Zeitschrift: L'Enseignement Mathématique
Herausgeber: Commission Internationale de l'Enseignement Mathématique
Band: 56 (2010)

Artikel: A note on lower bounds for frobenius traces
Autor: Bombieri, Enrico / Katz, Nicholas M.
DOI: https://doi.org/10.5169/seals-283519

L'Enseignement Mathématique (2) 56 (2010), 203–227

A NOTE ON LOWER BOUNDS FOR FROBENIUS TRACES

by Enrico BOMBIERI and Nicholas M. KATZ

1. INTRODUCTION

This paper grew out of the following question. Given an ordinary elliptic curve $E/\mathbf{F}_q$ over a finite field $\mathbf{F}_q$ of characteristic p, consider the sequence of integers $A(n)$, $n \geq 1$, defined by

$$\#E(\mathbf{F}_{q^n}) = q^n + 1 - A(n)\,.$$

Is it true that as n grows we have $|A(n)| \to \infty$?

Without the hypothesis "ordinary" the answer can be no, because for a supersingular elliptic curve one can have $A(n) = 0$ on entire arithmetic progressions of n. On the other hand, all the $A(n)$ in the supersingular case are divisible, as algebraic integers, by $q^{n/2}$, so the non-zero $A(n)$ must have $|A(n)| \geq q^{n/2}$. If instead $E/\mathbf{F}_q$ is ordinary, then all the $A(n)$ are not zero because they are all prime to p, so this vanishing problem at least disappears.

The $A(n)$ are the traces of the iterates of a certain Frobenius endomorphism F and this leads to the more general question of when we can assert that in the sequence $|\mathrm{Trace}(F^n)|$, $n \geq 1$, the non-zero terms tend to ∞.

The purpose of this note is to explain how classical results on recurrent sequences answer these questions. Because of the "culture gap" between the communities of those who know these classical results and those who are interested in traces of Frobenius, we have written this note so to make it accessible to members of both communities, at the risk that readers may find parts of this note overly detailed.

We will use three different methods to approach the problem. The Skolem-Mahler-Lech theorem on recurrent sequences is easy to prove and provides a "soft" answer, soft in the sense that it gives no estimate of the rate at which the non-zero terms tend to ∞. The other two methods lie much deeper. A theorem

due independently to Evertse and to van der Poorten and Schlickewei, itself based on an improved version of Schmidt's subspace theorem, gives such a rate, albeit ineffective in certain parameters. For elliptic curves (and some other exponential sums, including classical Kloosterman sums), the Baker-Wüstholz theorem gives an even better rate, this time effective in all parameters.

The problem of obtaining effective lower bounds in the most general case remains unsolved and probably lies very deep.

ACKNOWLEDGEMENTS. It is a pleasure to thank Umberto Zannier for his helpful comments on an earlier version of this paper.

2. UNBOUNDEDNESS, VIA SKOLEM'S METHOD

We begin by recalling the relevant version of the Skolem-Mahler-Lech theorem. For the convenience of the reader, we also recall its proof.

THEOREM 2.1. *Let K be an algebraically closed field of characteristic zero. Fix an integer $n \geq 1$, n numbers $\alpha_1, \ldots, \alpha_n$ in $K^\times$, the "eigenvalues", and n polynomials $\lambda_1(x), \ldots, \lambda_n(x)$ in $K[x]$, the "coefficients", not all of which are zero. For each integer $k \geq 1$, define*

$$A(k) := \sum_{i=1}^{n} \lambda_i(k)\, \alpha_i^k .$$

Then we have the following results.

(i) *Suppose that no ratio α_i/α_j, $i \neq j$, is a root of unity. Then there are only finitely many integers $k \geq 1$ for which $A(k) = 0$.*

(ii) *The integers $k \geq 1$ for which $A(k) = 0$ are the union of a (possibly empty) finite set together with a finite number, possibly zero, of arithmetic progressions to some common modulus D; we can take D to be the order of the group of roots of unity generated by all those roots of unity which are of the form α_i/α_j for some i, j.*

(iii) *Suppose that for some index i_0, $\lambda_{i_0}(x) \neq 0$ and, for any $j \neq i_0$, the ratio α_j/α_{i_0} is not a root of unity. Then there are only finitely many integers $k \geq 1$ for which $A(k) = 0$.*

(iv) *Suppose that no α_i is a root of unity. Then for any $\mu \neq 0$ in K, there are at most finitely many integers $k \geq 1$ with $A(k) = \mu$.*

Proof. (i) Let Λ be the set of coefficients of the polynomials $\lambda_i(x)$. It is standard that for almost all primes p we can embed the finitely generated ring

$$\mathbf{Z}[\alpha_1, 1/\alpha_1, \ldots, \alpha_n, 1/\alpha_n, \Lambda, \mu]$$

into the ring of integers $\mathcal{O}_\mathcal{P}$ in a finite extension $E_\mathcal{P}$ of $\mathbf{Q}_p$, cf. [C1] for an elementary proof or [Ka96], 5.9.3. (In Cassels it is shown that if K is any finitely generated field of zero characteristic and C is a finite subset of $K^\times$ then there is a set of primes p of positive density such that, for each p in this set, there is an embedding of K in the p-adic field $\mathbf{Q}_p$ in which all elements of C are units.)

We choose such an embedding, denote by $\pi \in \mathcal{O}_\mathcal{P}$ a uniformizing parameter, by $|\ |_\mathcal{P}$ the extension of the usual p-adic absolute value to $E_\mathcal{P}$, by $\mathrm{ord}_\mathcal{P}$ the associated additive valuation, and by L the cardinality of the finite group $\mathcal{O}_\mathcal{P}^\times/(1+p\pi\mathcal{O}_\mathcal{P})$.

For each i, we have

$$(\alpha_i)^L \in 1 + p\pi\mathcal{O}_\mathcal{P}\,.$$

Hence in each arithmetic progression $\{a+kL\}_{k\in\mathbf{Z}}$ modulo L, we have

$$A(a+kL) = \sum_{i=1}^{n} \alpha_i^a \lambda_i(a+kL)\,(\alpha_i^L)^k\,,$$

which we can view as the case where the eigenvalues are the α_i^L and the coefficients are $\alpha_i^a\lambda_i(a+xL)$. Notice that the new eigenvalues α_i^L continue to satisfy the condition that their ratios are not roots of unity.

Looking at each of these progressions separately, it suffices to prove (i) under the additional hypothesis that the n numbers α_i each lie in $1+p\pi\mathcal{O}_\mathcal{P}$. The key observation is that the functions

$$\log(1+z) = \sum_{m=1}^{\infty} (-1)^{m-1}\frac{z^m}{m}$$

and

$$\exp(z) = \sum_{m=0}^{\infty} \frac{z^m}{m!}$$

are a pair of inverse group isomorphisms between the multiplicative group $1+p\pi\mathcal{O}_\mathcal{P}$ and the additive group $p\pi\mathcal{O}_\mathcal{P}$. (Indeed, for any element $\mu \in \mathcal{O}_\mathcal{P}$ with $\mu^{p-1} \in p\mathcal{P}\mathcal{O}_\mathcal{P}$, log and exp are inverse group isomorphisms between the multiplicative group $1+\mu\mathcal{O}_\mathcal{P}$ and the additive group $\mu\mathcal{O}_\mathcal{P}$, see [DGS], p. 52.)

Thus distinct elements $\alpha_i \in 1 + p\pi\mathcal{O}_\mathcal{P}$ have distinct logarithms

$$\beta_i := \log(\alpha_i) = \sum_{m=1}^{\infty} (-1)^{m-1} \frac{(\alpha_i - 1)^m}{m} \in p\pi\mathcal{O}_\mathcal{P} .$$

The power functions $n \mapsto \alpha_i^n = \exp(\beta_i)^n = \exp(\beta_i n)$ are interpolated by the functions $z \mapsto \exp(\beta_i z)$, whose power series are easily seen to lie in $\mathcal{O}_\mathcal{P}[[\pi z]]$.

We next show that these n analytic functions $\exp(\beta_i z)$ have power series that are linearly independent over $E_\mathcal{P}[z]$. For completeness, we repeat here the standard proof. Suppose that $P_i(z)$, $(i = 1, \ldots, n)$, are non-zero polynomials in $E_\mathcal{P}[z]$, of degree δ_i, which we may and will assume to be monic. We will show that the n power series $f_i(z) := P_i(z)\, e^{\beta_i z}$ are linearly independent over $E_\mathcal{P}$. It suffices to show that their Wronskian

$$\Delta := \det\left(\left(\frac{\mathrm{d}}{\mathrm{d}z}\right)^{j-1} f_i(z)\right)_{i,j=1,\ldots,n}$$

is not zero. The (i,j)-th entry of the matrix is easily calculated to be

$$(\beta_i^{j-1} z^{\delta_i} + \text{lower degree terms})\, e^{\beta_i z} .$$

Therefore, the determinant is

$$\Delta = \left\{ z^{\sum \delta_i} + \text{lower degree terms} \right\} \mathrm{Vand}(\beta_1, \ldots, \beta_n)\, e^{(\sum \beta_i) z}$$

with Vand the Vandermonde determinant. The β_i, $i = 1, \ldots, n$, are distinct, hence the Vandermonde determinant is not 0.

