COMPUTING ZETA FUNCTIONS OF ARTIN-SCHREIER CURVES OVER FINITE FIELDS

ALAN G.B. LAUDER AND DAQING WAN

1. Introduction

We present a low-degree polynomial-time algorithm for computing the zeta function of certain Artin-Schreier curves defined over finite fields. One consequence is a practical method for computing the order of the Jacobian of an elliptic curve in characteristic 2, and more generally, for any hyperelliptic curve whose affine equation is of a particular form. Hyperelliptic curves have been proposed for use in public key cryptosystems by Koblitz [2, 14]. Our algorithm provides the first method of finding "random" hyperelliptic curves of arbitrary genus defined over large finite fields of characteristic 2 whose Jacobians have orders suitable for cryptographic use. Our method extends in a natural way to general Artin-Schreier curves, thus yielding a practical algorithm for arbitrary hyperelliptic curves in characteristic 2. We plan to present this generalisation in a sequel paper.

We now introduce some notation which will allow us to explain our results. Let p denote a prime number, and a a positive integer. Define $q = p^a$ and denote by $\mathbb{F}_q$ the finite field with q elements. Fix an algebraic closure $\bar{\mathbb{F}}_q$ of $\mathbb{F}_q$ and let $\mathbb{F}_{q^k}$ be the unique subfield of order q^k. We write $\bar{\mathbb{F}}_q^*$ for the set of non-zero elements in $\bar{\mathbb{F}}_q$. The Artin-Schreier curves over $\mathbb{F}_q$ we consider in this paper are defined by an equation of the form

$$Z^p - Z = f(X) \tag{1}$$

where $f \in \mathbb{F}_q[X, X^{-1}]$ is a Laurent polynomial. (More generally f may be a rational function, as we shall consider in the sequel paper.) Specifically, denote by C_f the curve embedded in $\bar{\mathbb{F}}_q^* \times \bar{\mathbb{F}}_q$ with equation (1) and let $\tilde{C}_f$ be the unique smooth projective curve birational to C_f. Let d denote the

Date: November 7, 2001.

Alan Lauder gratefully acknowledges the support of the EPSRC (Grant GR/N35366/01) and St John's College, Oxford. Daqing Wan is partially support by the NSF. The research was partly undertaken when the authors were visiting the Lorentz Center, University of Leiden, and the second author the Institute for Mathematical Sciences, National University of Singapore. We wish to thank Hendrik Lenstra and Harald Niederreiter, respectively, for their support during these visits, and also Pierrick Gaudry and Frederik Vercauteren for helpful comments on an earlier version of the paper.

Mathematics Subject Classification 2000: 11Y16, 11T99, 14Q15 *Key words and phrases*: Artin-Schreier, hyperelliptic, finite field, zeta function, algorithm, cryptography.

largest absolute value of any exponent which occurs in a non-zero term of f. For example, if $f \in \mathbb{F}_q[X]$ this is just the degree. Our main theorem is

Theorem 1. *The zeta function of the smooth projective curve $\tilde{C}_f$ may be computed deterministically in $\tilde{\mathcal{O}}(p^{4c}a^{3c}d^{4+c})$ with space complexity $\tilde{\mathcal{O}}(p^4a^3d^3)$ bits. (Here c is the exponent for multiplication in certain rings, and thus $1 \leq c \leq 2$ depending upon the method used.)*

Here we use Soft-Oh notation $\tilde{\mathcal{O}}$ which ignores logarithmic factors, as in [15, Section 6.3]. Using FFT methods one may take $c = 1$ giving a complexity of $\tilde{\mathcal{O}}(p^4a^3d^5)$.

We now explain how our algorithm may be applied to certain hyperelliptic curves in characteristic 2. Let C denote the affine curve with equation

$$Y^2 + X^mY = h(X)$$

where $h(X) \in \mathbb{F}_{2^a}[X]$ is of degree $2g+1$ and m is a non-negative integer not greater than g. Let $\tilde{C}$ be the unique smooth projective curve birational to C. Then $\tilde{C}$ is birational to an Artin-Schreier curve, as explained in Note 5, and thus one may compute the zeta function of $\tilde{C}$ in the complexity bounds of Theorem 1. From this it follows

Corollary 2. *The order of the Jacobian of the curve $\tilde{C}$ may be computed deterministically in $\tilde{\mathcal{O}}(a^3g^5)$ bit operations with space complexity $\tilde{\mathcal{O}}(a^3g^3)$ bits.*

With regard to the dependence on a, we note that our method when restricted to elliptic curves has comparable time complexity to [18]. Moreover, it is the first practical algorithm for hyperelliptic curves in characteristic 2 which has polynomial-time growth in both the field size and genus. (The problem of polynomial-time computability for arbitrary varieties in small characteristic was already solved in [15], but the general algorithm there is not very practical.) As previously mentioned, our algorithm can be extended to arbitrary hyperelliptic curves in characteristic 2, but we focus on the simplest cases in this paper. We refer to the references in [2] for the large literature on point counting, including [7, 19], and the more recent work [8, 9, 10, 11, 12, 17, 18, 22, 23, 25].

Sections 2, 3, 4 and 5 lay the mathematical foundation of our algorithm: it is based mainly upon an extension of the work of Dwork [6] due to Adolphson and Sperber [1]. Section 6 contains a statement of the algorithm for what we call Type 1 Artin-Schreier curves, and Section 7 describes exactly how to perform the main steps. The complexity analysis is tied up in Section 8, and Section 9 discusses the remaining type of Artin-Schreier curve in a more condensed fashion. As in [15] we aim to give a largely self-contained presentation.

2. L-functions and Artin-Schreier Curves

2.1. **General theory.** Let $\bar{\mathbb{Q}}$ denote an algebraic closure of the rationals $\mathbb{Q}$. Let $\Psi : \mathbb{F}_p \to \bar{\mathbb{Q}}$ be a non-trivial additive character, and $\mathrm{Tr}_k : \mathbb{F}_{q^k} \to \mathbb{F}_p$ the absolute trace map. A specific Ψ will be constructed in Section 4.2 but for now it may be arbitrary. Define $\Psi_k : \mathbb{F}_{q^k} \to \bar{\mathbb{Q}}$ to be the non-trivial additive character $\Psi \circ \mathrm{Tr}_k$.

For $f \in \mathbb{F}_q[X, X^{-1}]$ define

$$S_k^*(f, \Psi) := \sum_{x \in \mathbb{F}_{q^k}^*} \Psi_k(f(x)) \tag{2}$$

$$L^*(f, \Psi, T) := \exp\left(\sum_{k=1}^{\infty} \frac{S_k^*(f, \Psi)}{k} T^k\right). \tag{3}$$

For simplicity we shall omit the Ψ in this notation. Let C_f be the curve embedded in $\bar{\mathbb{F}}_q^* \times \bar{\mathbb{F}}_q$ with equation

$$Z^p - Z = f(X).$$

Let $\tilde{C}_f$ denote the unique smooth projective curve birational to C_f.

Lemma 3. *For each $x \in \mathbb{F}_{q^k}^*$ there are exactly n_x points of the form $(x, z) \in \mathbb{F}_{q^k}^* \times \mathbb{F}_{q^k}$ on C_f where*

$$n_x = \begin{cases} p & \text{if } \mathrm{Tr}_k(f(x)) = 0 \\ 0 & \text{if } \mathrm{Tr}_k(f(x)) \neq 0. \end{cases}$$

Proof. Follows from [16, Theorem 2.25]. □

Denote by $C_f(\mathbb{F}_{q^k})$ the set of $\mathbb{F}_{q^k}$-rational points on C_f. From Lemma 3 one deduces

$$\#(C_f(\mathbb{F}_{q^k})) = \sum_{a=0}^{p-1} \sum_{x \in \mathbb{F}_{q^k}^*} \Psi_k(af(x)) = \sum_{\theta \in G} \theta(S_k^*(f)) + (q^k - 1),$$

where G is the Galois group of $\mathbb{Q}(\zeta)$ over $\mathbb{Q}$, with $\zeta := \Psi(1)$ a primitive pth root of unity. Writing $Z(C_f, T)$ for the zeta function [15, Section 1] of C_f it follows that

$$Z(C_f, T) = \frac{\{\prod_{\theta \in G} \theta(L^*(f, T))\}(1 - T)}{(1 - qT)}. \tag{4}$$

Here G acts on power series $\mathbb{Q}(\zeta)[[T]]$ coefficient-wise, fixing monomials T^k.

To proceed further it is necessary to split the possible Laurent polynomials f into three types.

