On the discrete logarithm problem for plane curves

Claus Diem

June 10, 2011

Abstract

In this article the discrete logarithm problem in degree 0 class groups of curves over finite fields given by plane models is studied. It is proven that the discrete logarithm problem in degree 0 class groups of non-hyperelliptic curves of genus 3 (given by plane models of degree 4) can be solved in an expected time of $\tilde{\mathcal{O}}(q)$, where q is the cardinality of the ground field. Moreover, it is proven that for every fixed natural number $d \geq 4$ the following holds: We consider the discrete logarithm problem for curves given by plane models of degree d for which there exists a line which defines a divisor which splits completely into distinct $\mathbb{F}_q$-rational points. Then this problem can be solved in an expected time of $\tilde{\mathcal{O}}(q^{2-\frac{2}{d-2}})$. This holds in particular for curves given by reflexive plane models.

1 Introduction

This article is concerned with the complexity of the discrete logarithm problem in degree 0 class groups of curves over finite fields. (Unless stated otherwise, a curve is always assumed to be proper, non-singular and geometrically irreducible.) In various works on the subject, the complexity of the computation is expressed in terms of the genus and the cardinality of the ground field. For example, it is proven in [3] that for any fixed $g \in \mathbb{N}, g \geq 2$, the discrete logarithm problem in degree 0 class groups of curves of genus g can be solved in an expected time of

$$\tilde{\mathcal{O}}(q^{2-\frac{2}{g}}) \ .$$

Here and in the following, q is the cardinality of the ground field.

In this article, we study the complexity of the problem from a different point of view: We assume that the curve is given by a plane model, by which we mean a possibly singular plane curve which is birational to the curve in question. We then express the complexity in terms of the degree of the model and the cardinality of the ground field.

Our first result concerns non-hyperelliptic genus 3 curves. Via the canonical embedding, every such curve can be given as a plane curve of degree 4. By using such a model, we obtain the following result:

Theorem 1 *The discrete logarithm problem in the degree 0 class groups of non-hyperelliptic curves of genus 3 can be solved in an expected time of* $\tilde{\mathcal{O}}(q)$.

To state our second result, we need some notation: First, we set $\mathbb{P}^2_{\mathbb{F}_q} := \mathrm{Proj}(\mathbb{F}_q[X,Y,Z])$. Let $\mathcal{C}$ be a curve over $\mathbb{F}_q$, $\mathcal{C}_{pm}$ a plane model of $\mathcal{C}$, and $\pi : \mathcal{C} \longrightarrow \mathcal{C}_{pm}$ a birational morphism. Let $\mathcal{O}_{\mathcal{C}}(1) := \pi^*(\mathcal{O}(1))$, let for a linear form $W \in \mathbb{F}_q[X,Y,Z]_1 = \Gamma(\mathbb{P}^2_{\mathbb{F}_q}, \mathcal{O}(1))$ $W_{|\mathcal{C}} := \pi^*(W) \in \Gamma(\mathcal{C}, \mathcal{O}_{\mathcal{C}}(1))$, and let $\mathrm{div}(W_{|\mathcal{C}}) = \pi^*(\mathrm{div}(W))$ be the divisor of zeroes of $W_{|\mathcal{C}}$ on $\mathcal{C}$. Now let $\mathfrak{d}$ be the linear system on $\mathcal{C}$ "cut out by lines" (more precisely: obtained by pull-back of lines), that is,

$$\mathfrak{d} := \{ \mathrm{div}(W_{|\mathcal{C}}) \mid W \in \mathbb{F}_q[X,Y,Z]_1 = \Gamma(\mathbb{P}^2_{\mathbb{F}_q}, \mathcal{O}(1)) \} .$$

We say that an effective divisor *splits completely* if its support contains only $\mathbb{F}_q$-rational points.

The second result is as follows:

Theorem 2 *Let $d \geq 4$ be fixed. Then the discrete logarithm problem in the degree 0 class groups of curves given by plane models of degree d such that the linear system $\mathfrak{d}$ contains a divisor which splits completely into distinct points can be solved in an expected time of*

$$\tilde{\mathcal{O}}(q^{2-\frac{2}{d-2}}) .$$

We also show the following proposition on curves given by plane models, a result we consider to be of independent interest.

Proposition 1 *Let $d \geq 3$ be fixed. We consider curves of genus ≥ 1 given by plane models of degree d, where for $d > 4$ we restrict ourselves to reflexive plane models. Then the number of divisors in $\mathfrak{d}$ which split completely into distinct points is in $\frac{1}{d!}q^2 + \mathcal{O}(q^{\frac{3}{2}})$.*

Briefly, reflexivity means that the classical duality theory holds. We note that for characteristic 2, no plane model is reflexive, and if the degree is larger than the characteristic, every plane model is reflexive. More information on the notion of reflexivity can be founded below.

Theorem 2 and the proposition give:

Theorem 3 *Let $d \geq 4$ be fixed. Then the discrete logarithm problem in the degree 0 class groups of curves given by reflexive plane models of degree d can be solved in an expected time of*

$$\tilde{\mathcal{O}}(q^{2-\frac{2}{d-2}}) .$$

Let us consider curves of a fixed genus g given by plane models of a fixed degree d such that $\mathfrak{d}$ contains a divisor which splits completely into distinct points. Then under the condition that $g \geq d-1$, the second theorem improves upon the result of [3] mentioned above. The algorithms then have storage requirements of $\tilde{\mathcal{O}}(q^{1-\frac{1}{g}+\frac{1}{(d-2)\cdot g}})$.

The underlying computational model is throughout a randomized random access machine with logarithmic cost function. We refer the reader to [4] for a discussion on random access machines.

Reflexive plane models

We briefly review the notion of reflexivity and give some characterizations of a plane model being reflexive.

Let k be some field which for the moment we assume to be algebraically closed. Let X be an irreducible closed subvariety of $\mathbb{P}^n_k$, and let X_{ns} be the non-singular (=smooth) part of X. The *conormal variety* $C(X) \subseteq \mathbb{P}^n_k \times \mathbb{P}^{n*}_k$ of X is the closure of pairs (P, H), where P is a closed point of X_{ns} and H is a hyperplane of $\mathbb{P}^n_k$ which meets X in P tangentially. The *dual variety* X^* of X is the image of $C(X)$ in $\mathbb{P}^{n*}_k$. Analogous definitions can be made for subvarieties of $\mathbb{P}^{n*}_k$. Now X is called *reflexive* if – up to change of factors of $\mathbb{P}^n_k \times \mathbb{P}^{n*}_k$ and identification of $\mathbb{P}^{n**}_k$ with $\mathbb{P}^n_k$ – $C(X) = C(X^*)$. Reflexive varieties are (trivially) bidual, that is, $X^{**} = X$. In characteristic 0, every variety X is reflexive, but this is not anymore the case if the characteristic is positive.

We are interested in the case that $n = 2$ and X is one dimensional and not a line. Let these conditions be satisfied, and let $\rho : C(X) \longrightarrow X^*$ be the canonical map to the dual variety. The following conditions are equivalent:

- X is not reflexive.
- ρ is not birational.
- $\rho^* : k(X^*) \longrightarrow k(C(X))$ is inseparable.
- For sufficiently general pairs (P, H), where P is a closed point on $\mathcal{C}$ and H is a hyperplane which meets $\mathcal{C}$ in P tangentially, the intersection number of H and X at P is equal to the degree of inseparability of ρ^*.
- $\mathrm{char}(k) = 2$ or for sufficiently general pairs (P, H), where P is a closed point on $\mathcal{C}$ and H is a hyperplane which meets $\mathcal{C}$ in P tangentially, the intersection number of H and X at P is > 2.
- Let $X = V(F)$ for some irreducible homogeneous polynomial $F(X_1, X_2, X_3)$. Then all polynomials $F_i^2 F_{jj} + F_j^2 F_{ii} - 2F_i F_j F_{ij}$ for

$i, j = 1, 2, 3$ vanish on X. Here for a polynomial $f \in k[X_1, X_2, X_3]$ and $i = 1, 2, 3$, $f_i = \frac{\partial f}{\partial X_i}$.

The equivalence of the first three items is called the *Segre-Wallace criterion*; a proof can be found in [18] or [11]. The further equivalences are established in [9] and [8].

By the fifth (or sixth) item, if $\mathrm{char}(k) = 2$, then X cannot be reflexive. Other examples of non-reflexive varieties in the projective plane are the so called strange curves, that is, 1 dimensional irreducible varieties whose tangents all pass through a common point. In this case the dual variety is a line, and the bidual is a point.

Moreover, by the forth item, if $\mathrm{char}(k) > \deg(X)$, then X is reflexive.

Let now k be a perfect field, and let X be a geometrically irreducible and geometrically reduced closed subvariety of $\mathbb{P}^n_k$. Then $C(X_{\overline{k}})$ and $(X_{\overline{k}})^*$ descent to k as subvarieties of the corresponding surrounding spaces, and ρ descends too. The statements given above then still hold.

Notation and representation

Let us fix some notation we use throughout the article, and we discuss the representation of the basic objects for algorithmic purposes.

As already mentioned, the input curve $\mathcal{C}/\mathbb{F}_q$ is represented by a plane model $\mathcal{C}_{pm}$ in $\mathbb{P}^2_{\mathbb{F}_q}$ of a fixed degree $d \geq 4$. Concretely, we assume that $\mathcal{C}_{pm}$ (and therefore also $\mathcal{C}$) is given by a homogeneous polynomial $F(X, Y, Z) \in \mathbb{F}_q[X, Y, Z]$. Let $\pi : \mathcal{C} \longrightarrow \mathcal{C}_{pm}$ be a fixed birational morphism to the plane model. We also denote the composition of π with the inclusion $\mathcal{C}_{pm} \hookrightarrow \mathbb{P}^2_{\mathbb{F}_q}$ by π. Moreover, we denote the non-singular part of $\mathcal{C}_{pm}$ by $\mathcal{C}_{ns}$ (rather than by $(\mathcal{C}_{pm})_{ns}$), and we identify $\mathcal{C}_{ns}$ with its preimage in $\mathcal{C}$. We call the dual variety of $\mathcal{C}_{pm}$ the *dual model* and denote it by $\mathcal{C}^*_{pm}$.

To represent divisors on $\mathcal{C}$, we use an ideal theoretic representation, following [10]. Recall that in [10] divisors are represented by ideals of two orders of the function field of $\mathcal{C}$: a so-called finite and a so-called infinite order. We call this the *joint ideal representation*. Alternatively, one can use the *free ideal representation*, where the prime divisors are represented by prime ideals of these orders and the ideals themselves are represented by formal sums of the prime divisors (in sparse representation).

