ARIZONA WINTER SCHOOL 2016
SATO-TATE DISTRIBUTIONS

ANDREW V. SUTHERLAND

1. An introduction to Sato-Tate distributions

Before discussing Sato-Tate conjecture and Sato-Tate distributions in general, let us start in the more familiar setting of Artin motives (otherwise known as the Galois theory of number fields).

1.1. **A first example.** Let $f \in \mathbf{Z}[x]$ be a squarefree polynomial of degree d; for example, we may take $f(x) = x^3 - x + 1$. Since f has integer coefficients, we can reduce them modulo any prime p to obtain a polynomial f_p with coefficients in the finite field $\mathbf{Z}/p\mathbf{Z} \simeq \mathbf{F}_p$. For each prime p define

$$N_f(p) := \#\{x \in \mathbf{F}_p : f_p(x) = 0\},$$

which we note is an integer between 0 and d. We would like to understand how $N_f(p)$ varies with p. The table below shows the values of $N_f(p)$ when $f(x) = x^3 - x + 1$ for $p < 60$:

p:	2	3	5	7	11	13	17	19	23	29	31	37	41	43	47	53	59
$N_f(p)$	0	0	1	1	1	0	1	1	2	0	0	1	0	1	0	1	3

There does not appear to be any obvious pattern (and we should know not to expect one, the Galois group lurking behind the scenes is nonabelian). The prime $p = 23$ is exceptional because it divides $\mathrm{disc}(f)$, which means that $f_p(x)$ has a double root. As we are interested in the distribution of $N_f(p)$ as p tends to infinity, we are happy to ignore such primes, which are necessarily finite in number.

Looking at such a small dataset does not tell us much, so let us increase the bound B on the primes p that we are considering and count how often we see $N_f(p) = 0, 1, 2, 3$. Define

$$c_i(B) := \frac{\#\{p \le B : N_f(p) = i\}}{\#\{p \le B\}},$$

for $i = 0, 1, 2, 3$. We may then compute the following statistics:

B	$c_0(B)$	$c_1(B)$	$c_2(B)$	$c_3(B)$
10^3	0.323353	0.520958	0.005988	0.155689
10^4	0.331433	0.510586	0.000814	0.157980
10^5	0.333646	0.502867	0.000104	0.163487
10^6	0.333185	0.500783	0.000013	0.166032
10^7	0.333360	0.500266	0.000002	0.166373
10^8	0.333337	0.500058	0.000000	0.166605
10^9	0.333328	0.500016	0.000000	0.166656
10^{10}	0.333334	0.500003	0.000000	0.166663
10^{11}	0.333333	0.500001	0.000000	0.166666
10^{12}	0.333333	0.500000	0.000000	0.166666

Based on these statistics we may conjecture that the limiting values of $c_i(B)$ as $B \to \infty$ are

$$c_0 = 1/3, \quad c_1 = 1/2, \quad c_2 = 0, \quad c_3 = 1/6.$$

There is of course a natural motivation for this conjecture (which is in fact a theorem), that is completely independent of the statistics we just computed. Let us fix an algebraic closure $\overline{\mathbf{Q}}$ of $\mathbf{Q}$. The absolute Galois group $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ acts on the roots of $f(x)$ by permuting them. This allows us to define a *Galois representation* (continuous homomorphism)

$$\rho_f : \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_d(\mathbf{C}),$$

whose image is a subgroup of the permutation matrices in $\mathrm{O}_d(\mathbf{C}) \subseteq \mathrm{GL}_d(\mathbf{C})$; here O_d denotes the orthogonal group (we could replace $\mathbf{C}$ with any field of characteristic zero). Note that $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ and $\mathrm{GL}_d(\mathbf{C})$ are topological groups (the former has the Krull topology), and homomorphisms of topological groups are understood to be continuous. In order to associate a permutation of the roots of $f(x)$ to a matrix in $\mathrm{GL}_d(\mathbf{C})$ we need to fix an ordering of the roots; this amounts to choosing a basis for the vector space $\mathbf{C}^d$, which means that our representation ρ_f is really defined only up to conjugacy.

The value ρ_f takes on $\sigma \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ depends only on the restriction of σ to the splitting field L of f, so we could restrict our attention to $\mathrm{Gal}(L/\mathbf{Q})$. This makes ρ_f an *Artin representation*: a continuous representation $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_d(\mathbf{C})$ that factors through a finite quotient (by an open subgroup). But in the more general settings we wish to consider we may not know what L is (or even its degree), so it is better to work with $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$.

To facilitate this, we associate to each prime p an *absolute Frobenius element*

$$\mathrm{Frob}_p \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$$

which may be defined as follows. Fix an embedding $\overline{\mathbf{Q}}$ in $\overline{\mathbf{Q}}_p$ and use the valuation ideal $\mathfrak{P}$ of $\overline{\mathbf{Q}}_p$ (the maximal ideal of its ring of integers) to define a compatible system of primes $\mathfrak{q}_L := \mathfrak{P} \cap L$, where L ranges over all finite extensions of $\mathbf{Q}$. For each prime $\mathfrak{q}_L$, let $D_{\mathfrak{q}_L} \subseteq \mathrm{Gal}(L/K)$, denote its decomposition group, $I_{\mathfrak{q}_L} \subseteq D_{\mathfrak{q}_L}$ its inertia group, and $\mathbf{F}_{\mathfrak{q}_L} := \mathbf{Z}_L/\mathfrak{q}_L$ its residue field, where $\mathbf{Z}_L$ denotes the ring of integers of L. Taking the inverse limit of the exact sequences

$$1 \to I_{\mathfrak{q}_L} \to D_{\mathfrak{q}_L} \to \mathrm{Gal}(\mathbf{F}_{\mathfrak{q}_L}/\mathbf{F}_p) \to 1$$

over finite extensions $L/\mathbf{Q}$ gives an exact sequence of profinite groups

$$1 \to I_p \to D_p \to \mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p) \to 1.$$

We then pick $\mathrm{Frob}_p \in D_p \subseteq \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ to be any lift of the Frobenius automorphism $x \to x^p$ in $\mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p)$. Note that we have made two arbitrary choices in our definition of Frob_p, we chose an element in the inverse image of the Frobenius automorphism under $D_p \to \mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p)$, and we picked an embedding of $\overline{\mathbf{Q}}$ into $\overline{\mathbf{Q}}_p$, so Frob_p is in no way canonical, but it certainly exists. Its key property is that if $L/\mathbf{Q}$ is a finite Galois extension in which p is unramified, then the conjugacy class $\mathrm{conj}_L(\mathrm{Frob}_\mathfrak{p})$ in $\mathrm{Gal}(L/\mathbf{Q})$ of the restriction of $\mathrm{Frob}_\mathfrak{p} : \overline{\mathbf{Q}} \to \overline{\mathbf{Q}}$ to L is uniquely determined, independent of the choices we made. One can think of $\mathrm{Frob}_\mathfrak{p}$ as defining a map $L \to \mathrm{conj}_L(\mathrm{Frob}_\mathfrak{p})$ that assigns to each finite Galois extension $L/\mathbf{Q}$ the conjugacy class of $\mathrm{Gal}(L/\mathbf{Q})$ corresponding to the Frobenius automorphism when p is unramified in L. Everything we have said applies *mutatis mutandi* if we replace $\mathbf{Q}$ by a number field K: put $\overline{K} := \overline{\mathbf{Q}}$, replace p by a prime $\mathfrak{p}$ of K (by which we mean a nonzero prime ideal of $\mathbf{Z}_K$), and replace $\mathbf{F}_p$ by the residue field $\mathbf{F}_\mathfrak{p} := \mathbf{Z}_K/\mathfrak{p}$.

We now make the following observation: for any prime p that does not divide $\operatorname{disc} f$ we have

$$N_f(p) = \operatorname{tr} \rho_f(\mathrm{Frob}_p). \tag{1}$$

This follows from the fact that the trace of a permutation matrix counts its fixed points. Since p is unramified, the inertia group $I_p \subseteq \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ acts trivially on the roots of f, and the action of Frob_p on the roots of f coincides (up to conjugation) with the action of the Frobenius automorphism $x \to x^p$ on the roots of $f_p(x)$, both of which are described by the permutation matrix $\rho(\mathrm{Frob}_p)$. The Chebotarev density theorem implies that we can compute c_i by applying (1) and simply counting the number of matrices in $\rho_f(\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}))$ that have trace i; it is enough to determine the trace and size of each conjugacy class.

Theorem 1.1. CHEBOTAREV DENSITY THEOREM *Let L/K be a finite Galois extension of number fields with Galois group $G := \mathrm{Gal}(L/K)$. For every subset C of G stable under conjugation we have*

$$\lim_{B\to\infty} \frac{\#\{N(\mathfrak{p}) \le B : \mathrm{conj}_L(\mathrm{Frob}_\mathfrak{p}) \subseteq C\}}{\#\{N(\mathfrak{p}) \le B\}} = \frac{\#C}{\#G},$$

where $\mathfrak{p}$ ranges over primes of K and $N(\mathfrak{p}) := \#\mathbf{F}_\mathfrak{p}$ is the cardinality of the residue field $\mathbf{F}_\mathfrak{p} := \mathbf{Z}_K/\mathfrak{p}$.

Proof. See Lecture 2. □

Remark 1.2. Typically one takes C to be a single conjugacy class (general result follows easily from this case). The asymptotic ratio that appears in the theorem depends only on degree-1 primes (those with prime residue field), since these make up all but a negligible proportion of the primes $\mathfrak{p}$ for which $N(\mathfrak{p}) \le B$ (this follows from earlier density results that are easy to prove). In our statement of the theorem we do not exclude primes of K that are ramified in L because they are finite in number and no matter what value $\mathrm{conj}_L(\mathrm{Frob}_\mathfrak{p})$ takes on these primes it will not change the limiting ratio.

In our example with $f(x) = x^3 - x + 1$, one finds that $G_f := \rho_f(\overline{\mathbf{Q}}/\mathbf{Q})$ is isomorphic to S_3 (the Galois group of its splitting field), and its three conjugacy classes, represented by the matrices

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

have traces 0, 1, 3 and sizes 2, 3, 1, respectively. It follows that

$$c_0 = 1/3, \quad c_1 = 1/2, \quad c_2 = 0, \quad c_3 = 1/6,$$

just as we conjectured.

If we endow the group G_f with the discrete topology it becomes a compact group, and therefore has a *Haar measure* μ that is uniquely determined once we scale it so that $\mu(G_f) = 1$, which we always assume to be the case. Recall that a Haar measure on a compact group G is a translation-invariant Radon measure (so $\mu(gS) = \mu(Sg) = \mu(S)$ for any measurable set S and $g \in G$) and is unique up to a scalar factor.[1] For finite groups the Haar measure μ is just the normalized counting measure. We can compute the expected value of trace (and many other statistical quantities of interest) by integrating against the Haar measure, which in this case just amounts to summing over the finite group G_f.

$$\mathrm{E}[\mathrm{tr}(g)] = \int_{G_f} \mathrm{tr}(g)\,\mu = \sum_{g\in G_f} \mathrm{tr}(g) = \sum_{i=0}^{d} c_i i.$$

[1] For locally compact groups G one distinguishes left and right Haar measures, but the two coincide when G is compact; see [14] for more background on Haar measures.

The Chebotarev density theorem implies that this is also the average value of $N_f(p)$, that is,

$$\lim_{B\to\infty} \frac{\sum_{p\le B} N_f(p)}{\sum_{p\le B} 1} = \mathrm{E}[\mathrm{tr}(g)].$$

This average is 1 in our example (c.f. Exercise 1.1)

The quantities c_i define a probability distribution, namely the probability distribution on $\mathrm{tr}(g)$ induced by the Haar measure μ, which we can view as a probability distribution on $N_f(p)$. That is, picking a random prime p in some large interval $[1,B]$ and computing $N_f(p)$ is the same thing as picking a random matrix g in H_f and computing $\mathrm{tr}(g)$. More precisely, the sequence $(N_f(p))_p$ indexed by primes p is *equidistributed* with respect to the pushforward of the Haar measure μ on to the space $\{\mathrm{tr}(g) : g \in G_f\}$; we will formally define precisely what this means in the next lecture.