We now return to the proof of part (i) of the theorem. Since not all coefficients $\lambda_i(x)$ vanish, the function

$$F(z) := \sum_{i=1}^{n} \lambda_i(z) \exp(\beta_i z)$$

is not zero in $\mathcal{O}_\mathcal{P}[[\pi z]]$. It follows that $F(z)$ has at most finitely many zeroes in $\mathcal{O}_\mathcal{P}$ and *a fortiori* has at most finitely many integer zeroes, which will prove what we want. This is an easy consequence of the Weierstrass Preparation Theorem applied to the power series ring $\mathcal{O}_\mathcal{P}[[\pi z]]$ (see Lang [La], Thm. 9.2), or of the theory of Newton polygons (see, for example, Dwork [Dw], Thm. 1.1 or Dwork, Gerotto, Sullivan [DGS], II.2.1). In its most elementary form, this finiteness of the number of zeroes follows from Strassmann's Theorem:

If $f(z) = \sum a_m z^m$ is convergent for $|z|_\mathcal{P} \le 1$ and not identically 0, and M is the largest index m for which $|a_m|_\mathcal{P}$ reaches its maximum, then the equation $f(z) = 0$ has at most M zeroes ζ with $\mathrm{ord}_p(\zeta) \ge 0$.

The following simple proof by induction on M can be found in Cassels [C2], Thm. 4.1. Since $\sum a_m z^m$ is convergent for $|z|_{\mathcal{P}} \leq 1$, we have $|a_m|_{\mathcal{P}} \to 0$, hence M exists. If $M = 0$, there is nothing to prove. Now if $f(\zeta) = 0$ we have

$$\begin{aligned} f(z) = f(z) - f(\zeta) &= \sum_{m=1}^{\infty} a_m (z^m - \zeta^m) \\ &= (z-\zeta) \sum_{m=1}^{\infty} \sum_{j=0}^{m-1} a_m z^j \zeta^{m-1-j} = (z-\zeta) \sum_{m=0}^{\infty} b_j z^j \\ &= (z-\zeta) g(z) , \end{aligned}$$

say, with

$$b_j = \sum_{m=j+1}^{\infty} a_m \zeta^{m-1-j} .$$

From this, it is clear (we are dealing with an ultrametric valuation) that $|b_j|_{\mathcal{P}} \to 0$ as $j \to \infty$. Moreover, it is immediate that $|b_j|_{\mathcal{P}} \leq |a_M|_{\mathcal{P}}$ for all j, $|b_{M-1}|_{\mathcal{P}} = |a_M|_{\mathcal{P}}$, and $|b_j|_{\mathcal{P}} < |a_M|_{\mathcal{P}}$ if $j > M$; the result follows by induction applied to $g(z) = \sum b_j z^j$, which we may because $|b_j|_{\mathcal{P}} \to 0$, so the sum is convergent in $|z|_{\mathcal{P}} \leq 1$.

A refinement of Strassmann's Theorem is the p-adic Rouché theorem (see [DGS], IV.4.2 and its more general formulation for quotients of analytic functions[1]), rather than just power series in $E_{\mathcal{P}}[[z]]$):

Let $f(z) = \sum a_m z^m \in E_{\mathcal{P}}[[z]]$ be a power series convergent in $|z|_{\mathcal{P}} \leq 1$ and let $\|f\| := \max_m |a_m|_{\mathcal{P}}$. If $h(z) \in E_{\mathcal{P}}[[z]]$ is another power series convergent in $|z|_{\mathcal{P}} \leq 1$ and with $\|h\| < \|f\|$, then f and $f+h$ have the same finite number of zeroes in the disk $|z|_{\mathcal{P}} \leq 1$.

Once we have (i), we get (ii) and (iii) by partitioning the eigenvalues α_i into equivalence classes according to the equivalence relation where $a \equiv b$ if and only if b/a is a root of unity. By renumbering, we may assume that $\alpha_1, \ldots, \alpha_r$ are representatives of these equivalence classes, and that the class of α_i consists of $\zeta_{i,j}\alpha_i$, for $j = 1, \ldots, n_i$, with suitable roots of unity $\zeta_{i,j}$ of order dividing some positive integer D.

[1]) This extension is important, because analytic continuation in a p-adic field cannot be done by Weierstrass's method using Taylor series.

Then for a fixed integer $0 \leq a < D$, and any integer $k \geq 1$, the sequence $k \mapsto A(a + kD)$ is of the same form, with r eigenvalues α_i^D, $i = 1, \ldots, r$, except that now it may be the case that all the coefficients vanish. We do not care about the exact formulas for these coefficients, except to note that for each equivalence class which is a singleton, say α_{i_0}, the new coefficient of α_{i_0} is $\alpha_{i_0}^a \lambda_{i_0}(a + xD)$. If all coefficients vanish, then we have vanishing on the entire progression. If not, then by (i) we only have finitely many vanishing terms in the progression. This gives (ii) and (iii).

Suppose now that no α_i is a root of unity. We get (iv) by applying (iii) to the situation with $n+1$ eigenvalues $(\alpha_1, \ldots, \alpha_n, 1)$ and coefficients $(\lambda_1(x), \ldots, \lambda_n(x), -\mu)$, for here the equivalence class of the eigenvalue 1 is a singleton, whose coefficient $-\mu$ is not zero.

COROLLARY 2.2. *Let K be an algebraically closed field of characteristic zero, $n \geq 1$ an integer, and $F \in GL(n, K)$ an $n \times n$ invertible matrix whose reversed characteristic polynomial* $\det(I - FT)$ *has integer coefficients. Suppose that no eigenvalue of F is a root of unity. Define a sequence of integers $A(n)$ by*

$$A(n) := \mathrm{Trace}(F^n), \quad n \geq 1.$$

Then the non-zero $A(n)$ have $|A(n)| \to \infty$. More precisely, for any integer $M \geq 1$, there exists an integer $k_M \geq 1$ such that if $k > k_M$, then either $|A(k)| > M$ or $A(k) = 0$.

Proof. Apply Theorem 2.1 (iv), to the eigenvalues α_i of F, taking all $\lambda_i = 1$. For any integer $k \geq 0$, $A(k)$ is an integer, by the integrality assumption on the coefficients of the characteristic polynomial. There are at most finitely many integers $k \geq 0$ for which $0 < |\mathrm{Trace}(F^k)| \leq M$, hence taking k_M to be the largest of these, we get the assertion.

Here is another corollary. As before, K is an algebraically closed field of characteristic zero, $n \geq 1$ an integer, and $F \in GL(n, K)$ is an $n \times n$ invertible matrix whose reversed characteristic polynomial $P(T) := \det(I - FT)$ has integer coefficients. Given an integer $k \geq 1$, we say that an element $G \in GL(n, K)$ is an *integral form of F^k* if the following two conditions hold. Let $I \in GL(n, K)$ be the identity element. Then

(i) the reversed characteristic polynomial $\det(I - GT)$ has integer coefficients;

(ii) for some integer $d \geq 1$, we have $\det(I - G^d T) = \det(I - F^{dk} T)$.

COROLLARY 2.3. *Let K be an algebraically closed field of characteristic zero, $n \geq 1$ an integer, and $F \in GL(n,K)$ an $n \times n$ invertible matrix whose reversed characteristic polynomial* $\det(I-FT)$ *has integer coefficients. Suppose that no eigenvalue of F is a root of unity. Then for any integer $M \geq 1$, there exists an integer $k_M \geq 1$ such that for $k > k_M$, and for any integral form G of F^k, either* $\mathrm{Trace}(G) = 0$ *or* $|\mathrm{Trace}(G)| > M$.

Proof. Denote by α_i the eigenvalues of F. An integral form G of F^k has eigenvalues $\zeta_i \alpha_i^k$, for some choice of roots of unity ζ_i. We claim that given F, there is an integer $D \geq 1$ such for any $k \geq 1$ and any integral form G of F^k, the possible ζ_i are all D-th roots of unity. Granting this claim, we get the result by applying Theorem 2.1 (iv), to the α_i and to each of the D^n n-tuples $(\lambda_1, \ldots, \lambda_n)$ with λ_i a D-th root of unity.

To prove the claim, we argue as follows. Since $\det(I - FT)$ has integer coefficients, the α_i are algebraic numbers, so lie in some finite Galois extension $K_0/\mathbf{Q}$. If we pick a prime p which splits completely in K_0, and a prime $\mathcal{P}$ of K_0 lying over p, then the $\mathcal{P}$-adic completion of K_0 is just the p-adic field $\mathbf{Q}_p$. So we can view all the α_i as lying in the p-adic field $\mathbf{Q}_p$. The fact that $\det(I - GT)$ has integer coefficients shows that each product $\zeta_i \alpha_i^k$ is algebraic of degree at most n over $\mathbf{Q}$, and hence of degree at most n over $\mathbf{Q}_p$. On the other hand, $\alpha_i^k \in \mathbf{Q}_p$, so ζ_i lies in an extension of $\mathbf{Q}_p$ of degree at most n. Since $\mathbf{Q}_p$ has only finitely many extensions of given degree, the ζ_i lie in a single finite extension, say $E_{\mathcal{P}}$, of $\mathbf{Q}_p$, and any such finite extension contains only finitely many roots of unity.

