L-FUNCTIONS IN ARITHMETIC STATISTICS

ALINA BUCUR

Arithmetic statistics is the study of number-theoretic objects in aggregates, rather than in isolation. This study takes many different forms, but in this paper we will concentrate on (some) instances where the behavior of L-functions plays an important role.

Perhaps the best known result in arithmetic statistics where L-functions determine what happens is the Prime Number Theorem. It states that, as $X \to \infty$,

$$\#\{p < X; p \text{ prime}\} \sim \mathrm{Li}(x), \quad \text{where } \mathrm{Li}(x) = \int_2^\infty \frac{dt}{\log t}. \tag{1}$$

While L-functions are not mentioned at all in the statement, the proof of Hadamard and de la Vallée Poussin is based on the analytic properties of the Riemann's zeta function

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p\text{prime}} \left(1 - p^{-s}\right)^{-1} \text{ for } \mathrm{Re}(s) > 1$$

and its connection with prime numbers given by the Euler product.

Some typical questions in arithmetic statistics are the following. What is the probability that a random integer is squarefree? or is prime? How many points with integer coordinates are there on an elliptic curve defined over $\mathbb{Q}$? How many number fields of degree d are there with discriminant of absolute value at most X? What does the class group of a random quadratic field look like? Many aspects of the subject are well-understood, but many more remain the subject of conjectures, like the Cohen-Lenstra heuristics or Malle's conjecture.

Our starting point is the theme of rational points on curves defined over finite fields, which is the subject of the Serre's book [54]. It is its publication that was celebrated during the 2021 conference organized by Alp Bassa, Joan-Carles Lario, Elisa Lorenzo García, Christophe Ritzenthaler and René Schoof and that led to the publication of this volume. We are concerned with the following question.

Question 2. *What can we say about the number of $\mathbb{F}_q$-points on a curve C as either the base field $\mathbb{F}_q$ varies and/or the curve C varies?*

The goal of this paper it to provide an overview of the current state of knowledge about the subject in three different directions, which all lead to questions in arithmetic statistics and are related to the behavior of L-functions. The first direction will be the *geometric* direction where we look at the moduli space of curves of genus g for a fixed g and let the number of elements in the field $\mathbb{F}_q$ grow to infinity. The second direction is the *arithmetic* direction where we start with a fixed curves defined over the rationals and reduce it modulo larger and larger primes. Finally, we will discuss the *probabilistic* situation, where we fix the field $\mathbb{F}_q$ and look at curves of genus g where g grows to infinity.

Date: September 15, 2021.

2010 *Mathematics Subject Classification.* 11G05, 11R44.

supported by the Simons Foundation (collaboration grant #524015) and NSF (DMS 2002716).

1. Setup and notation

Denote by $\mathbb{F}_q$ the finite field with q elements (necessarily q is a prime power) and by C a smooth, complete, geometrically irreducible projective curve of genus g defined over $\mathbb{F}_q$. We are trying to understand how $\#C(\mathbb{F}_q)$ varies as C varies and/or as q varies.

Each such curve has a zeta function that is defined as

$$\zeta_C(s) = \exp\left(\sum_{m=1}^{\infty} \frac{\#C(\mathbb{F}_{q^m})}{m} q^{-ms}\right).$$

Perhaps the most significant fact that allows one to make progress on this question are the Weil Conjectures, formulated by André Weil in 1949 [58]. Weil himself proved [57] his conjectures for curves. Thus we know that for a curve C

$$\zeta_C(s) = Z_C(q^{-s})$$

where $Z_C(u)$ is a rational function of u of the form

$$Z_C(u) = \frac{P_C(u)}{(1-u)(1-qu)}.$$

The numerator $P_C(u)$ is a polynomial of degree $2g$, and its zeros have absolute value $q^{-1/2}$. This last assertion is the equivalent of the Riemann hypothesis in this context. Moreover, we know that the number of $\mathbb{F}_q$-points of C is related to the trace of the Frobenius operator associated to the curve C via the relation

$$\#C(\mathbb{F}_q) = q + 1 - \mathrm{Trace}(\mathrm{Frob}_C) \tag{3}$$

and that the eigenvalues of Frob_C are the reciprocals of the the zeros of $Z_C(u)$.

The next big leap in our understanding came from Appendix A of Serre's book *Abelian ℓ-adic representations and elliptic curves* [48] published in 1968. In it, Serre uses the Peter-Weyl theorem from representation theory and shows how equidistribution results follow from analytic properties of (certain) L-functions, namely from their analytic continuation. The use the analytic properties of L-functions to get results about the distributions of arithmetic objects of interest goes as far back as the proof of the Prime Number Theorem by Hadamard and de la Vallée-Poussin. In that case, the asymptotic (1) is obtained from the meromorphic continuation of the Riemann zeta function $\zeta(s)$ and its nonvanishing on the line $\mathrm{Re}(s) = 1$, except for the pole at $s = 1$. Similarly, the Chebotarev density theorem follows from the analytic properties of certain Artin L-functions that can be viewed as associated to representations of some finite groups. This point of view was generalized by Serre to compact groups.

Let G be a compact group and denote by $\mathrm{Conj}(G)$ the space of conjugacy classes of G. Let $(x_v)_{v\in\Sigma}$ be a countable family in $\mathrm{Conj}(G)$. Let ρ be an irreducible representation of G with character χ. Define the L-function associated to ρ as

$$L(s,\rho) = \prod_{v\in\Sigma} \frac{1}{\det\left(1 - \rho(x_v)\mathrm{N}v^{-s}\right)}$$

where $v \mapsto \mathrm{N}v$ is a function from the set Σ to the integers that will be made more precise later.

Since the determinant is invariant under conjugation, the above expression depends only on the character χ so we can also write

$$L(s,\chi) = L(s,\rho).$$

Namely in [48] he proves the following two results. We will make three assumptions about the function $v \mapsto \mathrm{N}v$, namely

- The function $\mathrm{N}v$ takes values in $\mathbb{Z}_{\geq 2}$;
- The (Euler) product
$$\prod_{v\in\Sigma} \frac{1}{1-\mathrm{N}v^{-s}}$$
converges absolutely for $\mathrm{Re}(s) > 1$, has meromorphic continuation to $\mathrm{Re}(s) \geq 1$, and has no zeros and no poles on the line $\mathrm{Re}(s) = 1$ except for a simple pole at $s = 1$;
- The infinite product $L(s,\rho) = L(s,\chi)$ converges absolutely for $\mathrm{Re}(s) > 1$, has meromorphic continuation to $\mathrm{Re}(s) \geq 1$, and has no zeros and no poles on the line $\mathrm{Re}(s) = 1$ except for possibly a pole at $s = 1$.

In [48], Serre proves the following two results.

Theorem 4 (Serre). *Under these conditions,*

(1)
$$\#\{v \in \Sigma; \mathrm{N}v \leq X\} \sim \frac{X}{\log X}$$

(2)
$$\sum_{\mathrm{N}v\leq X} \chi(x_v) = c(\chi)\frac{X}{\log X} + o\left(\frac{X}{\log X}\right),$$

where $c(\chi)$ denotes the order of the pole of $L(s,\rho)$ at $s = 1$.

As usual, $f(X) = o(g(X))$ means that $\lim_{X\to\infty} \frac{f(X)}{g(X)} = 0$ and we take $c(\chi)$ to be negative if $L(s,\rho)$ has a zero at $s = 1$. Note that the same way the Riemann hypothesis would give a power-saving error term in the Prime Number Theorem, the Riemann hypothesis (or a suitable zero-free region) for the function $L(s,\chi)$ would give a power saving in the error term in Theorem 4.

Theorem 5 (Serre). *If in addition there exists a constant $M > 0$ such that*

$$\#\{v \in \Sigma; \mathrm{N}v = X\} < M \textit{ for all } X,$$

then $(x_v)_{v\in\Sigma}$ are equidistributed with respect to the normalized Haar measure on the compact group G if and only if $L(s,\chi)$ is holomorphic and nonzero at $s = 1$ for all nontrivial irreducible characters χ of G.

2. Geometric situation

For this section, C will be a curve of genus g, with g fixed, defined over $\mathbb{F}_q$. We want to understand how $\#C(\mathbb{F}_q)$ varies as we let $q \to \infty$. In view of the relation (3), this would follow from understanding the distribution of the zeros of the zeta function $Z_C(u)$ associated to the curve C, but understanding the distribution of the zeros would be strictly more information than just understanding $\#C(\mathbb{F}_q)$.

In 1968, Birch [5] begins to answer this question in the case of $g = 1$. Namely, he computes the even moments of the Trace(Frob$_E$) as E varies over the elliptic curves defined over a field $\mathbb{F}_p$ with a prime $p \geq 5$ number of elements. This paper offers a glimpse of the deep connection between elliptic curves and modular forms as the Ramanujan τ-function appears in the tenth moment.

Namely, Birch uses the Selberg trace formula to compute

$$\begin{aligned}
\mathbb{E}(\mathrm{Trace}(\mathrm{Frob}_E)^2) &= p^2\\
\mathbb{E}(\mathrm{Trace}(\mathrm{Frob}_E)^4) &= 2p^3 - 3p\\
\mathbb{E}(\mathrm{Trace}(\mathrm{Frob}_E)^6) &= 5p^4 - 9p^2 - 5p\\
\mathbb{E}(\mathrm{Trace}(\mathrm{Frob}_E)^8) &= 14p^5 - 28p^3 - 20p^2 - 7p\\
\mathbb{E}(\mathrm{Trace}(\mathrm{Frob}_E)^{10}) &= 42p^6 - 90p^4 - 75p^3 - 35p^2 - 9p - \tau(p)
\end{aligned}$$

as well as the general formula

$$\mathbb{E}(\mathrm{Trace}(\mathrm{Frob}_E)^{2k}) \sim \frac{(2k)!}{k!(k+1)!} p^{k+2} \text{ as } k \to \infty.$$

Then in 1973, Yoshida [62] employs the analytic properties of L-functions of elliptic curves and their symmetric powers (which he relates to fiber products of E with itself) to prove an analogue of the Sato-Tate conjecture (see Section 4) in function fields.

