STRANGE CONGRUENCES ON POINT COUNTS

NICHOLAS M. KATZ AND YUTA NAKAYAMA

Abstract. We discuss some strange congruences on counts of rational points on (certain) curves over finite fields, raise some questions, and give applications to monodromy.

Contents

1. Introduction

We did some computer experiments using Magma [BCP] which suggested some unexpected congruences. For simplicity, here is the simplest case. We took an odd prime p, a power q of p, and a finite extension $k/\mathbb{F}_q$. We then considered the one parameter family, parameter t, of hyperelliptic curves C_t given by the affine equation

$$y^2 = x^q + x^2 + t,$$

which for each $t \neq 0$ is (the complement of a single point at ∞ in) a projective, smooth, geometrically connected curve of genus $g = (q-1)/2$. For $t \in k^\times$, write

$$\#C_t(k) = \#k + 1 - a(k,t).$$

In terms of the quadratic character χ_2 of $k^\times$ (extended to k by decreeing $\chi_2(0) = 0$), we have the well known formula

$$a(k,t) = -\sum_{x \in k} \chi_2(x^q + x^2 + t).$$

In fact, we also computed this sum when $t = 0$. What we found empirically was the congruence

$$a(k,t) \equiv 1 \bmod p$$

for every $t \in k$. In fact, we found, again empirically, the further congruence

$$a(k,t) \equiv 1 \bmod q$$

for every $t \in k$. Along these same lines, we looked at the families

$$y^2 = x^{q^n} + x^2 + t.$$

Key words and phrases. curves over finite fields, point counts, character sums, monodromy.

For these families, we found, again empirically, the further congruence

$$a(k,t) \equiv 1 \bmod q^n$$

provided that k was an extension of $\mathbb{F}_{q^n}$.

Of these congruences, only the case $q = 3$ has a "classical" explanation: in characteristic 3, the Hasse invariant of an elliptic curve of equation $y^2 = x^3 + a_2x^2 + a_4x + a_6$ is the coefficient a_2 of x^2. In general, the congruence means that precisely one of the $2g = q - 1$ Frobenius eigenvalues $(\alpha_1 \ldots, \alpha_{2g})$ is nonzero mod p, say α_1, and that α_1 is 1 mod p. On the one hand, we infer that $\#C_t(k) \equiv 0 \bmod p$ for each $k/\mathbb{F}_q$ and each $t \in k^\times$, but we also infer that $\#\mathrm{Jac}_{C_t}(k) \equiv 0 \bmod p$ (because this cardinality is $\prod_i(1 - \alpha_i)$). But only in the $q = 3$ case is the set of rational points on the curve, when viewed in the Jacobian by $P \mapsto$ class of $[P] - [\infty]$, a subgroup.

We did further computer experiments of the following kind. We looked at "superelliptic" curves of equation

$$y^a = x^q + x^a + \text{lower terms, say } y^a = x^q + x^a + g(x), \deg(g) < a.$$

We took q to be 1 mod a, a finite extension $k/\mathbb{F}_q$ with $g(x) \in k[x]$, and computed the mod p sum

$$-\sum_{x \in k}(x^q + x^a + g(x))^{(\#k-1)/a},$$

an element of k, and the "exact" sum

$$-\sum_{x \in k}\mathrm{Teich}((x^q + x^a + g(x))^{(\#k-1)/a}),$$

where we denote by Teich the Teichmüller lift from k to the ring of Witt vectors $W(k)$. What we found empirically was that the "exact" sum was again 1 mod p. We also found that once q was 1 mod a, then if we looked at

$$y^a = x^{q^n} + x^a + g(x), \deg(g) < a$$

over any finite extension $k/\mathbb{F}_{q^n}$, then the "exact" sum was 1 mod p^n.

What we prove in this paper are the mod p congruences discussed above, see Theorem 4.1. The congruences mod higher powers of p remain open. Also open is the "meaning", if any, of these congruences. We also give applications to the determination of some monodromy groups.

2. The basic set up

Given a perfect field k of characteristic $p > 0$, and a proper, smooth, geometrically connected curve C/k, of genus $g \geq 1$, the (absolute) Cartier operator $\mathcal{C}$ is a p^{-1}-linear endomorphism of the g-dimensional k-vector space $H^0(C, \Omega^1_{C/k})$. Given a strictly positive power $q = p^f$ of p, we denote by $\mathcal{C}_q$ the f-fold iterate of $\mathcal{C}$. It is thus a q^{-1}-linear endomorphism of $H^0(C, \Omega^1_{C/k})$.

Let us recall, in a simple case, how to compute $\mathcal{C}_q$. For this, it is convenient to introduce the Dwork-inspired q^{-1}-linear operator Ψ_q on the polynomial ring $k[x]$, defined by

$$\Psi_q(\sum_n a_n x^n) := \sum_n (a_{nq})^{1/q} x^n.$$

Thus $\Psi_q(x^n)$ vanishes unless $q|n$, in which case it is $x^{n/q}$. Its relevance is the simple formula

$$\mathcal{C}_q(x^n dx/x) = \Psi_q(x^n)dx/x, \quad \mathcal{C}_q(x^n dx/(xf^q)) = \Psi_q(x^n)dx/(xf).$$

Consider a "superelliptic" curve C/k with affine equation of the form

$$y^a = f_d(x),$$

with $a \geq 2$ prime to p, and $f := f_d \in k[x]$ a polynomial of degree $d \geq 2$. We make the following two assumptions:

$$\gcd(f, f') = 1 \text{ and } \gcd(a, d) = 1.$$

The first assumption, that f has d distinct zeroes in $\overline{k}$, is the condition that this affine curve is smooth over k. The second, that $\gcd(a, d) = 1$, is that the complete nonsingular model of this affine curve has a single point at ∞. At ∞, the function x has a pole of order a, and the function y has a pole of order d. Thus dx has a pole of order $a+1$ at ∞, i.e., dx/x has a simple pole at ∞.

The space $H^0(C, \Omega^1_{C/k})$ has dimension $(a-1)(d-1)/2$, and a k-basis is given by the differentials

$$x^{i-1}dx/y^j = x^i dx/(xy^j), \text{ with } 1 \leq j \leq a-1, \ 1 \leq i \text{ such that } jd - ia - 1 \geq 0,$$

the inequalities being the condition of holomorphy at ∞.

