BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 46, Number 3, July 2009, Pages 413–457
S 0273-0979(09)01257-9
Article electronically published on March 27, 2009

LANG–TROTTER REVISITED

NICHOLAS M. KATZ

Dedicated to the memory of Serge Lang

Abstract. The Lang–Trotter Conjecture(s) concern elliptic curves over the field $\mathbb{Q}$ of rational numbers. We first explain the broader number-theoretic context into which they fit. Then we turn to formulating their "function field" analogues. We explain how these analogues can be proven in some very special cases, and we speculate about what might be true in the general function field case.

Table of contents

0. Preface

The Lang–Trotter Conjecture(s), first published in 1976 [L-T] but formulated a few years earlier, specifically concern elliptic curves over the field $\mathbb{Q}$ of rational numbers. These conjectures are best understood in a much broader context of what "should" be true, and of what might be true. We discuss this context at length in the Introduction to this paper; indeed, we don't state any versions of the conjectures themselves until we are two-thirds of the way through the Introduction. After this leisurely Introduction, we turn in Section 2 to the consideration of versions of these same Lang–Trotter Conjectures, but now reformulated so that they make sense when the field $\mathbb{Q}$ is replaced by a function field over a finite field,[1] e.g. by $\mathbb{F}_p(t)$, the field of rational functions in one variable over the finite field $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$. Even in that setting there is little we can say in general.

Received by the editors December 21, 2008, and, in revised form, February 23, 2009.
2000 *Mathematics Subject Classification.* Primary 11F80, 11G05, 14G35.

[1]We do this fully mindful of the witticism that "the function field case is the last refuge of a scoundrel".

However, there are certain beautiful and long-studied elliptic curves over function fields, namely the universal elliptic curves over modular curves,[2] where it turns out that we can settle affirmatively all these function field conjectures.[3] We do this in Sections 3–6. In Section 7, we make a transition back to considering quite general elliptic curves over function fields, and their "Galois images". In Section 8 we discuss the possibility of having "exact" point count formulas in the general case, which depend only on the Galois image. This hope is inspired by Gekeler's beautiful product formula, valid for certain universal elliptic curves over modular curves (and possibly for all; this remains an open question). It turns out, thanks to an argument of Deligne, that this hope is overly optimistic in general; we end the section by asking if some asymptotic consequence of it is correct. Much remains to be done.

This paper is partly an exposition of open problems, some of which have entirely elementary statements, partly an exposition of known results, and partly an exposition of new results. We have tried to make the exposition accessible to people with a wide range of backgrounds; the reader will judge how well we have succeeded.

1. Introduction

Given a polynomial $f(X_1, ..., X_n) \in \mathbb{Z}[X_1, ..., X_n]$, the question of describing the set

$$\{x = (x_1, ..., x_n) \in \mathbb{Z}^n \mid f(x) = 0\}$$

of all[4] integer solutions of the equation $f = 0$ goes back at least to Diophantus, some 1750 years ago. Here one wants to prove either that a) there are no solutions, or b) there are only finitely many solutions (and ideally specify both how many and how large) or c) there are infinitely many solutions (and ideally give an asymptotic formula for how many there are of "size" at most h, as $h \to \infty$). Thus, for example, Fermat's Last Theorem was a problem of type a), the Mordell Conjecture of type b), and Pell's equation of type c).

Sometimes one can prove the nonexistence of solutions by finding either an Archimedean obstruction or a congruence obstruction. For example, the equation

$$x^2 + y^2 + 691 = 0$$

has no integer solutions because it has no $\mathbb{R}$ solutions; the equation

$$x^2 + y^2 = 691,$$

and more generally any equation of the form

$$x^2 + y^2 = 4n + 3,$$

has no integer solutions because it has no solutions mod 4; and the equation

$$y^2 + x^4 + 2 = 0$$

has no integer solutions both because it has no $\mathbb{R}$ solutions and because it has no mod 5 solutions.

[2]Perhaps the simplest example is this: the ground field is $\mathbb{F}_p(t)$, any odd prime p, and the elliptic curve has the equation $y^2 = (x + t)(x^2 + x + t)$. This is the universal curve with a point of order 4, namely the point $(0, t)$.

[3]Unfortunately, these universal elliptic curves over modular curves seem to have no analogue in the world of elliptic curves over number fields.

[4]If the polynomial f is homogeneous of some degree $d \geq 1$, we allow only integer solutions $(x_1, ..., x_n) \in \mathbb{Z}^n$ with $\gcd(x_1, ..., x_n) = 1$.

Even in the possible presence of an Archimedean obstruction, it can still be interesting to ask, given f, modulo which primes p the equation $f = 0$ has an $\mathbb{F}_p$ solution. For example, the study of the equation in one variable,

$$x^2 + 1 = 0,$$

mod odd primes p, amounts to the determination of the "quadratic character of -1 mod p", and led Euler to the theorem, already stated a century earlier by Fermat, that all primes of the form $4n + 1$, but none of the form $4n - 1$, are sums of two squares. In this example, the number N_p of mod p solutions is either 0 or 2; if we write

$$N_p = 1 + a_p,$$

then $a_p = \pm 1$, and the result is that $a_p = 1$ if p is of the form $4n + 1$, and $a_p = -1$ if p is of the form $4n - 1$.

Still with this $x^2 + 1 = 0$ example, we might ask whether $a_p = 1$ (resp. $a_p = -1$) holds for infinitely many primes. That it does, for both choices of sign, amounts to the special case of Dirichlet's theorem that there are infinitely many primes in each of the two arithmetic progressions $4n \pm 1$.

Now let us consider an equation in two variables. For simplicity, we take it to be of the form

$$y^2 = h(x)$$

with $h(x) \in \mathbb{Z}[x]$ monic of some odd degree $2g + 1$, such that h has $2g + 1$ distinct zeroes in $\mathbb{C}$. The $\mathbb{C}$ solutions, together with a single "point at ∞", form a compact Riemann surface of genus g. The discriminant $\Delta \in \mathbb{Z}$ of the polynomial $h(x)$ is nonzero. For any "good" prime, i.e., any odd prime p which does not divide Δ, the $\mathbb{F}_p$-solutions of this equation, together with a single "point at ∞", form the $\mathbb{F}_p$ points $C(\mathbb{F}_p)$ of a (projective, smooth, geometrically connected) curve $C/\mathbb{F}_p$ of genus g over $\mathbb{F}_p$. In this case, for each good prime p we have

$$\#C(\mathbb{F}_p) = 1 + \#\{\mathbb{F}_p \text{ solutions of } y^2 = h(x)\},$$

and we define the integers a_p by

$$\#C(\mathbb{F}_p) = p + 1 - a_p.$$

In the $x^2 + 1$ example with its a_p, we knew a priori that a_p was either ± 1, and the two questions were a) how a_p depended on p and b) were there infinitely many p with a given choice of a_p.

In the curve case, we almost never know a "simple" rule for how a_p depends on p (short of literally computing it for each given p, more or less cleverly). We do have an Archimedean bound, the celebrated Weil bound

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

and since a curve cannot have a negative number of points, we have the Archimedean inequality

$$a_p \leq p + 1,$$

which for large genus g and small prime p, say $2g > \sqrt{p}$, does not follow from the Weil bound.

What else do we know about the numbers a_p for a given curve? Remarkably little (outside the trivial case of genus $g = 0$, where all a_p vanish), but we have a plethora

of open problems and conjectures about them, some of which have strikingly elementary formulations, or at least consequences which have strikingly elementary formulations.

Here is one example of an easy-to-state open problem. Suppose we are given the numbers $a_p/p^{1/2}$ for all good p, but are not told what curve they came from, or even its genus. By the Weil bound, we have

$$a_p/p^{1/2} \in [-2g, 2g].$$

Is it true that we can recover $2g$ as the lim sup of the numbers $|a_p|/p^{1/2}$? Or weaker, is it true that the inequality

$$|a_p|/p^{1/2} > 2g - 2$$

holds for infinitely many p? Weaker yet, does it hold for at least one good p? If this were the case, then $2g$ would be the smallest even integer such that $|a_p|/p^{1/2} \leq 2g$ for all good p.

The truth of the strong form, that $2g$ is the lim sup of the numbers $|a_p|/p^{1/2}$, is implied by a general Sato–Tate conjecture about the real numbers $a_p/p^{1/2}$ attached to a curve C of genus $g \geq 1$. To formulate it, denote by $USp(2g) \subset Sp(2g, \mathbb{C})$ a maximal compact subgroup of the complex symplectic group. [So $USp(2)$ is just $SU(2)$.] The conjecture[5] is that for a given curve C there is a compact subgroup $K \subset USp(2g)$ with the property that, roughly speaking, the numbers $a_p/p^{1/2}$ are distributed like the traces of random elements of K. More precisely, denote by dk the Haar measure on K of total mass one, and denote by

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

the trace map, for the tautological $2g$-dimensional representation of K. Any continuous function

$$F : [-2g, 2g] \to \mathbb{C}$$

gives rise to a continuous function on K by $k \mapsto F(\text{Trace}(k))$, so we can form the integral

$$\int_K F(\text{Trace}(k))dk.$$

The conjecture is that for any such F, we can compute this integral by averaging F over more and more of the $a_p/p^{1/2}$; i.e., we have the limit formula

$$\lim_{T\to\infty} \frac{\sum_{\text{good } p\leq T} F(a_p/p^{1/2})}{\#\{\text{ good } p \leq T\}} = \int_K F(\text{Trace}(k))dk.$$

If Sato–Tate holds for C, then we will recover $2g$ as the lim sup of the numbers $|a_p|/p^{1/2}$. Given a real $\epsilon > 0$, take for F a continuous $\mathbb{R}$-valued function on $[-2g, 2g]$ which is nonnegative, supported in $[2g - \epsilon, 2g]$ and identically 1 on $[2g - \epsilon/2, 2g]$. [For instance, take F piecewise linear.] Because the set

$$U_{\epsilon/2} := \{k \in K \mid \text{Trace}(k) > 2g - \epsilon/2\}$$

[5]Strictly speaking, what we are formulating is "merely" the consequence for traces of the actual Sato–Tate conjecture, which asserts the equidistribution of unitarized Frobenius conjugacy classes in the space $K^{\#}$ of conjugacy classes of K, with respect to Haar measure; cf. [Se-Mot, 13.5]. Only in genus 1 are they equivalent.

is an open neighborhood of the identity element, it has strictly positive Haar measure, and therefore the integral

$$\int_K F(\mathrm{Trace}(k))dk \geq \int_{U_{\epsilon/2}} F(\mathrm{Trace}(k))dk = \int_{U_{\epsilon/2}} dk > 0.$$

So if Sato–Tate holds, there must be infinitely many p for which $|a_p|/p^{1/2} \geq 2g - \epsilon$.

The Sato–Tate conjecture is now known for all elliptic curves over $\mathbb{Q}$ whose j-invariant is not an integer, where the group K is $SU(2)$ itself [H-SB-T, Thm. A], and is expected to hold, still with $K = SU(2)$, so long as the curve does not have complex multiplication. It has been known for elliptic curves over $\mathbb{Q}$ with complex multiplication for over fifty years, thanks to work of Deuring [Deu-CM] and Hecke [He]. In the CM case, the K is the normalizer in $SU(2)$ of its maximal torus.

In higher genus, Sato–Tate is hardly ever known.[6] For certain hyperelliptic curves $y^2 = h(x)$ as above, we can be more precise in its formulation. Denote by G the Galois group of (the splitting field $L/\mathbb{Q}$ of) the polynomial $h(x)$. If $g \geq 2$ and if G is either the full symmetric group S_{2g+1} or the alternating group A_{2g+1}, then Sato–Tate should[7] hold, with $K = USp(2g)$.

Now let us turn to considering, for a given curve C, the integers a_p themselves. Here we ask two questions. First, for which integers A will we have $A = a_p$ for infinitely many p? Second, for an A which does occur as a_p for infinitely many p, give an asymptotic formula for the number of p up to X for which $A = a_p$.

Of course these same questions make sense for other naturally occurring sequences of integers a_p. For example, if we take, instead of a curve, a projective smooth hypersurface $H \subset \mathbb{P}^{n+1}$ of degree d, then for good primes p we define integers a_p by

$$\#H(\mathbb{F}_p) = \sum_{i=0}^{n} p^i + a_p.$$

Here the Weil bound is replaced by Deligne's bound

$$|a_p| \leq \mathrm{prim}(n,d)p^{n/2},$$

with $\mathrm{prim}(n,d)$ the constant $((d-1)/d)((d-1)^{n+1} - (-1)^{n+1})$.

Or we might wish to consider the sequence $a_p = \tau(p)$, where Ramanujan's $\tau(n)$ are the coefficients in

$$q\prod_{n\geq 1}(1-q^n)^{24} = \sum_{n\geq 1}\tau(n)q^n.$$

Here we have Deligne's bound

$$|a_p| \leq 2p^{11/2}.$$

[6] However, it is (trivially) known for a genus 2 curve whose Jacobian is isogenous to $E \times E$, for an elliptic curve E for which Sato–Tate is known. For example, take $h(x) = x^3 + \lambda(x^2 + x) + 1$ to be a palindromic cubic with all distinct roots, i.e., $\lambda \neq -1, 3$. Then $C :=$ (the complete nonsingular model of) $y^2 = h(x^2)$ has its Jacobian isogenous to $E \times E$ for E the elliptic curve of equation $y^2 = h(x)$, by the two maps $C \to E$ given by $(x,y) \mapsto (x^2, y)$ and $(x,y) \mapsto (1/x^2, y/x^3)$. In particular, for each good p, the a_p's of these curves are related by $a_{p,C} = 2a_{p,E}$. This last identity has an elementary proof.

[7] There is a conjectural description of K in terms of the ℓ-adic representations attached to C, and having $K = USp(2g)$ is conjecturally equivalent to the property that for every ℓ, the ℓ-adic representation has a Zariski-dense image in $GSp(2g, \mathbb{Q}_\ell)$. That this property holds for the curves $y^2 = h(x)$ whose G is either S_{2g+1} or A_{2g+1} is a striking result of Zarhin [Z].

The Lang–Trotter approach to these questions is based in part on a simple probabilistic model. For each (good) prime p, we have an integer a_p in a finite set

$$X_p \subset \mathbb{Z}.$$

In the curve case, $X_p = \mathbb{Z} \cap [-2g\sqrt{p}, 2g\sqrt{p}]$. In the hypersurface case, $X_p = \mathbb{Z} \cap [-\operatorname{prim}(n,d)p^{n/2}, \operatorname{prim}(n,d)p^{n/2}]$. In the Ramanujan τ case, $X_p = \mathbb{Z} \cap [-2p^{11/2}, 2p^{11/2}]$.

The sets X_p are increasing, in the sense that $X_{p_1} \subset X_{p_2} \subset \mathbb{Z}$ if $p_1 \leq p_2$, and their union, in this simple model, is all of $\mathbb{Z}$. Our collection of a_p is an element in the product space

$$X := \prod_{\text{good } p} X_p.$$

We endow each X_p with counting measure, normalized to have total mass one; i.e., each point x_p in X_p has mass $1/\#X_p$.

We then endow X with the product measure. The basic idea is that, in the absence of any special information, the particular element $(a_p)_p$ of X should behave like a "random" element of X, in the sense that any "reasonable" property of elements of X which holds on a set of measure one should hold for the particular element $(a_p)_p$. For example, fix an integer A, and consider the set of points $x = (x_p)_p \in X$ that have the property that $A = x_p$ for infinitely many p. If this set has measure one, then we will "expect" that $A = a_p$ for infinitely many p, and if for some explicit function $g : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$, the set of $x = (x_p)_p \in X$ for which the asymptotic formula

$$\#\{p \leq T | A = x_p\} \sim g(T) \text{ as } T \to \infty$$

holds is a set of measure one, then we "expect" that we have the asymptotic formula

$$\#\{p \leq T | A = a_p\} \sim g(T) \text{ as } T \to \infty.$$

Let us recall the basic results which address these questions.

Lemma 1.1. *Fix $A \in \mathbb{Z}$. The following properties are equivalent.*

(1) *The set of points $x = (x_p)_p \in X$ which have the property that $A = x_p$ for infinitely many p has measure one.*
(2) *The series $\sum_p 1/\#X_p$ diverges.*

Proof. Given A, consider the set $Z_A \subset X$ of those $x = (x_p)_p \in X$ for which $A = x_p$ holds for only finitely many p. So (1) for A is the statement that this set Z_A has measure zero. This set Z_A is the increasing union of the sets

$$Z_{n,A} := \{x \in X \mid x_p \neq A \ \forall p \geq p_n\}.$$

So Z_A has measure zero if and only if each $Z_{n,A}$ has measure zero. But the measure of $Z_{n,A}$ is the product $\prod_{p \geq p_n}(1 - 1/\#X_p)$, which is zero if and only (2) holds. □

As a special case of the strong law of large numbers, we get a quantitative version of the previous result.

Lemma 1.2. *Suppose the series $\sum_p 1/\#X_p$ diverges. Fix an integer A, and an increasing sequence b_p of positive real numbers with $b_p \to \infty$ such that the series $\sum_p 1/\#X_p(b_p)^2$ converges. Then for $x \in X$ in a set of measure one, we have*

$$\#\{p \leq p_n \mid x_p = A\} = \sum_{p \leq p_n} 1/\#X_p + o(b_{p_n}).$$

Proof. This is the strong law of large numbers [Ito, Thm. 4.5.1], applied to the independent sequence of L^2 functions $\{f_p\}_p$ on X given by $f_p(x) := \delta_{x_p,A}$. The mean $E(f_p)$ of f_p is $1/\#X_p$, and its variance $V(f_p)$ is bounded above by $1/\#X_p + 1/(\#X_p)^2 \leq 2/\#X_p$. So by hypothesis the series $\sum_p V(f_p)/b_p^2$ converges. Then the strong law of large numbers tells us that on a set of measure one, we have

$$\lim_{n\to\infty} (1/b_{p_n}) \sum_{p\leq p_n} (f_p - E(f_p)) = 0.$$

Making explicit the f_p, we recover the assertion of the lemma. □

Let us see what this gives in the cases we have looked at above.

In the case of a curve C, we have $\#X_p \sim 4g\sqrt{p}$. The series $\sum_p 1/\sqrt{p}$ diverges, and one knows that

$$\sum_{p\leq T} 1/\sqrt{p} \sim \sqrt{T}/\log T.$$

Here we can take $b_p = p^{(1+\epsilon)/4}$ for any fixed real $\epsilon > 0$. So we get

$$\#\{p \leq T \mid x_p = A\} = \sum_{p\leq T} 1/\#X_p + o(T^{(1+\epsilon)/4}) \sim \sqrt{T}/4g\log T$$

on a set of measure one.

In the case of a smooth hypersurface of dimension n, we have $\#X_p \sim 2\,\mathrm{prim}(n,d)p^{n/2}$. So for $n \geq 3$, the series $\sum_p 1/\#X_p$ converges. Similarly for the Ramanujan τ, we have $\#X_p \sim 2p^{11/2}$, and again the series $\sum_p 1/\#X_p$ converges. So in both of these cases we don't expect any A to occur as a_p infinitely often.

The remaining example case is that of a smooth surface in $\mathbb{P}^3$. Here $\#X_p \sim 2\,\mathrm{prim}(2,d)p$, so the series $\sum_p 1/\#X_p$ diverges, but very slowly: one knows that

$$\sum_{p\leq T} 1/p \sim \log\log T.$$

So while the probabilistic heuristic suggests that a given A might occur infinitely often as an a_p, it also suggests that no computer experiment could ever convince us of this.

Let us now return to the case of a (projective, smooth, geometrically connected) curve $C/\mathbb{Q}$, and introduce the second heuristic on which the Lang–Trotter approach is based. This is the notion of a congruence obstruction. If a given integer A occurs as a_p for infinitely many p, then whatever the modulus $N \geq 2$, the congruence $A \equiv a_p \bmod N$ will hold for infinitely many p.

Here is the simplest example of a congruence obstruction. Take a hyperelliptic curve C of the equation $y^2 = h(x)$ with $h(x) \in \mathbb{Z}[x]$ monic of degree $2g+1 \geq 3$, with $2g+1$ distinct roots in $\mathbb{C}$. Suppose in addition that all these $2g+1$ roots lie in $\mathbb{Z}$. Then for any good (so necessarily odd) p, a_p will be even. [Here is the elementary proof, based on the character sum formula for a_p. Denote by $\chi_{\mathrm{quad},p}$ the quadratic character $\chi_{\mathrm{quad},p} : \mathbb{F}_p^\times \to \pm 1$ (so $\chi_{\mathrm{quad},p}$ takes the value 1 precisely on squares) and extend it to all of $\mathbb{F}_p$ by setting $\chi_{\mathrm{quad},p}(0) := 0$. Then for any $b \in \mathbb{F}_p$, $1 + \chi_{\mathrm{quad},p}(b)$ is the number of square roots of b in $\mathbb{F}_p$. So the number of $\mathbb{F}_p$ points on C is

$$1(\text{the point at } \infty) + \sum_{x\in\mathbb{F}_p} (1 + \chi_{\mathrm{quad},p}(h(x))) = p + 1 + \sum_{x\in\mathbb{F}_p} \chi_{\mathrm{quad},p}(h(x)).$$

So we have the formula

$$a_p = -\sum_{x\in\mathbb{F}_p} \chi_{\mathrm{quad},p}(h(x)).$$

In this formula, the reductions mod p of the $2g+1$ roots of h are the $2g+1$ distinct (because p is a good prime) elements of $\mathbb{F}_p$ at which h mod p vanishes; at all other points of $\mathbb{F}_p$, h is nonzero. So a_p is the sum of an even number $p-(2g+1)$ of nonzero terms, each ± 1, so is even.] So in this example, no odd integer A can ever be an a_p for a good prime p.

In the special case of an elliptic curve $E/\mathbb{Q}$, say with good reduction outside of some Δ, there is another visible source of congruence obstructions, namely torsion points, based on the fact that the set $E(\mathbb{Q})$ has the structure of an abelian group. Suppose that the group $E(\mathbb{Q})$ contains a point P of finite order $N \geq 2$. For every odd prime p not dividing Δ, it makes sense to reduce this point mod p, and we obtain a point of the **same** order N in $E(\mathbb{F}_p)$. Therefore N divides $\#E(\mathbb{F}_p)$, so we have the congruence

$$a_p \equiv p+1 \mod N.$$

From this congruence, we see that among odd primes p not dividing Δ, $A=1$ can never occur as a_p unless $N|p$, i.e., unless N is itself an odd prime, in which case we might have $a_p = 1$ for $p = N$, but for no other; cf. [Maz, pp. 186–188].