2.2. **Three types of Artin-Schreier curves.** *Type 1*: Assume that $f \in \mathbb{F}_q[X]$ has degree d not divisible by p. Here we can also define

$$S_k(f) := \sum_{x \in \mathbb{F}_{q^k}} \Psi_k(f(x)) = S_k^*(f) + \Psi(k\mathrm{Tr}_1(f(0))) \tag{5}$$

$$L(f,T) := \exp\left(\sum_{k=1}^{\infty} \frac{S_k(f)}{k} T^k\right) = L^*(f,T)(1 - \Psi(\mathrm{Tr}_1(f(0)))T)^{-1}. \tag{6}$$

Hence from (4)

$$Z(C_f, T) = \frac{\{\prod_{\theta \in G} \theta(L(f,T))\}(1-T)}{(1-qT)\{\prod_{\theta \in G} \theta(1 - \Psi(\mathrm{Tr}_1(f(0)))T)\}}.$$

(The latter term on the denominator is $(1-T)^{p-1}$ or $1 + T + \ldots + T^{p-1}$ depending upon whether $\mathrm{Tr}_1(f(0))$ is zero or not, although this does not concern us.) By Weil, the L-function $L(f,T)$ is a polynomial of degree $d-1$, and all of its reciprocal roots have complex absolute value $\sqrt{q}$ under all complex embeddings of $\bar{\mathbb{Q}}$. Thus $\prod_{\theta \in G} \theta(L(f,T))$ is a polynomial of degree $(p-1)(d-1)$, pure of weight 1 (see [24, Section 3] for definitions).

Let $Z(\tilde{C}_f, T)$ denote the zeta function of the smooth projective curve $\tilde{C}_f$. Let g be the genus of $\tilde{C}_f$. Again, by Weil, we know that

$$Z(\tilde{C}_f, T) = \frac{P(\tilde{C}_f, T)}{(1-T)(1-qT)},$$

where the numerator is a polynomial of degree $2g$, pure of weight 1. Since C_f and $\tilde{C}_f$ are birational they differ by a finite number of points, and hence their zeta functions differ by a factor of weight 0. Comparing the pure weight 1 parts in $Z(C_f, T)$ and $Z(\tilde{C}_f, T)$, we deduce

$$P(\tilde{C}_f, T) = \prod_{\theta \in G} \theta(L(f,T)). \tag{7}$$

In particular, the genus g of the curve $\tilde{C}_f$ is given by the formula

$$g = (p-1)(d-1)/2.$$

Type 2: Assume that $f \in \mathbb{F}_q[X^{-1}]$ has negative degree d^- not divisible by p. That is, d^- is the lowest exponent occurring in f. Then C_f is birational to C_{f^*} where $f^* := f(X^{-1})$, and we have reduced to type 1.

Type 3: Assume that $f \in \mathbb{F}_q[X, X^{-1}]$ but not $\mathbb{F}_q[X] \cup \mathbb{F}_q[X^{-1}]$. Let the negative degree be d^- and the positive degree be d^+. Assume p does not divide $d^- d^+$. In this case, by Weil, $L^*(f,T)$ is a polynomial of degree $d^+ - d^-$, pure of weight 1. By comparing the pure weight 1 parts in $Z(C_f, T)$ and $Z(\tilde{C}_f, T)$ we find that

$$P(\tilde{C}_f, T) = \prod_{\theta \in G} \theta(L^*(f,T)), \tag{8}$$

is the numerator of $Z(\tilde{C}_f, T)$. In particular, the genus g of the curve $\tilde{C}_f$ is given by the formula

$$g = (p-1)(d^+ - d^-)/2.$$

Thus in all cases the computation of the zeta function of the smooth projective curve $\tilde{C}_f$ reduces to the evaluation of the L-function of certain one variable exponential sums.

Note 4 The degree conditions on f are essential for cohomological arguments; however, given a Laurent polynomial f which does not satisfy them one may replace it by a new polynomial $\tilde{f}$ which does such that C_f and $C_{\tilde{f}}$ are isomorphic. For each term $a_{jp}X^{jp}$, the isomorphism $Z \to Z + bX^j$ $(X \to X)$ shows that we can replace the term $a_{jp}X^{jp}$ with b_jX^j where $b := a_{jp}^{1/p} \in \mathbb{F}_q$. Repeat this procedure until no term has a non-zero exponent divisible by p. The resulting polynomial $\tilde{f}$ has no terms with non-zero exponents divisible by p, and thus it certainly satisfies any necessary degree restrictions.

Note 5 Let C be the curve in Corollary 2. Then C is birational to the curve C_f with equation $Z^2 + Z = f(X)$ where $f := X^{-2m}h$. This can be seen by making the change of variable $Y = ZX^m$. Thus $\tilde{C}$, as in Corollary 2, is birational to $\tilde{C}_f$ and the zeta function of $\tilde{C}_f$ can then be computed using Note 4, to get f in the correct form, and the algorithm we will present. By the special value formula for the zeta function at $T = 1$, the order of the Jacobian is the numerator of the zeta function evaluated at $T = 1$ [2, page 175]. Similar comments apply to curves over $\mathbb{F}_q$ of the form $Y^p - X^mY = h(X)$ where $q = p^a$ and $p-1$ divides m. Also, general hyperelliptic curves in characteristic 2 are birational to Artin-Schreier curves of the form $Z^2 + Z = f(X)$ for f a rational function. As such they may be tackled using a generalisation of our approach.

2.3. **Type 1 Artin-Schreier curves.** In Sections 3, 4, 5, 6, 7 and 8 we shall denote by f a polynomial in $\mathbb{F}_q[X]$ of degree d not divisible by p. We write $f = \sum_{j \in J} a_jX^j$ where $a_j \neq 0$. In these sections we will explain how to compute the zeta function of the smooth projective curve $\tilde{C}_f$. That is, we shall cover classical "type 1" Artin-Schreier curves. In Section 9 we shall discuss the modifications required when $f \in \mathbb{F}_q[X, X^{-1}]$ is a Laurent polynomial with both positive and negative exponents. These are "type 3" Artin-Schreier curves. As mentioned before, "type 2" may be reduced to "type 1" and we shall not discuss them again.

3. p-ADIC THEORY

3.1. **p-adic rings.** Let $\mathbb{Q}_p$ denote the p-adic numbers with ring of integers $\mathbb{Z}_p$. Fix Ω the completion of an algebraic closure of $\mathbb{Q}_p$. Denote by ϵ a primitive $(q-1)$th root of unity in Ω, and $\pi \in \Omega$ an element which satisfies $\pi^{p-1} = -p$. Define

$$A := \mathbb{Z}_p[\epsilon, \pi]. \tag{9}$$

In particular, A has residue field $\mathbb{F}_q$. Elements in A may be represented via π-adic expansions whose coefficients are from some distinct set of representatives for the quotient $A/(\pi)$ of size q. For any primitive pth root of unity ζ we have that both $\mathbb{Q}_p(\zeta)$ and $\mathbb{Q}_p(\pi)$ are the unique totally ramified extensions of $\mathbb{Q}$ of degree $p-1$ (since $\mathrm{ord}(1-\zeta) = \mathrm{ord}(\pi) = 1/(p-1)$). Thus $\mathbb{Z}_p[\zeta] = \mathbb{Z}_p[\pi]$ and so A contains the exponential sums defined in the previous section. Let G denote the group of automorphisms $\theta_j : \mathbb{Q}_p(\pi) \to \mathbb{Q}_p(\pi)$ for $1 \le j \le p-1$ where each θ_j fixes $\mathbb{Q}_p$ and

$$\theta_j(\pi) := \eta^j \pi. \tag{10}$$

Here $\eta \in \mathbb{Z}_p$ is a primitive $(p-1)$th root of unity. Then G is the Galois group of the extension $\mathbb{Q}_p(\pi)/\mathbb{Q}_p$ (its action on ζ is $\theta_j : \zeta \mapsto \zeta^{\eta^j \bmod p}$, although we shall not need this explicitly).

It will be convenient to work in complete rings which contain arbitrary roots π^r of π, for r a rational number. To this end, let $\Pi = \cup_{z \in \mathbb{N}} \{\pi^{1/z}\}$ and $\tilde{A}$ denote the completion of the ring $\mathbb{Z}_p[\epsilon, \Pi]$. Here $\mathbb{N}$ denotes the positive integers.

Denote by τ the map on $\tilde{A}$ defined as

$$\tau(\epsilon) := \epsilon^p, \ \tau \text{ fixes } \mathbb{Z}_p[\Pi] \text{ and is continuous.} \tag{11}$$

Let ord and $|.|_p$ denote the p-adic valuation and norm on $\tilde{A}$ normalised so that $\mathrm{ord}(p) = 1$ and $|p|_p = 1/p$.

Note that the ring A can be constructed and computed in easily; see [15, Section 3] for more details on p-adic fields. (The larger ring $\tilde{A}$ is only introduced for mathematical convenience; all of our computations are performed in A.)

3.2. **A weight function.** As in [15, Section 4] we define a weight function: For each non-negative integer $u \in \mathbb{Z}_{\ge 0}$ let

$$\mathrm{wt}(u) = \lceil u/d \rceil.$$

Here $\lceil * \rceil$ is the least integer not less than $*$ and d the degree of the polynomial $f \in \mathbb{F}_q[X]$. Define $\tilde{\mathrm{wt}}(u) := u/d$ and so $\lceil \tilde{\mathrm{wt}}(u) \rceil = \mathrm{wt}(u)$. Notice that $\pi^{\mathrm{wt}(r)} \in A$ and $\pi^{\tilde{\mathrm{wt}}(r)} \in \tilde{A}$ for every $r \in \mathbb{Z}_{\ge 0}$.

