

## COUNTING POINTS ON CURVES USING A MAP TO $\mathbf{P}^1$

JAN TUITMAN

ABSTRACT. We introduce a new algorithm to compute the zeta function of a curve over a finite field. This method extends Kedlaya’s algorithm to a very general class of curves using a map to the projective line. We develop all the necessary bounds, analyse the complexity of the algorithm and provide some examples computed with our implementation.

### 1. INTRODUCTION

Let  $\mathbf{F}_q$  denote the finite field of characteristic  $p$  and cardinality  $q = p^n$ . Moreover, let  $\mathbf{Q}_p$  denote the field of  $p$ -adic numbers and  $\mathbf{Q}_q$  its unique unramified extension of degree  $n$ . As usual, let  $\sigma \in \text{Gal}(\mathbf{Q}_q/\mathbf{Q}_p)$  denote the unique element that lifts the  $p$ -th power Frobenius map on  $\mathbf{F}_q$ . Finally, let  $\mathbf{Z}_q$  denote the ring of integers of  $\mathbf{Q}_q$ , so that  $\mathbf{Z}_q/p\mathbf{Z}_q \cong \mathbf{F}_q$ . Suppose that  $X$  is a smooth proper algebraic curve of genus  $g$  over  $\mathbf{F}_q$ . Recall that the zeta function of  $X$  is defined as

$$Z(X, T) = \exp \left( \sum_{i=1}^{\infty} |X(\mathbf{F}_{q^i})| \frac{T^i}{i} \right).$$

It follows from the Weil conjectures that  $Z(X, T)$  is of the form

$$\frac{\chi(T)}{(1-T)(1-qT)},$$

with  $\chi(T) \in \mathbf{Z}[T]$  a polynomial of degree  $2g$ , the inverse roots of which have complex absolute value  $q^{\frac{1}{2}}$  and are permuted by the map  $t \rightarrow q/t$ . Moreover, by the Lefschetz formula for rigid cohomology, we have that

$$\chi(T) = \det \left( 1 - T F_p^n | H_{\text{rig}}^1(X) \right),$$

where  $F_p$  denotes the  $p$ -th power Frobenius map.

In [Ked01], Kedlaya showed how  $Z(X, T)$  can be determined efficiently, in the case when  $X$  is a hyperelliptic curve and the characteristic  $p$  is odd, by explicitly computing the action of  $F_p$  on  $H_{\text{rig}}^1(X)$ . His algorithm was then extended to characteristic 2 [DV06b] and also to superelliptic curves [GG01],  $C_{ab}$  curves [DV06a] and nondegenerate curves [CDV06]. However, for  $C_{ab}$  and nondegenerate curves these algorithms have proved a lot less efficient in practice than for hyperelliptic and superelliptic curves. The main reason for this is that the algorithms for  $C_{a,b}$  and nondegenerate curves use a more complicated Frobenius lift that no longer sends  $x$  to  $x^p$ . Moreover, in the case of nondegenerate curves, the linear algebra that is used to compute in the cohomology is not very efficient and when the curve admits a

---

Received by the editor March 7, 2014 and, in revised form, September 10, 2014.  
 2010 *Mathematics Subject Classification*. Primary 11Y99, 11M38.

low degree map to the projective line, as is the case for most nondegenerate curves, this is not fully exploited.

In this paper we propose a new algorithm for computing  $Z(X, T)$  that avoids these problems and can be applied to more general curves as well. Our approach combines Kedlaya's original algorithm and Lauder's fibration method [Lau06]. In the work of Lauder, the Frobenius lift is computed by solving a  $p$ -adic differential equation. For curves it turns out to be more efficient to compute the Frobenius lift directly by Hensel lifting as in Kedlaya's algorithm, especially since this allows one to avoid the radix conversions that take up most of the time in the examples of the fibration method computed by Walker in his thesis [Wal10].

