsymbol{\theta}) \le 1$ for all $\boldsymbol{\theta}$.

(c) $\psi_\rho(\boldsymbol{\theta}) = 1$ for $\boldsymbol{\theta} \in T^{-1}((\alpha + \rho\Delta/r, \beta - \rho\Delta/r))$.

(d) $\psi_\rho(\boldsymbol{\theta}) = 0$ for $\boldsymbol{\theta} \in T^{-1}(\mathbb{R} \setminus [\alpha - \rho\Delta/r, \beta + \rho\Delta/r])$.

(e) $c_{\mathbf{0}}(\psi_\rho) = c_{\mathbf{0}}(\psi_0)$.

(f) For $\boldsymbol{m} \neq 0$,

$$c_{\boldsymbol{m}}(\psi_\rho) = c_{\boldsymbol{m}}(\psi_0)\nu(\boldsymbol{m}, \boldsymbol{\delta})^\rho.$$

The initial function ψ_0 satisfies all these properties. Now assume that $\psi_{\rho-1}$ also satisfies them. Then it is clear that ψ_ρ will satisfy the first two. In order to prove (c), note that for $\boldsymbol{z} \in R(\boldsymbol{\delta})$, the multivariate mean value theorem gives

$$|T(\boldsymbol{\theta} + \boldsymbol{z}) - T(\boldsymbol{\theta})| \le K|z| \le K\sqrt{q}\,\delta = \frac{\Delta}{r}. \tag{3.10}$$

Let $\boldsymbol{\theta} \in T^{-1}((\alpha+\rho\Delta/r, \beta-\rho\Delta/r))$. By (3.10), we see that $\boldsymbol{\theta}+\boldsymbol{z} \in T^{-1}((\alpha+(\rho-1)\Delta/r, \beta-(\rho-1)\Delta/r))$ and therefore

$$\psi_\rho(\boldsymbol{\theta}) = \frac{1}{(2\delta)^q} \int_{R(\boldsymbol{\delta})} \psi_{\rho-1}(\boldsymbol{\theta} + \boldsymbol{z})\, d\boldsymbol{z} = \frac{1}{(2\delta)^q} \int_{R(\boldsymbol{\delta})} d\boldsymbol{z} = 1,$$

where in the middle equality we have used the induction hypothesis. The proof of (d) is analogous. Properties (e) and (f) are immediate from Lemma 3.3.

Note that (f), (3.7), and (3.8) imply that

$$|c_{\boldsymbol{m}}(\psi_\rho)| \le |c_{\boldsymbol{m}}(\psi_0)| \le \min\left\{|c_{\mathbf{0}}(\psi_0)|, \frac{C}{\pi m}\right\}.$$

To conclude, take $D := \psi_r$, and the proposition follows from (f) and the fact that, for $m_j \neq 0$, we have

$$|\nu(m_j, \delta)| \le \min\left\{1, \frac{rK\sqrt{q}}{2\pi|m_j|\Delta}\right\}. \qquad \blacksquare$$

3.3. The Cartan subgroup

As in the previous sections, A denotes an abelian variety of dimension g defined over the number field k. From now on we will assume moreover that its Sato–Tate group $\mathrm{ST}(A)$ is connected.

Since $\mathrm{ST}(A)$ is reductive, its complexified Lie algebra $\mathfrak{g}$ is the product of a semisimple Lie algebra $\mathfrak{s}$ and an abelian Lie algebra $\mathfrak{a}$. Recall the notations from Section 3.1 relative to $\mathfrak{s}$; in particular, $\mathfrak{h}$ is a Cartan subalgebra for $\mathfrak{s}$, and h denotes the rank of $\mathfrak{h}$. Given $(\theta_1, \dots, \theta_g) \in \mathbb{R}^g$, set

$$d(\theta_1, \dots, \theta_g) := \mathrm{diag}(e^{2\pi i\theta_1}, \dots, e^{2\pi i\theta_g}, e^{-2\pi i\theta_1}, \dots, e^{-2\pi i\theta_g}).$$

Let a denote the rank of $\mathfrak{a}$ and let $q = h + a$ be the rank of $\mathfrak{g}$. As in Section 3.2, write $\boldsymbol{\theta}$ to denote $(\theta_1, \dots, \theta_q) \in \mathbb{R}^q$. We may choose $\boldsymbol{a}_{q+1}, \dots, \boldsymbol{a}_g \in \mathbb{Z}^q$ such that the image H of the map

$$\iota: \mathbb{R}^q \to \mathrm{ST}(A), \quad \iota(\boldsymbol{\theta}) = d(\theta_1, \dots, \theta_q, \boldsymbol{\theta} \cdot \boldsymbol{a}_{q+1}, \dots, \boldsymbol{\theta} \cdot \boldsymbol{a}_g), \tag{3.11}$$

has complexified Lie algebra isomorphic to $\mathfrak{h} \times \mathfrak{a}$. We then say that H is a *Cartan subgroup* of $\mathrm{ST}(A)$. For notational purposes, it will be convenient to let $\boldsymbol{a}_1, \dots, \boldsymbol{a}_q$ denote the standard basis of $\mathbb{Z}^q$. Let $a_{l,j}$ denote the j-th component of $\boldsymbol{a}_l$.

Consider the map

$$T: \mathbb{R}^q \xrightarrow{\iota} H \subseteq \mathrm{ST}(A) \xrightarrow{\text{Trace}} [-2g, 2g], \quad T(\boldsymbol{\theta}) = \sum_{j=1}^{g} 2\cos(2\pi \boldsymbol{a}_j \cdot \boldsymbol{\theta}). \tag{3.12}$$

In the next section, we will apply the construction of a Vinogradov function attached to the map T, as seen in Section 3.2. In order to control $|\nabla(T)|$ we need to control the size of $\boldsymbol{a}_{q+1}, \dots, \boldsymbol{a}_g$. The following form of the Mumford–Tate conjecture serves such a purpose.

By a *Cartan Hodge circle* we will mean the image of any homomorphism

$$\varphi: \mathbb{R} \to H$$

such that $\varphi(\theta)$ has g eigenvalues equal to $e^{2\pi i\theta}$ and g eigenvalues equal to $e^{-2\pi i\theta}$. The following statement is a refinement of the "Hodge condition" included among the "Sato–Tate axioms" stated in [12, Proposition 3.2], [13, Remark 2.3] (see also [32, Section 8.2.3.6(i)]).

Theorem 3.5. *Suppose that the Mumford–Tate conjecture holds for A. Then the group H is generated by Cartan Hodge circles.*

Proof. If $\mathrm{ST}(A)$ is abelian, then it is equal to H and the claim is that $\mathrm{ST}(A)$ itself is generated by Hodge circles. This follows from [12, Proposition 3.2] as augmented in [13, Remark 2.3].

We next reduce the general case to the previous paragraph, by arguing as in the proof of Deligne's theorem on absolute Hodge cycles. Recall that the Mumford–Tate group of A is the smallest $\mathbb{Q}$-algebraic subgroup of $\mathrm{GL}(H_1(A^{\mathrm{top}}_{\mathbb{C}}, \mathbb{Q}))$ whose base extension to $\mathbb{R}$ contains the action of the Deligne torus $\mathrm{Res}_{\mathbb{C}/\mathbb{R}}(\mathbb{G}_{\mathrm{m}})$ coming from the Hodge structure. Under our hypotheses on A, we may recover $\mathrm{ST}(A)$ by taking the Mumford–Tate group, taking the kernel of the determinant to get the Hodge group, then taking a maximal compact subgroup.

By the proof of [10, Proposition 6.1], there exists an algebraic family of abelian varieties containing A as a fiber such that, on the one hand, the generic Mumford–Tate group is equal to that of A, and on the other hand there is a fiber B whose Mumford–Tate group is a maximal torus in A. Using the previous paragraph, we see that the desired assertion for A follows from the corresponding assertion for B, which we deduce from the first paragraph. ∎

Lemma 3.6. *Suppose that the Mumford–Tate conjecture holds for A. Then $|a_{l,j}| = O(1)$.*

Proof. Write $\mathcal{A}$ for the matrix $(a_{l,j})_{l,j}$. Giving a Cartan Hodge circle amounts to giving a vector $\boldsymbol{v} \in \{\pm 1\}^q$ such that

$$\mathcal{A}\boldsymbol{v}^t = \boldsymbol{u}^t, \tag{3.13}$$

where $\boldsymbol{u} \in \{\pm 1\}^g$ (and $\boldsymbol{v}^t$, $\boldsymbol{u}^t$ denote the transposes of $\boldsymbol{v}$, $\boldsymbol{u}$). By Theorem 3.5, there exist q linearly independent vectors $\boldsymbol{v}$ satisfying an equation of the type (3.13). Let $\boldsymbol{v}_j$, for $j = 1, \dots, q$, denote these vectors, and let $\boldsymbol{u}_j \in \{\pm 1\}^g$ denote the corresponding constant terms in the equation that they satisfy. Let $v_{j,l}$ (resp. $u_{j,l}$) denote the l-th component of $\boldsymbol{v}_j$ (resp. $\boldsymbol{u}_j$). Write $\mathcal{V}$ (resp. $\mathcal{U}$) for the matrix $(v_{l,j})_{j,l}$ (resp. $(u_{l,j})_{j,l}$). Since $\mathcal{V}$ is invertible, we have

$$\mathcal{A} = \mathcal{U}\mathcal{V}^{-1}.$$

The lemma now follows immediately from the fact that all the entries of $\mathcal{V}$ and of $\mathcal{U}$ are ± 1. ∎

Lemma 3.7. *Suppose that the Mumford–Tate conjecture holds for A. Then the map $T\colon \mathbb{R}^q \to [-2g, 2g]$ from (3.12) satisfies conditions* (1)–(3) *of Proposition 3.4. Moreover, the constants K and C appearing in* (2) *and* (3) *respectively are both $O(1)$.*

Proof. An easy computation shows that for every $\boldsymbol{\theta} \in \mathbb{R}^q$ we have

$$\nabla(T)(\boldsymbol{\theta}) = -4\pi\Big(\sum_{j=1}^{g} \sin(2\pi \boldsymbol{a}_j \cdot \boldsymbol{\theta}) a_{j,1}, \dots, \sum_{j=1}^{g} \sin(2\pi \boldsymbol{a}_j \cdot \boldsymbol{\theta}) a_{j,g}\Big),$$

from which the desired bound $|\nabla(T)(\boldsymbol{\theta})| = O(1)$ is a consequence of Lemma 3.6.