1.2. **Moment sequences.** There is another way to characterize the probability distribution on $\mathrm{tr}(g)$ given by the c_i; we can compute its *moment sequence*:

$$\mathrm{M}[\mathrm{tr}(g)] := (\mathrm{E}[\mathrm{tr}(g)^n])_{n\ge 0},$$

where

$$\mathrm{E}[\mathrm{tr}(g)^n] = \int_{G_f} \mathrm{tr}(g)^n \mu.$$

It might seem silly to include the zeroth moment $\mathrm{E}[\mathrm{tr}(g)^0] = \mathrm{E}[1] = 1$ in our sequence, but we will see later why this is useful. In our example we have (beginning with the zeroth moment $\mathrm{E}[(\mathrm{tr}(g)^0] = 1$):

$$\mathrm{M}[\mathrm{tr}(g)] = (1,1,2,5,14,41,\dots,\tfrac{1}{2}(3^{n-1}+1),\dots).$$

The sequence $\mathrm{M}[\mathrm{tr}(g)]$ uniquely determines[2] the probability distribution on $\mathrm{tr}(g)$, thus it captures all the information encoded in the c_i. It may not seem very useful to replace a finite set of rational numbers with an infinite sequence of integers, but when dealing with continuous probability distributions (which we are forced to do as soon as we leave our weight zero setting), it is a convenient tool.

If we pick another cubic polynomial $f \in \mathbf{Z}[x]$, we will generally obtain the same result as we did in our example; when ordered by height almost all cubic polynomials f have $G_f \simeq S_3$. But there are exceptions. Certainly if f is not irreducible over $\mathbf{Q}$ then G_f will be isomorphic to a proper subgroup of S_3, and this also occurs when the splitting field of f is a cyclic cubic extension (this happens whenever the discriminant of $f(x)$ is a square; take $f(x) = x^3 - 3x - 1$ for example). Up to conjugacy there are four subgroups of S_3, and each corresponds to a different distribution of $N_f(p)$.

$f(x)$	G_f	c_0	c_1	c_2	c_3	$\mathrm{M}[\mathrm{tr}(g)]$
x^3-x	1	0	0	0	1	$(1,3,9,27,81,\dots)$
x^3+x	C_2	0	1/2	0	1/2	$(1,2,5,14,41,\dots)$
x^3-3x-1	C_3	2/3	0	0	1/3	$(1,1,3,19,27,\dots)$
x^3-x+1	S_3	1/3	1/2	0	1/6	$(1,1,2,\ 5,14,\dots)$

One can do the same thing with polynomials of degree $d > 3$. For $d \le 19$ the results are exhaustive: for every transitive subgroup G of S_d the database of Klüners and Malle [38] contains at least one polynomial $f \in \mathbf{Z}[x]$ with $G_f \simeq G$ (including all 1954 transitive subgroups of S_{16}). The non-transitive cases can be constructed as products (of groups and of polynomials) of transitive cases of lower degree.

[2]Not all moment sequences uniquely determine an underlying probability distribution, but all the moment sequence we shall consider satisfy *Carleman's condition* [39, p. 126], which ensures that they do.

It is an open question whether this can be done for all d (even in principle); this amounts to a strong form of the inverse Galois problem over $\mathbf{Q}$ (here we are asking not only whether every finite group can be realized as a Galois group over $\mathbf{Q}$, but whether every permutation group of degree d can be realized as the Galois group of the splitting field of a polynomial of degree d).

1.3. **Zeta functions.** For polynomials f of degree $d = 3$ there is a one-to-one correspondence between subgroups of S_d and distributions of $N_f(p)$. This is not true for $d \geq 4$. For example, the polynomials $f(x) = x^4 - x^3 + x^2 - x + 1$ with $G_f \simeq C_4$ and $g(x) = x^4 - x^2 + 1$ with $G_g \simeq C_2 \times C_2$ both have $c_0 = 3/4$, $c_1 = c_2 = c_3 = 0$, and $c_4 = 1/4$, corresponding to the moment sequence $\mathrm{M}[\mathrm{tr}(g)] = (1, 1, 4, 16, 64, \ldots)$. But we can distinguish these cases if, in addition to considering the distribution of $N_f(p)$, we also consider the distribution of

$$N_f(p^r) := \#\{x \in \mathbf{F}_{p^r} : f_p(x) = 0\}.$$

We have $N_g(p^2) = 4$ for almost all p, whereas $N_f(p^2)$ is 4 or 2 depending on whether p is a square modulo 5 or not. In terms of the matrix group G_f we have

$$N_f(p^r) = \mathrm{tr}\big(\rho_f(\mathrm{Frob}_p)^r\big) \tag{2}$$

for all primes p that do not divide $\mathrm{disc}\, f$. To see this, note that the permutation matrix $\rho_f(\mathrm{Frob}_p)^r$ corresponds to the permutation of the roots of $f_p(x)$ given by the rth power of the Frobenius automorphism $x \mapsto x^p$. Its fixed points are precisely the roots of $f_p(x)$ that lie in $\mathbf{F}_{p^r}$; taking the trace counts these roots, which is, by definition $N_f(p^r)$.

This naturally leads to the definition of the local *zeta function* of f at p:

$$Z_{f_p}(T) := \exp\left(\sum_{r=1}^{\infty} N_f(p^r)\frac{T^r}{r}\right), \tag{3}$$

which can be viewed as a generating function for the sequence $(N_f(p), N_f(p^2), N_f(p^3), \ldots)$.

Remark 1.3. The identity (2) can be viewed as very special case of the Grothendieck-Lefschetz Trace Formula. It allows us to express the zeta function $Z_{f_p}(T)$ as a sum over powers of the traces of the image of Frob_p under the Galois representation ρ_f. In general one considers the trace of the Frobenius endomorphism acting on étale cohomology, but in our setting the only relevant cohomology is H^0.

While defined as a power series, in fact $Z(f_p; T)$ is a rational function of the form

$$Z_{f_p}(T) = \frac{1}{L_p(T)}$$

where $L_p(T)$ is an integer polynomial whose roots lie on the unit circle. This can be viewed as a consequence of the Weil conjectures in dimension zero,[3] but in fact it follows directly from (2). Indeed, for any matrix $A \in \mathrm{GL}_d(\mathbf{C})$ we have the identity

$$\exp\left(\sum_{r=1}^{\infty} \mathrm{tr}\, A^r \frac{T^r}{r}\right) = \det(1 - AT)^{-1}, \tag{4}$$

[3]Provided one accounts for the fact that $f(x) = 0$ does not define an irreducible variety unless $\deg(f) = 1$; in this case $N_f(p^r) = 1$ and $L_p(T) = 1 - T$, which is consistent with the usual formulation of the Weil conjectures (see Theorem 1.7).

which can be proved by expressing the coefficients on both sides as symmetric functions in the eigenvalues of A (see Exercise 1.2). Applying (2) and (4) to the definition of $Z_{f_p}(T)$ in (3) yields

$$Z_{f_p}(T) = \frac{1}{\det(1-\rho_f(\mathrm{Frob}_p)T)},$$

thus

$$L_p(T) = \det(1-\rho_f(\mathrm{Frob}_p)T),$$

and we see that $L_p(T)$ is precisely the polynomial that appears in the Euler factor at p of the (partial) *Artin L-function* $L(\rho_f, s)$ for the representation ρ_f:

$$L(\rho_f, s) := \prod_p L_p(p^{-s})^{-1},$$

at least for primes p that do not divide $\mathrm{disc}(f)$; for the definition of the Euler factors at ramified primes (and the Gamma factors at archimedean places), see [41, Ch. 2]. We shall not be concerned with the Euler factors at ramified primes, other than to note that they are all holomorphic and nonvanishing. We should note that the L-function $L(\rho_f, s)$ is not primitive, because ρ_f is not irreducible; one can always remove at least a factor of $\zeta(s)$ (the Riemann zeta function).

Returning to our interest in equidistribution, the Haar measure μ on $G_f = \rho_f(\mathrm{Frob}_p)$ allows us to determine the distribution of L-polynomials $L_p(T)$ that we see as p varies. The polynomial $L_p(T)$ is just the reciprocal polynomial (reversed coefficients) of the characteristic polynomial of $\rho_f(\mathrm{Frob}_p)$. If we fix a polynomial $P(T)$ of degree $d = \deg f$, and pick a prime p at random from some large interval, the probability that $L_p(T) = P(T)$ is equal the the probability that the reciprocal polynomial $T^d P(1/T)$ is the characteristic polynomial of a random element of G_f (this probability will be zero unless $P(T)$ has a particular form, see Exercise 1.3).

Remark 1.4. For $d \le 5$ the distribution of characteristic polynomials uniquely determines each subgroup of S_d (up to conjugacy). This is not true for $d \ge 6$, and for $d \ge 8$ one can find non-isomorphic subgroups of S_d with the same distribution of characteristic polynomials (the transitive permutation groups `8T10` and `8T11` which arise for $x^8 - 13x^6 + 44x^4 - 17x^2 + 1$ and $x^8 - x^5 - 2x^4 + 4x^2 + x + 1$ are an example).

1.4. **Computing zeta functions in dimension zero.** Let us make a few remarks on the practical question of how one goes about computing the zeta function $Z_{f_p}(T)$, which amounts to computing the integer polynomial $L_p(T)$. It suffices to compute the integers $N_f(p^r)$ for $r \le d$, which is equivalent to determining the degrees of the irreducible polynomials appearing in the factorization of $f_p(x)$ in $\mathbf{F}_p[x]$; these determine the cycle type, and therefore the conjugacy class, of the permutation of the roots of $f_p(x)$ induced by the action of the Frobenius automorphism $x \mapsto x^p$, which in turn determines the characteristic polynomial of $\rho_f(\mathrm{Frob}_p)$ and $L_p(T) = \det(1-\rho_f(\mathrm{Frob}_p)T)$ (see Exercise 1.3). To determine the factorization pattern of $f_p(x)$ we may apply the following algorithm.

Algorithm 1.5. Given a polynomial $g \in \mathbf{F}_p[x]$ of degree d, compute the number n_i of degree i factors of g in $\mathbf{F}_q[x]$ for $1 \le i \le d$ as follows:

1. Let $g_1(x)$ be $g(x)$ made monic.
2. For i from 1 to d:
   a. If $\deg(g_i) < i$ then set $n_j := 0$ for $i \le j \le d$ and proceed to step 3.
   b. Using binary exponentiation in the ring $\mathbf{F}_p[x]/(g_i)$, compute $r_i(x) := x^{p^i} \bmod g_i(x)$.

c. Compute $h_i(x) = \gcd(g_i, r_i(x) - x) = \gcd(g_i(x), x^{p^i} - x)$ using the Euclidean algorithm.

d. Set $n_i := \deg h_i$ and compute $g_{i+1} := g_i / h_i$ using exact division.

3. Output $n_1, \ldots, n_d$.

Algorithm 1.5 makes repeated use of the fact that

$$x^{p^i} - x = \prod_{a \in \mathbf{F}_{p^i}} (x - a)$$

is equal to the product of all the monic degree i polynomials in $\mathbf{F}_p[x]$. By removing factors from $g(x)$ in increasing order by degree we ensure that the only degree i factors in $g_i(x)$ are irreducible. One represents $\mathbf{F}_p[x]$ as $\mathbf{Z}/p\mathbf{Z}$ using integer representatives in $[0, p-1]$. The algorithm generalizes to arbitrary finite fields $\mathbf{F}_q$ (replace p by q and represent $\mathbf{F}_q$ as $\mathbf{F}_p[x]/(h)$ for some irreducible polynomial $h \in \mathbf{F}_p[x]$). Using fast algorithms for integer and polynomial arithmetic and the fast Euclidean algorithm (see [20, §8-11], for example), one can show that this algorithm uses $O((d \log p)^{2+o(1)})$ bit operations, a running time that is quasi-quadratic in the $O(d \log p)$ bit-size of its input $g \in \mathbf{F}_p[x]$. In practical terms, it is extremely efficient. For example, the table of $c_i(B)$ values for our example polynomial $f(x) = x^3 - x + 1$ took less than two minutes to create using the `smalljac` software library [36, 58], which includes an efficient implementation of basic finite field arithmetic. The NTL [54] and FLINT [23, 24] libraries provide more comprehensive functionality; FLINT is incorporated into Sage.