3.3. **Banach modules.** Let B denote a complete subring of Ω. A Banach module over B is an ultrametrically normed complete module E over B, such that $||re|| = |r|_p ||e||$ for $r \in B$ and $e \in E$, where $||.||$ is the module norm. An orthonormal basis for E is a set $\{e_i \,|\, i \in \mathbb{N}\}$ such that every element in E can be written uniquely in the form $\sum_i b_i e_i$ where $b_i \in B$ with $|b_i|_p \to 0$ as $i \to \infty$. (See [20] and [4, Section A].) In the case that B is a field we call it a Banach space.

Definition 6. *For each rational number $\delta > 0$, let $\tilde{L}(\delta)$ be the Banach module over $\tilde{A}$ whose orthonormal basis consists of all terms $\pi^{\tilde{\mathrm{wt}}(r)\delta}X^r$ for $r \in \mathbb{Z}_{\geq 0}$. Let $A\{X\}$ ($\tilde{A}\{X\}$ respectively) denote the Banach module over A ($\tilde{A}$ resp.) whose orthonormal basis consists of all terms X^r for $r \in \mathbb{Z}_{\geq 0}$.*

Note that $\tilde{L}(\delta') \subset \tilde{L}(\delta)$ for $\delta' > \delta$ and all the above spaces lie in $\tilde{A}\{X\}$. One may check that all the above spaces are closed under multiplication, and are in fact rings. Extend τ to act on each power series in the ring $\tilde{A}\{X\}$ by taking

$$\tau(X) := X, \ \tau \text{ is } \tilde{A}\text{-linear and continuous.} \tag{12}$$

4. Analytic representation of additive characters

4.1. **Dwork's splitting functions.** Let $\theta(t)$ denote the splitting function [15, Section 4.1]

$$\theta(t) := \theta_1(t) = \exp(\pi(t - t^p)). \tag{13}$$

Write

$$\hat{f} := \sum_{j \in J} \hat{a}_j X^j \tag{14}$$

for the polynomial over A obtained by taking the Teichmuller lifting of each coefficient of f.

Lemma 7. *For each term $\hat{a}_j X^j$ in $\hat{f}$, we have $\theta(\hat{a}_j X^j) \in \tilde{L}(\delta)$ for any*

$$\delta < \left(\frac{p-1}{p}\right)^2$$

Proof. Writing $\theta(t) =: \sum_{r=0}^{\infty} \lambda_r t^r$ we see from [15, Equation (3)] that $\mathrm{ord}(\lambda_r) \geq (p-1)r/p^2$. Terms in $\theta(\hat{a}_j X^j)$ are of the form $\lambda_r \hat{a}_j^r X^{jr}$ for r a non-negative integer. Now $\mathrm{ord}(\lambda_r \hat{a}_j^r) \geq (p-1)r/p^2$ since $\mathrm{ord}(\hat{a}_j) = 0$. Also $0 \leq j \leq d$ and so $\tilde{\mathrm{wt}}(jr) = jr/d \leq r$. Thus

$$\mathrm{ord}(\lambda_r \hat{a}_j^r) \geq \frac{p-1}{p^2} r \geq \frac{p-1}{p^2} \tilde{\mathrm{wt}}(jr).$$

Also for $\delta < ((p-1)/p)^2$ we have

$$\mathrm{ord}(\pi^{\delta \tilde{\mathrm{wt}}(jr)}) < \frac{p-1}{p^2} \tilde{\mathrm{wt}}(jr)$$

and the result follows. □

Definition 8. *Let F and $F^{(a)}$ be defined as*

$$F := \prod_{j \in J} \theta(\hat{a}_j X^j) \tag{15}$$

$$F^{(a)} := \prod_{i=0}^{a-1} \tau^i(F(X^{p^i})). \tag{16}$$

Recall here that $q = p^a$. Both F and $F^{(a)}$ are one-way infinite power series in $A\{X\}$. By Lemma 7 and the fact that each $\tilde{L}(\delta)$ is a ring we have

Lemma 9. *The power series $F \in \tilde{L}(\delta)$ for any*

$$\delta < \left(\frac{p-1}{p}\right)^2$$

4.2. **Dwork's additive character.** For $k \geq 1$ define

$$\Phi_k(t) := \prod_{i=0}^{ak-1} \theta(t^{p^i}) \in \mathbb{Z}_p[\pi][[t]]$$

and let

$$\Phi(t) := \prod_{i=0}^{a-1} \theta(t^{p^i}) \in \mathbb{Z}_p[\pi][[t]].$$

Denote by Teich the Teichmüller lifting map from $\bar{\mathbb{F}}_q$ to $\mathbb{Z}_p^{\mathrm{unram}}$, where $\mathbb{Z}_p^{\mathrm{unram}}$ is the unramified integral closure of $\mathbb{Z}_p$ in Ω. Then

$$\Psi_k := \Phi_k \circ \mathrm{Teich}, \Psi := \Phi \circ \mathrm{Teich}$$

are non-trivial characters from $\mathbb{F}_{q^k}$ and $\mathbb{F}_p$, respectively, to $\mathbb{Z}_p[\pi]$. We have that $\Psi_k(x) = \Psi(\mathrm{Tr}_k(x))$ where Tr_k is the trace map from $\mathbb{F}_{q^k}$ to $\mathbb{F}_p$. (See [15, Lemma 6].)

The following lemma is proved exactly as [15, Proposition 9]. It gives an analytic expression for the exponential sum $S_k^*(f) := S_k^*(f, \Psi)$

Lemma 10. *Let $S_k^*(f)$ be the exponential sum defined in (2) using Dwork's additive character Ψ, and $F^{(a)}(X)$ the power series from Definition 8. Then*

$$S_k^*(f) = \sum_{x^{q^k-1}=1} F^{(a)}(x)F^{(a)}(x^q)\dots F^{(a)}(x^{q^{k-1}}).$$

Here the sum is over the Teichmüller liftings in $\mathbb{Z}_p^{\mathrm{unram}}$ of the points on the torus $\mathbb{F}_{q^k}^*$.

4.3. **Completely continuous maps.** The next step is to introduce operators on the module L so that the righthand side of the above expression can be interpreted as the "trace" of a map on a Banach module.

Definition 11. *Let ψ_p be the map on $\tilde{A}\{X\}$ which acts on monomials as*

$$\psi_p(X^r) := \begin{cases} X^{r/p} & \text{if } p \text{ divides } r \\ 0 & \text{otherwise} \end{cases}$$

and extends to all of $\tilde{A}\{X\}$ by τ^{-1}-linearity and continuity. Specifically, $\psi_p(\sum_r A_r X^r) = \sum_{r,\,p|r} \tau^{-1}(A_r)X^{r/p}$. Write $\psi_q := \psi_p^a$, a linear map since τ^{-a} is the identity on $\tilde{A}$. Let $\alpha := \psi_p \circ F$ and $\alpha_a := \psi_q \circ F^{(a)}$. Precisely, α is multiplication by F followed by the map ψ_p, and likewise for α_a (see [15, Definitions 20, 21]*).*

Lemma 12. *The maps satisfy* $\alpha_a = \alpha^a$.

Proof. Proved exactly as in [15, Lemma 22]. □

Lemma 13. *The map* α *is stable on* $\tilde{L}(\delta)$ *for*

$$\delta < \frac{(p-1)^2}{p}$$

Proof. We first claim that $\psi_p(\tilde{L}(\delta)) \subseteq \tilde{L}(p\delta)$ for any rational $\delta > 0$: For this it is enough to observe that $\psi_p(\pi^{\tilde{\mathrm{wt}}(r)\delta} X^r) = \pi^{\tilde{\mathrm{wt}}(r/p)p\delta} X^{r/p}$ for any r divisible by p. Now let $G \in \tilde{L}(\delta)$ where δ satisfies the necessary inequality, depending on p. Then by Lemma 9 $F \in \tilde{L}(\delta/p)$ and so since also $G \in \tilde{L}(\delta/p)$ by closure under multiplication we find that $FG \in \tilde{L}(\delta/p)$. Hence $\psi_p(FG) \in \tilde{L}(\delta)$, that is $\alpha(G) \in \tilde{L}(\delta)$ as required. □

Definition 14. *Let* L *be defined as*

$$L := \begin{cases} \tilde{L}(1) \cap A\{X\} & \text{if } p > 2 \\ (\tilde{L}(1/(2+(1/4d))) \cap A\{X\}) \otimes \mathbb{Q} & \text{if } p = 2. \end{cases}$$

Thus for $p > 2$ we have that L is just the Banach module over A with orthonormal basis the terms $\pi^{\mathrm{wt}(u)} X^u$ for non-negative integers u. For $p = 2$ it is the Banach space over $A \otimes \mathbb{Q}$ with orthonormal basis the terms $\pi^{\lceil \delta \tilde{\mathrm{wt}}(u) \rceil} X^u$ for non-negative u, where $\delta = 1/(2+(1/4d))$. (Note that any $\delta < 1/2$ would probably do, and our choice is just enough for an adhoc argument we present later.)