As for (3), let $1 \le j \le q$, and fix $\pi_j(\boldsymbol{\theta}) \in \mathbb{R}^{q-1}$ and $\gamma \in \mathbb{R}$. Suppose that $\vartheta \in [0, 1]$ satisfies

$$T(\theta_1, \dots, \theta_{j-1}, \vartheta, \theta_{j+1}, \dots, \theta_q) = \gamma.$$

This means that there exist real numbers r_l depending exclusively on $\pi_j(\boldsymbol{\theta})$ such that

$$\sum_{l=1}^{g} 2\cos(2\pi a_{lj}\vartheta + r_l) = \gamma.$$

Let $N = \max_l a_{lj}$. By the identity

$$\cos(2\pi a_{lj}\vartheta + r_l) = \cos(2\pi a_{lj}\vartheta)\cos(r_l) - \sin(2\pi a_{lj}\vartheta)\sin(r_l)$$

and de Moivre's formula, we deduce that there exist polynomials $p, q \in \mathbb{R}[x]$ of degree $\leq N$ such that

$$p(\cos(2\pi\vartheta)) - q(\sin(2\pi\vartheta)) = \gamma.$$

If we write $q(x) = \sum_n b_n x^n$, the above equality implies that $\cos(2\pi\vartheta)$ is a root of

$$r(x) = \Big(\gamma - p(x) + \sum_n b_{2n}(1-x^2)^n\Big)^2 - (1-x^2)\Big(\sum_n b_{2n+1}(1-x^2)^n\Big)^2.$$

Since $r(x)$ has degree $\leq 2N$, we find that $\cos(2\pi\vartheta)$ is limited to $2N$ values. This implies that ϑ is limited to $4N$ values, and we conclude by applying Lemma 3.6, which shows that $N = O(1)$. ∎

3.4. Main theorem

In this section we prove an effective version of the Sato–Tate conjecture building on the results obtained in all of the previous sections.

Let μ be the pushforward of the Haar measure of $\mathrm{ST}(A)$ on $[-2g, 2g]$ via the trace map. We refer to [32, Sections 8.1.3, 8.4.3] for properties and the structure of this measure. It admits a decomposition $\mu = \mu^{\mathrm{disc}} + \mu^{\mathrm{cont}}$, where μ^{disc} is a finite sum of Dirac measures and μ^{cont} is a measure having a continuous, integrable, and even $\mathcal{C}^\infty$ density function with respect to the Lebesgue measure outside a finite number of points. Since we will assume that $\mathrm{ST}(A)$ is connected, we will in fact find that μ^{disc} is trivial (see [32, Section 8.4.3.3]).

Attached to the Lie algebra $\mathfrak{g}$ of $\mathrm{ST}(A)$, let $\varepsilon := \varepsilon_{\mathfrak{g}}$ be as defined in (1.2) and $\nu := \nu_{\mathfrak{g}}\colon \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ be as defined in (1.3). For an interval $I \subseteq [-2g, 2g]$, recall that we denote by δ_I the characteristic function of I.

Theorem 3.8. *Let k be a number field and g a positive integer. Let A be an abelian variety defined over k of dimension g, absolute conductor N, and such that $\mathrm{ST}(A)$ is connected. Suppose that the Mumford–Tate conjecture holds for A and that Conjecture 2.4 holds for every irreducible character χ of $\mathrm{ST}(A)$. For each prime $\mathfrak{p}$ not dividing N, let $\overline{a}_{\mathfrak{p}}$ denote the normalized Frobenius trace of A at $\mathfrak{p}$. Then for all nonempty subintervals I of $[-2g, 2g]$, we have*

$$\sum_{\mathrm{Nm}(\mathfrak{p}) \leq x} \delta_I(\overline{a}_{\mathfrak{p}}) = \mu(I)\,\mathrm{Li}(x) + O\left(\frac{x^{1-\varepsilon_{\mathfrak{g}}}\log(Nx)^{2\varepsilon_{\mathfrak{g}}}}{\log(x)^{1-4\varepsilon_{\mathfrak{g}}}}\right) \quad \text{for } x \geq x_0,$$

where $x_0 = O(\nu_{\mathfrak{g}}(|I|)\log(2N)^2\log(\log(4N))^4)$.

Let us resume the notations of Section 3.1 relative to the semisimple algebra $\mathfrak{s}$. Thus, $\mathfrak{h}$ is a Cartan subalgebra for $\mathfrak{s}$ of rank h, $\mathcal{R} \subseteq \mathfrak{h}_0^*$ is the lattice of integral weights, $\mathcal{W}$ is the Weyl group of $\mathfrak{s}$, $\mathcal{C}$ denotes the integral weights in a Weyl chamber, and $\omega_1, \ldots, \omega_h$ are the fundamental weights. Let a denote the rank of $\mathfrak{a}$, so that $q = h + a$. Before starting the proof we introduce some additional notations.

Recall the map $\iota\colon \mathbb{R}^q \to \mathrm{ST}(A)$ from (3.11). Without loss of generality, we may assume that the decomposition $\mathbb{R}^q = \mathbb{R}^h \times \mathbb{R}^a$ is such that the complexification of the Lie algebra

of $\iota(\mathbb{R}^h)$ (resp. $\iota(\mathbb{R}^a)$) is $\mathfrak{h}$ (resp. $\mathfrak{a}$). Let us write $\boldsymbol{\theta}_h$ (resp. $\boldsymbol{\theta}_a$) for the projection of $\boldsymbol{\theta}$ onto $\mathbb{R}^h$ (resp. $\mathbb{R}^a$).

From now on we fix a $\mathbb{Z}$-basis $\psi_1, \dots, \psi_q$ of the character group $\hat{H}$ of H: for $1 \le j \le h$, the character ψ_j is induced by the fundamental weight ω_j of $\mathfrak{s}$; for $h+1 \le j \le q$, we set

$$\psi_j(\iota(\boldsymbol{\theta})) = e^{2\pi i \theta_j}.$$

The action of $\mathcal{W}$ on $\mathfrak{h}_0^*$ induces an action of $\mathcal{W}$ on the character group $\hat{H}$ of H. We may define an action of $\mathcal{W}$ on $[0,1]^q$ by transport of structure: given $w \in \mathcal{W}$, let $w(\boldsymbol{\theta})$ be defined by

$$\psi_j(\iota(w(\boldsymbol{\theta}))) = w(\psi_j)(\iota(\boldsymbol{\theta})) \quad \text{for all } j = 1, \dots, q. \tag{3.14}$$

Of course the action of $\mathcal{W}$ restricts to the first factor of the decomposition $[0,1]^q = [0,1]^h \times [0,1]^a$. Note that the map ι from (3.11) induces an isomorphism

$$\iota\colon [0,1]^q/\mathcal{W} \xrightarrow{\sim} \mathrm{Conj}(\mathrm{ST}(A)).$$

Recall the elements $y_{\mathfrak{p}} \in \mathrm{Conj}(\mathrm{ST}(A))$ introduced in Section 2. Let $\boldsymbol{\theta}_{\mathfrak{p}} \in [0,1]^q/\mathcal{W}$ be the preimage of $y_{\mathfrak{p}}$ by the above isomorphism.

Consider the map $T\colon \mathbb{R}^q \to [-2g, 2g]$ defined in (3.12). Note that $T(\boldsymbol{\theta}_{\mathfrak{p}})$ is well-defined since T factors through $[0,1]^q/\mathcal{W}$, and it is equal to the normalized Frobenius trace $\overline{a}_{\mathfrak{p}}$. Let K and C denote the constants of Lemma 3.7 relative to the map T.

Let the interior of I be of the form (α, β) for $-2g \le \alpha < \beta \le 2g$. Let $\Delta > 0$ be any real number satisfying the constraint (3.6) relative to α, and β (arbitrary for the moment and to be specified in the course of the proof of Theorem 3.8).

Let $D := D_{\Delta,I}\colon \mathbb{R}^q \to \mathbb{R}$ be the Vinogradov function produced by Proposition 3.4, when applied to α, β, Δ, and T, and relative to the choice of a positive integer $r \ge 1$ (arbitrary for the moment and to be specified in the course of the proof of Theorem 3.8). Define

$$F := F_{\Delta,I}\colon \mathbb{R}^q \to \mathbb{R}, \quad F(\boldsymbol{\theta}) := \frac{1}{\#\mathcal{W}} \sum_{w \in \mathcal{W}} D(w(\boldsymbol{\theta})). \tag{3.15}$$

Notice that $F(\boldsymbol{\theta}_{\mathfrak{p}})$ is well-defined since F has been defined as an average over $\mathcal{W}$. In consonance with Remark 2.6, we make the convention that sums involving the elements $\boldsymbol{\theta}_{\mathfrak{p}}$ run over primes $\mathfrak{p}$ not dividing N.

Lemma 3.9. *If part* (i) *of Conjecture 2.4 holds for every irreducible character χ of* $\mathrm{ST}(A)$*, then*

$$\sum_{\mathrm{Nm}(\mathfrak{p}) \le x} \delta_I(\overline{a}_{\mathfrak{p}}) = \sum_{\mathrm{Nm}(\mathfrak{p}) \le x} F_{\Delta,I}(\boldsymbol{\theta}_{\mathfrak{p}}) + O(\Delta \,\mathrm{Li}(x))$$

for every Δ satisfying (3.6) *and every $x \ge 2$.*

Proof. Let Y_α, Y_β denote the preimages of α, β by the map T in $[0,1]^q$. Let $\mathcal{S} = \{\boldsymbol{s} \in [0,1]^q;\ \nabla(T)(\boldsymbol{s}) = 0\}$ denote the set of critical points of T. Let $\mathcal{R}$ be the set

$$\{\boldsymbol{\theta} \in [0,1]^q;\ \pi_j(\boldsymbol{\theta}) = \pi_j(\boldsymbol{s}) \text{ for some } 1 \le j \le q,\ \boldsymbol{s} \in \mathcal{S}\}.$$

Since T satisfies property (3) of Proposition 3.4, by Lemma 3.7 the intersections of Y_α, Y_β with $\mathcal{R}$ are finite (and, in fact, even of cardinality $O(1)$). Let W_α, W_β denote the intersections of Y_α, Y_β with the complement of $\mathcal{R}$.