Let us explain briefly the general mechanism by which congruence obstructions arise. Taking for Δ the product of the primes which are bad for our curve C, we get a proper smooth curve $\mathcal{C}/\mathbb{Z}[1/\Delta]$. For each integer $N \geq 2$, we have the "mod N representation" attached to $\mathbb{C}/\mathbb{Q}$, or more precisely to its Jacobian $\mathrm{Jac}(C)/\mathbb{Q}$. This is the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on the group $\mathrm{Jac}(C)(\overline{\mathbb{Q}})[N]$ of points of order dividing N. This group is noncanonically $(\mathbb{Z}/N\mathbb{Z})^{2g}$, and it is endowed with a Galois-equivariant alternating autoduality toward the group $\mu_N(\overline{\mathbb{Q}})$ of Nth roots of unity. Because C is a proper smooth curve $\mathcal{C}/\mathbb{Z}[1/\Delta]$, the mod N representation is unramified outside of $N\Delta$, so we may view it as a homomorphism

$$\rho_N : \pi_1(\mathrm{Spec}(\mathbb{Z}[1/N\Delta])) \to GSp(2g, \mathbb{Z}/N\mathbb{Z})$$

toward the group $GSp(2g,\mathbb{Z}/N\mathbb{Z})$ of mod N symplectic similitudes. The key compatibility is that for any prime p not dividing $N\Delta$, the *arithmetic* Frobenius conjugacy class

$$\mathrm{Frob}_p \in \pi_1(\mathrm{Spec}(\mathbb{Z}[1/N\Delta]))$$

has

$$\mathrm{Trace}(\rho_N(\mathrm{Frob}_p)) \equiv a_p \bmod N, \quad \det(\rho_N(\mathrm{Frob}_p)) \equiv p \mod N.$$

Now consider the image group $\mathrm{Im}(\rho_N) \subset GSp(2g,\mathbb{Z}/N\mathbb{Z})$. If this group contains at least one element whose trace is $A \mod N$, then by Chebotarev the set of primes p not dividing $N\Delta$ for which $a_p \equiv A \mod N$ has a strictly positive Dirichlet density, so in particular is infinite. On the other hand, if the image group $\mathrm{Im}(\rho_N) \subset GSp(2g,\mathbb{Z}/N\mathbb{Z})$ contains *no* element whose trace is A mod N, then $a_p \equiv A$ mod N can hold at most for one of the finitely many primes p dividing N. It is precisely in this second case that A has a congruence obstruction at N (to having $a_p = A$ for infinitely many primes p).

Lang and Trotter conjecture[8] that, for curves, it is **only** congruence obstructions which prevent an A from being a_p infinitely often:

[8]Lang and Trotter make this conjecture explicitly only for elliptic curves.

Conjecture 1.3 (Weak Lang–Trotter)**.** *Let $C/\mathbb{Q}$ be a projective, smooth, geometrically connected curve, with good reduction outside of Δ. Given an integer A, suppose that for every modulus $N \geq 2$, A has no congruence obstruction at N; i.e., the congruence $A \equiv a_p \bmod N$ holds for infinitely many p. Then we have $A = a_p$ for infinitely many p.*

In the case of a non-CM elliptic curve E, Lang and Trotter also formulate, for any A which has no congruence obstructions, a precise conjectural asymptotic for how often A is an a_p. Given such an A, they define a nonzero real constant $c_{A,E}$ and make the following precise conjecture.

Conjecture 1.4 (Strong Lang–Trotter for elliptic curves)**.** *Let $E/\mathbb{Q}$ be a non-CM elliptic curve. Then as $T \to \infty$,*

$$\#\{p \leq T \mid a_p = A\} \sim c_{A,E}(2/\pi)\sqrt{T}/\log T.$$

Here is their recipe for the constant $c_{A,E}$. For each integer $N \geq 2$, consider the finite group

$$G_N := \mathrm{Im}(\rho_N) \subset GL(2, \mathbb{Z}/N\mathbb{Z}).$$

For each $a \in \mathbb{Z}/N\mathbb{Z}$, we have the subset $G_{N,a} \subset G_N$ defined as

$$G_{N,a} := \{\text{elements } \gamma \in G_N \text{ with } \mathrm{Trace}(\gamma) = a\},$$

whose cardinality we denote by

$$g_{N,a} := \#G_{N,a}.$$

We define

$$g_{N,\mathrm{avg}} := (1/N) \sum_{a \bmod N} g_{N,a} = (1/N)\#G_N$$

to be the average, over a, of $g_{N,a}$. For an A with no congruence obstruction, Lang and Trotter show that as N grows multiplicatively, the ratio

$$g_{N,A}/g_{N,\mathrm{avg}}$$

(which Lang and Trotter write as $Ng_{N,A}/\#G_N$) tends to a nonzero (Archimedean) limit, which they define to be $c_{A,E}$. [If we apply this recipe to an A which has a congruence obstruction, then for all sufficiently divisible N, we have $g_{N,A} = 0$, so the limit exists, but it is 0.]

In this vein, we have the following "intermediate" conjecture, for any[9] curve C of any genus $g \geq 1$ which is "strongly non-CM" in the sense that for every ℓ, the ℓ-adic representation has Zariski dense image in $GSp(2g, \mathbb{Q}_\ell)$.

Conjecture 1.5 (Intermediate Lang–Trotter)**.** *Let $C/\mathbb{Q}$ be a projective, smooth, geometrically connected curve, with good reduction outside of Δ, such that for every ℓ, the ℓ-representation has Zariski dense image in $GSp(2g, \mathbb{Q}_\ell)$. Suppose the integer A has no congruence obstruction mod any N. Then for every real $\epsilon > 0$, there exists a constant $c(C, A, \epsilon)$ such that for $T \geq c(C, A, \epsilon)$, we have*

$$\sqrt{T}^{1-\epsilon} \leq \#\{p \leq T \mid a_p = A\} \leq \sqrt{T}^{1+\epsilon}.$$

[9]Without some sort of "non-CM" hypothesis, we can have $a_p = 0$ for a set of primes p of positive Dirichlet density; cf. the example following Conjecture 1.7. Perhaps for nonzero A the conjecture remains reasonable for any $C/\mathbb{Q}$.

There are no cases whatever of a pair (C, A) for which this conjecture is known. In the case of elliptic curves, there are some results on upper bounds with $\epsilon = 1/2$, some under GRH [Se-Cheb, 8.2, Thm. 20], and some on average; cf. [Da-Pa], [Ba], [Co-Shp].

Are there other situations where one should expect congruence obstructions to be the only thing preventing a given integer A from occurring as a_p infinitely often? A natural context for this question is that of a compatible system of ℓ-adic representations of some $\pi_1(\mathrm{Spec}(\mathbb{Z}[1/\Delta]))$. Let us recall one version of this notion. We are given an integer $n \geq 1$ and, for each prime ℓ, a homomorphism

$$\rho_{\ell^\infty} : \pi_1(\mathrm{Spec}(\mathbb{Z}[1/\ell\Delta])) \to GL(n, \mathbb{Z}_\ell).$$

The compatibility condition is that for every prime p not dividing Δ, there is an integer polynomial $P_p(T) \in \mathbb{Z}[T]$ such that for every prime $\ell \neq p$, the reversed characteristic polynomial

$$\det(1 - T\rho_{\ell^\infty}(\mathrm{Frob}_p)) \in \mathbb{Z}_\ell[T]$$

lies in $\mathbb{Z}[T]$ and is equal to $P_p(T)$. We are then interested in the $a_p := \mathrm{Trace}(\mathrm{Frob}_p)$ (trace in any ℓ-adic representation with $\ell \neq p$) for good (i.e., prime to Δ) primes p. Reducing mod powers ℓ^ν of ℓ, we get the representations

$$\rho_{\ell^\nu} : \pi_1(\mathrm{Spec}(\mathbb{Z}[1/\ell\Delta])) \to GL(n, \mathbb{Z}/\ell^\nu\mathbb{Z}).$$

Putting these together, we get for each integer $N \geq 2$ a mod N representation

$$\rho_N : \pi_1(\mathrm{Spec}(\mathbb{Z}[1/N\ell\Delta])) \to GL(n, \mathbb{Z}/N\mathbb{Z}).$$

Exactly as in the case of curves, A has no congruence obstruction at N, i.e., $A \equiv a_p$ mod N holds for infinitely many p, if and only if there is an element in the image group $\mathrm{Im}(\rho_N) \subset GL(n, \mathbb{Z}/N\mathbb{Z})$ whose trace is A mod N. In this case the set of p for which $A \equiv a_p$ mod N has positive Dirichlet density.

In the case of curves, these representations are "pure of weight-1" in the sense that for each good p, when we factor $P_p(T) = \prod_i(1 - \alpha_i T)$ over $\mathbb{C}$, each α_i has $|\alpha_i| = p^{1/2}$. This in turn implies the estimate

$$|a_p| \leq np^{1/2}.$$

The Lang–Trotter idea is that for any compatible system which is pure of weight-1, it is only congruence obstructions which prevent an integer A from being a_p for infinitely many primes p. As Serre has remarked [Se-Cheb, 8.2, Remarques (3)], all of the image groups $\mathrm{Im}(\rho_N) \subset GL(n, \mathbb{Z}/N\mathbb{Z})$ contain the identity, and hence its trace, the integer n, has no congruence obstruction. Specializing to the case of curves, we get the following conjecture, which in genus $g \geq 1$ seems to be entirely open. [It is of course trivially correct in genus zero, where every a_p vanishes.]

Conjecture 1.6. *Let $C/\mathbb{Q}$ be a projective smooth geometrically connected curve of genus g. Then there are infinitely many good primes p with $a_p = 2g$.*

Already very special cases of this conjecture are extremely interesting. Consider the special $g = 1$ case when $E/\mathbb{Q}$ is the lemniscate curve $y^2 = x^3 - x$, which has good reduction outside of 2. Here we know the explicit "formula" for a_p; cf. [Ir-Ros, Chpt.18, §4, Thm. 5]. If $p \equiv 3$ mod 4, then $a_p = 0$. If $p \equiv 1$ mod 4, then we can write $p = \alpha^2 + \beta^2$ with integers α, β, α odd, β even, and $\alpha \equiv 1 + \beta$ mod 4. This specifies α uniquely, and it specifies $\pm\beta$. [More conceptually, the two Gaussian integers $\alpha \pm \beta i$ are the unique Gaussian primes in $\mathbb{Z}[i]$ which are 1 mod $2 + 2$i and

which lie over p.] Then $a_p = 2\alpha$. So we have $a_p = 2$ precisely when there is a Gaussian prime of the form $1+\beta i$ with $1 \equiv 1+\beta \bmod 4$, i.e. with $\beta = 4n$ for some integer n. Thus $a_p = 2$ precisely when there exists an integer n with

$$p = 1 + 16n^2.$$

So the conjecture for this particular curve is the statement that there are infinitely many primes of the form $1 + 16n^2$.

There is another element common to all the mod N image groups. Embeddings of $\overline{\mathbb{Q}}$ into $\mathbb{C}$ determine "complex conjugation" elements in $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, all in the same conjugacy class, denoted $\mathrm{Frob}_{\mathbb{R}}$. In the curve case, $\mathrm{Frob}_{\mathbb{R}}$ has g eigenvalues 1 and g eigenvalues -1 in every ℓ-adic representation. Therefore $\mathrm{Frob}_{\mathbb{R}}$ has trace zero in every ℓ-adic representation, and consequently in every mod N representation. So we are led to the following conjecture, which in genus $g = 1$ is a celebrated result of Elkies; cf. [Elkies-Real] and [Elkies-SS].

Conjecture 1.7. *Let $C/\mathbb{Q}$ be a projective smooth geometrically connected curve of genus g. Then there are infinitely many good primes p with $a_p = 0$.*

This conjecture is trivially true in some cases. For example, take an odd $\mathbb{Q}$-polynomial $h(x) = -h(-x)$ with all distinct roots, and the curve $y^2 = h(x)$. Then the character sum formula for a_p shows that $a_p = 0$ for all good $p \equiv 3 \bmod 4$. But for an irreducible h of degree $d \geq 5$ whose Galois group is either S_d or A_d, and the curve $y^2 = h(x)$, this conjecture seems to be entirely open.

What should we expect for compatible systems which are pure of weight-2, i.e., each $|\alpha_i| = p$? In this weight-2 case, the probabilistic model has sets $X_p = \mathbb{Z} \cap [-np, np]$ of size $2np + 1$. So the series $\sum_p 1/\#X_p \sim (1/(2n)) \sum_p 1/p$ diverges slowly, and the model allows $A = a_p$ to hold about $(1/(2n)) \log\log T$ times for primes up to T. But in weight-2 there may be more than congruence obstructions to having a given A being a_p infinitely often. Here is the simplest example. Start with an elliptic curve $E/\mathbb{Q}$, say with good reduction at primes p not dividing some integer Δ, and its compatible system of weight-1 representations

$$\rho_{\ell^\infty} : \pi_1(\mathrm{Spec}(\mathbb{Z}[1/\ell\Delta])) \to GL(2, \mathbb{Z}_\ell).$$

In each of these, $\mathrm{Frob}_{\mathbb{R}}$ has eigenvalues 1 and -1. Now consider the compatible system

$$\mathrm{Sym}^2(\rho_{\ell^\infty}) : \pi_1(\mathrm{Spec}(\mathbb{Z}[1/\ell\Delta])) \to GL(3, \mathbb{Z}_\ell).$$

In each of these, $\mathrm{Frob}_{\mathbb{R}}$ has two eigenvalues 1 and one eigenvalue -1, so has trace 1, and hence has trace 1 in every mod N representation $\mathrm{Sym}^2(\rho_N)$. Thus $A = 1$ has no congruence obstruction for the compatible system of $\mathrm{Sym}^2(\rho_{\ell^\infty})$'s. Denote by A_p the trace of Frob_p in this Sym^2 system. Then A_p is related to the original a_p by the formula

$$A_p = (a_p)^2 - p.$$

So $A_p = 1$ is equivalent to $(a_p)^2 - p = 1$, i.e. to

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

which is trivially impossible for $p \geq 5$.

It would be interesting to understand, even conjecturally, what "should" be true about compatible weight-2 systems, for instance for the a_p of a weight-3 newform[10]

[10]The weight in the sense of modular forms is one more than the weight in the sense of compatible systems.

with integer coefficients on some congruence subgroup $\Gamma_1(N)$. Here we are dealing with a compatible system of 2-dimensional representations, so in particular $A = 2$ has no congruence obstruction. It may well be that no fixed nonzero integer A is a_p for infinitely many p; no computer experiment can convince us either way. Nonetheless, we report on some computer experiments below. Caveat emptor.

The simplest examples of weight-3 newforms with integer coefficients are obtained by taking a (K-valued, type $(1,0)$) weight-1 grossencharacter ρ of a quadratic imaginary field K of class number one and inducing its square down to $\mathbb{Q}$. The common feature they exhibit is that for a certain integer $D \geq 1$, we have $a_p = 2$ if and only if the pair of simultaneous equations

$$x^2 + Dy^2 = p, x^2 - Dy^2 = 1$$

has an integer solution. Here are some examples.

($D = 1$) Here $K = \mathbb{Q}(i)$, and ρ attaches to an odd prime ideal $\mathcal{P}$ of $\mathbb{Z}[i]$ the unique generator $\pi = \alpha + \beta i \equiv 1 \bmod (2 + 2i)$. This ρ is the grossencharacter attached to the elliptic curve $y^2 = x^3 - x$; cf. [Ir-Ros, Chpt. 18, Thm. 5]. Inducing ρ^2 gives a weight-3 newform on $\Gamma_1(16)$ whose nebentypus character is the mod 4 character of order 2. [This is 16k3A[1,0]1 in Stein's tables [St].] See [Ka-TLFM, 8.8.10-11] for another occurrence, in the cohomology of a certain elliptic surface.] For this form, we have $a_p = 0$ unless $p \equiv 1 \bmod 4$. When $p \equiv 1 \bmod 4$, choose a $\mathcal{P}$ lying over p, and write $\rho(\mathcal{P}) = \pi = \alpha + \beta i$. Then

$$a_p = \mathrm{Trace}_{\mathbb{Q}(i)/\mathbb{Q}}((\pi)^2) = 2(\alpha^2 - \beta^2) = 2(\alpha - \beta)(\alpha + \beta).$$

So no odd A is ever a_p. For a fixed nonzero even A, the pair of integers $(\alpha - \beta, \alpha + \beta)$ is on the finite list of factorizations in $\mathbb{Z}$ of $A/2$. Solving for (α, β), we see that (α, β) is itself on a finite list. So $p = \alpha^2 + \beta^2$ is on a finite list, and hence $a_p = A$ holds for at most finitely many primes p. In this particular example, $A = 2$ is never an a_p, since the only integer solutions of $\alpha^2 - \beta^2 = 1$ are $(\pm 1, 0)$. This $D = 1$ case is the only case where we can **prove** that for any fixed nonzero A, $a_p = A$ holds for at most finitely many primes p.

($D = 2$) Here $K = \mathbb{Q}(\sqrt{-2})$, and ρ attaches to an odd prime ideal $\mathcal{P}$ of $\mathbb{Z}[\sqrt{-2}]$ the unique generator $\pi = \alpha + \beta\sqrt{-2}$ with $\alpha \equiv 1 \bmod 4$. Inducing ρ^2 gives a weight-3 newform on $\Gamma_1(8)$ whose nebentypus character is the mod 8 character of order 2 whose kernel is $\{1, 3\}$. [This is 8k3A[1,1]1 in Stein's tables [St].] For odd p, a_p vanishes unless $p \equiv 1$ or $3 \bmod 8$. When $p \equiv 1$ or $3 \bmod 8$, choose either $\mathcal{P}$ lying over p, and write $\rho(\mathcal{P}) = \pi = \alpha + \beta\sqrt{-2}$. Then $p = \mathrm{Norm}_{\mathbb{Q}(\sqrt{-2})/\mathbb{Q}}(\pi) = \alpha^2 + 2\beta^2$, and

$$a_p = \mathrm{Trace}_{\mathbb{Q}(\sqrt{-2})/\mathbb{Q}}((\pi)^2) = 2(\alpha^2 - 2\beta^2).$$

($D = 3$) Here $K = \mathbb{Q}(\zeta_3)$, and ρ attaches to a prime-to-6 prime ideal $\mathcal{P}$ of $\mathbb{Z}[\zeta_3]$ the unique generator $\pi = \alpha + \beta\sqrt{-3}$ which lies in the order $\mathbb{Z}[\sqrt{-3}]$ and which has $\alpha \equiv 1 \bmod 3$. Inducing ρ^2 gives a weight-3 newform on $\Gamma_1(12)$ whose nebentypus character is the mod 3 character of order 2. [This is 12k3A[0,1]1 in Stein's tables [St].] For p prime to 6, a_p vanishes unless $p \equiv 1 \bmod 3$. If $p \equiv 1 \bmod 3$, choose a $\mathcal{P}$ lying over p, and write $\rho(\mathcal{P}) = \pi = \alpha + \beta\sqrt{-3}$. Then $p = \mathrm{Norm}_{\mathbb{Q}(\zeta_3)/\mathbb{Q}}(\pi) = \alpha^2 + 3\beta^2$, and

$$a_p = \mathrm{Trace}_{\mathbb{Q}(\zeta_3)/\mathbb{Q}}((\pi)^2) = 2(\alpha^2 - 3\beta^2).$$

($D = 27$) Here $K = \mathbb{Q}(\zeta_3)$, and ρ attaches to a prime-to-3 prime ideal $\mathcal{P}$ of $\mathbb{Z}[\zeta_3]$ the unique generator $\pi = \alpha + \beta(3\zeta_3)$ which lies in the order $\mathbb{Z}[3\zeta_3]$ and has $\alpha \equiv 1 \bmod 3$. This ρ is the grossencharacter attached to the elliptic curve $y^2 = x^3 + 16$; cf. [Ir-Ros, Chpt. 18, Thm. 4]. Inducing ρ^2 gives a weight-3 newform on $\Gamma_1(27)$ whose nebentypus character is the mod 3 character of order 2. [This is 27k3A[9]1 in Stein's tables [St].] For p prime to 3, a_p vanishes unless $p \equiv 1 \bmod 3$. If $p \equiv 1 \bmod 3$, choose a $\mathcal{P}$ lying over p, and write $\rho(\mathcal{P}) = \pi = \alpha + 3\beta\zeta_3$. Then $p = \mathrm{Norm}_{\mathbb{Q}(\zeta_3)/\mathbb{Q}}(\pi) = \alpha^2 - 3\alpha\beta + 9\beta^2$ and

$$a_p = \mathrm{Trace}_{\mathbb{Q}(\zeta_3)/\mathbb{Q}}((\pi)^2) = 2\alpha^2 - 6\alpha\beta - 9\beta^2.$$

So if a_p is even, then β must be even, say $\beta = 2B$, and our equations become

$$p = (\alpha - 3B)^2 + 27B^2, \quad a_p = 2((\alpha - 3B)^2 - 27B^2).$$

($D = 7,\ 11,\ 19,\ 43,\ 67,\ 163$) Here $K = \mathbb{Q}(\sqrt{-D})$, and ρ attaches to a prime-to-D prime ideal $\mathcal{P}$ of $\mathbb{Z}[(1 + \sqrt{-D})/2]$ the unique generator $\pi = \alpha_0 + \beta_0(1 + \sqrt{-D})/2$ which mod $\sqrt{-D}$ is a square mod D. Inducing ρ^2 gives a weight-3 newform on $\Gamma_1(D)$ whose nebentypus character is the mod D character of order 2. [This is Dk3A[(D-1)/2]1 in Stein's tables [St].] For $p \neq D$, a_p vanishes unless p is a square mod D. If p is a square mod D, choose either $\mathcal{P}$ lying over p, and write $\rho(\mathcal{P}) = \pi = \alpha_0 + \beta_0(1 + \sqrt{-D})/2$. Then $p = \mathrm{Norm}_{\mathbb{Q}(\sqrt{-D})/\mathbb{Q}}(\pi) = \alpha_0^2 + \alpha_0\beta_0 + ((D+1)/4)\beta_0^2$ and $a_p = \mathrm{Trace}_{\mathbb{Q}(\sqrt{-D})/\mathbb{Q}}(\pi^2) = 2\alpha_0^2 + 2\alpha_0\beta_0 - ((D-1)/2)\beta_0^2$. Here $(D-1)/2$ is odd, so if a_p is even, then β_0 must be even: π lies in the order $\mathbb{Z}[\sqrt{-D}]$. Rewrite this π as $\alpha + \beta\sqrt{-D}$ with α a square mod D. So if a_p is even, then $p = \alpha^2 + D\beta^2$ and $a_p = 2(\alpha^2 - D\beta^2)$.