Our approach can be summarised as follows. We start with a finite separable map  $x$  from the curve  $X$  to the projective line. After removing the ramification locus of  $x$  from the curve, we can choose a Frobenius lift that sends  $x$  to  $x^p$ , which we compute by Hensel lifting as in Kedlaya's algorithm. We then compute in the cohomology as in Lauder's fibration method to find the matrix of Frobenius and the zeta function of  $X$ .

Let  $x : X \rightarrow \mathbf{P}_{\mathbf{F}_q}^1$  be a finite separable map of degree  $d_x$  and  $y : X \rightarrow \mathbf{P}_{\mathbf{F}_q}^1$  a rational function that generates the function field of  $X$  over  $\mathbf{F}_q(x)$ , such that  $Q(x, y) = 0$  where  $Q \in \mathbf{F}_q[x, y]$  is irreducible and monic in  $y$  (of degree  $d_x$ ). The polynomial  $Q$  is the natural input to our algorithm. The degree of  $Q$  in  $x$  will be denoted by  $d_y$ . The time complexity of the algorithm is then  $\tilde{O}(pd_x^6 d_y^4 n^3)$  by Theorem 4.6 and the space complexity  $\tilde{O}(pd_x^4 d_y^3 n^3)$  by Theorem 4.7.

When  $Q$  is nondegenerate with respect to its Newton polygon  $\Gamma$ , which is common in the sense of [CDV06, §7.2], we have that  $d_x d_y \in O(g)$ . The time and space complexity of the algorithm are then  $\tilde{O}(pg^6 n^3)$  and  $\tilde{O}(pg^4 n^3)$ , respectively. Note that this slightly improves the complexity estimate from [CDV06]. Now if additionally we fix  $d_x$ , then  $d_y \in O(g)$ , so that the time and space complexities of the algorithm are  $\tilde{O}(pg^4 n^3)$  and  $\tilde{O}(pg^3 n^3)$ , respectively. This extends the complexity estimate from [Ked01] from the case where  $d_x = 2$  to the case where  $d_x$  is only fixed.

Note that the time and space complexities of our algorithm are quasilinear in  $p$  and hence not polynomial in the size of the input which is  $\log(p)d_x d_y n$ . This is also the case for Kedlaya's algorithm and the algorithm from [CDV06]. However, for hyperelliptic curves, the dependence on  $p$  of the time and space complexities of Kedlaya's algorithm has been improved to  $\tilde{O}(p^{1/2})$  [Har07] and average polynomial time [Har14] by Harvey. It is an interesting problem whether these ideas can be used to improve the dependence on  $p$  of the complexity of our algorithm as well.

We need some assumptions for the algorithm to work. First, we assume that we have a lift  $\mathcal{Q} \in \mathbf{Z}_q[x, y]$  of the polynomial  $Q$  such that Assumption 1 below is satisfied. This basically means that over  $\mathbf{Q}_q$  the branch points of the map  $x$  and the points lying over it are all distinct modulo  $p$ . Second, we assume that the zero locus of  $\mathcal{Q}$  in the affine plane with coordinates  $x, y$  is smooth. The first of these assumptions is essential, but the second one can probably be removed, as sketched in Section 4. Finally, we suppose that we can compute certain integral bases in function fields and exclude the time and space required to do so from our complexity estimates.

We have written a publicly available implementation of our algorithm in the computer algebra package Magma [BCP97]. This implementation turns out to be

quite practical and seems to work for almost all polynomials  $Q$  as illustrated by the example files that come with the code. This should be contrasted with the algorithm from [CDV06], which was never fully implemented because it was expected not to be practical. Indeed, in some special cases where we have compared our new algorithm against our experimental implementation of the algorithm from [CDV06], the new algorithm runs faster by at least two orders of magnitude.

The author was supported by FWO-Vlaanderen. We thank the referees for their useful comments and suggestions.