We claim that W_α, W_β have volume $O(1)$ as $(q-1)$-dimensional Riemannian submanifolds of $[0,1]^q$. Before showing the claim, we note that it implies the lemma. Indeed, as functions over $[0,1]^q$, the characteristic function of $T^{-1}(I)$ and $F_{\Delta,I}$ only differ (by construction of the latter) over the $\mathcal{W}$-translates of tubular neighborhoods $B(Y_\alpha, r_\Delta)$ and $B(Y_\beta, r_\Delta)$ of Y_α and Y_β of radii $r_\Delta = O(\Delta)$. If W_α, W_β have volume $O(1)$, then $B(Y_\alpha, r_\Delta)$, $B(Y_\beta, r_\Delta)$ have volume $O(\Delta)$. Weyl's integration formula [2, Chapter IX, Section 6, Corollary 2, p. 338] together with the fact that the absolute value of Weyl's density function is $O(1)$ (see [2, Chapter IX, Section 6, p. 335]) imply that the Haar measure of the $\mathcal{W}$-translates of $B(Y_\alpha, r_\Delta)$ and $B(Y_\beta, r_\Delta)$ is $O(\Delta)$. Then the lemma follows from the equidistribution of $\theta_{\mathfrak{p}}$ implied by part (i) of Conjecture 2.4 and the prime number theorem.

We now show that W_α has volume $O(1)$ (the same argument applies to W_β). For $1 \le j \le q$, define

$$\mathcal{V}_j = \left\{ \boldsymbol{\theta} \in [0,1]^q ; \frac{\partial T}{\partial \theta_j}(\boldsymbol{\theta}) \ge \frac{\partial T}{\partial \theta_l}(\boldsymbol{\theta}) \text{ for every } 1 \le l \le q \right\}.$$

It suffices to show that $W_\alpha \cap \mathcal{V}_j$ has volume $O(1)$ for every j. By symmetry, we may assume that $j = q$, which will be convenient for notational purposes. Let $\mathcal{Z}_{\alpha,q}$ denote the interior of the image of $W_\alpha \cap \mathcal{V}_q$ by the projection map $\pi_q \colon [0,1]^q \to [0,1]^{q-1}$. For $\boldsymbol{\vartheta} \in \mathcal{Z}_{\alpha,q}$, choose $\tilde{\boldsymbol{\vartheta}} \in W_\alpha \cap \mathcal{V}_q$ such that $\pi_q(\tilde{\boldsymbol{\vartheta}}) = \boldsymbol{\vartheta}$. By the implicit function theorem there exist a neighborhood $\mathcal{U}_{\boldsymbol{\vartheta}} \subseteq \mathcal{Z}_{\alpha,q}$ of $\boldsymbol{\vartheta}$ and a differentiable function $g_{\boldsymbol{\vartheta}} \colon \mathcal{U}_{\boldsymbol{\vartheta}} \to \mathbb{R}$ such that

$$\tilde{\boldsymbol{\vartheta}} = (\boldsymbol{\vartheta}, g(\boldsymbol{\vartheta})) \quad \text{and} \quad (\boldsymbol{t}, g(\boldsymbol{t})) \in W_\alpha \cap \mathcal{V}_q \quad \text{for every } \boldsymbol{t} \in \mathcal{U}_\alpha.$$

The lifts $\tilde{\boldsymbol{\vartheta}}$ can be compatibly chosen so that the functions $g_{\boldsymbol{\vartheta}}$ glue together into a differentiable function $g \colon \mathcal{Z}_{\alpha,q} \to W_\alpha \cap \mathcal{V}_q$. Then Lemma 3.7 provides the following bound for the volume of $W_\alpha \cap \mathcal{V}_q$:

$$O\left(\int_{\mathcal{Z}_{\alpha,q}} \prod_{j=1}^{q-1} \left(1 + \left(\frac{\partial g}{\partial \theta_j}(\boldsymbol{\vartheta}) \right)^2 \right)^{1/2} d\boldsymbol{\vartheta} \right)$$
$$= O\left(\int_{\mathcal{Z}_{\alpha,q}} \prod_{j=1}^{q-1} \left(1 + \left(\frac{\partial T}{\partial \theta_q} \right)^{-2} \left(\frac{\partial T}{\partial \theta_j} \right)^2 (\boldsymbol{\vartheta}, g(\boldsymbol{\vartheta})) \right)^{1/2} d\boldsymbol{\vartheta} \right) = O(1),$$

which completes the proof. ∎

Proof of Theorem 3.8. The choice of a basis of fundamental weights $\omega_1, \dots, \omega_h$ gives an isomorphism

$$\mathbb{Z}^h \simeq \mathcal{R} = \mathcal{W} \cdot \mathcal{C},$$

by means of which, from now on, we will view integral weights of $\mathfrak{s}$ as elements in $\mathbb{Z}^h$. Similarly, the choice of the basis elements of (3.14) provides an isomorphism between the lattice of integral weights of $\mathfrak{a}$ and $\mathbb{Z}^a$. For a weight $\boldsymbol{m} \in \mathbb{Z}^q$, let $\boldsymbol{m}_h$ and $\boldsymbol{m}_a$ denote the projections to $\mathbb{Z}^h$ and $\mathbb{Z}^a$. For $\boldsymbol{m}_h \in \mathbb{Z}^h$, define

$$f_{\boldsymbol{m}_h}(\boldsymbol{\theta}_h) = \frac{1}{t_{\boldsymbol{m}_h}} \sum_{w \in \mathcal{W}} e^{2\pi i \boldsymbol{m}_h \cdot w(\boldsymbol{\theta}_h)},$$

where $t_{\boldsymbol{m}_h}$ denotes the size of the stabilizer of $\boldsymbol{m}_h$ under the action of $\mathcal{W}$. If $\Gamma_{\boldsymbol{n}_h}$ denotes the representation of highest weight $\boldsymbol{n}_h$, then

$$\operatorname{Trace}(\Gamma_{\boldsymbol{n}_h}(\boldsymbol{\theta}_h)) = \sum_{\boldsymbol{m}_h \preceq \boldsymbol{n}_h} m_{\boldsymbol{n}_h}^{\boldsymbol{m}_h} f_{\boldsymbol{m}_h}(\boldsymbol{\theta}_h),$$

where the sum runs over weights $\boldsymbol{m}_h \in \mathcal{C}$. Equivalently, we have

$$f_{\boldsymbol{m}_h}(\boldsymbol{\theta}_h) = \sum_{\boldsymbol{n}_h \preceq \boldsymbol{m}_h} d_{\boldsymbol{m}_h}^{\boldsymbol{n}_h} \operatorname{Trace}(\Gamma_{\boldsymbol{n}_h}(\boldsymbol{\theta}_h)). \tag{3.16}$$

We remark that Proposition 3.1 ensures that, for each $\boldsymbol{m}_h$, the number of nonzero coefficients $d_{\boldsymbol{m}_h}^{\boldsymbol{n}_h}$ in the above equation, as well as the size of each of them, is $O(1)$. By taking the Fourier expansion of D, we obtain

$$\begin{aligned} F(\boldsymbol{\theta}) &= \frac{1}{\#\mathcal{W}} \sum_{\boldsymbol{m} \in \mathbb{Z}^q} c_{\boldsymbol{m}} t_{\boldsymbol{m}_h} f_{\boldsymbol{m}_h}(\boldsymbol{\theta}_h) e^{2\pi i \boldsymbol{m}_a \cdot \boldsymbol{\theta}_a} \\ &= \frac{1}{\#\mathcal{W}} \sum_{\boldsymbol{m} \in \mathcal{C} \times \mathbb{Z}^a} \Big(\sum_{w \in \mathcal{W}} c_{w(\boldsymbol{m})} \Big) t_{\boldsymbol{m}_h} f_{\boldsymbol{m}_h}(\boldsymbol{\theta}_h) e^{2\pi i \boldsymbol{m}_a \cdot \boldsymbol{\theta}_a}. \end{aligned}$$

Let $M \geq 1$ be a positive integer (arbitrary for the moment and to be determined later). Let $\mathcal{C}^{\leq M}$ denote the subset of $\mathcal{C} \times \mathbb{Z}^a$ made of weights $\boldsymbol{m}$ whose components have absolute value $\leq M$. Note that if $\boldsymbol{m} \in \mathcal{C}^{\leq M}$, then in particular we have $\|\boldsymbol{m}_h\|_{\mathrm{fund}} \leq M$. Let $\mathcal{C}^{>M}$ denote the complement of $\mathcal{C}^{\leq M}$ in $\mathcal{C} \times \mathbb{Z}^a$.