Lemma 2.1. *Choose $q = p^f$ so that $q \equiv 1 \bmod a$. Then*

$$\mathcal{C}_q(x^i dx/(xy^j)) = \Psi_q(x^i f(x)^{j(q-1)/a}) dx/xy^j.$$

Proof. Indeed,

$$x^i dx/(xy^j) = x^i(dx/(xy^{jq}))y^{j(q-1)} = x^i(dx/(xy^{jq}))f(x)^{j(q-1)/a} = (x^i f(x)^{j(q-1)/a})dx/(xy^{jq}),$$

whose image under $\mathcal{C}_q$ is visibly $\Psi_q(x^i f(x)^{j(q-1)/a})dx/(xy^j)$. □

Remark 2.2. Suppose $q \equiv 1 \bmod a$. Then the group $\mu_a(\mathbb{F}_q)$ of ath roots of unity in $\mathbb{F}_q$ acts on the curve C, with $\zeta \in \mu_a$ mapping (x, y) to $(x, \zeta y)$, and $\mathcal{C}_q$ commutes with this action. The decomposition of $H^0(C, \Omega^1_{C/k})$ into eigenspaces for the action of μ_a is the decomposition by the power $1 \leq j \leq a-1$ of y in the denominator of $x^i dx/(xy^j)$, on which μ_a acts by the $-j$'th power of the "identical" character $\zeta \mapsto \zeta$. Each of these eigenspaces is stable by $\mathcal{C}_q$. For each j with $1 \leq j \leq a-1$, let us denote the corresponding eigenspace as

$$H^0(C, \Omega^1_{C/k})_j := \text{ the span of the } x^i dx/(xy^j), \ 1 \leq i \text{ such that } jd - ia - 1 \geq 0.$$

Notice that this eigenspace vanishes unless $jd \geq a+1$, i.e., unless $j \geq (a+1)/d$. As a/d is not an integer (because $\gcd(a, d) = 1$), it is equivalent to say that this eigenspace vanishes unless $j \geq a/d$.

Lemma 2.3. *Suppose $q \equiv 1 \bmod a$ and C is defined over $\mathbb{F}_q$, i.e. the defining equation $y^a = f_d(x)$ has $f_d \in \mathbb{F}_q[x]$. Then $\mathcal{C}_q$ is an $\mathbb{F}_q$-linear endomorphism of $H^0(C, \Omega^1_{C/\mathbb{F}_q})$, which stabilizes each $H^0(C, \Omega^1_{C/k})_j$ subspace, and for each $1 \leq j \leq a-1$ we have the identity*

$$\mathrm{Trace}(\mathcal{C}_q | H^0(C, \Omega^1_{C/k})_j) = -\sum_{x \in \mathbb{F}_q} (f_d(x))^{j(q-1)/a}, \text{ equality in } \mathbb{F}_q.$$

Proof. First we deal with the case when $j < a/d$. Then the j-eigenspace vanishes, and $(f_d(x))^{j(q-1)/a}$ has degree $\leq jd(q-1)/a < q-1$, in which case the asserted sum over x vanishes as well, cf. the next paragraph.

The key point is that for $g(x) := x^n$, $-\sum_{x \in \mathbb{F}_q} g(x)$ in $\mathbb{F}_q$ vanishes unless $n \geq 1$ and $(q-1)|n$, in which case the sum is 1. So $-\sum_{x \in \mathbb{F}_q} (f_d(x))^{j(q-1)/a}$ can be calculated as follows. Write

$$(f_d(x))^{j(q-1)/a} = \sum_n A_n x^n.$$

Then by the key point we have

$$-\sum_{x \in \mathbb{F}_q} (f_d(x))^{j(q-1)/a} = \sum_{n \geq 1} A_{n(q-1)}.$$

It suffices to treat the case when $j > a/d$, so that the space $H^0(C, \Omega^1_{C/k})_j$ is nonzero. Let us see how this sum is related to $\mathrm{Trace}(\mathcal{C}_q|H^0(C, \Omega^1_{C/k})_j)$. In the basis of $H^0(C, \Omega^1_{C/k})_j$ given by the $x^i dx/(xy^j)$, $1 \le i$ such that $jd - ia - 1 \ge 0$, the diagonal entries of the matrix of $\mathcal{C}_q$ are as follows. For each integer i with $i \ge 1, ia+1 \le jd$, the (i,i) entry is the coefficient of x^i in $\Psi_q((x^i f(x)^{j(q-1)/a})$, or equivalently the coefficient of x^{iq} in $x^i f(x)^{j(q-1)/a}$, or equivalently the coefficient of $x^{iq-i} = x^{i(q-1)}$ in $f(x)^{j(q-1)/a}$, which is the coefficient $A_{i(q-1)}$. So it remains only to see that the indices n with

$$0 < n(q-1) \le dj(q-1)/a$$

are precisely those with $n > 0$ and $jd - na - 1 \ge 0$, or equivalently with $jd - 1 \ge na > 0$. The first inequality is $0 < n \le dj/a$. But dj/a cannot be an integer: indeed, if $a|dj$, then because $\gcd(d,a) = 1$ we would have $a|j$, which is impossible because $1 \le j \le a-1$. Thus the first inequality is $0 < n < dj/a$, i.e., $0 < na < dj$, which we rewrite as $0 < na \le dj - 1$. Thus the (i,i) diagonal entry of the matrix of $\mathcal{C}_q|H^0(C, \Omega^1_{C/k})_j$ is precisely the coefficient $A_{i(q-1)}$, and the allowed $i > 0$ run over the possible $n > 0$ for which $A_{n(q-1)}$ occurs in $(f_d(x))^{j(q-1)/a}$. □

