Counting points on hyperelliptic curves using Monsky-Washnitzer cohomology

Kiran S. Kedlaya
University of California, Berkeley

Midwest Arithmetical Geometry in Cryptography
November 2, 2001

Preprint available at

`http://www.math.berkeley.edu/~kedlaya`;

to appear in *Journal of the Ramanujan Mathematical Society*.

Notation

$\mathbb{F}_q$ — the finite field of $q = p^n$ elements (p odd)

$\mathbb{Z}_q$ — the unramified extension of $\mathbb{Z}_p$ with residue field $\mathbb{F}_q$ (a/k/a $W(\mathbb{F}_q)$)

$\mathbb{Q}_q$ — the fraction field of $\mathbb{Z}_q$

$\overline{P}(x)$ — a monic polynomial of degree $2g+1$ over $\mathbb{F}_q$ with distinct roots over $\overline{\mathbb{F}_q}$

$X : y^2 = \overline{P}(x)$ — a hyperelliptic curve of genus g with a rational Weierstrass point

The problem

Problem: find $\alpha_1, \ldots, \alpha_{2g} \in \mathbb{C}$ such that $|\alpha_i| = \sqrt{q}$, $\alpha_i \alpha_{g+i} = q$ and

$$\#X(\mathbb{F}_{q^r}) = q^{rn} + 1 - \alpha_1^r - \cdots - \alpha_{2g}^r,$$

guaranteed to exist by the Weil Conjectures. Then

$$\#\,\mathrm{Jac}(X)(\mathbb{F}_{q^r}) = \prod_{i=1}^{2g} (\alpha_i^r - 1).$$

The characteristic polynomial of Frobenius:

$$Q(t) = \prod_i (t - \alpha_i)$$

has integer coefficients, which are bounded in terms of q and g.

Approaches for $g = 1$

ℓ-adic approach [Schoof, Atkin, Elkies]: compute $\#X(\mathbb{F}_q)$ modulo enough small primes ℓ to determine $Q(t)$ by the Chinese Remainder Theorem (using division polynomials).

p-adic approach [Satoh]: compute $\#X(\mathbb{F}_q)$ modulo a sufficiently high power of p (using Serre-Tate lift).

The ℓ-adic approach is better for p large, the p-adic for p small.

Trouble in higher genus

ℓ-adic approach [Pila]: requires explicit equations for the Jacobian of the curve, which are extremely cumbersome to write down.

p-adic approach [Satoh]: uses formal groups, again cumbersome to write down.

Both approaches are exponential in the genus.

Our goal: give a p-adic computation that is also polynomial in the genus. (Another approach is due to Lauder and Wan; see also Harley's talk.)

The strategy

We construct a vector space H ("cohomology") that obeys a Lefschetz trace formula:

$$\#X(\mathbb{F}_{q^r}) = q^r + 1 - \text{Trace}(F^r, H)$$

for some linear map $F : H \to H$ ("Frobenius"). Then compute $Q(t)$ (resp., the α_i) as the characteristic polynomial (resp., the eigenvalues) of F.

We will use Monsky-Washnitzer cohomology, which will produce vector spaces over $\mathbb{Q}_q$. This cohomology is defined for any *affine* variety of characteristic p, so we must "puncture" the curve to use it.

The Monsky-Washnitzer cohomology of X

Choose a monic polynomial $P(x)$ over $\mathbb{Z}_q$ congruent to $\overline{P}(x)$ modulo p. The ring R consists of power series

$$\sum_{i=0}^{\infty} \sum_{j=-\infty}^{\infty} a_{ij} x^i y^j$$

over $\mathbb{Z}_q$ such that $v_p(a_{ij})/(i + |j|)$ is eventually bounded away from 0 ("overconvergence"), modulo the relation $y^2 = P(x)$. (One can assume $a_{ij} = 0$ for $i > 2g$.)

Geometrically, this is the "weak completion" of the coordinate ring of the affine variety

$$V = \mathrm{Spec}\, \mathbb{Z}_q[x, y, y^{-1}]/(y^2 - P(x)).$$

(That ring has the same reduction modulo p as R does.)

Overconvergence in detail

The series

$$\sum_{i=0}^{\infty} x^{p^i}$$

is not p-adically convergent.

The series

$$\sum_{i=0}^{\infty} p^i x^{p^i}$$

is convergent, but not overconvergent.

The series

$$\sum_{i=0}^{\infty} p^i x^i$$

is overconvergent.