On the one hand, by invoking the bounds from part (iii) of Proposition 3.4, we have

$$\begin{aligned} F_{>M}(\boldsymbol{\theta}) &:= \frac{1}{\#\mathcal{W}} \sum_{\boldsymbol{m} \in \mathcal{C}^{>M}} \Big(\sum_{w \in \mathcal{W}} c_{w(\boldsymbol{m})} \Big) t_{\boldsymbol{m}_h} f_{\boldsymbol{m}_h}(\boldsymbol{\theta}_h) e^{2\pi i \boldsymbol{m}_a \cdot \boldsymbol{\theta}_a} \\ &= O\Big(\sum_{m > M} m^{q-1} \frac{1}{m} \Big(\frac{rK\sqrt{q}}{2\pi m \Delta} \Big)^r \Big) = O\Big(\frac{1}{M^{r-q+1} \Delta^r} \Big(\frac{rK\sqrt{q}}{2\pi} \Big)^r \Big). \end{aligned} \tag{3.17}$$

On the other hand, consider the class function

$$\begin{aligned} F_{\leq M}(\boldsymbol{\theta}) &:= \frac{1}{\#\mathcal{W}} \sum_{\boldsymbol{m} \in \mathcal{C}^{\leq M}} \Big(\sum_{w \in \mathcal{W}} c_{w(\boldsymbol{m})} \Big) t_{\boldsymbol{m}_h} f_{\boldsymbol{m}_h}(\boldsymbol{\theta}_h) e^{2\pi i \boldsymbol{m}_a \cdot \boldsymbol{\theta}_a} \\ &= \delta(F_{\leq M}(\boldsymbol{\theta})) + \frac{1}{\#\mathcal{W}} \sum_{\boldsymbol{m} \in \mathcal{C}^{\leq M}} \Big(\sum_{w \in \mathcal{W}} c_{w(\boldsymbol{m})} \Big) t_{\boldsymbol{m}_h} \sum_{\boldsymbol{0} \neq \boldsymbol{n}_h \preceq \boldsymbol{m}_h} d_{\boldsymbol{m}_h}^{\boldsymbol{n}_h} \operatorname{Trace}(\Gamma_{\boldsymbol{n}_h}(\boldsymbol{\theta}_h)) e^{2\pi i \boldsymbol{m}_a \cdot \boldsymbol{\theta}_a}. \end{aligned} \tag{3.18}$$

In the above expression $\delta(F_{\leq M}(\boldsymbol{\theta}))$ stands for the multiplicity of the identity representation in $F_{\leq M}(\boldsymbol{\theta})$. Note that $F_{\leq M}$ is a finite linear combination of irreducible characters of $\mathrm{ST}(A)$, and by Proposition 3.1 we may assume that M is large enough that $\delta(F_{\leq M}(\boldsymbol{\theta})) = \delta(F(\boldsymbol{\theta}))$, which we will do from now on.

The next step is to bound the virtual dimension of the nontrivial part of $F_{\leq M}(\boldsymbol{\theta})$ in order to be able to apply Proposition 2.5. More precisely, if $p_{\boldsymbol{n}_h}$ denotes the coefficient of $\mathrm{Trace}(\Gamma_{\boldsymbol{n}_h}(\boldsymbol{\theta}_h))$ in (3.18), then

$$\begin{aligned}\sum_{\boldsymbol{m}\in\mathcal{C}^{\leq M}} \sum_{\mathbf{0}\neq\boldsymbol{n}_h\preceq\boldsymbol{m}_h} |p_{\boldsymbol{n}_h}| \dim(\Gamma_{\boldsymbol{n}_h}) &= O\Big(\sum_{\mathbf{0}\neq\boldsymbol{m}_h\in\mathcal{C}^{\leq M}} c_{\boldsymbol{m}_h} \dim(\Gamma_{\boldsymbol{m}_h})\Big) \\ &= O\Big(\sum_{0<m\leq M} m^{q-1} \frac{1}{m} \Big(\frac{rK\sqrt{q}}{2\pi m\Delta}\Big)^{\rho} m^{\varphi}\Big).\end{aligned}$$

In the above computation we have used: Proposition 3.1 to bound the size and number of nonzero entries in the inverse of the matrix of weight multiplicities; Proposition 3.2 to control the dimension of the representations of weight lower than a given one and to bound the dimension of the representation $\Gamma_{\boldsymbol{m}_h}$ in terms of $\|\boldsymbol{m}_h\|_{\mathrm{fund}}$; and Proposition 3.4 (iii) to bound the Fourier coefficients for an unspecified (for the moment) $1 \leq \rho \leq r$. We will now distinguish two cases, depending on whether φ is zero or not.

Suppose first that φ is nonzero. Take $r = \rho = q + \varphi - 1$, which we note satisfies $r \geq 1$. Then

$$\sum_{\boldsymbol{m}\in\mathcal{C}^{\leq M}} \sum_{\mathbf{0}\neq\boldsymbol{n}_h\preceq\boldsymbol{m}_h} |p_{\boldsymbol{n}_h}| \dim(\Gamma_{\boldsymbol{n}_h}) = O\Big(\sum_{0<m\leq M} \frac{1}{m\Delta^{q+\varphi-1}}\Big) = O\Big(\frac{\log(M)}{\Delta^{q+\varphi-1}}\Big). \tag{3.19}$$

Let $L > 0$ be the implied constant in the bound of Proposition 3.2 (iii) for the motivic weight, so that for $\boldsymbol{m}_h \in \mathcal{C}^{\leq M}$, we have $w_{\Gamma_{\boldsymbol{m}_h}} \leq LM$. Using the decomposition

$$F(\boldsymbol{\theta}) = F_{\leq M}(\boldsymbol{\theta}) + F_{>M}(\boldsymbol{\theta}),$$

the tail (3.17) and virtual dimension (3.19) bounds, and applying Proposition 2.5, we obtain

$$\begin{aligned}\sum_{\mathrm{Nm}(\mathfrak{p})\leq x} F(\boldsymbol{\theta}_{\mathfrak{p}}) &= \delta(F(\boldsymbol{\theta}))\,\mathrm{Li}(x) + O\Big(\frac{\log(M)}{\Delta^{q+\varphi-1}}\sqrt{x}\log\big(N(x+LM)\big)\Big) \\ &\quad + O\Big(\frac{\mathrm{Li}(x)}{M^{\varphi}\Delta^{q+\varphi-1}}\Big).\end{aligned} \tag{3.20}$$

It follows from the proof of Lemma 3.9 that

$$\delta(F(\boldsymbol{\theta})) = \mu(I) + O(\Delta). \tag{3.21}$$

Therefore, to conclude the proof, it will suffice to balance the error terms in (3.20) with $O(\Delta\,\mathrm{Li}(x))$. If φ is nonzero, we may take

$$\Delta := x^{-\varepsilon}\log(x)^{4\varepsilon}\log(Nx)^{2\varepsilon}, \quad M = \lceil \Delta^{-(q+\varphi)/\varphi} \rceil, \tag{3.22}$$

where $\varepsilon = \varepsilon_{\mathfrak{g}}$ is as defined in (1.2). In view of Lemma 3.9, this concludes the proof, provided that we verify that this choice of Δ satisfies the constraint (3.6). This amounts to $2\Delta \le |I|$, or equivalently to

$$x \ge \frac{2^{\varepsilon^{-1}}}{|I|^{\varepsilon^{-1}}} \log(x)^4 \log(Nx)^2.$$

By the elementary Lemma 3.10 below, this is easily seen to be the case as long as $x \ge x_0$, where

$$x_0 = O\big(v_{\mathfrak{g}}(|I|) \log(2N)^2 \log(\log(4N))^4\big). \tag{3.23}$$

Suppose now that $\varphi = 0$. We take $r = q$ and use the tail bound (3.17) as in the previous case. To bound the Fourier coefficients of $F_{\le M}(\boldsymbol{\theta})$, we use the bound $|c_m| = O(1/m)$ if $q = 1$ and the bound corresponding to $\rho = q - 1$ otherwise (as in Proposition 3.4 (iii)). We obtain

$$\sum_{\mathrm{Nm}(\mathfrak{p}) \le x} F(\boldsymbol{\theta}_{\mathfrak{p}}) = \delta(F(\boldsymbol{\theta})) \operatorname{Li}(x) + O\left(\frac{\log(M)}{\Delta^{q-1}} \sqrt{x} \log\big(N(x + LM)\big)\right) + O\left(\frac{\operatorname{Li}(x)}{M\Delta^q}\right). \tag{3.24}$$

To balance the error terms in the above equation with $O(\Delta \operatorname{Li}(x))$, we may take

$$\Delta := x^{-1/(2q)} \log(x)^{2/q} \log(Nx)^{1/q}, \quad M = \lceil \Delta^{-q-1} \rceil.$$

Since $\varepsilon_{\mathfrak{g}} = 1/(2q)$ in this case, this yields precisely the error term of the statement of the theorem. Again, Δ satisfies the constraint (3.6) as soon as $x \ge x_0$, where x_0 is as in (3.23). ∎

We leave the proof of the following to the reader.

Lemma 3.10. *For integers $r, N \ge 1$, with r even, and a real number $A > 0$, we have*

$$A \log(x)^r \log(Nx)^2 < x$$

provided that $x > C \log(2N)^2 \log(\log(4N))^r \max\{1, A \log(A)^{r+2}\}$ for some $C > 0$ depending exclusively on r.

Remark 3.11. To simplify the statement of Theorem 3.8, we have assumed Conjecture 2.4 for every irreducible character χ of $\mathrm{ST}(A)$. It is however clear from the proof that this hypothesis can be relaxed: it suffices to assume Conjecture 2.4 for those representations $\Gamma_{\boldsymbol{m}}$ with $\boldsymbol{m} \in \mathcal{C}^{\le M}$, where M is as in (3.22).

Remark 3.12. The choice of the exponent of x in the error term in Theorem 3.8 is dictated by the balancing of $O(\Delta \operatorname{Li}(x))$ with the first of the two error terms in (3.20). The balancing with the second error term only affects the logarithmic factors.

4. Applications

In this section we discuss three applications of Theorem 3.8. In Section 4.1 we consider an interval variant of Linnik's problem for abelian varieties. Given an abelian variety A

defined over k of dimension g and a subinterval I of $[-2g, 2g]$, this asks for an upper bound on the least norm of a prime $\mathfrak{p}$ not dividing N such that the normalized Frobenius trace $\overline{a}_{\mathfrak{p}}(A)$ lies in I. In Section 4.2 we consider a sign variant of Linnik's problem for a pair of abelian varieties A and A' defined over the number field k and such that $\mathrm{ST}(A \times A') \simeq \mathrm{ST}(A) \times \mathrm{ST}(A')$. This asks for an upper bound on the least norm of a prime $\mathfrak{p}$ such that $a_{\mathfrak{p}}(A)$ and $a_{\mathfrak{p}}(A')$ are nonnegative and have opposite sign. Finally, in Section 4.3, when A is an elliptic curve with CM, we conditionally determine (up to constant multiplication) the asymptotic number of primes for which $a_{\mathfrak{p}}(A) = \lfloor 2\sqrt{\mathrm{Nm}(\mathfrak{p})} \rfloor$.