3. The congruence formula for the L-function mod p

In fact, the trace formula of Lemma 2.3 is a consequence of an identity of characteristic polynomials. For a superelliptic curve C of equation $y^a = f_d(x)$ over $\mathbb{F}_q$ with $q \equiv 1 \bmod a$, a fixed character $\chi_j : \zeta \mapsto \zeta^j$ of $\mu_a(\mathbb{F}_q)$, denote by $H^1(C, \mathcal{O}_C)(\chi_j)$ (or more clumsily $H^1(C, \mathcal{O}_C)_{-j}$) the corresponding eigenspace in $H^1(C, \mathcal{O}_C)$. One knows [Ka-Int, 3.1] that the action of Frob_q on $H^1(C, \mathcal{O}_C)$ is the linear dual of the action of $\mathcal{C}_q$ on $H^0(C, \Omega^1_{C/\mathbb{F}_q})$. Passing to χ-components, the action of Frob_q on $H^1(C, \mathcal{O}_C)(\chi_j)$ is the linear dual of the action of $\mathcal{C}_q$ on the χ_{-j} eigenspace $H^0(C, \Omega^1_{C/\mathbb{F}_q})_j$ of $H^0(C, \Omega^1_{C/\mathbb{F}_q})$. Consider now the L-function of C, with the Teichmüller lifting $\mathrm{Teich}(\chi_j) : \mu_a(\mathbb{F}_q) \to \mu_a(W(\mathbb{F}_q))$. One knows that in terms of crystalline cohomology $H^1_{\mathrm{cris}}(C/W)$, one has

$$L(C/\mathbb{F}_q, \mathrm{Teich}(\chi_j)) = \det(1 - T\mathsf{Frob}_q|H^1_{\mathrm{cris}}(C/W)(\mathrm{Teich}(\chi_j))),$$

an identity in $W(\mathbb{F}_q)[T]$. Reducing mod p, we have an identity

$$L(C/\mathbb{F}_q, \chi_j) = \det(1 - T\mathsf{Frob}_q|H^1_{DR}(C/\mathbb{F}_q)(\chi_j)), \text{ equality in } \mathbb{F}_q[T].$$

Consider the χ_j component of the Hodge filtration short exact sequence

$$0 \to H^0(C, \Omega^1_{C/\mathbb{F}_q})_{-j} \to H^1_{DR}(C/\mathbb{F}_q)(\chi_j) \to H^1(C, \mathcal{O}_C)(\chi_j) \to 0.$$

The map Frob_q kills the first term $H^0(C, \Omega^1_{C/\mathbb{F}_q})_{-j}$, so

$$\det(1 - T\mathsf{Frob}_q|H^1_{DR}(C/\mathbb{F}_q)(\chi_j)) = \det(1 - T\mathsf{Frob}_q|H^1(C, \mathcal{O}_C)(\chi_j) = \det(1 - T\mathcal{C}_q|H^0(C, \Omega^1_{C/\mathbb{F}_q})_j),$$

the final equality by duality.

Thus we have the congruence formula:

Theorem 3.1. *Suppose $q \equiv 1 \bmod a$ and C is defined over $\mathbb{F}_q$, i.e. the defining equation $y^a = f_d(x)$ has $f_d \in \mathbb{F}_q[x]$. Then $\mathcal{C}_q$ is an $\mathbb{F}_q$-linear endomorphism of $H^0(C, \Omega^1_{C/\mathbb{F}_q})$, which stabilizes each $H^0(C, \Omega^1_{C/k})_j$ subspace, and for each $1 \le j \le a-1$ we have the identity*

$$L(C/\mathbb{F}_q, \mathrm{Teich}(\chi_j)) \bmod p = \det(1 - T\mathcal{C}_q|H^0(C, \Omega^1_{C/\mathbb{F}_q})_j).$$

Remark 3.2. The coefficient of $-T$ in $L(C/\mathbb{F}_q, \mathrm{Teich}(\chi_j))$ is the sum

$$-\sum_{x \in \mathbb{F}_q} \mathrm{Teich}(f(x)^{j(q-1)/a}),$$

whose reduction mod p is precisely the sum

$$-\sum_{x\in\mathbb{F}_q} f(x)^{j(q-1)/a}$$

which was obtained in Lemma 2.3 as $\mathrm{Trace}(\mathcal{C}_q|H^0(C,\Omega^1_{C/\mathbb{F}_q})_j)$.

4. Some special curves

In this section, we fix an integer $a \geq 2$ which is prime to p, a strictly positive power q of p which has $q \equiv 1 \bmod a$, a finite extension $k/\mathbb{F}_q$, and a superelliptic curve C defined over k with affine equation of the special form

$$y^a = f_q,\ f_q := x^q + x^a + g(x),\ \deg(g) < a.$$

Theorem 4.1. *For $j = 1$, we have*

$$\mathrm{Trace}(\mathcal{C}_{\#k}|H^0(C,\Omega^1_{C/k})_1) = 1, \text{ equality in } k,$$

or, equivalently,

$$-\sum_{x\in k}(f_q(x))^{(\#k-1)/a} = 1.$$

Furthermore, we have the mod p identity

$$L(C/k, \mathrm{Teich}(\chi_1)) \bmod p = 1 - T.$$

Proof. Let us write $\#k = q^f$, so that $\mathcal{C}_{\#k} = \mathcal{C}_{q^f}$ is the f-fold interate of $\mathcal{C}_q$. [Remember that $\mathcal{C}_q$ is q^{-1}-linear, but its f-fold iterate is k-linear.] We will use the basis of $H^0(C,\Omega^1_{C/k})_1$ given by the $x^i dx/(xy)$, $1 \leq i$ such that $q - ia - 1 \geq 0$, which is to say $1 \leq i \leq (q-1)/a$.

We next define an increasing filtration

$$W_1 \subset W_2 \ldots \subset W_{(q-1)/a} = H^0(C,\Omega^1_{C/k})_1$$

as follows: W_r is the subspace spanned by the basis elements $x^i dx/(xy)$ with $i \leq r$.

We will establish the following three statements.

1. The "matrix" of $\mathcal{C}_q$ in this basis is upper triangular, meaning precisely that each subspace W_r is $\mathcal{C}_q$-stable.
2. For $r > 1$, $\mathcal{C}_q$ maps W_r to W_{r-1}.
3. The only nonzero diagonal entry in this matrix, namely the $(1,1)$ entry, is 1.

Once these points are established, any iterate of $\mathcal{C}_q$ on $H^0(C,\Omega^1_{C/k})_1$ in this same basis also has properties 1), 2), 3). Applying this to the f-fold iterate, we get the assertion, as a consequence of Lemma 2.3.

Fix an index i with $1 \leq i \leq (q-1)/a$. We ask which basis elements $x^n dx/(xy)$ can occur with nonzero coefficient in $\mathcal{C}_q(x^i dx/(xy))$. This is equivalent to asking which powers x^n can occur in $\Psi_q(x^i(f_q)^{(q-1)/a})$, or equivalently which powers x^{nq} occur in $x^i(f_q)^{(q-1)/a}$, or, finally, which powers x^{qn-i} occur in $(f_q)^{(q-1)/a}$.