Lemma 15. *The maps* α *and* α_a *are stable on* L.

Proof. First suppose $p > 2$. Then α is stable on $\tilde{L}(1)$ by Lemma 13. Certainly α is stable on the ring of convergent power series $A\{X\}$, since $F \in A\{X\}$. Thus α is stable on $\tilde{L}(1) \cap A\{X\} = L$. That α_a is stable on L now follows from Lemma 12. Second, consider the case $p = 2$. Putting $p = 2$ in Lemma 13 we find α is stable on $\tilde{L}(1/(2+1/(4d)))) \cap A\{X\}$. Given $G \in L$ we have that $mG \in \tilde{L}(1/2+(1/4d))) \cup A\{X\}$ for some p-adic integer m. The result now follows easily. □

Note 16 The rings $\tilde{A}$ and $\tilde{L}(\delta)$ and the function $\tilde{\mathrm{wt}}$ were introduced to prove the above result in as simple a manner as possible; we shall have little further need for them, working from now on mainly with A and L.

For certain classes of linear maps on Banach modules the trace and determinant are defined. This is done in the usual way, via matrices for the maps with respect to an orthonormal basis. We refer to [20] and [4, Section A2] for definitions. The key result is the Dwork Trace Formula.

Theorem 17. *With* $S_k^*(f)$ *and* $L^*(f,T)$ *the exponential sum and* L*-function defined as in (2) and (3) using Dwork's additive character* Ψ *(Section 4.2),* *and α_a the map on L given in Definition 11 we have*

$$S_k^*(f) = (q^k - 1)\mathrm{Tr}(\alpha_a^k | L).$$

Thus we have

$$L^*(f,T) = \frac{\det(1 - T\alpha_a|L)}{\det(1 - Tq\alpha_a|L)}.$$

Here the trace and determinant are defined via matrices for the maps with respect to the orthonormal basis of L.

Proof. This is in essence a case of [15, Theorem 25] (with "$n = 0$"). Note that the matrix "M_a" in [15, Theorem 25] is that for the map α_a with respect to a "formal basis" [15, Section 5.2] of the form $\{X^i \,|\, i \in \mathbb{Z}_{\geq 0}\}$. For the above formulae we require a matrix with respect to the orthonormal basis $\{\pi^{\mathrm{wt}(i)} X^i \,|\, i \in \mathbb{Z}_{\geq 0}\}$ (when $p > 2$). One may verify, using for example [15, Equation (12)], that the trace of powers of these two matrices are the same. $\square$

These formulae may be used to compute the zeta function in a similar fashion to [15]. Namely, a finite matrix may be computed which represents the map α acting on some appropriate modular reduction of L. This matrix is then used to compute the characteristic polynomial of α_a itself up to a necessary p-adic accuracy. Note that one may use the simpler splitting function θ_1 for $p = 2$ in this approach. However, this "chain level" method results, for example, in a time complexity of $\tilde{\mathcal{O}}(a^{4.38})$ with space $\tilde{\mathcal{O}}(a^4)$ bits using the fastest methods for matrix multiplication and ring arithmetic. Using some homological algebra one can derive a better "cohomological" formula leading to an improved algorithm. This is what we do in this paper.

At this stage since f is an ordinary polynomial rather than a Laurent polynomial with negative and positive terms, we can do a little more work to derive a better chain level formula: Let $L_{>0}$ denote the Banach module comprising of those power series in L with zero constant term. For $p > 2$ the module $L_{>0}$ is defined over A, and for $p = 2$ over $A \otimes \mathbb{Q}$. It follows easily from Lemma 15 that

Lemma 18. *The maps α and α_a are stable on $L_{>0}$.*

Now one may check via a matrix for the map α_a with respect to the orthonormal basis $\{\pi^{\mathrm{wt}(r)} X^r\}_{r \geq 0}$ for $p > 2$, and $\{\pi^{\lceil \delta \tilde{\mathrm{wt}}(r) \rceil} X^r\}_{r \geq 0}$ for $p = 2$ (where $\delta = 1/(2 + 1/(4d))$), that

$$\det(1 - T\alpha_a|L) = (1 - F^{(a)}(0)T)\det(1 - T\alpha_a|L_{>0}).$$

Here

$$F^{(a)}(0) = \Psi(\mathrm{Tr}_1(a_0))$$

is a root of unity, where Ψ is Dwork's additive character from $\mathbb{F}_p$ to $\mathbb{Z}_p[\pi]$. Let $S_k(f)$ and $L(f,T)$ be as in (5) and (6). Then one has $S_k(f) = S_k^*(f) + \Psi(k\mathrm{Tr}_1(a_0))$ and so $L(f,T) = L^*(f,T)(1 - F^{(a)}(0)T)^{-1}$. Hence we have

Lemma 19. *Let $L(f,T)$ be the exponential sum over the affine line from (5). Then*

$$L(f,T) = \frac{\det(1 - T\alpha_a | L_{>0})}{\det(1 - qT\alpha_a | L)}.$$

5. Dwork cohomology

Let H be the polynomial in L defined as

$$H = \pi \hat{f} \tag{17}$$

(In fact $H \in \tilde{L}(1) \cap A\{X\}$ in all cases, and this latter ring equals L for $p > 2$, and lies strictly within L for $p = 2$.)

Let D be the operator on L defined as

$$D = X\frac{d}{dX} + H_X$$

where

$$H_X := X\frac{dH}{dX}.$$

Here d/dX is the usual differential operator on polynomials extended to power series by continuity. Now $X(d/dX)$ is stable on L, and L is a ring. From this it follows that D is stable on L, and in fact maps L to $L_{>0}$.

Note 20 We pause to explain the motivation behind the above definitions: Define $\hat{\theta}(t) := \prod_{i=0}^{\infty} \theta(t^{p^i})$. One may check that $\hat{\theta}(t) = \exp(\pi t)$. Now $F = \prod_{j \in J} \theta(\hat{a}_j X^j)$. Defining $\hat{F}(X) := \prod_{j \in J} \hat{\theta}(\hat{a}_j X^j)$ we find that this equals $\exp(\pi \hat{f})$ which is just $\exp(H)$, with H as in (17). Since $\hat{a}_j^q = \hat{a}_j$ and $\tau^i(\hat{a}_j) = \hat{a}_j^{p^i}$ for each $j \in J$ it follows from Definition 8 that $F^{(a)}(X) = \hat{F}(X)/\hat{F}(X^q)$. From this we see $\alpha_a = \psi_q \circ F^{(a)} = \exp(-H) \circ \psi_q \circ \exp(H)$. Define the operator $E := X(d/dX)$, and so $E \circ (q\psi_q) = \psi_q \circ E$. Then with D as above one may check $D = E + E(H) = \exp(-H) \circ E \circ \exp(H)$. Thus D and α_a are obtained from E and ψ_q by some kind of twisting; also, it now follows that $D \circ (q\alpha_a) = \alpha_a \circ D$, which is the crucial relation. (See [6, pages 55-60] and [21, pages 267-270] for more details.)

Let $\mathcal{L}$ be the complex

$$0 \longrightarrow L \xrightarrow{D} L_{>0} \longrightarrow 0.$$

This is a complex of A-modules when $p > 2$ and $A \otimes \mathbb{Q}$-spaces for $p = 2$. Denote by H_1 and H_0 the kernel and co-kernel of the map D. In particular, $H_0 := L_{>0}/D(L)$.

Proposition 21. *The map D is injective and so $H_1 = 0$. Moreover, H_0 is a finite free A-module of rank $d-1$. A basis for H_0 may be taken as the set of terms*

$$\{\pi^{\mathrm{wt}(i)} X^i \,|\, 0 < i < d\} = \{\pi X, \pi X^2, \ldots, \pi X^{d-1}\}.$$

Proof. Over the formal power series ring $\Omega[[X]]$, the formal solutions of the first order linear differential equation $D = 0$ is the one-dimensional subspace generated by $\exp(-H)$. But the power series $\exp(-H) \notin L$ (the decay rate of coefficients is too slow). This shows that the restriction of the operator D to L is injective. The second part of the proposition follows from the normal form computations in section 7. ☐

By Note 20 we have

$$D \circ q\alpha_a = \alpha_a \circ D.$$

Thus the map α_a defines a chain map on $\mathcal{L}$:

$$\begin{array}{ccccccc} 0 & \longrightarrow & L & \xrightarrow{D} & L_{>0} & \longrightarrow & 0 \\ & & \downarrow q\alpha_a & & \downarrow \alpha_a & & \\ 0 & \longrightarrow & L & \xrightarrow{D} & L_{>0} & \longrightarrow & 0. \end{array}$$

Denote by $H_0(\alpha_a)$ and $H_1(q\alpha)$ the maps induced on the homology H_0 and H_1 by this chain map, and $\det(1 - H_0(\alpha_a)T)$, $\det(1 - H_1(q\alpha_a)T)$ the corresponding determinants.