We have already noted that in the $D = 1$ example, we never have $a_p = 2$. In the other examples, it is a simple matter to do a computer search for primes p with $a_p = 2$. We run through the solutions $(\pm x_n, \pm y_n)$ of Pell's equation $x^2 - Dy^2 = 1$ by computing the powers of the smallest real quadratic unit $u_D = x_1 + y_1\sqrt{D}$ of norm 1 with x_1, y_2 strictly positive integers. Then $u_D^n = x_n + y_n\sqrt{D}$ and we test the primality of $x_n^2 + Dy_n^2$. But a simple algebra lemma[11] shows that if $x_n^2 + Dy_n^2$ is prime, then n is itself a power 2^a of 2. Indeed, if n has an odd divisor $d \geq 3$, say $n = dm$, the lemma applied to u_D^m shows that $x_n^2 + Dy_n^2$ is divisible by $x_m^2 + Dy_m^2$, so is certainly not prime. In a naive probabilistic model, the probability that $x_{2^a}^2 + Dy_{2^a}^2$ is prime is

$$1/\log(x_{2^a}^2 + Dy_{2^a}^2) \sim 1/\log(u_D^{2^{a+1}}) = 1/(2^{a+1}\log(u_D)).$$

The series $\sum_{a \geq 0} 1/(2^{a+1}\log(u_D))$ converges. So we "expect" that $x_{2^a}^2 + Dy_{2^a}^2$ is prime for at most finitely many values of a. In other words, for any squarefree integer $D > 0$, we expect that there are only finitely many primes p such that the simultaneous equations

$$x^2 + Dy^2 = p, \qquad x^2 - Dy^2 = 1$$

[11]The lemma is this. In the polynomial ring $\mathbb{Z}[X, Y, \sqrt{D}]$ in 3 variables $X, Y, \sqrt{D}$, write $(X + Y\sqrt{D})^n = X_n + Y_n\sqrt{D}$ with X_n, Y_n in the subring $\mathbb{Z}[X, Y, D]$. If n is odd, then $X_n^2 + DY_n^2$ is divisible by $X^2 + Dy^2$ in $\mathbb{Z}[X, Y, D]$. To prove it, notice that $X^2 + Dy^2$ is X_2 and (hence) that $X_n^2 + DY_n^2$ is X_{2n}, so we reduce to the (easy) statement, applied to $(X + Y\sqrt{D})^2$, that X divides X_n in $\mathbb{Z}[X, Y, D]$ if n is odd.

have an integer solution. In particular, for each of our example newforms, we should have $a_p = 2$ for at most finitely many primes p.

Here is a table of search results. The column headed "T" specifies the search range: all $n = 2^a \le T, a \ge 0$. In this search range, we will find all primes $p \le 10^X$, i.e., all primes with at most X decimal digits, for which $a_p = 2$. This is the meaning of the "X" column. The next to last column, #, tells how many primes p in the search range had $a_p = 2$, and the last column tells which powers of u_D gave those p.

D	u_D	T	X	#	for $n =$
2	$3 + 2\sqrt{D}$	32768	50170	3	$1, 2, 4$
3	$2 + \sqrt{D}$	32768	37482	3	$1, 2, 8$
27	$26 + 5\sqrt{D}$	∞	∞	0	none
7	$8 + 3\sqrt{D}$	32768	78801	3	$1, 2, 16$
11	$10 + 3\sqrt{D}$	16384	42596	2	$1, 2$
19	$170 + 39\sqrt{D}$	8192	41475	0	none
43	$3482 + 531\sqrt{D}$	8192	62961	0	none
67	$48842 + 5967\sqrt{D}$	8192	81753	2	$4, 32$
163	$64080026 + 5019135\sqrt{D}$	8192	132837	0	none

That there are provably none for $D = 27$ results from the fact that u_{27} is the cube of u_3. For the amusement of the reader, we give below, for $D = 2, 3, 7, 11$, the two or three primes p with $a_p = 2$ in our search range.

D	p_1	p_2	p_3
2	17	577	665857
3	7	19	708158977
7	127	32257	1500381713949050304320032818543397109 77
11	199	79201	no third one

[For $D = 67$, the first of the two primes found in our search range with $a_p = 2$ was

$$p = 41453144812389737831066275128882623112 97.$$

The second prime found with $a_p = 2$ had 320 digits; it was too big for Mathematica to certify its primality.]

2. Lang–Trotter in the function field case: generalities and what we might hope for

We now turn to a discussion of the Lang–Trotter conjecture for elliptic curves in the function field case; cf. [Pa] for an earlier discussion (but note that his Proposition 4.4 is incorrect). Thus we let k be a finite field $\mathbb{F}_q$ of some characteristic $p > 0$, X/k a projective, smooth, geometrically connected curve, K the function field of X, and E/K an elliptic curve over K. Then E has good reduction at all but finitely many closed points $\mathcal{P} \in X$; more precisely, its Neron model $\mathcal{E}/X$ is, over some dense open set $U \subset X$, a one-dimensional abelian scheme. For each closed point $\mathcal{P} \in U$, with residue field $\mathbb{F}_\mathcal{P}$ of cardinality $\mathbb{N}(\mathcal{P})$, we have the elliptic curve $\mathcal{E}_\mathcal{P}/\mathbb{F}_\mathcal{P} := \mathcal{E} \otimes_U \mathbb{F}_\mathcal{P}/\mathbb{F}_\mathcal{P}$, and the integer $A_\mathcal{P}$, defined by

$$\#\mathcal{E}_\mathcal{P}(\mathbb{F}_\mathcal{P}) = \mathbb{N}(\mathcal{P}) + 1 - A_\mathcal{P}.$$

Exactly as in the number field case, the idea is to try to guess for which integers A there should exist infinitely many closed points $\mathcal{P} \in U$ with $A_\mathcal{P} = A$, and if

possible to be more precise about how many such closed points there are of any given degree. We will try to do this when both of the following two hypotheses hold.

(NCj) The j-invariant $j(E/K) \in K$ is nonconstant, i.e., does not lie in k.

(Ord) For each $\mathcal{P} \in U$, the elliptic curve $\mathcal{E} \otimes_U \mathbb{F}_{\mathcal{P}}/\mathbb{F}_{\mathcal{P}}$ is ordinary; i.e., the integer $A_{\mathcal{P}}$ is prime to $p := char(K)$.

Remark 2.1. The reason we assume (NCj) is this. If (NCj) does not hold, i.e., if our family has constant j, then for any *nonzero* integer A, the equality $A_{\mathcal{P}} = A$ holds for at most finitely many $\mathcal{P}$. Why is this so? If this constant j is supersingular (:= not ordinary), then for each $\mathcal{P}$, the elliptic curve $\mathcal{E} \otimes_U \mathbb{F}_{\mathcal{P}}/F_{\mathcal{P}}$ is supersingular. So the integer $A_{\mathcal{P}}$ is divisible, as an algebraic integer, by $\mathbb{N}(\mathcal{P})^{1/2}$, and hence either $A_{\mathcal{P}} = 0$ or we have the inequality $|A_{\mathcal{P}}| \geq \mathbb{N}(\mathcal{P})^{1/2}$. As there are only finitely many $\mathcal{P}$ of any given norm, the result follows. If, on the other hand, the constant j is ordinary, then $A_{\mathcal{P}}$ is never zero (because it is prime to p), and one knows [B-K, 2.10] that $|A_{\mathcal{P}}| \to \infty$ as $\deg(\mathcal{P}) \to \infty$. So in this ordinary case as well, for any given integer A, the equality $A_{\mathcal{P}} = A$ holds for at most finitely many $\mathcal{P}$.

Remark 2.2. When (NCj) holds, any U of good reduction contains at most finitely many closed points $\mathcal{P}$ which are supersingular (:= not ordinary) [simply because the values at all supersingular points of the nonconstant function j lie in the finite set $\mathbb{F}_{p^2}$]. Removing the supersingular points gives us a smaller dense open $U \subset X$ over which (Ord) holds and does not affect which integers A occur as $A_{\mathcal{P}}$ for infinitely many $\mathcal{P}$.

So we now let k be a finite field $\mathbb{F}_q$ of some characteristic $p > 0$, U/k a smooth, geometrically connected curve with function field K, and $\mathcal{E}/U$ an elliptic curve over U whose j-invariant is nonconstant and which is fibre-by-fibre ordinary. There are slight differences from the number field case which we must take into account.

The first is that inside the fundamental group $\pi_1(U)$ we have the normal subgroup $\pi_1^{\mathrm{geom}}(U) := \pi_1(U \otimes_k \overline{k})$, which sits in a short exact sequence

$$\{1\} \to \pi_1^{\mathrm{geom}}(U) \to \pi_1(U) \xrightarrow{\deg} \mathrm{Gal}(\overline{k}/k) \cong \hat{\mathbb{Z}} \to \{1\}.$$

For each finite extension field $\mathbb{F}_Q/k$, and each $\mathbb{F}_Q$-valued point $u \in U(\mathbb{F}_Q)$, we have its arithmetic Frobenius conjugacy class $\mathrm{Frob}_{u,\mathbb{F}_Q} \in \pi_1(U)$, whose image in $\mathrm{Gal}(\overline{k}/k)$ is the Qth power automorphism of $\overline{k}$. For a closed point $\mathcal{P}$ of U of some degree $d \geq 1$, viewed as a $\mathrm{Gal}(\overline{k}/k)$-orbit of length d in $U(\overline{k})$, we have the arithmetic Frobenius conjugacy class $\mathrm{Frob}_{\mathcal{P}} \in \pi_1(U)$, equal to the class of $\mathrm{Frob}_{u,\mathbb{F}_Q}$, for $\mathbb{F}_Q$ the residue field $\mathbb{F}_{q^d}$ of $\mathcal{P}$ and for $u \in U(\mathbb{F}_Q)$ any point in the orbit which "is" $\mathcal{P}$. For any element $F \in \pi_1(U)$ of degree one, e.g., $\mathrm{Frob}_{u,k}$, if there exists a k-rational point of U, we have a semidirect product description

$$\pi_1^{\mathrm{geom}}(U) \rtimes \langle F \rangle \xrightarrow{\sim} \pi_1(U),$$

where $\langle F \rangle \xrightarrow{\sim} \hat{\mathbb{Z}}$ is the pro-cyclic group generated by F.

The second difference from the number field case is that only for integers $N_0 \geq 2$ which are prime to p is the group scheme $\mathcal{E}[N_0]$ a finite étale form of $\mathbb{Z}/N_0\mathbb{Z} \times \mathbb{Z}/N_0\mathbb{Z}$. So it is only for integers $N_0 \geq 2$ which are prime to p that we get a mod N_0 representation

$$\rho_{N_0} : \pi_1(U) \to (GL(2, \mathbb{Z}/N_0\mathbb{Z}).$$

For a finite extension field $\mathbb{F}_Q/k$, and an $\mathbb{F}_Q$-valued point $u \in U(\mathbb{F}_Q)$, we have an elliptic curve $\mathcal{E}_{u,\mathbb{F}_Q}/\mathbb{F}_Q$, the number of whose $\mathbb{F}_Q$-rational points we write

$$\mathcal{E}_{u,\mathbb{F}_Q}(\mathbb{F}_Q) = Q + 1 - A_{u,\mathbb{F}_Q}.$$

The fundamental compatibility is that for each $N_0 \geq 2$ that is prime to p, we have

$$\mathrm{Trace}(\rho_{N_0}(\mathrm{Frob}_{u,\mathbb{F}_Q})) \equiv A_{u,\mathbb{F}_Q} \mod N_0, \det(\rho_{N_0}(\mathrm{Frob}_{u,\mathbb{F}_Q})) \equiv Q \mod N_0.$$

In particular, for a closed point $\mathcal{P}$ of U, we have

$$\mathrm{Trace}(\rho_{N_0}(\mathrm{Frob}_{\mathcal{P}})) \equiv A_{\mathcal{P}} \mod N_0, \det(\rho_{N_0}(\mathrm{Frob}_{\mathcal{P}})) \equiv \mathbb{N}(\mathcal{P}) \mod N_0.$$

The third difference from the number field case is that, because $\mathcal{E}/U$ is fibre-by-fibre ordinary, the p-divisible group $\mathcal{E}[p^\infty]$ sits in a short exact sequence

$$0 \to \mathcal{E}[p^\infty]^0 \to \mathcal{E}[p^\infty] \to \mathcal{E}[p^\infty]^{et} \to 0,$$

in which the quotient $\mathcal{E}[p^\infty]^{et}$ is a form of $\mathbb{Q}_p/\mathbb{Z}_p$, and the kernel $\mathcal{E}[p^\infty]^0$ is the dual form of μ_{p^∞}. So the quotient $\mathcal{E}[p^\infty]^{et}$ gives us a homomorphism

$$\rho_{p^\infty} : \pi_1(U) \to \mathrm{Aut}_{gp}(\mathbb{Q}_p/\mathbb{Z}_p) \cong GL(1,\mathbb{Z}_p) \cong \mathbb{Z}_p^\times.$$

On Frobenius elements, this p-adic character ρ_{p^∞} of $\pi_1(U)$ gives the p-adic unit eigenvalue of Frobenius: the fact that the integer $A_{u,\mathbb{F}_Q}$, resp. $A_{\mathcal{P}}$, is prime to p implies that the integer polynomial

$$X^2 - A_{u,\mathbb{F}_Q}X + Q, \text{ resp. } X^2 - A_{\mathcal{P}}X + \mathbb{N}(\mathcal{P}),$$

has a unique root in $\mathbb{Z}_p^\times$, namely $\rho_{p^\infty}(\mathrm{Frob}_{u,\mathbb{F}_Q})$, resp. $\rho_{p^\infty}(\mathrm{Frob}_{\mathcal{P}})$. More concretely, we have identities in $\mathbb{Z}_p$,

$$A_{u,\mathbb{F}_Q} = \rho_{p^\infty}(\mathrm{Frob}_{u,\mathbb{F}_Q}) + Q/\rho_{p^\infty}(\mathrm{Frob}_{u,\mathbb{F}_Q}),$$
$$A_{\mathcal{P}} = \rho_{p^\infty}(\mathrm{Frob}_{\mathcal{P}}) + \mathbb{N}(\mathcal{P})/\rho_{p^\infty}(\mathrm{Frob}_{\mathcal{P}}).$$

Given a prime-to-p integer A, and a power Q of p, we denote by $\mathrm{unit}_Q(A) \in \mathbb{Z}_p^\times$ the unique root in $\mathbb{Z}_p^\times$ of the polynomial $X^2 - AX + Q$. We have

$$X^2 - AX + Q = (X - \mathrm{unit}_Q(A))(X - Q/\mathrm{unit}_Q(A)).$$

Thus

$$\rho_{p^\infty}(\mathrm{Frob}_{u,\mathbb{F}_Q}) = \mathrm{unit}_Q(A_{u,\mathbb{F}_Q}),$$
$$\rho_{p^\infty}(\mathrm{Frob}_{\mathcal{P}}) = \mathrm{unit}_{\mathbb{N}(\mathcal{P})}(A_{\mathcal{P}}).$$

If $Q \geq p^\nu$, resp. if $\mathbb{N}(\mathcal{P}) \geq p^\nu$, then we have the congruences

$$\mathrm{unit}_Q(A_{u,\mathbb{F}_Q}) \equiv A_{u,\mathbb{F}_Q} \mod p^\nu,$$
$$\mathrm{unit}_{\mathbb{N}(\mathcal{P})}(A_{\mathcal{P}}) \equiv A_{\mathcal{P}} \mod p^\nu.$$

For a fixed power p^ν of p, $\nu \geq 0$, we denote by

$$\rho_{p^\nu} : \pi_1(U) \to (\mathbb{Z}_p/p^\nu\mathbb{Z}_p)^\times$$

the reduction mod p^ν of ρ_{p^∞}, with the convention that for $\nu = 0$, ρ_{p^0} is the trivial representation toward the trivial group. Thus if $Q \geq p^\nu$, resp. if $\mathbb{N}(\mathcal{P}) \geq p^\nu$, then we have the congruences

$$\rho_{p^\nu}(\mathrm{Frob}_{u,\mathbb{F}_Q}) \equiv A_{u,\mathbb{F}_Q} \mod p^\nu,$$
$$\rho_{p^\nu}(\mathrm{Frob}_{\mathcal{P}}) \equiv A_{\mathcal{P}} \mod p^\nu.$$

Given an integer A, we can of course ask if $A = A_{\mathcal{P}}$ for infinitely many closed points $\mathcal{P}$. But in the function field case there are two additional questions we can ask.

(1) For a given finite extension $\mathbb{F}_Q/k$ is there a closed point $\mathcal{P}$ with residue field $\mathbb{F}_Q$, i.e. with $\mathbb{N}(\mathcal{P}) = Q$, and with $A = A_{\mathcal{P}}$? If so, how many such closed points are there?
(2) For a given finite extension $\mathbb{F}_Q/k$ is there an $\mathbb{F}_Q$-valued point $u \in U(\mathbb{F}_Q)$ with $A = A_{u,\mathbb{F}_Q}$? If so, how many such $\mathbb{F}_Q$-valued points are there?

To describe conjectural answers to these questions, we need some notation. Given an integer $N \geq 2$, factor it as

$$N = N_0 p^\nu$$

with N_0 prime to p and $\nu \geq 0$. Then form the product representation

$$\rho_N := \rho_{N_0} \times \rho_{p^\nu} : \pi_1(U) \to GL(2, \mathbb{Z}/N_0\mathbb{Z}) \times (\mathbb{Z}/p^\nu\mathbb{Z})^\times.$$

We will write an element of the product group as

$$(g_{N_0}, \gamma_{p^\nu}) \in GL(2, \mathbb{Z}/N_0\mathbb{Z}) \times (\mathbb{Z}/p^\nu\mathbb{Z})^\times.$$

We define its determinant in $(\mathbb{Z}/N_0\mathbb{Z})^\times$ by

$$\det(g_{N_0}, \gamma_{p^\nu}) := \det(g_{N_0}) \in \mathbb{Z}/N_0\mathbb{Z},$$

and its trace in $\mathbb{Z}/N\mathbb{Z}$ by

$$\mathrm{Trace}(g_{N_0}, \gamma_{p^\nu}) := (\mathrm{Trace}(g_{N_0}), \gamma_{p^\nu}) \in \mathbb{Z}/N_0\mathbb{Z} \times \mathbb{Z}/p^\nu\mathbb{Z} \xleftarrow{\sim} \mathbb{Z}/N\mathbb{Z},$$

the last arrow being "simultaneous reduction" mod N_0 and p^ν.

In analogy to the number field case, we denote by G_N the image group

$$G_N := \rho_N(\pi_1(U)) \subset GL(2, \mathbb{Z}/N_0\mathbb{Z}) \times (\mathbb{Z}/p^\nu\mathbb{Z})^\times.$$

But in the function field case, we must consider also the normal subgroup $G_N^{\mathrm{geom}} \lhd G_N$ defined as

$$G_N^{\mathrm{geom}} := \rho_N(\pi_1^{\mathrm{geom}}(U)).$$

For each strictly positive power $Q = (\#k)^d$ of $\#k$, we define $G_{N,\det=Q} \subset G_N$ to be the coset of G_N^{geom} defined by

$$G_{N,\det=Q} := \rho_N(\pi_1(U)_{\deg=d}) = \rho_N(F^d \pi_1^{\mathrm{geom}}(U)) = \rho_N(F)^d G_N^{\mathrm{geom}},$$

for any element $F \in \pi_1(U)$ of degree one.

For each integer A mod N, we define $G_N(A, Q) \subset G_{N,\det=Q}$ as follows. If N is prime to p, i.e., if $N = N_0$, then $G_N(A, Q)$ is the subset of $G_{N_0,\det=Q}$ consisting of those elements whose trace is A mod N_0. If $p|N$, then $G_N(A, Q)$ is empty if $p|A$. If $p|N$ and A is prime to p, it is the subset of $G_{N,\det=Q}$ consisting of those elements whose trace is $(A \bmod N_0,\ \mathrm{unit}_Q(A) \bmod p^\nu)$ in $\mathbb{Z}/N_0\mathbb{Z} \times \mathbb{Z}/p^\nu\mathbb{Z}$. [This makes sense, because, for any fixed Q as above, if an integer A is invertible mod p, then $\mathrm{unit}_Q(A)$ mod p^ν depends only on A mod p^ν. But only for $Q \geq p^\nu$ will we have $\mathrm{unit}_Q(A) \equiv A$ mod p^ν.]

For later use, we define

$$g_{N,\det=Q} := \#G_{N,\det=Q},$$
$$g_N(A, Q) := \#G_N(A, Q),$$
$$g_N(\mathrm{avg}, Q) := (1/N) \sum_{A \bmod N} g_N(A, Q) = (1/N) g_{N,\det=Q}.$$

The relevance of the subsets $G_N(A,Q) \subset G_{N,\det=Q} \subset G_N$ is this. Suppose we are given an integer A prime to p, and a power Q of $\#k$. If there is an $\mathbb{F}_Q$-valued point $u \in U(\mathbb{F}_Q)$ with $A_{u,\mathbb{F}_Q} = A$, resp. a closed point $\mathcal{P}$ with norm Q and $A_{\mathcal{P}} = A$, then for every $N \geq 2$, $\rho_N(\mathrm{Frob}_{u,\mathbb{F}_Q})$, resp. $\rho_N(\mathrm{Frob}_{\mathcal{P}})$, lies in $G_N(A,Q)$.

We say that the data (A,Q), A an integer prime to p and Q a (strictly positive) power of $\#k$, has a congruence obstruction at N if the set $G_N(A,Q)$ is empty, and we say that (A,Q) has an Archimedean obstruction if $A^2 > 4Q$.

The most optimistic hope is that if (A,Q) has neither Archimedean nor congruence obstruction (i.e., A is prime to p, $|A| < 2\sqrt{Q}$, and for all $N \geq 2$ the set $G_N(A,Q)$ is nonempty), then there should be a closed point $\mathcal{P}$ with norm Q and $A_{\mathcal{P}} = A$. [We might even speculate about how many, at least if Q is suitably large.] Unfortunately, this hope is false for trivial reasons; we can remove from U all its closed points of any given degree and obtain now a new situation where the groups G_N, being birational invariants, are unchanged, but where there are no closed points whatever of the given degree. What is to be done? One possibility is to make this sort of counterexample illegal: go back to the projective smooth geometrically connected curve X/k with function field K in which U sits as a dense open set, and replace U by the possibly larger open set $U_{\max} \subset X$ that we obtain by removing from X **only** those points at which the Neron model of $\mathcal{E}_K/K$ has either bad reduction or supersingular reduction. But even this alleged remedy is insufficient, as we will see below. It is still conceivable that if (A,Q) has neither Archimedean nor congruence obstruction there is an $\mathbb{F}_Q$-point $u \in U_{\max}$ such that $\mathrm{Frob}_{u,\mathbb{F}_Q}$ gives rise to (A,Q); the counterexample below does not rule out this possibility.

