SIMPLE THINGS WE DON'T KNOW

NICHOLAS M. KATZ

Abstract. This is a quite faithful rendering of a Colloquio De Giorgi I had the honor to give at Scuola Normale Superiore on March 21, 2012. The idea was to explain some open problems in arithmetic algebraic geometry which are simple to state but which remain shrouded in mystery.

1. An interactive game: dimension zero

Suppose I give you an integer $N \geq 2$, and tell you that I am thinking of a monic integer polynomial $f(X) \in \mathbb{Z}[X]$ whose discriminant $\Delta(f)$ divides some power of N. I tell you further, for every prime number p not[1] dividing N, the number

$$n_p(f) := \#\{x \in \mathbb{F}_p | f(x) = 0 \text{ in } \mathbb{F}_p\}$$

of its solutions in the prime field $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$. You must then tell me the degree of the polynomial f.

In this "infinite" version, where I tell you the $n_p(f)$ for every good prime, your task is simple; the degree of f is simply the largest of the $n_p(f)$. Indeed, $n_p(f) = \deg(f)$ precisely when p is a prime which splits completely in the number field $K_f := \mathbb{Q}(\text{the roots of f})$. By Chebotarev, this set of primes is infinite, and has density $1/\#Gal(K_f/\mathbb{Q})$.

If you do not have infinite patience, you may hope that you can specify a constant X_N, depending only on N, such that it will be enough for me to tell you $n_p(f)$ only for the good primes which are $\leq X_N$. Alas, this cannot be done. Whatever constant X_N you choose, I will pick an integer $a \geq 2$ such that $N^a > X_N$, and take for my f the cyclotomic polynomial $\Phi_{N^a}(X)$, whose roots are the primitive N^a'th roots of unity. With this choice of f, we have $n_p(f) = 0$ for all good primes $p \leq N^a$. Indeed, with this choice of f, $n_p(f)$ vanishes for a good prime p **unless** $p \equiv 1 \mod N^a$, in which case $n_p(f) = \deg(f) (= \phi(N^a) = N^{a-1}\phi(N)$, ϕ being Euler's ϕ function.). But the condition that p be congruent to 1 mod N^a certainly forces $p > N^a$. In particular, we have $n_p(f) = 0$ for all good primes $p \leq N^a$.

[1]We will call such a prime a good prime (for this problem).

2. An interactive game: curves

In this game, I give you an integer $N \geq 2$, and tell you that I am thinking of a (proper, smooth) curve $C/\mathbb{Z}[1/N]$ (with geometrically connected fibres). Once again I tell you, for every good prime p, the number

$$n_p(C) := \#C(\mathbb{F}_p),$$

the number of its points with values in $\mathbb{F}_p$. You must then tell me the common genus g of the geometric fibres of C.

What is your strategy? You know that, by Weil [Weil], the number $n_p(C)$ is approximately $p+1$. More precisely, if we write

$$n_p(C) = p + 1 - a_p(C),$$

so that the data of the integers $n_p(C)$ is equivalent to the data of the integers $a_p(C)$, then we have the Weil bound

$$|a_p| \leq 2g\sqrt{p}.$$

A natural guess is that you can recover the integer $2g$ as the limsup of the ratios $|a_p|/\sqrt{p}$ as p varies over all good primes. You might even hope to recover $2g$ as the limsup of the ratios $a_p/\sqrt{p}$. Or you might make the more modest guess that you can recognize $2g$ as being the largest even integer such that, on the one hand, we have $|a_p|/\sqrt{p} \leq 2g$ **but** for at least one good prime we have $|a_p|/\sqrt{p} > 2g-2$. Or you may be more ambitious and require that there are infinitely many good primes p with $|a_p|/\sqrt{p} > 2g-2$.

The sad truth is that, except in some very special cases, none of these guesses is known to be correct. Let us first discuss the two cases where something is known, namely $g=0$ and $g=1$.

In the case of genus 0, then $a_p(C) = 0$ for every good p, and all guesses are hence correct.

In the case of genus one, the modest guess that we will have $a_p(C) \neq 0$ for infinitely many good p is easy to establish. First, we may replace our genus one curve C, which may not have a $\mathbb{Q}$-rational point, by its Jacobian, without changing the number of mod p points. Now $C/\mathbb{Z}[1/N]$ has a group(scheme) structure. In particular, each set $C(\mathbb{F}_p)$ has the structure of a finite abelian group. For any prime p not dividing $3N$ which splits completely in the number field $\mathbb{Q}(C[3] :=$ the points of order 3), we know both that $C(\mathbb{F}_p)$ contains a subgroup of order 9, namely all the nine points of order dividing 3, and that p must be congruent to 1 mod 3 (this last fact because by the e_n pairing, once we have all the points of order any given n invertible in our field, that same field contains all the n'th roots of unity). So from the

equality $n_p(C) = p + 1 - a_p(C)$, we get the congruences

$$a_p(C) \equiv p + 1 \mod 9, \quad p \equiv 1 \mod 3,$$

which together give the congruence

$$a_p(C) \equiv 2 \mod 3$$

for every prime p not dividing $3N$ which splits completely in $\mathbb{Q}(C[3])$.