We now give some applications to varieties over finite fields, and to isotrivial[2]) families of such varieties. All of these applications have a common structure, that there is only one cohomology group we do not know in advance. Let us explain in a bit more detail. To begin with, suppose we are given a proper, smooth, geometrically connected variety X over a finite field $\mathbf{F}_q$ of characteristic $p > 0$. We choose a prime number $\ell \neq p$. Then we have Grothendieck's ℓ-adic étale cohomology groups[3]) $H^i_{\acute{e}t}(X \otimes_{\mathbf{F}_q} \overline{\mathbf{F}}_q, \overline{\mathbf{Q}}_\ell)$. In order to simplify notation, we shall write here $\overline{X}$ for $X \otimes_{\mathbf{F}_q} \overline{\mathbf{F}}_q$ (thus $\overline{X}$ is X after base change from $\mathbf{F}_q$ to $\overline{\mathbf{F}}_q$) and we shall write H^i for $H^i_{\acute{e}t}(\overline{X}, \overline{\mathbf{Q}}_\ell)$ if X is clear from the context.

---

[2]) A family $X \to S$ is *isotrivial* if it becomes a product $S' \times Y \to S'$ with trivial projection on the first factor, after a suitable finite étale base extension $S' \to S$.

[3]) In this paper, if K is a field, we denote by $\overline{K}$ a choice of an algebraic closure of K.

These cohomology groups are finite-dimensional $\overline{\mathbf{Q}}_\ell$-vector spaces, which vanish for i outside the interval $[0, 2\dim(X)]$. On each group H^i we have the Frobenius automorphism $\mathrm{Frob}_{\mathbf{F}_q}$, and according to the Lefschetz Fixed Point Formula [Gr] the number $\#X(\mathbf{F}_{q^n})$ of points of X defined over the field $\mathbf{F}_{q^n}$ is given, for each integer $n \geq 1$, by the formula

$$\#X(\mathbf{F}_{q^n}) = \sum_i (-1)^i \mathrm{Trace}(\mathrm{Frob}_{\mathbf{F}_q}^n | H^i).$$

This appears at first sight to be an equality of an integer $\#X(\mathbf{F}_{q^n})$ with an alternating sum of terms $\mathrm{Trace}(\mathrm{Frob}_{\mathbf{F}_q}^n | H^i)$ on the right, each of which is *a priori* only an element of $\overline{\mathbf{Q}}_\ell$. However, Deligne, [De2] proved that each individual trace term $\mathrm{Trace}(\mathrm{Frob}_{\mathbf{F}_q}^n | H^i)$ on the right is itself an integer, and moreover that this integer is independent of the auxiliary choice of the prime number $\ell \neq p$. Equivalently, for each i the reversed characteristic polynomial $\det(I - T\,\mathrm{Frob}_{\mathbf{F}_q} | H^i)$ is independent of $\ell \neq p$ and has integer coefficients. Moreover, he proved in the same paper that each eigenvalue of $\mathrm{Frob}_{\mathbf{F}_q}$ on H^i has complex absolute value $q^{i/2}$. See the review [Ka94] for a slight elaboration of this summary; for the purpose of this paper, it suffices to know only that such a cohomology theory exists and that it has the above properties.

All this becomes much more concrete and explicit in a diophantine setting when our variety X is either a curve or a complete intersection, because for such an X, say of dimension d, there is only one of its cohomology groups, namely the middle dimensional group H^d, which is difficult to understand completely. More precisely, for $0 \leq i \leq 2d$ and $i \neq d$, we have by [DK], XI, 1.6,

(i) if i is odd, then $H^i = 0$,

(ii) if i is even, say $i = 2r$, then $\dim(H^{2r}) = 1$, and $\mathrm{Frob}_{\mathbf{F}_q}$ acts on it by multiplication by q^r.

Thus if X is a (proper, smooth, geometrically connected) curve, we have

$$\#X(\mathbf{F}_{q^n}) = 1 + q^n - \mathrm{Trace}(\mathrm{Frob}_{\mathbf{F}_q}^n | H^1).$$

If X is a (proper, smooth, geometrically connected) complete intersection of odd dimension d, we have

$$\#X(\mathbf{F}_{q^n}) = 1 + q^n + q^{2n} + \cdots + q^{dn} - \mathrm{Trace}(\mathrm{Frob}_{\mathbf{F}_q}^n | H^d).$$

Now if X is a complete intersection of even dimension d, then H^d contains a one-dimensional subspace which is $\mathrm{Frob}_{\mathbf{F}_q}$-stable and on which $\mathrm{Frob}_{\mathbf{F}_q}$ acts with eigenvalue $q^{d/2}$ ([DK], XI, 1.6(iv)); the quotient of H^d by this

one-dimensional subspace is denoted by Prim^d. So here we have the formula

$$\#X(\mathbf{F}_{q^n}) = 1 + q^n + q^{2n} + \cdots + q^{dn} + (-1)^d \,\mathrm{Trace}(\mathrm{Frob}_{\mathbf{F}_q}^n | \mathrm{Prim}^d)\,.$$

We unify these last two formulas by defining $\mathrm{Prim}^d := H^d$ in the case when d is odd; then the last formula is valid when our X is a complete intersection of any dimension d.

The above definition of Prim^d has been somewhat simplified here for our purposes, so it is worthwhile to spend a few words to introduce the general notion of *the primitive part of the cohomology*, which is quite interesting and important in the study of projective varieties. This has no role for the limited results in our paper and the reader may skip the more technical definitions which follow.

Let X be a proper, smooth, geometrically connected variety of dimension d over a finite field $\mathbf{F}_q$ of characteristic $p > 0$ and let Y be a smooth hyperplane section[4]) of X. Let ℓ be a prime number $\ell \neq p$. The embedding $Y \hookrightarrow X$ induces restriction homomorphisms

$$H^i_{\acute{e}t}(\overline{X}, \overline{\mathbf{Q}}_\ell) \to H^i_{\acute{e}t}(\overline{Y}, \overline{\mathbf{Q}}_\ell)\,.$$

Using Poincaré duality, we get dual homomorphisms

$$L\colon H^i_{\acute{e}t}(\overline{Y}, \overline{\mathbf{Q}}_\ell) \to H^{i+2}_{\acute{e}t}(\overline{X}, \overline{\mathbf{Q}}_\ell(1))\,,$$

called the *Gysin homomorphisms*, where $\overline{\mathbf{Q}}_\ell(k)$ denotes the k-th Tate twist[5]). The image $\eta \in H^2_{\acute{e}t}(\overline{X}, \overline{\mathbf{Q}}_\ell(1))$ obtained applying the Gysin map to the class

[4]) The alert reader may correctly object at this point that over the given ground field $\mathbf{F}_q$, every hyperplane section might be singular. For instance, this is the case if $d = 2n$ is even and X is the smooth hypersurface in $\mathbf{P}^{2n+1}$ given by the equation in homogeneous coordinates $\sum_{i=0}^{n}(x_{2i}^q x_{2i+1} - x_{2i} x_{2i+1}^q) = 0$, see [Ka99], Question 10, pp. 621–622. One way around this difficulty is to use the fact that over every finite extension of sufficiently large degree of our ground field there do exist smooth hyperplane sections. Indeed, the singular hyperplane sections are a proper closed subscheme (the *dual variety*) $X^\vee$ of the projective space $\mathbf{P}^\vee$ of all hyperplane sections, cf. [DK], Exp. XVII, 3.1.4; hence the complement $\mathbf{P}^\vee \setminus X^\vee$ (the variety of smooth hyperplane sections) is not empty, smooth and geometrically connected, so has points in all finite extensions of large enough degree, by a well-known result of Lang and Weil [LW]. However, there is a more elegant geometric approach to the question. Poonen [P] has shown that for a given X as above, there exist smooth degree D hypersurface sections over the given ground field if D is large enough. (See also Gabber [Ga] for an independent proof if in addition D is divisible by the characteristic p.) Using these results, we can proceed in either of two ways. Suppose we are given a smooth hypersurface section Y of X of some degree D. We can use the *D-fold Veronese embedding* (via all monomials of degree D) to get a new projective embedding of X in which the previous degree D hypersurface sections now become hyperplane sections; for this projective embedding, there do exist smooth hyperplane sections over $\mathbf{F}_q$. Alternatively, in the arguments which follow we can use the $\overline{\mathbf{Q}}_\ell$-cohomology class $(1/D)1_Y$, instead of the class 1_Y of Y itself, to obtain the desired $\overline{\mathbf{Q}}_\ell$-cohomology class of a hyperplane section.

[5]) The *k-th Tate twist* $\overline{\mathbf{Q}}_\ell(k)$ is a certain one-dimensional Galois module over $\overline{\mathbf{Q}}_\ell$ for the action of the absolute Galois group of $\mathbf{F}_p$. The effect on the eigenvalues of the action of $\mathrm{Frob}_{\mathbf{F}_q}$ due to the twist is to multiply the eigenvalues by q^{-k}.

$\mathbf{1}_Y \in H^0_{\acute{e}t}(\overline{Y}, \overline{\mathbf{Q}}_\ell)$ corresponding to Y (the so-called *fundamental class of* Y) is the class of a hyperplane section of $\overline{X}$.