Here is the simplest counterexample. Take any prime power $q = p^\nu \geq 4$, take for $U = U_{\max}$ the (ordinary part of the) Igusa curve $Ig(q)^{\mathrm{ord}}/\mathbb{F}_q$, and take for $\mathcal{E}/U$ the corresponding universal elliptic curve. For a finite field (or indeed for any perfect field) L/k, an L-valued point $u \in Ig(q)^{\mathrm{ord}}(L)$ is an L-isomorphism class of pairs $(E/L, P \in E[q](L))$ consisting of an elliptic curve E/L together with an L-rational point of order q. Now consider the data $(A = 1-2q, Q = q^2)$. The key fact is that any $E_2/\mathbb{F}_{q^2}$ with trace $A_2 = 1-2q$ is isomorphic to the extension of scalars of a unique $E_1/\mathbb{F}_q$ with trace $A_1 = 1$, as will be shown in Lemma 4.1. But any such $E_1/\mathbb{F}_q$ has q rational points, so the group $E_1(\mathbb{F}_q)$ is cyclic of order q, and hence every point of order q in $E_1(\overline{\mathbb{F}_q})$, and a fortiori every point of order q in $E_1(\mathbb{F}_{q^2})$, is already $\mathbb{F}_q$-rational. So although the data $(A = 1-2q, Q = q^2)$ occurs from an $\mathbb{F}_{q^2}$-point, and hence has no congruence obstruction, it does not occur from a closed point of degree 2.

There are three plausible hopes one might entertain in the function field case. Let $\mathcal{E}/U$ be as above (fibrewise ordinary, nonconstant j-invariant). Here are the first two.

HOPE (1) Given a prime-to-p integer A, there exists a real constant $C(A, \mathcal{E}/U)$ with the following property. If Q is a power of $\#k$ with $Q \geq C(A, \mathcal{E}/U)$, and if (A,Q) has neither Archimedean nor congruence obstruction, then there exists a closed point $\mathcal{P}$ with norm Q and $A_{\mathcal{P}} = A$.

HOPE (2) Given a prime-to-p integer A, and a real number $\epsilon > 0$, there exists a real constant $C(A, \epsilon, \mathcal{E}/U)$ with the following property. If Q is a power of $\#k$ with $Q \geq C(A, \epsilon, \mathcal{E}/U)$, and if (A,Q) has neither Archimedean nor congruence obstruction, then for the number $\pi_{A,Q}$ of closed points

with norm Q and $A_{\mathcal{P}} = A$ and for the number $n_{A,Q}$ of $\mathbb{F}_Q$-valued points $u \in U(\mathbb{F}_Q)$ with $A_{u,\mathbb{F}_Q} = A$ we have the inequalities

$$Q^{\frac{1}{2}-\epsilon} < \pi_{A,Q} \leq n_{A,Q} < Q^{\frac{1}{2}+\epsilon}.$$

To describe the final hope, we must discuss another, weaker, notion of congruence obstruction. Given a prime-to-p integer A, suppose there are infinitely many closed points $\mathcal{P}$ with $A_{\mathcal{P}} = A$. Then as there are only finitely many closed points of each degree, it follows that there are infinitely many powers Q_i of $\#k$ for which (A, Q_i) has no congruence obstruction (and of course no Archimedean obstruction either). For a fixed $N = N_0 p^\nu$, if Q_i is sufficiently large ($Q_i \geq p^\nu$ being the precise condition), then G_N contains an element whose trace is A mod N.

So we are led to a weaker notion of congruence obstruction, which is the literal analogue of the number field condition: we say that the prime-to-p integer A has a congruence obstruction at N if G_N contains no element whose trace is A mod N, and we say that A has a congruence obstruction if it has one at N for some N. This brings us to the third hope.

HOPE (3) Suppose the prime-to-p integer A has no congruence obstruction. Then there exist infinitely many closed points $\mathcal{P}$ with $A_{\mathcal{P}} = A$.

Notice, however, that the assumption that A has no congruence obstruction is, at least on its face, much weaker than the assumption that there are infinitely many powers Q_i of $\#k$ for which (A, Q_i) has no congruence obstruction.

3. Lang–Trotter in the function field case: the case of modular curves

In the number field case, there is no elliptic curve where we know Lang–Trotter for even a single nonzero integer A. But over any finite field k, we will show that there are infinitely many examples of situations $\mathcal{E}/U/k$, nonconstant j-invariant and fibrewise ordinary, where all three of our hopes are provably correct. These examples are provided by modular curves over finite fields, and the universal families of elliptic curves they carry.

Let us first describe the sorts of level structures we propose to deal with in a given characteristic $p > 0$. We specify three prime-to-p positive integers (L, M, N_0) and a power $p^\nu \geq 1$ of p. We assume that (L, M, N_0) are pairwise relatively prime.

Given this data, we work over a finite extension $k/\mathbb{F}_p$ given with a primitive N_0th root of unity $\zeta_{N_0} \in k$, and consider the moduli problem, on k-schemes S/k, of S-isomorphism classes of fibrewise ordinary elliptic curves E/S endowed with all of the following data, which for brevity we will call an $\mathcal{M}$-structure on E/S:

(1) a cyclic subgroup of order L, i.e., a $\Gamma_0(L)$-structure on E/S,
(2) a point P_M of order M, i.e., a $\Gamma_1(M)$-structure on E/S,
(3) a basis (Q, R) of $E[N_0]$ with $e_{N_0}(Q, R) = \zeta_{N_0}$, i.e., an oriented $\Gamma(N_0)$-structure on E/S,
(4) a generator T of $\mathrm{Ker}(V^\nu : E^{(p^\nu/S)} \to E)$, i.e., an $Ig(p^\nu)$-structure on E/S.

Having specified a finite extension $k/\mathbb{F}_p$ given with a primitive N_0th root of unity $\zeta_{N_0} \in k$ and the data (L, M, N_0, p^ν) above, we make the further assumption that

at least one of the following three conditions holds:

(1) $M \geq 4$,
(2) $N_0 \geq 3$,
(3) $p^\nu \geq 4$.

This assumption guarantees that the associated moduli problem is representable by a smooth, geometrically connected k-curve $\mathcal{M}^{\mathrm{ord}}$ over which we have the corresponding universal family $\mathcal{E}^{\mathrm{univ}}/\mathcal{M}^{\mathrm{ord}}$. For this situation, points of $\mathcal{M}^{\mathrm{ord}}$ have a completely explicit description.

For any k-scheme S/k, the S-valued points of $\mathcal{M}^{\mathrm{ord}}$ are precisely the S-isomorphism classes of fibrewise ordinary elliptic curves E/S endowed with an $\mathcal{M}$-structure. In particular, for $\mathbb{F}_Q/k$ a finite overfield, an $\mathbb{F}_Q$-valued point of $\mathcal{M}^{\mathrm{ord}}$ is an $\mathbb{F}_Q$-isomorphism class of pairs

$$(\text{an ordinary elliptic curve } E/\mathbb{F}_Q, \text{ an } \mathcal{M}\text{-structure on it}).$$

What about closed points $\mathcal{P}$ of $\mathcal{M}^{\mathrm{ord}}$ with norm $\mathbb{N}(\mathcal{P}) = Q$? These are precisely the orbits of $\mathrm{Gal}(\mathbb{F}_Q/k)$ on the set $\mathcal{M}^{\mathrm{ord}}(\mathbb{F}_Q)$ which contain $\deg(\mathbb{F}_Q/k)$ distinct $\mathbb{F}_Q$-valued points. In more down-to-earth terms, an $\mathbb{F}_Q$-valued point lies in the orbit of a closed point of norm $\mathbb{N}(\mathcal{P}) = Q$ if and only it is not (the extension of scalars of) a point with values in a proper subfield $k \subset \mathbb{F}_{Q_1} \subsetneqq \mathbb{F}_Q$. Let us denote by

$$\mathcal{M}^{\mathrm{ord}}(\mathbb{F}_Q)^{\mathrm{prim}} \subset \mathcal{M}^{\mathrm{ord}}(\mathbb{F}_Q)$$

those $\mathbb{F}_Q$-valued points which lie in no proper subfield. So we have the tautological formula

$$\#\{\text{closed points with norm } Q\} = \frac{\#\mathcal{M}^{\mathrm{ord}}(\mathbb{F}_Q)^{\mathrm{prim}}}{\deg(\mathbb{F}_Q/k)}.$$

4. Counting ordinary points on modular curves by class number formulas

In this section, we recall the use of class number formulas in counting ordinary points. In a later section, we will invoke the Brauer–Siegel theorem (but only for quadratic imaginary fields, so really Siegel's theorem [Sie]) and its extension to quadratic imaginary orders to convert these class number formulas into the explicit upper and lower bounds asserted in Hope (2). These class number formulas go back to Deuring [Deu]; cf. also Waterhouse [Wat]. As Howe points out [Howe], the story is considerably simplified if we make use of Deligne's description [De-VA] of ordinary elliptic curves over a given finite field. Let $\mathbb{F}_q$ be a finite field and $E/\mathbb{F}_q$ an ordinary elliptic curve. We have

$$\#E(\mathbb{F}_q) = q + 1 - A,$$

for some prime-to-p integer A satisfying

$$A^2 < 4q.$$

Conversely, given any prime-to-p integer A satisfying

$$A^2 < 4q,$$

one knows by Honda–Tate, cf. [Honda] and [Tate], that there is at least one ordinary elliptic curve $E/\mathbb{F}_q$ with

$$\#E(\mathbb{F}_q) = q + 1 - A.$$

The first question, then, is to describe, for fixed (A, q) as above (i.e., A prime-to-p with $A^2 < 4q$) the category of all ordinary elliptic curves $E/\mathbb{F}_q$ with

$$\#E(\mathbb{F}_q) = q + 1 - A,$$

the morphisms being $\mathbb{F}_q$-homomorphisms. We denote by $\mathbb{Z}[F]$ the ring $\mathbb{Z}[X]/(X^2 - AX + q)$. Since $A^2 < 4q$, this ring $\mathbb{Z}[F]$ is an order in a quadratic imaginary field, whose ring of integers we will denote $\mathcal{O}$. [In the general Deligne story we would need to work with the ring $\mathbb{Z}[F, q/F]$, but here q/F is already present, namely $q/F = A - F$.] Deligne provides an explicit equivalence of categories (by picking (!) an embedding of the ring of Witt vectors $W(\mathbb{F}_q)$ into $\mathbb{C}$ and then taking the first integer homology group of the Serre–Tate canonical lifting; cf. [Mes, V 2.3, V 3.3, and Appendix]) of this category with the category of $\mathbb{Z}[F]$-modules H which as $\mathbb{Z}$-modules are free of rank 2 and such that the characteristic polynomial of F acting on H is

$$X^2 - AX + q.$$

In this equivalence of categories, suppose an ordinary $E/\mathbb{F}_q$ gives rise to the $\mathbb{Z}[F]$-module H. For any prime-to-p integer N, the group $E[N](\overline{\mathbb{F}_q})$ as $\mathbb{Z}[F]$-module, F acting as the arithmetic Frobenius Frob_q in $\mathrm{Gal}(\overline{\mathbb{F}_q}/\mathbb{F}_q)$, is just the $\mathbb{Z}[F]$-module H/NH. For a power p^ν of p, the group $E[p^\nu](\overline{\mathbb{F}_q})$ as $\mathbb{Z}[F]$-module is obtained from H as follows. We first write the $\mathbb{Z}_p[F]$-decomposition

$$H \otimes_{\mathbb{Z}} \mathbb{Z}_p = H^{et} \oplus H^{\mathrm{conn}},$$
$$H^{et} := \mathrm{Ker}(F - \mathrm{unit}_q(A)), \quad H^{\mathrm{conn}} := \mathrm{Ker}(F - q/\mathrm{unit}_q(A)),$$

of $H \otimes_{\mathbb{Z}} \mathbb{Z}_p$ as the direct sum of two free $\mathbb{Z}_p$-modules of rank one, of which the first is called the "unit root subspace". Then for each power p^ν of p, we have

$$E[p^\nu](\overline{\mathbb{F}_q}) \cong H^{et}/p^\nu H^{et}.$$

An equivalent, but less illuminating, description of $H^{et}/p^\nu H^{et}$ is as the image of F^ν in $H/p^\nu H$ (because $H/p^\nu H$ is the direct sum $H^{et}/p^\nu H^{et} \oplus H^{\mathrm{conn}}/p^\nu H^{\mathrm{conn}}$, and F^ν is an isomorphism on the first factor but kills the second factor).

Here is an application of Deligne's description.

Lemma 4.1. *Suppose $\mathbb{E}_2/\mathbb{F}_{q^2}$ is an elliptic curve with trace $A_2 = 1 - 2q$. Then there exists a unique elliptic curve $\mathbb{E}_1/\mathbb{F}_q$ with trace $A_1 = 1$ which gives rise to $\mathbb{E}_2/\mathbb{F}_{q^2}$ by extension of scalars.*

Proof. Denote by F_2 the Frobenius for $\mathbb{E}_2/\mathbb{F}_{q^2}$. Then F_2 satisfies

$$F_2^2 - (1 - 2q)F_2 + q^2 = 0, \text{ i.e. } F_2 = (F_2 + q)^2.$$

Thus $F_1 := F_2 + q$ is a square root of F_2, and it satisfies the equation

$$F_1^2 - F_1 + q = 0, \text{ i.e. } F_2 = F_1 - q.$$

This last equation shows that $\mathbb{Z}[F_2] = \mathbb{Z}[F_1]$. In terms of the $\mathbb{Z}[F_2]$-module H_2 attached to $\mathbb{E}_2/\mathbb{F}_{q^2}$, $\mathbb{E}_1/\mathbb{F}_q$ is the unique curve over $\mathbb{F}_q$ corresponding to the same H_2, now viewed as a $\mathbb{Z}[F_1]$-module. □

Class number formulas are based on the following "miracle" of complex multiplication of elliptic curves. [We say "miracle" because the analogous statements

can be false for higher-dimensional abelian varieties.] Given a $\mathbb{Z}[F]$-module H as above, we can form a possibly larger order R,

$$\mathbb{Z}[F] \subset R \subset \mathcal{O},$$

defined as

$$R := \mathrm{End}_{\mathbb{Z}[F]}(H).$$

Of course this R is just the $\mathbb{F}_q$-endomorphism ring of the corresponding $E/\mathbb{F}_q$, thanks to the equivalence. So tautologically H is an R-module. The miracle is that H is an **invertible** R-module; cf. [Sh, 4.11, 5.4.2]. Of course any order $\mathbb{Z}[F] \subset R \subset \mathcal{O}$ can occur as H varies, since one could take R itself as an H. So if we separate the ordinary elliptic curves $E/\mathbb{F}_q$ with given data (A, q) by the orders which are their $\mathbb{F}_q$-endomorphism rings, then for a given order R the $\mathbb{F}_q$-isomorphism classes with that particular R are the isomorphism classes of invertible R-modules, i.e., the elements of the Picard group $\mathrm{Pic}(R)$, whose order is called the class number $h(R)$ of the order R.

Suppose that we now fix not only (A, q) but also the endomorphism ring R. Then for any ordinary elliptic curve $E/\mathbb{F}_q$ with this data, the question of exactly how many $\mathcal{M}$-structures $E/\mathbb{F}_q$ admits is determined entirely by the data consisting of (A, q) and R. Indeed, if $E/\mathbb{F}_q$ gives rise to H, then H is an invertible R-module. Now for **any** invertible R-module H_1, and for any integer $N_1 \geq 1$, the invertible R/N_1R-module H_1/N_1H_1 is R-isomorphic to R/N_1R (simply because R/N_1R, being finite, is semi-local, so has trivial Picard group), and hence a fortiori is $\mathbb{Z}[F]$-isomorphic to R/N_1R. Taking H_1 to be H and N_1 to be LMN_0p^ν, we conclude that $H/(LMN_0p^\nu)H$ is $\mathbb{Z}[F]$-isomorphic to $R/(LMN_0p^\nu)R$. Translating back through Deligne's equivalence, we see that $E[LMN_0p^\nu]((\overline{\mathbb{F}_q})$ is $\mathbb{Z}[F]$-isomorphic to

$$R/LMN_0R \times F^\nu(R/p^\nu R).$$

Thus we have the following dictionary:

(1) $\Gamma_0(L)$-structure: a cyclic subgroup of R/LR of order L which is $\mathbb{Z}[F]$-stable.
(2) $\Gamma_1(M)$-structure: a point $P \in R/MR$ which has additive order M and which is fixed by F.
(3) **unoriented** $\Gamma(N_0)$-structure: a $\mathbb{Z}/N_0\mathbb{Z}$-basis of R/N_0R consisting of points fixed by F. An unoriented $\Gamma(N_0)$-structure exists if and only if F acts as the identity on R/N_0R. If an unoriented $\Gamma(N_0)$-structure exists, there are precisely $\#GL(2, \mathbb{Z}/N_0\mathbb{Z})$ of them. Of these, precisely $\#SL(2, \mathbb{Z}/N_0\mathbb{Z})$ are oriented (for a chosen ζ_{N_0}).
(4) $Ig(p^\nu)$-structure: a $\mathbb{Z}/p^\nu\mathbb{Z}$-basis of $F^\nu(R/p^\nu R)$ ($\cong H^{et}/p^\nu H^{et}$) which is fixed by F, or equivalently, an F-fixed point in $R/p^\nu R$ which has additive order p^ν.

Thus we see explicitly that how many $\mathcal{M}$-structures $E/\mathbb{F}_q$ admits is determined entirely by the data consisting of (A, q) and R. Let us denote this number by

$$\#\mathcal{M}(A, q, R).$$

Notice also that for such an $E/\mathbb{F}_q$ giving rise to (A, q) and R, the automorphism group of $E/\mathbb{F}_q$ is the group $R^\times$ of units in the endomorphism ring R. Recall that $\mathbb{F}_q$ points on the modular curve $\mathcal{M}^{\mathrm{ord}}$ are $\mathbb{F}_q$-isomorphism classes of pairs (ordinary $E/\mathbb{F}_q$, $\mathcal{M}$-structure on $E/\mathbb{F}_q$). So the number of $\mathbb{F}_q$ points on $\mathcal{M}^{\mathrm{ord}}$ whose

underlying ordinary elliptic curve gives rise to the data (A, q, R) is the product

$$\#\mathcal{M}(A, q, R)h(R)/\#R^\times.$$

For given (ordinary) data (A, q), with $\mathbb{Z}[F] := \mathbb{Z}[X]/(X^2 - AX + q)$ and ring of integers $\mathcal{O} \subset \mathbb{Q}[F]$, let us denote by

$$\mathcal{M}^{\mathrm{ord}}(\mathbb{F}_q, A) \subset \mathcal{M}^{\mathrm{ord}}(\mathbb{F}_q)$$

the set of $\mathbb{F}_q$ points on $\mathcal{M}^{\mathrm{ord}}$ whose underlying ordinary elliptic curve gives rise to the data (A, q). Then $\#\mathcal{M}^{\mathrm{ord}}(\mathbb{F}_q, A)$ is a sum, over all orders R between $\mathbb{Z}[F]$ and $\mathcal{O}$:

$$\#\mathcal{M}^{\mathrm{ord}}(\mathbb{F}_q, A) = \sum_{\text{orders } Z[F]\subset R\subset\mathcal{O}} \#\mathcal{M}(A, q, R)h(R)/\#R^\times.$$

Before we try to count $\mathcal{M}$-structures, let us record the congruences and inequalities which necessarily hold when such structures exist.

Lemma 4.2. *Let $k/\mathbb{F}_p$ be a finite extension, given with a primitive N_0th root of unity $\zeta_{N_0} \in k$, $\mathbb{F}_q/k$ a finite extension, and $E/\mathbb{F}_q$ an ordinary elliptic curve which gives rise to the data (A, q, R). Suppose that $E/\mathbb{F}_q$ admits an $\mathcal{M}$-structure. Then $q \equiv 1 \mod N_0$, and we have the following additional congruences.*

1. *There exists $a \in (\mathbb{Z}/L\mathbb{Z})^\times$ satisfying $a^2 - Aa + q \equiv 0 \mod L$; i.e., the polynomial $X^2 - AX + q$ factors completely $\mod L$. Equivalently, there exists $a \in (\mathbb{Z}/L\mathbb{Z})^\times$ such that $A \equiv a + q/a \mod L$.*
2. *$q + 1 \equiv A \bmod MN_0^2p^\nu$.*

Moreover, we have $q \geq p^\nu$ if p is odd. When $p = 2$, we also have $q \geq p^\nu$ except in the two exceptional cases $(q, p^\nu) = (2, 4)$ and $(q, p^\nu) = (4, 8)$; in those two cases we have $A = -1$ and $A = -3$, respectively.

Proof. That $q \equiv 1 \bmod N_0$ results from the fact that $\mathbb{F}_q$ contains a primitive N_0th root of unity. To prove (1), suppose we have an F-stable $\mathbb{Z}/L\mathbb{Z}$ subgroup $\Gamma_0 \subset R/LR$. Then F, being an automorphism of R/LR, acts on this subgroup by multiplication by some unit $a \in (\mathbb{Z}/L\mathbb{Z})^\times$. But $F^2 - AF + q$ annihilates R, so it annihilates R/LR. As $\Gamma_0 \subset R/LR$ is F-stable, and F acts on Γ_0 by a, we get that $a^2 - Aa + q \in \mathbb{Z}/L\mathbb{Z}$ annihilates this cyclic group of order L, so $a^2 - Aa + q = 0$ in $\mathbb{Z}/L\mathbb{Z}$. The existence of such an a is equivalent to the polynomial $X^2 - AX + q$ factoring mod L, and to the congruence $A \equiv a + q/a \bmod L$ (then the factorization is $(X - a)(X - q/a) \bmod L$). The congruence (2) is just the point-count divisibility that follows from having an $\mathcal{M}$-structure. To prove the "moreover" statement, we exploit the fact that, by (2), p^ν divides $q + 1 - A$. We argue by contradiction. If $p^\nu > q$, then $p^\nu \geq pq$ (since q is itself a power of p). So p^ν is divisible by pq, and hence pq divides $q + 1 - A$. By the Weil bound and ordinarity, $q + 1 - A$ is nonzero (indeed $q + 1 - A > (\sqrt{q} - 1)^2 > 0$), so from the divisibility we get the inequality

$$q + 1 - A \geq pq.$$

Again by the Weil bound, we have $(\sqrt{q} + 1)^2 > q + 1 - A$, so we get

$$q + 1 + 2\sqrt{q} = (\sqrt{q} + 1)^2 > pq = (p - 2)q + q + q.$$

Adding $1 - 2\sqrt{q} - q$ to both sides, we get

$$2 > (p - 2)q + (\sqrt{q} - 1)^2.$$

This is nonsense if $p \geq 3$. If $p = 2$, this can hold, precisely in the indicated cases. □

To say more about how this works explicitly, we need to keep track, for given ordinary data (A, q), of the orders between $\mathbb{Z}[F]$ and the full ring of integers $\mathcal{O}$. The orders $R \subset \mathcal{O}$ are the subrings of the form $\mathbb{Z} + f\mathcal{O}$, with $f \geq 1$ an integer. The integer $f \geq 1$ is called the conductor of the order; it is the order of the additive group $\mathcal{O}/R$. Because (A, q) is given, the particular order $\mathbb{Z}[F] \subset \mathcal{O}$ is given, and we will denote by $f_{A,q}$ its conductor:

$$f_{A,q} := \text{conductor of } \mathbb{Z}[F].$$

An order $R \subset \mathcal{O}$ contains $\mathbb{Z}[F]$ if and only if its conductor f_R divides $f_{A,q}$. For an intermediate order $\mathbb{Z}[F] \subset R \subset \mathcal{O}$, we define its **co-conductor** f_R^c to be the quotient:

$$f_R^c := f_{A,q}/f_R = \#(R/\mathbb{Z}[F]).$$

Of course this notion of co-conductor only makes sense because we have specified the particular order $\mathbb{Z}[F]$. Just as the conductor measures how far "down" an intermediate order is from $\mathcal{O}$, so its co-conductor measures how far "up" it is from $\mathbb{Z}[F]$.