The next big step in our quest comes in 1980 when Deligne proves his influential equidistribution theorem in [21] as part of his proof of the Weil Conjectures for general projective varieties over $\mathbb{F}_q$. Start with a family of smooth proper varieties X_T indexed by a finite type space T over a finite field $\mathbb{F}_q$. Fix a positive integer m and take the m-weight part of the zeta function of X_T. Then there is a group G (the monodromy group of the family) and a sequence of elements (defined up to conjugation) in G which correspond to the closed points of the base space such that when one averages (in the correct manner that stamps out the possible oscillatory behavior) over the Conj(G) the sequence of points will converge weakly (in distribution) to the measure induced by the Haar measure of G. This will imply that the factors of the zeta function will have the same distribution as the random matrices in that group.

Let us make this more precise in the case where the base space is a curve. Let K be the function field of a curve over $\mathbb{F}_q$ and let S be a finite set of places of K. We will denote by K_S^{ur} the maximal extension of K (inside some fixed algebraic closure) that is unramified at all places outside S. Let ρ be a representation of $\mathrm{Gal}(K_S^{ur}/K)$ on the $\mathbb{Q}_\ell$-vector space V with ℓ not dividing the characteristic of K. The arithmetic monodromy group G_{arith} of ρ is defined as the Zariski closure of the image of ρ in $\mathrm{GL}(V)$. We will think of it as a complex Lie group via an embedding of $\overline{\mathbb{Q}}_\ell$ into $\mathbb{C}$. Its geometric monodromy group G_{geom} is the Zariski closure of $\rho(\mathrm{Gal}(K_S^{ur}/K\overline{\mathbb{F}}_q))$ in $\mathrm{GL}(V)$.

We will assume for simplicity that $G_{\mathrm{geom}} = G_{\mathrm{arith}} = G$, which allows us to avoid the aforementioned subtleties in the averaging process. Let Γ be a maximal compact subgroup of G and denote by $\varphi_v \in \mathrm{Gal}(K_S^{ur}/K)$ the arithmetic Frobenius (or its conjugacy class) at each place $v \notin S$.

Theorem 6 (Deligne). *Let $\rho : \mathrm{Gal}(K_S^{ur}/K) \to \mathrm{GL}(V)$ be an ℓ-adic representation which is pure of weight 0. Assume that the arithmetic and geometric monodromy groups coincide, i.e. $G_{\mathrm{geom}} = G_{\mathrm{arith}} = G$. Then, as v runs over the places of K outside the set S, the conjugacy classes $\theta(v)$ in Γ corresponding to $\rho(\varphi_v)^{\mathrm{semisimple}}$ are weakly equidistributed in $\mathrm{Conj}(\Gamma)$ with respect to the direct image of normalized Haar measure on Γ. In particular, for any nontrivial, irreducible, complex representation Λ of Γ (or of G) and any $r \in \mathbb{Z}_{>0}$, we have*

$$\left| \frac{\displaystyle\sum_{v\notin S, \deg v | r} (\deg v) \operatorname{Trace} \Lambda \left(\theta(v)^{r/(\deg v)}\right)}{\displaystyle\sum_{v\notin S, \deg v | r} \deg v} \right| = O\left(\frac{\dim \Lambda}{q^{r/2}}\right).$$

Serre himself makes use of his ideas in [48] to study in [52] the (normalized) eigenvalues of the Hecke operator T_p on the space of parabolic modular forms of weight k on the congruence group $\Gamma_0(N)$ are asymptotically equidistributed with respect to the measure

$$\mu(x) = \frac{p+1}{\pi} \frac{(4-x^2)^{1/2}}{2(p+2+p^{-1}-x^2)} dx$$

on [-2,2] as $k + N \to \infty$ for k even and $p \nmid N$.

In 1998, Katz and Sarnak published their seminal work [35]. Their starting point is to use Deligne's equidistribution theorem to prove that the eigenvalues of the Frobenius endomorphism associated to the hyperelliptic curves of genus g (where g is fixed) defined over $\mathbb{F}_q$ considered over larger and larger finite extensions of $\mathbb{F}_q$ (i.e. $\mathbb{F}_{q^m}$ with $m \to \infty$) are asymptotically distributed like the eigenvalues of random matrices in $\mathrm{USp}(2g)$. Here random means random with respect to the Haar measure. Since the cup pairing imposes a certain symmetry on the eigenvalues of Frobenius associated to curves in the moduli space $\mathcal{M}_g$ of all curves of genus g, it follows that the hyperelliptic curves are generic in $\mathcal{M}_g$ itself.

They also formulated a more general philosophy and proved it in a host of instances. Namely, they predict that the distribution of eigenvalues of Frobenius for a family of curves defined over $\mathbb{F}_q$ approaches, in the large q limit, like the eigenvalues of some ensemble of random matrices and that ensemble is dictated by the monodromy group of the family.

3. Probabilistic situation

We will now look at the mirror situation, that is the large genus limit. For this, we will fix a finite field $\mathbb{F}_q$ and study the distribution of $\#C(\mathbb{F}_q)$ where C is a curve of genus g as $g \to \infty$.

The case of hyperelliptic curves, or alternatively $\mathbb{Z}/2\mathbb{Z}$ covers of the projective line $\mathbb{P}^1$ over a field $\mathbb{F}_q$ with q odd, was studied by Kurlberg and Rudnick [37] in 2009. The technique used by the two authors is to relate the $\#C(\mathbb{F}_q)$ to sums of the quadratic character associated to the cover itself. This was later expanded to $\mathbb{Z}/\ell\mathbb{Z}$ covers of the projective line $\mathbb{P}^1$ over a field $\mathbb{F}_q$ with $q \equiv 1 \pmod{\ell}$ by various authors. In 2010 Bucur, David, Feigon, and Lalín [10] looked at certain connected components of the moduli space of all such covers. Then in 2015, Cheong, Wood, and Zaman [17] employed Kummer theory to study the family $y^\ell = f(x)$ where $f(x)$ is an ℓ-power free monic polynomial whose degree goes to infinity. The case of the whole moduli space of $\mathbb{Z}/\ell\mathbb{Z}$ covers of the projective line was studied by Bucur, David, Feigon, Kaplan, Lalín, Ozman, and Wood [12]. They proved the following result.

Theorem 7 (Bucur, David, Feigon, Kaplan, Lalín, Ozman, Wood)**.** *Let $\mathcal{H}_{g,\ell}$ be the moduli space of $\mathbb{Z}/\ell\mathbb{Z}$ Galois covers of $\mathbb{P}^1$ of genus g. Then, as $g \to \infty$,*

$$\frac{|\{C \in \mathcal{H}_{g,\ell}(\mathbb{F}_q) : \#C(\mathbb{F}_q) = m\}|'}{|\mathcal{H}_{g,\ell}(\mathbb{F}_q)|'} = \mathrm{Prob}\left(X_1 + \dots X_{q+1} = m\right) + O_\ell\left(\frac{1}{g}\right),$$

where the X_i's are independent identically distributed random variables such that

$$X_i = \begin{cases} 0 & \textit{with probability } \dfrac{(\ell-1)q}{\ell(q+\ell-1)}, \\[2ex] 1 & \textit{with probability } \dfrac{\ell-1}{q+\ell-1}, \\[2ex] \ell & \textit{with probability } \dfrac{q}{\ell(q+\ell-1)}. \end{cases}$$

In the formula, as usual, the $'$ notation means that the covers C on the moduli space are counted with the usual weights $1/\#\operatorname{Aut}(C)$.

Note that the special case $\ell = 2$ was proved in [37]. The $q+1$ random variables corresponds to the $\mathbb{F}_q$-points of $\mathbb{P}^1$ and the proof essentially shows that the number of points in each fiber of the covering map $C \to \mathbb{P}^1$ is independent of what happens in the other fibers. Since $\mathbb{Z}/\ell\mathbb{Z}$ is a simple group, the number of points in each fiber will have to be one of $0, 1$ or ℓ.

Going beyond the case of simple groups, Wood [59] discussed S_3 covers of $\mathbb{P}^1$ and Lorenzo García, Meleleo and Millione [40] discussed the case of biquadratic covers of the projective line. The latter paper uses a modification of the character sum method from [37], while Wood relates trigonal curves to cubic extensions of function fields, and then uses the work of Datskovsky and Wright [20] to count cubic extensions with every possible fiberwise behavior above each rational point of the base curve.

Theorem 8 (Lorenzo García, Meleleo, Millione)**.** *Let $\mathcal{B}_g(\mathbb{F}_q)$ be the family of genus g quartic non-cyclic cover of the projective line $\mathbb{P}^1_{\mathbb{F}_q}$, and consider the following decomposition*

$$\mathcal{B}_g(\mathbb{F}_q) = \bigcup_{g_1+g_2+g_3=g} \mathcal{B}_{(g_1,g_2,g_3)}(\mathbb{F}_q)$$

where $\mathcal{B}_{(g_1,g_2,g_3)}(\mathbb{F}_q)$ denotes the subfamily of curves $C \in \mathcal{B}_g(\mathbb{F}_q)$ such that the three hyperelliptic quotients of C have genera g_1, g_2 and g_3.