A fundamental theorem (the *Hard Lefschetz Theorem*) which goes back to Lefschetz for varieties over the complex field and classical "Betti" cohomology with coefficients in $\mathbf{C}$ and proved by Deligne [De3], 4.1.1, for ℓ-adic cohomology (hence applicable in our setting) is:

Let X be a proper, smooth, geometrically connected, projective variety of dimension d, over a finite field field $\mathbf{F}_q$ of characteristic $p > 0$. Let $\ell \neq p$ and let η be the hyperplane class in $H^2_{\acute{e}t}(\overline{X}, \overline{\mathbf{Q}}_\ell(1))$. Then the homomorphism

$$\eta^k \colon H^{d-k}_{\acute{e}t}(\overline{X}, \overline{\mathbf{Q}}_\ell) \to H^{d+k}_{\acute{e}t}(\overline{X}, \overline{\mathbf{Q}}_\ell(k))$$

given by cup-product with the class η^k is an isomorphism.

In particular, the eigenvalues of $\mathrm{Frob}_{\mathbf{F}_q}$ on $H^{d+k}_{\acute{e}t}(\overline{X}, \overline{\mathbf{Q}}_\ell)$ are equal to q^k times the eigenvalues of $\mathrm{Frob}_{\mathbf{F}_q}$ on $H^{d-k}_{\acute{e}t}(\overline{X}, \overline{\mathbf{Q}}_\ell)$. On the other hand, multiplying by η once more, the map

$$\eta^{k+1} \colon H^{d-k}_{\acute{e}t}(\overline{X}, \overline{\mathbf{Q}}_\ell) \to H^{d+k+2}_{\acute{e}t}(\overline{X}, \overline{\mathbf{Q}}_\ell(k+1))$$

may have a non-trivial kernel. This kernel

$$\mathrm{Prim}^{d-k}(\overline{X}, \overline{\mathbf{Q}}_\ell) := \ker(\eta^{k+1} | H^{d-k}_{\acute{e}t}(\overline{X}, \overline{\mathbf{Q}}_\ell))$$

is the *primitive part* of the cohomology group $H^{d-k}_{\acute{e}t}(\overline{X}, \overline{\mathbf{Q}}_\ell)$. This subspace

$$\mathrm{Prim}^{d-k}(\overline{X}, \overline{\mathbf{Q}}_\ell) \subset H^{d-k}_{\acute{e}t}(\overline{X}, \overline{\mathbf{Q}}_\ell)$$

is stable by the action of $\mathrm{Frob}_{\mathbf{F}_q}$, whose eigenvalues on $\mathrm{Prim}^{d-k}(\overline{X}, \overline{\mathbf{Q}}_\ell)$ are hence among its eigenvalues on $H^{d-k}_{\acute{e}t}(\overline{X}, \overline{\mathbf{Q}}_\ell)$. The remaining eigenvalues of $\mathrm{Frob}_{\mathbf{F}_q}$ on $H^{d-k}_{\acute{e}t}(\overline{X}, \overline{\mathbf{Q}}_\ell)$ can be recovered from its eigenvalues on $H^{d-k-2}_{\acute{e}t}(\overline{X}, \overline{\mathbf{Q}}_\ell)$; they are equal to q times the eigenvalues of $\mathrm{Frob}_{\mathbf{F}_q}$ acting on $H^{d-k-2}_{\acute{e}t}(\overline{X}, \overline{\mathbf{Q}}_\ell)$.

This shows the importance of the primitive part of the cohomology: its knowledge is sufficient, *via* the Hard Lefschetz Theorem, to compute the eigenvalues of the action of $\mathrm{Frob}_{\mathbf{F}_q}$ on the whole ℓ-adic cohomology of X. Moreover, by the *Weak Lefschetz Theorem*, cf. [De3], 4.1.6, the cohomology groups H^i of a smooth projective variety X of dimension d are isomorphic, for $i \leq d-2$, to the cohomology groups of any smooth hyperplane section. The cohomology group H^{d-1} of X can be recovered as a suitable "gcd" of the

groups H^{d-1} of "all" smooth hyperplane sections[6]) of $\bar{X}$. The cohomology groups for $i > d$ can of course be recovered by Poincaré duality from those with $i < d$. Inductively, this leaves only the middle dimensional cohomology H^d of X to be computed. The interest reader who wants a quick introduction to this deep theory may consult Danilov's article [Dan], §7 and §8.

Going back to curves or complete intersections, there is a single cohomology group, H^1 or Prim^d respectively, which we do not know explicitly. It is with the traces of iterates of Frobenius on this single unknown group that we will now be concerned. These traces are, as noted above, integers, and we will want to know cases when they are all not zero. One way to insure their being not equal to zero is to know that they are not zero modulo p, the characteristic of the finite field $\mathbf{F}_q$ over which we are working. For this, we can make use of the following congruence formula in [DK], XXII, 3.1. For any proper $X/\mathbf{F}_q$, we have its coherent cohomology groups $H^i(X, \mathcal{O}_X)$, on which the q-th power map Fr_q induces an $\mathbf{F}_q$-linear endomorphism. Then we have an identity in $\mathbf{F}_q$,

$$\#X(\mathbf{F}_q) \pmod p = \sum_i (-1)^i \mathrm{Trace}(Fr_q | H^i(X, \mathcal{O}_X)) .$$

In the case when our $X/\mathbf{F}_q$ is either a curve or a complete intersection of dimension $d > 0$ which is proper, smooth, and geometrically connected, we have

(i) $H^0(X, \mathcal{O}_X) = \mathbf{F}_q$, with $Fr_q = id$;

(ii) for $i \neq 0$ or d, we have $H^i(X, \mathcal{O}_X) = 0$.

So when our X is a curve, we get

$$\#X(\mathbf{F}_q) \pmod p = 1 - \mathrm{Trace}(Fr_q | H^1(X, \mathcal{O}_X))$$

and when our X is a complete intersection of dimension $d > 0$, we get

$$\#X(\mathbf{F}_q) \pmod p = 1 + (-1)^d \mathrm{Trace}(Fr_q | H^d(X, \mathcal{O}_X)) .$$

If we compare the Lefschetz Fixed Point Formula with the congruence formula, we get mod p congruences, namely: when our (proper, smooth, geometrically connected) X is a curve,

$$\mathrm{Trace}(\mathrm{Frob}_{\mathbf{F}_q} | H^1) \equiv \mathrm{Trace}(Fr_q | H^1(X, \mathcal{O}_X)) \pmod p$$

and when our (proper, smooth, geometrically connected) X is a complete intersection of dimension $d > 0$,

$$\mathrm{Trace}(\mathrm{Frob}_{\mathbf{F}_q} | \mathrm{Prim}^d) \equiv \mathrm{Trace}(Fr_q | H^d(X, \mathcal{O}_X)) \pmod p .$$

---

[6]) One needs the consideration of a *Lefschetz pencil* of hyperplane sections and delicate monodromy calculations, see [De3], 4.5.1.

With this background established, we now return to giving applications of our previous results to varieties over finite fields. We begin with the case of curves over finite fields.

THEOREM 2.4. *Let $X/\mathbf{F}_q$ be a proper, smooth, geometrically connected curve over a finite field $\mathbf{F}_q$ of characteristic $p > 0$. Define a sequence of integers $A(n), n \geq 1$ by*

$$\#X(\mathbf{F}_{q^n}) = q^n + 1 - A(n).$$

Then the non-zero $A(n)$ satisfy $|A(n)| \to \infty$.

Proof. This follows from Corollary 2.3 above, applied with K taken to be $\overline{\mathbf{Q}}_\ell$ for some $\ell \neq p$ and with F taken to be the action of the geometric Frobenius $\mathrm{Frob}_{\mathbf{F}_q}$ on $H^1_{\acute{e}t}(\overline{X}, \overline{\mathbf{Q}}_\ell)$. By the Lefschetz Fixed Point Formula [Gr], we have $A(n) = \mathrm{Trace}(F^n)$. By Weil's Riemann hypothesis for curves over finite fields [W1], p. 70, the eigenvalues of F all have archimedean absolute value $q^{1/2}$, so are not roots of unity.

COROLLARY 2.5. *Let $X/\mathbf{F}_q$ be a proper, smooth, geometrically connected curve over a finite field $\mathbf{F}_q$ of characteristic $p > 0$. Suppose that one of the following three conditions holds.*

(i) *The genus g is 1 and $X/\mathbf{F}_q$ is ordinary*[7]).

(ii) *The genus g of X is prime to p, and the q-th power map on $H^1(X, \mathcal{O}_X)$ is the identity (i.e., the Hasse-Witt matrix*[8]*) relative to $\mathbf{F}_q$ is the identity $g \times g$ matrix over $\mathbf{F}_q$), or, equivalently, the group of p-torsion rational points of the Jacobian $Jac(X)(\mathbf{F}_q)$ has order p^g.*

(iii) *For some integer $N \geq 1$ which is prime to p and modulo which $2g$ is not zero, there are N^{2g} points of order dividing N in $Jac(X)(\mathbf{F}_q)$.*

Then for all $n \geq 1$, we have $A(n) \neq 0$, hence $|A(n)| \to \infty$.

Proof. In case (i), each $A(n)$, $n \geq 1$, is prime to p, so is not zero. In case (ii), the congruence formula [DK], XXII, 3.1, shows that for $n \geq 1$, we have $A(n) \equiv g \pmod{p}$, so again $A(n) \neq 0$. In case (iii), we have $A(n) \equiv 2g \pmod{N}$ for all $n \geq 1$, so again $A(n) \neq 0$.

---

[7]) An elliptic curve over a finite field $\mathbf{F}_q$ of characteristic p is *ordinary* if its group of p-division points has order p. In the only other possible case, namely order 1, the curve is called *supersingular*.