Remark 1.6. Note that Algorithm 1.5 does *not* output the factorization of $g(x)$, just the degrees of its irreducible factors. The algorithm can be extended to a *probabilistic* algorithm that outputs the complete factorization of $g(x)$, see [20, Alg. 14.8], with an expected running time that is again quasi-quadratic in d and $\log p$. However, no *deterministic* polynomial-time algorithm for factoring polynomials over finite fields is known, not even in the case $d = 2$. This is a famous open problem. One approach to solving it is to first prove the generalized Riemann hypothesis (GRH), which would at least address the case $d = 2$, and many others, but it is not known whether the GRH is sufficient to address all cases.[4]

1.5. **Arithmetic schemes.** We now want to generalize our example. Let us replace our equation f with an *arithmetic scheme* X, a scheme of finite type over $\mathbf{Z}$; in the case we have been considering $X = \operatorname{Spec} A$ where $A = \mathbf{Z}[x]/(f)$. For each prime p the fiber X_p of $X \to \operatorname{Spec} \mathbf{Z}$ is a scheme of finite type over $\mathbf{F}_p$ and we let $N_X(p) := X_p(\mathbf{F}_p)$ count its $\mathbf{F}_p$-points; equivalently, we may define $N_X(p)$ as the number of closed points (maximal ideals) $x \in X$ whose residue field has cardinality p, and similarly define $N_X(q)$ for prime powers $q = p^r$. The local zeta function of X at p is defined in the same way as $Z_{f_p}(T)$:

$$Z_{X_p}(T) := \exp\left(\sum_{r=1}^{\infty} N_X(p^r) \frac{T^r}{r}\right).$$

The local zeta functions of X can be combined to into a single global zeta-function

$$\zeta_X(s) := \prod_p Z_{X_p}(p^{-s}).$$

In our example with $X = \operatorname{Spec} \mathbf{Z}[x]/(f)$, the global zeta function $\zeta_X(s)$ agrees with the Artin L-function $L(\rho_f, s) = \prod L_p(s)^{-1}$ up to a finite set of factors at primes p that divide $\operatorname{disc}(f)$.

Finally, let us note that everything we have done over $\mathbf{Q}$ works over any number field K: replace $\mathbf{Q}$ by K, replace $\mathbf{Z}$ by $\mathcal{O}_K$, replace p by a prime $\mathfrak{p}$ of K (nonzero prime ideal of $\mathcal{O}_K$), replace the finite field $\mathbf{F}_p$ by the finite field $\mathbf{F}_\mathfrak{p} := \mathcal{O}_K/\mathfrak{p}$ and order primes $\mathfrak{p}$ by $N(\mathfrak{p})$ (we can break ties arbitrarily), so that rather

[4] On the plus side, if you succeed with this first step the Clay institute will provide you with funding for the remaining work.

that summing over $p \le B$ we sum over $\mathfrak{p}$ for which $N(\mathfrak{p}) \le B$. As noted in Remark 1.2, we can generally restrict our attention to degree-1 primes $\mathfrak{p}$ (those with prime residue field).

1.6. **A second example.** We now leave the world of Artin motives (motives of weight 0 arising from varieties of dimension 0) and consider our first example in weight 1 and dimension 1, an elliptic curve $E/\mathbf{Q}$, which is the setting in which the Sato–Tate conjecture was originally formulated. Such a curve can always be written in the form

$$E\colon y^2 = x^3 + Ax + B,$$

with $A, B \in \mathbf{Z}$. This equation is understood to define a smooth projective curve in $\mathbf{P}^2$ (homogenize the equation by introducing a third variable z), which has a single projective point $P_\infty := (0 : 1 : 0)$ at infinity that we take as the identity element of the group law. Recall that an elliptic curve is not just a curve, it comes equipped with a distinguished rational point, which after a suitable translation we may assume is the point P_∞.

The group operation on E can be defined via the usual chord-and-tangent law (three points on a line sum to zero), which can be used to derive explicit formulas with coefficients in $\mathbf{Q}$, or in terms of the divisor class group $\mathrm{Pic}^0(E)$ (divisors of degree zero modulo principal divisors), in which every divisor class can be uniquely represented by a divisor of the form $P - P_\infty$, where P is a point on the curve. This latter view is more useful in the sense that it generalizes to curves of genus $g > 1$, whereas the chord-and-tangent law does not. The Abel–Jacobi map $P \mapsto P - P_\infty$ gives a bijection between points on E and points on $\mathrm{Jac}(E)$ that commutes with the group operation, so the two approaches are isomorphic.

For each prime p that does not divide the discriminant $\Delta := -16(4A^3 + 27B^2)$ we can reduce E modulo p to obtain an elliptic curve $E_p/\mathbf{F}_p$; in this case we say that p is a *prime of good reduction for* E (or simply a *good prime*). We should note that the discriminant Δ is not necessarily minimal, the curve E may have another model that has good reduction at primes that divide Δ (including 2), but as we are happy to ignore a finite set of primes, we shall not be concerned with this.[5]

For every prime p of good reduction for E we have[6]

$$N_E(p) = \#E_p(\mathbf{F}_p) = p + 1 - t_p,$$

where the integer t_p satisfies the Hasse-bound $|t| \le 2\sqrt{p}$. Notice that, unlike the situation in our first example, the integers $N_E(p)$ tend to infinity with p, and to first order it is equal to $p + 1$. In order to study how the error term t_p varies with p we define the normalized value

$$x_p := t_p/\sqrt{p} \in [-2, 2].$$

We are now in a position to conduct the following experiment: given an elliptic curve $E/\mathbf{F}_p$, compute x_p for all good primes $p \le B$ and see how the x_p are distributed over the real interval $[-2, 2]$.

One can see an example in Figure 1 for the curve $y^2 = x^3 + x + 1$. The figure shows a histogram in which the x-axis ranges over the interval $[-2, 2]$. In each frame this interval is subdivided into approximately $\sqrt{\pi(B)}$ subintervals, each of which contains a colored bar whose height is proportional to the number of x_p (for $p \le B$) that lie in the subinterval. The gray line shows the height of the uniform distribution for scale (note that the vertical and horizontal scales are not the same, they were chosen

[5] All elliptic curves over $\mathbf{Q}$ have a global minimal model, but it is not necessarily of the form $y^2 = x^3 + Ax + B$. Over number fields global minimal models do not always exist (they do when the class number is one).

[6] The trace of Frobenius t_p is usually denoted a_p, but we wish to avoid a future collision in notation.

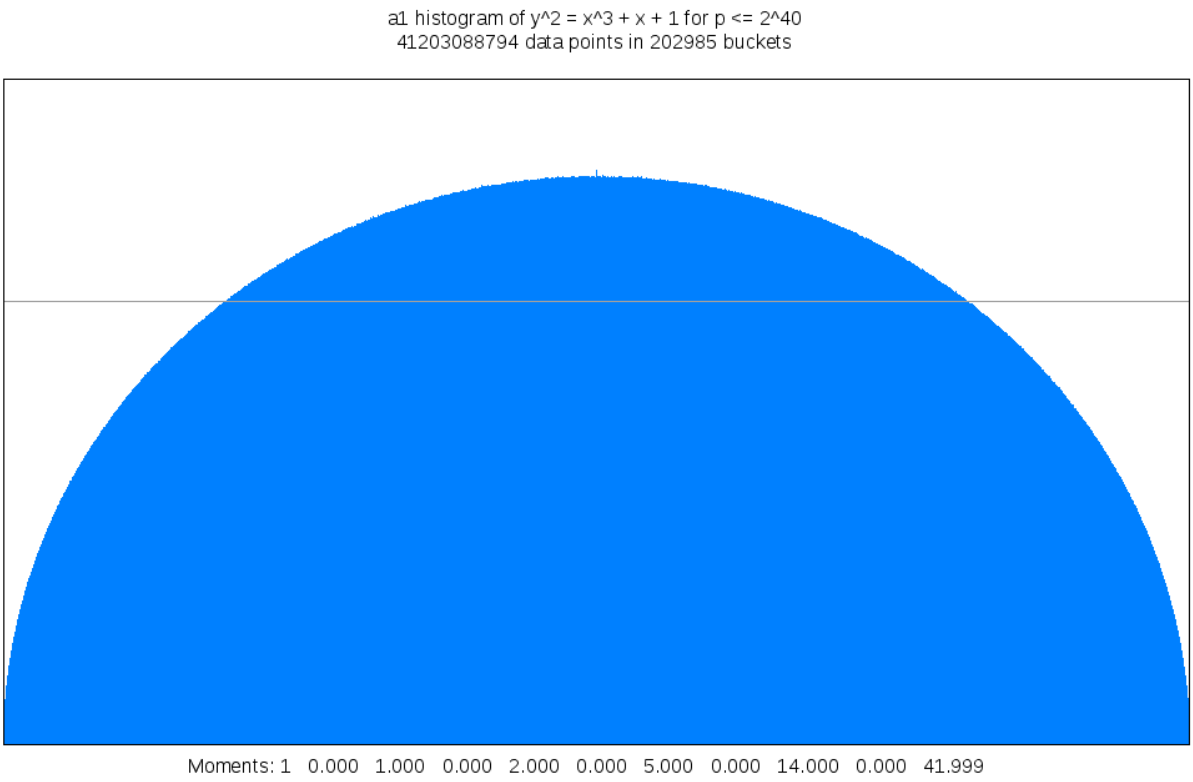

FIGURE 1. Click image to animate (requires Adobe Reader), or visit this web page.

judiciously). For $0 \le n \le 10$, the moment statistics

$$M_n := \frac{\sum_{p \le B} x_p^n}{\sum_{p \le B} 1},$$

are shown below the histogram. They appear to be converging to $1, 0, 1, 0, 2, 0, 5, 0, 14, 0, 42$ (the start of sequence A126120 in the Online Encyclopedia of Integer Sequences (OEIS) [44]).

The Sato–Tate conjecture for elliptic curves over $\mathbf{Q}$ (now a theorem) implies that for almost all elliptic curves $E/\mathbf{Q}$, if we run the same experiment we will see the same asymptotic distribution of Frobenius traces that is visible in the figure above (and the same limiting sequence of moments). In order to make this precise we would like to explain where the conjectured distribution comes from. In our first example we had a compact matrix group G_f associated to our scheme $X = \operatorname{Spec} \mathbf{Z}[x]/(f)$ whose Haar measure governed the distribution of $N_f(p)$. In fact we showed that more is true: there is a direct relationship between characteristic polynomials of elements of G_f and the L-polynomials $L_p(T)$ appearing in the local zeta functions $Z_{f_p}(T)$.

The same is true here. In order to identify a candidate group G_E whose Haar measure controls the distribution of normalized Frobenius traces x_p we should first look at the local zeta functions $Z_{E_p}(T)$. Let us recall what the Weil conjectures [67] (proved by Deligne [12, 13]) tell us about the zeta function of a variety over a finite field. The case of one-dimensional varieties (curves) was proved by Weil [65], who also proved an analogous result for abelian varieties [66], and these cover all the cases we shall

consider, but let us state the general result. Recall that for a compact manifold over $\mathbf{C}$, its *Betti number* b_i is the rank of the singular homology group $H_i(X,\mathbf{Z})$, and its *Euler characteristic* is $\chi = \sum(-1)^i b_i$.

Theorem 1.7 (WEILCONJECTURES)**.** *Let X be a geometrically irreducible non-singular projective variety of dimension n defined over a finite field $\mathbf{F}_q$. Let*

$$Z_X(T) := \exp\left(\sum_{r=1}^{\infty} N_E(q^r)\frac{T^r}{r}\right).$$

The following hold:

(1) ***Rationality**: $Z_X(T)$ is a rational function of the form*

$$Z_X(T) = \frac{P_1(T)\cdots P_{2n-1}(T)}{P_0(T)\cdots P_{2n}(T)},$$

with $P_i \in \mathbf{Z}[T]$ and $P_j(0) = 1$.