For q large enough, $\mathcal{C}$ has an $\mathbb{F}_q$-valued point. To represent the input elements, we fix a point $P_0 \in \mathcal{C}(\mathbb{F}_q)$ and represent the divisor classes by along P_0 reduced divisors. Here we can use either the free or the joint representation of divisors. Later in the algorithm the free representation of divisors is crucial.

For more information on these issues, in particular on computational aspects, we refer to [10], [2, Chapter 2] and [3, Section 2].

Overview and historical comments

We give an overview on the proof of the theorems.

The algorithms follow the *index calculus* strategy. Briefly, this means: First a so-called *factor base* is fixed. This is a set of prime divisors (closed points); in our case this is a subset $\mathcal{F}$ of $\mathcal{C}(\mathbb{F}_q)$. Now in a basic index calculus, one searches for relations between input elements, factor base elements and maybe some further elements of $\mathrm{Cl}^0(\mathcal{C})$. If enough relations have been obtained, one eliminates the factor base elements and tries to obtain a non-trivial relation between input elements. From this relation one then tries to derive the sought-after discrete logarithm.

For curves of a fixed genus ≥ 3, it has already been argued in [16] that by using a *large prime variation*, one can obtain a reduction in the expected running time which is superpolynomial in the input length. A further reduction can be obtained by using a *double large prime variation*; this has been studied in [6], [13] and [3].

In a double large prime variation, one fixes a set $\mathcal{L}$ of so called *large primes*; in our case this is $\mathcal{C}(\mathbb{F}_q) - \mathcal{F}$. Then one searches for relations involving up to two large primes. Such relations are stored in a graph on the set of vertices $\mathcal{L} \,\dot\cup\, \{*\}$. Here a relation involving one large prime P is stored as a labeled edge between $*$ and P, and a relation involving two distinct large primes P and Q is stored as a labeled edge between P and Q. Later, this graph is then used to generate relations between input elements and factor base elements.

The result in [3] given at the beginning of this article is proven with the construction of a tree of large prime relations (an idea which goes back to [6]). Moreover, in order to control the depth of the tree, similarly to the algorithm in [13], the tree is constructed in stages.

The algorithms of this work also use a tree of large prime relations, which is again constructed in stages. The essential difference to the algorithm in [3] is that we construct the tree and also the factor base in a different way: We generate relations by intersecting the plane model with lines. Let $D_\infty := \mathrm{div}(Z_{|\mathcal{C}})$. The crucial (but trivial) observation is that if $D = \sum_{P \in \mathcal{C}} n_P P$ is a divisor in the linear system $\mathfrak{d}$, then D is linearly equivalent to D_∞, and we have a relation

$$\sum_{P \in \mathcal{C}} n_P [P] = [D_\infty] \,. \tag{1}$$

In [1] we gave an algorithm in which a graph of large prime relations is generated by intersecting the curve with lines running trough two points of the factor base. On a heuristic basis, we already argued that for any fixed $d \geq 4$, one can with this algorithm solve the discrete logarithm problem in degree 0 class groups of curves given by plane models of degree d in an expected time of

$$\tilde{\mathcal{O}}(q^{2-\frac{2}{d-2}}) \ .$$

Further evidence, including experimental data, that the result is valid for non-hyperelliptic curves of genus 3 (given by plane curves of degree 4) is given in [5]. However, even in this restricted case, no proof has been given until now.

In order to obtain the two theorems mentioned above, we modify the algorithm in [1] in some ways. As already mentioned, we employ a stage-wise construction of a tree of large prime relations. In contrast, in [1] we first constructed a "full" graph of large prime relations. Moreover, during the construction of the tree, we also repeatedly enlarge the factor base (and shrink the set of large primes). The enlargements are done at the beginning of each stage in a randomized manner; we perform these enlargements to generate new randomness for each stage of the construction of the tree. To our knowledge, such enlargements of the factor base have not been suggested before.

The rest of this work is organized as follows: In the next section, we establish asymptotic results on the number of divisors in $\mathfrak{d}$ which split completely into distinct points. In particular, we prove Proposition 1. In the third section, we give the algorithm. By our previous work [3], we only have to give an algorithm for the construction of a suitable tree of large prime relations. The final subsection of this section contains the analysis of this algorithm, based on combinatorial and probabilistic techniques.

2 Estimates on completely split divisors

The purpose of this subsection is to give asymptotic estimates on the number of completely split divisors in $\mathfrak{d}$ which split completely into distinct points. We establish two results. The first one is an asymptotic lower bound under the condition that there exists at least one such divisor. The second result is Proposition 1.

We consider curves $\mathcal{C}/\mathbb{F}_q$ represented by plane models of degree a fixed degree d. For the moment, we do not make any assumption on the plane models.

The following definition is convenient:

Definition 2 Let X be an infinite countable set and $(a_x)_{x\in X}, (b_x)_{x\in X} \in \mathbb{R}^X$ with $b_x \neq 0$ for all $x \in X$. Then we write

$$a_x \gtrsim b_x$$

if $\liminf\limits_{x\in X} \dfrac{a_x}{b_x} \geq 1$.

Remark 3 In the following applications of this definition, the elements of X consist of isomorphism classes of the following data: A curve $\mathcal{C}/\mathbb{F}_q$ with a fixed map $\pi : \mathcal{C} \longrightarrow \mathbb{P}^2_{\mathbb{F}_q}$ which is birational onto its image and additionally a tuple of points in $\mathcal{C}(\mathbb{F}_q)$ or a subset $S \subseteq \mathcal{C}(\mathbb{F}_q)$ (or both). Isomorphisms are isomorphisms of curves over $\mathbb{F}_q$ respecting the maps to the projective plane and additionally the points or the subset.

Let such a set X be fixed. Let for some element $x \in X$ q_x be the cardinality of the ground field. If for some prime power q and some $x \in X$ we have $q_x = q$, we say that x is *over* $\mathbb{F}_q$. Now let $(a_x)_{x\in X}, (b_x)_{x\in X} \in \mathbb{R}^X$ with $b_x \neq 0$ for all $x \in X$. As for each prime power q there are only finitely many elements in X over $\mathbb{F}_q$ we have: $\liminf_{x\in X} \frac{a_x}{b_x} \geq 1$ if and only if there exists a function f from the set of prime powers to $\mathbb{R}$ which converges to 1 such that $\frac{a_x}{b_x} \geq f(q_x)$ for all x. In this case, it is also justified to state that $\frac{a_x}{b_x} \gtrsim 1$ for $q \longrightarrow \infty$.

Similar considerations also apply to the use of the $\mathcal{O}$-notation, etc.

Asymptotic lower bounds

Proposition 4 *For $P \in \mathcal{C}(\mathbb{F}_q)$ such that there exists a divisor in $\mathfrak{d}$ containing P which splits completely into distinct points, the number of divisors in $\mathfrak{d}$ which split completely into distinct points and contain P is $\gtrsim \frac{1}{(d-1)!} \cdot q$.*

Proof. Let P be such a point. Let $c : \mathcal{C} \longrightarrow \mathbb{P}^1_{\mathbb{F}_q}$ be defined by the central projection with center P. (c is unique up to an automorphism of $\mathbb{P}^1_{\mathbb{F}_q}$.) Then c is a covering of degree $\leq d-1$. (The degree is $d-1$ if and only if P does not lie over a singular point of $\mathcal{C}_{pm}$.) Note that by adding $\pi^{-1}(\pi(P))$, we get a bijection between the pull-backs of the $\mathbb{F}_q$-rational points of $\mathbb{P}^1_{\mathbb{F}_q}$ to $\mathcal{C}$ and the divisors of $\mathfrak{d}$ containing P.

We denote points on curves and the corresponding places of function fields in the same way, and we now consider the extension of function fields $\mathbb{F}_q(\mathcal{C})|\mathbb{F}_q(\mathbb{P}^1)$ corresponding to c. Now $c(P)$ is unramified and completely split in $\mathbb{F}_q(\mathcal{C})$. The fact that it is unramified implies that the extension is separable; let M be a Galois closure.

Recall that a place of degree 1 of $\mathbb{F}_q(\mathbb{P}^1)$ splits completely in $\mathbb{F}_q(\mathcal{C})$ if and only if it splits completely in M.

This implies that $c(P)$ splits completely in M, and this implies that $\mathbb{F}_q$ is the exact constant field of M. Now with the effective Chebotaryov density theorem from [12], we conclude that the number of places of $\mathbb{F}_q(\mathbb{P}^1)$ of degree 1 which split completely in M (or in $\mathbb{F}_q(\mathcal{C})$) is in $\frac{1}{\deg(c)!}q + \mathcal{O}(q^{\frac{1}{2}})$.

This gives the proposition. □

Proposition 5 *Let us assume that there exists at least one divisor in $\mathfrak{d}$*

which splits completely into distinct points. Then there are $\gtrsim \frac{1}{d\cdot(d-2)!\cdot(d-1)!}\cdot q^2$ *such divisors.*

Proof. Let $P \in \mathcal{C}(\mathbb{F}_q)$ be a point which is contained in a divisor which splits completely into distinct points. By the previous proposition, we have $\gtrsim \frac{1}{(d-1)!}\cdot q^2$ divisors which split completely into distinct points and contain P. Altogether, these divisors contain $\gtrsim \frac{d-1}{(d-1)!}\cdot q^2 = \frac{1}{(d-2)!}\cdot q^2$ points of $\mathcal{C}(\mathbb{F}_q)$. Now for each point in such a divisor, we apply the previous proposition again. We obtain in this way $\gtrsim \frac{1}{(d-2)!\cdot(d-1)!}\cdot q^2$ distinct tuples $(Q, D) \in \mathcal{C}(\mathbb{F}_q) \times \mathfrak{d}$, where D splits completely into distinct points and contains Q. This gives $\gtrsim \frac{1}{d\cdot(d-2)!\cdot(d-1)!}\cdot q^2$ divisors in $\mathfrak{d}$ which split completely into distinct points. □

Proof of Proposition 1

Let us explain the general strategy for the proof.

For each point $P \in \mathcal{C}_{ns}(\mathbb{F}_q)$ we wish to estimate the number of divisors in $\mathfrak{d}$ which split completely into distinct points and contain P. For this, we proceed similarly to the proof of Proposition 4.