2. LIFTING THE CURVE AND FROBENIUS

Recall that  $X$  is a smooth proper algebraic curve of genus  $g$  over the finite field  $\mathbf{F}_q$  of characteristic  $p$  and cardinality  $q = p^n$ . Let  $x : X \rightarrow \mathbf{P}_{\mathbf{F}_q}^1$  be a finite separable map of degree  $d_x$  and  $y : X \rightarrow \mathbf{P}_{\mathbf{F}_q}^1$  a rational function that generates the function field of  $X$  over  $\mathbf{F}_q(x)$ , such that  $Q(x, y) = 0$  where  $Q \in \mathbf{F}_q[x, y]$  is irreducible and monic in  $y$  (of degree  $d_x$ ). The degree of  $Q$  in  $x$  will be denoted by  $d_y$ . Let  $\mathcal{Q} \in \mathbf{Z}_q[x, y]$  be a lift of  $Q$  that contains the same monomials in its support as  $Q$  and is still monic in  $y$ .

**Proposition 2.1.** *The ring  $\mathcal{A} = \mathbf{Z}_q[x, y]/(\mathcal{Q})$  is a free module of rank  $d_x$  over  $\mathbf{Z}_q[x]$  and a basis is given by  $[1, y, \dots, y^{d_x-1}]$ .*

*Proof.* This follows from the fact that  $\mathcal{Q}$  is monic in  $y$ . □

**Definition 2.2.** We let  $\Delta(x) \in \mathbf{Z}_q[x]$  denote the discriminant of  $\mathcal{Q}$  with respect to the variable  $y$  and  $r(x) \in \mathbf{Z}_q[x]$  the squarefree polynomial  $r = \Delta/(\gcd(\Delta, \frac{d\Delta}{dx}))$ . Note that  $\Delta(x) \not\equiv 0 \pmod{p}$  since  $x : X \rightarrow \mathbf{P}_{\mathbf{F}_q}^1$  is separable. We denote

$$\mathcal{S} = \mathbf{Z}_q[x, 1/r], \quad \mathcal{R} = \mathbf{Z}_q[x, 1/r, y]/(\mathcal{Q}),$$

and write  $\mathcal{V} = \text{Spec } \mathcal{S}$ ,  $\mathcal{U} = \text{Spec } \mathcal{R}$ , so that  $x$  defines a finite étale morphism from  $\mathcal{U}$  to  $\mathcal{V}$ . Finally, we let  $U = \mathcal{U} \otimes_{\mathbf{Z}_q} \mathbf{F}_q$ ,  $V = \mathcal{V} \otimes_{\mathbf{Z}_q} \mathbf{F}_q$  denote the special fibres and  $\mathbb{U} = \mathcal{U} \otimes_{\mathbf{Z}_q} \mathbf{Q}_q$ ,  $\mathbb{V} = \mathcal{V} \otimes_{\mathbf{Z}_q} \mathbf{Q}_q$  the generic fibres of  $\mathcal{U}$  and  $\mathcal{V}$ , respectively.

**Assumption 1.** *We will assume that:*

- (1) *There exists a smooth proper curve  $\mathcal{X}$  over  $\mathbf{Z}_q$  and a smooth relative divisor  $\mathcal{D}_{\mathcal{X}}$  on  $\mathcal{X}$  such that  $\mathcal{U} = \mathcal{X} \setminus \mathcal{D}_{\mathcal{X}}$ .*
- (2) *There exists a smooth relative divisor  $\mathcal{D}_{\mathbf{P}^1}$  on  $\mathbf{P}_{\mathbf{Z}_q}^1$  such that  $\mathcal{V} = \mathbf{P}_{\mathbf{Z}_q}^1 \setminus \mathcal{D}_{\mathbf{P}^1}$ .*

We write  $\mathbb{X} = \mathcal{X} \otimes \mathbf{Q}_q$  for the generic fibre of  $\mathcal{X}$ .

*Remark 2.3.* A relative divisor  $\mathcal{D}$  on a smooth curve over  $\mathbf{Z}_q$  is smooth over  $\mathbf{Z}_q$  if and only if it is reduced and all of the points in its support are smooth over  $\mathbf{Z}_q$ , or equivalently if and only if it reduces modulo  $p$  to a reduced divisor  $D$ . Hence by Assumption 1, all branch points of the map  $x$  restricted to  $\mathbb{X}$ , and all points on  $\mathbb{X}$  lying over these branch points, are distinct modulo  $p$ .