The monomials that can possibly appear in $(f_q)^{(q-1)/a}$ can be described as follows. Write out

$$f_q = x^q + \sum_{m\leq a} A_m x^m,$$

where we temporarily forget that $A_a = 1$. If a monomial x^{qn-i} occurs in $(f_q)^{(q-1)/a}$, then for some expression

$$(q-1)/a = \alpha + \sum_{m \le a} \beta_m$$

with nonnegative integers $(\alpha, \beta_0, \ldots, \beta_a)$ we must have

$$qn - i = q\alpha + \sum_{1 \le m \le a} m\beta_m,$$

as the degree zero term β_0 does not contribute to the degree. The second term $\sum_{1 \le m \le a} m\beta_m$ is bounded by $a \sum_{1 \le m \le a} \beta_m \le a((q-1)/a) = q-1$. Thus

$$qn - i = q\alpha + (\text{ a nonnegative term } \le q-1).$$

We next show that $\alpha = n-1$. To see this, first write this last equality in the crude form

$$q(\alpha - n) = -i - \text{nonnegative},$$

to infer that $\alpha < n$. Then write

$$q(n-\alpha) = i + (\text{ a nonnegative term } \le q-1).$$

Here $i \le (q-1)/a$, so we have the inequality

$$q(n-\alpha) \le (q-1)/a + (q-1),$$

so trivially $q(n-\alpha) < 2q$. Thus $n-\alpha$ is a strictly positive integer which is < 2, hence is 1, and $n - \alpha = 1$.

Using that $\alpha = n-1$, we then have

$$qn - i = q(n-1) + \sum_{1 \le m \le a} m\beta_m,$$

which we rewrite as

$$q-i = \sum_{1 \le m \le a} m\beta_m \le a(\sum_{1 \le m \le a} \beta_m) \le a(\sum_{0 \le m \le a} \beta_m) \le a((q-1)/a - \alpha) = (q-1) - a\alpha = (q-1) - a(n-1),$$

giving

$$-i \le -1 - a(n-1),$$

i.e.,

$$a(n-1) \le i - 1.$$

Recalling that $a \ge 2$, we see that if $i = 1$, then $n = 1$, while if $i \ge 2$, then $n - 1 \le (i-1)/a < i-1$, in which case $n < i$.

It remains to show that the monomial x^{q-1} occurs with coefficient 1 in $(x^q + x^a + \sum_{m<a} A_m x^m)^{(q-1)/a}$. For any expression

$$(q-1)/a = \alpha + \sum_{m \le a} \beta_m$$

with nonnegative integers $(\alpha, \beta_0, \ldots, \beta_m)$ we must have

$$q - 1 = q\alpha + \sum_{0 \le m \le a} m\beta_m.$$

Thus $\alpha = 0$ and

$$(q-1)/a = \sum_{m \le a} \beta_m.$$

But

$$q-1=\sum_{0\le m\le a} m\beta_m \le a(\sum_{0\le m\le a}\beta_m)=(q-1),$$

hence we have the equality

$$\sum_{0\le m\le a} m\beta_m = a(\sum_{0\le m\le a}\beta_m)=q-1,$$

and thus

$$\sum_{m\le a}(m-a)\beta_m=0.$$

But each $m-a\le 0$, and each $\beta_m\ge 0$. So each summand $(m-a)\beta_m\le 0$, hence each summand $(m-a)\beta_m=0$. For $m<a$, this forces $\beta_m=0$. Thus $\beta_a=(q-1)/a$, all other β_m vanish, as does α. So our x^{q-1} occurs entirely as the $(q-1)/a$ power of x^a, with coefficient 1.

From the upper triangular shape of the "matrix" of $\mathcal{C}_q$, and hence of $\mathcal{C}_{\#k}$ as well, with zeros on the diagonal except for an entry 1 in the $(1,1)$ position, the congruence for the L function results from Theorem 3.1.

□

5. Nonsingularity often does not matter

Let k be a field of characteristic p, and C_0/k the affine curve of equation

$$y^a=f_d(x),$$

with $a\ge 2$ prime to p, and $f:=f_d\in k[x]$ a polynomial of degree d. We assume that

$$\gcd(a,d)=1.$$

This condition ensures that C_0 is geometrically irreducible, whatever the polynomial f_d. As noted in our earlier discussion just before Lemma 2.1, if f_d has d distinct zeroes in $\overline{k}$, then C_0 is the complement of a single point at ∞ in a projective, smooth, geometrically connected curve C of genus $(a-1)(d-1)/2$. But, for example, in the extreme case when $f_d=x^d$, C_0 is a rational curve, with $C_0\smallsetminus(0,0)$ being the group $\mathbb{G}_m$, by the map $t\mapsto(x=t^a,y=t^d)$.

Lemma 5.1. *For C_0 as above, and any $\ell\ne p$, the compact cohomology groups*

$$H^i_c(C_0):=H^i_c(C_0\otimes_k\overline{k},\overline{\mathbb{Q}_\ell})$$

are given by $H^0_c=0, H^2_c=\overline{\mathbb{Q}_\ell}(-1)$, and $H^i_c=0$ for $i>2$.

Proof. The only nonobvious point is that $H^0_c(C_0)=0$. To see this, denote by C the projective closure in $\mathbb{P}^2$ of C_0, i.e., the curve of equation $Z^{d-a}Y^a=F_d(X,Z)$ if $d>a$, or the curve of equation $Y^a=Z^{a-d}F_d(X,Z)$ if $a>d$. In either case, there is precisely one point at ∞, i.e., one point with $Z=0$, simply because $a\ne d$. Then C is geometrically irreducible, so its $H^0_c(C)$ is $\overline{\mathbb{Q}_\ell}$. The excision sequence for the inclusion $C_0\subset C$ begins with

$$0\to H^0_c(C_0)\to H^0_c(C)\to H^0_c(\text{the single point }\infty),$$

in which the final arrow is an isomorphism. □

Suppose now that the field k is a finite field containing the ath roots of unity. Then we compute $\#C_0(k)$ as the sum

$$\sum_{x\in k}(\text{the number of }a\text{th roots of }f_d(x)\text{ in }k).$$

For any element $z \in k$, the number of its ath roots in k is the sum

$$1 + \sum_{\text{nontrivial characters } \chi \text{ of } k^\times \text{ with } \chi^a = \mathbb{1}} \chi(z),$$

with the usual convention that for $\chi \neq \mathbb{1}$, $\chi(0) = 0$. Thus if we denote by $\mathsf{Char}_{\text{nontriv}}(a)$ the set of χ being summed over, we have

$$\#C_0(k) = \#k + \sum_{\chi \in \mathsf{Char}_{\text{nontriv}}(a)} \sum_{x \in k} \chi(f_d(x)).$$