While Section 4.1 is a direct consequence of Theorem 3.8, both Section 4.2 and Section 4.3 require slight variations of it. We will explain how to modify the proof of Theorem 3.8 to obtain these versions.

4.1. Interval variant of Linnik's problem for abelian varieties

Theorem 3.8 has the following immediate corollary.

Corollary 4.1. *Assume the hypotheses and notations of Theorem 3.8. For every nonempty subinterval I of $[-2g, 2g]$, there exists a prime $\mathfrak{p}$ not dividing N with*

$$\mathrm{Nm}(\mathfrak{p}) = O\big(v_{\mathfrak{g}}(\min\{|I|, \mu(I)\}) \log(2N)^2 \log(\log(4N))^4\big)$$

such that $\overline{a}_{\mathfrak{p}} \in I$.

Proof. There exist constants $K_1, K_2 > 0$ such that, for

$$x \geq K_2 v_{\mathfrak{g}}(|I|) \log(2N)^2 \log(\log(4N))^4,$$

the number of primes $\mathfrak{p}$ such that $\mathrm{Nm}(\mathfrak{p}) \leq x$ and $\overline{a}_{\mathfrak{p}} \in I$ is at least

$$\mu(I) \operatorname{Li}(x) \left(1 - \frac{K_1}{\mu(I)} \Delta\right),$$

where Δ is as in (3.22). This count will be positive if $K_1 \Delta < \mu(I)$, or equivalently if

$$x > \frac{K_1^{\varepsilon^{-1}}}{\mu(I)^{\varepsilon^{-1}}} \log(x)^4 \log(Nx)^2.$$

One easily verifies that this condition is satisfied for $x \geq x_0$, for some

$$x_0 = O\big(v_{\mathfrak{g}}(\mu(I)) \log(2N)^2 \log(\log(4N))^4\big),$$

and the corollary follows. ∎

4.2. Frobenius sign separation for pairs of abelian varieties

In this section we will provide an answer to the Frobenius sign separation problem for pairs of abelian varieties using a variation of Theorem 3.8. Resume the notations of Section 3.4; additionally, let A' be an abelian variety defined over k and let g', N', μ', etc.

denote the corresponding notions. We will make the hypothesis that the natural inclusion of $\mathrm{ST}(A \times A')$ in the product $\mathrm{ST}(A) \times \mathrm{ST}(A')$ is an isomorphism.

Hypothesis 4.2. *We have* $\mathrm{ST}(A \times A') \simeq \mathrm{ST}(A) \times \mathrm{ST}(A')$.

Theorem 4.3 shows that under the conjectures of Section 2, this hypothesis ensures the existence of a prime $\mathfrak{p}$ not dividing NN' such that

$$a_{\mathfrak{p}}(A) \cdot a_{\mathfrak{p}}(A') < 0 \tag{4.1}$$

and, in fact, determines the asymptotic density of such primes. Corollary 4.4, which gives an upper bound on the least norm of such a prime, is then an immediate consequence. Note that requiring A and A' not to be isogenous does not guarantee the existence of a prime satisfying (4.1), as is shown by the trivial example in which A' is taken to be a proper power of A.

Write the complexified Lie algebra of $\mathrm{ST}(A)$ (resp. $\mathrm{ST}(A')$) as $\mathfrak{g} = \mathfrak{s} \times \mathfrak{a}$ (resp. $\mathfrak{g}' = \mathfrak{s}' \times \mathfrak{a}'$), where $\mathfrak{s}, \mathfrak{s}'$ are semisimple and $\mathfrak{a}, \mathfrak{a}'$ are abelian. Throughout this section, write

$$\varepsilon_{\mathfrak{g},\mathfrak{g}'} := \frac{1}{2(q + q' + \varphi + \varphi' - 1)}, \tag{4.2}$$

where φ (resp. φ') is the size of the set of positive roots of $\mathfrak{s}$ (resp. $\mathfrak{s}'$) and q (resp. q') is the rank of $\mathfrak{g}$ (resp. $\mathfrak{g}'$). Define

$$\nu_{\mathfrak{g},\mathfrak{g}'}\colon \mathbb{R}_{>0} \to \mathbb{R}_{>0}, \quad \nu_{\mathfrak{g},\mathfrak{g}'}(z) = \max\left\{1, \frac{\log(z)^8}{z^{1/\varepsilon_{\mathfrak{g},\mathfrak{g}'}}}\right\}.$$

Theorem 4.3. *Let k be a number field, and let g and g' positive integers ≥ 1. Let A (resp. A') be an abelian variety defined over k of dimension g (resp. g'), absolute conductor N (resp. N'), and such that $\mathrm{ST}(A)$ (resp. $\mathrm{ST}(A')$) is connected. Assume that Hypothesis 4.2 holds. Suppose that the Mumford–Tate conjecture holds for $A \times A'$, and that Conjecture 2.4 holds for every product $\chi \cdot \chi'$ of irreducible characters χ of $\mathrm{ST}(A)$ and χ' of $\mathrm{ST}(A')$. For each prime $\mathfrak{p}$ not dividing NN', let $\overline{a}_{\mathfrak{p}}$ (resp. $\overline{a}'_{\mathfrak{p}}$) denote the normalized Frobenius trace of A (resp. A') at $\mathfrak{p}$. Then for all nonempty subintervals I of $[-2g, 2g]$ and I' of $[-2g', 2g']$, we have*

$$\sum_{\mathrm{Nm}(\mathfrak{p}) \leq x} \delta_I(\overline{a}_{\mathfrak{p}})\delta_{I'}(\overline{a}'_{\mathfrak{p}}) = \mu(I)\mu'(I')\,\mathrm{Li}(x) + O\left(\frac{x^{1-\varepsilon_{\mathfrak{g},\mathfrak{g}'}}\log(NN'x)^{2\varepsilon_{\mathfrak{g},\mathfrak{g}'}}}{\log(x)^{1-6\varepsilon_{\mathfrak{g},\mathfrak{g}'}}}\right)$$

for $x \geq x_0$, where $x_0 = O(\nu_{\mathfrak{g},\mathfrak{g}'}(\min\{|I|, |I'|\})\log(2NN')^2 \log(4NN')^6)$.

Proof. Let (α, β) and (α', β') denote the interiors of I and I', respectively. For a common choice of $\Delta > 0$, define $F_{\Delta,I}(\boldsymbol{\theta})$ and $F'_{\Delta,I'}(\boldsymbol{\theta}')$ relative to undetermined positive integers r and r' in a manner analogous to (3.15). Let $M \geq 1$ be a positive integer (arbitrary for the moment and to be determined later). In analogy with the definition of $L > 0$ in the line following (3.19), let $L' > 0$ be the implied constant in the bound $w_{\Gamma_{\boldsymbol{m}'_h}} = O(\|\boldsymbol{m}'_h\|_{\mathrm{fund}})$. Let L'' denote $\max\{L, L'\}$.

Suppose that $\varphi + \varphi'$ is nonzero. Choose $r = q + \varphi - 1$ and $r' = q' + \varphi' - 1$. Analogues of (3.17) and (3.19) give

$$\sum_{\mathrm{Nm}(\mathfrak{p}) \le x} F(\boldsymbol{\theta}_{\mathfrak{p}}) F'(\boldsymbol{\theta}'_{\mathfrak{p}}) = \delta(F(\boldsymbol{\theta}))\delta(F'(\boldsymbol{\theta}))\,\mathrm{Li}(x) + O\left(\frac{\log(M)^2}{\Delta^{q+q'+\varphi+\varphi'-2}} \sqrt{x} \log\big(NN'(x + L''M)\big)\right) + O\left(\frac{\mathrm{Li}(x)}{M^{\varphi+\varphi'}\Delta^{q+q'+\varphi+\varphi'-2}}\right). \tag{4.3}$$

Here we have used the fact that the multiplicity of the trivial representation $\delta(\Gamma_{\boldsymbol{m}_h} \otimes \Gamma_{\boldsymbol{m}'_h})$ is zero unless both $\boldsymbol{m}_h$ and $\boldsymbol{m}'_h$ are $\mathbf{0}$, as follows from Hypothesis 4.2. We also used the fact that the conductor of $A \times A'$ is $O(NN')$. By the proof of Lemma 3.9, we have

$$\delta(F(\boldsymbol{\theta})F'(\boldsymbol{\theta}')) = \delta(F(\boldsymbol{\theta}))\delta(F'(\boldsymbol{\theta}')) = \mu(I)\mu'(I') + O(\Delta).$$

If $\varphi + \varphi'$ is nonzero, take $\varepsilon := \varepsilon_{\mathfrak{g},\mathfrak{g}'}$ as in (4.2) and

$$\Delta := x^{-\varepsilon} \log(x)^{6\varepsilon} \log(NN'x)^{2\varepsilon}, \quad M = \lceil \Delta^{-\frac{q+q'+\varphi+\varphi'-1}{\varphi+\varphi'}} \rceil, \tag{4.4}$$

which balance the error terms in (4.3) with $O(\Delta\,\mathrm{Li}(x))$.