Lemma 4.3. *Let $k/\mathbb{F}_p$ be a finite extension, given with a primitive N_0th root of unity $\zeta_{N_0} \in k$, $\mathbb{F}_q/k$ a finite extension, and $E/\mathbb{F}_q$ an ordinary elliptic curve which gives rise to the data (A, q, R). Suppose that the following congruences hold.*

(1) *There exists $a \in (\mathbb{Z}/L\mathbb{Z})^\times$ satisfying $a^2 - Aa + q \equiv 0 \mod L$.*
(2) *$q + 1 \equiv A \mod MN_0p^\nu$.*

Then we have the following conclusions.

(1) *Whatever the order R, $E/\mathbb{F}_q$ admits precisely $\phi(p^\nu)$ $Ig(p^\nu)$-structures.*
(2) *If R has co-conductor prime to L, then $E/\mathbb{F}_q$ admits at least one $\Gamma_0(L)$-structure.*
(3) *If R has co-conductor prime to M, then $E/\mathbb{F}_q$ admits precisely $\phi(M)$ $\Gamma_1(M)$-structures.*
(4) *If R has co-conductor divisible by N_0, then $E/\mathbb{F}_q$ admits precisely $\#SL(2, \mathbb{Z}/N_0\mathbb{Z})$ oriented $\Gamma(N_0)$-structures. Otherwise, $E/\mathbb{F}_q$ admits none.*

Proof. (1) Since $E/\mathbb{F}_q$ is ordinary, the group $E(\overline{\mathbb{F}_q})[p^\infty]$ is noncanonically $\mathbb{Q}_p/\mathbb{Z}_p$. So the p-power torsion subgroup of $E(\mathbb{F}_q)$ is cyclic, and its order is the highest power of p which divides $\#E(\mathbb{F}_q) = q + 1 - A$. Because this cardinality is divisible by p^ν, $E(\mathbb{F}_q)[p^\nu]$ is cyclic of order p^ν, and its $\phi(p^\nu)$ generators are precisely the $Ig(p^\nu)$-structures on $E/\mathbb{F}_q$.

(2) and (3) The existence of a $\Gamma_0(L)$ (resp. $\Gamma_1(M)$) structure depends only upon R/LR (resp. R/MR) as a $\mathbb{Z}[F]$-module. If R has co-conductor prime to L (resp. M), then the inclusion $\mathbb{Z}[F] \subset R$ induces a $\mathbb{Z}[F]$-isomorphism $\mathbb{Z}[F]/L\mathbb{Z}[F] \cong R/LR$ (resp. $\mathbb{Z}[F]/M\mathbb{Z}[F] \cong R/MR$). So it suffices to treat the single case when $R = \mathbb{Z}[F]$. We will now show in $\mathbb{Z}[F]/L\mathbb{Z}[F]$ (resp. $\mathbb{Z}[F]/M\mathbb{Z}[F]$) that the kernel of $F - a$ (resp. $F - 1$) is a cyclic subgroup of order L (resp. M). Once we show this, then the kernel of $F - a$ in $\mathbb{Z}[F]/L\mathbb{Z}[F]$ is the asserted $\Gamma_0(L)$-structure, and the $\phi(M)$ generators of the kernel of $F - 1$ in $\mathbb{Z}[F]/M\mathbb{Z}[F]$ are all the $\Gamma_1(M)$-structures. The assertion about the kernels results from the fact (elementary divisors) that for an endomorphism Λ of a finite free $\mathbb{Z}/L\mathbb{Z}$-module (resp. of a finite free $\mathbb{Z}/M\mathbb{Z}$-module), $\mathrm{Ker}(\Lambda)$ and $\mathrm{Coker}(\Lambda)$ are isomorphic abelian groups. [In fact, as Bill Messing explained to me, the kernel and cokernel of an endomorphism of any finite abelian group are isomorphic abelian groups, but we will not need that finer statement

here.] Applying this to the endomorphisms $F - a$ of $\mathbb{Z}[F]/L\mathbb{Z}[F]$ and $F - 1$ of $\mathbb{Z}[F]/M\mathbb{Z}[F]$, we find that the relevant kernels are the cyclic groups underlying the quotient rings

$$\begin{aligned}\mathbb{Z}[F]/(L, F-a) &:= \mathbb{Z}[X]/(L, X^2 - aX + q, X - a)\\ &\cong \mathbb{Z}/(L, a^2 - aA + q) \cong \mathbb{Z}/L\mathbb{Z}\end{aligned}$$

and

$$\begin{aligned}\mathbb{Z}[F]/(M, F-1) &:= \mathbb{Z}[X]/(M, X^2 - aX + q, X - 1)\\ &\cong \mathbb{Z}/(M, 1 - A + q) \cong \mathbb{Z}/M\mathbb{Z}.\end{aligned}$$

(4) We have $q \equiv 1 \bmod N_0$ because $\mathbb{F}_q$ contains a primitive N_0th root of unity; by assumption N_0^2 divides $q + 1 - A$. We must show that all the points of order dividing N_0 are $\mathbb{F}_q$-rational if and only if R has co-conductor divisible by N_0. All the points of order dividing N_0 are $\mathbb{F}_q$-rational if and only if $F - 1$ kills R/NR, i.e., if and only if $(F-1)/N$, which a priori lies in the fraction field of $\mathcal{O}$, lies in R. [Let us remark in passing that in order for $(F-1)/N$ to lie in $\mathcal{O}$, it is necessary and sufficient that its norm and trace down to $\mathbb{Q}$ lie in $\mathbb{Z}$. But its norm down to $\mathbb{Q}$ is $(q+1-A)/N_0^2$ and its trace down to $\mathbb{Q}$ is $(A-2)/N_0 = (q-1)/N_0 + (A-q-1)/N_0$.] Thus there exist $\Gamma(N_0)$-structures if and only if R contains the order $\mathbb{Z}[(F-1)/N_0]$. This last order visibly has co-conductor N_0, so the orders containing it are precisely those whose co-conductor is divisible by N_0. Once any (possibly unoriented) $\Gamma(N_0)$ structure exists, there are precisely $\#SL(2, \mathbb{Z}/N_0\mathbb{Z})$ oriented $\Gamma(N_0)$-structures. □

Remark 4.4. In the above lemma, we don't specify how many $\Gamma_0(L)$-structures there are,"even" when R has co-conductor prime to L, and we don't say when any exist for other R. We also don't say how many $\Gamma_1(M)$-structures there are for other R. For these R, we will be able to make do with the trivial inequalities, valid for any R,

$$\begin{aligned}0 \le \#\{\Gamma_0(L)\text{-structures on } R/LR\} &\le \#\mathbb{P}^1(\mathbb{Z}/L\mathbb{Z}),\\ 0 \le \#\{\Gamma_1(M)\text{-structures on } R/MR\} &\le \phi(M)\#\mathbb{P}^1(\mathbb{Z}/M\mathbb{Z}).\end{aligned}$$

5. Interlude: Brauer–Siegel for quadratic imaginary orders

The following minor variant of Siegel's theorem for quadratic imaginary fields is certainly well known to the specialists. We give a proof here for lack of a suitable reference. For a quadratic imaginary order, i.e., an order R in a quadratic imaginary field, we denote by d_R its discriminant, by $h(R) := \#\operatorname{Pic}(R)$ its class number, and by

$$h^\star(R) := h(R)/\#R^\times$$

its "normalized" class number. [We should warn the reader that in Gekeler [Ge, 2.13, 2.14] his $h^\star$ and his $H^\star$ are twice ours.]

Theorem 5.1. *Given a real $\epsilon > 0$, there exists a real constant $C_\epsilon > 0$ such that for any quadratic imaginary order R with $|d_R| \ge C_\epsilon$, we have the inequalities*

$$|d_R|^{\frac{1}{2}-\epsilon} \le h^\star(R) \le |d_R|^{\frac{1}{2}+\epsilon}.$$

Proof. Given a quadratic imaginary order R, denote by f_R its conductor, K its fraction field, and $\mathcal{O}_K$ the ring of integers of K. Then the discriminant d_R of $R = \mathbb{Z} + f_R\mathcal{O}_K$ is related to the discriminant $d_{\mathcal{O}_K}$ by the simple formula

$$d_R = f_R^2 d_{\mathcal{O}_K}.$$

Their normalized class numbers are related as follows (cf. [Cox, 7.2.6 and exc. 7.30(a)] or [Sh, p. 105, exc. 4.12]):

$$\frac{h^\star(R)}{h^\star(\mathcal{O}_K)} = \frac{\#(\mathcal{O}_K/f_R\mathcal{O}_K)^\times}{\#(Z/f_R\mathbb{Z})^\times}.$$

We rewrite this as follows. Given the quadratic imaginary field K, denote by χ_K the associated Dirichlet character: for a prime number p, $\chi_K(p) := 1$ if p splits in K, $\chi_K(p) := 0$ if p ramifies in K, and $\chi_K(p) := -1$ if p is inert in K. We then define the multiplicative function ϕ_K on strictly positive integers by

$$\phi_K(1) = 1,\ \phi_K(nm) = \phi_K(n)\phi_K(m) \text{ if } \gcd(n,m) = 1,$$
$$\phi_K(p^\nu) = p^{\nu-1}(p - \chi_K(p)) \text{ if } \nu \geq 1.$$

In terms of this function, we can rewrite the relation of normalized class numbers as

$$h^\star(R) = \phi_K(f_R)h^\star(\mathcal{O}_K).$$

By Siegel's theorem, applied with $\epsilon/2$, there exist real constants $A_\epsilon > 0$ and $B_\epsilon > 0$ such that for all quadratic imaginary fields K we have

$$(**_{\epsilon/2}):\ A_\epsilon |d_{\mathcal{O}_K}|^{\frac{1}{2}-\epsilon/2} \leq h^\star(\mathcal{O}_K) \leq B_\epsilon |d_{\mathcal{O}_K}|^{\frac{1}{2}+\epsilon/2}.$$

[This is true without A and B for $|d|$ large; A and B take care of the small $|d|$. Conversely, if we know $(**_{\epsilon/2})$ for all $|d|$, we get $(**_\epsilon)$ for large $|d|$ with $A = B = 1$.]

In view of the formulas

$$h^\star(R) = \phi_K(f_R)h^\star(\mathcal{O}_K)$$

and

$$d_R = f_R^2 d_{\mathcal{O}_K},$$

it suffices to show that there exist real constants $A'_\epsilon > 0$ and $B'_\epsilon > 0$ such that for every quadratic imaginary field K and every integer $f \geq 1$, we have

$$A'_\epsilon f^{1-\epsilon} \leq \phi(f) \leq B'_\epsilon f^{1+\epsilon}.$$

In view of the definition of ϕ_K, this is immediate from the following two observations. First, for large (how large depending on ϵ) primes p, we have

$$p^{1-\epsilon} \leq p - 1 \leq \phi_K(p) \leq p + 1 \leq p^{1+\epsilon}.$$

Second, for the finitely many, say N, small primes p where this fails, we can find real constants $A''_\epsilon > 0$ and $B''_\epsilon > 0$ such that

$$A''_\epsilon p^{1-\epsilon} \leq p - 1 \leq \phi_K(p) \leq p + 1 \leq B''_\epsilon p^{1+\epsilon}$$

holds for these N primes. We define

$$A'_\epsilon := (A''_\epsilon)^N, \qquad B'_\epsilon := (B''_\epsilon)^N.$$

Then we have the desired inequality

$$A'_\epsilon f^{1-\epsilon} \leq \phi_K(f) \leq B'_\epsilon f^{1+\epsilon}.$$

Once we have this, we combine it with Siegel's theorem for quadratic imaginary fields to conclude that for every quadratic imaginary order R we have

$$A_\epsilon A'_\epsilon |d_R|^{\frac{1}{2}-\epsilon/2} \leq h^\star(R) \leq B_\epsilon B'_\epsilon |d_R|^{\frac{1}{2}+\epsilon/2}.$$

Then as soon as $|d_R|$ is large enough that

$$1 \leq A_\epsilon A'_\epsilon |d_R|^{\epsilon/2}$$

and

$$B_\epsilon B'_\epsilon |d_R|^{-\epsilon/2} \leq 1,$$

we get the assertion of the theorem. □

It is also convenient to introduce the (normalized) Kronecker class number of a quadratic imaginary order R, $H^\star(R)$, defined as the sum of the normalized class numbers of all orders between R and the ring of integers $\mathcal{O}$ in its fraction field:

$$H^\star(R) := \sum_{\text{orders } R \subset R' \subset \mathcal{O}} h^\star(R').$$

Corollary 5.2. *Given a real $\epsilon > 0$, there exists a real constant $C_\epsilon > 0$ such that for any quadratic imaginary order R with $|d_R| \geq C_\epsilon$, we have the inequalities*

$$|d_R|^{\frac{1}{2}-\epsilon} \leq H^\star(R) \leq |d_R|^{\frac{1}{2}+\epsilon}.$$

Proof. We trivially have $H^\star(R) \geq h^\star(R)$, so we get the asserted lower bound for $H^\star(R)$. To get the lower bound, recall from the proof of the previous theorem that for any quadratic imaginary order R', we have

$$h^\star(R') \leq B_\epsilon B'_\epsilon |d_{R'}|^{\frac{1}{2}+\epsilon/2}.$$

So we get

$$H^\star(R) \leq \sum_{\text{orders } R \subset R' \subset \mathcal{O}} B_\epsilon B'_\epsilon |d_{R'}|^{\frac{1}{2}+\epsilon/2}.$$

The co-conductors $f^c_{R'} := f_R/f_{R'}$ of these intermediate orders with respect to R are precisely the divisors of f_R, and we have

$$d_{R'} = d_R/(f^c_{R'})^2.$$

Thus we have

$$H^\star(R) \leq \sum_{n | f_R} B_\epsilon B'_\epsilon |d_R/n^2|^{\frac{1}{2}+\epsilon/2}.$$

But the sum

$$\sum_{n \geq 1} 1/n^{1+\epsilon}$$

converges, to $\zeta(1+\epsilon)$, so we get the inequality

$$H^\star(R) \leq B_\epsilon B'_\epsilon \zeta(1+\epsilon) |d_R|^{\frac{1}{2}+\epsilon/2}$$

for all quadratic imaginaries R, and we need only take $|d_R|$ large enough that

$$B_\epsilon B'_\epsilon \zeta(1+\epsilon) |d_R|^{-\epsilon/2} \leq 1$$

to insure the asserted upper bound. □

6. Point-count estimates

We now return to the modular curve $\mathcal{M}^{\mathrm{ord}}/k$. Recall that we fix a characteristic $p > 0$, three prime-to-p positive integers (L, M, N_0) and a power $p^\nu \geq 1$ of p. We assume that (L, M, N_0) are pairwise relatively prime. We assume that either $M \geq 4$ or $N_0 \geq 3$ or $p^\nu \geq 4$. We work over a finite extension $k/\mathbb{F}_p$ given with a primitive N_0th root of unity $\zeta_{N_0} \in k$. We have the smooth, geometrically connected modular curve $\mathcal{M}^{\mathrm{ord}}/k$, which parameterizes isomorphism classes of fibrewise ordinary elliptic curves over k-schemes endowed with a $\Gamma_0(L)$-structure, a $\Gamma_1(M)$-structure, a $\Gamma(N_0)$-structure, and an $Ig(p^\nu)$-structure.

For a finite extension $\mathbb{F}_q/k$, and a prime-to-p integer A with $|A| < 2\sqrt{q}$, we denote by $\mathbb{Z}[F] := \mathbb{Z}[X]/(X^2 - AX + q)$ and by $\mathcal{M}^{\mathrm{ord}}(\mathbb{F}_q, A)$ the set of $\mathbb{F}_q$-points on $\mathcal{M}^{\mathrm{ord}}$ whose underlying ordinary elliptic curve gives rise to the data (A, q). We have already noted, in Lemma 4.2, that $q \equiv 1 \bmod N_0$, and that $\mathcal{M}^{\mathrm{ord}}(\mathbb{F}_q, A)$ is empty unless (A, q) satisfies both of the following conditions:

(1) $X^2 - AX + q$ factors completely mod L,
(2) $A \equiv q + 1 \bmod MN_0^2p^\nu$.

Lemma 6.1. *Denote by $D_0 = D_0(L, M, N_0, p^\nu)$ and $D_1 = D_1(L, M, N_0, p^\nu)$ the nonzero constants*

$$D_0 := \phi(M)\#SL(2, \mathbb{Z}/N_0\mathbb{Z})\phi(p^\nu),$$
$$D_1 := \#\mathbb{P}^1(\mathbb{Z}/L\mathbb{Z})\#\mathbb{P}^1(\mathbb{Z}/M\mathbb{Z})D_0,$$

with the convention that when any of L, M, N_0, p^ν is 1, the corresponding factor is 1. For (A, q) with A prime to p, $|A| < 2\sqrt{q}$, and $q \equiv 1 \bmod N_0$ satisfying the two conditions

(1) $X^2 - AX + q$ *factors completely* mod L,
(2) $A \equiv q + 1 \bmod MN_0^2p^\nu$,

we have the inequalities

$$D_0 h^\star(\mathbb{Z}[(F-1)/N_0]) \leq \#\mathcal{M}^{\mathrm{ord}}(\mathbb{F}_q, A) \leq D_1 H^\star(\mathbb{Z}[(F-1)/N_0]).$$

Proof. This is immediate from Lemma 4.3 and the identity

$$\#\mathcal{M}^{\mathrm{ord}}(\mathbb{F}_q, A) = \sum_{\text{orders } Z[F] \subset R \subset \mathcal{O}} \#\mathcal{M}(A, q, R) h^\star(R).$$

□

Lemma 6.2. *Given a prime-to-p integer A, suppose there exists an $\mathbb{F}_q/k$ with $q > A^2/4$ such that (A, q) satisfies the conditions of the previous lemma. If $p = 2$, suppose further that $q \geq 8$. Then there exist infinitely many powers Q of q such that (A, Q) satisfies these same conditions.*

Proof. We first observe that the "moreover" part of Lemma 4.2, and the assumption that $q \geq 8$ if $p = 2$, insures that $q \geq p^\nu$. So the p-part of the second condition is simply that $A \equiv 1 \bmod p^\nu$, and this will hold whatever power Q we take. The other conditions depend only on $q \bmod LMN_0^2$. As q is invertible mod LMN_0^2, we have $q^e \equiv 1 \bmod LMN_0^2$ for some divisor e of $\phi(LMN_0^2)$. Then every power $Q := q^{1+ne}$, $n \geq 1$ has $Q \equiv q \bmod LMN_0^2$. □

Theorem 6.3. *Given a prime-to-p integer A, suppose there exists an $\mathbb{F}_q/k$ with $q > A^2/4$ such that (A,q) satisfies the conditions of Lemma 6.1. If $p = 2$, suppose further that $q \geq 8$. Given a real number $\epsilon > 0$, there exists a real constant $C(A,\epsilon,\mathcal{M}^{\mathrm{ord}}/k)$ such that whenever $\mathbb{F}_Q/k$ is a finite extension with $Q \geq C(A,\epsilon,\mathcal{M}^{\mathrm{ord}}/k)$ such that (A,Q) satisfies the conditions of Lemma 6.1, then we have the inequalities*

$$Q^{\frac{1}{2}-\epsilon} \leq \#\mathcal{M}^{\mathrm{ord}}(A,Q) < Q^{\frac{1}{2}+\epsilon}.$$

Proof. This is immediate from Lemma 6.1 and the Brauer–Siegel inequalities: the discriminant of $\mathbb{Z}[(F-1)/N_0]$, for F relative to $\mathbb{F}_Q$, is $(A^2-4Q)/N_0^2$, and A and N_0 are fixed while Q grows. □

We now explain how to pass from estimates for $\mathbb{F}_Q$-points to estimates for closed points of norm Q, with given A. Denote by $\mathcal{M}^{\mathrm{ord}}_{\mathrm{closed}}(A,Q)$ the set of closed points of norm Q giving rise to (A,Q), and by

$$\mathcal{M}^{\mathrm{ord}}(A,Q)^{\mathrm{prim}} \subset \mathcal{M}^{\mathrm{ord}}(A,Q)$$

the subset of those $\mathbb{F}_Q$-points which, viewed simply as points in $\mathcal{M}^{\mathrm{ord}}(\mathbb{F}_Q)$, come from no proper subfield $k \subset \mathbb{F}_{Q_1} \subsetneqq \mathbb{F}_Q$. As noted at the end of Section 3, we have

$$\#\mathcal{M}^{\mathrm{ord}}_{\mathrm{closed}}(A,Q) = \#\mathcal{M}^{\mathrm{ord}}(A,Q)^{\mathrm{prim}}/\log_{\#k}(Q).$$

So our basic task is to estimate $\#\mathcal{M}^{\mathrm{ord}}(A,Q)^{\mathrm{prim}}$.

Lemma 6.4. *Let A be a prime-to-p integer, Q a prime power, and $\mathbb{F}_q \subset \mathbb{F}_Q$ a subfield. There exists a list, depending on (A,Q,q), of at most six integers a such that if $E_0/\mathbb{F}_q$ is an elliptic curve with $\#E_0(\mathbb{F}_Q) = Q+1-A$, then $\#E_0(\mathbb{F}_q) = q+1-a$ for some a on the list.*

Proof. Since A is prime to p, any such $E_0/\mathbb{F}_q$ becomes ordinary over $\mathbb{F}_Q$, so is already ordinary. Denote by $n := \deg(\mathbb{F}_Q/\mathbb{F}_q)$, by F the Frobenius of $E_0 \otimes_{\mathbb{F}_q} \mathbb{F}_Q//\mathbb{F}_Q$, and by F_0 the Frobenius of $E_0/\mathbb{F}_q$. We have an inclusion of orders

$$\mathbb{Z}[F] \subset \mathbb{Z}[F_0].$$

These orders have the same fraction field K, and in K we have $(F_0)^n = F$. But K is quadratic imaginary, so it contains at most 6 roots of unity. So if F, a root of $X^2 - AX + q$ in K, has any nth roots in K, it has at most 6, since the ratio of any two is a root of unity in K. The list is then the list of traces, down to $\mathbb{Q}$, of all the nth roots of F. □

In fact, we will need only the following standard fact, whose proof we leave to the reader.