[8]) The *Hasse-Witt matrix* is obtained by looking at the action of the p-power map on a basis of $H^1(X, \mathcal{O}_X)$ and is explicitly computable. For a curve of genus 1 it reduces to a single element in $\mathbf{F}_q$, the *Hasse invariant*.

We get similar results for complete intersections over finite fields.

THEOREM 2.6. *Let $X/\mathbf{F}_q$ be a proper, smooth, geometrically connected complete intersection of dimension $d \geq 1$ over a finite field $\mathbf{F}_q$ of characteristic $p > 0$. Define a sequence of integers $A(n), n \geq 1$ by*

$$\#X(\mathbf{F}_{q^n}) = \sum_{i=0}^{d} q^{ni} + (-1)^d A(n).$$

Then the non-zero $A(n)$ have $|A(n)| \to \infty$.

Proof. This again follows from Corollary 2.5 above, applied with K taken to be $\overline{\mathbf{Q}}_\ell$ for some $\ell \neq p$ and with F taken to be the action of the geometric Frobenius $\mathrm{Frob}_{\mathbf{F}_q}$ on $\mathrm{Prim}^d_{\acute{e}t}(X \otimes_{\mathbf{F}_q} \overline{\mathbf{F}}_q, \overline{\mathbf{Q}}_\ell)$ (the "primitive part" $\mathrm{Prim}^d_{\acute{e}t}$ of the cohomology $H^d_{\acute{e}t}$ of a smooth complete intersection X is simply $H^d_{\acute{e}t}$ if d is odd and, if d is even, it is $H^d_{\acute{e}t}$ of X modulo the image of $H^d_{\acute{e}t}$ of the ambient projective space, see [DK], XI, 1.6(iv)). By the Lefschetz Fixed Point Formula [Gr] and the known cohomological structure of complete intersections [DK], XI, 1.6, we have $A(n) = \mathrm{Trace}(F^n)$. By Deligne's Riemann hypothesis for varieties over finite fields [De2], the eigenvalues of F have archimedean absolute value $q^{d/2}$, so are not roots of unity.

COROLLARY 2.7. *Let $X/\mathbf{F}_q$ be a proper, smooth, geometrically connected complete intersection of dimension $d \geq 1$ over a finite field $\mathbf{F}_q$ of characteristic $p > 0$. Suppose that $g := \dim(H^d(X, \mathcal{O}_X))$ is prime to p, and that the q-th power map on $H^d(X, \mathcal{O}_X)$ is the identity. Then for all $n \geq 1$, we have $A(n) \neq 0$, hence $|A(n)| \to \infty$.*

Proof. Again by the congruence formula [DK], XXII, 3.1, for $n \geq 1$ we have $A(n) \equiv g \pmod{p}$, so again $A(n) \neq 0$.

Here is a variant of the last result, when the geometric genus is 1.

COROLLARY 2.8. *Let $X/\mathbf{F}_q$ be a proper, smooth, geometrically connected complete intersection of dimension $d \geq 1$ over a finite field $\mathbf{F}_q$ of characteristic $p > 0$. Suppose that $\dim(H^d(X, \mathcal{O}_X)) = 1$, and that the q-th power map on $H^d(X, \mathcal{O}_X)$ is not zero, say is multiplication by $a \in \mathbf{F}_q^\times$. Then, for all $n \geq 1$, $A(n)$ is prime to p, so it is not zero, hence $|A(n)| \to \infty$.*

Proof. Again by the congruence formula [DK], XXII, 3.1, for $n \geq 1$, we have $A(n) \equiv a^n \pmod{p}$, hence for all n we have $A(n) \neq 0$.

We now turn to isotrivial families, and apply Corollary 2.8 above.

THEOREM 2.9. *Let $\mathbf{F}_q$ be a finite field of characteristic $p > 0$, let $S/\mathbf{F}_q$ be a smooth, geometrically connected $\mathbf{F}_q$-scheme of finite type with $S(\mathbf{F}_q)$ nonempty, and let $\pi: X \to S$ be a proper smooth morphism of relative dimension $d \geq 1$, all of whose geometric fibres are curves or, if $d \geq 2$, complete intersections. Suppose the morphism π is isotrivial, in the sense that when pulled back to a suitable finite étale S-scheme T/S it becomes constant. For each closed point $\mathcal{P}$ of S, with residue field denoted $\mathbf{F}_\mathcal{P}$, consider the fibre $X_{\mathbf{F}_\mathcal{P}} := X \otimes_{\mathcal{O}_S} \mathbf{F}_\mathcal{P}$ and define the integer $A_\mathcal{P}$ by*

$$\#X_{\mathbf{F}_\mathcal{P}}(\mathbf{F}_\mathcal{P}) = \sum_{i=0}^{d} \mathrm{Norm}(\mathcal{P})^i + (-1)^d A_\mathcal{P}\,.$$

Then the non-zero $A_\mathcal{P}$ have $|A_\mathcal{P}| \to \infty$ as $\deg(\mathcal{P}) \to \infty$. More precisely, for any integer $M \geq 1$, there exists an integer $k_M \geq 1$ such that for any $k > k_M$, and for any closed point $\mathcal{P}$ with $\deg(\mathcal{P}) = k$, either $A_\mathcal{P} = 0$ or $|A_\mathcal{P}| > M$.

Proof. We choose a point $s_0 \in S(\mathbf{F}_q)$, and denote by $X_0/\mathbf{F}_q$ the fibre of X/S over s_0. We choose a prime $\ell \neq p$, and take for F the action of geometric $\mathrm{Frob}_{\mathbf{F}_q}$ on $\mathrm{Prim}^d(\overline{X_0}, \overline{\mathbf{Q}}_\ell)$. By the isotriviality of X/S, for any closed point $\mathcal{P}$ of S, the fibre $X_{\mathbf{F}_\mathcal{P}}$ becomes isomorphic to $X_0 \otimes \mathbf{F}_\mathcal{P}$ after extension of scalars to some finite extension of $\mathbf{F}_\mathcal{P}$. Therefore the geometric Frobenius $\mathrm{Frob}_\mathcal{P}$ acting on $\mathrm{Prim}^d(\overline{X_{\mathbf{F}_\mathcal{P}}}, \overline{\mathbf{Q}}_\ell)$ is an integral form of $F^{\deg(\mathcal{P})}$. So the assertion results from Corollary 2.8 above.

COROLLARY 2.10. *If X/S as above is an isotrivial family of elliptic curves which are ordinary, i.e., if the constant j-invariant is ordinary, then all $A_\mathcal{P}$ are not zero (because prime to p), hence $|A_\mathcal{P}| \to \infty$ as $\deg(\mathcal{P}) \to \infty$.*

3. LOWER BOUNDS, VIA THE SUBSPACE THEOREM

Fix an integer $Q > 1$. In practice, Q will be a prime power p^m, but right now that is not important. An algebraic number $\alpha \in \overline{\mathbf{Q}}$ is called a *Q-Weil number* if, for every embedding $\iota: \overline{\mathbf{Q}} \subset \mathbf{C}$, we have $|\iota(\alpha)|_\mathbf{C} = Q^{1/2}$, for $|\ |_\mathbf{C}$ the usual complex absolute value $|x+iy|_\mathbf{C} := (x^2+y^2)^{1/2}$. A Q-Weil number is called *integral* if in addition it is an algebraic integer.

Lower bounds come from the following special case of a theorem of Evertse [Ev], Cor. 2, also due independently to van der Poorten and Schlickewei [PS], Theorem 3.

THEOREM 3.1. *Let $Q > 1$ and $n \geq 1$ be integers. Let $\alpha_1, \ldots, \alpha_n$ be integral Q-Weil numbers. For each integer $k \geq 1$, define*

$$A(k) := \sum_{i=1}^{n} \alpha_i^k.$$

Given a real number $\varepsilon > 0$, there exists a real constant $C_1 > 0$ such that for any integer $k \geq 1$, either $A(k) = 0$ or, for any archimedean absolute value on $\overline{\mathbf{Q}}$, we have

$$|A(k)| \geq C_1 Q^{k(1-\varepsilon)}.$$

Proof. This is the following special case of [Ev], Cor. 2. Take for K a number field containing all the α_i. Take for S the set of all places of K which are either archimedean or which lie over primes dividing Q. Take for $T \subset S$ a single archimedean place. Since the absolute norm of every α_i is a power of Q, the algebraic integers α_i are all S-units.

Then, for each integer $k \geq 1$ with $A(k) = 0$, simply apply [Ev], Cor. 2, to the S-units $x_i := \alpha_i^k$.

We can trivially make the constant C_1 disappear if we insist that k be sufficiently large.

COROLLARY 3.2. *Under the hypotheses of the theorem, given a real number $\varepsilon > 0$, there exists an integer k_0 such that for all integers $k \geq k_0$, either $A(k) = 0$ or, for any archimedean absolute value on $\overline{\mathbf{Q}}$, we have*

$$|A(k)| \geq Q^{k(1-2\varepsilon)}.$$

THEOREM 3.3. *Let $X/\mathbf{F}_q$ be a proper smooth variety over $\mathbf{F}_q$. Fix an integer $i \geq 1$, and a prime $\ell \neq p$. Consider the sequence of integers $A_i(n)$, $n \geq 1$, (independent of the auxiliary choice of ℓ, cf.* [De3], 3.3.9*) defined as*

$$A_i(n) := \mathrm{Trace}(\mathrm{Frob}_{\mathbf{F}_q}^n | H^i_{\acute{e}t}(\overline{X}, \overline{\mathbf{Q}}_\ell)).$$

Fix a real number $\varepsilon > 0$. Then for all sufficiently large n, either $A_i(n) = 0$ or

$$|A_i(n)| \geq \left(q^{in/2}\right)^{1-\varepsilon}.$$

Proof. This is an immediate consequence of Deligne's theorem [De3], 3.3.9, by applying Theorem 3.1 and Corollary 3.2 to the eigenvalues of $\mathrm{Frob}_{\mathbf{F}_q}$ on H^i, which are integral q^i-Weil numbers.