If the three genera g_1, g_2, g_3 go to infinity, then we have that

$$\frac{|\{C \in \mathcal{B}_{(g_1,g_2,g_3)}(\mathbb{F}_q) : \operatorname{Tr}(\operatorname{Frob}_C) = -M\}|'}{|\mathcal{B}_{(g_1,g_2,g_3)}(\mathbb{F}_q)|'} = \operatorname{Prob}\left(\sum_{j=1}^{q+1} X_j = M\right)$$

where the X_j are i.i.d. (identically independently distributed) random variables such that

$$X_i = \begin{cases} -1 & \textit{with probability } \frac{3(q+2)}{4(q+3)} \\ 1 & \textit{with probability } \frac{6}{4(q+3)} \\ 3 & \textit{with probability } \frac{q}{4(q+3)}. \end{cases}$$

Theorem 9 (Wood)**.** *Assume $\mathbb{F}_q$ has characteristic ≥ 5 and let*

$$T_g := \{\pi : C \to \mathbb{P}^1; C \textit{ is a smooth, geometrically integral, genus } g \textit{ curve with } \pi \textit{ degree } 3\}.$$

We have

$$\lim_{g\to\infty} \frac{\#\{C \in T_g(\mathbb{F}_q) \mid \#C(\mathbb{F}_q) = k\}|}{\#T_g(\mathbb{F}_q)} = \operatorname{Prob}(X_1 + \cdots + X_{q+1} = k),$$

where the X_i are independent identically distributed random variables and

$$X_i = \begin{cases} 0 & \textit{with probability } \frac{2q^2}{6q^2+6q+6} \\ 1 & \textit{with probability } \frac{3q^2+6}{6q^2+6q+6} \\ 2 & \textit{with probability } \frac{6q}{6q^2+6q+6} \\ 3 & \textit{with probability } \frac{q^2}{6q^2+6q+6}. \end{cases}$$

A slightly different approach comes from a geometric sieve introduced by Poonen [46].

Theorem 10 (Poonen)**.** *Let X be a quasi-projective subscheme of $\mathbb{P}^n$ over $\mathbb{F}_q$, Z finite subscheme of $\mathbb{P}^n$ such that $U = X \setminus (X \cap Z)$ is smooth of dimension m. Fix $T \subset H^0(Z, \mathcal{O}_Z)$. Given a homogeneous polynomial f of degree d, let $f|_Z$ denote the element of $H^0(Z, \mathcal{O}_Z)$ that on each connected component Z_i equals the restriction of $x_j^{-d} f$ to Z_i, where $j = j(i)$ is the smallest integer $0 \leq j \leq n$ such that the coordinate x_j is invertible on Z_i. Then*

$$\frac{\#\{f \in S_d; H_f \cap U \textit{ smooth, } f|_Z \in T\}}{\#S_d} \sim \frac{\#T}{\#H^0(Z, \mathcal{O}_Z)} \zeta_U(m+1)^{-1} \textit{ as } d \to \infty.$$

The strategy used by Poonen is based on the fact that one can compute the probability that H_f is smooth at a closed point P of the subscheme U and if these conditions were independent we would get that the probability that H_f is smooth was their product. For the proof Poonen uses a sieving argument that separately treats the closed points of X of low, medium, and high degree (as a function of d) and treat each of the three sets separately. For the points of low degree (including the points of Z), one proves that the conditions at each point are indeed independent. Their product gives the main term. For a single point in the middle range, one similarly shows that singularities manifest with the probability predicted by the local factor. One no longer has independence of these local conditions, but they together contribute so little (which is due to the Weil bounds) to the product that they can be controlled by crude estimates. For the points of high degree, one must use a global argument since there are too many points to control individually. Poonen introduces the clever device of writing the f_i so as to partially decouple the low-order Taylor coefficients; one then gets a suitable bound using Bézout's theorem. This trick is the cause of the explicit appearance of p in the error term as computed by Bucur, David, Feigon and Lalín in [11]; it relies on the fact that the derivative of a p-th power vanishes in characteristic p.

Theorem 11 (Bucur, David, Feigon, Lalín)**.** *Let* $X_1, \dots, X_{q^2+q+1}$ *be* q^2+q+1 *i.i.d. Bernoulli random variables taking the value* 1 *with probability* $(q+1)/(q^2+q+1)$ *and the value* 0 *with probability* $q^2/(q^2+q+1)$*. Then, for* $0 \le t \le q^2+q+1$*,*

$$\begin{aligned}\frac{\#\left\{F \in S_d^{\mathrm{ns}}; \#C_F(\mathbb{F}_q) = t\right\}}{\#S_d^{\mathrm{ns}}} &= \operatorname{Prob}\left(X_1 + \cdots + X_{q^2+q+1} = t\right) \\ &\times \left(1 + O\left(q^t\left(d^{-1/3} + (d-1)^2 q^{-\min\left(\left\lfloor \frac{d}{p}\right\rfloor+1, \frac{d}{3}\right)} + dq^{-\left\lfloor \frac{d-1}{p}\right\rfloor - 1}\right)\right)\right),\end{aligned}$$

where $\lfloor \cdot \rfloor$ *denotes the integer part.*

A generalization of Poonen sieve was developed by Bucur and Kedlaya [14] to study complete intersections in projective spaces.

Theorem 12 (Bucur-Kedlaya)**.** *Let* X *be a quasiprojective subscheme of dimension* $m \ge 0$ *of the projective space* $\mathbb{P}^n$ *over some finite field* $\mathbb{F}_q$ *of characteristic* p*. Let* Z *be a finite subscheme of* X *for which* $U = X \backslash Z$ *is smooth of dimension* m*, and define* $z = \dim_{\mathbb{F}_q} H^0(Z, \mathcal{O}_Z)$*. For any* k*-tuple of positive integers* $\mathbf{d} = (d_1, \dots, d_k)$ *we denote* $S_{\mathbf{d}} = S_{d_1} \times \cdots \times S_{d_k}$ *(*k*-tuples of homogeneous polynomials in* $n+1$ *variables) and for each* $\mathbf{f} = (f_1, \dots, f_k) \in S_{\mathbf{d}}$*, we will write* $H_{\mathbf{f}} = H_{f_1} \cap \cdots \cap H_{f_k}$*.*

Choose an integer $k \in \{1, \dots, m+1\}$*,* $z \le d_1 \le \cdots \le d_k$*, and a subset* T *of* $H^0(Z, \mathcal{O}_Z(\mathbf{d}))$*.*

$$\mathcal{P}_{\mathbf{d}} = \{\mathbf{f} \in S_{\mathbf{d}} : H_{\mathbf{f}} \cap U \text{ is smooth of dimension } m-k, \text{ and } \mathbf{f}|_Z \in T\}.$$

Then

$$\frac{\#\mathcal{P}_{\mathbf{d}}}{\#S_{\mathbf{d}}} = \frac{\#T}{q^z} \prod_{x \in U^\circ} \left(1 - q^{-k\deg(x)} + q^{-k\deg(x)} L(q^{\deg(x)}, m, k)\right) \\ + O((d_1 - z + 1)^{-(2k-1)/m} + d_k^m q^{-d_1/\max\{m+1,p\}}),$$

where

$$L(q, m, k) = \prod_{j=0}^{k-1} (1 - q^{-(m-j)})$$

denotes the probability that k *randomly chosen vectors in* $\mathbb{F}_q^m$ *are linearly independent.*

The argument follows the same path as Poonen's, separating the points into low, medium and high degree and computing the contribution of each type of points separately. However, Bucur and Kedlaya get an unexpected average number of points. For instance, for the intersection of two surfaces of degrees d_1 and d_2 in $\mathbb{P}^3$, the average number of points will be

$$q+1-\frac{q^{-2}(1+q^{-1})}{1+q^{-2}-q^{-5}}<q+1$$

In general, for $n-1$ hypersurfaces intersecting in $\mathbb{P}^n$, the average number of points is

$$(q+1)-(q+1)(1-q^{1-n})\frac{1-(1-q^{-n})\dots(1-q^{-3})}{1-q^{1-n}+q^{1-n}(1-q^{-n})\dots(1-q^{-2})}<q+1,$$

and it is of size $(q+1)\big(1+O(q^{-3})\big)$.

More generally, for a random smooth intersection of hypersurfaces of degrees $d_1,\dots,d_k$ in $\mathbb{P}^n$, the average number of $\mathbb{F}_q$-rational points tends to $q+1$ if $k=1$, but to a limit strictly less than $q+1$ if $k>1$. This can be seen as follows. One would get a limiting average of exactly $q+1$ if the local condition for smoothness at a point x were that the first-order Taylor approximations of $f_1,\dots,f_k$ had to be linearly independent. For $k=1$ (the case in [46]) this is the correct local condition, but for $k>1$, this condition is too restrictive when the sections do not all vanish at x. One possible explanation is that for $k>1$, the intersection of the hypersurfaces can be smooth without being geometrically integral. However, we suspect that this occurs with probability 0 as the d_i tend to infinity, and so does not account for the discrepancy. It will be interesting to understand exactly what is hiding behind this discrepancy.

Theorem 13 (Kurlberg, Wigman). *Fix a prime p. There exists a sequence of families $\{\mathcal{F}_i\}_{i=1}^{\infty}$ of smooth curves defined over $\mathbb{F}_p$ with the following properties such that $\#\mathcal{F}_i\to\infty$, the average number of points*

$$M_i:=\frac{1}{\#\mathcal{F}_i}\sum_{C\in\mathcal{F}_i}\#C(\mathbb{F}_p)\to\infty,$$

the variance

$$V_i:=\frac{1}{\#\mathcal{F}_i}\sum_{C\in\mathcal{F}_i}(\#C(\mathbb{F}_p)-M_i)^2\to infty$$

and, for all compact intervals I,

$$\frac{1}{\#\mathcal{F}_i}\left|\left\{C\in\mathcal{F}_i:\frac{\#C(\mathbb{F}_p)-M_i}{V_i^{1/2}}\in I\right\}\right|=\frac{1}{\sqrt{2\pi}}\int_I e^{-x^2/2}dx+o(1),$$

as $i\to\infty$.

Kurlberg and Wigman used Poonen's sieve to construct a family of curves for which the point count statistics over $\mathbb{F}_p$ becomes Gaussian for *p fixed.* In particular, the average number of $\mathbb{F}_p$-points tends to infinity.

Another generalization of Poonen's sieve is formulated by Erman and Wood [25] and deals with semiample, instead of ample, divisors.