In the $g = 1$ case, the truth of the Sato-Tate conjecture, established for non-CM[2] elliiptic curves by Harris, Taylor et al, cf [BGHT], [CHT],[HST], [T] leads easily to a proof that the most precise guess is correct in genus one. We will explain how this works in the next sections.

3. A "baby" version of the Sato-Tate conjecture

Let us begin with quick excursion into the world of compact Lie groups. For each even integer $2g \geq 2$, we denote by $USp(2g)$ the "compact symplectic group". We can see it concretely as the intersection of the complex symplectic group $Sp(2g, \mathbb{C})$ (take the standard symplectic basis e_i, f_i, $1 \leq i \leq g$ in which $(e_i, e_j) = (f_i, f_j) = 0$ for all i, j and $(e_i, f_j) = \delta_{i,j}$) with the unitary group $U(2g)$ (where the same e_i, f_i, $1 \leq i \leq g$ form an orthonormal basis). Or we can see $USp(2g)$ as a maximal compact subgroup of $Sp(2g, \mathbb{C})$, or we can see it as the "compact form" of $Sp(2g, \mathbb{C})$.

What is relevant here is that $USp(2g)$ is given to us with a $2g$-dimensional $\mathbb{C}$-representation, and in this representation every element has eigenvalues consisting of g pairs of complex conjugate numbers of absolute value one. Consequently, every element has its trace a real number which lies in the closed interval $[-2g, 2g]$.

For any closed subgroup K of $USp(2g)$, we also have a given $2g$-dimensional representation, whose traces lie in the closed interval $[-2g, 2g]$. Out of this data, we construct a "Sato-Tate measure" μ_K, a Borel probability measure on the closed interval $[-2g, 2g]$. Here are three equivalent descriptions of the measure μ_K. In all of them, we begin with the Haar measure $\mu_{Haar,K}$ on K of total mass one. We have the trace, which we view as a continuous map

$$\text{Trace} : K \to [-2g, 2g].$$

In the fancy version, we define $\mu_K := \text{Trace}_\star(\mu_{Haar,K})$. More concretely, for any continuous $\mathbb{R}$-valued function f on the closed interval $[-2g, 2g]$,

[2]In the CM case, Deuring proved that we are dealing with a Hecke character, and the required equidistribution for these goes back to Hecke.

we impose the integration formula

$$\int_{[-2g,2g]} f d\mu_K := \int_K f(\mathrm{Trace}(k)) d\mu_{Haar,K}.$$

For an interval $I \subset [-2g, 2g]$, indeed for any Borel-measurable set $I \subset [-2g, 2g]$, its measure is given by

$$\mu_K(I) := \mu_{Haar,K}(\{k \in K |\ \mathrm{Trace}(k) \in I\}).$$

With these definitions in hand, we can state the "baby"[3] Sato-Tate conjecture.

Conjecture 3.1. *Given an integer $N \geq 2$, and a projective smooth curve $C/\mathbb{Z}[1/N]$ with geometrically connected fibres of genus $g \geq 1$, there exists a compact subgroup $K \subset USp(2g)$ such that the sequence $\{a_p(C)/\sqrt{p}\}_{\text{good } p}$ is equidistributed in $[-2g, 2g]$ for the measure μ_K.*

This equidistribution means that for any continuous $\mathbb{R}$-valued function f on the closed interval $[-2g, 2g]$, we have the integration formula

$$\int_{[-2g,2g]} f d\mu_K = \lim_{X\to\infty} (1/\pi_{\text{good}}(X)) \sum_{p\leq X, p \text{ good}} f(a_p(C)/\sqrt{p}),$$

where we have written $\pi_{\text{good}}(X)$ for the number of good primes up to X.