We now turn to the situation with pure exponential sums. In nearly all examples, the situation is the following, which we describe first in technical terms, followed by simple explicit examples understandable by non-experts.

We are given an affine, smooth, geometrically connected variety $U/\mathbf{F}_q$ of some dimension $d \geq 1$, a prime number $\ell \neq p$, and a lisse $\overline{\mathbf{Q}}_\ell$-sheaf $\mathcal{F}$ on U which is *integral* (all local Frobeniuses have algebraic integer eigenvalues) and pure of some integer weight $w_0 \geq 0$. We have somehow proven that for all i, the "forget supports" map[9])

$$H^i_c(\overline{U}, \mathcal{F}) \to H^i(\overline{U}, \mathcal{F})$$

is an isomorphism. It then follows, cf. [De3], 3.3.6, and [Se], that $H^i_c = 0$ for $i \neq d$ and that, putting

$$w := d + w_0\,,$$

the Frobenius eigenvalues on H^d_c are integral $q^{w/2}$-Weil numbers. The sequence of algebraic integers

$$A(n) := \mathrm{Trace}(\mathrm{Frob}^n_{\mathbf{F}_q} | H^d_c(\overline{U}, \mathcal{F}))$$

is the sequence of exponential sums, over bigger and bigger finite extensions of $\mathbf{F}_q$, that we are interested in.

So in any such situation, Theorem 3.3 assures us that for any chosen embedding ι of the number field $\mathbf{Q}(\{\textit{eigenvalues of } \mathrm{Frob}_{\mathbf{F}_q}\})$ into $\mathbf{C}$, and any chosen real number $\varepsilon > 0$, we have that for all n sufficiently large either $A(n) = 0$ or $|\iota A(n)|_{\mathbf{C}} \geq (q^{nw/2})^{1-\varepsilon}$.

It is consequently of some interest to know in what situations of this type we know in addition that $A(n) \neq 0$ for n large. Here are three such situations which occur in practice, where in fact $A(n) \neq 0$ for all $n \geq 1$.

(i) The d *variable Kloosterman sums* $Kl_d(\psi, a, \mathbf{F}_q)$, for $d \geq 2$, ψ a nontrivial additive character of $\mathbf{F}_q$, and $a \in \mathbf{F}_q^\times$, defined by

$$(-1)^{d-1} Kl_d(\psi, a, \mathbf{F}_q) := \sum_{x_1 x_2 \ldots x_d = a,\ \textit{all}\ x_i \in \mathbf{F}_q} \psi(x_1 + \cdots + x_n)\,.$$

Only H^{d-1}_c is not zero, and the d Frobenius eigenvalues are integral q^{d-1}-Weil numbers [De1], 7.1.3, 7.4. This sum lies in $\mathbf{Z}[\zeta_p]$ and never vanishes, because modulo the unique prime ideal $\mathfrak{p}$ of $\mathbf{Z}[\zeta_p]$ lying over p we have

$$(-1)^{d-1} Kl_d(a, \mathbf{F}_q) \equiv (q-1)^{d-1} \equiv (-1)^{d-1} \pmod{\mathfrak{p}}$$

[9]) Here H^*_c denotes *cohomology with compact support.*

(simply because ψ is trivial $(\mathrm{mod}\ \mathfrak{p})$). Here the sequence of $A(n)$ is

$$A(n) = Kl_d(\psi \circ \mathrm{Trace}_{\mathbf{F}_{q^n}/\mathbf{F}_q}, a, \mathbf{F}_{q^n}).$$

Therefore, for any given real $\varepsilon > 0$ we have the lower bound

$$|Kl_d(\psi \circ \mathrm{Trace}_{\mathbf{F}_{q^n}/\mathbf{F}_q}, a, \mathbf{F}_{q^n})| \geq \left(q^{n(d-1)/2}\right)^{1-\varepsilon}$$

for all n sufficiently large.

(ii) Start with the projective line $\mathbf{P}^1/\mathbf{F}_q$ and remove a nonempty set S of $\mathbf{F}_q$-rational points, with $\#S - 1$ invertible $(\mathrm{mod}\ p)$. We take $U := \mathbf{P}^1 \setminus S$. On U, we take a regular function $f \in H^0(U, \mathcal{O}_U)$ whose pole orders e_s at the points $s \in S$ are all prime to p. For ψ a nontrivial additive character of $\mathbf{F}_q$, we have the sum

$$S(\psi, f, \mathbf{F}_q) := -\sum_{u \in U(\mathbf{F}_q)} \psi(f(u)).$$

Only the first cohomology group with compact support H^1_c is not zero, and the $\#S - 2 + \sum_{s \in S} e_s$ Frobenius eigenvalues are integral q-Weil numbers [W2]. This sum lies in $\mathbf{Z}[\zeta_p]$ and never vanishes, because modulo the unique prime ideal $\mathfrak{p}$ of $\mathbf{Z}[\zeta_p]$ lying over p, it is congruent to $-(q + 1 - \#S) \equiv \#S - 1$, which by assumption is not zero mod p. The sequence $A(n)$ in this case is

$$A(n) = S(\psi \circ \mathrm{Trace}_{\mathbf{F}_{q^n}/\mathbf{F}_q}, f, \mathbf{F}_{q^n}).$$

Hence for any given real $\varepsilon > 0$ we have the lower bound

$$|S(\psi \circ \mathrm{Trace}_{\mathbf{F}_{q^n}/\mathbf{F}_q}, f, \mathbf{F}_{q^n})| \geq \left(q^{n/2}\right)^{1-\varepsilon}$$

for all sufficiently large n.

(iii) Here we have a slight variant on example (ii) above. Take for U the affine line $\mathbf{A}^1/\mathbf{F}_q$ and $f \in \mathbf{F}_q[X]$ a polynomial of degree $d \geq 1$. Under the hypothesis that

$$p \equiv 1 \pmod{d},$$

Sperber [Sp], 3.11, shows that the $d - 1$ Frobenius eigenvalues on H^1_c have all distinct $\mathcal{P}$-adic valuations at any prime lying over p; their $\mathcal{P}$-adic orders, normalized so that q has $\mathrm{ord}_{\mathcal{P}}(q) = 1$, are $1/d, 2/d, \ldots, (d-1)/d$. Here the $A(n)$ are

$$A(n) = -S(\psi \circ \mathrm{Trace}_{\mathbf{F}_{q^n}/\mathbf{F}_q}, f, \mathbf{F}_{q^n}),$$

they never vanish, and we have the same conclusion as in (ii) above.

4. Effective lower bounds, via Baker's method

In some cases there are only two Frobenius eigenvalues, they are complex conjugates of each other, and their ratio is *not* a root of unity. These cases include an ordinary elliptic curve over $\mathbf{F}_q$, and also the classical Kloosterman sums, denoted $Kl_2(\psi, a, \mathbf{F}_q)$ in the previous section. In both of these cases, the two Frobenius eigenvalues are integral q-Weil numbers, say α and $\overline{\alpha}$, with $\alpha\overline{\alpha} = q$. After we fix a complex embedding, we can write the two eigenvalues as $q^{1/2}e^{\pm i\theta}$ for a unique $\theta \in [0, \pi]$. Then the $A(n)$ are given by

$$A(n) := \alpha^n + \overline{\alpha}^n = 2q^{n/2}\cos(n\theta).$$

Here is the key technical result, an immediate application of the deep Baker-Wüstholz theorem [BW]. For the definition of *height*, we refer to [BG], §1.5.

THEOREM 4.1. *Let $\theta \in [0, \pi]$. Suppose that $e^{2i\theta}$ is not a root of unity, but is an algebraic number, algebraic of degree d over $\mathbf{Q}$. Define*

$$\begin{aligned} C(N, d) &:= 18(N+1)!N^{N+1}(32d)^{N+2}\log(2Nd), \\ h'(e^{2i\theta}) &:= \max\bigl(\log(H((1 : e^{2i\theta}))), \theta/d, 1/d\bigr), \\ h'(-1) &:= \pi/d, \end{aligned}$$

where $H((x_0 : \ldots : x_r))$ is the Weil height of an (algebraic) point $(x_0 : \ldots : x_r)$ in projective space $\mathbf{P}^r$. Then for any integer $n \geq 1$ and any integer k we have the inequality

$$\log(|2n\theta - k\pi|) > -C(2, d)\,h'(e^{2i\theta})\,h'(-1)\log(2n).$$

Proof. Fix $n \geq 1$. Since $\theta \in [0, \pi]$, we have $2n\theta \in [0, 2n\pi]$. So the closest approach of $n\theta$ to an integer multiple of π occurs for some $k \in [0, 2n]$. (Indeed, for any integer k outside of this interval, we trivially have $|2n\theta - k\pi| \geq \pi$, and $\log\pi > 0$.) Because $e^{2i\theta}$ is not a root of unity, $\log(e^{2i\theta}) = 2i\theta$ and $\log(-1) = i\pi$ are linearly independent over $\mathbf{Q}$. Now apply the Baker-Wüstholz theorem, with the $N = 2$ algebraic numbers $e^{2i\theta}$ and -1, to the linear combination of logarithms $n\log(e^{2i\theta}) - k\log(-1)$.

