Exponential Sums and Finite Groups

Nicholas M. Katz

Princeton University

Princeton, June 8, 2021

It is a great pleasure to be here, albeit virtually, to honor Luc Illusie, whom I have known and admired for the past half century.

This is a a report on joint work with Pham Huu Tiep and Antonio Rojas Léon.

Abhyankar's insight

For $C/\mathbb{C}$ a compact Riemann surface of genus g and $S \subset C$ a finite set of points, we have known for at least 90 years that its fundamental group $\pi_1(C \setminus S)$ is a free group on $2g + \#S - 1$ generators. Hence the finite quotient groups G of $\pi_1(C \setminus S)$ are those finite groups generatable by $2g + \#S - 1$ elements.

Now consider the same situation with $\mathbb{C}$ replaced by an algebraically closed field of characteristic $p > 0$.

For a finite group G, denote by $G_p \lhd G$ the normal subgroup generated by its p-Sylow subgroups.

Abhyankar had the insight that the finite groups G which were quotients of $\pi_1(C \setminus S)$ should be precisely those such that G/G_p was generatable by $2g + \#S - 1$ elements.

In particular, for $\mathbb{A}^1$, precisely those G's with $G = G_p$, and for $\mathbb{G}_m$ those G with G/G_p cyclic.

This was proven by Raynaud for $\mathbb{A}^1$ and extended to the general case by Harbater and also by Pop.

Suppose we are given a finite group G which can occur in characteristic p on $C \setminus S$, together with a faithful (complex) representation ρ of G.

Because G is finite, there is always some number field K such that the image of ρ lands in $\mathrm{GL}_n(K)$. If we now choose a prime number ℓ and an embedding of K into $\overline{\mathbb{Q}_\ell}$, we can view ρ as a representation $\rho : G \to \mathrm{GL}_n(\overline{\mathbb{Q}_\ell})$.

Since G is a quotient of $C \setminus S$, we can compose

$$\pi_1(C \setminus S) \twoheadrightarrow G \to \mathrm{GL}_n(\overline{\mathbb{Q}_\ell}),$$

to get a continuous ℓ-adic representation of $\pi_1(C \setminus S)$, i.e., an ℓ-adic *local system* of rank n on $C \setminus S$, whose image is the finite group G.

In the paragraph above, ℓ could have been any prime. But in order to apply the rich theory of ℓ-adic cohomology, we will always choose $\ell \neq p$, and our local systems will be ℓ-adic ones.

Working backwards, in three steps, from local systems to groups

In practice, our $C \setminus S$ comes from a $C_0 \setminus S_0$ over a finite extension field $\mathbb{F}_q$ of $\mathbb{F}_p$, and we look at a geometrically irreducible local system $\mathcal{H}_0$ on $C_0 \setminus S_0$ which is pure of some weight $w \geq 0$. We will know (in the sense of “have a formula for") the trace function of $\mathcal{H}_0$: for each finite extension E/k, and each point $x \in (C_0 \setminus S_0)(E)$, we will know $\mathrm{Trace}(Frob_{x,E}|\mathcal{H}_0)$. In all cases we consider below, this trace will lie in the cyclotomic integer ring $\mathbb{Z}[\zeta_p, \zeta_{q-1}]$.

Then $\det(\mathcal{H}_0)$ is geometrically of finite order, so by an α^{-deg} twist we may reduce to the case when $\det(\mathcal{H}_0 \otimes \alpha^{-deg})$ is arithmetically of finite order. In favorable cases, we can take $\alpha = \sqrt{q}^w$; then the Tate-twisted $\mathcal{H}_0(w/2)$ is both pure of weight zero and has determinant of arithmetically finite order.

Suppose we are in this favorable case where we can take $\alpha = \sqrt{q}^{w}$.

We have both G_{geom}, the Zariski closure in the ambient GL_n of the image of (the geometric monodromy of) $\mathcal{H}_0(w/2)$, and the larger group G_{arith}, the Zariski closure for the arithmetic monodromy of $\mathcal{H}_0(w/2)$.

By Grothendieck's global version of his local monodromy theorem, G_{geom} is a semisimple algebraic group over $\overline{\mathbb{Q}_\ell}$. In general we have $G_{\mathrm{geom}} \lhd G_{\mathrm{arith}}$. It is easy to see that G_{geom} is finite if and only if G_{arith} is finite.