(2) ***Functional Equation**: the roots of $P_i(T)$ are the same as the roots of $T^{\deg P_{2n-i}} P_{2n-i}(1/(q^n T))$.*[7]

(3) ***Riemann Hypothesis**: the complex roots of $P_i(T)$ all have absolute value $q^{-i/2}$.*

(4) ***Betti Numbers**: if X is the reduction of a non-singular variety Y defined over a number field $K \subseteq \mathbf{C}$, then the degree of P_i is equal to the Betti number b_i of $Y(\mathbf{C})$.*

The curve E_p is a curve of genus $g = 1$, so we may apply the Weil conjectures in dimension $n = 1$, with Betti numbers $b_0 = b_2 = 1$ and $b_1 = 2g = 2$. This implies that its zeta function has the form

$$Z_{E_p}(T) = \frac{L_p(T)}{(1-T)(1-pT)}, \tag{5}$$

where $L_p \in \mathbf{Z}[T]$ is a polynomial of the form

$$L_p(T) = pT^2 + c_1 T + 1,$$

with $|c_1| \le 2\sqrt{p}$ (by the Riemann Hypothesis). If we expand both sides of (5) as power series in $\mathbf{Z}[[T]]$ we obtain

$$1 + N_E(p)T^2 + \cdots = 1 + (p + 1 + c_1)T + \cdots,$$

so we must have $N_E(p) = p + 1 + c_1$, and therefore

$$c_1 = N_E(p) - p - 1 = -t_p.$$

It follows that $N_E(p)$ determines the entire zeta function $Z_{E_p}(T)$.

Corresponding to our normalization $x_p = t_p/\sqrt{p}$, we define the *normalized L-polynomial*

$$\bar{L}_p(T) := L_p(T/\sqrt{p}) = T^2 + a_1 T + 1,$$

where $a_1 = c_1/\sqrt{p} = -x_p$ is a real number in the interval $[-2,2]$ and the roots of $\bar{L}_p(T)$ lie on the unit circle. In our first example we obtained the group G_f as a subgroup of the permutations in $\mathrm{GL}_d(\mathbf{C})$. Here we want a subgroup of $\mathrm{GL}_2(\mathbf{C})$ whose elements have eigenvalues that are

- inverses (the functional equation requires $\bar{L}_p(T)$ to be self-reciprocal);
- lie on the unit circle (by the Riemann hypothesis).

[7] Moreover, one has $Z_X(T) = \pm q^{-n\chi/2} T^{-\chi} Z_X(1/(q^n T))$, where χ is the Euler characteristic of X, which is defined as the intersection number of the diagonal with itself in $X \times X$.

This makes it clear that the group G_E we are looking for, the group whose Haar measure we expect to control the distribution of $\bar{L}_p(T)$ as p varies, should be a subgroup of SU(2). As in the weight zero case, we expect that G_E should generically be as large as possible, that is, $G_E = \mathrm{SU}(2)$.

We now consider what it means for an elliptic curve to be generic.[8] Recall that the endomorphism ring of an elliptic curve always contains a subgroup isomorphic to $\mathbf{Z}$, corresponding to the multiplication-by-n maps $P \mapsto nP$. Here

$$nP = P + \cdots + P$$

denotes repeated addition under the group law, and we take the additive inverse if n is negative. For elliptic curves over fields of characteristic zero, this typically accounts for all the endomorphisms, but in special cases the endomorphism ring $\mathrm{End}(E)$ may be larger and contain elements that are not multiplication-by-n maps. One can show that the characteristic polynomials of these extra endomorphisms are necessarily quadratic, with negative discriminants, and over $\mathbf{C}$ they correspond to multiplication by a root of this polynomial (which must lie in an imaginary quadratic field). We say that such elliptic curves have *complex multiplication* (CM).

We can now state the Sato-Tate conjecture, as independently formulated by Mikio Sato and John Tate in the mid 1960's and finally proved in the late 2000's [5, 6, 22].

Theorem 1.8 (Sato–Tate conjecture)**.** *Let $E/\mathbf{Q}$ be an elliptic curve without CM. The sequence of normalized Frobenius traces x_p associated to E is equidistributed with respect to the pushforward of the Haar measure on* SU(2) *under the trace map. In particular, for every subinterval $[a,b]$ of $[-2,2]$ we have*

$$\lim_{B\to\infty} \frac{\#\{p \le B : x_p \in [a,b]\}}{\#\{p \le B\}} = \frac{1}{2\pi}\int_a^b \sqrt{4-t^2}\,dt.$$

We have not defined x_p for primes of bad reduction, but there is no need to do so; this theorem is purely an asymptotic statement. To see where the expression in the integral comes from, we need to understand the Haar measure on SU(2) (or rather its pushforward onto conjugacy classes). A conjugacy class in SU(2) can be described by an *eigenangle* $\theta \in [0,\pi]$; its eigenvalues are then $e^{\pm i\theta}$ (a conjguate pair on the unit circle, as required). In terms of eigenangles, the Haar measure on SU(2) is given by

$$\mu = \frac{2}{\pi}\sin^2\theta\, d\theta,$$

and the trace is $t = 2\cos\theta$; from this one can deduce the measure $\frac{1}{2\pi}\sqrt{4-t^2}dt$ on the trace that appears in Theorem 1.8. We can also use the Haar measure to compute the nth moment of the trace

$$\mathrm{E}[t^n] = \frac{2}{\pi}\int_0^\pi (2\cos\theta)^n \sin^2\theta\, d\theta = \begin{cases} 0 & \text{if } n \text{ is odd,} \\ \frac{1}{m+1}\binom{2m}{m} & \text{if } n = 2m \text{ is even,} \end{cases} \tag{6}$$

and find that the $2m$th moment is the mth Catalan number.[9]

[8] The criterion given here in terms of endomorphism rings suffices for elliptic curves (and curves of genus $g \le 3$ or abelian varieties of dimension $g \le 3$), but in general one wants the Galois image to be generic (as large as possible), which is a strictly stronger condition for $g > 3$. This issue will be discussed further in Lecture 3.

[9] This gives yet another way to define the Catalan numbers, one that does not appear to among the 214 combinatorial interpretations enumerated in [57].

1.7. **Exercises.**

Exercise 1.1. Let $f \in \mathbf{Z}[x]$ be a nonconstant squarefree polynomial. Prove that the average value of $N_f(p)$ over $p \le B$ converges to the number of factors of f in $\mathbf{Z}[x]$ as $B \to \infty$.

Exercise 1.2. Prove that (4) holds for all matrices $A \in \mathrm{GL}_d(\mathbf{C})$.

Exercise 1.3. Let $f_p \in \mathbf{F}_p[x]$ denote a squarefree polynomial of degree $d > 0$ and let $L_p(T)$ denote the denominator of the zeta function $Z_{f_p}(T)$. We know that the roots of $L_p(T)$ lie on the unit circle in the complex plane; show that each is in fact an nth root of unity for some $n \le d$. Give a one-to-one correspondence between (1) cycle-types of degree-d permutations, (2) possible factorization patterns of f_p in $\mathbf{F}_p[x]$, and (3) the possible polynomials $L_p(T)$. Explain why non-conjugate elements of $\rho_f(\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}))$ may have the same characteristic polynomial (give an explicit example).

Exercise 1.4. Construct a (not necessarily irreducible) quintic polynomial $f \in \mathbf{Z}[x]$ with no roots in $\mathbf{Q}$ for which $f_p(x)$ has a root in $\mathbf{F}_p$ for every prime p (hint: think about what its Galois group must look like). Compute $c_0, \dots, c_5$ and G_f.

Exercise 1.5. Let X be the arithmetic scheme $\operatorname{Spec} \mathbf{Z}[x,y]/(f,g)$, where

$$f(x,y) := y^2 - 2x^3 + 2x^2 - 2x - 2, \qquad g(x,y) := 4x^2 - 2xy + y^2 - 2.$$

By computing $Z_{X_p}(T) = L_p(T)^{-1}$ for sufficiently many small primes p, construct a list of the polynomials $L_p \in \mathbf{Z}[T]$ that you believe occur infinitely often, and estimate their relative frequencies. Use this data to derive a candidate for the matrix group $G_X := \rho_X(\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, where ρ_X is the Galois representation defined by the action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ on $X(\overline{\mathbf{Q}})$. You may want to use of computer algebra system such as Sage or Magma to facilitate these calculations.

2. Equidistribution, L-functions, and moment sequences

2.1. **Equidistribution.** Let us now formally define the notion of equidistribution, following [49, §1A]. For a compact Hausdorff space X, we use $C(X)$ denote its Banach space of complex-valued continuous functions $f : X \to \mathbf{C}$ equipped with the sup-norm $\|f\| := \sup_{x\in X} |f(x)|$. Note that $C(X)$ closed under pointwise addition and multiplication, and contains the constant functions, it is thus a commutative $\mathbf{C}$-algebra with unit $\mathbb{1}_X$ (the function $x \mapsto 1$).[10] The subset of real-valued functions in $C(X)$ is a distributive lattice under the order relation $f \leq g$, defined whenever $f(x) \leq g(x)$ for all $x \in X$; real-valued functions $f \geq 0$ are called *positive*.

Definition 2.1. A (positive normalized Radon) *measure* on a compact Hausdorff space X is a continuous $\mathbf{C}$-linear map $\mu\colon C(X) \to \mathbf{C}$ that is positive and of total mass 1. Here *positive* means that $\mu(f) \in \mathbf{R}_{\geq 0}$ for all positive $f \in C(X)$, and *total mass* 1, means $\mu(\mathbb{1}_X) = 1$ where $\mathbb{1}_X$ is the indicator function of X.

Example 2.2. For each point $x \in X$ the map $f \mapsto f(x)$ defines a measure δ_x, the *Dirac measure*.

The value of μ on $f \in C(X)$ is often denoted using integral notation

$$\int_X f\mu := \mu(f),$$

and we shall use the two interchangeably.

It is tempting to now define the measure of a set $S \subseteq X$ as the measure of its indicator function $\mathbb{1}_S$, but in general the function $\mathbb{1}_S$ will not lie in $C(X)$; this occurs if and only if S is both open and closed (which we note applies to $S = X$). Instead, for each open set $S \subseteq X$ we define

$$\mu(S) = \sup\left\{\mu(f) : 0 \leq f \leq 1_S, \text{ postive } f \in C(X)\right\} \in [0,1],$$

and for each closed set $S \subseteq X$ define

$$\mu(S) = 1 - \mu(X - S) \in [0,1].$$

If $S \subseteq X$ is contained in some open set U of measure $\mu(U) \leq \epsilon$ for every $\epsilon \geq 0$ then we define $\mu(S) = 0$ and say that S has measure zero. If the boundary $\partial S := \overline{S} - S^0$ of a set S has measure zero, then $\mu(S^0) = \mu(\overline{S})$, and we define $\mu(S)$ to be this common value; such sets are said to be *μ-quarrable*.

For the purpose of studying equidistribution, we shall restrict our attention to μ-quarrable sets S. This does typically does not include all measurable sets (an element of the Borel σ-algebra Σ of X generated by the open and closed sets of X under countable unions and intersections); see Exercise 2.1. On the other hand, if we are given an inner regular Borel measure μ on X of total mass 1, by which we mean a countably additive function $\mu\colon \Sigma \to \mathbf{R}_{\geq 0}$ for which $\mu(S) = \inf\{\mu(U) : S \subseteq U,\ U \text{ open}\}$ and $\mu(X) = 1$, it is easy to check that defining $\mu(f) := \int_X f\mu$ for each $f \in C(X)$ yields a measure in the sense of Definition 2.1 (see [30, §1] for this and related results). This implies that the (normalized) Haar measure of any compact group is a measure under Definition 2.1.

Definition 2.3. A sequence $(x_1, x_2, x_3, \ldots)$ in X is said to be *equidistributed with respect to μ* (or simply *μ-equidistributed*) if for every $f \in C(X)$

$$\mu(f) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} f(x_i).$$

[10] In fact it is a commutative C^*-algebra with complex conjugation as its involution, but we will not make use of this here.