At every point  $P \in \mathcal{X} \setminus \mathcal{U}$ , we let  $z_P$  denote an étale local coordinate on  $\mathcal{X}$ . By a slight abuse of notation, we write  $\text{ord}_P(\cdot)$  for the discrete valuation on  $\mathcal{O}_{\mathbb{X}, P}$ . We let  $e_P$  denote the ramification index of the map  $x$ . Note that the  $e_P$  are the same on  $X$  as on  $\mathbb{X}$ , since they can only increase under reduction modulo  $p$ , but add up to  $d_x$  in every fibre.

**Assumption 2.** *We will assume that the zero locus of  $\mathcal{Q}(x, y)$  in  $\mathbf{A}_{\mathbf{Q}_q}^2$  is smooth.*

**Proposition 2.4.** *The element*

$$s(x, y) = r(x) / \frac{\partial \mathcal{Q}}{\partial y}$$

of  $\mathbf{Q}_q(x, y)$  is contained in  $\mathcal{A}$ .

*Proof.* For  $k \in \mathbf{N}$ , we let  $W_k$  denote the free  $\mathbf{Z}_q[x]$ -module of polynomials in  $\mathbf{Z}_q[x, y]$  of degree at most  $k - 1$  in the variable  $y$ . Let  $\Sigma$  be the matrix of the  $\mathbf{Z}_q[x]$ -module homomorphism:

$$(1) \quad W_{d_x-1} \oplus W_{d_x} \rightarrow W_{2d-1}, \quad (a, b) \mapsto a\mathcal{Q} + b\frac{\partial \mathcal{Q}}{\partial y},$$

with respect to the bases  $[1, y, \dots, y^{d_x-2}]$ ,  $[1, y, \dots, y^{d_x-1}]$  and  $[1, y, \dots, y^{2d_x-2}]$ . By definition we have  $\Delta = \det(\Sigma)$ , so that  $\Delta$  is contained in the image of (1) and  $\Delta(x) / \frac{\partial \mathcal{Q}}{\partial y}$  is contained in  $\mathcal{A}$ . By Assumption 2, the ring  $\mathcal{A} \otimes \mathbf{Q}_q$  is the integral closure of  $\mathbf{Q}_q[x]$  in  $\mathbf{Q}_q(x, y)$ . Note that the basis  $[1, y, \dots, y^{d_x-1}]$  of  $\mathcal{A} \otimes \mathbf{Q}_q$  is therefore an integral basis for  $\mathbf{Q}_q(x, y)$  over  $\mathbf{Q}_q[x]$ . Since  $\mathcal{Q}$  is monic in  $y$ , for any irreducible polynomial  $\pi \in \mathbf{Q}_q[x]$  the element  $\frac{\partial \mathcal{Q}}{\partial y} / \pi$  of  $\mathbf{Q}_q(x, y)$  is not integral at the place  $(\pi)$ , and hence its inverse  $\pi / \frac{\partial \mathcal{Q}}{\partial y}$  is integral (even zero) at  $(\pi)$ . Hence  $s$  is contained in  $\mathcal{A}$ . □

**Definition 2.5.** We denote the ring of over-convergent functions on  $\mathcal{U}$  by

$$\mathcal{R}^\dagger = \mathbf{Z}_q\langle x, 1/r, y \rangle^\dagger / (\mathcal{Q}).$$

Note that  $\mathcal{R}^\dagger$  is a free module of rank  $d_x$  over  $S^\dagger = \mathbf{Z}_q\langle x, 1/r \rangle^\dagger$  and that a basis is given by  $[y^0, \dots, y^{d_x-1}]$ . A Frobenius lift  $F_p : \mathcal{R}^\dagger \rightarrow \mathcal{R}^\dagger$  is defined as a  $\sigma$ -semilinear ring homomorphism that reduces modulo  $p$  to the  $p$ -th power Frobenius map.