Theorem 22. *The L-function from Theorem 17 satisfies*

$$L(f, T) = \det(1 - H_0(\alpha_a)T).$$

Proof. We have

$$\det(1 - Tq\alpha_a|L)^{-1} \det(1 - T\alpha_a|L_{>0}) = \det(1 - H_1(q\alpha_a)T)^{-1} \det(1 - H_0(\alpha_a)T).$$

This identity is proved in the same way that [20, Proposition 9] is derived from [20, Lemma 2]. Now $H_1 = 0$ and so the first factor on the righthand side is 1. The expression for the L-function now follows from Theorem 17. ☐

Corollary 23. *The zeta function $Z(\tilde{C}_f, T)$ of the smooth projective curve $\tilde{C}_f$ birational to the affine curve with equation $Z^p - Z = f(X)$ satisfies*

$$Z(\tilde{C}_f, T) = \frac{\prod_{j=1}^{p-1} \theta_j(\det(1 - H_0(\alpha_a)T))}{(1 - T)(1 - qT)}.$$

Here θ_j are the automorphisms (10) of $\mathbb{Z}_p[\pi]$ extended to act on polynomials by fixing T. The numerator is a polynomial of degree $(p-1)(d-1)$.

Proof. Follows from Theorem 22 and equation (7). ☐

Thus the strategy of the algorithm is to compute the determinant of the map $H_0(\alpha_a)$ on the zeroth homology H_0, up to a suitable modular precision. This may be done efficiently via the following lemma, which is an immediate consequence of Lemma 12.

Lemma 24. *Let $H_0(\alpha)$ denote the map induced on H_0 by α. Then $H_0(\alpha_a) = H_0(\alpha)^a$.*

It will be enough to compute the coefficients of the characteristic polynomial of Frobenius modulo p^N for

$$N = \lfloor (p-1)(d-1)(1+a/2) + 1 \rfloor + \varepsilon$$

where $\varepsilon = 0$ for $p > 2$ and 1 for $p = 2$. This follows since the L-function of the exponential sum $L(f,T)$ has reciprocal roots whose complex absolute values are $\sqrt{q}$. Thus the coefficient of T^k in the polynomial $\prod_{\theta_j \in G} \theta_j(L(f,T))$ are integers of absolute value at most $\binom{2g}{k} p^{ak/2} \leq 2^{2g} p^{ak/2}$. Since the polynomial $\prod_{\theta_j \in G} \theta_j(L(f,T))$ has degree $2g = (p-1)(d-1)$, it follows that working modulo p^N for $N > (p-1)(d-1)(1+a/2)$ is sufficient. The extra ε arises for $p = 2$ since in this case computations will involve p-adic numbers in $A \otimes \mathbb{Q}$ of order ≥ -1, rather than p-adic integers in A (see Corollary 30).

6. The Algorithm

We now present our point counting algorithm for Type 1 Artin-Schreier curves (see Section 2 for our classification of Artin-Schreier curves).

Algorithm 25 [Artin-Schreier - Type 1]
Input: An equation $Z^p - Z = f(X)$ over $\mathbb{F}_q$, where $f \in \mathbb{F}_q[X]$ and $q = p^a$.
Output: The zeta function $Z(\tilde{C}_f, T)$ of the unique smooth projective curve birational to the affine curve defined by this equation.

Step 0: Replace f by a polynomial all of whose terms have exponents not divisible by p, in the manner explained in Note 4. Denote this new polynomial also by f. This will not change the zeta function. Set $N := \lfloor (p-1)(d-1)(1+a/2) + 1 \rfloor + \varepsilon$, where d is the degree of f, and $\varepsilon = 0$ or 1 according to whether p is greater or equal to 2. We shall compute the coefficients of the numerator of the zeta function modulo p^N.

Step 1: Compute the power series F given in Definition 8 with coefficients determined modulo p^N. Let α be the map on the ring L (Definition 6) defined as $\alpha = \psi_p \circ F$ (Definition 11). Let $H_0(\alpha)$ be the map induced on the zeroth homology H_0 of the complex $\mathcal{L}$ by α.

Step 2: Let $\pi X, \pi X^2, \dots, \pi X^{d-1}$ be the basis for the zeroth homology H_0. For each basis element e, compute the image $H_0(\alpha)(e) \in H_0$ with coefficients determined modulo p^N. Construct M, defined as the matrix representing the map $H_0(\alpha)$ with respect to the basis, with coefficients determined modulo p^N.

Step 3: Compute

$$M_a := \tau^{a-1}(M)\tau^{a-2}(M)\dots\tau(M)M$$

modulo p^N, where the map τ is the lifting of Frobenius to A as given in (11). Thus M_a is a matrix for the map $H_0(\alpha_a)$.

Step 4: Output the rational function

$$Z(\tilde{C}_f, T) := \frac{\prod_{j=1}^{p-1} \theta_j(\det(I - M_a T))}{(1-T)(1-qT)}$$

where θ_j is the automorphism from (10) extended to act on $\mathbb{Z}_p[\pi][T]$ by fixing monomials.

The correctness of the algorithm follows from Corollary 23. The matrix M is called the absolute Frobenius matrix, and $\det(I - M_a T)$ the characteristic polynomial of Frobenius. In Section 7 we will describe exactly how this first matrix is computed, allowing us to give a complexity analysis of the algorithm in Section 8. This will complete the proof of Theorem 1 for Type 1 and 2 Artin-Schreier curves. We present the algorithm for Type 3 curves in Section 9.

Note 26 The above algorithm can be practically improved by using the functional equation

$$T^{2g} q^g P(\tilde{C}_f, \frac{1}{qT}) = P(\tilde{C}_f, T).$$

This functional equation shows that we can take the smaller integer accuracy $N' = \lfloor (p-1)(d-1)(1 + a/4) + 1 \rfloor$, computing only the first half of the coefficients in $P(\tilde{C}_f, T)$, and then recovering the second half of the coefficients by the functional equation.

7. Performing the main steps

We shall work with elements in L with coefficients determined modulo p^N. (In the case $p = 2$ by this we mean that the coefficients a_i of p^i in their p-adic expansions are known for $i < N$.) If an element is given to this accuracy we say it lies in $L \bmod p^N$. Similarly, H_0 is the free module over A when $p > 2$, and $A \otimes \mathbb{Q}$ when $p = 2$, spanned by the basis monomials

$$\{\pi X, \pi X^2, \ldots, \pi X^{d-1}\} \tag{18}$$

and we write $s \in H_0 \bmod p^N$ if an element s is given in H_0 with coefficients modulo p^N. In the next two sections we shall explain how given an element $G \in L_{>0} \bmod p^N$ we can compute $s \in H_0 \bmod p^N$ such that

$$G = D(r) + s \bmod p^N$$

for some $r \in L$. We call this process finding a "normal form" in $L_{>0} \bmod p^N$ and say that s is "cohomologous" to G.

7.1. **Normal forms in $L_{>0}/D(L)$ - Case $p>2$.** In the case $p>2$ because of the decay rate of power series in L working in $L_{>0}/D(L)$ is particularly simple. From (14) and (17)

$$H=\pi\sum_{j=0}^{d}\hat{a}_jX^j$$

and thus

$$H_X=\pi\sum_{j=1}^{d}\hat{a}_j jX^j$$

where $\mathrm{ord}(\hat{a}_d d)=0$ since $d\neq 0$ in $\mathbb{F}_q$. Consider the basis monomial $\pi^{\mathrm{wt}(u)}X^u$ for the Banach module L where $u\geq d$. We have the trivial identity

$$\pi^{\mathrm{wt}(u)}X^u=(\pi(\sum_{j=1}^{d}\hat{a}_j jX^j)+X\frac{d}{dX})((\hat{a}_d d)^{-1}\pi^{\mathrm{wt}(u)-1}X^{u-d})$$
$$-(\pi(\sum_{j=1}^{d-1}\hat{a}_j jX^j)+X\frac{d}{dX})((\hat{a}_d d)^{-1}\pi^{\mathrm{wt}(u)-1}X^{u-d})=:D(r)+r'$$

where $r\in L$ and $r'\in L_{>0}$. Moreover, r' is a sum of monomials of degree less than u and greater or equal to $u-d+1$. Now let $G\in L_{>0}\bmod p^N$ and assume that $b\pi^{\mathrm{wt}(u)}X^u$ is the highest term which occurs in G for some $b\in A$. (Note that by the decay rate on the coefficients of elements in L we see that $u=\mathcal{O}(Npd)$.) We may suppose for our purposes that $u\geq d$. We can write this term as $D(br)+br'$ where br' is a sum of monomials of degree less than u but not less than $u-d+1$. Write $G=G'+b\pi^{\mathrm{wt}(u)}X^u$. Then in $L_{>0}/D(L)$ we find that G is "cohomologous" to $G_1:=G'+br'$. To compute G_1 requires d multiplications in $A\bmod p^N$, and the same number of additions (plus a little precomputation which can be ignored). Now continue in this way until the highest term in some G_m has degree less than d. Precisely, we need m at most $u-d+1=\mathcal{O}(Npd)$.