Theorem 14 (Erman, Wood). *Let X be a smooth projective variety over $\mathbb{F}_q$, with a very ample divisor A and a globally generated divisor E. Let π be the map given by the complete linear series on E, $\pi\colon X\xrightarrow{|E|}\mathbb{P}^M$.*

There exists an n_0, depending only on $\dim X$ and $\operatorname{char}(\mathbb{F}_q)$, such that for $n \geq n_0$, the probability of smoothness for a random $D \in |nA + dE|$ as $d \to \infty$ is given by the product of local probabilities taken over the fibers of π:

$$\operatorname{Prob}(D \textit{ is smooth}) = \prod_{P \in \mathbb{P}^M} \operatorname{Prob}(D \textit{ is smooth at all points of } \pi^{-1}(P)).$$

The product on the right converges, is zero only if some factor is zero, and is always non-zero for n sufficiently large.

Note that the case where $A = E$ is the one studied by Poonen in [46]. When E is not very ample, each fiber of π may consist of many points and the fibers may have different dimensions. Singularity at points of a single fiber of π will generally be dependent but the theorem shows that this is the only dependence as $d \to \infty$ and we still get independence between fibers. Erman and Wood use this to compute the distribution of points on certain curves on Hirzebruch surfaces.

Theorem 15 (Erman, Wood)**.** *For fixed $n \geq 3$ and $d \to \infty$, the probability that a curve of bidegree (n, d) in a Hirzebruch surface X is smooth is*

$$\prod_{P \in \mathbb{P}^1_{\mathbb{F}_q}} (1 - q^{-2\deg(P)})(1 - q^{-3\deg(P)}) = \zeta_{\mathbb{P}^1_{\mathbb{F}_q}}(2)^{-1} \zeta_{\mathbb{P}^1_{\mathbb{F}_q}}(3)^{-1} = (1 - q^{-1})(1 - q^{-2})^2(1 - q^{-3}).$$

One can also bring the idea of cohomological stability to bear on questions in arithmetic statistics, as Ellenberg, Venkatesh and Westerland [24] have done to prove some instances of the Cohen-Lenstra heuristics. They formulate the following principle.

Conjecture 16 (Ellenberg, Venkatesh, Westerland)**.** *Assume X_n is an algebraic variety over $\mathbb{F}_q$ of dimension growing with n. Then the quantity $q^{-\dim X_n} \# X_n(\mathbb{F}_q)$ should be expected to approach a limit as $n \to \infty$ precisely when the varieties X_n have stable homology.*

Achter, Erman, Kedlaya, Wood, and Zureick-Brown [1] use this principle to model the average number of points on a curve of genus g. Let M_g denote the fine moduli space of curves of genus g in the sense of Deligne and Mumford [22]; it is an object in the category of algebraic stacks over $\operatorname{Spec}(\mathbb{Z})$. The set $|M_g(\mathbb{F}_q)|$ of (isomorphism classes of) $\mathbb{F}_q$-rational points of M_g may then be identified with the set of isomorphism classes of smooth, projective, geometrically connected curves of genus g over $\mathbb{F}_q$. To simplify notation, let us further identify $|M_g(\mathbb{F}_q)|$ with a set consisting of one curve in each isomorphism class. For $C \in |M_g(\mathbb{F}_q)|$, let $\operatorname{Aut}(C)$ be the group of automorphisms of C as a curve over $\mathbb{F}_q$ (not over an algebraic closure over $\mathbb{F}_q$). We equip $|M_g(\mathbb{F}_q)|$ with the probability measure in which each point x is weighted proportionally to $1/\# \operatorname{Aut}(C)$. This is well-understood to be the most natural way to count objects with automorphisms, and matches the weighting of points in the Lefschetz trace formula for Deligne-Mumford stacks given by Behrend [4]. Assuming cohomological stability for certain moduli spaces of curves, they prove the following conditional result.

Conjecture 17 (Achter, Erman, Kedlaya, Wood, Zureick-Brown)**.** *Put $\lambda := \lambda(q) = q + 1 + 1/(q - 1)$.*

(a) *For all nonnegative integers n,*

$$\lim_{g \to \infty} \operatorname{Prob}(\# C(\mathbb{F}_q) = n : C \in |M_g(\mathbb{F}_q)|) = \frac{\lambda^n e^{-\lambda}}{n!}.$$

(b) *For all positive integers n,*

$$\lim_{g\to\infty} \mathbb{E}(\#C(\mathbb{F}_q)^n : C \in |M_g(\mathbb{F}_q))|) = \sum_{i=1}^{n} \left\{ {n \atop i} \right\} \lambda^i,$$

where $\left\{ {n \atop i} \right\}$ denotes a Stirling number of the second kind (i.e., the number of unordered partitions of $\{1, \ldots, n\}$ into i disjoint sets).

A review of the high genus situation would not be complete without mentioning the work of Drinfel'd and Vlăduţ [23]. For any curve of genus g the Weil bounds imply that

$$q + 1 - 2g\sqrt{q} \leq \#C(\mathbb{F}_q) \leq q + 1 + 2g\sqrt{q}. \tag{18}$$

Now set

$$N_q(g) = \sup\{\#C(\mathbb{F}_q); C \text{ curve of genus } g \text{ over } \mathbb{F}_q\}.$$

The Weil bound (18) implies that

$$N_q(g) \leq q + 1 + 2g\sqrt{q}.$$

Serre [54, Theorem 2.1.1] improved this estimate to

$$N_q(g) \leq q + 1 + g\lfloor 2\sqrt{q} \rfloor.$$

Drinfel'd and Vlăduţ [23] improve on it even further in the large genus limit.

Theorem 19 (Drinfel'd, Vlăduţ)**.**

$$\limsup_{g\to\infty} \frac{N_q(g)}{g} \leq \sqrt{q} - 1.$$

This result has a very different flavor, as it uses linear programming.

4. Arithmetic situation

For E an elliptic curve over a number field K and $\mathfrak{p}$ a prime ideal of K at which E has good reduction, let $a_\mathfrak{p} = a_\mathfrak{p}(E)$ be the Frobenius trace of E at $\mathfrak{p}$, so that $\mathrm{Norm}(\mathfrak{p}) + 1 - a_\mathfrak{p}$ is the number of rational points on the reduction of E modulo $\mathfrak{p}$. Then define the Frobenius angle $\theta_\mathfrak{p} = \theta_\mathfrak{p}(E) \in [0, \pi]$ by the formula

$$1 - a_\mathfrak{p}(E)T + \mathrm{Norm}(\mathfrak{p})T^2 = (1 - \mathrm{Norm}(\mathfrak{p})^{1/2}e^{i\theta}T)(1 - \mathrm{Norm}(\mathfrak{p})^{1/2}e^{-i\theta}T).$$

Let μ_{ST} denote the Sato-Tate measure, so that

$$\mu_{\mathrm{ST}}(f) = \int_0^\pi \frac{2}{\pi} \sin^2\theta f(\theta)\, d\theta.$$

For I an interval, let δ_I denote the characteristic function. The Sato-Tate conjecture, formulated originally for elliptic curves in the 1960s, states the following.

Conjecture 20 (Sato-Tate)**.** *Let E be an elliptic curve over a number field K without complex multiplication. Let N denote the absolute conductor of E. Then for any closed subinterval I of $[0, \pi]$,*

$$\sum_{\mathrm{Norm}(\mathfrak{p})\leq x, \mathfrak{p}\nmid N} \delta_I(\theta_\mathfrak{p}) \sim \mu_{\mathrm{ST}}(I)\,\mathrm{Li}(x).$$

In 1994, Serre [51] gave the ultimate formulation of the Sato-Tate conjecture in terms of motives. We start by recalling the conjectural properties of motivic L-functions following Serre [51]. Fix two number fields K, L. Let $\mathcal{M}$ be a pure motive of weight w over K with coefficients in L. For each prime ideal $\mathfrak{p}$ of K, let $G_\mathfrak{p}$ be a decomposition subgroup of $\mathfrak{p}$ inside the absolute Galois group G_K, let $I_\mathfrak{p}$ be the inertia subgroup of $G_\mathfrak{p}$, and let $\mathrm{Frob}_\mathfrak{p} \in G_\mathfrak{p}/I_\mathfrak{p}$ be the Frobenius element. The *Euler factor* of $\mathcal{M}$ at $\mathfrak{p}$ (for the automorphic normalization) is the function

$$L_\mathfrak{p}(s, \mathcal{M}) = \det(1 - \mathrm{Norm}(\mathfrak{p})^{-s-w/2}\,\mathrm{Frob}_\mathfrak{p}, V_v(\mathcal{M})^{I_\mathfrak{p}} \otimes_{L_v} \mathbb{C})^{-1}$$

for v a finite place of L equipped with an embedding $L_v \hookrightarrow \mathbb{C}$ and $V_v(\mathcal{M})$ the v-adic étale realization of $\mathcal{M}$ equipped with its action of $G_\mathfrak{p}$. It is clear that this definition does not depend on the choice of $G_\mathfrak{p}$; it is conjectured also not to depend on v or the embedding $L_v \hookrightarrow \mathbb{C}$, and this is known when $\mathcal{M}$ has good reduction at $\mathfrak{p}$ (which excludes only finitely many primes).

The ordinary L-function of $\mathcal{M}$ is the Euler product $L(s, \mathcal{M}) = \prod_\mathfrak{p} L_\mathfrak{p}(s, \mathcal{M})$. For each infinite place ∞ of K, there is also an archimedean Euler factor defined as follows. Put

$$\Gamma_\mathbb{R}(s) = \pi^{-s/2}\Gamma(s/2), \qquad \Gamma_\mathbb{C}(s) = 2^{-s}\pi^{-s}\Gamma(s).$$

Form the Betti realization of $\mathcal{M}$ at ∞ and the spaces $H^{p,q}$ for $p + q = w$, and put $h^{p,q} = \dim H^{p,q}$. Note that complex conjugation takes $H^{p,q}$ to $H^{q,p}$ and thus acts on $H^{w/2,w/2}$; let h^+ and h^- be the dimensions of the positive and negative eigenspaces (both taken to be 0 if w is odd). Then put

$$L_\infty(s, \mathcal{M}) = \Gamma_\mathbb{R}(s)^{h^+}\Gamma_\mathbb{R}(s+1)^{h^-} \prod_{p+q=w, p<q} \Gamma_\mathbb{C}(s + w/2 - p)^{h^{p,q}}.$$

The completed L-function is then defined as

$$\Lambda(s, \mathcal{M}) = N^{s/2} L(s, \mathcal{M}) L_\infty(s, \mathcal{M}),$$

for N the absolute conductor of $\mathcal{M}$ (i.e., the norm from K to $\mathbb{Q}$ of the conductor ideal of $\mathcal{M}$).