**Theorem 2.6.** *There exists a Frobenius lift  $F_p : \mathcal{R}^\dagger \rightarrow \mathcal{R}^\dagger$  for which  $F_p(x) = x^p$ .*

*Proof.* Define sequences  $(\alpha_i)_{i \geq 0}, (\beta_i)_{i \geq 0}$ , with  $\alpha_i \in S^\dagger$  and  $\beta_i \in \mathcal{R}^\dagger$ , by the following recursion:

$$\begin{aligned} \alpha_0 &= \frac{1}{r^p}, \\ \beta_0 &= y^p, \\ \alpha_{i+1} &= \alpha_i(2 - \alpha_i r^\sigma(x^p)) && (\text{mod } p^{2^{i+1}}), \\ \beta_{i+1} &= \beta_i - \mathcal{Q}^\sigma(x^p, \beta_i) s^\sigma(x^p, \beta_i) \alpha_i && (\text{mod } p^{2^{i+1}}). \end{aligned}$$

Then one easily checks that the  $\sigma$ -semilinear ring homomorphism  $F_p : \mathcal{R}^\dagger \rightarrow \mathcal{R}^\dagger$  defined by

$$F_p(x) = x^p, \quad F_p(1/r) = \lim_{i \rightarrow \infty} \alpha_i, \quad F_p(y) = \lim_{i \rightarrow \infty} \beta_i,$$

is a Frobenius lift. □

**Proposition 2.7.** *Let  $G \in M_{d_x \times d_x}(\mathbf{Z}_q[x, 1/r])$  denote the matrix such that*

$$d(y^j) = \sum_{i=0}^{d_x-1} G_{i+1, j+1} y^i dx,$$

for all  $0 \leq j \leq d_x - 1$ . Then we can write  $G = M/r$  with  $M \in M_{d_x \times d_x}(\mathbf{Z}_q[x])$ .

*Proof.* This follows from the formula

$$(2) \quad d(y^j) = -jy^{(j-1)} \left(\frac{s}{r}\right) \frac{\partial Q}{\partial x} dx. \quad \square$$

In the terminology of the fibration method,  $Gdx$  is the matrix of the Gauss–Manin connection  $\nabla$  on the 0-th higher direct image  $\mathbf{R}^0x_*(\mathcal{O}_{\mathbb{U}})$  with respect to the basis  $[1, y, \dots, y^{d_x-1}]$ . By Proposition 2.7, this matrix has at most a simple pole at all points  $\neq \infty$  in the support of  $\mathcal{D}_{\mathbf{P}^1}$ . At  $x = \infty$  we will have to make a change of basis for this to be the case.

**Assumption 3.** *We will assume that a matrix  $W^\infty \in Gl_{d_x}(\mathbf{Z}_q[x, x^{-1}])$  is known such that if we denote  $b_j^\infty = \sum_{i=0}^{d_x-1} W_{i+1, j+1}^\infty y^i$  for all  $0 \leq j \leq d_x - 1$ , then  $[b_0^\infty, \dots, b_{d_x-1}^\infty]$  is an integral basis for  $\mathbf{Q}_q(x, y)$  over  $\mathbf{Q}_q[x^{-1}]$ .*

**Proposition 2.8.** *Let  $G^\infty \in M_{d_x \times d_x}(\mathbf{Z}_q[x, x^{-1}, 1/r])$  denote the matrix such that*

$$db_j^\infty = \sum_{i=0}^{d_x-1} G_{i+1, j+1}^\infty b_i^\infty dx,$$

*for all  $0 \leq j \leq d_x - 1$ . Then  $G^\infty dx$  has at most a simple pole at  $x = \infty$ .*