In this way we may find a "normal form" for any element in $L_{>0}$. That is, given $G\in L_{>0}\bmod p^N$ we can write it as

$$G=D(r)+s$$

where $r\in L$ and $s\in H_0\bmod p^N$ is a linear combination with coefficients in $A\bmod p^N$ of the basis monomials

$$\{\pi X,\pi X^2,\dots,\pi X^{d-1}\}.$$

The process above has time complexity

$$\tilde{\mathcal{O}}((Npd)d(Npa)^c) \tag{19}$$

bit operations where c is the exponent for multiplication as defined in Section 8.1.

7.2. **Normal forms in $L_{>0}/D(L)$ - Case $p = 2$.** The reduction process for $p = 2$ is in essence the same; namely, one uses a trivial identity to reduce the degree at each step. However, the justification that it works is somewhat more involved, since the power series in L decay rather more slowly.

We shall assume that $d\hat{a}_d = 1$; the more general situation just involves some notational complications, the essential point being that we always have $\mathrm{ord}(d\hat{a}_d) = 0$. Let $J_1 := J - \{d\}$ be the support set J of f excluding the element d, and $b_j := j\hat{a}_j$. In a similar manner to before we have the identity

$$X^u = D\left(\frac{1}{\pi}X^{u-d}\right) - \left(\frac{(u-d)}{\pi}X^{u-d} + \sum_{j\in J_1} b_j X^{u-d+j}\right). \tag{20}$$

This is used to reduce a power series given in finite precision in $L_{>0}$ to its normal form, that is, a polynomial over $A \otimes \mathbb{Q}$ of degree less than d with no constant term. As before the complexity is (19).

We must also address one theoretical problem. Let $G := \sum_u a_u X^u \in L$, and suppose that $a_u X^u = D(r_u) + s_u$ with s_u a polynomial of degree less than d over $A \otimes \mathbb{Q}$. Then $G = D(\sum_u r_u) + \sum_u s_u$ provided that $|r_u|, |s_u| \to 0$ as $u \to \infty$. To show that these sequences indeed converge, and to get a bound on their p-adic orders, requires a more careful analysis which we now perform in a rather adhoc fashion.

Observe that $\mathrm{ord}(b_j) = \mathrm{ord}(j) + \mathrm{ord}(\hat{a}_j) = \mathrm{ord}(j)$ for $j \in J_1$, and so in particular if j is even then $\mathrm{ord}(b_j) \geq 1$. Iterating the identity (20) κ times we find that X^u can be written as $D(a_\kappa) + b_\kappa$ where $a_\kappa = a_{\kappa-1} + (\tilde{b}_{\kappa-1}/\pi X^d)$ and, ignoring signs, b_κ is a sum of terms of the form

$$b(S, B, u)X^{u-n(S,B)}. \tag{21}$$

Here $\tilde{b}_{\kappa-1}$ is the set of terms in $b_{\kappa-1}$ of degree greater or equal to d, with $b_0 = X^u$ and $a_0 = 0$. The coefficient $b(S, B, u)$ and exponent $n(S, B)$ are defined in the following manner: Let S be a sequence of length k over the set $\{\circ, \Delta\}$. Let $k_\circ$ denote the number of $\circ$'s in the sequence and B a sequence of length $k_\circ$ over J_1. Let S_B denote the sequence of length k over $J_1 \cup \{\Delta\}$ in which the ith $\circ$ in S has been replaced by the ith element in the sequence B. Suppose that

$$S_B = b_{j_1} \dots b_{j_{s_1}} \Delta b_{j_{s_1+1}} \dots b_{j_{s_2}} \Delta \dots$$

with $0 \leq s_1 \leq s_2 \leq \dots \leq k_\circ$. Define $b(S, B, v)$ to be

$$\left(\prod_{i=1}^{s_1} b_{j_i}\right) \frac{(v - \sum_{i=1}^{s_1}(d - j_i) - d)}{\pi} \left(\prod_{i=s_1+1}^{s_2} b_{j_i}\right) \frac{(v - \sum_{i=1}^{s_2}(d - j_i) - 2d)}{\pi} \dots$$

and

$$n(S, B) := \sum_{i=1}^{k_\circ}(d - j_i) + (k - k_\circ)d = kd - \sum_{i=1}^{k_\circ} j_i.$$

Here v is any integer such that $v - n(S,B) > 0$. In (21) we take $v := u$, and S_B is any sequence of length $k \leq \kappa$ such that $u - n(S,B) > 0$. (In fact, b_κ is the sum $(-1)^k b(S,B,u)X^{u-n(S,B)}$ over all such sequences B_S.) We now examine such sequences in greater detail with the aim of getting a bound on the p-adic order of $b(S,B,u)$ in terms of $n(S,B)$.

Let S be a sequence over $\{\circ, \Delta\}$ which contains an even number, 2ℓ say, of Δ's and has last symbol Δ. We call such sequences *good*. Bracket the symbols uniquely by grouping the next two Δ's which occur, moving left to right, along with any leading or intervening $\circ$'s. For example, we might have the bracketing

$$(\Delta\Delta)(\circ\Delta\Delta)(\Delta\Delta)(\circ\Delta\circ\circ\Delta)(\Delta\Delta)(\Delta\Delta)(\circ\circ\Delta\Delta)(\Delta\Delta).$$

where $\ell = 8$. If more than $\lceil l/2 \rceil$ of the brackets are of the form $(\Delta\Delta)$ then there must be two consecutive brackets containing these two letters. If this is the case then rebracket some of these pairs in the form $(\Delta\Delta\Delta\Delta)$ until no more than $\lceil \ell/2 \rceil$ brackets $(\Delta\Delta)$ are left. The above string would be rebracketed

$$(\Delta\Delta)(\circ\Delta\Delta)(\Delta\Delta)(\circ\Delta\circ\circ\Delta)(\Delta\Delta\Delta\Delta)(\circ\circ\Delta\Delta)(\Delta\Delta).$$

We shall call this a Δ-bracketing of a good sequence. The Δ-bracketing of a good sequence S induces a bracketing of B_S, for suitable B, and a factorisation of $b(S,B,v)$, for suitable v. One may bound the p-adic order of $b(S,B,v)$ by obtaining bounds on the order of each term in this factorisation. To this end, we examine the sequences and p-adic numbers which occur in such decompositions. In the proofs below we shall repeatedly use the fact that d is odd and $\pi = -2$.

Lemma 27. *Let B_S be the sequence*

$$b_{j_1} \ldots b_{j_{s_1}} \Delta b_{j_{s_1+1}} \ldots b_{j_{s_2}} \Delta$$

where $s_2 \geq 0$. We have $\operatorname{ord}(b(S,B,v)) \geq -1$ for any $v > n(S,B)$. Moreover, if $s_2 > 0$ then $n(S,B) \geq 2d+1$, and for $s_2 = 0$ we have $n(S,B) = 2d$.

Proof. The claims on $n(S,B)$ are straightforward. For the other part, if $s_2 = 0$ then the sequence is of the form $\Delta\Delta$ and so $b(S,B,v) = (v-d)(v-2d)/\pi^2$ which has order at least -1. So suppose $s_2 > 0$. Since each b_j has $\operatorname{ord}(b_j) = \operatorname{ord}(j) \geq 0$ we can restrict to the subsequence $\Delta b_{j_{s_1+1}} \ldots b_{j_{s_2}} \Delta$. If $s_2 - s_1 = 0$ then this is just $\Delta\Delta$ and the claim follows as before. Thus it remains to consider sequences B_S of the form $\Delta b_{j_1} \ldots b_{j_s} \Delta$, where $s > 0$, and the corresponding $b(S,B,v)$: If $v - d$ is even we are done. So suppose that $v-d$ is odd. If all $d - j_i$ for $1 \leq i \leq s$ are even then $v - \sum_{i=1}^{s}(d-j_i) - 2d$ is even and we are done. Otherwise, at least one $d - j_i$ is odd, say for $i = m$. Then j_m is even and $\operatorname{ord}(b_{j_m}) = \operatorname{ord}(j_m) \geq 1$. The result now follows. $\square$

Lemma 28. *If S is the sequence $\Delta\Delta\Delta\Delta$ then $\operatorname{ord}(b(S,\emptyset,v)) \geq -1$ for any $v > 4d$, and $n(S,\emptyset) = 4d$.*

Proof. The claim on $n(S,\emptyset)$ is immediate. For the other part

$$b(S,\emptyset,v) = (u-d)(u-2d)(u-3d)(u-4d)/\pi^4$$

and the numerator is divisible by 2^3. □

We now put these results together.