When we view the same χ as a character of the group μ_a, by precomposing with $z \mapsto z^{(\#k-1)/a}$, then under the μ_a action on C_0 given by $\zeta : (x, y) \mapsto (x, \zeta y)$, we can break $H^1_c(C_0)$ into eigenspaces under $\mathsf{Char}(a)$, the group of all characters of order dividing a. For each $\chi \in \mathsf{Char}_{\text{nontriv}}(a)$, the Lefschetz trace formula then gives

$$\text{Trace}(\mathsf{Frob}_k | H^1_c(C_0)^\chi) = -\sum_{x \in k} \chi(f_d(x)),$$

while $H^1_c(C_0)^{\mathbb{1}} = 0$ and $H^2_c(C_0)$ has trivial μ_a action, and Frobenius trace $\#k$.

The summands $\chi(f_d(x))$ lie in the cyclotomic ring $\mathbb{Z}[\zeta_a]$, which we may embed in the Witt vector ring $W(k)$. Viewed in the Witt vector ring, it makes sense to ask about p-adic congruences for the sums $-\sum_{x \in k} \chi(f_d(x))$.

Lemma 5.2. *Suppose that $k := \mathbb{F}_q$ contains the ath roots of unity. For each j with $1 \le j \le a-1$, denote by $\mathcal{V}_j$ the k-span of the monomials x^i, with exponents $i \ge 1$, for which $ia \le jd - 1$. Denote by*

$$\Psi_q \circ f_d(x)^{j(q-1)/d}$$

the $\mathbb{F}_q$-linear endomorphism of $\mathbb{F}_q[x]$ given by

$$g \mapsto \Psi_q(g(x) f_d(x)^{j(q-1)/a}).$$

Then $\Psi_q \circ f_d(x)^{j(q-1)/d}$ maps $\mathcal{V}_j$ to itself, and we have the trace formula

$$\text{Trace}(\Psi_q \circ f_d(x)^{j(q-1)/d} | \mathcal{V}_j) = -\sum_{x \in \mathbb{F}_q} (f_d(x))^{j(q-1)/a}, \text{ equality in } \mathbb{F}_q.$$

Proof. Repeat verbatim the Ψ_q part of the proof of Lemma 2.3. □

Lemma 5.3. *Suppose that $\mathbb{F}_q$ contains the ath roots of unity, and $k = \mathbb{F}_{q^n}$. Then the $\mathbb{F}_{q^n}$-linear endomorphism*

$$\Psi_{q^n} \circ f_d(x)^{j(q^n-1)/a} | \mathcal{V}_j$$

is the n-fold iterate of the q^{-1}-linear endomorphism

$$\Psi_q \circ f_d(x)^{j(q-1)/a} | \mathcal{V}_j.$$

Proof. Indeed, as additive endomorphisms of $k[x]$, we have the composition rule

$$\Psi_{q^a} \circ h(x) \circ \Psi_q \circ k(x) = \Psi_{q^{a+1}} \circ (h(x)^q k(x)).$$

The assertion then results by inductively applying this composition rule, since

$$f_d(x)^{j(q^n-1)/d} = \prod_{i=0}^{n-1} (f_d(x)^{j(q-1)/d})^{q^i}.$$

□

Consider now the following special case of the situation above. We have the integer $a \geq 2$ which is prime to p, a strictly positive power q of p which has $q \equiv 1 \mod a$, a finite extension $\mathbb{F}_{q^n}/\mathbb{F}_q$, and an affine curve $C_0/\mathbb{F}_{q^n}$ defined by an affine equation of the special form

$$y^a = f_q,\ f_q := x^q + x^a + g(x),\ \deg(g) < a.$$

Theorem 5.4. *For $j = 1$, we have*

$$\mathrm{Trace}(\Psi_{q^n} \circ f_q(x)^{(q-1)/a} | \mathcal{V}_1) = 1,\ \text{equality in } k,$$

or, equivalently,

$$-\sum_{x \in \mathbb{F}_{q^n}} (f_q(x))^{(q^n-1)/a} = 1.$$

Proof. Repeat verbatim the $\Psi_{\#k}$ part of the proof of Theorem 4.1. □

Remark 5.5. The main casualty of not requiring nonsingularity is that we lose the congruence formula given by Theorem 3.1. To see what the problem is, consider what is arguably the worst case, when we take a to be $q-1$, and the curve

$$y^{q-1} = x^q + x^{q-1},$$

which, after the change of variable $(x, y) \mapsto (x, y/x)$ is just the rational curve $y^{q-1} = x + 1$, whose $H^1 = 0$, while $\oplus_{1 \leq j \leq q-2} W^j$ has dimension $(q-1)(q-2)/2$.

6. Some open questions, in the nonsingular case

In the context of Theorem 4.1, for each nontrivial character χ_j, the χ_j component of H^1_{cris} has dimension $q-1$, so each L function is a polynomial of that degree, and one can ask what is its Newton polygon. In the case of χ_1, we have seen that the Newton polygon has a single slope 0, all other slopes are strictly positive. [We might remark in passing that knowing this requires the full congruence formula for the L function, otherwise we might have had, say $p+1$ unit eigenvalues, each of which was 1 mod p, instead of a single one. Of course this possible ambiguity can only arise when $q - 1 \geq p$, i.e., when $q > p$.]

By general semicontinuity, when we look at nonsingular curves of the form $y^a = x^q + x^a + g(x)$ with $\deg(g) < a$, on an open dense set of g's the Newton polygon of the L function for each χ_j is constant. What is it, especially for χ_1, where we know the Newton polygon has a single slope 0?