A basis in cohomology

The Monsky-Washnitzer cohomology space H is the set of differentials $r\,dx + s\,dy$ for $r, s \in R[\frac{1}{p}]$, modulo df and $f(2y\,dy - P'(x)\,dx)$ for any $f \in R[\frac{1}{p}]$.

This is a vector space over $\mathbb{Q}_q$, which we can split into plus and minus eigenspaces for the map $y \mapsto -y$ of R. The resulting spaces have bases

$$H_- : \quad \frac{x^i\,dx}{y} \quad (i = 0, \ldots, 2g)$$

$$H_+ : \quad \frac{x^i\,dx}{y^2} \quad (i = 0, \ldots, 2g-1).$$

(We'll see why later.)

Whither overconvergence?

Had we not imposed the overconvergence condition, then

$$\sum_{i=0}^{\infty} p^i x^{p^{2i}-1}\,dx$$

would not be zero in H, because its integral

$$\sum_{i=0}^{\infty} p^{-i} x^{p^{2i}}$$

does not converge. In fact, H would be infinite dimensional in this case.

A typical example in R is

$$\sum_{i=1}^{\infty} p^i x^{i-1}\,dx = d\left(\sum_{i=1}^{\infty} (p^i/i)x^i\right).$$

A trace formula

Let $F : R \to R$ be any map lifting the q-power Frobenius map $x \mapsto x^q$ modulo p and commuting with $y \mapsto -y$. The Monsky-Washnitzer theory implies the following Lefschetz trace formula:

$$\#X(\mathbb{F}_{q^r}) = q^r + 1 - \mathrm{Trace}(F^r, H_-).$$

Our plan: compute such a map, compute the matrix by which F acts on some basis of H_- to sufficient p-adic precision, and read off the characteristic polynomial.

To simplify matters, we'll construct F as the n-th power of a map F_p lifting the p-power Frobenius $x \mapsto x^p$.

The trace formula in detail

Let Z be the set of Weierstrass points of X and $Y = X - Z$. Then Y is affine and $Y = \mathrm{Spec}(R/pR)$, so Monsky-Washnitzer predicts that

$$\begin{aligned} \#X(\mathbb{F}_{q^r}) - \#Z(\mathbb{F}_{q^r}) &= \#Y(\mathbb{F}_{q^r}) \\ &= q^r - \mathrm{Trace}(F^r, H). \end{aligned}$$

Let $\pi : X \to \mathbb{P}^1$ be the quotient by $y \mapsto -y$. Then $\mathbb{P}^1 - \pi(Z)$ is again affine, so

$$\begin{aligned} q^r + 1 - \#Z(\mathbb{F}_{q^r}) &= \#(\mathbb{P}^1 - \pi(Z))(\mathbb{F}_{q^r}) \\ &= q^r - \mathrm{Trace}(F^r, H_+). \end{aligned}$$

Subtracting these two yields the trace formula given above:

$$\#X(\mathbb{F}_{q^r}) = q^r + 1 - \mathrm{Trace}(F^r, H_-).$$

The Frobenius map

Recall that $\mathbb{Z}_q$ has a canonical map σ lifting the map $t \mapsto t^p$ modulo p. The "Frobenius" map $F_p : R \to R$ is given by

$$\begin{aligned} x &\mapsto x^p \\ y &\mapsto y^p(1+pE)^{1/2} \\ &= y^p \sum_{i=0}^{\infty} \binom{1/2}{i} p^i E^i, \end{aligned}$$

where $E = (P(x)^\sigma - P(x)^p)/(py^{2p})$. (The valuation of $\binom{1/2}{i}$ is $O(\log i)$, so this is overconvergent.) Extend F_p to H by sending $f\,dx$ to $f^\sigma\, d(x^p) = px^{p-1} f^\sigma\, dx$.

In practice, compute $F_p(y)$ using a Newton iteration, not the power series.

What is the matrix?

We'd like to compute the matrix by which F_p, and by extension $F = F_p^n$, act on the basis $x^i\,dx/y$ $(i = 0, \ldots, 2g-1)$ of H_-.

To do this, we first apply F_p to each differential, then rewrite the result as an exact differential (which maps to 0 in H_-) plus a linear combination of basis elements.

So we need a procedure to "reduce" an arbitrary differential to a linear combination of basis elements, plus an exact differential.