Remark 2.4. When we speak of equidistribution, we are talking about a *sequence* of elements (x_i) listed in a particular order, as indicated by the indices $i \in \mathbf{Z}_{>0}$. It does not make sense to say that a set is equidistributed. For example, suppose we took the set of odd primes and arranged them in the sequence $(5, 13, 3, 17, 29, 7, \ldots)$ so that we listed two primes congruent to 1 modulo 4 followed by one prime congruent to 3 modulo 4, and continued in this fashion. The sequence obtained by reducing this sequence modulo 4 is not equidistributed with respect to the uniform measure on $(\mathbf{Z}/4\mathbf{Z})^\times$, but the sequence of odd primes in their usual order is. However, local rearrangements that change the index of an element by at most a bounded amount do not change equidistribution (or the lack thereof). This applies to sequences indexed by primes in a number field that are ordered by norm (where we may break ties arbitrarily), but not to sequences indexed by primes in a global function field.

If (x_i) is a sequence in X, for each $f \in C(X)$, we define the kth-*moment* of the sequence $(f(x_i))$ by

$$\mathrm{M}_k[(f(x_i)] := \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} f^k(x_i).$$

If these limits exist for all $k \geq 0$ we define the *moment sequence*

$$\mathrm{M}[f(x_i)] := (\mathrm{M}_0[(f(x_i)], \mathrm{M}_1[(f(x_i)], \mathrm{M}_2[(f(x_i)], \ldots).$$

Note that if (x_i) is μ-equidistributed, then we have $\mathrm{M}_k[f(x_i)] = \mu(f^k)$ and the moment sequence

$$\mathrm{M}[f(x_i)] = (\mu(f^0), \mu(f^1), \mu(f^2), \ldots), \tag{7}$$

is independent of the sequence (x_i), it depends only on the function f and the measure μ.

Remark 2.5. There is a partial converse that is relevant to some of our applications. To simplify matters, let us restrict our attention to real-valued functions; so for the purposes of this remark, let $C(X)$ denote the Banach algebra of real-valued functions on X and replace $\mathbf{C}$ with $\mathbf{R}$ in Definition 2.1. Let (x_i) be a sequence in X and let $f \in C(X)$. Then $f(X)$ is a compact subset of $\mathbf{R}$, and we can view $(f(x_i))$ as a sequence in $f(X)$. If the moments $\mathrm{M}_k[f(x_i)]$ exist for all $k \geq 0$, then there is a unique measure on $f(X)$ with respect to which the sequence $(f(x_i))$ is equidistributed; this follows from the Stone-Weierstrass theorem. If μ is a measure on $C(X)$, we may define the pushforward measure $\mu_f(g) := \mu(g \circ f)$ on $C(f(X)$, and we see that the sequence $(f(x_i))$ is μ_f-equidistributed if and only if (7) holds. This gives a necessary (but in general not sufficient condition) for (x_i) to be μ-equidistributed that can be checked by comparing moment sequences. If we have a collection of functions $f_i \in C(X)$ such that the pushforward measures μ_{f_i} uniquely determine μ, we obtain a necessary and sufficient condition involving the moment sequences of the f_i with respect to μ. One can generalize this remark to the complex-valued case using the theory of C^*-algebras.

More generally, we have the following lemma.

Lemma 2.6. *Let (f_j) be a family of functions whose linear combinations are dense in $C(X)$. If (x_i) is a sequence in X for which the limit $\lim_{n\to\infty}\sum_{i=1}^{n} f_j(x_i)$ exists for every f_j then there is a unique measure μ on X for which (x_i) is μ-equidistributed.*

Proof. This is [49, Lemma A.1, p. I-19]. □

Proposition 2.7. *If (x_i) is a μ-equidistributed sequence in X and S is a μ-quarrable set in X then*

$$\mu(S) = \lim_{n\to\infty} \frac{\#\{x_i \in S : i \leq n\}}{n}.$$

Proof. See Exercise 2.2. □

Example 2.8. If $X = [0,1]$ and μ is the Lebesgue measure then a sequence (x_i) is μ-equidistributed if and only if for every $0 \le a < b \le 1$ we have

$$\lim_{n\to\infty} \frac{\#\{x_i \in [a,b] : i \le n\}}{n} = b - a$$

More generally, if X is a compact subset of $\mathbf{R}^n$ and μ is the normalized Lebesgue measure then (x_i) is μ-equidistributed if and only if for every measurable set $S \subseteq X$ we have $\lim_{n\to\infty} \frac{1}{n}\#\{x_i \in S : i \le n\} = \mu(S)$.

2.2. **Equidistribution in compact groups.** We now specialize to the case where $X := \mathrm{conj}(G)$ is the space of conjugacy classes of a compact group G, obtained by taking the quotient of G as a topological space under the equivalence relation defined by conjugacy; let $\pi : G \to X$ denote the quotient map. We then equip X with the pushforward of the Haar measure μ on G (normalized so that $\mu(G) = 1$), which we also denote μ. Explicitly, π induces a natural map of Banach spaces

$$\begin{aligned} C(X) &\to C(G) \\ f &\mapsto f \circ \pi \end{aligned}$$

and the value of μ on $C(X)$ is defined by

$$\mu(f) := \mu(f \circ \pi).$$

We say that a sequence (x_i) in X or a sequence (g_i) in G is *equidistributed* if it is μ-equidistributed (when we speak of equidistribution in a compact group without referencing a measure, we always mean the normalized Haar measure).

Proposition 2.9. *Let G be a compact group with Haar measure μ, and let $X = \mathrm{conj}(G)$. A sequence (x_i) in X is μ-equidistributed if and only if for every irreducible character χ of G we have*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \chi(x_i) = \mu(\chi)$$

Proof. As explained in [49, Prop. A.2], this follows from Lemma 2.6 and the Peter-Weyl theorem, since the irreducible characters χ of G generate a dense subset of $C(X)$. □

Corollary 2.10. *Let G be a compact group with Haar measure μ, and let $X = \mathrm{conj}(G)$. A sequence (x_i) in X is μ-equidistributed if and only if for every nontrivial irreducible character χ of G we have*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \chi(x_i) = 0.$$

Proof. For the trivial character we have $\mu(1) = \mu(G) = 1$, and for any nontrivial irreducible character χ we must have $\mu(\chi) = \int_G \chi\mu = 0$; the corollary then follows from the previous proposition. □

As an exercise, let us apply Corollary 2.10 to prove an equidistribution result for elliptic curves over finite fields that will be useful later. We first recall some basic facts. Let $E/\mathbf{F}_q$ be an elliptic curve. Recall that the *Frobenius endomorphism* π_E is defined by

$$(x : y : z) \mapsto (x^q : y^q : z^q).$$

Like all endomorphisms of elliptic curves, π_E has a characteristic polynomial of the form

$$\lambda^2 - (\mathrm{tr}\,\pi_E)\lambda + \deg \pi_E$$

that is satisfied by both π_E and its dual[11] $\hat{\pi}_E$, where $\operatorname{tr}\pi_E = \pi_E + \hat{\pi}_E$ and $q = \deg \pi_E = \pi_E \hat{\pi}_E$ are both integers. The set $E(\mathbf{F}_q)$ is, by definition, the subset of $E(\overline{\mathbf{F}_q})$ fixed by π_E; equivalently, the kernel of the endomorphism $\pi_E - 1$. One can show that $\pi_E - 1$ is a separable, hence

$$\#E(\mathbf{F}_q) = \#\ker(\pi_E - 1)\deg(\pi_E - 1) = (\pi_E - 1)(\hat{\pi}_E - 1) = \hat{\pi}_E\pi_E + 1 - (\hat{\pi}_E + \pi_E)q + 1 - \operatorname{tr}\pi_E.$$

It follows that $t_q := q + 1 - \#E(\mathbf{F}_q)$ is the *trace of Frobenius* $\operatorname{tr}\pi_E$. As we showed in the previous lecture in the case $q = p$, the zeta function of E can then be written as

$$Z_E(T) = \frac{qT^2 - t_qT + 1}{(1-T)(1-qT)},$$

where the complex roots of $qT^2 - t_qT + 1$ have absolute value $q^{-1/2}$. This implies that we can write $t_q = \alpha + \bar{\alpha}$ for some $\alpha \in \mathbf{C}$ with $|\alpha| = q^{1/2}$, and we have $\#E(\mathbf{F}_q) = q + 1 - (\alpha + \bar{\alpha})$. We now observe that for any $r \geq \mathbf{Z}_{\geq 1}$ the set $E(\mathbf{F}_{q^r})$ is the subset of $E(\overline{\mathbf{F}_q})$ fixed by π_E^r (which is induced by the q^r-power Frobenius automorphism); it follows that

$$E(\mathbf{F}_{q^r}) = q^r + 1 - (\alpha^r + \bar{\alpha}^r).$$

Finally, we recall that $E/\mathbf{F}_q$ is said to be *ordinary* if t_q is not zero modulo p (the characteristic of $\mathbf{F}_q$).

As an application of Corollary 2.10, we now prove the following result, taken from [15, Prop 2.2].

Proposition 2.11. *Let $E/\mathbf{F}_q$ be an ordinary elliptic curve and for $r \geq \mathbf{Z}_{\geq 1}$ define*

$$x_r := \frac{q^r + 1 - \#E(\mathbf{F}_{q^r})}{q^{r/2}}.$$

The sequence $(x_i) \in X := [-2, 2]$ is equidistributed with respect to the measure

$$\mu = \frac{1}{\pi}\frac{dz}{\sqrt{4 - z^2}},$$

where dz is the Lebesgue measure on $[-2, 2]$.

Proof. Let $\alpha := \alpha_q$ be as above, with $|\alpha| = q^{1/2}$ and $\operatorname{tr}\pi_E = \alpha + \bar{\alpha}$. Then $x_r = (\alpha^r + \bar{\alpha}^r)/q^{r/2}$ for all $r \geq 1$. Let $\mathrm{U}(1) := \{u \in \mathbf{C}^\times : u\bar{u} = 1\}$ be the unitary group. For $u = e^{i\theta}$, the Haar measure on $\mathrm{U}(1)$ corresponds to the uniform measure on $\theta \in [-\pi, \pi)$, this follows immediately the translation invariance of the Haar measure. Let us compute the pushforward of the Haar measure of $\mathrm{U}(1)$ to $[-2, 2]$ via the map $u \mapsto z := u + \bar{u} = 2\cos\theta$. We have $dz = 2\sin\theta\, d\theta$, and see that the pushforward is precisely μ.

The nontrivial irreducible characters of $\phi_a \colon \mathrm{U}(1) \to \mathbf{C}^\times$ are of the form $\phi_a = (u^a)$ for some nonzero $a \in \mathbf{Z}$. For each such ϕ_a we have

$$\lim_{n\to\infty}\frac{1}{n}\sum_{r=1}^{n}\phi_a(\alpha^r/q^{r/2}) = \lim_{n\to\infty}\frac{1}{n}\sum_{r=1}^{n}(\alpha/q^{1/2})^{ra} = \lim_{n\to\infty}\frac{1}{n}\frac{(\alpha/q^{1/2})^{a(n+1)} - (\alpha/q^{1/2})^a}{(\alpha/q^{1/2})^a - 1} = 0.$$

The hypothesis that E is ordinary guarantees that $\alpha/q^{1/2}$ is not a root of unity, thus $(\alpha/q^{1/2})^a - 1$ is nonzero for all nonzero $a \in \mathbf{Z}$. The proposition then follows from Corollary 2.10. □

See [2] for a generalization to smooth projective curves $C/\mathbf{F}_q$ of arbitrary genus $g \geq 1$.

[11] By the *dual* of an endomorphism of a principally polarized abelian variety we mean the Rosati dual, which for elliptic curves we may identify with the dual isogeny.

2.3. **Equidistribution for L-functions.** As above, let G be a compact group and let $X := \mathrm{conj}(G)$. Let K be a number field, and let $(\mathfrak{p}_1, \mathfrak{p}_2, \mathfrak{p}_3, \ldots)$ be a sequence consisting of all but finitely many primes $\mathfrak{p}$ of K ordered by norm; this means that $N(\mathfrak{p}_i) \le N(\mathfrak{p}_j)$ for all $i \le j$. Let (x_i) be a sequence in X. Given an irreducible representation $\rho : G \to \mathrm{GL}_d(\mathbf{C})$, we define the L-function

$$L(\rho, s) := \prod_{i \ge 1} \det(1 - \rho(x_i) N(\mathfrak{p}_i)^{-s})^{-1},$$

for $s \in \mathbf{C}$ with $\mathrm{Re}(s) > 1$.