We consider a covering $c : \mathcal{C} \longrightarrow \mathbb{P}^1_{\mathbb{F}_q}$ defined by the central projection with center P. This is a covering of degree $d-1$, and by adding P, we get a bijection between the pull-backs of the $\mathbb{F}_q$-rational points of $\mathbb{P}^1_{\mathbb{F}_q}$ to $\mathcal{C}$ and the divisors of $\mathfrak{d}$ containing P. We are therefore interested in the number of $\mathbb{F}_q$-rational points Q of $\mathbb{P}^1_{\mathbb{F}_q}$ which are unramified under the covering c such that $c^{-1}(Q)$ splits into distinct $\mathbb{F}_q$-rational points of $\mathcal{C}$ where none of these points is equal to P.

As above, we consider the extension of function fields $\mathbb{F}_q(\mathcal{C})|\mathbb{F}_q(\mathbb{P}^1)$ corresponding to c. A first *necessary* condition in order that there is any $\mathbb{F}_q$-rational point Q of $\mathbb{P}^1_{\mathbb{F}_q}$ such that $c^{-1}(Q)$ splits completely into distinct distinct $\mathbb{F}_q$-rational points is that the extension of function fields is separable. This condition is satisfied if and only if there are only finitely many closed points of $\mathbb{P}^1_{\mathbb{F}_q}$ which are ramified with respect to c.

Let us now assume that the extension is indeed separable. Let M be a Galois closure of the extension. Recall again that a place of $\mathbb{F}_q(\mathbb{P}^1)$ splits completely in $\mathbb{F}_q(\mathcal{C})$ if and only if it splits completely in M.

As already mentioned above, a second *necessary* condition in order that there is any place of degree 1 of $\mathbb{F}_q(\mathbb{P}^1)$ which splits completely in M is that $\mathbb{F}_q$ is the exact constant field of M. On the other hand, if this condition is satisfied, by the effective Chebotaryov density theorem from [12], the number of such places is in $\frac{1}{(d-1)!}\cdot q + \mathcal{O}(q^{\frac{1}{2}})$.

The condition that $\mathbb{F}_q$ is the exact constant field of M is satisfied if and only if $[M : \mathbb{F}_q(\mathbb{P}^1)] = [\overline{\mathbb{F}}_q M : \overline{\mathbb{F}}_q(\mathbb{P}^1)]$, and this is in particular satisfied if $[\overline{\mathbb{F}}_q M : \overline{\mathbb{F}}_q(\mathbb{P}^1)] = (d-1)!$, that is, $\mathrm{Gal}(\overline{\mathbb{F}}_q M|\overline{\mathbb{F}}_q(\mathbb{P}^1)) \approx S_{d-1}$.

Now the argument is different according to whether $d = 4$ or $d > 4$ and the plane model is assumed to be reflexive.

$\boldsymbol{d = 4}$

This case was already considered in [5]. Note first that as the degree of a covering c as above is prime and the genus of $\mathcal{C}$ is ≥ 1 by assumption, the extension of function fields is indeed separable (see [7, Proposition 2.5]).

The essential observation is now: If $c_{\overline{\mathbb{F}}_q} : \mathcal{C}_{\overline{\mathbb{F}}_q} \longrightarrow \mathbb{P}^1_{\overline{\mathbb{F}}_q}$ does not have a non-trivial automorphism, then the corresponding extension of function fields $\overline{\mathbb{F}}_q(\mathcal{C})|\overline{\mathbb{F}}_q(\mathbb{P}^1)$ is not Galois, and therefore $\mathrm{Gal}(\overline{\mathbb{F}}_q M|\overline{\mathbb{F}}_q(\mathbb{P}^1)) \approx S_3$.

We therefore obtain:

Proposition 6 *Let $P \in \mathcal{C}(\mathbb{F}_q)$ be such that the corresponding covering $\mathcal{C}_{\overline{\mathbb{F}}_q} \longrightarrow \mathbb{P}^1_{\overline{\mathbb{F}}_q}$ over $\overline{\mathbb{F}}_q$ does not have a non-trivial automorphism. Then the number of divisors in $\mathfrak{d}$ which split completely into distinct points is in*

$$\frac{1}{6} \cdot q + \mathcal{O}(q^{\frac{1}{2}}) \ .$$

As we assumed that genus of $\mathcal{C}$ be ≥ 1, the number of automorphisms of degree 3 of $\mathcal{C}$ is in $\mathcal{O}(1)$. So the number of points of $\mathcal{C}(\mathbb{F}_q)$ for which the assumption does not hold is in $\mathcal{O}(1)$. The proposition then easily implies Proposition 1.

$\boldsymbol{\mathcal{C}_{pm}}$ reflexive

We use the following general result:

Proposition 7 *Let $L|K$ be a finite separable extension of fields of degree n such that $L|K$ does not contain an intermediate field distinct from K and L and there is a discrete place Q of K which splits in L in the form $2P_1 + P_2 + P_3 + \cdots + P_{n-1}$ for distinct places P_i of L.[1] Then the monodromy group of $L|K$, that is, the Galois group of a Galois closure of $L|K$, is isomorphic to S_n.*

Sketch of a proof. Let M be a Galois closure of the extension. Then $\mathrm{Gal}(M|K)$ acts on the set of embeddings of L into M; we consider $\mathrm{Gal}(M|K)$ as a permutation group on this set. This operation is of course transitive. Moreover, the condition that $L|K$ does not contain a proper subfield is equivalent to the permutation group being primitive.

[1] The additive notation is of course unusual in this general setting. We use it here for consistency.

Now the non-trivial automorphism of $M|K$ which fixes P_1 acts as a transposition. We therefore have a transitive primitive permutation group with a transposition. With [19, Theorem 13.3] we conclude that the group is the full symmetric group. □

From this general proposition the following result follows immediately:

Proposition 8 *Let $\overline{k}$ be an algebraically closed field, and let $L|\overline{k}(x)$ be a finite extension of degree n. Let us assume that there are only finitely many places of $\overline{k}(x)$ over $\overline{k}$ which are ramified in L and that every such place splits in L as $2P_1 + P_2 + P_3 + \cdots + P_{n-1}$ for distinct places P_i. Then the extension $L|\overline{k}(x)$ is separable and its monodromy group is isomorphic to S_n.*

Proof. Under the given conditions the extension is obviously separable. Moreover, it cannot contain an intermediate field different from L and $\overline{k}(x)$. For, let N be an intermediate field distinct from $\overline{k}(x)$. As every finite extension of $\overline{k}(x)$ is ramified, there exists a place Q of the function field $\overline{k}(x)$ which is ramified in N. Let us fix such a place, let R be such a ramified place of N over Q, and let r be the ramification degree. Furthermore, let $a := [L : N]$. Then the conorm of P in L has the form $rD + \tilde{D}$ for effective divisors $D, \tilde{D}$ of the function field L with $\deg(D) = a$. By our assumption it follows that $a = 1$ (and $r = 2$).

Now the statement follows with the previous proposition. □

We now make use of the classical theory of plane curves (geometrically irreducible and geometrically reduced 1-dimensional varieties in the projective plane in our terminology), including duality theory. A good reference for this classical theory in characteristic 0 is [17], the key statements we need for reflexive curves in positive characteristic can be found in [8].

For a closed point P of $\mathcal{C}_{\overline{\mathbb{F}}_q}$, the multiplicity of the divisor $\pi^{-1}(\pi(P))$ at P is called the *order* of P (with respect to fixed plane model $\mathcal{C}_{pm}$ and the map π).

Now for each closed point P of $\mathcal{C}_{\overline{\mathbb{F}}_q}$ (including points lying over singular points of $(\mathcal{C}_{pm})_{\overline{\mathbb{F}}_q}$) there is exactly one line L in $\mathbb{P}^2_{\overline{\mathbb{F}}_q}$ such that the multiplicity of the divisor $\pi^{-1}(L)$ at P is larger than the order of P. Following [17, IV, 5.3], we call this line the *tangent* at P. In [17] the tangent at P is characterized as the tangent to the local branch of $(\mathcal{C}_{pm})_{\overline{\mathbb{F}}_q}$ corresponding to P. There is the following alternative proof of existence and uniqueness of the tangent without power series which is important for our applications: We consider a covering $c : \mathcal{C}_{\overline{\mathbb{F}}_q} \longrightarrow \mathbb{P}^1_{\overline{\mathbb{F}}_q}$ defined by central projection with center $\pi(P)$. There exists exactly one closed point Q of $\mathbb{P}^1_{\overline{\mathbb{F}}_q}$ whose pull-back to $\mathcal{C}_{\overline{\mathbb{F}}_q}$ contains P. Now exactly for this point Q, the multiplicity of P in the divisor $\pi^{-1}(\pi(P)) + c^{-1}(Q) \in \mathfrak{d}$ is larger than the order of P. The line which defines this divisor is the tangent at P.

If L is the tangent at P, the multiplicity of $\pi^{-1}(L)$ at P is called the *class* of P; cf. [17]. If $\pi(P)$ is non-singular, P is called a *flex point* if and only if the class of P is greater than 2.

Let $\tau : \mathcal{C} \longrightarrow \mathcal{C}^*_{pm}$ be the canonical map from $\mathcal{C}$ to the dual model associated to $\mathcal{C}_{pm}$ and π. This means that for every closed point P of $\mathcal{C}_{\overline{\mathbb{F}}_q}$, $\tau(P)$ is the point corresponding to the tangent at P.

A tangent is called *ordinary* if it is the tangent of exactly one closed point of $\mathcal{C}_{\overline{\mathbb{F}}_q}$ and the intersection multiplicity at this point is 2 (that is, the class of the point is 2).

Lemma 9 *Let P be a closed point of $(\mathcal{C}_{ns})_{\overline{\mathbb{F}}_q}$. Then the tangent at P is ordinary if and only if the point $\tau(P)$ is a non-singular point of the dual model $(\mathcal{C}^*_{pm})_{\overline{\mathbb{F}}_q}$.*

Proof. It is obvious that more than one point of $\mathcal{C}_{\overline{\mathbb{F}}_q}$ lies over $\tau(P)$ if and only if the tangent at P is also the tangent of another point of $\mathcal{C}_{\overline{\mathbb{F}}_q}$.

We claim that P is a flex point of $(\mathcal{C}_{pm})_{\overline{\mathbb{F}}_q}$ if and only if the order of P with respect to $(\mathcal{C}^*_{pm})_{\overline{\mathbb{F}}_q}$ and τ is greater than 1.