By purity, one further knows that G_{arith} is finite if and only if for every finite extension $E/\mathbb{F}_q$, and every point $x \in (C_0 \setminus S_0)(E)$, the Frobenius trace $\mathrm{Trace}(Frob_{x,E}|\mathcal{H}_0(w/2))$ is an algebraic integer.

This trace is, by definition of the Tate-twist,

$$(1/\sqrt{\#E}^{w})\mathrm{Trace}(Frob_{x,E}|\mathcal{H}_0).$$

Its only possible nonintegrality comes from the division of a cyclotomic integer by a power of $\sqrt{q}$. Concretely, then, the criterion for finitenes of G_{arith} is that for each p-adic ord_p on the field $\mathbb{Q}(\zeta_p, \zeta_{q-1})$, each finite extension $E/\mathbb{F}_q$, and every point $x \in (C_0 \setminus S_0)(E)$, we have

$$\mathrm{ord}_p(\mathrm{Trace}(Frob_{x,E}|\mathcal{H}_0)) \geq \mathrm{ord}_p(\sqrt{\#E}^{w}).$$

Step 1: find “interesting" local systems $\mathcal{H}_0$ as inputs.

Step 2: for each, either prove G_{arith} is finite, or prove that it is not finite.

Step 3: if G_{arith} is finite, determine G_{geom} and G_{arith}. If G_{arith} is not finite, determine G_{geom} and G_{arith}.

Three examples of interesting irreducible local systems on open curves over $\mathbb{F}_q$

Ambient setting: ψ and χ:

ψ is a nontrivial additive character of $\mathbb{F}_p$ (viewed as having values in $\overline{\mathbb{Q}_\ell}$); leads to Artin-Schreier sheaf $\mathcal{L}_\psi$ on $\mathbb{A}^1/\mathbb{F}_p$. Trace at points of $k/\mathbb{F}_p$ by composition with $\mathrm{Trace}_{k/\mathbb{F}_p}$.

χ is a (possibly trivial) character of $\mathbb{F}_q^\times$ (viewed as having values in $\overline{\mathbb{Q}_\ell}$); leads to Kummer sheaf $\mathcal{L}_\chi$ on $\mathbb{G}_m/\mathbb{F}_q$. Trace at points of $k^\times$ for $k/\mathbb{F}_q$ by composition with $\mathrm{Norm}_{k/\mathbb{F}_q}$.

Example 1. On $\mathbb{G}_m/\mathbb{F}_q$, we have the hypergeometric sheaves

$$\mathcal{H}(\chi_1, ..., \chi_n; \rho_1, ..., \rho_m)$$

with $n > m \geq 0$, each χ_i and each ρ_j is a (possibly trivial) character of $\mathbb{F}_q^\times$, and no χ_i is any ρ_j. We know that **if** G_{geom} is finite, then we can take $\alpha = \sqrt{q}^w$. [But this need not be true for a hypergeometric whose G_{geom} is infinite.]

Example 2. On $\mathbb{A}^1/\mathbb{F}_q$, we have the local systems whose trace functions are

$$t \mapsto -\sum_x \psi(f(x) + tx)\chi(x),$$

with $f(x) \in \mathbb{F}[x]$ a polynomial of prime to p degree $n \geq 2$ and χ either the trivial or, if p is odd, the quadratic character of $\mathbb{F}_q^\times$. We can take $\alpha = \sqrt{q}$. [This can be false for other χ, already in the “baby case" when $f(x) = x^2$.]

Example 3. On a hyperelliptic curve $U := C \setminus \{\infty\}$ with equation $y^2 = f_{2g+1}(x)$, in odd characteristic p, we have the local systems whose trace functions are

$$t \mapsto -\sum_{(x,y)\in U} \psi(yg(x) + tx)\chi(x),$$

with both f_{2g+1} and g in $\mathbb{F}_q[x]$ of degree $\leq n$ and χ either the trivial or the quadratic character of $\mathbb{F}_q^\times$, and with the proviso that $2g + 1 + 2\deg(g)$, the order of pole at ∞ of $yg(x)$, is prime to p. We can take $\alpha = \sqrt{q}$. [This can be false for other χ, already in the “baby case" when $g(x) = 1$ and $f(x) = x$.]

An open problem in Step 2

In each of the three examples, the local system is pure of weight one; this is Weil's theorem for curves, and after replacing $\mathcal{H}_0$ by $\mathcal{H}_0(1/2)$ we have finite G_{geom} if and only if $\mathcal{H}_0(1/2)$ has all its Frobenius traces algebraic integers (which we have seen is equivalent to all these traces being p integral for all p-adic places). The question is how long we have to wait, i.e., how many traces we need to compute, to decide if in fact all Frobenius traces are p-integral.