We now explain how the truth of this baby Sato-Tate conjecture for a given curve $C/\mathbb{Z}[1/N]$ implies that

$$2g = limsum_{\text{good } p}\ \ a_p(C)/\sqrt{p}.$$

We must show that for any real $\epsilon > 0$, there are infinitely many good primes p for which $a_p(C)/\sqrt{p}$ lies in the interval $(2g-\epsilon, 2g]$. For this we argue as follows. In any probability space, there are at most countably many points ("atoms") which have positive measure, cf. [Feller, page 135]. So at the expense of replacing the chosen ϵ by a smaller one, we may further assume that the point $2g - \epsilon$ is not an atom for the measure μ_K[4]. The the open set $(2g - \epsilon, 2g]$ has a boundary of measure zero, and hence [Serre, Prop. 1, I-19] we may apply the integration

[3] "Baby" because it is the trace consequence of the "true" Sato-Tate conjecture, which we will not go into here. See [FKRS] for a plethora of numerical evidence in the case $g = 2$.

[4] The point $2g$ is never an atom for the measure μ_K. Indeed, the only element of a compact subgroup $K \subset USp(2g)$ with trace $2g$ is the identity, and this point has positive measure in K if and only if K is finite. But if K were finite, then equidistribution would imply that $a_p(C)/\sqrt{p} = 2g$ for infinitely many good primes p. But the equality $a_p(C) = 2g\sqrt{p}$ holds for no p, simply because a_p is an integer, while $2g\sqrt{p}$ is not.

formula above to the characteristic function of this open set. Thus we get

$$\mu_K((2g-\epsilon, 2g] = \lim_{X\to\infty} \frac{\#\{p \le X, p \text{ good}, \ \mathrm{a_p(C)}/\sqrt{\mathrm{p}} > 2\mathrm{g} - \epsilon\}}{\#\{p \le X, p \text{ good}\}}.$$

But the set $\{k \in K, \mathrm{Trace}(k) > 2g - \epsilon\}$ is open in K and contains the identity, so has strictly positive mass for Haar measure; this mass is, by definition, the μ_K measure of $(2g - \epsilon, 2g]$. Thus we have

$$\lim_{X\to\infty} \frac{\#\{p \le X, p \text{ good}, \ \mathrm{a_p(C)}/\sqrt{\mathrm{p}} > 2\mathrm{g} - \epsilon\}}{\#\{p \le X, p \text{ good}\}} > 0,$$

and hence there are infinitely many good p for which $a_p(C)/\sqrt{p} > 2g - \epsilon$[5].

4. Integrality consequences

For any compact subgroup $K \subset USp(2g)$, the moments $M_{n,K} := \int_{[-2g,2g]} x^n d\mu_K$, $n \ge 0$, of the measure μ_K are nonnegative integers. Indeed, the n'th moment $M_{n,K}$ is the integral $\int_K (\mathrm{Trace}(k))^n d\mu_{Haar,K}$, which is the multiplicity of the trivial representation $\mathbb{1}$ in the n'th tensor power $std_{2g}^{\otimes n}$ of the given $2g$-dimensional representation std_{2g} of K. So the baby Sato-Tate conjecture predicts that for our curve $C/\mathbb{Z}[1/N]$, for each integer $n \ge 0$, the sums

$$(1/\pi_{\text{good}}(X)) \sum_{p\le X, p \text{ good}} (a_p(C)/\sqrt{p})^n$$

not only have a limit as $X \to \infty$, but also that this limit is a nonnegative integer. Indeed, the baby Sato-Tate conjecture for $C/\mathbb{Z}[1/N]$ holds with a specified compact subgroup $K \subset USp(2g)$ if an only if[6] the above limit exist for each $n \ge 0$ and is equal to the n'th moment $M_{n,K}$.

---

[5]If we keep our original ϵ, and allow the possibility that $2g - \epsilon$ is an atom, we can argue as follows. We take a continuous function f with values in $[0, 1]$ which is 1 on the interval $[2g - \epsilon/2, \ 2g]$ and which is 0 in $[-2g, 2g - \epsilon]$. Then we have the inequality

$$\liminf_{X\to\infty} \frac{\#\{p \le X, p \text{ good}, \ \mathrm{a_p(C)}/\sqrt{\mathrm{p}} > 2\mathrm{g} - \epsilon\}}{\#\{p \le X, p \text{ good}\}} \ge$$

$$\ge \int_{[-2g,2g]} f d\mu_K \ge \int_{(2g-\epsilon/2, \ 2g]} f d\mu_K = \mu_K((2g - \epsilon/2, \ 2g]) > 0$$

and we conclude as above.

[6]On a closed interval, a Borel probability measure is determined by its moments.

5. Some test cases

There is a conjectural recipe for the compact subgroup $K \subset USp(2g)$ attached to a given $C/\mathbb{Z}[1/N]$ in terms of the ℓ-adic representation of $Gal(\overline{\mathbb{Q}}/\mathbb{Q})$ on the ℓ-adic Tate module of the Jacobian of $C/\mathbb{Z}[1/N]$. We will not go into that recipe here, except to say that it predicts that we will have $K = USp(2g)$ precisely when this ℓ-adic representation has an open image in the group $GSp(2g, \mathbb{Q}_\ell)$ of symplectic similitudes. A marvelous theorem [Zarhin] of Zarhin asserts that this is the case for any hyperelliptic curve $y^2 = h(x)$ with $h(x) \in \mathbb{Q}[x]$ a polynomial of degree $n \geq 5$ whose Galois group over $\mathbb{Q}$ is either the alternating group A_n or the full symmetric group S_n.