Suppose now that $\varphi = \varphi' = 0$. Choose $r = q$ and $r' = q'$. As in (3.24), we apply Proposition 3.4 (iii) with $\rho = r$ (resp. $\rho' = r'$) to bound the Fourier coefficients of $F_{>M}$ (resp. $F'_{>M}$); for the Fourier coefficients of $F_{\ge M}$ (resp. $F'_{\ge M}$) we use the bound $c_m = O(1/m)$ (resp. $c'_m = O(1/m)$) if $q = 1$ (resp. $q' = 1$) and the bound corresponding to $\rho = q - 1$ (resp. $\rho' = q' - 1$) if $q > 1$ (resp. $q' > 1$). We obtain

$$\sum_{\mathrm{Nm}(\mathfrak{p}) \le x} F(\boldsymbol{\theta}_{\mathfrak{p}}) F'(\boldsymbol{\theta}'_{\mathfrak{p}}) = \delta(F(\boldsymbol{\theta}))\delta(F'(\boldsymbol{\theta}))\,\mathrm{Li}(x) + O\left(\frac{\log(M)^2}{\Delta^{q+q-2}} \sqrt{x} \log\big(NN'(x + L''M)\big)\right) + O\left(\frac{\mathrm{Li}(x)}{M^2\Delta^{q+q'}}\right).$$

In order to balance the error terms of the above expression with $O(\Delta\,\mathrm{Li}(x))$, we take

$$\Delta := x^{-1/(q+q'-1)} \log(x)^{3/(q+q'-1)} \log(NN'x)^{1/(q+q'-1)}, \quad M = \lceil \Delta^{-(q+q'-1)/2} \rceil.$$

This yields the error term in the statement of the theorem, since $\varepsilon_{\mathfrak{g},\mathfrak{g}'} = 1/(2(q + q' - 1))$ when $\varphi = \varphi' = 0$. It only remains to determine the set of x for which the constraint $2\Delta \le \min\{|I|, |I'|\}$, or equivalently the inequality

$$x > \frac{2^{\varepsilon^{-1}}}{(\min\{|I|, |I'|\})^{\varepsilon^{-1}}} \log(x)^6 \log(Nx)^2$$

is satisfied. As follows from Lemma 3.10, this happens if $x \ge x_0$, where x_0 is as in the statement of the theorem. ∎

Corollary 4.4. *Assume the hypotheses of Theorem 4.3. Then there exists a prime $\mathfrak{p}$ not dividing NN' with*

$$\mathrm{Nm}(\mathfrak{p}) = O\big(\log(2NN')^2 \log(\log(4NN'))^6\big)$$

such that $a_{\mathfrak{p}}(A)$ and $a_{\mathfrak{p}}(A')$ are nonzero and of opposite sign.

Proof. In Theorem 4.3, take the subintervals $I = (\delta, 2g-\delta)$ and $I' = (-2g+\delta, -\delta)$ for $\delta = 1/2$. There exist constants $K_1, K_2, K_3 > 0$ such that, for

$$x \geq K_3 \log(2NN')^2 \log(\log(4NN'))^6,$$

the number of primes $\mathfrak{p}$ such that $\mathrm{Nm}(\mathfrak{p}) \leq x$, $\overline{a}_{\mathfrak{p}} \in I$, and $\overline{a}'_{\mathfrak{p}} \in I'$ is at least

$$K_1 \,\mathrm{Li}(x)(1 - K_2\Delta),$$

where Δ as in (4.4). This count will be positive provided that $K_2\Delta < 1$, or equivalently if

$$x > K_2^{\varepsilon^{-1}} \log(x)^6 \log(NN'x)^2.$$

One easily verifies that this condition is satisfied for $x \geq x_0$, for some

$$x_0 = O(\log(2NN')^2 \log(\log(4NN'))^6). \qquad \blacksquare$$

Remark 4.5. Under the current assumption that $\mathrm{ST}(A)$ and $\mathrm{ST}(A')$ are connected, one may wonder when Hypothesis 4.2 is satisfied. According to [1, Lemma 6.10] this should happen rather often when $\mathrm{Hom}(A_{\overline{\mathbb{Q}}}, A'_{\overline{\mathbb{Q}}}) = 0$. More precisely, if both A and A' satisfy the Mumford–Tate conjecture, $\mathrm{Hom}(A_{\overline{\mathbb{Q}}}, A'_{\overline{\mathbb{Q}}}) = 0$, A has no factors of type IV, and either

(i) A' is of CM type, or

(ii) A' has no factors of type IV,

then Hypothesis 4.2 holds.

4.3. CM elliptic curve reductions with maximal number of points

In this section we prove a variation of Theorem 3.8 when the interval I *varies with* x. We determine (up to constant multiplication and under the assumption of Conjecture 2.4) the number of primes at which the Frobenius trace of an elliptic curve defined over k with potential CM achieves the integral part of the Weil bound. We will start by assuming that A has CM already defined over k, that is, $\mathrm{ST}(A) \simeq \mathrm{U}(1)$.

Throughout this section let $x \geq 2$ and $y \geq 2^{2/3}$ be real numbers. Let I_y denote the subinterval $[2 - y^{-1/2}, 2]$ of $[-2, 2]$.

Lemma 4.6. *For $\mu = dz/(\pi\sqrt{4-z^2})$, we have*

$$\mu(I_y) = \frac{1}{\pi y^{1/4}} + O\left(\frac{1}{y^{3/4}}\right) \quad \textit{for every } y \geq 2^{2/3}.$$

Proof. Recall the map from (3.12), which in this case is simply

$$T\colon \mathbb{R} \to [-2,2], \quad T(\theta) = 2\cos(2\pi\theta).$$

We first determine the preimage $[-\theta_y, \theta_y] := T^{-1}(I_y) \cap [-1/2, 1/2]$. We easily find

$$\theta_y = \frac{1}{2\pi}\arccos\left(1 - \frac{y^{-1/2}}{2}\right) = \frac{1}{2\pi y^{1/4}} + O\left(\frac{1}{y^{3/4}}\right) \quad \text{for every } y \geq 2^{2/3}.$$

Since μ is the pushforward via T of the uniform measure on $[0,1]$, we see that $\mu(I_y)$ is the length of $[-\theta_y, \theta_y]$, from which the lemma follows. ∎

Proposition 4.7. *Let A be an elliptic curve with CM defined over k of absolute conductor N. Suppose that Conjecture 2.4 holds for every character[2] of $\mathrm{ST}(A) \simeq \mathrm{U}(1)$. For each prime $\mathfrak{p}$ not dividing N, let $\overline{a}_\mathfrak{p}$ denote the normalized Frobenius trace of A at $\mathfrak{p}$. For every $x \geq 2$, we have*

$$\sum_{\mathrm{Nm}(\mathfrak{p}) \leq x} \delta_{I_y}(\overline{a}_\mathfrak{p}) = \frac{1}{\pi y^{1/4}}\mathrm{Li}(x) + O\big(\sqrt{x}\log(Nx)\log(x)\big) \quad \textit{for every } x^{2/3} \leq y \leq x.$$

Proof. Let us start by choosing $\Delta = y^{-1/2-\nu}$ for some $\nu > 0$ so that hypothesis (3.6) for Δ and I_y is satisfied. Let us choose the function $D = D_{\Delta, I_y}$ from Proposition 3.4 relative to $r = 1$. Proceeding exactly as in the case $\varphi = 0$ of the proof of Theorem 3.8 we arrive at (3.24) (the fact that the exponent of Δ in the mid error term of (3.24) is $q-1$ is precisely what makes the case $q = 1$ special: this allowed us to choose Δ beforehand and arbitrarily small). As seen in the proof of Lemma 3.9, we have $\delta(F(\theta)) = \mu(I_y) + O(\Delta)$. Then, the choice of $M = \Delta^{-2}$ gives

$$\sum_{\mathrm{Nm}(\mathfrak{p}) \leq x} F(\theta_\mathfrak{p}) = \mu(I_y)\mathrm{Li}(x) + O\big(\sqrt{x}\log(Nx)\log(x)\big).$$

By Lemmas 3.9 and 4.6 we have

$$\sum_{\mathrm{Nm}(\mathfrak{p}) \leq x} \delta_{I_y}(\overline{a}_\mathfrak{p}) = \frac{1}{2\pi y^{1/4}}\mathrm{Li}(x) + O\left(\frac{\mathrm{Li}(x)}{y^{3/4}}\right) + O(\sqrt{x}\log(Nx)\log(x)).$$

The proposition now follows from the fact that if $y \geq x^{2/3}$, then the $O(\mathrm{Li}(x)/y^{3/4})$ term is subsumed in the error term of the statement. ∎

For every $x \geq 2$, define

$$R(x) := \{\mathfrak{p} \nmid N \text{ prime of } k;\ \mathrm{Nm}(\mathfrak{p}) \leq x \text{ and } |\overline{a}_\mathfrak{p} - 2| < x^{-1/2}\},$$

and for every $x^{2/3} \leq y \leq x$, define

$$S(y,x) := \{\mathfrak{p} \nmid N \text{ prime of } k;\ y < \mathrm{Nm}(\mathfrak{p}) \leq x \text{ and } |\overline{a}_\mathfrak{p} - 2| < y^{-1/2}\}.$$

Lemma 4.6 and Proposition 4.7 have the following corollary.

[2] In other words, we assume that GRH holds for the Hecke L-function attached to every integral power of the Grossencharacter attached to A.

Corollary 4.8. *Assume the same hypotheses as in Proposition 4.7. For every $x \geq 2$, we have:*

(i) $\#R(x) = \frac{1}{\pi x^{1/4}} \operatorname{Li}(x) + O(\sqrt{x}\log(Nx)\log(x))$.

(ii) $\#S(y,x) = \frac{1}{\pi y^{1/4}}(\operatorname{Li}(x) - \operatorname{Li}(y)) + O(\sqrt{x}\log(Nx)\log(x))$ *for every* $x^{2/3} \leq y \leq x$.