Theorem 2.12. *Let G, X, P, and (x_i) be as above, and suppose that every irreducible representation ρ of G has a meromorphic L-function $L(\rho, s)$ on $\mathrm{Re}(s) \ge 1$ with no zeros or poles except possibly at $s = 1$. The sequence (x_i) is μ-equidistributed if and only if for every irreducible nontrivial representation ρ of G the L-function $L(\rho, s)$ extends to a holomorphic function on $\mathrm{Re}(s) \ge 1$ that is nonzero at $s = 1$.*

Proof. See the corollary to [49, Thm. A.2], or see [15, Thm. 2.3]. □

A notable case where the hypothesis of Theorem 2.12 is known to hold is when $L(\rho, s)$ corresponds to an Artin L-function. As in Lecture 1 that to each prime $\mathfrak{p}$ in K we associate an absolute Frobenius element $\mathrm{Frob}_\mathfrak{p} \in \mathrm{Gal}(\overline{K}/K)$, and for each finite Galois extension L/K we use $\mathrm{conj}_L(\mathrm{Frob}_\mathfrak{p})$ denote the conjugacy class in $\mathrm{Gal}(L/K)$ of the restriction of $\mathrm{Frob}_\mathfrak{p}$ to L.

Corollary 2.13. *Let L/K be a finite Galois extension with $G := \mathrm{Gal}(L/K)$, let $(\mathfrak{p}_i)$ be the sequence of unramified primes of K ordered by norm (break ties arbitrarily). The sequence $(\mathrm{conj}_L(\mathrm{Frob}_{\mathfrak{p}_i}))$ is equidistributed in $\mathrm{conj}(G)$; in particular, the Chebotarev density theorem (Theorem 1.1) holds.*

Proof. For the trivial representation $L(1, s)$ agrees with the Dedekind zeta function $\zeta_K(s)$ up to a finite number of holomorphic nonvanishing factors, and (as originally proved by Hecke) $\zeta_K(s)$ is holomorphic and nonvanishing on $\mathrm{Re}(s) \ge 1$ except for a simple pole at $s = 1$; see [43, Cor. VII.5.11], for example. For every nontrivial irreducible representation $\rho : G \to \mathrm{GL}_d(\mathbf{C})$, the L-function $L(\rho, s)$ agrees with the corresponding Artin L-function for ρ, up to a finite number of holomorphic nonvanishing factors, and (as originally proved by Artin) $L(\rho, s)$ is known to be holomorphic and nonvanishing on $\mathrm{Re}(s) \ge 1$; see [9, p.225], for example. The corollary then follows from Theorem 2.12. □

2.4. **Sato–Tate for CM elliptic curves.** As a second application of Theorem 2.12, let us prove an equidistribution result for CM elliptic curves. To do so we need to introduce Hecke L-function, which are closely related (via Artin reciprocity) to Artin L-functions of one-dimensional representations. For the sake of concreteness we give the classical ideal-theoretic definition (there is of course an idelic version).

Definition 2.14. Let K be number field, let $\mathfrak{n}$ be a $\mathbf{Z}_K$-ideal, let $I_\mathfrak{n}$ denote the group of fractional $\mathbf{Z}_K$-ideals prime to $\mathfrak{n}$, and let $l \ge 1$ be an integer. An *algebraic Hecke character* ψ of K of modulus $\mathfrak{n}$ and *infinity type* l is a group homomorphism

$$\psi : I_\mathfrak{n} \to \mathbf{C}^\times$$

such that $\psi(\alpha \mathbf{Z}_k) = \alpha^l$ for all $\alpha \in K^\times$ for which $\alpha \equiv^\times 1 \bmod \mathfrak{n}$ (here the *multiplicative congruence* notation $\alpha \equiv^\times \beta \bmod \mathfrak{n}$ means $v_\mathfrak{p}(\alpha/\beta - 1) \ge v_p(\mathfrak{n})$ for every prime $\mathfrak{p}|\mathfrak{n}$).

If ψ and ψ' are algebraic Hecke characters of moduli $\mathfrak{n}$ and $\mathfrak{n}'$ of the same infinity type with $\mathfrak{n} \subseteq \mathfrak{n}'$ such that ψ is the restriction of ψ' to $I_\mathfrak{n}$ then we say that ψ is an extension of ψ'. If ψ is not the extension of any $\psi' \ne \psi$, then we set that ψ is primitive and call $\mathfrak{n}$ the conductor of ψ. We extend ψ to all fractional $\mathbf{Z}_K$-ideals by defining $\psi(\mathfrak{a}) = 0$ if $\mathfrak{p} + \mathfrak{a} \ne \mathbf{Z}_K$.

Each algebraic Hecke character ψ has an associated *Hecke L-function*

$$L(\psi, s) := \prod_{\mathfrak{p}} (1 - \psi(\mathfrak{p}) N(\mathfrak{p})^{-s})^{-1},$$

where the product runs over primes of K; the Euler factor is trivial when $\mathfrak{p}$ divides the modulus $\mathfrak{n}$ of ψ.

Now let us restrict to the case where K is an imaginary quadratic field, and let ψ be a primitive algebraic Hecke character of modulus $\mathfrak{n}$ and infinity type $l = 1$. Let $(\mathfrak{p}_1, \mathfrak{p}_2, \mathfrak{p}_3, \ldots)$ be the primes $\mathfrak{p}$ of K that do not divide $\mathfrak{n}$, ordered by norm (break ties arbitrarily). Fix an embedding $K \hookrightarrow \mathbf{C}$, so that we have $\mathfrak{p}\bar{\mathfrak{p}} = N(\mathfrak{p})$, where $\bar{\cdot}$ denotes complex conjugation. Then $\mathfrak{p}^{-1} = N(\mathfrak{p})^{-1}\bar{\mathfrak{p}}$, and for $\mathfrak{p} \nmid \mathfrak{n}$ we must have

$$\psi(\mathfrak{p})\psi(\bar{\mathfrak{p}}) = \psi(\mathfrak{p})\overline{\psi(\mathfrak{p})} = N(\mathfrak{p}),$$

since $\psi : I_{\mathfrak{n}} \to \mathbf{C}_\times$ is a group homomorphism. Therefore $|\psi(\mathfrak{p})| = N(\mathfrak{p})^{1/2}$ for all $\mathfrak{p} \nmid \mathfrak{n}$. We now define

$$x_i := \frac{\psi(\mathfrak{p}_i)}{N(\mathfrak{p}_i)^{1/2}} \in \mathrm{U}(1).$$

Lemma 2.15. *The sequence (x_i) is equidistributed in* U(1).

Proof. As in the proof of Proposition 2.11, the nontrivial characters of U(1) are those of the form $\phi_a(z) = z^a$ with $a \in \mathbf{Z}$ nonzero, each of which corresponds, via Artin reciprocity, to a 1-dimensional Artin representation ρ whose L-function $L(\rho, s)$ is known to be holomorphic and nonvanishing on $\mathrm{Re}(s) \geq 1$; see the proof of Corollary 2.13. And as in the proof of Corollary 2.13, $L(1, s)$ is holomorphic and nonvanishing on $\mathrm{Re}(s)$ except for a simple pole at $s = 1$, because the same is true of $\zeta_K(s)$. The lemma then follows from Theorem 2.12. □

As an application of Corollary 2.15, we can now prove the Sato-Tate conjecture for CM elliptic curves; for the sake of simplicity we restrict ourselves to the case where K is an imaginary quadratic field and E/K is an elliptic curve with CM field K (recall that the CM field is the field defined by the characteristic polynomial of any element of $\mathrm{End}(E) - \mathbf{Z}$). The case of an elliptic curve $E/\mathbf{Q}$ can then be treated by basechange to K.

Let $\mathfrak{n}$ be the *conductor* of E; this is an ideal in $\mathbf{Z}_K$ supported on the primes of bad reduction for E; [55, §IV.10] for a precise definition. A classical result of Deuring guarantees the existence of a primitive algebraic Hecke character ψ_E of K of modulus $\mathfrak{n}$ and infinity type 1 such that for ever prime $\mathfrak{p} \nmid \mathfrak{n}$ we have

$$\psi_E(\mathfrak{p}) + \overline{\psi_E(\mathfrak{p})} = t_{\mathfrak{p}},$$

where $t_{\mathfrak{p}} := \mathrm{tr}\, \pi_E = N(\mathfrak{p}) + 1 - \#E_{\mathfrak{p}}(\mathbf{F}_{\mathfrak{p}}) \in \mathbf{Z}$ is the trace of Frobenius of the reduction of E modulo $\mathfrak{p}$.

Proposition 2.16. *Let K be an imaginary quadratic field and let E/K be a CM elliptic curve of conductor $\mathfrak{n}$ with CM field K. Let $(\mathfrak{p}_1, \mathfrak{p}_2, \ldots)$ be the primes of K that do not divide $\mathfrak{n}$ ordered by norm (break ties arbitrarily), and let $x_i := t_{\mathfrak{p}_i}/N(\mathfrak{p}_i)^{1/2} \in [-2, 2]$ be the normalized Frobenius trace of $E_{\mathfrak{p}_i}$. The sequence (x_i) is equidistributed on $[-2, 2]$ with respect to the measure*

$$\mu_{\mathrm{cm}} := \frac{1}{\pi} \frac{dz}{\sqrt{4 - z^2}}.$$

Proof. By the previous lemma, the sequence $(\psi_E(\mathfrak{p}_i)/N(\mathfrak{p}_i)^{1/2})_{i \geq 1}$ is equidistributed in U(1). As shown in the proof of Proposition 2.11, the measure μ_{cm} is the pushforward of the Haar measure on U(1) to $[-2, 2]$ under the map $u \mapsto u + \bar{u}$. The image of $\psi_E(\mathfrak{p}_i)/N(\mathfrak{p}_i)^{1/2}$ under this map is

$$\frac{\psi_E(\mathfrak{p}_i)}{N(\mathfrak{p}_i)^{1/2}} + \frac{\overline{\psi_E(\mathfrak{p}_i)}}{N(\mathfrak{p}_i)^{1/2}} = \frac{t_{p_i}}{N(\mathfrak{p}_i)} = x_i,$$

and the proposition follows. □

Figure 2 shows a trace histogram for the CM elliptic curve $y^2 = x^3 + 1$ over its CM field $\mathbf{Q}(\sqrt{-3})$.