Let s be the class of P. As P is non-singular, there exist homogeneous coordinates such that with respect to this coordinate system P is given by $[0:0:1]$ and a reduced parametrization

$$[t \,:\, b_s t^s + b_{s+1}t^{s+1} + b_{s+2}t^{s+2} + \cdots \,:\, 1]\ ,$$

where $b_s \neq 0$. Now the image of this $\overline{\mathbb{F}}_q((t))$-valued point in the dual curve is given by the vector product of the parametrization and its derivative; this is

$$[b_s s t^{s-1} + \cdots \,:\, -1 \,:\, b_s(1-s)t^s + \cdots]\ .$$

(The corresponding result in characteristic 0 is a classical result from duality theory of curves. As shown in [8], the result also holds in positive characteristic.)

As the map $\tau : \mathcal{C}_{pm} \longrightarrow \mathcal{C}^*_{pm}$ is birational, it induces an isomorphism of function fields, so we again have a reduced parametrization.

Let p be the characteristic of the ground field. Then the order of P with respect to $\mathcal{C}^*_{pm}$ and τ is $s-1$ except if $p|s$ in which case the order is $\geq s$. As $p > 2$, we conclude: P is a flex point if and only if its order with respect to $\mathcal{C}^*_{pm}$ and τ is > 1. □

Lemma 10 *The number of non-ordinary tangents of $\mathcal{C}_{\overline{\mathbb{F}}_q}$ is in $\mathcal{O}(1)$.*

Proof. The number of closed points of $\mathcal{C}_{\overline{\mathbb{F}}_q}$ which lie over a singular point of $(\mathcal{C}_{pm})_{\overline{\mathbb{F}}_q}$ is $< \frac{(d-1)(d-2)}{2}$. So the number of tangents running through these points is also bounded by this number.

By the previous lemma, we now have to bound the number of points lying over singular points of the dual model. It is a classical result that the degree of the dual model $\mathcal{C}_{pm}^*$ is bounded by $d \cdot (d-1)$. (Briefly, the argument is as follows: The degree of $\mathcal{C}_{pm}^*$ is given by the number of intersection points of $(\mathcal{C}_{pm}^*)_{\overline{\mathbb{F}}_q}$ with a line in the dual plane which does not run through the singularities of $(\mathcal{C}_{pm}^*)_{\overline{\mathbb{F}}_q}$. Such a line corresponds to a point $P \in \mathbb{P}^2(\overline{\mathbb{F}}_q)$, and the intersection points of the line with $(\mathcal{C}_{pm}^*)_{\overline{\mathbb{F}}_q}$ correspond to the tangents of $\mathcal{C}_{\overline{\mathbb{F}}_q}$ passing through P (which are all ordinary). The corresponding points on $(\mathcal{C}_{pm})_{\overline{\mathbb{F}}_q}$ are contained in the intersection of $(\mathcal{C}_{pm})_{\overline{\mathbb{F}}_q}$ with the polar curve for $(\mathcal{C}_{pm})_{\overline{\mathbb{F}}_q}$ and P, which has degree $d-1$. The result now follows with Bezout's theorem.)

This implies that the arithmetic genus of $\mathcal{C}_{pm}^*$ is bounded by $\frac{1}{2} \cdot (d-3) \cdot (d-2)^2 \cdot (d-1)$, and in particular the number of closed points of $\mathcal{C}_{\overline{\mathbb{F}}_q}$ lying over singular points of of $(\mathcal{C}_{pm}^*)_{\overline{\mathbb{F}}_q}$ is bounded by this number. □

The previous lemma implies immediately:

Lemma 11 *The number of closed points P of $\mathcal{C}_{\overline{\mathbb{F}}_q}$ such that some non-ordinary tangent passes through P is in $\mathcal{O}(1)$. (We consider all tangents of all closed points of $\mathcal{C}_{\overline{\mathbb{F}}_q}$ which pass through P, not only the unique tangent at P.)*

Proposition 12 *Let $P \in \mathcal{C}_{ns}(\mathbb{F}_q)$ such that only ordinary tangents pass through P. Then the number of divisors of $\mathfrak{d}$ which contain P and split completely into distinct points is in*

$$\frac{1}{(d-1)!} \cdot q + \mathcal{O}(q^{\frac{1}{2}}) \ .$$

Proof. We consider a covering $c_{\overline{\mathbb{F}}_q} : \mathcal{C}_{\overline{\mathbb{F}}_q} \longrightarrow \mathbb{P}^1_{\overline{\mathbb{F}}_q}$ defined by P. By assumption, for every closed point Q of $\mathbb{P}^1_{\overline{\mathbb{F}}_q}$, the divisor $c^{-1}(Q)+P$ either splits into distinct points or is of the form $2P_1 + P_2 + \cdots + P_{d-1}$ for distinct points P_i. Moreover, the second case only happens for finitely many points (see proof of Lemma 10). Therefore, there are only finitely many ramified closed points of $\mathbb{P}^1_{\overline{\mathbb{F}}_q}$, and for every such point Q, we have $c^{-1}(Q) = 2\tilde{P}_1 + \tilde{P}_2 + \cdots + \tilde{P}_{d-2}$ for distinct points $\tilde{P}_i$. By Proposition 8 and the general remarks at the beginning of the proof, the result follows.

This proposition and Lemma 11 easily imply the statement in Proposition 1 under the assumption the plane model is reflexive.

3 The algorithm

3.1 General considerations

As indicated above, we first construct a tree of large prime relations. Let us formally define what we mean by such a tree.

Let $\mathcal{C}$ be a curve over a finite field $\mathbb{F}_q$. Let $\mathcal{F}$ be a set of prime divisors of $\mathcal{C}$, and let $\mathcal{L}$ be another set of prime divisors of $\mathcal{C}$ which is disjoint from $\mathcal{F}$. We call $\mathcal{F}$ the *factor base* and $\mathcal{L}$ the *set of large primes*. Furthermore, let some elements $c_1, \ldots, c_u \in \mathrm{Cl}^0(\mathcal{C})$ be given.

Now a tree of large prime relations for the given data is an undirected labeled rooted tree whose vertices are contained in $\mathcal{L} \,\dot{\cup}\, \{*\}$ with root $*$, where the edges are labeled as follows:

Each label is a tuple $((r_F)_{F \in \mathcal{F}}, (s_j)_{j=1,\ldots,r})$, where either each entry is an integer or each entry is a residue class modulo the group order which defines in the following way a relation:

- If the edge connects $*$ and a prime divisor P, the equality $\sum_{F \in \mathcal{F}} r_F[F] + [P] = \sum_j s_j c_j$ holds.
- If the edge connects two distinct prime divisors P and Q, the equality $\sum_{F \in \mathcal{F}} r_F F + [P] + [Q] = \sum_j s_j c_j$ holds.

We only consider trees where the relations are given modulo the group order, and we store the labels in sparse representation. Furthermore, as already mentioned, we represent divisors and divisor classes as described in [3, Section 2].

If $\mathcal{T}$ is such a tree, we denote its set of vertices by $V(\mathcal{T})$. Now, by following the arguments in [3, Section 3], one can obtain:

Proposition 13 *Let g and $c \in \mathbb{N}$ with $g, c \geq 2$ be fixed. Then there is an algorithm such that the following holds:*

The input consists of

- *a curve $\mathcal{C}$ of genus g, given by a plane model of bounded degree*
- *the group order of $\mathrm{Cl}^0(\mathcal{C})$*
- *two elements $a, b \in \mathrm{Cl}^0(\mathcal{C})$ with $b \in \langle a \rangle$,*
- *elements $c_1, \ldots, c_u \in \mathrm{Cl}(\mathcal{C})$ whose degrees are bounded, where u is polynomially bounded in $\log(q)$*
- *a factor base $\mathcal{F} \subseteq \mathcal{C}(\mathbb{F}_q)$ of size $\tilde{\mathcal{O}}(q^{1-\frac{1}{c}})$*
- *a tree of large prime relations $\mathcal{T}$ for factor base $\mathcal{F}$, set of large primes $\mathcal{C}(\mathbb{F}_q) - \mathcal{F}$, and classes $c_1, \ldots, c_u$*

- *of a depth which is polynomially bounded in* $\log(q)$
- *with* $\#(\mathcal{F} \cup V(\mathcal{T})) \geq q^{1-\frac{1}{g}+\frac{1}{cg}}$
- *such that the number of non-trivial residue classes involved in each label is polynomially bounded in* $\log(q)$.

Upon this input the algorithm computes the discrete logarithm of b with respect to a in an expected time of $\tilde{\mathcal{O}}(q^{2-\frac{2}{c}})$. *The algorithm has storage requirements of* $\tilde{\mathcal{O}}(\#(\mathcal{F} \cup V(\mathcal{T})) \cdot \log(q))$.

Remark 14 We are going to apply this proposition with $c = d - 2$, where d is the degree of the plane model under consideration. In our application, d is fixed, but not the genus, g, is not. As for fixed degree there are only finitely many possibilities for the genus, there still exists an algorithm with the specified properties.

In [3] this is proven for $c = g$, but the general case is no more difficult than the statement in [3]. We briefly recall the algorithm and its analysis. For simplicity, we focus on the case the group order N is prime and generated by a.

We repeatedly select uniformly randomly chosen elements $\alpha, \beta \in \mathbb{Z}/N\mathbb{Z}$ and compute the unique along P_0 reduced divisor D with

$$[D] - \deg(D) \cdot [P_0] = \alpha a + \beta b$$

in free representation. If D splits over the factor base and the vertices of the tree, we use the tree to obtain a "full relation", that is, a relation between factor base elements, $c_1, \ldots, c_u$ and a, b. We stop this procedure if we have obtained $\#\mathcal{F}+u+1$ full relations. After that we try to compute the discrete logarithm via linear algebra. If this fails, we repeat the whole procedure.

Using an algorithm from sparse linear algebra, the linear algebra computation can be performed in an expected time of $\tilde{O}(q^{2-\frac{2}{c}})$. We have to show that the relation generation can also be performed in an expected time of $\tilde{\mathcal{O}}(q^{2-\frac{2}{c}})$.