*Proof.* We denote  $t = 1/x$  and let  $H \in M_{d_x \times d_x}(\mathbf{Q}_q(t))$  be defined by  $H(t)dt = G^\infty(x)dx$ . Note that  $\text{ord}_P(dt/t) = -1$  at every point  $P \in \mathcal{X} \setminus \mathcal{U}$  lying over  $t = 0$ . At every such  $P$  and for all  $0 \leq i \leq d_x - 1$  we clearly have  $\text{ord}_P(db_i^\infty) \geq 0$ , so that  $\text{ord}_P(tdb_i^\infty) - \text{ord}_P(dt) \geq 1$ . Since  $[b_0^\infty, \dots, b_{d_x-1}^\infty]$  is an integral basis for  $\mathbf{Q}_q(x, y)$  over  $\mathbf{Q}_q[t]$ , we conclude that  $tH$  does not have a pole at  $t = 0$ , so that  $Hdt$  has at most a simple pole there.  $\square$

**Definition 2.9.** Let  $x_0 \neq \infty$  be a geometric point of  $\mathbf{P}^1(\bar{\mathbf{Q}}_q)$ . The exponents of  $Gdx$  at  $x_0$  are defined as the eigenvalues of the residue matrix  $(x - x_0)G|_{x=x_0}$ . Moreover, the exponents of  $G^\infty dx$  at  $x = \infty$  are defined as its exponents at  $t = 0$ , after substituting  $x = 1/t$ .

**Proposition 2.10.** *The exponents of  $Gdx$  at any point  $x_0 \neq \infty$  and the exponents of  $G^\infty dx$  at  $x = \infty$  are elements of  $\mathbf{Q} \cap \mathbf{Z}_p$  and are contained in the interval  $[0, 1)$ .*

*Proof.* Let  $\lambda \in \bar{\mathbf{Q}}_q$  denote an exponent of  $Gdx$  at  $x_0 \neq \infty$ . Then there exists  $f = \sum_{i=0}^{d_x-1} a_i y^i$  with  $a_0, \dots, a_{d_x-1} \in \bar{\mathbf{Q}}_q$  such that

$$(3) \quad df = \left(\frac{\lambda f}{x - x_0} + g\right) dx$$

as 1-forms on  $\mathbb{U} \otimes \bar{\mathbf{Q}}_q$ , where  $g \in \mathcal{O}(\mathbb{U} \otimes \bar{\mathbf{Q}}_q)$  satisfies  $\text{ord}_P(g) \geq 0$  at all points  $P \in x^{-1}(x_0)$ . Note that for at least one  $P \in x^{-1}(x_0)$  we have  $\text{ord}_P(f) < \text{ord}_P(x - x_0)$ , since otherwise  $f/(x - x_0)$  would be integral over  $\mathbf{Q}_q[x]$ , contradicting Assumption 2. For such a  $P$ , dividing by  $f$  in (3) and taking residues, we obtain

$$\text{ord}_P(f) = \lambda \text{ord}_P(x - x_0) = \lambda e_P.$$

Since  $0 \leq \text{ord}_P(f) < \text{ord}_P(x - x_0)$ , we see that  $\lambda \in \mathbf{Q} \cap [0, 1)$ . By Assumption 1, elements of  $\mathcal{S}$  have  $p$ -adically integral Laurent series expansions at  $x_0$ , so that  $(x - x_0)G|_{x=x_0} \in M_{d_x \times d_x}(\mathbf{Z}_q)$ . Since  $p$ -adically integral matrices have  $p$ -adically integral eigenvalues, we conclude that  $\lambda \in \mathbf{Z}_p$ . To obtain the same result for the exponents of  $G^\infty dx$  at  $x = \infty$ , replace  $x_0$  by  $\infty$  and  $(x - x_0)$  by  $t = 1/x$  in the argument.  $\square$

**Definition 2.11.** For a geometric point  $x_0 \in \mathbf{P}^1(\bar{\mathbf{Q}}_q)$ , we let  $\text{ord}_{x_0}(\cdot)$  denote the discrete valuation on  $\bar{\mathbf{Q}}_q(x)$  corresponding to  $x_0$ . We extend these definitions to matrices over  $\bar{\mathbf{Q}}_q(x)$  by taking the minimum over their entries.