Let $M_k(x)$ denote the set of primes $\mathfrak{p}$ of k not dividing N with $\mathrm{Nm}(\mathfrak{p}) \leq x$ such that $a_\mathfrak{p} = \lfloor 2\sqrt{\mathrm{Nm}(\mathfrak{p})} \rfloor$, or equivalently such that $|\overline{a}_\mathfrak{p} - 2| < 1/\sqrt{\mathrm{Nm}(\mathfrak{p})}$. Let $2 < x_n < x_{n-1} < \cdots < x_1 = x$ be real numbers. Note that

$$R(x) \subseteq M_k(x) \subseteq \bigcup_{j=1}^{n-1} S(x_{j+1}, x_j) \cup \{\mathfrak{p} \nmid N \text{ prime of } k;\ \mathrm{Nm}(\mathfrak{p}) \leq x_n\}. \tag{4.5}$$

Proposition 4.9. *Assume the same hypotheses as in Proposition 4.7. Then*

$$\#M_k(x) \asymp_N \frac{x^{3/4}}{\log(x)} \quad \text{as } x \to \infty.$$

Proof. From (4.5) and Corollary 4.8, we immediately obtain

$$x^{3/4}/\log(x) = O_N(\#M_k(x)).$$

To show that $\#M_k(x) = O_N(x^{3/4}/\log(x))$, for $j = 1, \ldots, n := \lfloor x^{1/16} \rfloor$, define $x_j := x/j^4$. Since $x_n = O(x^{3/4})$, by (4.5) and Corollary 4.8 we have

$$\begin{aligned}\#M_k(x) &\leq \sum_{j=1}^{n-1} \frac{j+1}{\pi x^{1/4}}(\operatorname{Li}(x_j) - \operatorname{Li}(x_{j+1})) + \operatorname{Li}(x_n) + O(x^{1/2+1/16}\log(Nx)\log(x)) \\ &= \frac{2}{\pi x^{1/4}}\operatorname{Li}(x) + \frac{1}{\pi x^{1/4}}\sum_{j=2}^{n-1}\operatorname{Li}(x_j) - \frac{n}{\pi x^{1/4}}\operatorname{Li}(x_n) + O_N\left(\frac{x^{3/4}}{\log(x)}\right).\end{aligned}$$

In view of the above, the proposition will follow from the fact that

$$\sum_{j=2}^{n} \frac{x^{3/4}}{j^4 \log(x/j^4)} = O\left(\frac{x^{3/4}}{\log(x)}\right).$$

But the change of variable $z = x/y^4$ gives

$$\sum_{j=2}^{n} \frac{x^{3/4}}{j^4\log(x/j^4)} = O\left(\int_2^{x^{1/16}} \frac{x^{3/4}}{y^4\log(x/y^4)}\,dy\right) = O\left(\int_{x^{3/4}}^{x/16} \frac{1}{z^{1/4}\log(z)}\,dz\right).$$

Set

$$f(z) = 1/(z^{1/4}\log(z))$$

and

$$\theta(z) = 4z^{3/4}/3,$$

so that integration by parts yields

$$F(x) := \int_{x^{3/4}}^{x/16} f(z)\,dz = \frac{\theta(x/16)}{\log(x/16)} - \frac{\theta(x^{3/4})}{\log(x^{3/4})} + \int_{x^{3/4}}^{x/16} \frac{\theta(z)}{z\log(z)^2}\,dz.$$

Since the first term on the right-hand side above is $O(x^{3/4}/\log(x))$, and the second term is bounded by $F(x)/\log(x^{3/4})$, we deduce that $F(x) = O(x^{3/4}/\log(x))$, which concludes the proof. ∎

Corollary 4.10. *Let A be an elliptic curve with potential CM (say by an imaginary quadratic field K) not defined over k. Under Conjecture 2.4 for every character of $\mathrm{ST}(A_{kK}) \simeq \mathrm{U}(1)$, we have*

$$\#M_k(x) \asymp_N \frac{x^{3/4}}{\log(x)} \quad \text{as } x \to \infty.$$

Proof. Consider the base change A_{kK} and the set of primes of kK defined as

$$M_{kK}^{\mathrm{split}}(x) := \{\mathfrak{P} \nmid N \text{ prime of } kK \text{ split over } k; \\ \mathrm{Nm}(\mathfrak{P}) \le x \text{ and } a_{\mathfrak{P}}(A_{kK}) = \lfloor 2\sqrt{\mathrm{Nm}(\mathfrak{P})} \rfloor\}.$$

Since the number of primes of kK nonsplit over k of norm up to x is $O(\sqrt{x})$, in view of Proposition 4.9 we have

$$\#M_{kK}(x) \sim \#M_{kK}^{\mathrm{split}}(x) \quad \text{as } x \to \infty.$$

On the other hand, the map

$$M_{kK}^{\mathrm{split}}(x) \to M_k(x), \quad \mathfrak{P} \mapsto \mathfrak{P} \cap k,$$

is 2-to-1, and thus

$$\#M_k(x) \sim \tfrac{1}{2}\#M_{kK}^{\mathrm{split}}(x) \quad \text{as } x \to \infty.$$ ∎

As noted in the introduction, it was shown unconditionally by James and Pollack [21, Theorem 1] that

$$\#M_k(x) \sim \frac{2}{3\pi} \frac{x^{3/4}}{\log(x)} \quad \text{as } x \to \infty.$$

That result, which gives a partial answer to a question of Serre [33, Chapter II, Question 6.7], builds on a conditional result of James et al. [22]; that result is similar to ours, except that it aggregates primes for which the Frobenius trace is extremal in both directions. The added ingredient in [21] is the use of unconditional estimates for the number of primes in an imaginary quadratic field lying in a sector; such an estimate has been given by Maknys [23], modulo a correction described in [21]. (For the Gaussian integers, see also [35].)

Remark 4.11. Let A be an abelian variety of dimension $g \ge 1$ defined over k. Let d denote the real dimension of $\mathrm{ST}(A)$. It follows from [32, Section 8.4.4.4] that

$$\mu([2g - x^{-1/2}, 2g]) \cdot \mathrm{Li}(x) \sim C \cdot \frac{x^{1-d/4}}{\log(x)} \quad \text{as } x \to \infty, \tag{4.6}$$

for some constant $C > 0$. When $d > 2$, the count (4.6) is subsumed in the error term of Theorem 3.8. There is thus no hope that the method of proof of Corollary 4.8 can be extended to the case $d > 2$ to obtain the analogous statement.

When $d = 1$ (in which case A is $\overline{\mathbb{Q}}$-isogenous to the power of a CM elliptic curve and $\mathrm{ST}(A) \simeq \mathrm{U}(1)$), it is not difficult to generalize Proposition 4.7 to show that the number of primes $\mathfrak{p}$ such that

$$a_{\mathfrak{p}}(A) = \lfloor 2g\sqrt{\mathrm{Nm}(\mathfrak{p})} \rfloor$$

is again $\asymp_N x^{3/4}/\log(x)$. Note that for these primes, the equality

$$\lfloor 2g\sqrt{\mathrm{Nm}(\mathfrak{p})} \rfloor = g\lfloor 2\sqrt{\mathrm{Nm}(\mathfrak{p})} \rfloor$$

needs to hold because of the Weil–Serre bound.

As Andrew Sutherland kindly explained to us, when $d = 2$ there are already examples of abelian surfaces A defined over $\mathbb{Q}$ for which there are no primes p of good reduction for A such that

$$a_p(A) = 2\lfloor 2\sqrt{p} \rfloor. \tag{4.7}$$

Indeed, let A be the product of two elliptic curves E_1 and E_2 defined over $\mathbb{Q}$ with CM by two nonisomorphic imaginary quadratic fields M_1 and M_2, respectively. Suppose there were a prime $p > 3$ satisfying (4.7) of good reduction for A. Then

$$a_p(E_1) = a_p(E_2) = \lfloor 2\sqrt{p} \rfloor$$

and p would be ordinary for both E_1 and E_2. This would force both M_1 and M_2 to be the splitting field of the local factor of E_1 (which coincides with that of E_2) at p, contradicting the fact that M_1 and M_2 are not isomorphic.

Acknowledgments. We thank Christophe Ritzenthaler for raising the question of the infiniteness of the set of primes at which the Frobenius trace attains the integral part of the Weil bound, and Jeff Achter for directing us to [21]. This occurred during the AGCCT conference held at CIRM, Luminy, in June 2019, where Bucur gave a talk based on this article; we also thank the organizers for their kind invitation. We thank Andrew Sutherland for providing the example of Remark 4.11 and numerical data compatible with Proposition 4.9. We thank Jean-Pierre Serre for sharing the preprint of [33] with us and for precisions on a previous version of this manuscript.

Funding. All three authors were supported by the Institute for Advanced Study during 2018–2019; this includes funding from National Science Foundation grant DMS-1638352. All three authors were additionally supported by the Simons Foundation grant 550033. Bucur was also supported by the Simons Foundation collaboration grant 524015, and by NSF grants DMS-2002716 and DMS-2012061. Kedlaya was additionally supported by NSF grants DMS-1501214, DMS-1802161, DMS-2053473 and by the UCSD Warschawski Professorship. Fité was additionally supported by the Ramón y Cajal fellowship RYC-2019-027378-I, by the María de Maeztu program CEX2020-001084-M, by the DGICYT grant MTM2015-63829-P, and by the ERC grant 682152.