There exists a constant $C > 0$ such that the number of elements of $\mathrm{Cl}^0(\mathcal{C})$ which are represented by an along P_0 reduced divisor which splits completely into factor base elements and vertices of the tree is

$$> \frac{1}{g!} \cdot q^{(1-\frac{1}{g}+\frac{1}{cg})\cdot g} - C \cdot q^{g-1} \ .$$

For q large enough, this is $\geq \frac{1}{2g!} \cdot q^{(1-\frac{1}{g}+\frac{1}{cg})\cdot g} = \frac{1}{2g!} \cdot q^{g-1+\frac{1}{c}}$. Now for q large enough the probability that a uniformly randomly chosen group element is represented by a divisor which splits over the factor base and the vertices of the tree is

$$> \frac{1}{4g!} \cdot q^{(g-1+\frac{1}{c})-g} = \frac{1}{4g!} \cdot q^{-(1-\frac{1}{c})} \ .$$

The expected number of tries until we have one relation is therefore $\leq 4g! \cdot q^{1-\frac{1}{c}}$. Consequently, the expected number of tries until we have $\#\mathcal{F} + u + 1$ relations is in $\mathcal{O}(q^{2-\frac{2}{c}})$, and the expected time is in $\tilde{\mathcal{O}}(q^{2-\frac{2}{c}})$.

In the general case, we first construct a "potential generating system" $c'_1, \ldots, c'_{u'}$. (In [3], $c_1, \ldots, c_u$ is already such a system, and we have $u = u'$ and $c_j = c'_j$ for all j.)

Then we try to generate relations between the factor base, a, b, $c_1, \ldots, c_u$ and $c_1, \ldots, c'_{u'}$ as follows:

We choose $s_1, \ldots, s_{u'}$ uniformly randomly in $\mathbb{Z}/N\mathbb{Z}$, we compute the unique along P_0 reduced divisor D with

$$[D] - \deg(D) \cdot [P_0] = \sum_j s'_j c'_j + \alpha a + \beta b$$

in free representation. Again if D splits over the factor base and the vertices of the tree, we use the tree to obtain a "full relation".

As above, if we have enough relations, we try to solve for the discrete logarithm. Moreover, we stop and restart the whole algorithm if a predefined time bound has been reached.

For details we refer to [3, Section 3]. □

3.2 Construction of the tree of large prime relations

We consider curves of genus ≥ 2 represented by plane models of a fixed degree $d \geq 4$ such that the linear system $\mathfrak{d}$ as defined above contains a divisor which splits completely into distinct points. We wish to construct a factor base and a tree of prime relations satisfying the assumptions of Proposition 13 with $c = d - 2$, $u = 1$ and $c_1 = [D_\infty]$. (In fact, we do not even have to store c_1 because it is anyway determined by the left hand side of the relations.) As already mentioned, to construct the tree, we intersect the curve with lines and consider relations of the form (1).

So let a curve $\mathcal{C}/\mathbb{F}_q$, represented by a plane model $\mathcal{C}_{pm}$ of degree d be given.

We start off by computing the L-polynomial and therefore also the group order. Using Pila's extension of Schoof's algorithm ([14], [15]), this can be achieved in a time which is polynomially bounded in $\log(q)$. Then we compute the genus; again we need a time which is polynomially bounded in $\log(q)$ (cf. [3]).

To establish the theorems, we have to give an algorithm to compute a factor base and a tree of large prime relations as specified which operates in an expected time of $\tilde{\mathcal{O}}(q^{2-\frac{2}{d-2}})$ under the assumptions of the theorems.

Let us first discuss some basic computations.

Lemma 15 *One can compute a uniformly randomly distributed point in* $\mathcal{C}(\mathbb{F}_q)$ *in an expected time which is polynomially bounded in* $\log(q)$.

This is Proposition 3.8 in [3].

Lemma 16 *Given a linear form* $W \in \mathbb{F}_q[X,Y,Z]_1$*, one can compute the divisor* $\mathrm{div}(W_{|\mathcal{C}}) \in \mathfrak{D}$ *in free representation in an expected time which is polynomially bounded in* $\log(q)$.

Sketch of a proof. We choose some linear form U such that the intersection between the two lines defined by W and U does not lie on $\mathcal{C}_{pm}$. Then we have $\mathrm{div}(W_{|\mathcal{C}}) = (\frac{W_{|\mathcal{C}}}{U_{|\mathcal{C}}})_+$, the positive part of the principal divisor $(\frac{W_{|\mathcal{C}}}{U_{|\mathcal{C}}})$. With these considerations, the computation can be performed with standard algorithms on ideal arithmetic. □

Remark 17 In fact, we only need an algorithm to determine if a line runs through $\mathcal{C}_{ns}$, defines a completely split divisor, and in this case to compute such a divisor. This task can easily be achieved by inserting the equation for the line into the curve equation, factoring the resulting polynomial and finally by checking for each root if all partial derivatives vanish.

Construction of the tree

For the construction of the tree, we use a "stage-wise procedure" which is based on successive enlargements of the factor base. As in the previous subsection, we denote the tree by $\mathcal{T}$ and its set of vertices by $V(\mathcal{T})$. The factor base is always denoted by $\mathcal{F}$.

The following algorithm has the desired expected running time. However, it is conceivable that this algorithm violates the desired storage requirements of $\tilde{\mathcal{O}}(q^{1-\frac{1}{g}+\frac{1}{(d-2)g}})$ for $g \geq d-1$. At the end of this section we point out modifications leading to an algorithm which also has the desired storage requirements.

Stage 1

The computation for Stage 1 is as follows:

We first determine a subset $\mathcal{F}_0$ of $\mathcal{C}_{ns}(\mathbb{F}_q)$ of size $\lceil \log(q) \cdot q^{1-\frac{1}{d-2}} \rceil$ such that through each point of $\mathcal{F}_0$ there passes at least one line which splits completely into distinct points of $\mathcal{C}(\mathbb{F}_q)$. For this, we repeatedly choose lines uniformly at random and compute the corresponding divisor. Note that the probability that a line gives rise to a divisor which splits completely into distinct points is in $\Omega(1)$ by Proposition 5, thus the expected running time is in $\tilde{\mathcal{O}}(q^{1-\frac{1}{d-2}})$.

Then we choose a subset $\mathcal{F}_1$ of $\mathcal{C}_{ns}(\mathbb{F}_q) - \mathcal{F}_0$ of size $\lceil (5 \cdot (d-1)!^2)^{\frac{1}{d-2}} \cdot q^{1-\frac{1}{d-2}} \rceil$ uniformly randomly from the set of all such subsets. By Lemma 15 this task can also be achieved in an expected time of $\tilde{\mathcal{O}}(q^{1-\frac{1}{d-2}})$.

We iterate over all lines passing through two distinct points of $\mathcal{F}_1$. For each such line, we compute the corresponding divisor D on $\mathcal{C}$ in free representation.

Now for every such divisor D, we check if it splits in the form

$$D = P_1 + \cdots + P_{d-1} + Q \tag{2}$$

with $P_i \in \mathcal{F}_0 \cup \mathcal{F}_1$ and $Q \in \mathcal{C}_{ns}(\mathbb{F}_q) - (\mathcal{F}_0 \cup \mathcal{F}_1)$. If this is the case, we store the divisor.

After the consideration of all lines we choose for each point Q as in (2) one divisor as above. If then we have $\geq \lceil q^{1-\frac{1}{d-2}} \rceil$ distinct points Q (and corresponding divisors), we set the factor base as $\mathcal{F} := \mathcal{F}_0 \cup \mathcal{F}_1$, and for each such divisor, we insert an edge from $*$ to Q with the data for the relation

$$[P_1] + \cdots + [P_{d-1}] + [Q] = [D_\infty]$$

into the tree.

If we do not have enough points, we choose another subset $\mathcal{F}_1$ and repeat.

Stages ≥ 2

At the beginning of Stage $s \geq 2$, we have a factor base $\mathcal{F} \subseteq \mathcal{C}_{ns}(\mathbb{F}_q)$ and a tree $\mathcal{T}$. We now choose a set $\mathcal{G} \subseteq \mathcal{C}_{ns}(\mathbb{F}_q) - (\mathcal{F} \cup (V(\mathcal{T}))$ of size $\lceil (5 \cdot (d-1)!^2)^{\frac{1}{d-2}} \cdot q^{1-\frac{1}{d-2}} \rceil$ uniformly randomly from the set of all such subsets. We then consider all lines through two distinct points of $\mathcal{G}$. For each such line, we check if it defines a divisor D of the form

$$D = P_1 + \cdots + P_{d-2} + P + Q \tag{3}$$

with $P_i \in \mathcal{G}$ for $i = 1, \ldots, d-2$, $P \in \mathcal{F} \cup V(\mathcal{T})$, and $Q \in \mathcal{C}(\mathbb{F}_q) - (\mathcal{F} \cup \mathcal{G} \cup V(\mathcal{T}))$.

After the consideration of all lines we choose for each point Q as in (3) one divisor as above. If then we have $\geq \lceil 2^s \cdot q^{1-\frac{1}{d-2}} \rceil$ distinct points Q (and corresponding divisors), we update $\mathcal{F}$ as $\mathcal{F} \cup \mathcal{G}$, and for each such divisor, we insert an edge from P to Q with a label for the relation

$$[P_1] + \cdots + [P_{d-2}] + [P] + [Q] = [D_\infty]$$

into the tree.

Otherwise, we restart the computation of Stage s with another set $\mathcal{G}$.

The end

We stop the computation after $\lceil \log_2(q) \cdot (\frac{1}{d-2} - \frac{1}{g} + \frac{1}{(d-2)g}) \rceil$ stages. Note that then the tree has $\geq q^{1-\frac{1}{g}+\frac{1}{(d-2)g}}$ leaves.

Clearly, the number of stages is in $\mathcal{O}(\log(q))$ (and thus so is the depth of the tree), and the size of the factor base is in $\mathcal{O}(q^{1-\frac{1}{d-2}})$. Moreover, by the lemmata above, the computation of each stage can be performed in an expected time of $\tilde{\mathcal{O}}(q^{2-\frac{2}{d-2}})$.

So, all we have to prove is:

There exists a constant $c > 0$ such that for q large enough and for all s, the expected value of repetitions in Stage s is $\leq c$.

This statement is proven in the next subsection (see Proposition 20 and Remark 21).

It follows a more formal description of the algorithm.

Algorithm: Construction of the tree of large prime relations

Input: A curve $\mathcal{C}/\mathbb{F}_q$, represented by a plane model of the fixed degree d.

Output: A factor base and a tree of large prime relations satisfying the requirements of Proposition 13 for $c = d - 2$.