Another question: in this χ_1 case, where there is a single "unit root", give a limit formula for it, along the lines of [Ka-Int, §8].

7. Applications to monodromy

Let k be a field. We say that a polynomial $f(x) \in k[x]$ is "very weakly supermorse", compare [Ka-ACT, 5.5.2], if it satisfies the following three conditions.

1. The second derivative $f''(x)$ is not identically zero.
2. The derivative $f'(x)$, of degree denoted δ, has δ distinct zeroes (in $\overline{k}$), say $\alpha_1, \ldots, \alpha_\delta$.
3. If $\delta > 1$, f separates the zeroes of f': if $1 \leq i < j \leq \delta$, then $f(\alpha_i) \neq f(\alpha_j)$.

Equivalently, f is very weakly supermorse if, for all but finitely many $\lambda \in \overline{k}$, the polynomial $f(x) - \lambda$ has $d := \deg(f)$ distinct zeroes, while for a finite nonempty set $\Lambda \subset \overline{k}$, and for $\lambda \in \Lambda$, $f(x) - \lambda$ has $d-1$ distinct zeroes, i.e., $d-2$ simple roots and one double root. The notion of "weakly supermorse" defined in [Ka-ACT, 5.5.2] required in addition that $d := \deg(f)$ be prime to p, so that δ is $d-1$ there.

The proof of [Ka-ACT, Lemma 5.15] may be repeated verbatim to prove the following lemma.

Lemma 7.1. *Let k be a field, and $f(x) \in k[x]$ a polynomial whose second derivative f'' is not identically zero. Then there exists a nonzero polynomial $R(t) \in k[t]$ such that for any extension field K/k, and and $a \in K$ with $R(a) \neq 0$, the polynomial $f(x) + ax$ is very weakly supermorse.*

Suppose now that $f(x)$ is very weakly supermorse. Denote by $S := \mathrm{CritVal}(f) \subset \mathbb{A}^1$ the finite set of critical values of f, i.e., the values of f on the δ zeroes of f'. Then on $\mathbb{A}^1 \setminus S$ we have the local system

$$\mathcal{F} := f_\star\overline{\mathbb{Q}_\ell}/\overline{\mathbb{Q}_\ell},$$

for any ℓ invertible in k. At each critical value $s \in S$, the local monodromy of $\mathcal{F}$ at s is a transposition in G_{geom} := the geometric monodromy group of the local system $\mathcal{F}$, viewed as a transitive subgroup of the symmetric group S_d, for $d := \deg(f)$. If d were invertible in k, then $\mathcal{F}$ would be tame at ∞, in which case G_{geom} would be generated by all conjugates of all local monodromies at points in S, so would be a transitive subgroup of S_d generated by transpositions, and hence G_{geom} would be S_d in its deleted permutation representation. As this representation of S_d is irreducible, we recover the fact, cf. [Ka-ACT, 5.5], that $\mathcal{F}$ is geometrically irreducible.

However, in the case of interest, f has degree $d = q := p^f$, and we are in characteristic p. In this case $\mathcal{F}$ is no longer tame at ∞. [In fact it is totally wild.] How, then, can we exploit having transpositions as the monodromies at finite distance (i.e., in S), or find some other way, to prove that $\mathcal{F}$ is geometrically irreducible?

There are two standard methods. The first is to show that the two variable polynomial

$$\Delta_f(x, y) := \frac{f(x) - f(y)}{x - y}$$

is geometrically irreducible (i.e., irreducible in $\overline{k}[x, y]$).

The second is to show that G_{geom} is a primitive subgroup of S_d, since by Jordan's theorem [Wielandt, Theorem 13.9], a primitive subgroup of S_d which contains an r-cycle for some prime $r \leq d - 3$ is either A_d or S_d. In particular, a primitive subgroup of S_d with $d \geq 5$ which contains a transposition must be S_d.

To show that G_{geom} is primitive, it is equivalent to show that the polynomial f is indecomposable, meaning that in $\overline{k}[x]$, it cannot be written as a composition $g(h(x))$ of polynomials g, h, both of degree ≥ 2. Notice that if f is decomposable, say $f = g \circ h$, then we have the "chain rule"

$$\Delta_{g\circ h}(x, y) = \Delta_g(h(x), h(y))\Delta_h(x, y).$$

Thus if $\Delta_f(x, y)$ is geometrically irreducible, then f is indecomposable.

As an application, notice that when $f(x) = x^q + x^2$, and p is odd, then $\frac{f(x)-f(y)}{x-y}$ is geometrically irreducible, being $(x - y)^{q-1} + 2(x + y)$, which in coordinates $u := x - y, v := x + y$ is $u^{q-1} + 2v$ in $k[u, v]$.

Remark 7.2. When f has degree p, it is trivially indecomposable. In this case, we "recover" the fact that a transitive subgroup of S_p, $p \geq 5$, which contains a transposition must be S_p. [Of course this is also true for $p = 3$: the subgroup generated by a transposition is not transitive, and any larger subgroup of S_3 is all of S_3, because $\#\mathsf{S}_3 = 2 \times 3$.]

We also have the following case.

Lemma 7.3. *Over $\overline{\mathbb{F}_p}$, suppose $f(x) = x^{p^2} + x^a + g(x)$ with $a \geq 2$, $p^2 > a > \deg(g)$ and $\gcd(a, p) = 1$, is very weakly supermorse. Then f is indecomposable.*

Proof. The proof depends on the following lemma.

Lemma 7.4. *Suppose f is very weakly supermorse and is decomposable, $f(x) = g(h(x))$. Then $x \mapsto g(x)$ is finite étale.*

Proof. We argue by contradiction. If g is not finite étale, then for some scalar α, $g(x) - \alpha$ has a multiple root, so when we factor it, we get

$$g(x) - \alpha = C_\alpha \prod_j (x - \beta_j)^{n_j}$$

with $C_\alpha \neq 0$ and some $n_j \geq 2$. Then

$$f(x) - \alpha = g(h(x)) - \alpha = C_\alpha \prod_j (h(x) - \beta_j)^{n_j}.$$

So $f(x) - \alpha$ is divisible by some $(h(x) - \beta_j)^2$, so has either more than one double root (if $(h(x) - \beta_j$ has at least two distinct roots), or it has root of multiplicity at least 4 (if $(h(x) - \beta_j$ has only one root, necessarily of multiplicity $\deg(h) \geq 2$). □

This lemma shows that in characteristic zero, no very weakly supermorse polynomial is decomposable, and that in any characteristic, no weakly supermorse polynomial is decomposable.