Lemma 6.5. *Let A be an integer, q a prime power, and $Q = q^2$. If $E_0/\mathbb{F}_q$ is an elliptic curve with $\#E_0(\mathbb{F}_{q^2}) = q^2 + 1 - A$, then $\#E_0(\mathbb{F}_q) = q + 1 - a$ with a one of the two roots of $X^2 - 2q = A$.*

Theorem 6.6. *Given a prime-to-p integer A, suppose there exists an $\mathbb{F}_q/k$ with $q > A^2/4$ such that (A,q) satisfies the conditions of Lemma 6.1. If $p = 2$, suppose further that $q \geq 8$. Given a real number $\epsilon > 0$, there exists a real constant $C'(A,\epsilon,\mathcal{M}^{\mathrm{ord}}/k)$ such that whenever $\mathbb{F}_Q/k$ is a finite extension with $Q \geq C'(A,\epsilon,\mathcal{M}^{\mathrm{ord}}/k)$ such that (A,Q) satisfies the conditions of Lemma 6.1, then we have the inequalities*

$$Q^{\frac{1}{2}-\epsilon} \leq \#\mathcal{M}^{\mathrm{ord}}(A,Q)^{\mathrm{prim}} < Q^{\frac{1}{2}+\epsilon}.$$

Proof. The statement only gets harder as ϵ shrinks, so it suffices to treat the case when $0 < \epsilon < 1/10$. If the degree of $\mathbb{F}_Q$ over k is odd, we will use only the trivial inequality

$$\#\mathcal{M}^{\mathrm{ord}}(A,Q) - \#\mathcal{M}^{\mathrm{ord}}(A,Q)^{\mathrm{prim}} \leq \sum_{k\subset\mathbb{F}_q\subsetneq\mathbb{F}_Q} \#\mathcal{M}^{\mathrm{ord}}(\mathbb{F}_q).$$

Whatever the value of q, we have a uniform upper bound of the form

$$\#\mathcal{M}^{\mathrm{ord}}(\mathbb{F}_q) \leq \sigma q,$$

for σ the sum of the Betti numbers of $\mathcal{M}^{\mathrm{ord}} \otimes_k \overline{k}$. But if $\deg(\mathbb{F}_Q/k)$ is odd, each of the at most $\log_{\#k}(Q)$ terms is at most $\sigma Q^{\frac{1}{3}}$, so this error is, for large Q, negligible with respect to $Q^{\frac{1}{2}-\epsilon}$.

If the degree of $\mathbb{F}_Q$ over k is even, we can still use the above crude argument to take care of imprimitive points which come from a subfield $k \subset \mathbb{F}_q \subsetneq \mathbb{F}_Q$ with $\deg(\mathbb{F}_Q/\mathbb{F}_q) \geq 3$.

But we must be more careful about imprimitive points in $\#\mathcal{M}^{\mathrm{ord}}(A,Q)$ which come from the subfield $\mathbb{F}_q \subset \mathbb{F}_Q$ over which $\mathbb{F}_Q$ is quadratic. If $X^2 - 2q = A$ has no integer solutions, then there are no such imprimitive points. If $X^2 - 2q = A$ has integer solutions, say $\pm a$, then the number of such imprimitive points in $\#\mathcal{M}^{\mathrm{ord}}(A,Q)$ is

$$\#\mathcal{M}^{\mathrm{ord}}(a,q) + \#\mathcal{M}^{\mathrm{ord}}(-a,q).$$

If we take Q so large that $\sqrt{Q}$ is large enough for Theorem 6.3 to apply to the sets $\mathcal{M}^{\mathrm{ord}}(\pm a,q)$, then these sets have size at most $Q^{\frac{1}{4}+\frac{\epsilon}{2}}$, again negligible with respect to $Q^{\frac{1}{2}-\epsilon}$. □

Combining this with the identity

$$\#\mathcal{M}^{\mathrm{ord}}_{\mathrm{closed}}(A,Q) = \#\mathcal{M}^{\mathrm{ord}}(A,Q)^{\mathrm{prim}}/\log_{\#k}(Q),$$

and noting that $\log_{\#k}(Q)$ is negligible with respect to Q^{ϵ}, we get the following corollary.

Corollary 6.7. *Given a prime-to-p integer A, suppose there exists an $\mathbb{F}_q/k$ with $q > A^2/4$ such that (A,q) satisfies the conditions of Lemma 6.1. If $p = 2$, suppose further that $q \geq 8$. Given a real number $\epsilon > 0$, there exists a real constant $C''(A,\epsilon,\mathcal{M}^{\mathrm{ord}}/k)$ such that whenever $\mathbb{F}_Q/k$ is a finite extension with $Q \geq C''(A,\epsilon,\mathcal{M}^{\mathrm{ord}}/k)$ such that (A,Q) satisfies the conditions of Lemma 6.1, then we have the inequalities*

$$Q^{\frac{1}{2}-\epsilon} \leq \#\mathcal{M}^{\mathrm{ord}}_{\mathrm{closed}}(A,Q) < \#\mathcal{M}^{\mathrm{ord}}(A,Q) < Q^{\frac{1}{2}+\epsilon}.$$

To end this section, we interpret its results in terms of the mod N Galois images $G_N := \rho_N(\pi_1(\mathcal{M}^{\mathrm{ord}}))$ and their subsets $G_N(A,Q) \subset G_N$ introduced in Section 2.

Theorem 6.8. *Given a prime-to-p integer A, suppose that for the single value $N := LMN_0^2p^\nu$, $A \mod N$ is the trace of some element of G_N. Then there exist infinitely many closed points $\mathcal{P}$ of $\mathcal{M}^{\mathrm{ord}}$ with $A_{\mathcal{P}} = A$.*

Proof. By Chebotarev, every conjugacy class in G_N is the image of $\mathrm{Frob}_{\mathcal{P}}$ for infinitely many closed points $\mathcal{P}$. In particular, every conjugacy class in G_N is the image of some $\mathrm{Frob}_{\mathcal{P}}$ with $\mathbb{N}(\mathcal{P}) := Q \geq \mathrm{Max}(A^2/4, 8)$. By Lemma 4.2, we have $Q \geq p^\nu$, and $(A_{\mathcal{P}}, Q)$ satisfies the two conditions of that lemma, namely

(1) $X^2 - A_{\mathcal{P}}X + Q$ factors completely mod L,
(2) $A_{\mathcal{P}} \equiv Q + 1 \bmod MN_0^2 p^\nu$.

But $A \equiv A_{\mathcal{P}} \bmod N$, and hence (A, Q) satisfies these same two conditions. The result now follows from Lemma 6.2 and Corollary 6.7, applied to (A, Q). □

Similarly, we have the following result.

Theorem 6.9. *Given a prime-to-p integer A and a power q of $\#k$ with $q \geq \mathrm{Max}(A^2/4, 8)$, suppose that for the single value $N := LMN_0^2 p^\nu$, the subset $G_N(A, q) \subset G_N$ is nonempty. Then there exist infinitely many closed points $\mathcal{P}$ of $\mathcal{M}^{\mathrm{ord}}$ with $A_{\mathcal{P}} = A$ and with $\mathbb{N}(\mathcal{P}) \equiv q \bmod LMN_0^2$.*

Proof. Pick an element γ in $G_N(A, q)$; its conjugacy class in G_N is the image of $\mathrm{Frob}_{\mathcal{P}}$ for infinitely many closed points $\mathcal{P}$, so is the image of some $\mathrm{Frob}_{\mathcal{P}}$ with $\mathbb{N}(\mathcal{P}) := Q \geq \mathrm{Max}(A^2/4, 8)$. Exactly as in the proof of the theorem above, $Q \geq p^\nu$ and $(A_{\mathcal{P}}, Q)$ satisfies the two conditions of Lemma 4.2. We write these now as three conditions, breaking the second one into a prime-to-p part and a p-part:

(1) $X^2 - A_{\mathcal{P}}X + Q$ factors completely mod L,
(2a) $A_{\mathcal{P}} \equiv Q + 1 \bmod MN_0^2$,
(2b) $A_{\mathcal{P}} \equiv Q + 1 \bmod p^\nu$.

As $\mathrm{Frob}_{\mathcal{P}}$ lands in $G_N(A, Q)$, we have the congruences $A \equiv A_{\mathcal{P}} \bmod N$ and $Q \equiv q \bmod LMN_0^2$. Both Q and q are $0 \bmod p^\nu$. Hence (A, Q) satisfies these same conditions, and we conclude as above. □

7. Exact and approximate determination of Galois images

If we take the inverse limit of the mod N representations as N grows multiplicatively, we get a representation

$$\rho = \rho_{not\ p} \times \rho_{p^\infty} : \pi_1(\mathcal{M}^{\mathrm{ord}}) \to GL(2, \hat{\mathbb{Z}}_{not\ p})_{det\ in\ (\#k)^{\hat{\mathbb{Z}}}} \times \mathbb{Z}_p^\times.$$

Here $\hat{\mathbb{Z}}_{not\ p} := \prod_{\ell \neq p} \mathbb{Z}_\ell$, and $(\#k)^{\hat{\mathbb{Z}}}$ is the closed subgroup of $(\hat{\mathbb{Z}}_{not\ p})^\times$ profinitely generated by $\#k$. Under this representation, the geometric fundamental group lands in $SL(2, \hat{\mathbb{Z}}_{not\ p}) \times \mathbb{Z}_p^\times$.

The following theorem is certainly well known to the specialists. We give a proof for lack of a suitable reference.

Theorem 7.1. *In suitable bases, the image group*

$$\rho(\pi_1(\mathcal{M}^{\mathrm{ord}})) \subset GL(2, \hat{\mathbb{Z}}_{not\ p})_{det\ in\ (\#k)^{\hat{\mathbb{Z}}}} \times \mathbb{Z}_p^\times$$

consists of those elements which $\bmod L$ have the shape $(\star, 0, \star, \star)$, $\bmod M$ have the shape $(1, 0, \star, \star)$, $\bmod N_0$ have the shape $(1, 0, 0, 1)$, $\bmod p^\nu$ have the shape (1) (i.e., the p-component is $1 \bmod p^\nu$). The image of the geometric fundamental group,

$$\rho(\pi_1^{\mathrm{geom}}(\mathcal{M}^{\mathrm{ord}})) \subset SL(2, \hat{\mathbb{Z}}_{not\ p}) \times \mathbb{Z}_p^\times,$$

is just the intersection of $\rho(\pi_1(\mathcal{M}^{\mathrm{ord}}))$ with $SL(2, \hat{\mathbb{Z}}_{not\ p}) \times \mathbb{Z}_p^\times$; i.e., it consists of those elements of $SL(2, \hat{\mathbb{Z}}_{not\ p}) \times \mathbb{Z}_p^\times$ with the imposed shapes.

Proof. In a basis adapted to the imposed level structures, every element of the image $\rho(\pi_1(\mathcal{M}^{\mathrm{ord}}))$ has the asserted shapes. To see this, denote by K the function field of $\mathcal{M}^{\mathrm{ord}}$, and by $\overline{K}$ an algebraic closure of K. Viewing $\overline{\eta} := \mathrm{Spec}(\overline{K})$ as a

geometric generic point of $\mathcal{M}^{\mathrm{ord}}$, $\pi_1(\mathcal{M}^{\mathrm{ord}})$ viewed as $\pi_1(\mathcal{M}^{\mathrm{ord}}, \overline{\eta})$ is a quotient of $\mathrm{Aut}(\overline{K}/K)$, and ρ on $\mathrm{Aut}(\overline{K}/K)$ is the action of this group on the profinite group

$$\prod_{all\ primes\ \ell} T_\ell(\mathcal{E}^{\mathrm{univ}}(\overline{K})) = T_{not\ p}(\mathcal{E}^{\mathrm{univ}}(\overline{K})) \times T_p(\mathcal{E}^{\mathrm{univ}}(\overline{K})).$$

Here $T_{not\ p}(\mathcal{E}^{\mathrm{univ}}(\overline{K}))$ is a free $\hat{\mathbb{Z}}_{not\ p}$-module of rank 2, and $T_p(\mathcal{E}^{\mathrm{univ}}(\overline{K}))$ is a free $\mathbb{Z}_p$-module of rank one. Then take any $\mathbb{Z}_p$-basis of T_p, and any $\hat{\mathbb{Z}}_{not\ p}$-basis of $T_{not\ p}$ adapted to the imposed $\Gamma_0(L)$, $\Gamma_1(M)$, and $\Gamma(N_0)$-structures. Then $\mathrm{Aut}(\overline{K}/K)$ acts through elements of the asserted shape.

We next explain that it suffices to show that the image

$$\rho(\pi_1^{\mathrm{geom}}(\mathcal{M}^{\mathrm{ord}})) \subset SL(2, \hat{\mathbb{Z}}_{not\ p}) \times \mathbb{Z}_p^\times$$

is as asserted. For if $F \in \pi_1(\mathcal{M}^{\mathrm{ord}}))$ is any element of determinant $\#k$, then $\rho(\pi_1(\mathcal{M}^{\mathrm{ord}}))$ is the semidirect product of $\rho(\pi_1^{\mathrm{geom}}(\mathcal{M}^{\mathrm{ord}}))$ with the $\hat{\mathbb{Z}}$ generated by F. [Such elements F exist: if $\mathcal{M}^{\mathrm{ord}}$ has a k-point, take its Frobenius, otherwise take the ratio of the Frobenii at two closed points whose large degrees differ by one.] The key point is that $\rho(F)$ is an element of $GL(2, \hat{\mathbb{Z}}_{not\ p})_{det\ in\ (\#k)^{\hat{\mathbb{Z}}}} \times \mathbb{Z}_p^\times$ of the asserted shape. By the explicit description of $\rho(\pi_1^{\mathrm{geom}}(\mathcal{M}^{\mathrm{ord}}))$, this semidirect product itself has the asserted description.

That the image

$$\rho(\pi_1^{\mathrm{geom}}(\mathcal{M}^{\mathrm{ord}})) \subset SL(2, \hat{\mathbb{Z}}_{not\ p}) \times \mathbb{Z}_p^\times$$

is as asserted is a geometric statement, so we may extend scalars from k to $\overline{k}$. Suppose first that either $M \geq 4$ or that $N_0 \geq 3$. In that case we can consider the moduli problem $\mathcal{M}_0/\overline{k}$, where we require $\Gamma_0(L)$, $\Gamma_1(M)$, and (oriented) $\Gamma(N_0)$-structures, but no longer impose either ordinarity or any further condition on p-power torsion. Suppose we know that for $\mathcal{M}_0^{\mathrm{ord}}$, the image

$$\rho(\pi_1^{\mathrm{geom}}(\mathcal{M}_0^{\mathrm{ord}})) \subset SL(2, \hat{\mathbb{Z}}_{not\ p}) \times \mathbb{Z}_p^\times$$

is as asserted. Then we argue as follows. By a fundamental theorem of Igusa, cf. [Ig] and [K-M, 12.6.2], at any supersingular point $s \in \mathcal{M}_0(\overline{k})$, the p-adic character ρ_{p^∞} restricted to the inertia group I_s at s has largest possible image:

$$\rho_{p^\infty}(I_s) = \mathbb{Z}_p^\times.$$

Therefore the covering $\mathcal{M}^{\mathrm{ord}} \to \mathcal{M}_0^{\mathrm{ord}}$ is finite étale Galois, with group $(\mathbb{Z}/p^\nu\mathbb{Z})^\times$. So $\rho(\pi_1^{\mathrm{geom}}(\mathcal{M}^{\mathrm{ord}})) \subset \rho(\pi_1^{\mathrm{geom}}(\mathcal{M}_0^{\mathrm{ord}}))$ is an open subgroup of index $\phi(p^\nu)$. But $\rho(\pi_1^{\mathrm{geom}}(\mathcal{M}^{\mathrm{ord}}))$ lies in the group it is asserted to be, and **that** group has the same index $\phi(p^\nu)$ in the known $\rho(\pi_1^{\mathrm{geom}}(\mathcal{M}_0^{\mathrm{ord}}))$. So we get the asserted description of $\rho(\pi_1^{\mathrm{geom}}(\mathcal{M}^{\mathrm{ord}}))$.

We now show that for $\mathcal{M}_0^{\mathrm{ord}}$, the image

$$\rho(\pi_1^{\mathrm{geom}}(\mathcal{M}_0^{\mathrm{ord}})) \subset SL(2, \hat{\mathbb{Z}}_{not\ p}) \times \mathbb{Z}_p^\times$$

is as asserted. By Igusa's theorem, at any supersingular point of $\mathcal{M}_0$, $\rho_{p^\infty}(I_s) = \mathbb{Z}_p^\times$. But the representation $\rho_{not\ p}$ is everywhere unramified on $\mathcal{M}_0$, so $\rho(I_s) = \{1\} \times \mathbb{Z}_p^\times$ in the product $SL(2, \hat{\mathbb{Z}}_{not\ p}) \times \mathbb{Z}_p^\times$. So we are reduced to showing that the image

$$\rho_{not\ p}(\pi_1^{\mathrm{geom}}(\mathcal{M}_0^{\mathrm{ord}})) \subset SL(2, \hat{\mathbb{Z}}_{not\ p})$$

is as asserted. This follows from the tame specialization theorem [Ka-ESDE, 8.17.14] and the corresponding result over $\mathbb{C}$. The moduli scheme $\mathcal{M}_0$ is one geometric fibre of the corresponding moduli scheme $\mathcal{M}_0$ over $\mathbb{Z}[\zeta_{N_0}][1/LMN_0]$, which one knows has a proper smooth compactification $\overline{\mathcal{M}_0}$ over $\mathbb{Z}[\zeta_{N_0}][1/LMN_0]$ with "infinity" a divisor which is finite étale over $\mathbb{Z}[\zeta_{N_0}][1/LMN_0]$. Extend scalars from $\mathbb{Z}[\zeta_{N_0}][1/LMN_0]$ to the Witt vectors $W(\overline{k})$. On this scheme $\mathcal{M}_0/W(\overline{k})$, the lisse $\hat{\mathbb{Z}}_{not\ p}$-sheaf which "is" $\rho_{not\ p}$ is (automatically) tamely ramified along "infinity", so by the tame specialization theorem its geometric monodromy is the same on the special fibre as on the geometric generic fibre obtained by choosing (!) an embedding of $W(\overline{k}) \subset \mathbb{C}$.

It remains to show that the image

$$\rho(\pi_1^{\mathrm{geom}}(\mathcal{M}^{\mathrm{ord}})) \subset SL(2, \hat{\mathbb{Z}}_{not\ p}) \times \mathbb{Z}_p^\times$$

is as asserted in the general case, i.e., without the assumption that either $M \geq 4$ or $N_0 \geq 3$. To treat this case, pick two distinct primes ℓ_1 and ℓ_2, both of which are odd and prime to LMN_0p. Consider the moduli problems $\mathcal{M}_1^{\mathrm{ord}}$, respectively $\mathcal{M}_2^{\mathrm{ord}}$, over $\overline{k}$, where in addition to imposing an $\mathcal{M}$-structure we impose also an oriented $\Gamma(\ell_1)$-structure, resp. an oriented $\Gamma(\ell_2)$-structure. The two groups $\rho(\pi_1^{\mathrm{geom}}(\mathcal{M}_i^{\mathrm{ord}}))$, $i = 1, 2$, are then known. Both are subgroups of $\rho(\pi_1^{\mathrm{geom}}(\mathcal{M}_i^{\mathrm{ord}}))$, and together they visibly generate the asserted candidate for$\rho(\pi_1^{\mathrm{geom}}(\mathcal{M}_i^{\mathrm{ord}}))$. □

Using this result, one can say something, again well known to the specialists, in the case of general families.

Theorem 7.2. *Let $\mathcal{E}/U/k$ be a family of fibrewise ordinary elliptic curves with nonconstant j-invariant over a base curve U/k, k a finite field, which is smooth and geometrically connected.*

(1) *The image $\rho(\pi_1^{\mathrm{geom}}(U))$ is open in $SL(2, \hat{\mathbb{Z}}_{not\ p}) \times \mathbb{Z}_p^\times$; there exists an integer $D = D_0 p^\nu$, D_0 prime to p and $\nu \geq 0$, such that $\rho(\pi_1^{\mathrm{geom}}(U))$ contains the subgroup*

$$\mathrm{Ker}(SL(2, \hat{\mathbb{Z}}_{not\ p}) \times \mathbb{Z}_p^\times \to SL(2, \mathbb{Z}/D_0\mathbb{Z}) \times (\mathbb{Z}/p^\nu\mathbb{Z})^\times),$$

which is the kernel of reduction $\bmod (D_0, p^\nu)$.

(2) *Denote by $G_D^{\mathrm{geom}} \subset SL(2, \mathbb{Z}/D_0\mathbb{Z}) \times (\mathbb{Z}/p^\nu\mathbb{Z})^\times$ the* $\bmod D$ *image $\rho_D(\pi_1^{\mathrm{geom}}(U))$, and denote by $G_D \subset GL(2, \mathbb{Z}/D_0\mathbb{Z}) \times (\mathbb{Z}/p^\nu\mathbb{Z})^\times$ the* $\bmod D$ *image $\rho_D(\pi_1(U))$. Then G_D^{geom} is a normal subgroup of G_D, and the quotient is cyclic, generated by the image $\rho_D(F)$ of any element $F \in \pi_1(U))$ such that $\rho_{not\ p}(F)$ has determinant $\#k$.*

(3) *An element $\gamma_0 \in SL(2, \hat{\mathbb{Z}}_{not\ p}) \times \mathbb{Z}_p^\times$ lies in $\rho(\pi_1^{\mathrm{geom}}(U))$ if and only if* $\bmod D$ *it lies in G_D^{geom}.*

(4) *Suppose given an element $\gamma \in GL(2, \hat{\mathbb{Z}}_{not\ p})_{det\ in\ (\#k)^{\hat{\mathbb{Z}}}} \times \mathbb{Z}_p^\times$, with $\det(\gamma_{not\ p}) = (\#k)^n$, $n \in \hat{\mathbb{Z}}$. This element lies in $\rho(\pi_1(U))$ if and only if γ* $\bmod D$ *lies in the coset $\rho_D(F^n)G_D^{\mathrm{geom}}$ of G_D.*

Proof. It suffices to prove (1). For (2) is universally true, and (1) and (2) together imply (3). For (4), the condition given is obviously necessary; applying (3) to $F^{-n}\gamma$ we see that it is sufficient.