Conjecture 21 (GRH for motivic L-functions)**.** *Let d be the dimension of the fixed subspace of the motivic Galois group of $\mathcal{M}(-w/2)$ (taken to be* 0 *if w is odd).*

(a) *The function $s^d(1-s)^d\Lambda(s, \mathcal{M})$ (which is defined* a priori *for* $\mathrm{Re}(s) > 1$*) extends to an entire function on $\mathbb{C}$ of order* 1 *which does not vanish at $s = 0, 1$. (Recall that an entire function $f : \mathbb{C} \to \mathbb{C}$ is of* order 1 *if $f(z)e^{-\mu|z|}$ is bounded for each $\mu > 1$.)*

(b) *Let $\mathcal{M}^*$ denote the Cartier dual of $\mathcal{M}$. Then there exists $\epsilon \in \mathbb{C}$ with $|\epsilon| = 1$ such that $\Lambda(1-s, \mathcal{M}) = \epsilon\Lambda(s, \mathcal{M}^*)$ for all $s \in \mathbb{C}$.*

(c) *The zeroes of $\Lambda(s, \mathcal{M})$ all lie on the line* $\mathrm{Re}(s) = 1/2$*.*

In order to state the Sato–Tate conjecture in full generality, a few more definitions are in order. Let A be an abelian variety defined over a number field K of dimension $g \geq 1$ and $\varrho_{A,\ell}$ the ℓ-adic representation attached to A. Let N denote the 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 Hasse-Weil bound asserts that the normalized trace $\overline{a}_\mathfrak{p} := a_\mathfrak{p}(\mathbb{N}(\mathfrak{p}))^{-1/2}$ lies on the interval $[-2g, 2g]$ where $\mathbb{N}(\mathfrak{p})$ is the absolute norm of $\mathfrak{p}$.

Following Serre [53, Chap. 8] 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 $\iota\colon \bar{\mathbb{Q}}_\ell \simeq \mathbb{C}$ and denote by $G_{\ell,\iota}^{1,\mathrm{Zar}}$ the base change $G_\ell^{1,\mathrm{Zar}} \times_{\mathbb{Q}_\ell,\iota} \mathbb{C}$. The Sato–Tate group $\mathrm{ST}(A)$ is a maximal compact subgroup of the group of $\mathbb{C}$-points of $G_{\ell,\iota}^{1,\mathrm{Zar}}$.

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

It follows from the definition that $\mathrm{ST}(A)$ has a faithful unitary symplectic representation $\varrho\colon \mathrm{ST}(A) \to \mathrm{GL}(V)$. Here V a $2g$-dimensional $\mathbb{C}$-vector space, which we call the standard representation of $\mathrm{ST}(A)$. Via this representation, we regard $\mathrm{ST}(A)$ as a compact real Lie subgroup of $\mathrm{USp}(2g)$.

Let μ be the pushforward of the Haar measure of $\mathrm{ST}(A)$ on $[-2g,2g]$ via the trace map. We refer to [53, §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. Note that if $\mathrm{ST}(A)$ happens to be connected, then μ^{disc} is trivial (see [53, §8.4.3.3]).

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

$$\sum_{\mathbb{N}(\mathfrak{p})\le x} \delta_I(\overline{a}_{\mathfrak{p}}) \sim \mu(I)\,\mathrm{Li}(x) \qquad \text{as } x\to\infty\,. \tag{23}$$

Following Theorem 4, one sees that (23) is implied by the conjectural nonvanishing and analyticity on the right halfplane $\mathrm{Re}(s) \ge 1$ of the (normalized) L-function $L(s,\chi)$ for every nontrivial irreducible character χ of $\mathrm{ST}(A)$.

Even though we have such a general formulation of the Sato-Tate conjecture, one has to do a nontrivial amount of work to precisely formulate which groups can appear for abelian varieties of a given dimension. For instance, in the case of elliptic curves (dimension 1), one gets only three possible distributions: μ_{ST} that comes from $SU(2)$ in the case of non-CM elliptic curves, the measure coming from $\mathrm{SO}(2)$ in the case of elliptic curves with complex multiplication over K itself, and the measure coming from the normalizer of $\mathrm{SO}(2)$ inside $\mathrm{SU}(2)$ in the case of elliptic curves with complex multiplication over an extension of K. But for abelian surfaces (dimension 2), Fité, Kedlaya, Rotger and Sutherland [27] showed that there are exactly 52 possibilities in general, 34 of which occur for elliptic curves defined over

$\mathbb{Q}$. In dimension 3, the situation is even more complicated. Fité, Kedlaya, and Sutherland [28] showed that there are 410 possible Sato-Tate groups for abelian three-folds.

The Sato-Tate conjecture is now known unconditionally when K is totally real, thanks to Barnet-Lamb, Gee, Geraghty, Harris, and Taylor [3] and Clozel, Harris, and Taylor [18]. The function field is due to Yoshida [62]. At the heart of the proof of Taylor et al. is the use of Serre's method formulated in [48] to reduce the asymptotic statement to a question of meromorphic continuation and correct analytic properties for all $L(s, \operatorname{Sym}^k E)$, $k \geq 0$, most essentially potential automorphy.

The more refined Lang-Trotter conjecture [39] formulated in 1976 generated interest in refined error terms in the Sato-Tate conjecture. The relevant (though not necessarily most famous) part of the Lang-Trotter conjecture states the following.

Conjecture 24 (Lang, Trotter)**.** *Let E be an elliptic curve defined over $\mathbb{Q}$ without complex multiplication, and let K be an imaginary quadratic extension of $\mathbb{Q}$. For each prime p, denote E_p the reduction of E at p.*

There is an explicit constant $C_{E,K} > 0$ such that, as $X \to \infty$,

$$\#\{p \leq X : p \text{ prime, } E \text{ has good reduction at } p, \operatorname{End}_{\mathbb{F}_p}(E_p) \otimes_{\mathbb{Z}} \mathbb{Q} = K\} \sim C_{E,K} \frac{X^{1/2}}{\log X}.$$

Unfortunately we do not have any power saving error terms in the Sato-Tate conjecture, as the present methods would require us to have a zero-free region inside the critical strip for the L-functions involved. In 1985, V.K. Murty [45] studied the implications of the generalized Riemann Hypothesis to this conjecture by employing Serre's method for a certain family of L-functions.

Theorem 25 (Murty)**.** *Let E be an elliptic curve over a number field K without complex multiplication. Let N denote the absolute conductor of E. Assume that $L(s, \operatorname{Sym}^k E)$ satisfies the generalized Riemann Hypothesis (Conjecture 21) for all $k \geq 0$. Then for any closed subinterval I of $[0, \pi]$,*

$$\sum_{\operatorname{Norm}(\mathfrak{p}) \leq x, \mathfrak{p} \nmid N} \delta_I(\theta_{\mathfrak{p}}) = \mu_{\mathrm{ST}}(I) \operatorname{Li}(x) + O([K : \mathbb{Q}]^{1/2} x^{3/4} (\log(Nx))^{1/2}).$$

Murty's proof is based on a result of Vinogradov [56, Lemma 12] about Fourier series approximations of the characteristic function of an interval. Using a different optimization in Vinogradov's result, Bucur and Kedlaya [15] are able to give a quantitative answer to a classical question about the arithmetic of elliptic curves. Let E_1 and E_2 be nonisogenous elliptic curves over K, neither having complex multiplication. The isogeny theorem of Faltings [26] implies that there exists a prime ideal $\mathfrak{p}$ of K at which E_1, E_2 both have good reduction and have distinct Frobenius traces. In particular, for any fixed prime ℓ, there exists a prime ideal $\mathfrak{p}$ of K at which the Frobenius traces of E_1, E_2 differ modulo ℓ. Assuming the generalized Riemann hypothesis (Conjecture 21) for Artin L-functions, one can use the effective form of the Chebotarev density theorem (as suggested by Serre in [49]) to show the least norm of such a prime ideal is

$$O((\log N)^2 (\log\log 2N)^b)$$

for some fixed $b \geq 0$. Assuming the generalized Riemann Hypothesis for L-functions of the form $L(s, \operatorname{Sym}^m E_1 \otimes \operatorname{Sym}^n E_2)$, Bucur and Kedlaya use the effective form of the generalized

Sato-Tate conjecture for the abelian surface $E_1 \times_K E_2$ to obtain a similar bound for the least norm of a prime ideal at which the Frobenius traces of E_1, E_2 have opposite sign (Theorem 26). In both cases, the optimal bound is most likely closer to $O(\log N)$, but by analogy with the problem of finding the least quadratic nonresidue modulo N, it is unlikely that one can do better than $O((\log N)^2)$ using L-function methods.