Lemma 29. *Let $p=2$ and $G \in L_{>0}$. Suppose that $aG \in \tilde{L}(1/(2+(1/4d)))$ for some p-adic integer $a \in A$. Then we may write $G = D(r)+s$ where $r \in L$ and s is a polynomial of degree less than d with no constant term over $A \otimes \mathbb{Q}$ such that πas has coefficients in A.*

Proof. Let S be a good sequence over $\{\circ,\Delta\}$ with an even number, 2ℓ say, of Δ's and suppose we have chosen a Δ-bracketing. Let m_1 denote the number of brackets $(\Delta\Delta)$, m_2 denote the number of brackets with two Δ's but not of this form, and m_3 the number of brackets $(\Delta\Delta\Delta\Delta)$. The total number of Δ's is $2m_1+2m_2+4m_3 = 2\ell$. Since $m_1 \leq \lceil \ell/2 \rceil$ it follows that $m_1 \leq m_2+2m_3+2$. Applying Lemmas 27 and 28 we find that for any B of the correct length and suitable u

$$\begin{array}{rcl} n(S,B) & \geq & 2dm_1+(2d+1)m_2+4dm_3 \\ \mathrm{ord}(b(S,B,u)) & \geq & -(m_1+m_2+m_3). \end{array}$$

Since $m_1 \leq m_2+2m_3+2$ the ratio $n(S,B)/(-\mathrm{ord}(b(S,B,u)))$ is at least $(2d+(1/4))$, unless both m_2 and m_3 are zero. Hence provided m_2 and m_3 are not both zero we have

$$\mathrm{ord}(b(S,B,u)) \geq (1/(2d+(1/4)))(-n(S,B)). \tag{22}$$

Note that when $m_2, m_3 = 0$ we must have $S = (\Delta\Delta)$. Now let S' be an arbitrary sequence. Then it can be written as $S' = S|S''$ where $S \neq (\Delta\Delta)$ is a good sequence, and S'' is either $(\Delta\Delta)$ or a sequence with one or zero Δ's. For any suitable B'' and u'' it is easily seen that

$$\mathrm{ord}(b(S'',B'',u'')) \geq -1 \geq (1/(2d+(1/4)))(-n(S'',B''))-1.$$

Combining this with inequality (22) we find that the following holds for any B' of correct length and u' of suitable size

$$\mathrm{ord}(b(S',B',u')) \geq (1/(2d+(1/4)))(-n(S',B'))-1. \tag{23}$$

Let $c_uX^u \in \tilde{L}(1/(2+(1/4d)))$ with $\mathrm{ord}(c_u) = (1/(2+(1/4d)))\tilde{\mathrm{wt}}(u) = u/(2d+(1/4))$ and $u>0$. Then for suitably large κ we can write $X^u = D(a_\kappa)+b_\kappa$ where b_κ is a sum of terms $b(S,B,u)X^{u-n(S,B)}$ with $0 < u - n(S,B) < d$. It follows from (23) that $\mathrm{ord}(b(S,B,u)) \geq (1/(2d+(1/4)))(-u)-1$. Now $c_uX^u = D(c_ua_\kappa)+c_ub_\kappa$. By the bounds on $\mathrm{ord}(c_u)$ and $b(S,B,u)$, each term in c_ub_κ has p-adic order at least -1. Also, a_κ is obtained from $a_{\kappa-1}$ and $b_{\kappa-1}$ and an induction argument shows that c_ua_κ lies in $\pi^{-2}\tilde{L}(1/(2+(1/4d)))$.

Finally, let G be an arbitrary element of $L_{>0}$. Then it is an infinite sum with coefficients tending to zero of terms of the form c_uX^u, as above. The claims in the lemma now follow easily. □

Corollary 30. *For $p = 2$ the set $\{\pi X, \dots, \pi X^{d-1}\}$ forms a basis for $L_{>0}/D(L)$. Given any basis element e in this set, the power series $\alpha(e)$ is cohomologous to a polynomial over $A \otimes \mathbb{Q}$ with coefficients of p-adic order at least -1.*

Proof. From the computations above we know that this set spans $L_{>0}/D(L)$. Also any basis cannot have fewer than $d-1$ elements by consideration of the degree of the L-function $L(f, \Psi, T)$, and so it must be a basis??? For the second part, since $e \in \tilde{L}(1/2 + (1/4d))$ we find that $\alpha(e)$ also lies in this space. The result now follows from Lemma 29 by taking $a = 1$. □

7.3. **Computing the absolute Frobenius matrix.** We now describe how to perform the main step of the algorithm, that is, constructing the matrix for the absolute Frobenius map with respect to the basis $\{\pi X, \dots, \pi X^{d-1}\}$. First, one may compute F with the coefficients determined modulo p^N directly from the formula in Definition 8 and the expression for $\theta(t)$ in 13.) Working with coefficients modulo p^N, for each basis element e the polynomial $\psi_p \circ F(e) \bmod p^N$ may be constructed. The reduction method of Sections 7.1 and 7.2 is then used to write this as a linear combination of the basis elements $\pi X, \dots, \pi X^{d-1}$. In this way the matrix M is found with coefficients determined modulo p^N.

7.4. **Finding the characteristic polynomial of Frobenius.** One may compute the matrix M_a via the formula

$$M_a = \prod_{i=0}^{a-1} \tau^{-i}(M) = \tau^{a-1}(M)\tau^{a-2}(M)\dots M. \tag{24}$$

This is proved from Lemma 24 in the same way as [15, Lemma 26] (see also the sentence following this lemma). The characteristic polynomial may then be found deterministically by computing $\mathrm{Tr}(M_a^k)$ for $1 \le k \le d$ and using the Newton identity

$$\det(I - M_a T) = \exp\Big(-\sum_{k=1}^{\infty} \frac{\mathrm{Tr}(M_a^k)}{k} T^k\Big).$$

(Alternatively, one could use an interpolation method.) The numerator of the zeta function may now be found by computing the conjugates (10) and taking a product to get a polynomial in $\mathbb{Z}_p[T]$.

8. Complexity Analysis

8.1. **FFT, Karatsuba and classical multiplication algorithms.** Let $B := A \bmod p^N$ and P the ring of polynomials in one variable over B. Denote by c the (deterministic) exponent for multiplication in both rings: Precisely, polynomials of degree δ in P can be multiplied in $\tilde{\mathcal{O}}(\delta^c)$ operations in B, and elements in B multiplied in $\tilde{\mathcal{O}}((\log|B|)^c)$ bit operations. Using classical methods we take $c = 2$, Karatsuba's algorithm gives $c = \log_2(3) < 1.59$, and $c = 1$ using Fast Fourier Transform (FFT) methods. (Note that we ignore logarithmic factors, and the space complexity is $\tilde{\mathcal{O}}(\delta)$, $\tilde{\mathcal{O}}(\log|B|)$ in all cases. For polynomial multiplication using FFT methods we refer to [3]; we assume as in [12, Section 5] and [10, Section 4.3] that FFT methods may be applied to B, although we do not know a convenient reference for this.) Similarly, let ω denote the exponent for deterministic multiplication of matrices over B; thus two $\delta \times \delta$ matrices can be multiplied in $\tilde{\mathcal{O}}(\delta^\omega)$ operations in B. It will transpire that the choice of $\omega(\leq 3)$ does not affect the overall complexity.

8.2. **Complexity of point counting algorithm.** First we must compute F with coefficients determined modulo p^N. This may be done by multiplying together the $\mathcal{O}(d)$ polynomials $\theta(\hat{a}_j X^j)$ in the ring "$\tilde{L}(\delta) \bmod p^N$" for δ as in Lemma 9. (Note that each $\theta(t)$ can itself be constructed via multiplication of truncated power series of the form $\exp(*t^{p^j})$, as in [15, Lemma 29]. This part is dominated by the computation of F itself.) From the decay rate of coefficients of power series in $\tilde{L}(\delta)$, we see that polynomials $\theta(\hat{a}_j X^j) \bmod p^N$ are linear combinations of the monomials $\pi^{\lceil \tilde{\mathrm{wt}}(i)\delta \rceil} X^i$ for $\tilde{\mathrm{wt}}(i)$ bounded so that

$$\lceil \tilde{\mathrm{wt}}(i)\delta \rceil \leq N(p-1).$$

There are $\mathcal{O}(Npd)$ such terms. Since the coefficients of the power series $\theta(\hat{a}_j X^j)$ lie in A it follows that the construction of F may be done in

$$\tilde{\mathcal{O}}((Npd)^c (Npa\log p)^c) = \tilde{\mathcal{O}}((N^2p^2ad)^c) \tag{25}$$

bit operations. (Here $\mathcal{O}(Npa\log p)$ is the bit-size of elements in the ring $A \bmod p^N$.) Second, finding $\psi_p \circ F(e)$ for all $d-1$ basis monomials requires

$$\tilde{\mathcal{O}}(d(Npd)(Npa)) = \tilde{\mathcal{O}}(N^2p^2ad^2) \tag{26}$$

bit operations. Here we use the quasi-linear time method to compute the map τ^{-1} on $A \bmod p^N$ suggested in [12, Section 5] and [10, Step 2] (namely, precomputation of the map on each basis element ε by Newton iteration). Third, we must compute a normal form for each such expression to find the coefficients in the matrix M. By the time estimate (19), we see that each column of M can be computed in