COROLLARY 4.2. *Let $\theta \in [0,\pi]$ be as in the theorem. Given a real number $q > 1$, define*

$$c = c(\theta, q) := C(2,d)h'(e^{2i\theta})h'(-1)/\log(q).$$

Then for all integers $n \geq 1$, we have the estimate

$$|q^{n/2}\cos(n\theta)| \geq (1/\pi)\, q^{n/2 - c\log(2n)}.$$

Proof. Fix $n \geq 1$. By the theorem, for any integer k, we have the inequality

$$|2n\theta - k\pi| \geq q^{-c\log(2n)}.$$

For k an odd integer, we have the trigonometric identity $\cos(n\theta) = \pm\sin(n\theta - k\pi/2)$ and for the odd integer k_0 which minimizes $|n\theta - k\pi/2|$ we have

$$0 < |n\theta - k_0\pi/2| < \pi/2.$$

Also, for real x with $|x| \leq \pi/2$, we have the well-known inequality

$$|\sin(x)| \geq (2/\pi)|x|.$$

Thus we find

$$|\cos(n\theta)| = |\sin(n\theta - k_0\pi/2)| \geq (2/\pi)|n\theta - k_0\pi/2| \geq (1/\pi)q^{-c\log(2n)},$$

completing the proof.

Let us make this explicit in the two cases of ordinary elliptic curves and of classical Kloosterman sums.

COROLLARY 4.3. (i) *Given an ordinary elliptic curve over $\mathbf{F}_q$, the sequence of its $A(n)$ has, for all $n \geq 1$, the archimedean lower bound*

$$|A(n)| \geq (2/\pi)\, q^{n/2 - 2^{37}\log(2n)}.$$

(ii) *Given a classical Kloosterman sum $Kl_2(\psi, a, \mathbf{F}_q)$ over $\mathbf{F}_q$, denote by p the characteristic of $\mathbf{F}_q$. If $p = 2$ or $p = 3$, the sequence of its $A(n)$ has, for all $n \geq 1$, the same archimedean lower bound as for ordinary elliptic curves,*

$$|A(n)| \geq (2/\pi)\, q^{n/2 - 2^{37}\log(2n)}.$$

If $p \geq 5$, the sequence of its $A(n)$ has, for all $n \geq 1$, the archimedean lower bound

$$|A(n)| \geq (2/\pi)\, q^{n/2 - c_p\log(2n)},$$

with c_p the constant $c_p = 2^{33}p^4\log p$.

Proof. We will compute, in the two cases, an explicit upper bound for the constant c of the previous corollary.

Denote by α and $\overline{\alpha}$ the two Frobenius eigenvalues. After possibly interchanging them, we have $\alpha/\overline{\alpha} = e^{2i\theta}$. Thus

$$H((1 : e^{2i\theta})) = H((\alpha : \overline{\alpha})) \leq q^{1/2},$$

simply because α and $\overline{\alpha}$ are integral q-Weil numbers.

In the case of an ordinary elliptic curve, α and $\overline{\alpha}$ lie in a quadratic imaginary field, and their ratio is irrational, so we have $d = 2$ in this case. Then

$$\begin{aligned} h'(e^{2i\theta}) &:= \max\bigl(\log(H((1 : e^{2i\theta}))), \theta/d, 1/d\bigr) \\ &= \max(\log(q)/2, \pi/2, 1/2) \leq 5\log(q)/2; \end{aligned}$$

the factor 5 takes care of the worst case $q = 2$. So the constant c of the previous corollary is bounded by

$$c \leq C(2,2)(5/2)(\pi/2) = 18 \cdot 3! \cdot 2^3 \cdot (64)^4 \cdot \log(8) \cdot (5\pi/4) \leq 2^{37}.$$

In the case of a classical Kloosterman sum, the sum itself lies in $\mathbf{Q}(\zeta_p)^+$, the real subfield of $\mathbf{Q}(\zeta_p)$, and α and $\overline{\alpha}$ lie in a CM quadratic extension. Again their ratio is irrational (otherwise it would be a rational number of absolute value one, so ± 1), hence in this case we have $2 \leq d \leq \max(p-1, 2)$. So again we have

$$h'(e^{2i\theta}) \leq 5\log(q)/2.$$

For $p = 2$ and $p = 3$, we have $d = 2$, giving the bound

$$c \leq 2^{37}.$$

For $p \geq 5$ the bound becomes dependent on p, namely

$$\begin{aligned} c &\leq C(2, p-1)(5\pi/4) \\ &= 18 \cdot 3! \cdot 2^3 \cdot (32(p-1))^4 \cdot \log(4(p-1)) \cdot (5\pi/4) < 2^{33} p^4 \log p. \end{aligned}$$

This completes the proof.

5. Concluding remarks

As mentioned in the introduction, the main open problem here is obtaining effective lower bounds. On the other hand, much is known about the number of zeroes in a linear recurrence sequence. A theorem of Evertse, Schlickewei, and Schmidt [ESS] states the following.

Let K be a field of characteristic 0, let Γ be a subgroup of $(K^\times)^n$ of finite $\mathbf{Q}$-rank r, and let $a_1, \dots, a_n \in K^\times$. Let $\mathcal{X}$ be the set of those solutions $(x_1, \dots, x_n) \in \Gamma$ of the equation

$$a_1 x_1 + \cdots + a_n x_n = 1$$

for which no proper subsum of $a_1 x_1 + \cdots + a_n x_n$ vanishes. Then $\mathcal{X}$ is a finite set of cardinality

$$\#\mathcal{X} \le e^{(6n)^{3n}(r+1)}.$$

This can be applied easily to obtain further information on the set of zeroes of the sequences $A(n)$ examined here, since in this case we have $r = 1$. The Skolem-Mahler-Lech theorem shows that the zero set of the sequence $A(n)$ is the union of a finite set S_0 of isolated solutions and of finitely many arithmetic progressions. Theorem 1.2 of [ESS] immediately shows that

$$\#S_0 + \#(\textit{arithmetic progressions}) \le e^{2(12g)^{6g}}.$$

Although this is not directly relevant to the applications we have treated in this paper, a similar result also holds for any linear recurrence of order n (where the coefficients λ_i are allowed to be polynomials), with a bound $\exp\exp\exp(3n\log n)$ for the corresponding number of isolated solutions and of arithmetic progressions, see Schmidt [Sc].

The proof of these results is difficult and rather intricate, but it is a remarkable fact that these bounds depend only on n and the rank of Γ. It is an interesting problem to determine the correct rate of growth for the number of solutions of such equations.

For $n = 2$ and rank $r = 1$, J. Berstel provided the following example with 6 solutions. Consider the equation $ax^m + by^m = 1$ for fixed x, y, and varying $m \in \mathbf{Z}$, corresponding to the group $(x, y)^{\mathbf{Z}}$ of rank 1. We may assume that $m = 0$ is a solution. If $m = 1$ is also a solution, the equation becomes

$$\frac{y-1}{y-x}\,x^m + \frac{1-x}{y-x}\,y^m = 1;$$

we can exclude $x = 1$, $y = 1$, $x = y$ as degenerate cases. If now we fix two more values for m, say m_1 and m_2, we can eliminate y and obtain an algebraic equation for x, leading to infinitely many choices of the pair (x, y) for which there are four solutions. The choice $m_1 = 2$ leads to a degenerate case and if $m_1 = 3$ the values $m_2 = 4, 5, 6, 7, 9$ must be excluded, leading to degenerate cases or a group of rank 0. However, taking $m_1 = 4$ and $m_2 = 6$

gives the equation

$$x^6 + x^5 + 2x^4 + 3x^3 + 2x^2 + x + 1 = 0$$

for x. For any root ξ of this equation, we see that taking $\eta = -1/(1+\xi+\xi^3)$ (which is another root of the equation), we have

$$\frac{\eta - 1}{\eta - \xi}\xi^m + \frac{1 - \xi}{\eta - \xi}\eta^m = 1$$

for $m = 0, 1, 4, 6, 13, 52$. It is expected that 6 is the maximum number of solutions for an equation with $n = 2$ and $r = 1$; Beukers and Schlickewei [BS] obtained the upper bound 61.

For general n and r, Erdős, Stewart, and Tijdeman [EST] proved the existence of equations with $n = 2$ and arbitrarily large r with at least $\exp((4-\varepsilon)r^{1/2}(\log r)^{-1/2})$ solutions for any fixed $\varepsilon > 0$, and conjectured that if $n = 2$ the exponent $1/2$ could be improved to $2/3 - \varepsilon$ for any fixed positive ε (of course, allowing a constant depending on ε in place of 4); they also conjectured that the exponent $2/3$ should be sharp. Although this remains unsolved, progress was made by Konyagin and Soundararajan [KS], who constructed equations for the case $n = 2$ and arbitrarily large r with at least $\exp(r^{2-\sqrt{2}-\varepsilon})$ solutions, for any fixed $\varepsilon > 0$. For arbitrary n and r a lower bound $\exp((n^2(n-1)^{-1} - \varepsilon)r^{1-1/n}(\log r)^{-1/n})$ for the maximum number of solutions was provided by Evertse, Moore, Stewart, and Tijdeman [EMST]; this may be compared with the upper bound simply exponential in r provided by Evertse, Schlickewei and Schmidt, *loc. cit.*.