Theorem 26 (Bucur, Kedlaya). *Let E_1, E_2 be two $\overline{\mathbb{Q}}$-nonisogenous elliptic curves over a number field K, neither having complex multiplication. Let N be the product of the absolute conductors of E_1 and E_2. For each prime ideal $\mathfrak{p}$ of K not dividing N, let $\theta_{1,\mathfrak{p}}, \theta_{2,\mathfrak{p}}$ be the Frobenius angles of E_1, E_2 at $\mathfrak{p}$. Assume that the L-functions $L(s, \operatorname{Sym}^i E_1 \otimes \operatorname{Sym}^j E_2)$ for $i, j = 0, 1, \dots$ all satisfy the generalized Riemann Hypothesis (Conjecture 21). Then for any closed subintervals I_1, I_2 of $[0, \pi]$,*

$$\sum_{\operatorname{Norm}(\mathfrak{p}) \le x, \mathfrak{p} \nmid N} \delta_{I_1}(\theta_{1,\mathfrak{p}})\delta_{I_2}(\theta_{2,\mathfrak{p}}) = \mu_{\mathrm{ST}}(I_1)\mu_{\mathrm{ST}}(I_2)\operatorname{Li}(x) + O([K:\mathbb{Q}]^{1/3}x^{5/6}(\log(Nx))^{1/3}).$$

The following estimate follows immediately,

Theorem 27 (Bucur, Kedlaya). *With hypotheses and notation as in Theorem 26, there exists a prime ideal $\mathfrak{p}$ not dividing N with $\operatorname{Norm}(\mathfrak{p}) = O([K:\mathbb{Q}]^2(\log N)^2(\log\log 2N)^2)$ such that $a_{\mathfrak{p}}(E_1)$ and $a_{\mathfrak{p}}(E_2)$ are nonzero and of opposite sign.*

Remark 28. Theorem 27, which distinguishes two Frobenius traces using their archimedean behavior, should be compared with similar results which distinguish the traces using their mod-ℓ behavior for some prime ℓ. For example, in [49, §8.3, Théorème 21], Serre shows that there exists a prime ideal $\mathfrak{p}$ not dividing N with $\operatorname{Norm}(\mathfrak{p}) = O((\log N)^2(\log\log 2N)^{12})$ such that $a_{\mathfrak{p}}(E_1)$ and $a_{\mathfrak{p}}(E_2)$ differ modulo some auxiliary prime ℓ.

Both this argument and Theorem 27 give upper bounds on the norm of a prime ideal $\mathfrak{p}$ for which $a_{\mathfrak{p}}(E_1)$ and $a_{\mathfrak{p}}(E_2)$ differ. However, Serre has subsequently remarked [50, p. 715, note 632.6] that by replacing the mod-ℓ argument with an ℓ-adic argument, one can improve these bounds to $O((\log N)^2)$.

Theorem 29 (Serre). *Let Γ be a group, let ℓ be a prime number, let r be a positive integer, and let $\rho_1, \rho_2 : \Gamma \to \mathrm{GL}_r(\mathbb{Z}_\ell)$ be two homomorphisms with distinct traces. Then there exist a finite quotient G of Γ and a nonempty subset C of G with the following properties.*

(a) *The order of G is at most $\ell^{2r^2} - 1$.*

(b) *For any $\gamma \in \Gamma$ whose image in G belongs to C, $\operatorname{Trace}(\rho_1(\gamma)) \neq \operatorname{Trace}(\rho_2(\gamma))$.*

Corollary 30 (Serre). *Assume the Riemann hypothesis (Conjecture 21) for Artin L-functions. Then there exists a prime ideal $\mathfrak{p}$ not dividing N with $\operatorname{Norm}(\mathfrak{p}) = O((\log N)^2)$ such that $a_{\mathfrak{p}}(E_1) \neq a_{\mathfrak{p}}(E_2)$.*

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 [31] 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 [47] carried out the analogous analysis with the Fourier coefficients replaced by normalized Hecke eigenvalues (this only makes a difference when the weights are distinct). The analogue of Serre's argument for modular forms was given by R. Murty [44] and subsequently extended to Siegel modular forms by Ghitza [29] for Fourier coefficients and Ghitza and Sayer [30] for Hecke eigenvalues.

Bucur, Fité and Kedlaya [13] proved a similar result for abelian varieties, not just elliptic curves.

Theorem 31 (Bucur, Fité, Kedlaya)**.** *Let A be an abelian variety defined over the number field K of dimension $g \geq 1$, absolute conductor N, and such that* $\mathrm{ST}(A)$ *is connected. Suppose that the Mumford–Tate conjecture (Conjecture 22) holds for A and that the generalized Riemann hypothesis (Conjecture 21) holds for $L(s,\chi)$ for every irreducible character χ of* $\mathrm{ST}(A)$*. Then there exists an $\varepsilon > 0$ such that for all subintervals I of $[-2g, 2g]$, we have*

$$\sum_{\mathrm{Norm}(\mathfrak{p})\leq x} \delta_I(\overline{a}_{\mathfrak{p}}) = \mu(I)\,\mathrm{Li}(x) + O\left(\frac{x^{1-\varepsilon}\log(Nx)^{2\varepsilon}}{\log(x)^{1-4\varepsilon}}\right), \tag{32}$$

where the sum runs over primes not dividing N and the implicit constant in the O-notation depends exclusively on K and g. Moreover, if we write the Lie algebra of $\mathrm{ST}(A)$ *as $\mathfrak{g} = \mathfrak{s} \times \mathfrak{a}$, where $\mathfrak{s}$ is semisimple and $\mathfrak{a}$ is abelian, then we can take ε as*

$$\varepsilon(\mathfrak{g}) := \frac{1}{2(q + |\Phi^+|)},$$

where $|\Phi^+|$ is the size of the set of positive roots of $\mathfrak{s}$ and q is the rank of $\mathfrak{g}$.

A key ingredient in this work is the construction of a multivariate Vinogradov function; this is a smooth 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 [32, Thm. 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 V.K. Murty (as presented by the Bucur and Kedlaya [15]) on truncated sums of an irreducible character χ over the prime ideals of K. The implicit 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, which is an interesting result in itself.

An interesting application of this result is a conditional partial answer to a question posed by Serre [54, Chap. II, Question 6.7] about elliptic curves with maximal number of points (that realize the Weil bound). Let $M(x)$ denote the number of $\mathfrak{p}$ not dividing N with norm up to x for which $a_{\mathfrak{p}} = \lfloor 2\sqrt{\mathrm{Norm}(\mathfrak{p})}\rfloor$, which would ensure that the number of points on the reduction of E at $\mathfrak{p}$ attains the maximum possible value. Vaguely formulated, a natural approach to compute (at least an asymptotic lower bound on) $M(x)$ 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 (32) to be still dominated by the "main term", which is now multiplied by the tiny quantity $\mu(I_x)$. This trade-off can be achieved when E is an elliptic curve with CM.

Corollary 33 (Bucur, Fité, Kedlaya)**.** *Let E be an elliptic curve defined over K with potential CM, that is, such that $E_{\overline{\mathbb{Q}}}$ has CM. Under the generalized Riemann hypothesis (Conjecture 21) for the L-function attached to every power of the Hecke character of E, we have*

$$M(x) \asymp \frac{x^{3/4}}{\log(x)} \qquad \text{as } x \to \infty\,.$$

The notation $f \asymp g$ for two functions f and g means that $f = O(g)$ and $g = O(f)$. Thus the corollary also incorporates an upper bound, which requires a more elaborate argument.

Interestingly, James and Pollack [33] proved *unconditionally* the result conjectured by Serre using different methods from analytic number theory.

Theorem 34 (James, Pollack)**.** *Let E be an elliptic curve defined over K with potential CM, that is, such that $E_{\overline{\mathbb{Q}}}$ has CM. Then*

$$M(x) \sim \frac{2}{3\pi}\frac{x^{3/4}}{\log(x)} \qquad \text{as } x \to \infty\,.$$

Theorem 35 (Bucur, Fité, Kedlaya)**.** *Let A (resp. A') be an abelian variety defined over K of dimension g (resp. g'), and with absolute conductor N (resp. N'). Suppose that the generalized Riemann hypothesis (Conjecture 21) holds for every irreducible constituent of a virtual character ψ of the Sato-Tate group* $\mathrm{ST}(A \times A')$ *and that* $\mathrm{Hom}(A, A') = 0$. *Then there exists a prime $\mathfrak{p}$ not dividing NN' with norm*

$$\mathrm{Norm}(\mathfrak{p}) = O(\log(NN')^2)$$

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

Note that this result extends Theorem 27 of Bucur and Kedlaya. Later, that same result was improved by Chen, Park, and Swaminathan [16, Thm. 1.3] who proved an upper bound of the form $O(\log(NN')^2)$ for A and A' two nonisogenous elliptic curves without CM and relaxing the generalized Riemann hypothesis to only a handful of symmetric powers of the two elliptic curves.

There are two aspects of the approach of Chen, Park, and Swaminathan in [16, Thm. 1.3] which we would like to highlight. On the one hand, their method avoids the use of the effective Sato–Tate conjecture. They replace V.K. Murty's estimate with one they obtain by integrating with respect to a kernel introduced by Bach [2]. While V.K. Murty's estimate seems to be most adequate to treat density questions, for existence problems Bach's estimate seems to provide more accurate answers. On the other hand, the condition $a_{\mathfrak{p}}(A) \cdot a_{\mathfrak{p}}(A') < 0$ is recast as the positivity of a certain polynomial expression in $\overline{a}_{\mathfrak{p}}(A)$ and $\overline{a}_{\mathfrak{p}}(A')$ that (conveniently weighted) is shown via Bach's estimate to become eventually positive when summed over primes in K.

Bucur, Fité and Kedlaya generalized the approach of [16] to obtain a version of Bach's estimate valid for general abelian varieties by reinterpreting the polynomial expression in $\overline{a}_{\mathfrak{p}}(A)$ and $\overline{a}_{\mathfrak{p}}(A')$ alluded to above as the character ψ of the natural virtual representation

$$\mathrm{ST}(A \times A') \to \mathrm{GL}((V^{\oplus -2g} \oplus V \otimes V) \otimes ((V')^{\oplus 2g'} \oplus V' \otimes V'))\,,$$

where V and V' denote the standard representations of $\mathrm{ST}(A)$ and $\mathrm{ST}(A')$. It is easy to see that if $\mathrm{Hom}(A, A') = 0$, then the multiplicity of the trivial character in ψ is *strictly positive*, which explains the eventual positivity of ψ when summed over primes in K.