When $K = USp(2g)$, all the odd moments vanish, and one knows exact formulas for the first few even moments: $M_{0,USp(2g)} = 1$, and for $0 < 2k \leq 2g$ one has

$$M_{2k,USp(2g)} = 1 \times 3 \times ... \times (2k-1).$$

By a theorem of Schur, the truncated exponential series $e_n(x) := \sum_{0 \leq k \leq n} x^k/k!$ has Galois group A_n if 4 divides n, and S_n otherwise, cf. [Coleman] for a beautiful exposition of Schur's theorem. Moreover, Coleman shows that the discriminant of $e_n(x)$ is $(-1)^{n(n-1)/2}(n!)^n$, so the only bad primes for $y^2 = e_n(x)$, whose genus is Floor((n−1)/2), are those $p \leq n$. By a theorem of Osada [Osada, Cor. 3], the galois group of $h_n(x) := x^n - x - 1$ is S_n. The discriminant of $h_n(x)$ has much less regular behavior. It grows very rapidly, and often is divisible by huge primes[7]. So the bad primes for $y^2 = h_n(x)$, of genus Floor((n − 1)/2), are somewhat erratic. In any case, for either of these curves, the baby Sato-Tate conjecture predicts that the sums

$$(1/\pi_{\text{good}}(X)) \sum_{p \leq X, p \text{ good}} (a_p(C)/\sqrt{p})^d$$

tend to 0 for d odd, and for $d = 2k \leq 2g$ tend to

$$M_{2k,USp(2g)} = 1 \times 3 \times ... \times (2k-1).$$

There is no integer $n \geq 5$ for which either of these statements is known, either for Schur's curve $y^2 = e_n(x)$ or for Osada's curve $y^2 = h_n(x)$.

Another striking but unknown consequence of baby Sato-Tate for these curves is this. Because the group $USp(2g)$ contains the scalar -1, the measure $\mu_{USp(2g)}$ on $[-2g, 2g]$ is invariant under $x \mapsto -x$. So for

[7]For example, with $n = 17$, the discriminant is the prime 808793517812627212561, for $n = 22$ the prime factorization of the discriminant is $5 \times 69454092876521107983605569601$.

these two curves, the sets $\{\text{good } p, a_p(C) > 0\}$ and $\{\text{good } p, a_p(C) < 0\}$ should[8]each have Dirichlet density $1/2$.

Much remains to be done.

References

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

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

[Coleman] Coleman, R., On the Galois groups of the exponential Taylor polynomials. Enseign. Math. (2) 33 (1987), no. 3-4, 183-189.

[Feller] Feller, W.,An Introduction to Probability Theory and its Applications, Volume II, John Wiley and Sons, 1966.

[FKRS] Fité, F.,Kedlaya, K., Rotger, V., Sutherland, A., Sato-Tate distributions and Galois endomorphism modules in genus 2, arXiv:1110.6638, to appear in Compositio.

[HST] Harris, M., Shepherd-Barron, N., Taylor, R., A family of Calabi-Yau varieties and potential automorphy. Ann. of Math. (2) 171 (2010), no. 2, 779-813.

[Osada] Osada, H., The Galois groups of the polynomials $X^n + aX^\ell + b$, J. Number Th. 25 (1987), 230-238.

[Serre] Serre, J.-P., Abelian ℓ-adic Representations, Addison-Wesley, 1989.

[T] Taylor, R., Automorphy for some ℓ-adic lifts of automorphic mod ℓ Galois representations. II. Publ. Math. Inst. Hautes Études Sci. No. 108 (2008), 183-239.

[Weil] Weil, Andé, Variétés abéliennes et courbes algébriques. Actualités Sci. Ind., no. 1064 = Publ. Inst. Math. Univ. Strasbourg 8 (1946). Hermann & Cie., Paris, 1948. 165 pp.

[Zarhin] Zarhin, Yuri G., Very simple 2-adic representations and hyperelliptic Jacobians, Mosc.Math. J. 2 (2002), no. 2, 403-431.

Princeton University, Mathematics, Fine Hall, NJ 08544-1000, USA
E-mail address: nmk@math.princeton.edu

[8]The measure $\mu_{USp(2g)}$ has no atoms. In particular 0 is not an atom.