Consider any of the three example collections of local systems, over a fixed $\mathbb{F}_q$ and with a fixed auxiliary integer n. Here is a mock theorem:

mock Theorem, correct but useless per se
There exists a constant $N = N(q, n)$ such that in each of these collections of local systems, if all Frobenius traces on a given $\mathcal{H}_0(1/2)$ are algebraic integers at all points in all extensions of degree $\leq N$, then this $\mathcal{H}_0(1/2)$ has all traces algebraic integers.

proof There are only finitely many local systems in question. For each of the finitely many with infinite G_{geom}, record the degree of an extension field over which some Frobenius has a non-integer trace. The the sup of these extension degrees works as the required $N = N(q, n)$.

The question is the extent to which $N(q, n)$ can be explicitly bounded as a function of the input data (q, n). With a triply exponential bound? With a bound which is polynomial in $n \log(q)$, the number of bits in the input data?

What we know and don't know in Step 3

the hypereometric case In joint work with Pham Huu Tiep and Antonio Rojas Léon, we have determined which of the 26 sporadic groups can possibly occur as G_{geom} for a hypergeometric sheaf, and exhibited for each of these groups a hypergeometric that realizes it **.**

With Pham Huu Tiep, we have done the same thing for finite groups of Lie type: analyzed which can possibly be realized by hypergeometric sheaves, and realized each.

the $\mathbb{A}^1$ case The general situation on $\mathbb{A}^1$ is less clear. When the polynomial $f(x)$ in $\psi(f(x) + tx)\chi(x)$ is the single monomial x^A with $A > 1$ and prime to p, with Pham Huu Tiep we have complete understanding.[But as soon as we allow more general $f(x)$, we know almost nothing.] For the local systems

$$t \mapsto -\sum_x \psi(x^A + tx)\chi(x),$$

here is the complete story.

Although the statements won't mention hypergeometric sheaves, their proofs depend completely on that theory.

If G_{geom} is finite, there are two sporadic cases:

$p = 5, A = 7, \chi = \mathbb{1}$, and G_{geom} is $2.\mathrm{J}_2$.

$p = 3, A = 23, \chi = \chi_2$, and G_{geom} is the Conway group Co_3.

In addition, there are four infinite families, in which q denotes a power of p and we are in characteristic p.

p odd, $A = (q+1)/2, \chi = \mathbb{1}$ or χ_2: G_{geom} is the image of $\mathrm{SL}_2(q)$ in a Weil representation.

p odd, $A = 2q-1, \chi = \chi_2$: $G_{\mathrm{geom}} = \mathsf{A}_{2q}$ in its deleted permutation representation.

$A = (q^n+1)/(q+1)$ with $n \geq 3$ odd and $\chi^{q+1} = \mathbb{1}$: G_{geom} is the image of $\mathrm{SU}_n(q)$ in a Weil representation (except for the special case $(n=3, q=2)$ of $\mathrm{SU}_3(2)$).

$A = q+1 = p^f+1$, and $\chi = \mathbb{1}$. If $p > 2$, G_{geom} is the Heisenberg group p_+^{1+2f} of order pq^2 and exponent p. If $p = 2$, G_{geom} is the extraspecial 2-group 2_-^{1+2f}. We also have the degenerate case $A = 2 = 1 + p^0$ with p odd, whose G_{geom} is the cyclic group of order p.

If G_{geom} is infinite then G_{geom} is

Sp_{A-1} if A is odd, and $\chi = \mathbb{1}$.

SO_A if $A \neq 7$ is odd, p is odd, and $\chi = \chi_2$.

G_2 if $A = 7$, p is odd, and $\chi = \chi_2$.

SL_A if A is odd and $\chi^2 \neq \mathbb{1}$.

SL_{A-1} if $A \geq 4$ is even and $\chi = \mathbb{1}$.(The $A = 2, \chi = \mathbb{1}$) case is on the finite list.)

SL_A if $A \geq 4$ is even and $\chi \neq \mathbb{1}$.

$\{g \in \mathrm{GL2} | det(g)^p = 1\}$ if $A = 2$ and $\chi \neq \mathbb{1}$.

the hyperelliptic case

Here we can be very brief: we know nothing. Does/should the answer depend on which hyperelliptic curve we work on?

MUCH REMAINS TO BE DONE.