Construct a set $\mathcal{F} \subseteq \mathcal{C}(\mathbb{F}_q)$ and a labeled rooted tree $\mathcal{T}$ with vertex set contained in $\mathcal{C}(\mathbb{F}_q) \,\dot\cup\, \{*\}$ as follows:

Let $\mathcal{T}$ consist only of the root $*$.
Determine a subset $\mathcal{F}_0$ of $\mathcal{C}_{ns}(\mathbb{F}_q)$ of size $\lceil \log(q) \cdot q^{1-\frac{1}{d-2}} \rceil$ as follows:
Repeat
  Choose a linear form $W \in \mathbb{F}_q[X,Y,Z]_1$ uniformly at random and compute the divisor $D := \mathrm{div}(W_{|\mathcal{C}})$.
  If D splits completely into distinct points of $\mathcal{C}_{ns}(\mathbb{F}_q)$, insert one of these points in $\mathcal{C}_{ns}(\mathbb{F}_q)$ into $\mathcal{F}_0$.
Until $\mathcal{F}_0$ has size $\lceil \log(q) \cdot q^{1-\frac{1}{d-2}} \rceil$.
Repeat
  Choose $\mathcal{F}_1 \subseteq \mathcal{C}_{ns}(\mathbb{F}_q)$ of size $\lceil (5 \cdot (d-1)!^2)^{\frac{1}{d-2}} \cdot q^{1-\frac{1}{d-2}} \rceil$ uniformly randomly from the set of all such subsets.
  Construct a list L of divisors in free representation as follows:

Iterate over all lines passing through two distinct points of $\mathcal{F}_1$.

Whenever such a line defines a divisor of the form

$$P_1 + \cdots + P_{d-1} + Q$$

with $P_i \in \mathcal{F}_0 \cup \mathcal{F}_1$ and $Q \in \mathcal{C}_{ns}(\mathbb{F}_q) - (\mathcal{F}_0 \cup \mathcal{F}_1)$, store the divisor in L.

Sort L for the points $Q \in \mathcal{C}_{ns}(\mathbb{F}_q) - (\mathcal{F}_0 \cup \mathcal{F}_1)$ occurring in the divisors, for each such point choose one divisor and delete the others.

Until L contains $\geq q^{1-\frac{1}{d-2}}$ divisors.

Let $\mathcal{F} \longleftarrow \mathcal{F}_0 \cup \mathcal{F}_1$.

For each divisor in L, insert an edge from $*$ to Q into $\mathcal{T}$, labeled with the data for the corresponding relation.

For $s = 1, \ldots, \lceil \log_2(q) \cdot (\frac{1}{d-2} - \frac{1}{g} + \frac{1}{(d-2)g}) \rceil$ do

Repeat

Choose $\mathcal{G} \subseteq \mathcal{C}_{ns}(\mathbb{F}_q) - V(\mathcal{T})$ of size $\lceil (5 \cdot (d-1)!^2)^{\frac{1}{d-2}} \cdot q^{1-\frac{1}{d-2}} \rceil$ uniformly randomly from the set of all such subsets.

Construct a list L of divisors in free representation as follows:

Iterate over all lines passing through two distinct points of $\mathcal{G}$.

Whenever such a line defines a divisor of the form

$$P_1 + \cdots + P_{d-2} + P + Q$$

with $P_i \in \mathcal{G}$, $P \in \mathcal{F} \cup V(\mathcal{T})$ and $Q \in \mathcal{C}_{ns}(\mathbb{F}_q) - (\mathcal{F} \cup \mathcal{G} \cup V(\mathcal{T}))$, store the divisor in L.

Sort L for the points $Q \in \mathcal{C}_{ns}(\mathbb{F}_q) - (\mathcal{F} \cup \mathcal{G})$ occurring in the divisors, for each such point choose one divisor and delete the others.

Until L contains $\geq 2^s \cdot q^{1-\frac{1}{d-2}}$ divisors.

Let $\mathcal{F} \longleftarrow \mathcal{F} \cup \mathcal{G}$.

For each divisor in L, insert an edge from P to Q into $\mathcal{T}$, labeled with the data for the corresponding relation.

Output $\mathcal{F}, \mathcal{T}$

On the storage requirements

We wish to have an algorithm with storage requirements of $\tilde{\mathcal{O}}(\max(q^{1-\frac{1}{d-2}}, q^{1-\frac{1}{g}+\frac{1}{(d-2)g}}))$ (which is $\tilde{\mathcal{O}}(q^{1-\frac{1}{g}+\frac{1}{(d-2)g}})$ for $g \geq d-3$). It is however conceivable that the size of the lists in the algorithm above does not satisfy the desired bound. The bound can be guaranteed with the following minor modification of each stage of the algorithm:

One maintains a sorted list L already during the construction (via a balanced binary search tree), where the sorting is for the points Q in the divisors, where Q is as in the algorithm. One always inserts at most one relation for each point of $\mathcal{C}(\mathbb{F}_q) - (\mathcal{F}_0 \cup \mathcal{F}_1)$ (for Stage 1) respectively $\mathcal{C}(\mathbb{F}_q) - (\mathcal{F} \cup V(\mathcal{T}))$ (for Stage $s > 1$) into the list. Moreover, one stops the construction of the list if $\#\mathcal{F} + \#V(\mathcal{T}) + \#L \geq q^{1-\frac{1}{g}+\frac{1}{(d-2)\cdot g}}$. One then inserts all the new points into the tree and stops the whole computation.

Just as described above, if at a particular stage all lines through distinct points of $\mathcal{F}_1$ (for Stage 1) respectively $\mathcal{G}$ (for Stage $s > 1$) have been considered, one repeats the whole stage with another set of points.

3.3 Analysis of the construction of the tree

We now analyze the construction of the tree. Let for this $d \geq 4$ still be fixed.

Proposition 18 *We consider curves $\mathcal{C}/\mathbb{F}_q$ with fixed plane models $\mathcal{C}_{pm}$ of degree d and birational morphisms $\pi : \mathcal{C} \longrightarrow \mathcal{C}_{pm}$ and subsets $S \subseteq \mathcal{C}_{ns}(\mathbb{F}_q)$ with $\#S \in o(q)$ such that for every point $P \in S$ there exists a divisor in $\mathfrak{d}$ which splits completely into distinct points and contains P.*

Let $S \subseteq \mathcal{C}_{ns}(\mathbb{F}_q)$ be such a set. Then there are $\gtrsim \frac{1}{2(d-1)!}q$ points $Q \in \mathcal{C}_{ns}(\mathbb{F}_q) - S$ such that there is are at least $\frac{1}{2(d-1)!} \cdot \#S$ divisors in $\mathfrak{d}$ which split completely into distinct points of $\mathcal{C}_{ns}(\mathbb{F}_q)$ and contain Q and exactly one point from S.

Proof. By Proposition 4, for q large enough, the following holds: For $P \in S$ the number of divisors in $\mathfrak{d}$ which split completely into distinct points of $\mathcal{C}_{ns}(\mathbb{F}_q)$ and contain P is $\gtrsim \frac{1}{(d-1)!} \cdot \#\mathcal{C}_{ns}(\mathbb{F}_q)$.

As the divisors in $\mathfrak{d}$ are defined by lines, for every $P \in S$, the number of divisors as above which also contain another point from S is $< \#S \in o(q)$. Thus for $P \in S$, the number of such divisors which do not contain another point from S is again $\gtrsim \frac{1}{(d-1)!} \cdot \#\mathcal{C}_{ns}(\mathbb{F}_q)$. Altogether, we have $\gtrsim \frac{1}{(d-1)!} \cdot \#\mathcal{C}_{ns}(\mathbb{F}_q) \cdot \#S$ divisors in $\mathfrak{d}$ which split completely into distinct points and contain exactly one point from S.

Every point outside of S is contained in at most $\#S$ such divisors. Let $c \cdot \#\mathcal{C}_{ns}(\mathbb{F}_q)$ be the number of points of $\mathcal{C}_{ns}(\mathbb{F}_q)$ which contain $\geq \frac{1}{2(d-1)!} \cdot \#S$ such divisors. Then altogether we have $< (c \cdot \#S + (1-c) \cdot \frac{1}{2(d-1)!} \cdot \#S) \cdot \#\mathcal{C}_{ns}(\mathbb{F}_q) = (\frac{1}{2(d-1)!} + (1 - \frac{1}{2(d-1)!}) \cdot c) \cdot \#S \cdot \#\mathcal{C}_{ns}(\mathbb{F}_q)$ such divisors. This implies that

$$\frac{1}{(d-1)!} \lesssim \frac{1}{2(d-1)!} + (1 - \frac{1}{2(d-1)!}) \cdot c\,.$$

This implies

$$c \gtrsim \frac{\frac{1}{2(d-1)!}}{1 - \frac{1}{2(d-1)!}} = \frac{1}{2(d-1)! - 1} > \frac{1}{2(d-1)!}\,.$$

□

Proposition 19 *Let $c > 0$ be fixed. We consider curves $\mathcal{C}/\mathbb{F}_q$ with fixed plane models $\mathcal{C}_{pm}$ of degree d and birational morphisms $\pi : \mathcal{C} \longrightarrow \mathcal{C}_{pm}$ and subsets $S \subseteq \mathcal{C}_{ns}(\mathbb{F}_q)$ with $q^{\frac{1}{d-2}} \in \mathcal{O}(\#S)$ and $\#S \in o(q)$ such that for every point $P \in S$ there exists a divisor in $\mathfrak{d}$ which splits completely into distinct points and contains P.*

Let $S \subseteq \mathcal{C}_{ns}(\mathbb{F}_q)$ be such a subset, and let $u := \lceil c \cdot q^{1-\frac{1}{d-2}} \rceil$. Let $Q \in \mathcal{C}_{ns}(\mathbb{F}_q) - S$ such that there are at least $\frac{1}{2(d-1)!} \cdot \#S$ divisors in $\mathfrak{d}$ which split completely into distinct points of $\mathcal{C}_{ns}(\mathbb{F}_q)$ and contain exactly one point from S and Q. Let U be a random subset of $\mathcal{C}_{ns}(\mathbb{F}_q) - S$ which is uniformly randomly distributed among all subsets of cardinality u.

Then the probability that there is a divisor which splits completely into distinct points of $\mathcal{C}_{ns}(\mathbb{F}_q)$, contains Q, exactly one point of S and otherwise only points of U is $\gtrsim \frac{c^{d-2}}{2(d-1)!} \cdot \frac{\#S}{q}$. (We do not impose a condition on Q not being in U.)

For later use, we will prove a more accurate result with an error term (see Equation (4)).

Proof. Let for some divisor $D \in \mathfrak{d}$ which splits completely into distinct points of $\mathcal{C}_{ns}(\mathbb{F}_q)$ and contains Q and exactly one point from S p_D be the probability that all points of D distinct from Q as well as the point in S lie in U.