Reduction of differentials, part 1

Given a differential $A(x)y^{-2s-1}\,dx$ with $\deg A \leq 2g$, write

$$A(x) = B(x)P(x) + C(x)P'(x)$$
$$\deg B \leq 2g-1, \qquad \deg C \leq 2g.$$

Since $P'(x)\,dx = 2y\,dy$, we get the relation

$$\frac{C(x)P'(x)\,dx}{y^{2s+1}} = \frac{2C(x)\,dy}{y^{2s}} = \frac{2C'(x)\,dx}{(2s-1)y^{2s-1}}$$

in H_-. Thus

$$\frac{A(x)}{y^{2s+1}}\,dx = \frac{(2s-1)B(x) + 2C'(x)}{(2s-1)y^{2s-1}}\,dx,$$

and we have reduced the exponent of y in the denominator.

Reduction of differentials, part 2

A similar argument allows one to eliminate large positive powers of y. Namely, given $A(x)y^{2s+1}\,dx$, first rewrite it as $A(x)P(x)^{s+1}y^{-1}\,dx$. Now use the relation

$$\begin{aligned} 0 &= d(x^m y) \\ &= mx^{m-1}y\,dx + x^m\,dy \\ &= \frac{2mx^{m-1}P(x) + x^m P'(x)}{2y}\,dx \end{aligned}$$

to successively eliminate the highest powers of x. (The coefficient of x^{2g+m} in the numerator is $2m + (2g-1) \neq 0$.)

The algorithm (sketch, part 1)

Step 1: use Newton's iteration to compute $F_p(y^{-1})$ to "high accuracy".

Step 2: for $i = 0, \ldots, 2g-1$, compute an approximation of

$$\begin{aligned} F_p(x^i\,dx/y) &= F_p(y^{-1})F_p(x)^i F_p(dx) \\ &= F_p(y^{-1})x^{pi}px^{p-1}\,dx. \end{aligned}$$

and rewrite it as $\sum_i A_i(x)y^{2i+1}\,dx$ where each A_i has degree at most $2g$.

Step 3: take the finite sum $\sum_i A_i(x)y^{2i+1}\,dx$ and use the reduction procedures to obtain an equivalent differential of the form $A(x)\,dx/y$ with $\deg A \leq 2g-1$.

The algorithm (sketch, part 2)

Step 4: let M be the matrix with M_{ij} equal to the coefficient of $x^i\,dx/y$ in the reduction of $F_p(x^j\,dx/y)$. Then the matrix by which $F = F_p^n$ acts on the basis $x^i\,dx/y$ of H_- is

$$M' = MM^\sigma \cdots M^{\sigma^{n-1}}.$$

Step 5: compute the characteristic polynomial of M' to high enough accuracy to determine $Q(t)$ and the α_i.

Precision

How accurate is accurate enough?

If $\deg A \leq 2g$ and A has coefficients in $\mathbb{Z}_q$, the reduction of $A(x)y^{2s+1}\,dx$ will have coefficients in $\mathbb{Z}_q$ if you multiply it by $p^{\log_p 2|s|+1}$.

So computing $O(gn)$ terms of

$$F_p(y^{-1}) = y^p \sum_{i=0}^{\infty} \binom{1/2}{i} p^i E^i$$

will give the needed $O(gn)$ places of precision in the final matrix. (And all constants are explicit.)

Asymptotics (for fixed p)

One must carry $O(gn)$ places of precision in all computations in order to get enough accuracy to uniquely determine the characteristic polynomial $Q(t)$. Moreover, each place of precision requires $O(n)$ bits, since we're working in $\mathbb{Z}_q$.

The limiting memory constraint is storing the approximation of $F(x^i\,dx/y)$. It consists of $O(gn)$ terms, each a polynomial of degree $2g$, each of whose coefficients requires $O(gn^2)$ bits to store. Thus it requires $O(g^3n^3)$ memory.

Asymptotics (for fixed p) continued

The limiting time constraint is reducing each of these $2g$ approximations. This requires eliminating $O(gn)$ powers of y; each such step requires $O(g^{1+\epsilon})$ operations on numbers of size $O(gn^2)$. This yields a total time of $O(g^{4+\epsilon}n^{3+\epsilon})$.

For $g = 1$, this is the same asymptotic as Satoh's algorithm. It is unclear how the two compare in practice, though.

Generalizations

The Monsky-Washnitzer theory applies to any affine scheme; the main difficulties are:

- Finding an easily computed Frobenius lift;
- Finding an efficient reduction procedure for differentials.

Gaudry and Gurel treat the "superelliptic" case $y^n = f(x)$, where n is coprime to p, by a similar approach.

The Lauder-Wan method is more general, but less practical. Can it be combined with this?

The End