$$\tilde{\mathcal{O}}(N^{c+1}p^{c+1}d^2a^c) \tag{27}$$

bit operations. We require $d-1 = \mathcal{O}(d)$ such computations. Fourth, computing the matrix for M_a may be done using Equation (24) and the fast exponentiation method of [15, Lemma 31] in

$$\tilde{\mathcal{O}}(d^\omega (Npa)^c) \tag{28}$$

bit operations. Finally, the computation of the characteristic polynomial takes

$$\tilde{\mathcal{O}}(d^{\omega+1}(Npa)^c) \tag{29}$$

bit operations. (Computation of the products of the conjugates using (10) is absorbed in the other estimates.) Adding (25), (26), $d \times$ (27), (28) and (29) together and putting $N = \mathcal{O}(pad)$ we get

$$\tilde{\mathcal{O}}(p^{4c}a^{3c}d^{4+c})$$

where c is the exponent for ring multiplication as discussed in Section 8.1. (Here we have assumed $\omega \leq 3$.) Using asymptotically fast methods we may take $c = 1$, with Karatsuba multiplication $c = 1.59$, and classical methods $c = 2$. We get a running time of

$$\begin{array}{ll} \tilde{\mathcal{O}}(p^4a^3d^5) & \text{with FFT methods} \\ \tilde{\mathcal{O}}(p^{6.36}a^{4.77}d^{5.59}) & \text{with Karatsuba's algorithm} \\ \tilde{\mathcal{O}}(p^8a^6d^6) & \text{with classical methods.} \end{array}$$

The space complexity is in all cases determined by the size of the ring $L \bmod p^N$ and also the polynomial $F \bmod p^N$. These are both

$$\tilde{\mathcal{O}}((Npd)(Npa)) = (p^4a^3d^3).$$

This completes the proof of Theorem 1 in the case of Type 1 Artin-Schreier curves, and also Type 2 since they are easily reduced to Type 1.

9. Artin-Schreier covers of the torus

In this section we present in a much more condensed fashion the algorithm for computing the zeta function of the Type 3 Artin-Schreier curves.

Let $f \in \mathbb{F}_q[X, X^{-1}]$ have negative degree $d^- < 0$ and positive degree $d^+ > 0$. We shall explain how to compute the L-function $L^*(f, T)$. From this one may easily compute the zeta functions of Type 3 Artin-Schreier curves using (8).

Define a weight function wt on $\mathbb{Z}$ by

$$\mathrm{wt}(u) := \begin{cases} \lceil u/|d^-| \rceil & \text{if } u < 0 \\ \lceil u/d^+ \rceil & \text{if } u \geq 0 \end{cases}$$

where $|.|$ is the usual (not p-adic) absolute value. Let L be the Banach module over A with orthonormal basis $\pi^{\mathrm{wt}(r)}X^r$ for $r \in \mathbb{Z}$. Note that L contains two-way infinite power series, and in fact still forms a ring because of the decay conditions on coefficients. Now define the power series F and $F^{(a)}$ in exactly the same manner as before, where the splitting function θ is chosen according to p. In this case they are two-way infinite power series. With α and α_a defined exactly as before, we find that both maps are stable on L. We have the formula

$$L^*(f, T) = \frac{\det(1 - T\alpha_a | L)}{\det(1 - Tq\alpha | L)}.$$

In this case one cannot remove any unit root factors. Let $\mathcal{L}$ be the (slightly "larger") complex of A-modules

$$0 \longrightarrow L \xrightarrow{D} L \longrightarrow 0.$$

Here D is defined in exactly the same manner as before. In this case $H_0 = L/D(L)$ and $H_1 = 0$, and one recovers the cohomological trace formula

$$L^*(f,T) = \det(1 - H_0(\alpha_a)T).$$

Once again we have the crucial relation $H_0(\alpha_a) = H_0(\alpha)^a$. A basis for H_0 can be taken as

$$\{\pi X^{d^-}, \pi X^{d^-+1}, \ldots, \pi X^{-1}, 1, \pi X, \ldots, \pi X^{d^+-1}\}$$

with $d^+ - d^-$ the dimension of this space. Computation of normal forms in $L/D(L)$ is done in a similar manner to before. Precisely, for $p > 2$ one first uses the method in Section 7.1 to find an element cohomologous to a given element whose leading term has degree less than d^+. One then performs a similar process to increase the degree of the lowest term so that it is not less than d^-. For $p = 2$ the process is similar, but small denominators are again introduced. Note that all computations are done modulo p^N for any N greater than $(p-1)(d^+ - d^-)(a/2+1)$ (plus 1?? if $p = 2$). The complexity of the above algorithm may be checked to be identical to that in the case when f is just an ordinary polynomial. Precisely, if we write "$d := d^+ - d^-$" then all the bounds in Section 8 are still true. This completes the proof of Theorem 1, and Corollary 2 follows from Note 5.

References

[1] A. Adolphson and S. Sperber, Exponential sums and Newton polyhedra: Cohomology and estimates, Ann. Math. 130, (1989) 367-406.

[2] I. Blake, G, Seroussi, N. Smart, Elliptic Curves in Cryptography, LMS Lecture Note Series 265, Cambridge University Press, 1999.

[3] D. Cantor and E. Kaltofen, On fast multiplication of polynomials over arbitrary algebras, Acta Inform. 28 (7), (1991), 693-701.

[4] R.F. Coleman, P-adic Banach spaces and families of modular forms, Invent. Math. 127, (1997), 417-479.

[5] B. Dwork, On the rationality of the zeta function of an algebraic variety, Amer. J. Math., 82, (1960), 631-648.

[6] B. Dwork, On the zeta function of a hypersurface, Pub. Math. IHES No. 12, 1962.

[7] N. Elkies, Elliptic and modular curves over finite fields and related computational issues, in "*Computational perspectives in number theory: Proceedings of a conference in honour of A.O.L. Atkin*", (D.A. Buell and J.T. Teitelbaum), American Mathematical Society International Press 7, 1998, 21-76.

[8] M. Fouquet, P. Gaudry and R. Harley, An extension of Satoh's algorithm and its implementation, J. Ramanujan Math. Soc. 15, (2000), 281-318.

[9] M. Fouquet, P. Gaudry and R. Harley, Finding secure curves with the Satoh-FGH algorithm and an early abort strategy, in B. Pfitzmann (ed), Advances in Cryptology - EUROCRYPT 2001, Lecture Notes in Computer Science 2045, Springer-Verlag, (2001), 14-29.

[10] P. Gaudry and N. Gürel, An extension of Kedlaya's algorithm for counting points on superelliptic curves, preprint.

[11] P. Gaudry and R. Harley, Counting points on hyperelliptic curves over finite fields, in B. Preneel (ed), Advances in Cryptology - EUROCRYPT 2000, Lecture notes in Computer Science 1807, Springer-Verlag, (2000), 19-34.

[12] K.S. Kedlaya, Counting points on hyperelliptic curves using Monsky-Washnitzer cohomology, preprint 2001.

[13] N. Koblitz, p-Adic Numbers, p-Adic Analysis and Zeta Functions, 2nd Edition, Springer, 1984.
[14] N. Koblitz, Hyperelliptic cryptosystems, J. of Cryptology 1, (1989), 139-150.
[15] A.G.B. Lauder and D. Wan, Counting points on varieties over finite fields of small characteristic, preprint 2001.
[16] R. Lidl and H. Niederreiter, Introduction to Finite Fields and Their Applications, Cambridge University Press, 1986.
[17] B. Poonen, Computational aspects of curves of genus at least 2, in "*Algorithmic Number Theory II*" (H. Cohen), Lecture Notes in Computer Science 1122, Springer, 1996, 283-306.
[18] T. Satoh, The canonical lift of an ordinary elliptic curve over a finite fields and its points counting, J. Ramanujan Math. Soc. 15, (2000), 247-270.
[19] R. Schoof, Elliptic curves over finite fields and the computation of square roots mod p, Math. Comp. 44, no. 170, (1985), 483-494.
[20] J.-P. Serre, Endomorphisms complètement continus des espaces de Banach p-adique, Pub. Math. IHES 12, (1962), 69-85.
[21] S. Sperber, On the p-adic theory of exponential sums, Amer. J. Math. 108, (1983), 255-296.
[22] F. Vercauteren, B. Preneel and J. Vandewalle, A memory efficient version of Satoh's algorithm, in B. Pfitzmann (ed) Advances in Cryptology - EUROCRYPT 2001, Lecture Notes in Computer Science 2045, Springer-Verlag, (2001), 1-13.
[23] D. Wan, Computing zeta functions over finite fields, Contemporary Mathematics, 225 (1999), 131-141.
[24] D. Wan, Pure L-functions from algebraic geometry over finite fields, in "Finite Fields and Applications" D. Jungnickel and H. Niederreiter (eds), Springer-Verlag, (2000), 437-461.
[25] D. Wan, Algorithmic theory of zeta functions over finite fields, preprint 2001.

Computing Laboratory, Oxford University, Oxford OX1 3QD, UK *E-mail address*: alan.lauder@comlab.ox.ac.uk

Department of Mathematics, University of California, Irvine, CA 92697, USA *E-mail address*: dwan@math.uci.edu