Similarly, let for two distinct divisors $D_1, D_2 \in \mathfrak{d}$, each splitting completely into distinct points of $\mathcal{C}_{ns}(\mathbb{F}_q)$ and containing Q and exactly one point from S p_{D_1,D_2} be the probability that all points of D_1 and D_2 distinct from Q as well as the points in S lie in U.

Let p be the probability we wish to estimate in the proposition. We have

$$p \geq \sum_D p_D - \frac{1}{2} \sum_{D_1,D_2} p_{D_1,D_2} \ ,$$

where the sums range over all divisors specified above. Now

$$p_D = \frac{\binom{\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S-(d-2)}{u-(d-2)}}{\binom{\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S}{u}} = \frac{(u-d+3)\cdot(u-d+4)\cdots u}{(\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S-d+3)\cdots(\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S)} \ ,$$

$$p_{D_1,D_2} = \frac{\binom{\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S-(2d-4)}{u-(2d-4)}}{\binom{\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S}{u}} = \frac{(u-2d+5)\cdot(u-2d+6)\cdots u}{(\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S-2d+5)\cdots(\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S)} \ .$$

We have

$$p_D \in [(\frac{u-d+3}{\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S})^{d-2}, (\frac{u}{\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S-d+3})^{d-2}]$$

and

$$p_{D_1,D_2} \in [(\frac{u-2d+5}{\#\mathcal{C}_{ns}(\mathbb{F}_q) - \#S})^{2d-4}, (\frac{u}{\#\mathcal{C}_{ns}(\mathbb{F}_q) - \#S - 2d + 5})^{2d-4}] \ .$$

Let N be the number of divisors in $\mathfrak{d}$ which split completely into distinct points of $\mathcal{C}_{ns}(\mathbb{F}_q)$ and contain Q and exactly one point from S. Note that

$$N \in [\frac{1}{2(d-1)!} \cdot \#S, \#S]$$

by assumption.

Then

$$p \geq N \cdot (\frac{u-d+3}{\#\mathcal{C}_{ns}(\mathbb{F}_q) - \#S})^{d-2} - \frac{1}{2} \cdot N^2 \cdot (\frac{u}{\#\mathcal{C}_{ns}(\mathbb{F}_q) - \#S - 2d + 5})^{2d-4}$$

$$\in N \cdot (\frac{u}{\#\mathcal{C}_{ns}(\mathbb{F}_q) - \#S})^{d-2} \cdot (1 + \mathcal{O}(\frac{1}{u})) + \mathcal{O}(N^2 \cdot (\frac{u}{\#\mathcal{C}_{ns}(\mathbb{F}_q) - \#S})^{2d-4}) \ .$$

Now

$$N \cdot (\frac{u}{\#\mathcal{C}_{ns}(\mathbb{F}_q) - \#S})^{d-2} \in \Theta(\#S \cdot q^{-1}),$$

$$(\frac{u}{\#\mathcal{C}_{ns}(\mathbb{F}_q) - \#S})^{d-2} \cdot \frac{1}{u} \in \Theta(\#S \cdot q^{-2+\frac{1}{d-2}}),$$

$$N^2 \cdot (\frac{u}{\#\mathcal{C}_{ns}(\mathbb{F}_q) - \#S})^{2d-4} \in \Theta(\#S^2 \cdot q^{-2}) \ .$$

We have $\#S \cdot q^{-2+\frac{1}{d-2}} \in \mathcal{O}(\#S^2 \cdot q^{-2})$ as $q^{\frac{1}{d-2}} \in \mathcal{O}(\#S)$ by assumption. So there exists some constant $C > 0$ (depending only on c) such that

$$p \geq N \cdot (\frac{u}{\#\mathcal{C}_{ns}(\mathbb{F}_q) - \#S})^{d-2} - C \cdot (N \cdot (\frac{u}{\#\mathcal{C}_{ns}(\mathbb{F}_q) - \#S})^{d-2})^2 \ . \qquad (4)$$

In particular

$$p \gtrsim \frac{c^{d-2}}{2(d-1)!} \cdot \frac{\#S}{q} \ .$$

□

Proposition 20 *Let $c > 0$ be fixed. We consider curves $\mathcal{C}/\mathbb{F}_q$ with fixed plane models $\mathcal{C}_{pm}$ of degree d and birational morphisms $\pi : \mathcal{C} \longrightarrow \mathcal{C}_{pm}$ and subsets $S \subseteq \mathcal{C}_{ns}(\mathbb{F}_q)$ such that $q^{1-\frac{1}{d-2}} \in o(\#S)$ and $\#S \in o(q)$. Then the following holds:*

Let $S \subseteq \mathcal{C}_{ns}(\mathbb{F}_q)$ be such a subset, and let $u := \lceil c \cdot q^{1-\frac{1}{d-2}} \rceil$. Let U be a random subset of $\mathcal{C}_{ns}(\mathbb{F}_q) - S$ which is uniformly randomly distributed among all subsets of cardinality u.

Then the probability that there are at least $\frac{c^{d-2}}{5(d-1)!^2} \cdot \#S$ *points* $Q \in \mathcal{C}_{ns}(\mathbb{F}_q) - (S \cup U)$ *such that there exists a divisor in* $\mathfrak{d}$ *which splits completely into distinct points of* $\mathcal{C}_{ns}(\mathbb{F}_q)$*, contains* Q*, exactly one point of* S *and otherwise points of* U *is asymptotically equal to 1 for* $q \longrightarrow \infty$*.*

Proof. Let $\mathcal{Q}$ be the set of points $Q \in \mathcal{C}_{ns}(\mathbb{F}_q) - S$ such that there exist at least $\frac{1}{2(d-1)!} \cdot \#S$ divisors in $\mathfrak{d}$ which split completely into distinct points of $\mathcal{C}_{ns}(\mathbb{F}_q)$, contain Q and exactly one point from $\#S$. By Proposition 18, $\#\mathcal{Q} \gtrsim \frac{1}{2(d-1)!} q$. We only consider points from $\mathcal{Q}$.

Let for such a point Q A_Q be the event that there exists at least one divisor which split completely into distinct points of $\mathcal{C}_{ns}(\mathbb{F}_q)$, contains Q, exactly one point from $\#S$ and otherwise points from U. (We do not impose a condition on Q not being in U.)

We have

$$\mathbb{E}[\sum_{Q \in \mathcal{Q}} \chi_{A_Q}] = \sum_{Q \in \mathcal{Q}} \mathbb{P}[A_Q] \gtrsim \frac{c^{d-2}}{4 \cdot (d-1)!^2} \cdot \#S \tag{5}$$

by Proposition 19.

Below we show that the standard deviation of $\sum_{Q \in \mathcal{Q}} \chi_{A_Q}$ is in $o(\mathbb{E}[\sum_{Q \in \mathcal{Q}} \chi_{A_Q}])$. Let us for the moment assume that we have already proven this result, and let us see how the statement in the proposition then follows. With the Chebyshov inequality and (5) we conclude that with a probability which is asymptotically equal to 1 for $q \longrightarrow \infty$, we have

$$\sum_{Q \in \mathcal{Q}} \chi_{A_Q} \geq \frac{c^{d-2}}{\frac{9}{2} \cdot (d-1)!^2} \cdot \#S \ .$$

Recall that $\sum_{Q \in \mathcal{Q}} \chi_{A_Q}$ is a lower bound on the number of points $Q \in \mathcal{C}_{ns}(\mathbb{F}_q) - S$ such that there exists a divisor in $\mathfrak{d}$ which splits completely into distinct points of $\mathcal{C}_{ns}(\mathbb{F}_q)$, contains Q, exactly one point from S and otherwise only points from U. Now, such a point Q might also be contained in U. As however $\#U \in o(\#S)$ by assumption, we conclude:

With a probability which is asymptotically equal to 1, there exist at least $\frac{c^{d-2}}{5 \cdot (d-1)!^2} \cdot \#S$ points in $Q \in \mathcal{C}_{ns}(\mathbb{F}_q) - (S \cup U)$ such that there exists a divisor in $\mathfrak{d}$ which splits completely into distinct points of $\mathcal{C}_{ns}(\mathbb{F}_q)$, contains Q, exactly one point from S and otherwise only points from U. This is the desired result.

It remains to be shown that the standard deviation of $\sum_{Q \in \mathcal{Q}} \chi_{A_Q}$ is in $o(\mathbb{E}[\sum_{Q \in \mathcal{Q}} \chi_{A_Q}])$.

The variance of $\mathbb{E}[\sum_{Q\in\mathcal{Q}}\chi_{A_Q}]$ is

$$\begin{aligned}
& \mathbb{E}[(\sum_{Q\in\mathcal{Q}}\chi_{A_Q})^2]-(\mathbb{E}[\sum_{Q\in\mathcal{Q}}\chi_{A_Q}])^2 \\
= & \sum_{Q_1,Q_2\in\mathcal{Q}}(\mathbb{P}[A_{Q_1}\cap A_{Q_2}]-\mathbb{P}[A_{Q_1}]\cdot\mathbb{P}[A_{Q_2}]) \\
\leq & \mathbb{E}[\sum_{Q\in\mathcal{Q}}\chi_{A_Q}]+\sum_{Q_1,Q_2\in\mathcal{Q},Q_1\neq Q_2}(\mathbb{P}[A_{Q_1}\cap A_{Q_2}]-\mathbb{P}[A_{Q_1}]\cdot\mathbb{P}[A_{Q_2}])
\end{aligned} \tag{6}$$

We now wish to establish a suitable upper bound on $\mathbb{P}[A_{Q_1}\cap A_{Q_2}]-\mathbb{P}[A_{Q_1}]\cdot\mathbb{P}[A_{Q_2}]$ for $Q_1\neq Q_2$. We use (4) to obtain a lower bound on the minuend. (Note that $q^{\frac{1}{d-2}}\leq q^{1-\frac{1}{d-2}}\in o(\#S)$ by assumption and because $d\geq 4$. Therefore the assumptions of Proposition 19 are satisfied.)

The task is now to establish a suitable upper bound on the subtrahend.

Let $Q_1,Q_2\in\mathcal{Q}$ with $Q_1\neq Q_2$ be fixed. Let for two divisors $D_1,D_2\in\mathfrak{d}$, each splitting completely into distinct points of $\mathcal{C}_{ns}(\mathbb{F}_q)$, such that D_1 contains Q_1 and D_2 contains Q_2 and both contain a point from S, p_{D_1,D_2} be the probability that the remaining points in both divisors are all contained in U.

Clearly,

$$\mathbb{P}[A_{Q_1}\cap A_{Q_2}]\leq\sum_{D_1,D_2}p_{D_1,D_2}\ ,$$

where the sum ranges over all pairs of divisors just specified.