Tab. 1. Table of notations

Notation	Meaning	First usage
a	Rank of Lie algebra $\mathfrak{a}$	Section 3.3
$a_{\mathfrak{p}}$	Frobenius trace of A at $\mathfrak{p}$	Section 1
$\overline{a}_{\mathfrak{p}}$	Normalized version of $a_{\mathfrak{p}}$	Section 1, (1.1)
A	Abelian variety over k	Section 1
$\mathfrak{a}$	Abelian Lie algebra, factor of $\mathfrak{g}$	Section 3.3
$\alpha_{\mathfrak{p},j}$	Reciprocal roots of local L-factor	Section 2.2
D	Vinogradov function associated to Δ, T	Proposition 3.4
d_χ	Degree of the character χ	Section 2.2
δ_I	Characteristic function of I	Section 1
Δ	Cutoff parameter in definition of D	Proposition 3.4
$\epsilon_{\mathfrak{g}}$	Dependence on $\mathfrak{g}$ in Theorem 1.1	(1.2)
$\epsilon_{\mathfrak{g},\mathfrak{g}'}$	Dependence on $\mathfrak{g}, \mathfrak{g}'$ in Theorem 4.3	(4.2)
F	Average of D over $\mathcal{W}$	(3.15)
g	Dimension of A	Section 1
$\mathfrak{g}$	Lie algebra of $\mathrm{ST}(A)$	Section 3.3
Γ_λ	Representation of $\mathfrak{g}$ with highest weight λ	Section 3.1
H	Cartan subgroup of $\mathrm{ST}(A)$	Section 3.3
h	Rank of Lie algebra $\mathfrak{h}$	Section 3.1
$\mathfrak{h}$	Cartan subalgebra of $\mathfrak{s}$	Section 3.1
I	Subinterval of $[-2g, 2g]$	(1.1)
$\lvert I \rvert$	Length of I	Section 1
k	Number field over which A is defined	Section 1
$\mathrm{Li}(x)$	Logarithmic integral	(1.1)
M	Cutoff parameter in weight space	Proof of Theorem 3.8
$M_k(x)$	Extremal primes of norm up to x	Section 1
m_λ^μ	Weight multiplicity	Section 3.1
μ	Pushforward of Haar measure on $\mathrm{ST}(A)$	Section 1
N	Absolute conductor of A	Section 1
$\nu_{\mathfrak{g}}$	Cutoff for O notation	(1.3)
$\mathfrak{p}$	Prime ideal of k	Section 3.1
φ	Size of set Φ^+	(1.2)
Φ	Root system for $\mathfrak{s}$	Section 3.1
q	Rank of Lie algebra $\mathfrak{g}$ $(= h + a)$	(1.2)
$\mathcal{R}$	Lattice of integral weights of $\mathfrak{s}$	Section 3.1
S	Simple roots of Φ	Section 3.1
$\mathfrak{s}$	Semisimple factor of $\mathfrak{g}$	Section 3.1, Section 3.3
$\mathrm{ST}(A)$	Sato–Tate group of A	Section 1, Section 2.1
T	Trace map on $\mathbb{R}^q$	Section 3.3, (3.12)
$T_\ell(A)$	Tate module of A	Section 1
$V_\ell(A)$	$T_\ell(A) \otimes \mathbb{Q}_\ell$	Section 1
w	Element of $\mathcal{W}$	Section 3.1
$\mathcal{W}$	Weyl group of $\mathfrak{s}$	Section 3.1
ω_j	Basis element of fundamental weights	Section 3.1

References

[1] Banaszak, G., Kedlaya, K. S.: An algebraic Sato–Tate group and Sato–Tate conjecture. Indiana Univ. Math. J. **64**, 245–274 (2015) Zbl 1392.11041 MR 3320526

[2] Bourbaki, N.: Lie groups and Lie algebras: Chapters 7–9. Elements of Mathematics (Berlin), Springer, Berlin (2005) Zbl 1139.17002 MR 2109105

[3] Brumer, A., Kramer, K.: The conductor of an abelian variety. Compos. Math. **92**, 227–248 (1994) Zbl 0818.14016 MR 1283229

[4] Bucur, A., Kedlaya, K. S.: An application of the effective Sato–Tate conjecture. In: Frobenius distributions: Lang–Trotter and Sato–Tate conjectures, Contemp. Math. 663, American Mathematical Society, Providence, RI, 45–56 (2016) Zbl 1417.11103 MR 3502938

[5] Cantoral-Farfán, V., Commelin, J.: The Mumford–Tate conjecture implies the algebraic Sato–Tate conjecture of Banaszak and Kedlaya. Indiana Univ. Math. J. **71**, 2595–2603 (2022) Zbl 1506.14094 MR 4530050

[6] Chen, E., Park, P. S., Swaminathan, A. A.: Elliptic curve variants of the least quadratic non-residue problem and Linnik's theorem. Int. J. Number Theory **14**, 255–288 (2018) Zbl 1428.11106 MR 3726253

[7] Chen, H., Jones, N., Serban, V.: The Lang–Trotter conjecture for products of non-CM elliptic curves. Ramanujan J. **59**, 379–436 (2022) Zbl 1517.11058 MR 4480293

[8] Cojocaru, A. C., Wang, T.: Bounds for the distribution of the Frobenius traces associated to products of non-CM elliptic curves. Canad. J. Math. **75**, 687–712 (2023) MR 4586829

[9] Cojocaru, A. C., Wang, T.: Bounds for the distribution of the Frobenius traces associated to a generic abelian variety. arXiv:2207.02913v1 (2022)

[10] Deligne, P., Milne, J. S., Ogus, A., Shih, K.-y.: Hodge cycles, motives, and Shimura varieties. Lecture Notes in Math. 900, Springer, Berlin (1982) Zbl 0465.00010 MR 0654325

[11] Duke, W.: Some problems in multidimensional analytic number theory. Acta Arith. **52**, 203–228 (1989) Zbl 0631.12008 MR 1031335

[12] Fité, F., Kedlaya, K. S., Rotger, V., Sutherland, A. V.: Sato–Tate distributions and Galois endomorphism modules in genus 2. Compos. Math. **148**, 1390–1442 (2012) Zbl 1269.11094 MR 2982436

[13] Fité, F., Kedlaya, K. S., Sutherland, A. V.: Sato–Tate groups of some weight 3 motives. In: Frobenius distributions: Lang–Trotter and Sato–Tate conjectures, Contemp. Math. 663, American Mathematical Society, Providence, RI, 57–101 (2016) Zbl 1411.11089 MR 3502939

[14] Fulton, W., Harris, J.: Representation theory. Grad. Texts in Math. 129, Springer, New York (1991) Zbl 0744.22001 MR 1153249

[15] Ghitza, A.: Distinguishing Hecke eigenforms. Int. J. Number Theory **7**, 1247–1253 (2011) Zbl 1253.11052 MR 2825970

[16] Ghitza, A., Sayer, R.: Hecke eigenvalues of Siegel modular forms of "different weights". J. Number Theory **143**, 125–141 (2014) Zbl 1302.11028 MR 3227338

[17] Goldfeld, D., Hoffstein, J.: On the number of Fourier coefficients that determine a modular form. In: A tribute to Emil Grosswald: number theory and related analysis, Contemp. Math. 143, American Mathematical Society, Providence, RI, 385–393 (1993) Zbl 0805.11040 MR 1210527

[18] Grothendieck, A.: Modèles de Néron et monodromie (with an appendix by M. Raynaud). In: Groupes de monodromie en géometrie algébrique, SGA 7, Exposé IX, Lecture Notes in Math. 288, Springer, Berlin, 313–523 (1970) Zbl 0248.14006

[19] Gupta, R. K.: Characters and the q-analog of weight multiplicity. J. London Math. Soc. (2) **36**, 68–76 (1987) Zbl 0649.17009 MR 0897675

[20] Hecke, E.: Eine neue Art von Zetafunktionen und ihre Beziehungen zur Verteilung der Primzahlen. Math. Z. **6**, 11–51 (1920) Zbl 47.0152.01 MR 1544392

[21] James, K., Pollack, P.: Extremal primes for elliptic curves with complex multiplication. J. Number Theory **172**, 383–391 (2017); see also Errata Zbl 1419.11091 MR 3573159

[22] James, K., Tran, B., Trinh, M.-T., Wertheimer, P., Zantout, D.: Extremal primes for elliptic curves. J. Number Theory **164**, 282–298 (2016) Zbl 1416.11081 MR 3474389

[23] Maknys, M.: On the distance between consecutive prime ideal numbers in sectors. Acta Math. Hungar. **42**, 131–138 (1983) Zbl 0525.10026 MR 0716559

[24] Murty, M. R.: Congruences between modular forms. In: Analytic number theory (Kyoto, 1996), London Math. Soc. Lecture Note Ser. 247, Cambridge University Press, Cambridge, 309–320 (1997) Zbl 0910.11018 MR 1694998

[25] Murty, V. K.: Explicit formulae and the Lang–Trotter conjecture. Rocky Mountain J. Math. **15**, 535–551 (1985) Zbl 0587.14009 MR 0823264

[26] Sengupta, J.: Distinguishing Hecke eigenvalues of primitive cusp forms. Acta Arith. **114**, 23–34 (2004) Zbl 1097.11020 MR 2067870

[27] Serre, J.-P.: Facteurs locaux des fonctions zêta des varietés algébriques (définitions et conjectures). In: Séminaire Delange–Pisot–Poitou, 11e année: 1969/70, Théorie des nombres, Fasc. 2, exp. 19, 15 pp., Secrétariat Math., Paris (1970) Zbl 0214.48403 MR 3618526

[28] Serre, J.-P.: Linear representations of finite groups. Grad. Texts in Math. 42, Springer, New York (1977) Zbl 0355.20006 MR 0450380

[29] Serre, J.-P.: Quelques applications du théorème de densité de Chebotarev. Inst. Hautes Études Sci. Publ. Math. **54**, 323–401 (1981) Zbl 0496.12011 MR 0644559

[30] Serre, J.-P.: Complex semisimple Lie algebras. Springer, New York (1987) Zbl 0628.17003 MR 0914496

[31] Serre, J.-P.: Propriétés conjecturales des groupes de Galois motiviques et des représentations l-adiques. In: Motives (Seattle, WA, 1991), Proc. Sympos. Pure Math. 55, American Mathematical Society, Providence, RI, 377–400 (1994) Zbl 0812.14002 MR 1265537

[32] Serre, J.-P.: Lectures on $N_X(p)$. Chapman & Hall/CRC Research Notes in Mathematics 11, CRC Press, Boca Raton, FL (2012) Zbl 1238.11001 MR 2920749

[33] Serre, J.-P. (ed.): Rational points on curves over finite fields. Doc. Math. (Paris) 18, Société Mathématique de France, Paris (2020) Zbl 1475.11002 MR 4242817

[34] Vinogradov, I. M.: The method of trigonometrical sums in the theory of numbers. Dover Publ., Mineola, NY (2004) Zbl 1093.11001 MR 2104806

[35] Zarzycki, P.: Distribution of primes of imaginary quadratic fields in sectors. J. Number Theory **37**, 152–160 (1991) Zbl 0717.11051 MR 1092601