With this lemma at hand, we argue as follows. As the degree is p^2, if f is decomposable as $g(h(x))$, then both g, h have degree p, and g is finite étale, so necessarily of the form

$$g(x) = c_p x^p + a_1 x + c_0, \quad c_p c_1 \neq 0,$$

and

$$h(x) = \sum_{n=0}^{p} b_n x^n, \quad b_p \neq 0.$$

Then

$$f(x) = g(h(x)) = c_p(\sum_{n=0}^{p} b_n^p x^{pn}) + c_1(\sum_{n=0}^{p} b_n x^n) + c_0.$$

Suppose first that x^a occurs in f with $p < a < p^2$ and $p \nmid a$. All monomials in f of degree $> p$ come from $h(x)^p$, none of whose nonzero exponents is prime to p, contradiction.

Suppose next that $2 \leq a < p$. Then the term x^a must be present in h, and hence h has the form

$$h(x) = b_p x^p + b_a x^a + \text{lower terms}.$$

Then $c_p h(x)^p$ has the monomial x^{pa} occurring, and (because $a \geq 2$) this monomial does not occur in h. Thus $f(x) = g(h(x))$ has the monomial x^{pa} occurring, contrary to the assumed shape of f. □

Remark 7.5. Here is a case where we get the geometric irreducibility of $\mathcal{F}$ by showing that for $I(\infty)$, the inertia group at ∞, $\mathcal{F}$ is $I(\infty)$-irreducible. Namely, for $f(x) = x^q + x^{q-1} + (\text{ lower terms})$, its $\mathcal{F}$ has $\mathsf{Swan}_\infty(\mathcal{F}) = 1$, and all ∞-slopes $1/(q-1)$, by [Ka-MMP, Theorem A6.1.5], hence is $I(\infty)$-irreducible. For an $f(x) = x^q + x^a + (\text{ lower terms})$ with $a < q$, $p \nmid a$, and p odd, which is very weakly supermorse, its $\mathcal{F}$ has $\mathsf{Swan}_\infty(\mathcal{F}) = q - a$ (just use the Euler-Poincaré formula upstairs and down) and is totally wild at ∞, but we do not know its slopes, or whether it is $I(\infty)$-irreducible.

Theorem 7.6. *(*compare [Ka-ACT, 5.4]*) Let $a \geq 2$ be a prime to p integer, $q = p^f$ a power of an odd prime p with $q \equiv 1 \bmod a$, $k/\mathbb{F}_q$ a finite extension, and $f \in k[x]$ a polynomial of the form*

$$f(x) = x^q + x^a + \text{lower terms},$$

which is very weakly supermorse and for which

$$\mathcal{F} := f_\star \overline{\mathbb{Q}_\ell} / \overline{\mathbb{Q}_\ell}$$

is geometrically irreducible (see Remark 7.2 and Lemma 7.3 for examples). For ρ a multiplicative character of (exact) order a, form the middle additive convolution

$$\mathcal{G} := \mathcal{G}_\rho := \mathcal{F} \star_{+,mid} \mathcal{L}_\rho.$$

Then $\mathcal{G}$ is lisse and geometrically irreducible on

$$\mathbb{A}^1 \smallsetminus S,\ S := \text{the critical values of } f.$$

Its rank is $q-1$, and at each critical value $s \in S$ of f, its local monodromy is a pseudoreflection of determinant $\chi_2\rho$. Its trace function is given as follows: For L/k a finite extension, and $t \in L$ not a critical value of f,

$$\mathrm{Trace}(\mathsf{Frob}_{t,L}|\mathcal{G}) = -\sum_{x \in L} \rho_L(t - f(x)).$$

If ρ has order 2, the geometric monodromy group G_{geom} of $\mathcal{G}$ is Sp_{q-1}. If ρ has order ≥ 3, then $G^\circ_{\mathrm{geom}} = \mathrm{SL}_{q-1}$, and for N the order of $\chi_2\rho$, $G_{\mathrm{geom}} = \{A \in \mathrm{GL}_{q-1} | \det(A)^N = 1\}$.

Proof. Our $\mathcal{F}$ has local monodromies at finite distance which are reflections. Then by [Ka-RLS, the "tame on $\mathbb{A}^1$" part of the proof of Corollary 3.3.6, 1)], $FT(\mathcal{F})(\infty)$ is the direct sum

(slopes > 1)$\oplus$ (the direct sum of the $\mathcal{L}_{\chi_2(t-s)} \otimes \mathcal{L}_{\psi(st)}$ over the finite singularities $s \in S$)
and

$$\mathcal{G} := \mathcal{F} \star_{+,mid} \mathcal{L}_\rho$$

is lisse on $\mathbb{A}^1 \smallsetminus S$ and (as $\mathcal{F}$ is geometrically irreducible) geometrically irreducible, and each local monodromy at finite distance is a pseudoreflection of determinant $\chi_2\rho$. If we can prove that $\mathcal{G}$ is is not induced, then by the trichotomy of [Ka-MG, Proposition 1], $\mathcal{G}$ is either Lie-irreducible or a tensor product

$$\mathcal{G} = \mathcal{H} \otimes \mathcal{K}$$

with $\mathcal{H}$ Lie-irreducible and $\mathcal{K}$ irreducible of rank ≥ 2 with finite $G_{\mathrm{geom},\mathcal{K}}$. We cannot have $\mathcal{H}$ of rank one, otherwise $\det(\mathcal{G})$ is $\mathcal{H}^{\otimes(q-1)} \otimes \det(\mathcal{K})$; but $\det(\mathcal{G})$ is geometrically of finite order (because $\mathcal{G}$ starts life over a finite field), and hence $\mathcal{H}$ would be geometrically of finite order, the very opposite of being Lie-irreducible. But at any of the critical values of f, the local monodromy of $\mathcal{G}$ is a pseudoreflection, and no pseudoreflection is nontrivially a tensor product. Thus, if $\mathcal{G}$ is not induced, it is Lie-irreducible, and $\mathrm{Lie}(G_{\mathrm{geom},\mathcal{G}})$ is an irreducible, semisimple lie subalgebra of $M_{q-1}(\overline{\mathbb{Q}_\ell})$ which is normalized by a pseudoreflection of determinant $\chi_2\rho$. The key result now is due to Kazhdan-Margulis, Gabber, and Beukers-Heckman, see [Ka-ESDE, Theorem 1.5]. If ρ has order 2, then $G^0_{\mathrm{geom},\mathcal{G}}$ is either Sp_{q-1} or SL_{q-1}. In this order 2 case, we are dealing with the trace function of H^1 of the family of hyperelliptic curves $y^2 = t - f(x)$, so we have an a priori inclusion of $G_{\mathrm{geom},\mathcal{G}}$ in Sp_{q-1}. In the case of ρ of order ≥ 3, $\chi_2\rho$ also has order ≥ 3, and then $G^0_{\mathrm{geom},\mathcal{G}}$ must be SL_{q-1}.