To prove (1) we argue as follows. The assertion is geometric, so we may extend scalars from k to $\overline{k}$. It suffices to prove it for some finite étale cover U_1 of U, since

$\pi_1(U_1)$ is a subgroup of $\pi_1(U)$. So we may choose an odd prime $\ell \neq p$ and reduce to the case when $\mathcal{E}/U$ has an oriented $\Gamma(\ell)$-structure. Then we have a classifying map $U \to \mathcal{M}^{\mathrm{ord}}(\ell)$, for $\mathcal{M}(\ell)$ the $\Gamma(\ell)$ modular curve. This map is nonconstant, because our family has nonconstant j-invariant. Therefore the image of $\pi_1(U)$ in $\pi_1(\mathcal{M}^{\mathrm{ord}}(\ell))$ is a closed subgroup of finite index, hence an open subgroup of finite index, in $\pi_1(\mathcal{M}^{\mathrm{ord}}(\ell))$, to which we apply the theorem. □

Given $\mathcal{E}/U/k$ as in the above theorem, fibrewise ordinary with nonconstant j-invariant, we say that any integer D for which part (1) of the corollary holds is a **modulus** for $\mathcal{E}/U/k$. Of course if D is a modulus, so is any multiple of D.

8. Gekeler's product formula

To motivate this section, we first explain the heuristic which underlies it. Given a family $\mathcal{E}/U/\mathbb{F}_q$ with nonconstant j-invariant, we know the Sato–Tate conjecture for this family; cf. Deligne [De-WeilII, 3.5.7]. Given a finite extension $\mathbb{F}_Q/\mathbb{F}_q$, attached to each point $u \in U(\mathbb{F}_Q)$ is an elliptic curve $E_{u,\mathbb{F}_Q}/\mathbb{F}_Q$, which gives rise to data $(A_{u,\mathbb{F}_Q}, Q)$, which in turn gives rise to the real number $t_{u,\mathbb{F}_Q} := A_{u,\mathbb{F}_Q}/(2\sqrt{Q}) \in [-1,1]$. The Sato–Tate theorem says that as Q grows, these $\#U(\mathbb{F}_Q)$ real numbers $t_{u,\mathbb{F}_Q} \in [-1,1]$ become equidistributed for the measure $\frac{2}{\pi}(\sqrt{1-t^2})dt$ on $[-1,1]$. This means that for any continuous $\mathbb{C}$-valued function f on $[-1,1]$, we can compute $\frac{2}{\pi}\int_{-1}^{1} f(t)(\sqrt{1-t^2})dt$ as the large Q limit of $(1/\#U(\mathbb{F}_Q))\sum_{u\in U(\mathbb{F}_q)} f(t_{u,\mathbb{F}_Q})$. For a fixed Q, we make the change of variable $t = A/(2\sqrt{Q})$, so the integral becomes

$$\frac{2}{\pi}\int_{-2\sqrt{Q}}^{2\sqrt{Q}} f(A/(2\sqrt{Q}))(\sqrt{1-A^2/4Q})dA/(2\sqrt{Q})$$
$$= \frac{2}{\pi}\frac{1}{4Q}\int_{-2\sqrt{Q}}^{2\sqrt{Q}} f(A/(2\sqrt{Q}))(\sqrt{4Q-A^2})dA,$$

and the approximating sum is

$$(1/\#U(\mathbb{F}_Q))\sum_{u\in U(\mathbb{F}_q)} f(A_{u,\mathbb{F}_Q}/(2\sqrt{Q})).$$

Since there are "about" Q points in $U(\mathbb{F}_Q)$, it is "as though" a given integer $A \in [-2\sqrt{Q}, 2\sqrt{Q}]$ occurs as an $A_{u,\mathbb{F}_Q}$ for "about"

$$\frac{1}{2\pi}\sqrt{4Q-A^2}$$

of the $u \in U(\mathbb{F}_q)$; this is the Sato–Tate heuristic.

We have already discussed how congruence obstructions can prevent some particular (A,Q) from occurring at all in a given family. The Gekeler product formula says that, at least in certain modular families, whenever (A,Q) is ordinary and has no Archimedean or congruence obstruction, we can use congruence considerations to compute the ratio

$$\frac{\#\{u \in U(\mathbb{F}_Q) \mid A_{u,\mathbb{F}_Q} = A\}}{\frac{1}{2\pi}\sqrt{4Q-A^2}}.$$

The prototypical example of computing such a ratio by "congruence considerations" is Dirichlet's class number formula for a quadatic imaginary field K:

$$L(1,\chi) = \frac{2\pi h_K}{w_K\sqrt{|d_K|}}.$$

Here $h_K = h(\mathcal{O}_K)$ is the class number, $w_K = \#\mathcal{O}_K^\times$ is the order of the group of roots of unity in K, and d_K is the discriminant of $\mathcal{O}_K$. So in terms of the normalized class number

$$h_K^\star := h_K/w_K,$$

the formula reads

$$L(1,\chi) = \frac{h_K^\star}{\frac{1}{2\pi}\sqrt{|d_K|}}.$$

In this example, the ratio is the value $L(1,\chi)$, and "congruence considerations" give the Euler factors, whose conditionally convergent product is $L(1,\chi)$. Indeed, it is precisely this class number formula which underlies Gekeler's, as we will see.

We return now to a family $\mathcal{E}/U/\mathbb{F}_q$ with nonconstant j-invariant. Fix data (A,Q) with A prime to p, $|A| < 2\sqrt{Q}$, and no congruence obstruction. We introduced, for every integer $N \geq 2$, the finite groups G_N, their normal subgroups $G_N^{\mathrm{geom}} \lhd G_N$, the cosets

$$G_{N,\det=Q} \subset G_N,$$

and their subsets

$$G_N(A,Q) \subset G_{N,\det=Q},$$

whose cardinalities were denoted $g_N(A,Q)$ and $g_{N,\det=Q}$. We also introduced the rational number

$$g_N(\mathrm{avg},Q) = (1/N)g_{N,\det=Q} = (1/N)\sum_{A \bmod N} g_N(A,Q).$$

Gekeler's idea is to consider the ratios

$$g_N(A,Q)/g_N(\mathrm{avg},Q) = Ng_N(A,Q)/g_{N,\det=Q},$$

to show they have a "large N limit", and then to show that this limit is the ratio

$$\frac{\#\{u \in U(\mathbb{F}_Q) \mid A_{u,\mathbb{F}_Q} = A\}}{\frac{1}{2\pi}\sqrt{4Q-A^2}}.$$

Let us be more precise. We have the following elementary lemma.

Lemma 8.1. *Let D be a modulus for $\mathcal{E}/U/\mathbb{F}_q$. Suppose we are given (A,Q) with Q a power of $\#k$ and A an integer prime to p with $A^2 < 4Q$. Suppose that (A,Q) has no congruence obstruction. Suppose $N \geq 2$ and $M \geq 2$ are relatively prime. Suppose further that N is relatively prime to D.*

(1) *Under the "reduction mod NM" map we have an isomorphism of groups*

$$G_{NM}^{\mathrm{geom}} \xrightarrow{\sim} G_N^{\mathrm{geom}} \times G_M^{\mathrm{geom}}.$$

(2) *We have a bijection of cosets*

$$G_{NM,\det=Q} \xrightarrow{\sim} G_{N,\det=Q} \times G_{M,\det=Q}.$$

(3) *We have a bijection of sets*

$$G_{NM}(A,Q) \xrightarrow{\sim} G_N(A,Q) \times G_M(A,Q).$$

(4) *For a prime number ℓ prime to D and an integer $n \geq 1$, we have*

$$G_{\ell^n}^{\mathrm{geom}} = SL(2,\mathbb{Z}/\ell^n\mathbb{Z}),$$
$$G_{\ell^n,\det=Q} = \{\gamma \in GL(2,\mathbb{Z}/\ell^n\mathbb{Z}) \mid \det(\gamma) = Q\},$$
$$G_{\ell^n}(A,Q) = \{\gamma \in GL(2,\mathbb{Z}/\ell^n\mathbb{Z}) \mid \det(\gamma) = Q,\ \mathrm{Trace}(\gamma) = A\}.$$

(5) *If the characteristic p is prime to D, then for any $n \geq 1$, we have*

$$G_{p^n}^{\mathrm{geom}} = (\mathbb{Z}/p^n\mathbb{Z})^\times,$$
$$G_{p^n,\det=Q} = (\mathbb{Z}/p^n\mathbb{Z})^\times,$$
$$G_{p^n}(A,Q) = \{\alpha \in (\mathbb{Z}/p^n\mathbb{Z})^\times |\ \alpha \equiv \mathrm{unit}_Q(A) \bmod p^n\}.$$

Proof. Assertions (1), (4) and (5) result from Theorem 7.2. Assertion (1) implies (2) by the definition of $G_{N,\det=Q}$ as a coset, and (2) implies (3) trivially. □

Remark 8.2. In the case of any of the moduli problems we have considered, Theorem 7.1 shows that the image group

$$\rho(\pi_1^{\mathrm{geom}}(\mathcal{M}^{\mathrm{ord}})) \subset SL(2,\hat{\mathbb{Z}}_{not\ p}) \times \mathbb{Z}_p^\times = \prod_{\ell \neq p} SL(2,\mathbb{Z}_\ell) \times \mathbb{Z}_p^\times$$

is the product over all primes ℓ, including $\ell = p$, of the images of the separate ℓ-adic representations. So for (the universal families over) these modular curves, assertions (1), (2) and (3) of Lemma 8.1 hold for every pair (N,M) of relatively prime integers.

Gekeler proves that for a prime number ℓ prime to pD, the sequence of ratios

$$\ell^n g_{\ell^n}(A,Q)/g_{\ell^n,\det=Q},\ n \geq 1,$$

i.e., the sequence of ratios

$$\frac{\ell^n \#\{\gamma \in GL(2,\mathbb{Z}/\ell^n\mathbb{Z}) \mid \mathrm{Trace}(\gamma) = A,\ \det(\gamma) = Q\}}{\#SL(2,\mathbb{Z}/\ell^n\mathbb{Z})},$$

becomes constant for large n, and he computes this constant explicitly [Ge, Thm. 4.4], calling it $\nu_\ell(A,Q)$. He also shows that so long as ℓ is, in addition, either prime to $A^2 - 4Q$ or split in $K := \mathbb{Q}(\sqrt{A^2-4Q})$, then

$$\nu_\ell(A,Q) = 1/(1 - \frac{\chi(\ell)}{\ell})$$

is the Euler factor at ℓ in $L(1,\chi)$ for χ the quadratic character attached to the field K. If the characteristic p does not divide D, it is immediate from the definitions that the sequence

$$p^n g_{p^n}(A,Q)/g_{p^n,\det=Q},\ n \geq 1,$$

is constant, with value

$$(p^n \times 1)/\phi(p^n) = 1/(1 - \frac{1}{p}),$$

the Euler factor at p of the same $L(1,\chi)$.

To go further, we must define a reasonable factor "at D". It should be true that for any sequence of positive integers M_n such that $M_n | M_{n+1}$ and such that $D^n | M_n$ for all $n \geq 1$, the sequence of ratios

$$M_n g_{M_n}(A,Q)/g_{M_n,\det=Q},\ n \geq 1,$$

becomes constant for large n, and that this constant is independent of the particular choice of the sequence of M_n as above. If this is true, we would call this constant $\nu_D(A,Q)$. [For the moduli problems we have been considering, the problem of defining the factor "at D" is reduced to the problem of defining the factor separately at each prime dividing D; cf. Remark 8.2.] The most optimistic conjecture would then be the following.

Conjecture 8.3. *Let $\mathcal{E}/U/\mathbb{F}_q$ be fibrewise ordinary with nonconstant j-invariant, D a modulus. Suppose that $U = U_{\max}$. If (A, Q) has neither Archimedean nor congruence obstruction, then*

$$\frac{\#\{u \in U(\mathbb{F}_Q) \mid A_{u,\mathbb{F}_Q} = A\}}{\frac{1}{2\pi}\sqrt{4Q - A^2}} = \nu_D(A, Q) \prod_{\ell \nmid D} \nu_\ell(A, Q),$$

where the conditionally convergent product is defined by

$$\prod_{\ell \nmid D} \nu_\ell(A, Q) := \lim_{X \to \infty} \prod_{\ell < X, \ell \nmid D} \nu_\ell(A, Q).$$

However, this conjecture is **false** in general, for reasons that emerged in a discussion with Deligne. The problem is that for given (A, Q), it asserts a formula for the number $\#\{u \in U(\mathbb{F}_Q) \mid A_{u,\mathbb{F}_Q} = A\}$ which depends only on the image

$$\rho(\pi_1^{\mathrm{geom}}(U)) \subset SL(2, \hat{\mathbb{Z}}_{not\ p}) \times \mathbb{Z}_p^\times = \prod_{\ell \neq p} SL(2, \mathbb{Z}_\ell) \times \mathbb{Z}_p^\times$$

and on the coset

$$\rho(F_Q \pi_1^{\mathrm{geom}}(U)) \subset \rho(\pi_1(U)) \subset \prod_{\ell \neq p} GL(2, \mathbb{Z}_\ell) \times \mathbb{Z}_p^\times,$$

$F_Q \in \pi_1(U)$ being any element of degree $d := \log Q / \log q$. Consider a situation $\mathcal{E}/U/\mathbb{F}_q$, $U = U_{\max}$, for which the conjecture holds, and for which we have potentially multiplicative reduction at all places of bad reduction. Suppose we have another smooth, geometrically connected curve $V/\mathbb{F}_q$ and a finite flat $\mathbb{F}_q$-map, $f : V \to U$, such that the induced maps of fundamental groups

$$f_\star : \pi_1^{\mathrm{geom}}(V) \to \pi_1^{\mathrm{geom}}(U), f_\star : \pi_1(V) \to \pi_1(U)$$

are both surjective. Now consider the pullback $\mathcal{E}_V/V/\mathbb{F}_q$ of $\mathcal{E}/U/\mathbb{F}_q$ by the map f. This pullback family also has $V = V_{\max}$ (because supersingular points, resp. points of potentially multiplicative reduction, downstairs have their inverse images upstairs supersingular, resp. points of potentially multiplicative reduction). Since the pullback family has the same Galois image data, the notion of congruence obstruction is the same for the original family and for its pullback. So the conjecture predicts that for each (A, Q) with no congruence obstruction, we have

$$\#\{u \in U(\mathbb{F}_Q) \mid A_{u,\mathbb{F}_Q} = A\} = \#\{v \in V(\mathbb{F}_Q) \mid A_{v,\mathbb{F}_Q} = A\}.$$

Fixing Q and summing over the allowed A, which are the same upstairs and down, we get the equality

$$\#U(\mathbb{F}_Q) = \#V(\mathbb{F}_Q).$$

As we will recall below, one condition that forces $f_\star$ to be surjective on fundamental groups is that it be fully ramified over some $\bar{\mathbb{F}}_q$-point, say over $u_0 \in U(\bar{\mathbb{F}}_q)$, with unique point $v_0 \in V(\bar{\mathbb{F}}_q)$ lying over u_0. Given such an f, we are to have

$$\#U(\mathbb{F}_Q) = \#V(\mathbb{F}_Q).$$

But of course this is nonsense in general. Here is the simplest example. Work over a prime field $\mathbb{F}_p$ with $p \equiv 1 \bmod 3$, pick a cube root of unity, and take for $\mathcal{E}/U/\mathbb{F}_p$ the universal family for the moduli problem of oriented $\Gamma(3)$-structures. Here the modular curve $\mathcal{M}_{\Gamma(3)}$ is $\mathbb{P}^1 \setminus \{\mu_3, \infty\}$, with parameter μ and universal family

$$X^3 + Y^3 + Z^3 = 3\mu XYZ.$$

In this family, we have multiplicative reduction at each missing point, and while there are $p-1$ supersingular points over $\bar{\mathbb{F}}_p$, no supersingular point is $\mathbb{F}_p$-rational (simply because over $\mathbb{F}_p$ with $p \geq 5$, supersingular points have $A = 0$, whereas our A's satisfy $A \equiv p+1 \bmod 9$, so are certainly nonzero). Thus

$$\mathcal{M}^{\mathrm{ord}}_{\Gamma(3)}(\mathbb{F}_p) = \mathbb{F}_p \setminus \{\mu_3\}$$

has $p-3$ points. Now consider the pullback family by the double covering "square root of μ" of $U := \mathbb{P}^1 \setminus \{\mu_3, \infty\}$ by $V := \mathbb{P}^1 \setminus \{\mu_6, \infty\}$. Explicitly, this is the family

$$X^3 + Y^3 + Z^3 = 3\mu^2 XYZ.$$

Here

$$V(\mathbb{F}_p) = \mathbb{F}_p \setminus \{\mu_6\}$$

has $p-6$ points.

Here is the precise surjectivity statement used above, whose proof we owe to Deligne. It is not as well known as it should be.

Lemma 8.4. *Let X and Y be connected, locally Noetherian schemes, and let $f : X \to Y$ be a morphism which is proper and flat. Suppose that for some geometric point y of Y, say with values in the algebraically closed field K, the fibre $X_y(K) := f^{-1}(y)(K)$ consists of a single K-valued point $x \in X(K)$. Then the map of fundamental groups*

$$f_\star : \pi_1(X, x) \to \pi_1(Y, y)$$

is surjective.

Proof. For any X-scheme $h : Z \to X$, we can also view Z as a Y-scheme, by $f \circ h$. So we have the fibres $Z_x(K) := h^{-1}(x)(K)$ and $Z_y(K) := (f \circ h)^{-1}(y)(K)$. The hypothesis that $f^{-1}(y)(K) = x$ insures that $Z_x(K) = Z_y(K)$.

Let $g : E \to Y$ be finite étale of some degree $d \geq 1$, with E connected, and denote by $g_X : E_X \to X$ its pullback to a finite étale cover of X, of the same degree d. We must show that E_X remains connected. Let $j : Z \subset E_X$ be the inclusion of a connected component Z of E_X, and let $W := f_E(Z) \subset E$ be its image in E. As f is proper and flat, so is f_E. As Z is both open and closed in E_X, its image $W := f_E(Z) \subset E$ is both closed and open in E, and hence $W = E$.

$$\begin{array}{ccc} Z & \xrightarrow{f_E|Z} & W \\ {\scriptstyle j}\downarrow & & \downarrow{\scriptstyle =} \\ E_X & \xrightarrow{f_E} & E \\ \downarrow{\scriptstyle g_X} & & \downarrow{\scriptstyle g} \\ X & \xrightarrow{f} & Y \end{array}$$

So the fibre $W_y(K)$ of W over y consists of d points. But Z maps onto W; hence, viewing Z as a Y-scheme (by $f \circ g_X \circ j$) and $f_E|Z : Z \to W$ as a Y-morphism, $Z_y(K)$ maps onto $W_y(K)$. Therefore $Z_y(K)$ consists of at least d points. But $Z_x(K) = Z_y(K)$, as noted above; hence $Z_x(K)$ has at least d points. But the entire fibre $(E_X)_x(K)$ has d points. Therefore Z and E_X have the same fibre over x; as both are finite étale over X and $Z \subset E_X$, we conclude that $Z = E_X$. Thus E_X is connected, as required. □

Despite the failure of the conjecture above in general, there are certain modular families of elliptic curves for which the conjecture holds.

Theorem 8.5 (Gekeler)**.** *The conjecture is true for (the universal fibrewise ordinary families over) the Igusa curves $Ig(p^\nu)/k/\mathbb{F}_p$, any $p^\nu \geq 4$.*

Proof. For this family, Theorem 7.1 tells us that

$$\rho(\pi_1^{\mathrm{geom}}(Ig(p^\nu))) = SL(2, \hat{\mathbb{Z}}_{not\ p}) \times (1 + p^\nu \mathbb{Z}_p)$$

and

$$\rho(\pi_1(Ig(p^\nu)) = GL(2, \hat{\mathbb{Z}}_{not\ p})_{det\ in\ (\#k)^{\hat{z}}} \times (1 + p^\nu \mathbb{Z}_p).$$

So here we can take $D = p^\nu$. Take an (A, Q) with neither congruence nor Archimedean obstruction. Then A is prime to p, $\mathrm{unit}_Q(A) \equiv 1 \bmod p^\nu$, and $A^2 < 4Q$. For $\ell \neq p$, the local factor is Gekeler's $\nu_\ell(A, Q)$.

What about the factor $\nu_{p^\nu}(A, Q)$ at $D = p^\nu$? For any integer $n \geq \nu$, $g_{p^n}(A, Q)$ is the ratio

$$\frac{p^n \#\{\gamma \in ((1 + p^\nu \mathbb{Z}_p)/(1 + p^n \mathbb{Z}_p)) \mid \gamma \equiv \mathrm{unit}_Q(A) \bmod p^n\}}{\#((1 + p^\nu \mathbb{Z}_p)/(1 + p^n \mathbb{Z}_p))},$$

which is just $= p^n \times 1/p^{n-\nu} = p^\nu$, which in turn is

$$\phi(p^\nu) \times (1 - 1/p)^{-1},$$

and hence

$$\nu_{p^\nu}(A, Q) = \phi(p^\nu) \times (1 - 1/p)^{-1}.$$

Gekeler proves [Ge, Cor. 5.4] (remember that his $H^\star$ is twice ours and he writes $H^\star(A^2 - 4Q)$ for $H^\star(\mathbb{Z}[F])$) that

$$(1 - 1/p)^{-1} \prod_{\ell \neq p} \nu_\ell(A, Q) = \frac{2\pi H^\star(\mathbb{Z}[F])}{\sqrt{4Q - A^2}}.$$

Hence we have

$$\nu_{p^\nu}(A, Q) \prod_{\ell \neq p} \nu_\ell(A, Q) = \frac{\phi(p^\nu) H^\star(\mathbb{Z}[F])}{\frac{1}{2\pi}\sqrt{4Q - A^2}}.$$

But the numerator $\phi(p^\nu) H^\star(\mathbb{Z}[F])$ is precisely

$$\#\{u \in Ig(p^v)(\mathbb{F}_Q) \mid A_{u, \mathbb{F}_Q} = A\};$$

cf. Lemma 4.3, (1). □

Here is a slight generalization of Gekeler's result.