A more delicate problem has also been treated, namely the study of the intersection of two distinct recurrences and the "total multiplicity" of a recurrence, namely $A(m) = B(n)$ and $A(m) = A(n)$ for $m \neq n$. Under certain natural conditions one can prove that the number of admissible pairs (m, n) for which these equations hold is finite, see Evertse [Ev], Thm. 3, for the equation $A(m) = A(n)$ with recurrences of order at least 2 (this avoids the example $A(n) = n2^n$), and Laurent [Lau] for qualitative results for the equation $A(m) = B(n)$. Quantitative results, but not as strong as those mentioned above for the cardinality of the zero-set of a recurrence, can be found in Schlickewei and Schmidt [SS].

The reader interested in recurrence sequences and associated problems may profitably read the book [EvSW], which also contains an impressive bibliography of 1382 items on the subject.

The extension of these results to larger classes of polynomial-exponential equations in several variables remains a central and very challenging open problem. As an example, the famous Ramanujan equation $m^2 + 7 = 2^k$ has only the solutions $(m,k) = (1,3),(3,4),(5,5),(11,7),(181,15)$ in positive integers, which is not difficult to prove using Skolem's method. The modified equation $m^2 + 7^n = 2^k + (r-1)3^r$ associated to the group of rank 3

$$\{(1^m, 2^k, 3^r, 7^n)\}_{m,k,r,n \in \mathbf{Z}}$$

has, besides the five solutions with $n = 1$ and $r = 1$ inherited from the Ramanujan equation, seven new solutions $(m,k,r,n) = (2,1,2,1)$, $(7,1,3,1)$, $(14,1,4,2)$, $(3,2,3,2)$, $(13,9,1,3)$, $(113,11,7,4)$, $(407,13,9,1)$. Are there any other solutions in positive integers to this equation?

REFERENCES

[BW] BAKER, A. and G. WÜSTHOLZ. Logarithmic forms and group varieties. *J. Reine Angew. Math. 442* (1993), 19–62.

[BS] BEUKERS, F. and H. P. SCHLICKEWEI. The equation $x + y = 1$ in finitely generated groups. *Acta Arith. 78* (1996), 189–199.

[BG] BOMBIERI, E. and GUBLER, W. *Heights in Diophantine Geometry*. Paperback reprint of the 2006 original. Reprinted with corrections. New Mathematical Monographs *4*. Cambridge University Press, Cambridge, 2007.

[C1] CASSELS, J. W. S. An embedding theorem for fields. *Bull. Austral. Math. Soc. 14* (1976), 193–198; Addendum: "An embedding theorem for fields", *ibidem*, 479–480.

[C2] —— *Local Fields*. London Math. Soc. Student Texts *3*. Cambridge University Press, Cambridge, 1986.

[Dan] DANILOV, V. I. Cohomology of algebraic varieties. In: *Algebraic Geometry, II*, I. R. Shafarevich ed., 1–125, 255–262. Encyclopaedia Math. Sci. *35*. Springer, Berlin, 1996.

[De1] DELIGNE, P. Applications de la formule des traces aux sommes trigonométriques. In: *SGA 4 1/2*, Séminaire de Géométrie Algébrique du Bois Marie par P. Deligne, avec la collaboration de J. F. Boutot, A. Grothendieck, L. Illusie, et J. L. Verdier, 168–232. Lecture Notes in Mathematics *569*. Springer-Verlag, Berlin–New York, 1977.

[De2] —— La conjecture de Weil. I. *Publ. Math. Inst. Hautes Études Sci. 43* (1974), 273–307.

[De3] —— La conjecture de Weil. II. *Publ. Math. Inst. Hautes Études Sci. 52* (1980), 137–252.

[DK] *Groupes de monodromie en géométrie algébrique*. Séminaire de Géométrie Algébrique du Bois-Marie 1967–1969 (SGA 7 II). Dirigé par P. Deligne et N. Katz. Lecture Notes in Mathematics *340*. Springer-Verlag, Berlin–New York, 1973.

[Dw] DWORK, B. On the zeta function of a hypersurface. *Inst. Hautes Études Sci. Publ. Math. 12* (1962), 5–68.

[DGS] DWORK, B., G. GEROTTO and F. J. SULLIVAN. *An Introduction to G-functions*. Annals of Mathematics Studies *133*. Princeton University Press, Princeton, NJ, 1994.

[EST] ERDŐS P., C. L. STEWART and R. TIJDEMAN. Some Diophantine equations with many solutions. *Compositio Math. 66* (1988), 37–56.

[EvSW] EVEREST, G., A. VAN DER POORTEN, I. SHPARLINSKI and T. WARD. *Recurrence Sequences*. Mathematical Surveys and Monographs *104*. Amer. Math. Soc., Providence, RI, 2003.

[Ev] EVERTSE, J.-H. On sums of S-units and linear recurrences. *Compositio Math. 53* (1984), 225–244.

[ESS] EVERTSE, J.-H., H. P. SCHLICKEWEI and W. M. SCHMIDT. Linear equations in variables which lie in a muliplicative group. *Ann. of Math. (2) 155* (2002), 807–836.

[EMST] EVERTSE, J.-H., P. MOREE, C. L. STEWART and R. TIJDEMAN. Multivariate Diophantine equations with many solutions. *Acta Arith. 107* (2003), 103–125.

[Ga] GABBER, O. On space filling curves and Albanese varieties. *Geom. Funct. Anal. 11* (2001), 1192–1200.

[Gr] GROTHENDIECK, A. Formule de Lefschetz et rationalité des fonctions L. In: *Séminaire Bourbaki*, Vol. 9, Exp. 279, 41–55. Soc. Math. France, Paris, 1995.

[Ka94] KATZ, N. M. Review of ℓ-adic cohomology. In: *Motives* (Seattle, WA, 1991), 21–30. Proc. Sympos. Pure Math. *55*, Part 1. Amer. Math. Soc., Providence, RI, 1994.

[Ka96] —— *Rigid Local Systems*. Annals of Mathematics Studies *139*. Princeton University Press, Princeton, NJ, 1996.

[Ka99] —— Space filling curves over finite fields. *Math. Res. Lett. 6* (1999), 613–624.

[KS] KONYAGIN, S. and K. SOUNDARARAJAN. Two S-unit equations with many solutions. *J. Number Theory 124* (2007), 193–199.

[LW] LANG, S. and A. WEIL. Number of points of varieties in finite fields. *Amer. J. Math. 76* (1954), 819–827.

[La] LANG, S. *Algebra*. Revised third edition. Graduate Texts in Mathematics *211*. Springer-Verlag, New York, 2002.

[Lau] LAURENT, M. Équations exponentielles polynômes et suites récurrentes linéaires. In: *Journées arithmétiques de Besançon* (Besançon, 1985), 121–139, 343–344. Astérisque *147–148*. Soc. Math. France, Paris, 1987.

[Le] LECH, C. A note on recurring series. *Ark. Mat. 2* (1953), 417–421.

[M] MAHLER, K. Eine arithmetische Eigenschaft der Taylor-Koeffizienten rationaler Funktionen. *Proc. Akad. Wet. Amsterdam 38* (1935), 50–60.

[P] POONEN, B. Bertini theorems over finite fields. *Ann. of Math. (2) 160* (2004), 1099–1127.

[PS] VAN DER POORTEN, A. J. and H. P. SCHLICKEWEI. Additive relations in fields. *J. Austral. Math. Soc. Ser. A 51* (1991), 154–170.

[SS] SCHLICKEWEI, H. P. and W. M. SCHMIDT. The intersection of recurrence sequences. *Acta Arith. 72* (1995), 1–44.

[Sc] SCHMIDT, W. M. The zero multiplicity of linear recurrence sequences. *Acta Math. 182* (1999), 243–282.

[Se] SERRE, J.-P. Majorations de sommes exponentielles. In: *Journées Arithmétiques de Caen* (Univ. Caen, Caen, 1976), 111–126. Astérisque *41–42*. Soc. Math. France, Paris, 1977.

[Sk] SKOLEM, TH. Ein Verfahren zur Behandlung gewisser exponentialer Gleichungen und diophantischer Gleichungen. In: *8. Skand. Mat. Kongr.*, Stockholm, 1934, 163–188 (1934).

[Sp] SPERBER, S. On the p-adic theory of exponential sums. *Amer. J. Math. 108* (1986), 255–296.

[W1] WEIL, A. *Sur les courbes algébriques et les variétés qui s'en déduisent.* Actualités Sci. Ind. *1041*. Publ. Inst. Math. Univ. Strasbourg (1945). Hermann et Cie., Paris, 1948.

[W2] —— On some exponential sums. *Proc. Nat. Acad. Sci. U.S.A. 34* (1948), 204–207.

(Reçu le 11 mars 2009)

E. Bombieri

School of Mathematics
Institute for Advanced Study
Princeton, New Jersey 08540
U. S. A.
e-mail: eb@ias.edu

N. M. Katz

Department of Mathematics
Princeton University
Princeton, New Jersey 08544
U. S. A.
e-mail: nmk@math.princeton.edu