For an upper bound on p_{D_1,D_2}, there are two cases to consider, depending on whether the two lines meet in $\mathcal{C}_{ns}(\mathbb{F}_q)-S$ or not. In the first case, D_1 and D_2 have one point outside of S in common, and we write $D_1\cap D_2\nsubseteq S$. In the second case, they do not have a point outside of S in common, and we write $D_2\cap D_2\subseteq S$.

We consider the case that the two lines meet in $\mathcal{C}_{ns}(\mathbb{F}_q)-S$ first. In this case, $D_1\cup D_2$ contains $2d-5$ points in $\mathcal{C}_{ns}(\mathbb{F}_q)-(S\cup\{Q_1,Q_2\})$. We have

$$p_{D_1,D_2}=\frac{\binom{\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S-(2d-5)}{u-(2d-5)}}{\binom{\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S}{u}}=\frac{(u-2d+6)\cdot(u-2d+7)\cdots u}{(\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S-2d+6)\cdots(\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S)}\ .$$

Thus

$$p_{D_1,D_2}\leq(\frac{u}{\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S-2d+6})^{2d-5}\in\mathcal{O}(q^{\frac{-(2d-5)}{d-2}})=\mathcal{O}(q^{-2+\frac{1}{d-2}})$$

in this case.

Note that for every divisor $D_1\in\mathfrak{d}$ which contains exactly one point from S and Q_1, there exist at most $d-2$ divisors $D_2\in\mathfrak{d}$ with $D_1\cap D_2\nsubseteq S$

which contain exactly one point from S and Q_2. (D_2 is determined by its intersection with D_1.) Thus

$$\sum_{\substack{D_1,\, D_2 \text{ with} \\ D_1 \cap D_2 \not\subseteq S}} p_{D_1,D_2} \in \mathcal{O}(\#S \cdot q^{-2+\frac{1}{d-2}}) \,. \tag{7}$$

We now consider pairs of divisors of the second type. For such divisors D_1, D_2, we have

$$p_{D_1,D_2} = \frac{\binom{\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S-(2d-4)}{u-(2d-4)}}{\binom{\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S}{u}} = \frac{(u-2d+5)\cdot(u-2d+6)\cdots u}{(\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S-2d+5)\cdots(\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S)} \,.$$

Thus

$$p_{D_1,D_2} \leq (\frac{u}{\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S-2d+5})^{2d-4} \in (\frac{u}{\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S})^{2d-4} \cdot (1+\mathcal{O}(\frac{1}{q})) \,.$$

Let for $i = 1, 2$ N_i be the number of divisors which split completely into distinct points of $\mathcal{C}_{ns}(\mathbb{F}_q)$, contain Q_i and exactly one element from S. Then

$$\sum_{\substack{D_1,\, D_2 \text{ with} \\ D_1 \cap D_2 \subseteq S}} p_{D_1,D_2} \leq N_1 N_2 (\frac{u}{\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S})^{2d-4} \cdot (1 + C_1 \cdot \frac{1}{q}) \tag{8}$$

for some constant $C_1 > 0$.

Altogether, we obtain by (4), (7) and (8):

$$\begin{aligned} & \mathbb{P}[A_{Q_1} \cap A_{Q_2}] - \mathbb{P}[A_{Q_1}] \cdot \mathbb{P}[A_{Q_2}] \\ \leq\ & C_0 \cdot \#S \cdot q^{-2+\frac{1}{d-2}} + N_1 N_2 (\frac{u}{\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S})^{2d-4} \cdot (1 + C_1 \cdot \frac{1}{q}) - \\ & (N_1 \cdot (\frac{u}{\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S})^{d-2} - C_2 \cdot (N_1 \cdot (\frac{u}{\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S})^{d-2})^2) \cdot \\ & (N_2 \cdot (\frac{u}{\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S})^{d-2} - C_2 \cdot (N_2 \cdot (\frac{u}{\#\mathcal{C}_{ns}(\mathbb{F}_q)-\#S})^{d-2})^2) \end{aligned}$$

for constants $C_0, C_1, C_2 > 0$. This is in

$$\mathcal{O}(\#S \cdot q^{-2+\frac{1}{d-2}} + \#S^2 \cdot q^{-3}) \subseteq \mathcal{O}(\#S \cdot q^{-2+\frac{1}{d-2}}) \,.$$

Because of this and (6) the variance of $\sum_{Q \in \mathcal{Q}} \chi_{A_Q}$ is in

$$\mathcal{O}(\mathbb{E}[\sum_{Q \in \mathcal{Q}} \chi_{A_Q}] + \#S \cdot q^{\frac{1}{d-2}}) \subseteq \mathcal{O}(\mathbb{E}[\sum_{Q \in \mathcal{Q}} \chi_{A_Q}]) + o(\#S^2) \subseteq o(\mathbb{E}[\sum_{Q \in \mathcal{Q}} \chi_{A_Q}]^2) \,.$$

Here, the second inclusion follows from (5). We obtain that the standard deviation of $\sum_{Q \in \mathcal{Q}} \chi_{A_Q}$ is in $o(\mathbb{E}[\sum_{Q \in \mathcal{Q}} \chi_{A_Q}])$, and this completes the proof.□

Remark 21 We remark how Proposition 20 is applied in the analysis of the algorithm. We set $c := (5 \cdot (d-1)!^2)^{\frac{1}{d-2}}$, and we apply the proposition with subsets $S \subseteq \mathcal{C}(\mathbb{F}_q)$ such that $\#S \geq \log(q) \cdot q^{1-\frac{1}{d-2}}$ and $\#S \leq 2 \cdot q^{1-\frac{1}{g}+\frac{1}{(d-2)g}}$. Now the proposition says that there exists a function f from the set of prime powers to $\mathbb{R}_{>0}$ which converges to 1 such that the following holds: For all sets S under consideration the probability that there are at least $\#S$ points $Q \in \mathcal{C}_{ns}(\mathbb{F}_q) - (S \cup U)$ such that there exists a divisor in $\mathfrak{d}$ which splits completely into distinct points of $\mathcal{C}_{ns}(\mathbb{F}_q)$, contains Q, exactly one point of S and otherwise points of U is $\geq f(q)$ (cf. Remark 3). Let us fix such a function f.

To analyze Step 1, we set $S := \mathcal{F}_0$ and $U := \mathcal{F}_1$. Note that the assumptions on the size of S are satisfied because $\#\mathcal{F}_0 = \lceil \log(q) \cdot q^{1-\frac{1}{d-2}} \rceil$. Note also that the statement in the proposition is on divisors of the form $P_1 + \cdots + P_{d-2} + P + Q$ with $P_i \in \mathcal{F}_1$, $P \in \mathcal{F}_0$ and $Q \in \mathcal{C}(\mathbb{F}_q) - (\mathcal{F}_0 \cup \mathcal{F}_1)$. However, the conclusion of course also remains valid if we consider more divisors. The conclusion is then that the probability that a particular repetition of Step 1 leads to success is $\geq f(q)$.

For Steps ≥ 2, we set $S := \mathcal{F} \cup V(\mathcal{T})$ and $U := \mathcal{G}$. It is obvious that the assumptions are satisfied, and in the algorithm we consider exactly the same divisors as in the proposition. The conclusion is the same as in Step 1: The probability that a particular repetition leads to success is $\geq f(q)$.

References

[1] C. Diem. An Index Calculus Algorithm for Plane Curves of Small Degree. In F. Hess, S. Pauli, and M. Pohst, editors, *Algorithmic Number Theory — ANTS VII*, LNCS 4076, pages 543 – 557, Berlin, 2006. Springer.

[2] C. Diem. On arithmetic and the discrete logarithm problem in class groups of curves, 2008. Habilitation thesis.

[3] C. Diem. On the discrete logarithm problem in class groups of curves. *Math.Comp.*, 80:443 – 475, 2011.

[4] C. Diem. On the notion of bit complexity. *Bull. Eur. Assoc. Theor. Comput. Sci. EATCS*, 103:35–52, 2011. In the “Complexity Column”.

[5] C. Diem and E. Thomé. Index calculus in class groups of non-hyperelliptic curves of genus three. *J. Cryptology*, 21:593–611, 2008.

[6] P. Gaudry, E. Thomé, N. Thériault, and C. Diem. A double large prime variation for small genus hyperelliptic index calculus. *Math. Comp.*, 76:475–492, 2007.

[7] R. Hartshorne. *Algebraic Geometry.* Springer-Verlag, 1977.

[8] A. Hefez. Non-reflexive curves. *Compositio Math.*, 69:3–35, 1989.

[9] A. Hefez and S. Kleiman. Notes on the duality of projective varieties. In *Geometry today (Rome, 1984)*, volume 60 of *Progr. Math.*, pages 143–183. Birkhäuser, 1985.

[10] F. Heß. Computing Riemann-Roch spaces in algebraic function fields and related topics. *J. Symb. Comput.*, 11, 2001.

[11] S. Kleiman. The enumerative theory of singularities. In *Real and complex singularities (Proc. Ninth Nordic Summer School/NAVF Sympos. Math., Oslo, 1976)*, pages 297–396. Sijthoff and Noordhoff, Alphen aan den Rijn, 1977.

[12] V.K. Murty and J. Scherk. Effective versions of the Chebotarev density theorem for funciton fields. *C. R. Acad. Sci.*, 319:523–528, 1994.

[13] K. Nagao. Index calculus attack for Jacobian of hyperelliptic curves of small genus using two large primes. *Japan J. Indust. Appl. Math.*, 24, 2007.

[14] J. Pila. Frobenius maps of abelian varieties and fining roots of unity in finite fields. *Math. Comp.*, 55:745–763, 1990.

[15] J. Pila. Counting points on curves over families in polynomial time. Available on the arXiv under math.NT/0504570, 1991.

[16] N. Thériault. Index calculus attack for hyperelliptic curves of small genus. In *Advances in Cryptology — ASIACRYPT 2003*, volume 2894 of *LNCS*, pages 75–92. Springer-Verlag, 2003.

[17] R. Walker. *Algebraic Curves.* Springer-Verlag, 1978.

[18] A. Wallace. Tangency and duality over arbitrary fields. *Proc. Lond. Math. Soc. (3)*, 6:321–342, 1956.

[19] H. Wieland. *Finite Permutation Groups.* Academic Press, New York, 1964.

Claus Diem
University of Leipzig
Mathematical Institute
Johannisgasse 26
04103 Leipzig
Germany
diem@math.uni-leipzig.de