5. Future directions

There are many open problems and directions that still need exploring. The biggest problem in the field is to find a way of proving equidistribution results over number fields, either by adapting the methods that were successful in function fields or by deducing them from the function fields results. The main problem with adapting the function fields method to number fields is the fact that one has a lot of geometry in the function fields picture (e.g. the geometric Fourier transform over finite fields) that is missing from the number fields world. But there are many more attainable goals.

For instance, one can try to better understand $N_q(g)$ for small g as done in part I of [54], especially in the postfaces to Chapters 2 and 4 by Howe and Ritzenthaler. One can ask in general for better bounds on $N_q(g)$, and even computational evidence for what $N_q(g)$ should be in all cases. As the Drinfel'd -Vlăduţ [23] results shows, as g grows we get rather far away from the Weil bound. It would be interesting to find out what the $\limsup N_q(g)/g$ really is, and for which genera it can actually be attained, or how close it can come. In a certain sense, the quantity $N_q(g)$ is related to lower order terms in the distribution of $\#C(\mathbb{F}_q)$ as C varies over the curves of genus g. Which brings us to another direction to explore, namely to compute moments of the distribution of $\#C(\mathbb{F}_q)$ as done by Birch [5] for $g = 1$, but this time in the fixed q, large g situation.

As formulated by Ritzenthaler during 2021 conference dedicated to [54], it would be nice to gather heuristic arguments for or against the possibility of always having a maximal curve of genus g over $\mathbb{F}_q$ at bounded "distance" of the Hasse-Weil-Serre bound (independently of q for a fixed g). Another problem is to find exactly for which g and q one can find a defect zero curve, for which the Hasse-Weil-Serre bound is sharp. Clearly the number of g's for which one can find any such curve (over any finite field) is finite, but can one find the exact values? And can one find an estimate for the distance between $N_q(g)$ and the Hasse-Weil-Serre bound?

Also regarding the number of points on varieties over finite fields, one knows from work of Madan and Pal [41] that there exist infinitely many simple abelian varieties defined over $\mathbb{F}_2$ with exactly one $\mathbb{F}_2$-point. Recent work of Kedlaya [36] shows that for every positive integer m, there exist infinitely many simple abelian varieties over $\mathbb{F}_2$ of order m. One can study a similar question of simple abelian varieties of prescribed order over any finite field $\mathbb{F}_q$. For $q \geq 5$ the Weil estimate tells us that, in general, for a given m there are at most finitely many abelian varieties of order m. For $q = 3, 4$ we know that the same is true if one restricts to simple abelian varieties. (See for instance [34].) We also know from work of van Bommel, Costa, Li, Poonen, and Smith [6] that for any given q every sufficiently large m is the order of some simple abelian variety defined over $\mathbb{F}_q$. The question becomes how many and an asymptotic would be highly interesting. Dealing with prime fields might be a good first step.

Another clear direction is to obtain statistics for the number of points G-covers of the projective line over $\mathbb{F}_q$ for all groups G. This would be a version of Malle's conjecture [42, 43] in this context. As we have seen in Section 3, we have results for some groups, e.g.

$\mathbb{Z}/\ell\mathbb{Z}$, $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, S_3. One option is to use Wright's approach [61] and the results known so far to settle this question for all finite abelian groups. A more difficult problem is to compute the distribution of the number of points for some nonabelian groups beyond S_3. A first example would be D_4 as this has already been studied in the number field context as in [19]. A hopeful development in the number field comes from the recent work of Wood that formulates a nonabelian version of the Cohen-Lenstra heuristics. See [60] for a really nice exposition and some results in this direction over function fields. One other problem is to study G-covers of a fixed arbitrary curve over $\mathbb{F}_q$, not just the projective line.

If one starts with a curve C defined over $\mathbb{F}_q$, one can look at the multiple (or even all) point counts $\#C(\mathbb{F}_{q^m})$ for $m > 0$. Of course these are not independent of each other. Take as an example hyperelliptic curves. As shown by Kurlberg and Rudnick [37] (see also [10]) the number of $\mathbb{F}_q$-points of such a curve is given by

$$\#C(\mathbb{F}_q) = q + 1 + S(\chi)$$

where χ is the quadratic character associated to the double cover C of $\mathbb{P}^1$. Say for instance that the curve C has an affine model

$$y^2 = f(x).$$

Then

$$S(\chi) = \sum_{x \in \mathbb{P}^1(\mathbb{F}_q)} \chi(f(x))$$

with

$$\chi(\alpha) = \begin{cases} 1 & \alpha \in (\mathbb{F}_q^\times)^2 \\ 0 & \alpha = 0 \\ -1 & \text{otherwise.} \end{cases}$$

But when one looks at $\mathbb{F}_{q^2}$-points of C one finds that all fibers above the $\mathbb{F}_q$-points of $\mathbb{P}^1$ have 2 points, while half the fibers above the degree 2 points of $\mathbb{P}^1$ will have 2 points and half of them no points. Hence once gets, on average, more points than if one looked at all the curves defined over $\mathbb{F}_{q^2}$. This phenomenon is related to the observation made by Brock and Granville in [8].

It would be interesting to study the joint distribution of the sequence $\{\#C(\mathbb{F}_{q^m})\}_{m\geq 1}$. It is unclear what the model for these sequences even is. We know that each term in the sequence is related to traces of certain random matrices, but then one has the correlations between the $\#C(\mathbb{F}_{q^m})$ and $\#C(\mathbb{F}_{q^{mn}})$ to take into account, as well as the fact that all these point counts should be nonnegative integers. Hence once get infinitely many discrete conditions that cut into the space of random matrices.

When it comes to the Sato-Tate conjecture, one possible goal is a proof of the Sato-Tate for abelian varieties. Serre's Theorem 4 and Theorem 5 imply that in order to prove Sato-Tate in this context one needs the meromorphic continuation and analytic properties for all irreducible rep of Sp(4). To give an idea of the magnitude of the task, just doing this for the trivial representation amounts to proving the paramodular conjecture [9].

Conjecture 36 (Brumer, Kramer). *Let A be an abelian surface over $\mathbb{Q}$ of conductor N such that $\mathrm{End}_{\mathbb{Q}}(A) = \mathbb{Z}$. Then there exists a cuspidal Siegel paramodular form $f \in S_2^{(2)}\left(\Gamma^{\text{para}}(N)\right)$, a newform and eigenform of degree 2, weight 2, and level N, such that $L(A, s) = L(f, s)$.*

Another interesting question is to see if the Sato-Tate distributions are different if one restricts to Jacobians of curves of genus g instead of looking at all abelian varieties of dimension g. We know that $g \geq 4$ the two spaces are different, but is there a reason for the distributions to be the same or be different? Even a conjecture in this direction would be interesting. A problem whose magnitude is only hinted at by [28] is to find all the possible Sato-Tate groups for abelian varieties of dimension $d \geq 4$, or at least count them. Since for $g = 1$ we get 3 groups, for $g = 2$ we get 52 possibilities and for $g = 3$ we have 410 possibilities, one suspects this might not be something to be done by hand. One can simplify the problem and ask only for the connected components of the identity of these groups in each case. We do know that the numbers in this case are 2 for $g = 1$, 6 for $g = 2$, and 14 for $g = 3$ as seen in [27, 28]. An estimate for how many possibilities there are for higher g would already constitute huge progress.

One can use the Sato-Tate conjecture to deduce various statistics for the traces of Frobenius. For instance, I am indebted to the anonymous referee to this paper for suggesting the following problem. Let C be a a smooth and projective curve defined over $\mathbb{Q}$ of genus $g > 0$. Then one can ask if there exits a prime p such that the Frobenius trace a_p is strictly positive. In genus $g = 1$, the Sato-Tate conjecture implies that in fact there is a positive density of such primes. By Section 1, the holomorphicity and nonvanishing at $\Re(s) = 1$ of the Hasse-Weil L-function $L(C, s)$ would imply the existence of infinitely many primes p for which $a_p > 0$. (Here we follow Serre's convention from Theorems 4 and 5 and use the analytic normalization for L-functions that sets the rightmost pole at $s = 1$ and such that functional equation sends $s \to 1 - s$.) This is known for $g = 2$ thanks to recent work by Boxer–Calegari–Gee–Pilloni [7]. It is provocative that for genus at least 3 we do not even know that a single prime p with $a_p > 0$ exists.

Since this paper has to end at some point, we will mention just one more problem, namely that of ordinary primes. Take an abelian variety X defined over $\mathbb{Q}$ and reduce it modulo p. We say that a prime p is ordinary (for X) if the reduction X_p of X modulo p is ordinary. One would like to show that for each X there are infinitely many ordinary primes. A harder problem would be to find the density of ordinary primes for a given X.

References

[1] J.D. Achter, D. Erman, K.S. Kedlaya, M.M. Wood, and D. Zureick-Brown, *A heuristic for the distribution of point counts for random curves over finite field*, Philos. Trans. Roy. Soc. A 373 (2015), no. 2040, 20140310, 12 pp.

[2] E. Bach, *Explicit bounds for primality testing and related problems*, Math. comp. 55, no. 191 (1990), 355–380.

[3] T. Barnet-Lamb, T. Gee, D. Geraghty, M. Harris, and R. Taylor, *A family of Calabi-Yau varieties and potential automorphy II* Publ. Res. Inst. Math. Sci. 47 (2011), no. 1, 29–98.

[4] K.A. Behrend, *The Lefschetz trace formula for algebraic stacks*, Invent. Math., 112(1), 1993, 127–149.

[5] B.J. Birch, *How the number of points of an elliptic curve over a fixed prime field varies*, J. London Math. Soc. 43 (1968), 57–60.

[6] R. van Bommel, E. Costa, W. Li, B. Poonen, and A. Smith, *Abelian varieties of prescribed order over finite fields*, preprint, arXiv:2106.13651.

[7] G. Boxer, F. Calegari, T. Gee, and V. Pilloni, *Abelian surfaces over totally real fields are potentially modular*, Publ. Math., Inst. Hautes Étud. Sci. 134, 153–501 (2021).