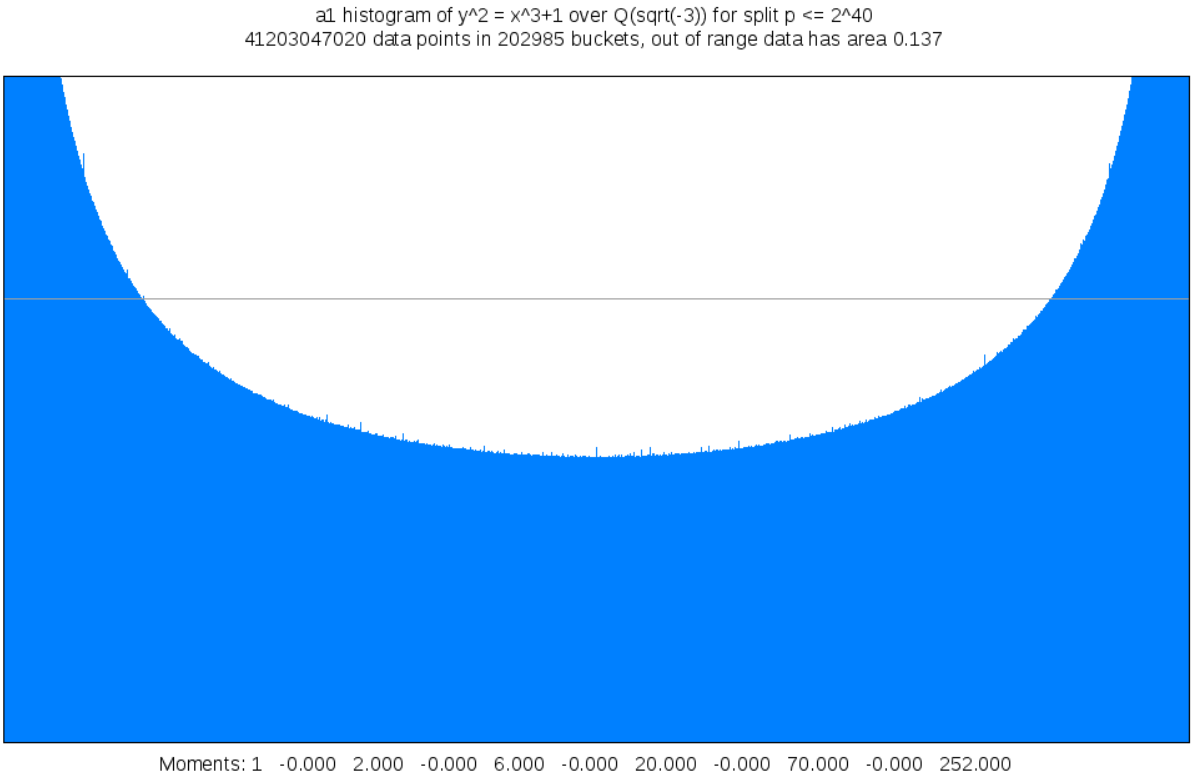

FIGURE 2. Click image to animate (requires Adobe Reader), or visit this web page.

Let us now consider the case of a CM elliptic curve $E/\mathbf{Q}$. For primes p of good reduction that are inert in the CM field K, the endomorphism algebra $\mathrm{End}^0(E_p) := \mathrm{End}(E_p) \otimes_{\mathbf{Z}} \mathbf{Q}$ of the reduced curve E_p will necessarily contain two distinct imaginary quadratic fields, one corresponding to the CM field and the other generated by the Frobenius endomorphism (the two cannot coincide because p is inert in K but the Frobenius endomorphism has norm p in $\mathrm{End}^0(E_p)$). It follows that $\mathrm{End}^0(E_p)$ must be a quaternion algebra, E_p is supersingular, and for $p > 3$ we must have $t_p = 0$, since $t_p \equiv 0 \bmod p$ and $|t_p| \le 2\sqrt{p}$; see [56, III,9,V.3] for these and other facts about the endomorphism ring of an elliptic curve.

At split primes $p = \mathfrak{p}\bar{\mathfrak{p}}$ the reduced curve E_p will be isomorphic to the reduction modulo $\mathfrak{p}$ of its base change to K (both of which will be elliptic curves over $\mathbf{F}_p = \mathbf{F}_{\mathfrak{p}}$), and they will have the same trace of Frobenius. It follows that the sequence of normalized Frobenius traces $x_i := t_p/\sqrt{p}$ is equidistributed on $[-2, 2]$ with respect to the measure that has mass $1/2$ concentrated at 0 with the remaining mass distributed according to $\frac{1}{2}\mu_{\mathrm{cm}}$ This can be seen in Figure 3, which shows a trace histogram for the CM elliptic curve $y^2 = x^3 + 1$ over $\mathbf{Q}$; the thin spike in the middle of the histogram at zero has area $1/2$ (one can also see that the nontrivial moments are half what they were in Figure 2).

2.5. **Sato–Tate for non-CM elliptic curves.** We can now state the Sato-Tate conjecture in the form originally proposed by Tate [61], as described in [49, §1A]. Tate's seminal paper [61] contains what is now known as the *Tate conjecture*, which comes in two conjecturally equivalent forms **T1** and **T2**, the

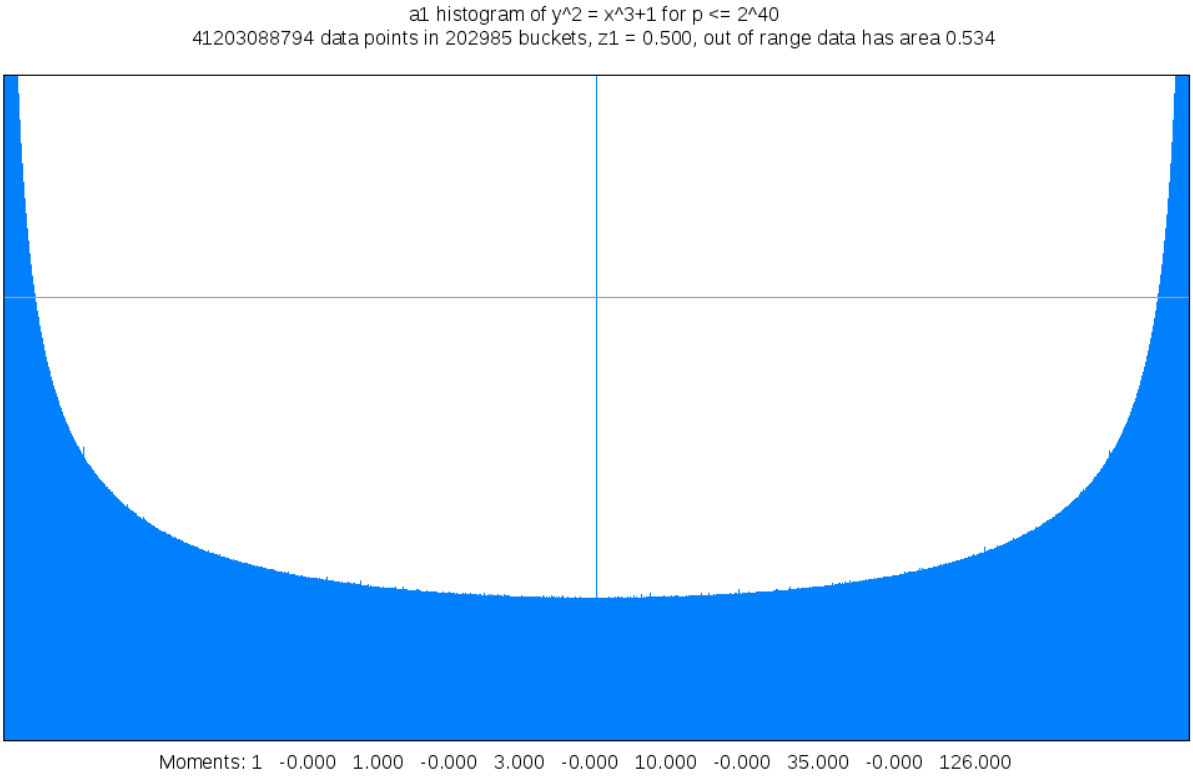

FIGURE 3. Click image to animate (requires Adobe Reader), or visit this web page.

latter of which is stated in terms of L-functions. The Sato-Tate conjecture is obtained by applying **T2** to all the powers of a fixed elliptic curve $E/\mathbf{Q}$ (as products of abelian varieties); see [46] for an overview.

Let G be the compact group $\mathrm{SU}(2)$ of 2×2 unitary matrices with determinant 1. The irreducible representations of G are the mth symmetric powers ρ_m of the natural representation ρ_1 of degree 2 given by the inclusion $\mathrm{SU}(2) \subseteq \mathrm{GL}_2(\mathbf{C})$.

Each element of $X := \mathrm{conj}(G)$ can be uniquely represented by a matrix of the form

$$\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix},$$

where $\theta \in [0, \pi]$ is the eigenangle of the conjugacy class. It follows that each $f \in C(X)$ can be viewed as a continuous function $f(\theta)$ on the compact set $[0, \pi]$. As noted in Lecture 1, the pushforward of the Haar measure on G to X is the measure

$$\mu = \frac{2}{\pi} \sin^2 \theta \, d\theta,$$

which we means that for each $f \in C(X)$ we have

$$\mu(f) = \int_0^\pi f(\theta) \sin^2 \theta \, d\theta.$$

Now let $E/\mathbf{Q}$ be an elliptic curve without CM, let $(p_1, p_2, p_3, \ldots)$ be the primes that do not divide the conductor N of E (in order), and define $x_i \in X$ to be the element of X corresponding to the unique

$\theta_i \in [0,\pi]$ for which we have $t_i = 2\cos\theta_i\sqrt{p_i}$, where $t_i := p_i + 1 - \#E(\mathbf{F}_{p_i})$ is the trace of Frobenius of the reduced curve E_{p_i}.

We are now in the setting of §2.3. We have a compact group $G := \mathrm{SU}(2)$, its space of conjugacy classes $X := \mathrm{conj}(G)$, a number field $K = \mathbf{Q}$, a sequence $(p_1, p_2, p_3, \ldots)$ containing all but finitely many primes of K ordered by norm, and a representation $\rho_m \colon G \to \mathrm{GL}_2(\mathbf{C})$, for each $m \geq 1$. The corresponding L-function is

$$L(\rho_m, s) := \prod_{j=1}^{\infty} \det(1 - \rho_m(x_j)p_j^{-s})^{-1} = \prod_{j=1}^{\infty}\prod_{k=0}^{m}(1 - e^{i(m-2k)\theta_j}p_j^{-s})^{-1}.$$

For each $p \nmid N$, define $\alpha_p := e^{i\theta_j}p_j^{1/2}$, where $p_j = p$. If we put $\alpha_j := e^{i\theta_j}p_j^{1/2}$ and define

$$L_m^1(s) = \prod_{p\nmid N}\prod_{j=0}^{m}(1 - \alpha_p^{m-j}\bar{\alpha}_p^j p^{-s})^{-1},$$

then for $m \geq 1$ we have

$$L_(\rho_m, s) = L_m^1(s - m/2).$$

Tate conjectured in [61] that $L_m^1(s)$ is holomorphic and nonvanishing on $\mathrm{Re}(s) \geq 1 + m/2$, which implies that each $L(\rho_m, s)$ is holomorphic and nonvanishing on $\mathrm{Re}(s) \geq 1$. Assuming this is true, Theorem 2.12 implies that the sequence (x_i) is μ-equidistributed. This is the Sato-Tate conjecture.

We now recall the *Modularity theorem* for elliptic curves $E/\mathbf{Q}$, which states that there is a one-to-one correspondence between isogeny classes of elliptic curves $E/\mathbf{Q}$ of conductor N and modular forms

$$f(z) = \sum_{n\geq 1} a_n e^{2\pi i n z} \in S_2(\Gamma_0(N))^{\mathrm{new}},$$

normalized so that $a_1 = 1$. Here $S_2(\Gamma_0(N))^{\mathrm{new}}$ denotes the space of cuspforms of weight 2 and level N (with trivial nebentypus) that are new at level N (they don't belong to $S_2(\Gamma_0(M))$ for any $M|N$).

The modular form $f(z)$ is a simultaneous eigenform for all the Hecke operators T_n, and for each prime $p \nmid N$ the eigenvalue a_p is equal to the trace of Frobenius of the reduced curve E_p. Moreover, the a_p also agree at the bad primes, that is, we have an equality of L-functions

$$L(E,s) = L(f,s) := \prod_p (1 - a_p p^{-s} + p^{1-2s})^{-1} = \prod_p\prod_{j=0}^{1}(1 - \alpha_p\bar{\alpha}_p p^{-s})^{-1},$$

where α_p and $\bar{\alpha}_p$ are the roots of $T^2 - a_p T + p$. This theorem was proved for semistable elliptic curves by Taylor and Wiles [62, 69], and then extended to all elliptic curves over $\mathbf{Q}$ by Breuil, Conrad, Diamond, and Taylor [7].

Theorem 2.17 (Barnet–Lamb, Gerhaghty, Harris, Taylor)**.** *Let $f(z) := \sum_{n\geq 1} a_n e^{2\pi i z n} \in S_2(\Gamma_0(N))$ be a weight 2 cuspform of level N, normalized so that $a_1 = 1$. For each prime $p \nmid N$ let $\alpha_p, \bar{\alpha}_p$ be the roots of $T^2 - a_p T + p$. If f does not have CM then*

$$\prod_{p\nmid N}\prod_{j=0}^{m}(1 - \alpha_p^{m-j}\bar{\alpha}_p^j p^{-s})^{-1}$$

is holomorphic and nonvanishing on $\mathrm{Re}(s) \geq 1$.