To determine $G_{\mathrm{geom},\mathcal{G}}$ exactly for ρ of order ≥ 3, we must show that $\det(\mathcal{G})$ is geometrically of order N :=the order of $\chi_2\rho$. In fact, we have a geometric isomorphism

$$\det(\mathcal{G}) = \mathcal{L} := \mathcal{L}_{\Lambda(\prod_{s \in S}(t-s))},\ \text{with } \Lambda := \chi_2\rho.$$

To see this, use the fact that, on the one hand, $\det(\mathcal{G})$ geometrically takes values in the group of roots of unity in $\mathbb{Q}(\rho)$ (by [De-Const, Theorem 9.8] or [Se-Ta, Theorem 2 (ii)] or [Ka-ACT, 5.2 bis 1)]), so has order prime to p, and hence is everywhere tame. At a finite singularity $s \in S$, $\det(\mathcal{G})$ is $\mathcal{L}_{\Lambda(t-s)}$. Thus $\det(\mathcal{G}) \otimes \mathcal{L}^{-1}$ is lisse on $\mathbb{A}^1$ and tame at ∞, so geometrically trivial. Once we have this determination of $\det(\mathcal{G})$, we see that $\det(\mathcal{G})$ has order dividing N (because Λ does), and that the image of each inertia group $I(s)$ is cyclic of order N.

So far, everything we have said about the determination of $G^0_{\mathrm{geom},\mathcal{G}}$ holds for any very weakly supermorse f whose

$$\mathcal{F} := f_\star\overline{\mathbb{Q}_\ell}/\overline{\mathbb{Q}_\ell}$$

is geometrically irreducible. We now make use use of the special congruences of Theorem 4.1. View the traces of $\mathcal{G}$ as lying in the cyclotomic ring $\mathbb{Z}[\zeta_a]$, and pick a p-adic place $\mathcal{P}$ of $\mathbb{Q}(\zeta_a)$. Then precisely one of the characters of order a of k is the Teichmüller lift of its reduction mod $\mathcal{P}$, call it ρ_1. For this ρ_1, over any extension L/k in which -1 is an ath power, all Frobenius traces of $\mathcal{G}_{\rho_1}$ are nonzero, because they are all 1 mod $\mathcal{P}$. [For any L/k, they are $\rho_{1L}(-1)$ mod $\mathcal{P}$, the -1 needed to change $t - f(x)$ into $f(x) - t$ and apply Theorem 4.1.] But every ρ of order a is a $\mathrm{Gal}(\mathbb{Q}(\zeta_a)/\mathbb{Q})$-conjugate of ρ_1, and the trace of a given Frobenius on $\mathcal{G}_\rho$ is the Galois conjugate of the trace of the same Frobenius on $\mathcal{G}_{\rho_1}$. Thus for every ρ of order a, $\mathcal{G}_\rho$ has all Frobenius traces nonzero. By [Ka-Sar, 10.2] or [KT30, proof of Proposition 4.4], a geometrically irreducible local system on a smooth, geometrically connected X/k with k a finite field, all of whose Frobenius traces are nonzero, is not induced. □

References

[BCP] Bosma, W, Cannon, J., and Playoust, C., The MAGMA algebra system I: The user language. *J. Symbolic Comput.* **24** (1997), 235–265.

[De-Const] Deligne, P., Les constantes des equations fonctionelles des fonctions L, (Modular functions of one variable, II (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972), pp. 501–597, Lecture Notes in Math., Vol. **349**, Springer, Berlin, 1973).

[Ka-ACT] Katz, N., Affine cohomological transforms, perversity, and monodromy. J. Amer. Math. Soc.6(1993), no.1, 149-222.

[Ka-ESDE] Katz, N., Exponential sums and differential equations. Annals of Mathematics Studies, **124**. Princeton Univ. Press, Princeton, NJ, 1990. xii+430 pp.

[Ka-Int] Katz, N., Internal reconstruction of unit-root F-crystals via expansion-coefficients. With an appendix by Luc Illusie Ann. Sci. École Norm. Sup. (4)18(1985), no.2, 245-285.

[Ka-MMP] Katz, N., Moments, Monodromy, and Perversity. Annals of Mathematics Studies, **159**. Princeton University Press, Princeton, NJ, 2005. viii+475 pp.

[Ka-MG] Katz, N., On the monodromy groups attached to certain families of exponential sums, *Duke Math. J.* **54** (1987), no. 1, 41–56.

[Ka-RLS] Katz, N., Rigid Local Systems. Annals of Mathematics Studies, **139**. Princeton University Press, Princeton, NJ, 1996. viii+223 pp.

[Ka-Sar] Katz, N., and Sarnak, P., Random matrices, Frobenius eigenvalues, and monodromy. American Mathematical Society Colloquium Publications, **45**. American Mathematical Society, Providence, RI, 1999. xii+419 pp.

[KT30] Katz, N., and Tiep, P.H., Generalized Kloosterman sheaves, (in preparation).

[Se-Ta] Serre, J.-P., and Tate, J., Good reduction of abelian varieties, Annals of Mathematics , Nov., 1968, Second Series, Vol. 88, No. 3 (Nov., 1968), pp. 492–517.

[Wielandt] Wielandt, H., Finite Permutation Groups, Academic Press, 1964.

Department of Mathematics, Princeton University, Princeton, NJ 08544
E-mail address: nmk@math.princeton.edu

Department of Mathematics, Princeton University, Princeton, NJ 08544
E-mail address: yn0920@princeton.edu