[8] B.W. Brock and A. Granville *More Points Than Expected on Curves over Finite Field Extensions*, Finite Fields and Their Applications, 7(1) 2001, 70–91.

[9] A. Brumer and K. Kramer,*Paramodular abelian varieties of odd conductor*, Trans. Amer. Math. Soc. 366.5 (2014), 2463–2516.

[10] A. Bucur, C. David, B. Feigon, and M. Lalín, *Statistics for traces of cyclic trigonal curves over finite fields*, Int. Math. Res. Not., 2010(5), 2010, 932–967.

[11] A. Bucur, C. David, B. Feigon, and M. Lalín, *Fluctuations in the number of points on smooth plane curves over finite fields*, J. Number Theory 130 (2010), 2528–2541.

[12] A. Bucur, C. David, B. Feigon, N. Kaplan, M. Lalín, E. Ozman, and M.M. Wood, *The distribution of $\mathbb{F}_q$-points on cyclic ℓ-covers of genus g*, Int. Math. Res. Not. IMRN 2016, no. 14, 4297–4340.

[13] A. Bucur, F. Fité, and K.S. Kedlaya, *Effective Sato-Tate conjecture for abelian varieties and applications*, preprint, arXiv:2002.08807.

[14] A. Bucur and K.S. Kedlaya, *The probability that a complete intersection is smooth*, J. Théor. Nombres Bordeaux 24 (2012), no. 3, 541–556.

[15] A. Bucur and K.S. Kedlaya, *An application of the effective Sato-Tate conjecture*, Frobenius distributions: Lang-Trotter and Sato-Tate conjectures, 45–56, Contemp. Math., 663, Amer. Math. Soc., Providence, RI, 2016.

[16] E. Chen, P.S. Park, A.A. Swaminathan, *Elliptic curve variants of the least quadratic non-residue problem and Linnik's theorem*, Int. J. of Number Theory, Vol. 14, No. 1 (2018), 255–288.

[17] G.Y. Cheong, M.M. Wood, and A. Zaman, *The distribution of points on superelliptic curves over finite fields* Proc. Amer. Math. Soc. 143 (2015), no. 4, 1365–1375.

[18] L. Clozel, M. Harris, and R. Taylor, *Automorphy for some ℓ-adic lifts of automorphic mod ℓ Galois representations*, with Appendix A, summarizing unpublished work of R. Mann, and Appendix B by M.-F. Vignéras. Publ. Math. Inst. Hautes Études Sci. No. 108 (2008), 1–181.

[19] H. Cohen, F.Diaz y Diaz, and M.Olivier, *Enumerating quartic dihedral extensions of $\mathbb{Q}$*. Compositio Math. 133.1 (2002), pp. 65–93.

[20] B. Datskovsky and D.J. Wright, *The adelic zeta function associated to the space of binary cubic forms. II. Local theory*, J. Reine Angew. Math. 367 (1986), 27–75.

[21] P. Deligne, *La conjecture de Weil. II*, Inst. Hautes Études Sci. Publ. Math. No. 52 (1980), 137–252.

[22] P.Deligne and D. Mumford, *The irreducibility of the space of curves of given genus*, Inst. Hautes Études Sci. Publ. Math., (36), 1969, 75–109.

[23] V.G. Drinfel'd and S.G. Vlăduţ , *The number of points of an algebraic curve* (Russian) Funktsional. Anal. i Prilozhen. 17 (1983), no. 1, 68–69.

[24] J.S. Ellenberg, A. Venkatesh, and C. Westerland, *Homological stability for Hurwitz spaces and the Cohen-Lenstra conjecture over function fields*, Ann. of Math. (2) 183 (2016), no. 3, 729–786.

[25] D. Erman and M.M. Wood, *Semiample Bertini theorems over finite fields*, Duke Math. J. 164 (2015), no. 1, 1–38.

[26] G. Faltings, *Endlichkeitssätze für abelsche Varietäten über Zahlkörpern*, Invent. Math. 73 (1983), 349–366.

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

[28] F. Fité, K.S. Kedlaya, and A.V. Sutherland, *Sato-Tate groups of abelian threefolds*, preprint, arXiv:2106.13759.

[29] A. Ghitza, *Distinguishing Hecke eigenforms*, Int. J. Num. Theory, 7 (2011), 1247–1253.

[30] A. Ghitza and R. Sayer, *Hecke eigenvalues of Siegel modular forms of "different weights"*, J. Num. Theory, **143** (2014), 125–141.

[31] D. Goldfeld and J. Hoffstein, *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**, Amer. Math. Soc., Providence, 1993, 385–393.

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

[33] K. James and P. Pollack, *Extremal primes for elliptic curves with complex multiplication*, J. Number Theory 172 (2017), 383–391.

[34] B. Kadets, *Estimates for the number of rational points on simple abelian varieties over finite fields*, Math. Z. 297 (2021) 465–473.

[35] N.M. Katz and P. Sarnak, *Random matrices, Frobenius eigenvalues, and monodromy* American Mathematical Society Colloquium Publications, 45. American Mathematical Society, Providence, RI, 1999.

[36] K.S. Kedlaya, *Abelian varieties over $\mathbb{F}_2$ of prescribed order*, to appear in Publications Mathématiques de Besançon (special issue for GTA 2021), arXiv:2107.12453.

[37] P. Kurlberg and Z. Rudnick, *The fluctuations in the number of points on a hyperelliptic curve over a finite field*, J. Number Theory, 129(3), 2009, 580–587.

[38] P. Kurlberg and I. Wigman, *Gaussian point count statistics for families of curves over a fixed finite field*, Int. Math. Res. Not. 2011, No. 10, 2217–2229.

[39] S. Lang and H. Trotter, *Frobenius distributions in* GL_2*-extensions*, Springer-Verlag, Berlin, 1976. Distribution of Frobenius automorphisms in GL_2-extensions of the rational numbers; Lecture Notes in Mathematics, Vol. 504, Springer-Verlag, Berlin-New York, 1976.

[40] E. Lorenzo García, G. Meleleo, and P. Milione, *Statistics for biquadratic covers of the projective line over finite fields* (with an appendix by A. Bucur), J. Number Theory, 173 (2017), 448–477.

[41] M.L. Madan and S. Pal, *Abelian varieties and a conjecture of R. M. Robinson*, J. Reine Angew. Math. 291 (1977), 78–91.

[42] G. Malle, *On the distribution of Galois groups*, J. Number Theory, 92(2), 2002, 315–329.

[43] G. Malle, *On the distribution of Galois groups, II*, Experimental Mathematics, 13(2), 2004, 129–135.

[44] M.R. Murty, *Congruences between modular forms*, in *Analytic Number Theory*, London Math. Soc. Lecture Note Series 247, Cambridge Univ. Press, Cambridge, 1997, 309–320.

[45] V. K. Murty, *Explicit formulae and the Lang-Trotter conjecture*, Rocky Mountain J. Math. 15 (1985), no. 2, 535–551.

[46] B. Poonen, *Bertini theorems over finite fields*, Ann. of Math. (2) 160 (2004), no. 3, 1099–1127.

[47] J. Sengupta, *Distinguishing Hecke eigenvalues of primitive cusp forms*, Acta Arithmetica 114 (2004), 23–34.

[48] J.-P. Serre, *Abelian ℓ-adic Representations and Elliptic Curves*, W.A. Benjamin Inc., 1968.

[49] J.-P. Serre, Quelques applications du théorème de densité de Chebotarev, *Publ. Math. IHÉS* **54** (1981), 123–201.

[50] J.-P. Serre, *Œuvres, Vol. III. 1972–1984*, Springer-Verlag, Berlin, 1986.

[51] J.-P. Serre, *Propriétés conjecturales des groupes de Galois motiviques et des représentations l-adiques*, in *Motives (Seattle, WA, 1991)*, Proceedings of Symposia in Pure Math. **55**, Amer. Math. Soc., 1994, 377–400.

[52] J.-P. Serre, *Répartition asymptotique des valeurs propres de l'opérateur de Hecke T_p*, J. Amer. Math. Soc. 10 (1997), no. 1, 75–102.

[53] J.-P. Serre, Lectures on $N_X(p)$, A.K. Peters, 2012.

[54] J.-P. Serre, *Rational points on curves over finite fields.* With contributions by E. Howe, J. Oesterlé and C. Ritzenthaler. Edited by A. Bassa, E. Lorenzo García, C. Ritzenthaler and R. Schoof. Documents Mathématiques 18. Paris. Société Mathématique de France (SMF), 187 pp. (2020).

[55] S. Türkelli, *Connected components of Hurwitz schemes and Malle's conjecture*, J. Number Theory 155 (2015), 163–201.

[56] I.M. Vinogradov, *The method of trigonometrical sums in the theory of numbers*, reprint of the 1954 translation. Dover Publications, Inc., Mineola, NY, 2004. x+180 pp.

[57] A. Weil, *Sur les courbes algébriques et les variétés qui s'en déduisent*, Publications del'Institut de Mathématique de l'Université de Strasbourg 7. Paris: Hermann et Cie., 1948.

[58] A. Weil, *Numbers of solutions of equations in finite fields*, Bull. Amer. Math. Soc. 55 (1949), 497–508.

[59] M.M. Wood, *The distribution of the number of points on trigonal curves over $\mathbb{F}_q$*, Int. Math. Res. Not. IMRN, (23), 2012, 5444–5456.

[60] M.M. Wood and P. M. Wood, *Nonabelian Cohen-Lenstra Moments*, Duke Math. J. 168, no. 3 (2019), 377–427.

[61] D.J. Wright, *Distribution of discriminants of abelian extensions*, Proc. London Math. Soc. (3) 58 (1989), no. 1, 17–50.

[62] H. Yoshida, *On an analogue of the Sato conjecture*, Invent. Math. 19 (1973), 261–277.

Department of Mathematics, University of California at San Diego, 9500 Gilman Dr #0112, La Jolla, CA 92093
E-mail address: alina@math.ucsd.edu