Theorem 8.6. *Let $N = N_0 p^\nu$ with N_0 prime to p. Supppose that either $N_0 \geq 3$ or that $p^\nu \geq 4$. Let $k/\mathbb{F}_p$ be a finite field containing a chosen primitive N_0th root of unity. Let $\mathcal{E}^{\mathrm{univ}}/\mathcal{M}^{\mathrm{ord}}/k$ be the universal family of ordinary elliptic curves endowed with both an oriented $\Gamma(N_0)$-structure and an $Ig(p^\nu)$-structure. The conjecture is true for this family.*

Proof. Suppose (A, Q) has neither Archimedean nor congruence obstruction. Our first task is to compute, for this family, the local factors $\nu_\ell(A, Q, \mathcal{M}^{\mathrm{ord}})$ at the primes ℓ dividing pN_0.

Suppose we have done this. Then we proceed as follows. We have seen in Lemma 4.3, (1) and (4), that

$$\begin{aligned}&\#\{u \in \mathcal{M}^{\mathrm{ord}}(\mathbb{F}_Q) \mid A_{u,\mathbb{F}_Q} = A\}\\ &= \phi(p^\nu)\#SL(2,\mathbb{Z}/N_0\mathbb{Z})H^\star(\mathbb{Z}[(F-1)/N_0]).\end{aligned}$$

So what we must show is that

$$\begin{aligned}&\frac{\phi(p^\nu)\#SL(2,\mathbb{Z}/N_0\mathbb{Z})H^\star(\mathbb{Z}[(F-1)/N_0])}{\frac{1}{2\pi}\sqrt{4Q-A^2}}\\ &= (\prod_{\ell|pN_0}\nu_\ell(A,Q,\mathcal{M}^{\mathrm{ord}}))(\prod_{\ell\nmid N_0}\nu_\ell(A,Q)).\end{aligned}$$

The factor $\nu_{p^\nu}(A,Q,\mathcal{M}^{\mathrm{ord}})$ at p is

$$\nu_{p^\nu}(A,Q,\mathcal{M}^{\mathrm{ord}}) = \phi(p^\nu)\times(1-1/p)^{-1},$$

exactly as in the proof of the previous theorem. According to that theorem, we have

$$(1-1/p)^{-1}\prod_{\ell\neq p}\nu_\ell(A,Q) = \frac{H^\star(\mathbb{Z}[F])}{\frac{1}{2\pi}\sqrt{4Q-A^2}}.$$

Comparing these two formulas, what we must show is that

$$\prod_{\ell|N_0}\frac{\nu_\ell(A,Q,\mathcal{M}^{\mathrm{ord}})}{\nu_\ell(A,Q)} = \#SL(2,\mathbb{Z}/N_0\mathbb{Z})\frac{H^\star(\mathbb{Z}[(F-1)/N_0])}{H^\star(\mathbb{Z}[F])}.$$

We next express the right-hand side as a product over the primes dividing N_0. Factoring $N_0 = \prod \ell_i^{n_i}$, we have

$$\#SL(2,\mathbb{Z}/N_0\mathbb{Z}) = \prod_{\ell_i|N_0}\#SL(2,\mathbb{Z}/\ell_i^{n_i}\mathbb{Z}).$$

We can factor the ratio

$$\frac{H^\star(\mathbb{Z}[(F-1)/N_0])}{H^\star(\mathbb{Z}[F])}$$

as follows. Denote by K the fraction field of $\mathbb{Z}[F]$, $\mathcal{O}_K$ its ring of integers, χ_K the quadratic character corresponding to $K/\mathbb{Q}$, ϕ_K the multiplicative function introduced in the proof of Theorem 5.1, and

$$f := \text{the conductor of the order } \mathbb{Z}[(F-1)/N_0].$$

We have the formulas

$$H^\star(\mathbb{Z}[(F-1)/N_0]) = \sum_{d|f}\phi_K(d)h^\star(\mathcal{O}_K),$$

$$H^\star(\mathbb{Z}[F]) = \sum_{d|fN_0}\phi_K(d)h^\star(\mathcal{O}_K).$$

Because ϕ_K is multiplicative, we have the formulas

$$\sum_{d|f}\phi_K(d) = \prod_{\ell|f}\sum_{a\geq 0,\ \ell^a|f}\phi_K(\ell^a),$$

$$\sum_{d|fN_0}\phi_K(d) = \prod_{\ell|fN_o}\sum_{a\geq 0,\ \ell^a|fN_0}\phi_K(\ell^a).$$

In these two products, the primes ℓ not dividing N_0 give rise to the same factor in each. So we get

$$\frac{H^\star(\mathbb{Z}[(F-1)/N_0])}{H^\star(\mathbb{Z}[F])} = \prod_{\ell | N_0} \frac{\sum_{a\geq 0,\ \ell^a | f} \phi_K(\ell^a)}{\sum_{a\geq 0,\ \ell^a | fN_0} \phi_K(\ell^a)}.$$

So we are reduced to showing that for each prime ℓ_i dividing $N_0 = \prod \ell_i^{n_i}$ we have

$$\frac{\nu_{\ell_i}(A,Q,\mathcal{M}^{\mathrm{ord}})}{\nu_{\ell_i}(A,Q)} = \#SL(2,\mathbb{Z}/\ell_i^{n_i}\mathbb{Z}) \frac{\sum_{a\geq 0,\ \ell_i^a | f} \phi_K(\ell_i^a)}{\sum_{a\geq 0,\ \ell_i^a | fN_0} \phi_K(\ell_i^a)}.$$

Fix one such $\ell_i := \ell$, and put $n := \mathrm{ord}_\ell(N_0)$, $\delta := \mathrm{ord}_\ell(f)$. Thus $n+\delta = \mathrm{ord}_\ell(fN_0)$. We now compute explicitly everything in sight, using the results [Ge, Thm.4.4] of Gekeler. To state them, we first establish a bit of notation. Given an arbitrary pair (A_1, Q_1) of integers with $A_1^2 - 4Q_1 < 0$, we wish to count the number $\alpha_{\ell^k}(A_1,Q_1)$ of 2×2 matrices $X \in M_2(\mathbb{Z}/\ell^k\mathbb{Z})$ with $\mathrm{Trace}(X) = A_1$ and $\det(X) = Q_1$, at least for k large. Attached to (A_1,Q_1) we have the quadratic imaginary order

$$\mathbb{Z}[F_1] := \mathbb{Z}[T]/(T^2 - A_1T + Q_1),$$

its fraction field K_1, ring of integers $\mathcal{O}_{K_1}$, and Dirichlet character χ_{K_1}. We denote by f_{A_1,Q_1} the conductor of the order $\mathbb{Z}[F_1]$, and we define

$$\delta_1 := \mathrm{ord}_\ell(f_{A_1,Q_1}).$$

Gekeler shows that for $k \geq 2\delta_1 + 2$, $\alpha_{\ell^k}(A_1,Q_1)$ is equal to

$$\begin{gathered}
\ell^{2k} + \ell^{2k-1} \quad \text{if } \chi_{K_1}(\ell) = 1,\\
\ell^{2k} + \ell^{2k-1} - (\ell+1)\ell^{2k-\delta_1-2} \quad \text{if } \chi_{K_1}(\ell) = 0,\\
\ell^{2k} + \ell^{2k-1} - 2\ell^{2k-\delta_1-1} \quad \text{if } \chi_{K_1}(\ell) = -1.
\end{gathered}$$

We apply this first to (A,Q). Then K_1 is just K, $f_{A,Q} = N_0 f$, and $\delta_1 = \delta + n$. So for large k, the factor

$$\nu_{\ell^k}(A,Q) := \ell^k \alpha_{\ell^k}(A,Q)/\#SL(2,\mathbb{Z}/\ell^k\mathbb{Z}) = \alpha_{\ell^k}(A,Q)/\ell^{2k-2}(\ell^2-1)$$

is easily calculated, as $\alpha_{\ell^k}(A,Q)$ is equal to

$$\begin{gathered}
\ell^{2k} + \ell^{2k-1} \quad \text{if } \chi_K(\ell) = 1,\\
\ell^{2k} + \ell^{2k-1} - (\ell+1)\ell^{2k-\delta-n-2} \quad \text{if } \chi_K(\ell) = 0,\\
\ell^{2k} + \ell^{2k-1} - 2\ell^{2k-\delta-n-1} \quad \text{if } \chi_K(\ell) = -1.
\end{gathered}$$

We next calculate the factor $\nu_{\ell^{k+2n}}(A,Q,\mathcal{M}^{\mathrm{ord}})$. Here the group $G_{\ell^{k+2n}}$ is the group of matrices of the shape $1+\ell^n X$, $X \in M_2(\mathbb{Z}/\ell^{n+k}\mathbb{Z})$, and $G_{\ell^{k+2n},\det=Q}$, being a coset of $G^{\mathrm{geom}}_{\ell^{k+2n}} = \{1+\ell^n X\} \cap SL(2,\mathbb{Z}/\ell^{2n+k}\mathbb{Z})$, has

$$\#G_{\ell^{k+2n},\det=Q} = \ell^{3(k+n)}.$$

How do we compute the number of $X \in M_2(\mathbb{Z}/\ell^{n+k}\mathbb{Z})$ such that $1+\ell^n X$ has trace A and determinant Q mod ℓ^{k+2n}? These conditions are

$$\begin{gathered}
2 + \ell^n \,\mathrm{Trace}(X) \equiv A \bmod \ell^{k+2n},\\
1 + \ell^n \,\mathrm{Trace}(X) + \ell^{2n} \det(X) \equiv Q \bmod \ell^{k+2n}.
\end{gathered}$$

Because (A, Q) has no congruence obstruction, we know that

$$A \equiv Q + 1 \bmod \ell^{2n},$$
$$Q \equiv 1 \bmod \ell^n.$$

Thus the conditions on X are

$$\mathrm{Trace}(X) \equiv (A-2)/\ell^n \bmod \ell^{k+n},$$
$$\det(X) \equiv (Q+1-A)/\ell^{2n} \bmod \ell^k.$$

To count these, we first consider $X_k \bmod \ell^k$, satisfying the above two conditions mod ℓ^k. The number of such $X_k \bmod \ell^k$ is

$$\alpha_{\ell^k}(A_1, Q_1),$$

with

$$A_1 := (A-2)/\ell^n,\ Q_1 := (Q+1-A)/\ell^{2n}.$$

Once we have such an $X_k \bmod \ell^k$, we lift it arbitrarily to some $X_{k+n} \bmod \ell^{k+n}$; then we can correct this lift by adding to it anything of the form $\ell^k Y$, $Y \bmod \ell^n$, so long as Y has the required trace mod ℓ^n, namely

$$\mathrm{Trace}(Y) \equiv (A_1 - \mathrm{Trace}(X_{k+n}))/\ell^k \bmod \ell^n.$$

So there are ℓ^{3n} possible Y, and hence the number of elements in $G_{\ell^{k+2n},\det=Q}$ with trace A and determinant Q is

$$\ell^{3n}\alpha_{\ell^k}(A_1, Q_1).$$

Thus we get

$$\nu_{\ell^{k+2n}}(A, Q, \mathcal{M}^{\mathrm{ord}}) = \ell^{k+2n}\ell^{3n}\alpha_{\ell^k}(A_1, Q_1)/\ell^{3(k+n)}.$$

For this (A_1, Q_1), $\mathbb{Z}[T]/(T^2 - A_T + Q_1)$ is just $\mathbb{Z}[(F-1)/\ell^n]$, and so its δ_1 is just δ, the ord_ℓ of the conductor of $\mathbb{Z}[(F-1)/N_0]$. Also its K_1 is just K. So for $k \geq 2\delta + 2$, $\alpha_{\ell^k}(A_1, Q_1)$ is given by Gekeler's formulas

$$\ell^{2k} + \ell^{2k-1} \quad \text{if } \chi_K(\ell) = 1,$$
$$\ell^{2k} + \ell^{2k-1} - (\ell+1)\ell^{2k-\delta-2} \quad \text{if } \chi_K(\ell) = 0,$$
$$\ell^{2k} + \ell^{2k-1} - 2\ell^{2k-\delta-1} \quad \text{if } \chi_K(\ell) = -1.$$

With this data at hand, it is straightforward but unenlightening to verify, case by case depending on the value of $\chi_K(\ell)$, the required identity

$$\frac{\nu_\ell(A, Q, \mathcal{M}^{\mathrm{ord}})}{\nu_\ell(A, Q)} = \#SL(2, \mathbb{Z}/\ell^n\mathbb{Z})\frac{\sum_{a=0}^{\delta}\phi_K(\ell^a)}{\sum_{a=0}^{\delta+n}\phi_K(\ell^a)}.$$

□

Remark 8.7. Perhaps with a more conceptual approach, one could also verify the conjecture for the more general moduli problems we considered, where we allow also a $\Gamma_0(L)$-structure and a $\Gamma_1(M)$-structure. What is the relation of the conjectured formula to the formula, in terms of orbital integrals, given by Kottwitz in [Ko1, §16, pp. 432-433] and [Ko2, p. 205], when that general formula is specialized to the case of elliptic curves; cf. also [Cl, §3,§4]?

Question 8.8. As explained above, the conjecture is false in general, because of its incompatibility with pullback by a map which is surjective on fundamental groups. Nonetheless, one could ask if the following consequence of it is asymptotically correct. Take one of the moduli problems $\mathcal{M}^{\mathrm{ord}}/k$ we have discussed above, and take a finite flat $f : V \to \mathcal{M}^{\mathrm{ord}}$, with V/k smooth and geometrically connected, such that $f_\star$ is surjective on fundamental groups (e.g., an f which is fully ramified over some point). Fix a prime-to-p integer A, and an extension $\mathbb{F}_q/k$ with $q \geq 8$ and $A^2 < 4q$. We know precisely what the congruence obstructions for (A, q) are, and that, if there are none, then there are $\mathbb{F}_q$-valued points of $\mathcal{M}^{\mathrm{ord}}$ whose Frobenii have trace A; cf. Lemmas 4.2, 4.3. We further know that if (A, q) has no congruence obstruction, then there are infinitely many extensions $\mathbb{F}_Q/k$ for which (A, Q) has no congruence obstruction, and for each of these there are $\mathbb{F}_Q$-valued points of $\mathcal{M}^{\mathrm{ord}}$ whose Frobenii have trace A; cf. Lemma 6.2. For each such Q, consider the ratio

$$\frac{\#\{v \in V(\mathbb{F}_Q) \mid A_{v,\mathbb{F}_Q} = A\}}{\#\{u \in \mathcal{M}^{\mathrm{ord}}(\mathbb{F}_Q) \mid A_{u,\mathbb{F}_Q} = A\}}.$$

Is it true that this ratio tends to 1 as Q tends archimedeanly to infinity over the Q's for which the denominator is nonzero?

About the author

Nicholas M. Katz is a professor at Princeton University. He is the author or co-author of several books about arithmetic algebraic geometry, monodromy, diophantine questions over finite fields, and their interactions.

References

[Ba] Baier, Stephan, The Lang–Trotter conjecture on average. J. Ramanujan Math. Soc. 22 (2007), no. 4, 299-314. MR2376806 (2008j:11065)

[B-K] Bombieri, E. and Katz, N., A Note on Lower bounds for Frobenius traces, 2008 preprint available at www.math.princeton.edu/~nmk.

[Cl] Clozel, Laurent, Nombre de points des variétés de Shimura sur un corps fini (d'après R. Kottwitz). Séminaire Bourbaki, Vol. 1992/93. Astérisque No. 216 (1993), Exp. No. 766, 4, 121-149. MR1246396 (95c:11075)

[Co-Shp] Cojocaru, Alina Carmen and Shparlinski, Igor E., Distribution of Farey fractions in residue classes and Lang–Trotter conjectures on average, Proc. Amer. Math. Soc. 136, Number 6 (2008), 1977-1986. MR2383504 (2009a:11035)

[Cox] Cox, David, Primes of the Form $x^2 + ny^2$, John Wiley and Sons, New York, 1989. MR1028322 (90m:11016)

[Da-Pa] David, Chantal and Pappalardi, Francesco, Average Frobenius distributions of elliptic curves, Internat. Math. Res. Notices 1999 (1999), no. 4, 165-183. MR1677267 (2000g:11045)

[De-VA] Deligne, Pierre, Variétés abéliennes ordinaires sur un corps fini. (French) Invent. Math. 8 (1969) 238-243. MR0254059 (40:7270)

[De-WeilII] Deligne, Pierre, La conjecture de Weil. II, Inst. Hautes Études Sci. Publ. Math. No. 52 (1980), 137-252. MR601520 (83c:14017)

[Deu] Deuring, M., Die Typen der Multiplikatorenringe elliptischer Funktionenkörper, Abh. Math. Sem. Hansischen Univ. 14 (1941), 197-272. MR0005125 (3:104f)

[Deu-CM] Deuring, M., Die Zetafunktion einer algebraischen Kurve vom Geschlechte Eins, Nachr. Akad. Wiss. Göttingen. Math.-Phys. Kl. Math.-Phys.-Chem. Abt. (1953), 85–94. MR0061133 (15:779d)

[Elkies-Real] Elkies, Noam D. Supersingular primes for elliptic curves over real number fields. Compositio Math. 72 (1989), no. 2, 165-172. MR1030140 (90i:11058)

[Elkies-SS] Elkies, Noam D. The existence of infinitely many supersingular primes for every elliptic curve over Q. Invent. Math. 89 (1987), no. 3, 561-567. MR903384 (88i:11034)

[Ge] Gekeler, Ernst-Ulrich, Frobenius distributions of elliptic curves over finite prime fields, IMRN 2003, no. 37, 1999-2018. MR1995144 (2004d:11048)

[H-SB-T] Harris, Michael, Shepherd-Barron, Nicholas, and Taylor, Richard, A family of Calabi-Yau varieties and potential automorphy, preprint available at www.math.harvard.edu/~rtaylor/cy3.pdf.

[He] Hecke, Erich, Über eine neue Art von Zetafunktionen und ihre Beziehungen zur Verteilungen der Primzahlen. Math. Z. 1 (1918), 357-376;4 (1920), 11-21. Reprinted in Mathematishe Werke, Vandenhoeck and Ruprecht, Göttingen, 1983. MR1544302

[Honda] Honda, Taira, Isogeny classes of abelian varieties over finite fields, J. Math. Soc. Japan 20 (1968), 83-95. MR0229642 (37:5216)

[Howe] Howe, Everett W., Principally polarized ordinary abelian varieties over finite fields. Trans. Amer. Math. Soc. 347 (1995), no. 7, 2361-2401. MR1297531 (96i:11065)

[Ig] Igusa, Jun-ichi, Class number of a definite quaternion with prime discriminant. Proc. Nat. Acad. Sci. U.S.A. 44 (1958), 312-314. MR0098728 (20:5183)

[Ir-Ros] Ireland, Kenneth F. and Rosen, Michael I., A Classical Introduction to Modern Number Theory. Revised edition of Elements of Number Theory. Graduate Texts in Mathematics, 84. Springer-Verlag, New York-Berlin, 1982. MR661047 (83g:12001)

[Ito] Ito, Kyosi, An Introduction to Probability Theory, The Clarendon Press, Oxford University Press, New York, 1984. MR753828 (85f:60001)

[Ka-ESDE] Katz, Nicholas M., Exponential sums and differential equations, Annals of Mathematics Studies, 124. Princeton University Press, Princeton, NJ, 1990. MR1081536 (93a:14009)

[Ka-TLFM] Katz, Nicholas M., Twisted L-Functions and Monodromy, Annals of Mathematics Studies, 150. Princeton University Press, Princeton, NJ, 2002. MR1875130 (2003i:11087)

[K-M] Katz, Nicholas M. and Mazur, Barry, Arithmetic moduli of elliptic curves. Annals of Mathematics Studies, 108. Princeton University Press, Princeton, NJ, 1985. MR772569 (86i:11024)

[Ko1] Kottwitz, Robert E., Points on some Shimura varieties over finite fields. J. Amer. Math. Soc. 5 (1992), no. 2, 373-444. MR1124982 (93a:11053)

[Ko2] Kottwitz, Robert E., Shimura varieties and λ-adic representations. Automorphic forms, Shimura varieties, and L-functions, Vol. I (Ann Arbor, MI, 1988), 161-209, Perspect. Math., 10, Academic Press, Boston, MA, 1990. MR1044820 (92b:11038)

[L-T] Lang, Serge and Trotter, Hale, Frobenius distributions in GL_2-extensions, Springer Lecture Notes in Mathematics 504, 1976. MR0568299 (58:27900)

[Maz] Mazur, Barry, Rational points of abelian varieties with values in towers of number fields. Invent. Math. 18 (1972), 183-266. MR0444670 (56:3020)

[Mes] Messing, William, The crystals associated to Barsotti-Tate groups: with applications to abelian schemes. Lecture Notes in Mathematics, Vol. 264. Springer-Verlag, Berlin-New York, 1972. MR0347836 (50:337)

[Pa] Pacheco, Amícar, Distribution of the traces of Frobenius on elliptic curves over function fields. Acta Arith. 106 (2003), no. 3, 255–263. MR1957108 (2004a:11046)

[Sch] Schoof, René, Nonsingular plane cubic curves over finite fields. J. Combin. Theory Ser. A 46 (1987), no. 2, 183-211. MR914657 (88k:14013)

[Se-Cheb] Serre, Jean-Pierre, Quelques applications du théorème de densité de Chebotarev, Publ. Math. IHES 54 (1981), 323-401. MR0644559 (83k:12011)

[Se-Mot] Serre, Jean-Pierre, Propriétés conjecturales des groupes de Galois motiviques et des représentations ℓ-adiques, Proceedings of Symposia in Pure Mathematics 55 (1994) Part I, 377-400. MR1265537 (95m:11059)

[Sh] Shimura, Goro, Introduction to the arithmetic theory of automorphic functions. Kanô Memorial Lectures, No. 1. Publications of the Mathematical Society of Japan, No. 11. Iwanami Shoten, Publishers, Tokyo; Princeton University Press, Princeton, N.J., 1971. MR0314766 (47:3318)

[Sie] Siegel, Carl Ludwig, Über die Classenzahl quadratischer Zahlkörper, Acta Arith. 1 (1935), 83-86.

[St] Stein, William, The Modular Forms Explorer, database available at http://modular.fas.harvard.edu.

[Tate] Tate, J., Classes d'isogénie des variétés abéliennes sur un corps fini, Séminaire Bourbaki 1968-69, exp. 352, 95-110.

[Wat] Waterhouse, William C., Abelian varieties over finite fields. Ann. Sci. École Norm. Sup. (4) 2 (1969) 521-560. MR0265369 (42:279)

[Z] Zarhin, Yuri G., Very simple 2-adic representations and hyperelliptic Jacobians, Mosc. Math. J. 2 (2002), no. 2, 403–431. MR1944511 (2003k:11098)

[Yo] Yoshida, Hiroyuki, On an analogue of the Sato conjecture. Invent. Math. 19 (1973), 261-277. MR0337977 (49:2746)

Department of Mathematics, Fine Hall, Princeton University, Princeton, New Jersey 08544-1000

E-mail address: nmk@math.princeton.edu