Proof. See [6, Theorem B]. □

Corollary 2.18. *The Sato-Tate conjecture holds.*

2.6. **Exercises.**

Exercise 2.1. Let X be a compact Hausdorff space. Show that the only set $S \subseteq X$ that are μ-quarrable for every measure μ on X are the sets that are both open and closed.

Exercise 2.2. Prove Proposition 2.7.

Exercise 2.3. Let G be a compact commutative Lie group (written multiplicatively) containing an element z such that the set $\{z^n : n \in \mathbf{Z}_{\geq 1}\}$ is dense in G. Show that the sequence $(z, z^2, z^3, \ldots)$ is equidistributed with respect to the Haar measure on G.

References

[1] J. Achter and J. Holden, *Notes on an analogue of the Fontaine-Mazur conjecture*, Journal de Théorie des Nombres de Bordeaux **15** (2003), 627–637.

[2] O. Ahmadi and I. E. Shparlinski, *On the distribution of the number of points on algebraic curves in extensions of finite fields*, Mathematical Research Letters **17** (2012), 689–699.

[3] G. Banaszak and K. S. Kedlaya, *An algebraic Sato-Tate group and Sato-Tate conjecture*, Indiana University Mathematics Journal **64** (2015), 245–274.

[4] G. Banaszak and K. S. Kedlaya, *Motivic Serre group, algebraic Sato-Tate group and Sato-Tate conjecture*, in *Frobenius Distributions on Curves*, Contemporary Mathematics, AMS, to appear.

[5] T. Barnet-Lamb, D. Geraghty, and T. Gee, *The Sato-Tate conjecture for Hilbert modular forms*, Journal of the AMS **24** (2011), 411–469.

[6] T. Barnet-Lamb, D. Geraghty, M. Harris, and R. Taylor, *A family of Calabi-Yau varieties and potential automorphy II*, Publications of the Research Institute for Mathematical Sciences **47** (2011), 29–98.

[7] C. Breuil, B. Conrad, F. Diamond, and R. Taylor, *On the modularity of elliptic curves over* $\mathbf{Q}$*: wild 3-adic exercises*, Journal of the AMS **14** (2001), 843–939.

[8] A. Bucur and K. S. Kedlaya, *An application of the effective Sato-Tate conjecture*, in *Frobenius Distributions on Curves*, Contemporary Mathematics, AMS, to appear.

[9] J. W. S. Cassels and A. Fröhlich, *Algebraic number theory*, second edition, London Mathematical Society, 2010.

[10] W. Castryck, A. Folsom, H. Hubrechts, and A. V. Sutherland, *The probability that the number of points on the Jacobian of a genus 2 curve is prime*, Proceedings of the London Mathematical Society **104** (2012), 1235–1270.

[11] A. C. Cojocaru, R. Davis, A. Silverberg, and K.Ė. Stange, *Arithmetic properties of the Frobenius traces defined by a rational abelian variety*, with two appendices by J-P. Serre, arXiv:1504.00902, 2015.

[12] P. Deligne, *La conjecture de Weil: I*, Publications Mathématiques IHÉS **43** (1974), 273–307.

[13] P. Deligne, *La conjecture de Weil: II*, Publications Mathématiques IHÉS **52** (1980), 173–252.

[14] J. Diestel and A. Spalsbury, *The joys of Haar measure*, Graduate Studies in Mathematics **150**, AMS, 2014.

[15] F. Fité, *Equidistribution, L-functions, and Sato–Tate groups*, Contemporary Mathematics **649** (2015), 63–88.

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

[17] F. Fité, K. S. Kedlaya, and A. V. Sutherland, *Sato-Tate groups of some weight 3 motives*, in *Frobenius Distributions on Curves*, Contemporary Mathematics **663**, AMS, to appear.

[18] F. Fité and A. V. Sutherland, *Sato-Tate distributions of twists of* $y^2 = x^5 - x$ *and* $y^2 = x^6 + 1$, Algebra and Number Theory **8** (2014), 543–585.

[19] F. Fité and A. V. Sutherland, *Sato-Tate groups of* $y^2 = x^8 + c$ *and* $y^2 = x^7 - cx$, in *Frobenius Distributions on Curves*, Contemporary Mathematics **663**, AMS, to appear.

[20] Joachim von zur Gathen and Jürgen Gerhard, *Modern computer algebra*, 3rd edition, Cambridge University Press, 2013.

[21] D. J. Grabiner and P. Magyar, *Random walks in Weyl chambers and the decomposition of tensor powers*, Journal of Algebraic Combinatorics **2** (1993), 239–260.

[22] M. Harris, N. Shepherd-Barron, and R. Taylor, *A family of Calabi-Yau varieties and potential automorphy*, Annals of Mathematics **171** (2010), 779–813.

[23] W. B. Hart, *Fast Library for Number Theory: An introduction*, in *Proceedings of the Third International Congress on Mathematical Software (ICMS 2010)*, LNCS 6327, Springer, 2010, 88–91.

[24] W. B. Hart, F. Johansson and S. Pancratz, *Fast Library for Number Theory*, version 2.5.2, http://flintlib.org, 2015.
[25] D. Harvey, *Kedlaya's algorithm in larger characteristic*, International Mathematics Research Notices **2007**.
[26] D. Harvey, *Counting points on hyperelliptic curves in average polynomial time*, Annals of Mathematics **179** (2014), 783–803.
[27] D. Harvey, *Computing zeta functions of arithmetic schemes*, Proceedings of the London Mathematical Society, to appear.
[28] D. Harvey and A. V. Sutherland, *Computing Hasse-Witt matrices of hyperelliptic curves in average polynomial time*, in *Algorithmic Number Theory 11th International Symposium (ANTS XI)*, LMS Journal of Computation and Mathematics **17** (2014), 257–273.
[29] D. Harvey and A. V. Sutherland, *Computing Hasse-Witt matrices of hyperelliptic curves in average polynomial time II*, in *Frobenius Distributions on Curves*, Contemporary Mathematics **663**, AMS, to appear.
[30] H. Heyer, *Probability measures on locally compact groups*, Springer, 1977.
[31] C. Johansson, *On the Sato-Tate conjecture for non-generic abelian surfaces*, with an appendix by Francesc Fité, Transactions of the AMS, to appear.
[32] I. Kaplansky, *Lattices of continuous functions*, Bulletin of the AMS **6** (1947), 617–623.
[33] N. M. Katz and P. Sarnak, *Random matrices, Frobenius eigenvalues, and monodromy*, Colloquium Publications **45**, AMS, 1999.
[34] K. S. Kedlaya, *Counting points on hyperelliptic curves using Monsky-Washnitzer cohomology*, Journal of the Ramanujan Mathematical Society **16** (2001), 323–338.
[35] K. S. Kedlaya, *Computing zeta functions via p-adic cohomology*, in *Algorithmic Number Theory 6th International Symposium (ANTS VI)*, Lecture Notes in Computer Science **3076**, Springer 2004, 1–17.
[36] K. S. Kedlaya and A. V. Sutherland, *Computing L-series of hyperelliptic curves*, in *Algorithmic Number Theory 8th International Symposium (ANTS VIII)*, Lecture Notes in Computer Science **5011**, Springer, 2008, 312–326.
[37] K. S. Kedlaya and A. V. Sutherland, *Hyperelliptic curves, L-polynomials, and random matrices*, in *Arithmetic Geometry, Cryptography, and Coding Theory (AGCCT-11)*, Contemporary Mathematics **487**, American Mathematical Society, 2000, 119–162.
[38] J. Klüners and G. Malle, *A database for field extensions of the rationals*, LMS J. Comput. Math. **4** (2001), 182–196.
[39] P. Koosis, *The logarithmic integral I*, Cambridge University Press, 1998.
[40] S. Lang and H. Trotter, *Frobenius distributions in GL_2-extensions*, Lecture Notes in Mathematics **504** (1976), Springer.
[41] M. R. Murty and V. K. Murty, *Non-vanishing of L-functions and applications*, Modern Birkhäuser Classics, 1997.
[42] V. K. Murty, *Explicit formulae and the Lang-Trotter conjecture*, Rocky Mountain Journal of Mathematics **15** (1985), 535–551.
[43] J. Neukirch, *Algebraic number theory*, Springer, 1999.
[44] OEIS Foundation Inc., *The On-Line Encyclopedia of Integer Sequences*, online database at http://oeis.org, 2016.
[45] J. Pila, *Frobenius maps of abelian varieties and finding roots of unity in finite fields*, Mathematics of Computation **55** (1990), 745–763.
[46] D. Ramakrishnan, *Remarks on the Tate Conjecture for beginners*, notes from the AIM Tate Conjecture Workshop, available at http://www.aimath.org/WWN/tateconjecture/tateconjecture.pdf, 2007.
[47] R. Schoof, *Elliptic curves over finite fields and the computation of square roots mod p*, Mathematics of Computation **44** (1995), 483–494.
[48] R. Schoof, *Counting points on elliptic curves over finite fields*, Journal de Théorie des Nombres de Bordeaux **7** (1995), 219–254.
[49] J.-P. Serre, *Abelian ℓ-adic representations and elliptic curves*, Research Notes in Mathematics **7**, A.K. Peters, 1998.
[50] J.-P. Serre, *Résumé des cours de 1985-1986, Annuaire du Collège de France, 1986, 95–99*; in *Oeuvres – Collected Papers, Volume IV*, Springer, 2003, 33–37.
[51] J.-P. Serre, *Lettre à Marie-France Vignéras du 10/2/1986*, in *Oeuvres – Collected Papers, Volume IV*, Springer, 2003, 38–55.
[52] J.-P. Serre, *Propriétés conjecturales des groups de Galois motiviques et des représentations ℓ-adiques*, in *Motives*, AMS Proceedings of Symposia in Pure Mathematics **55** (1994), 377–400.
[53] J.-P. Serre, *Lectures on $N_X(p)$*, Research Notes in Mathematics **11**, CRC Press, 2012.
[54] V. Shoup, *NTL: A Library for doing Number Theory*, version 9.6.4, available at http://www.shoup.net/ntl/, 2016.
[55] J. H. Silverman, *Advanced topics in the arithmetic of elliptic curves*, Springer, 1994.
[56] J. H. Silverman, *The arithmetic of elliptic curves*, second ed., Springer, 2009.
[57] R. Stanley, *Catalan numbers*, Cambridge University Press, 2015.

[58] A. V. Sutherland, `smalljac`, version 5.0, available at `http://math.mit.edu/~drew`, 2016.
[59] A. V. Sutherland, *Order computations in generic groups*, PhD thesis, Massachusetts Institute of Technology, 2007.
[60] A. V. Sutherland, *Structure computation and discrete logarithms in finite abelian p-groups*, Mathematics of Computation **80** (2011), 477–500.
[61] J. Tate, Algebraic cycles and poles of zeta functions, in *Arithmetical Algebraic Geometry (Proc. Conf. Purdue Univ., 1963)*, Harper & Row, New York, 1965.
[62] R. Taylor and A. Wiles, *Ring-theoretic properties of certain Hecke algebras*, Annals of Mathematics **141** (1995), 553–572.
[63] J. Thorner, *The error term in the Sato-Tate conjecture*, arXiv:1407.2656v2, 2015.
[64] P. van Wamelen, *On the CM character of the curves* $y^2 = x^q - 1$, Journal of Number Theory **64** (1997), 59–83.
[65] A. Weil, *Sur les courbes algébriques et les variétés qui s'en déduisent*, Publ. Inst. Math. Univ. Stasbourg **7** (1945).
[66] A. Weil, *Variétés abéliennes et courbes algébriques*, Publ. Inst. Math. Univ. Stasbourg **8** (1946).
[67] A. Weil, *Numbers of solutions of equations in finite fields*, Bull. Amer. Math. Soc. **55** (1949), 497–508.
[68] H. Weyl, *The classical groups: their invariants and representations*, Princeton University Press, 1966.
[69] A. Wiles, *Modular elliptic curves and Fermat's last theorem*, Annals of Mathematics **141** (1995), 443–551.