**Proposition 2.12.** *Let  $N \in \mathbf{N}$  be a positive integer.*

- (1) *The element  $F_p(1/r)$  of  $\mathcal{S}^\dagger$  is congruent modulo  $p^N$  to*

$$\sum_{i=p}^{pN} \frac{\rho_i(x)}{r^i},$$

where  $\rho_i \in \mathbf{Z}_q[x]$  satisfies  $\deg(\rho_i) < \deg(r)$  for all  $p \leq i \leq pN$ .

- (2) *For all  $0 \leq i \leq d_x - 1$ , the element  $F_p(y^i)$  of  $\mathcal{R}^\dagger$  is congruent modulo  $p^N$  to  $\sum_{j=0}^{d_x-1} \phi_{i,j}(x)y^j$ , where*

$$\phi_{i,j} = \sum_{k=0}^{p(N-1)} \frac{\phi_{i,j,k}(x)}{r^k}$$

for all  $0 \leq j \leq d_x - 1$  and  $\phi_{i,j,k} \in \mathbf{Z}_q[x]$  satisfies

$$\begin{aligned} \deg(\phi_{i,j,0}) &\leq -\text{ord}_\infty(W^\infty) - p \text{ord}_\infty((W^\infty)^{-1}), \\ \deg(\phi_{i,j,k}) &< \deg(r), \end{aligned}$$

for all  $0 \leq j \leq d_x - 1$  and  $1 \leq k \leq p(N - 1)$ .

- (3) *For all  $0 \leq i \leq d_x - 1$ , the element  $F_p(y^i/r)$  of  $\mathcal{R}^\dagger$  is congruent modulo  $p^N$  to  $\sum_{j=0}^{d_x-1} \psi_{i,j}(x)(y^j/r)$ , where*

$$\psi_{i,j} = \sum_{k=0}^{pN-1} \frac{\psi_{i,j,k}(x)}{r^k}$$

for all  $0 \leq j \leq d_x - 1$  and  $\psi_{i,j,k} \in \mathbf{Z}_q[x]$  satisfies

$$\begin{aligned} \deg(\psi_{i,j,0}) &\leq -\text{ord}_\infty(W^\infty) - p \text{ord}_\infty((W^\infty)^{-1}) - (p - 1) \deg(r), \\ \deg(\psi_{i,j,k}) &< \deg(r), \end{aligned}$$

for all  $0 \leq j \leq d_x - 1$  and  $1 \leq k \leq pN - 1$ .

*Proof.*

- (1) Since  $r^\sigma(x^p) \equiv r^p \pmod{p}$ , this follows from

$$F_p\left(\frac{1}{r}\right) = \frac{1}{r^\sigma(x^p)} = \frac{1}{r^p} \left(1 - \frac{r^p - r^\sigma(x^p)}{r^p}\right)^{-1} = \frac{1}{r^p} \sum_{i=0}^\infty \left(\frac{r^p - r^\sigma(x^p)}{r^p}\right)^i.$$

- (2) The matrix  $\Phi = (\phi_{i,j}) \in M_{d_x \times d_x}(\mathcal{S}^\dagger)$  defines a  $p$ -th power Frobenius structure on the higher direct image  $\mathbf{R}^0 x_* (\mathcal{O}_U)$ . By definition we have  $\text{ord}_p(\Phi) \geq 0$  and by Poincaré duality we find that  $\text{ord}_p(\Phi^{-1}) \geq 0$  as well. The result now follows from a theorem of Kedlaya and the author [KT12, Corollary 2.6] using Proposition 2.10.

- (3) Analogous to (2). □

3. COMPUTING (IN) THE COHOMOLOGY

**Definition 3.1.** The rigid cohomology of  $U$  in degree 1 can be defined as

$$H_{\text{rig}}^1(U) = \text{coker}(d : \mathcal{R}^\dagger \rightarrow \Omega^1(\mathbb{U}) \otimes \mathcal{R}^\dagger).$$

**Theorem 3.2.**

$$H_{\text{rig}}^1(U) \cong H_{dR}^1(\mathbb{U}).$$

*Proof.* This follows as a special case from the comparison theorem between rigid and de Rham cohomology of Baldassarri and Chiarellotto [BC94], since by Assumption 1  $\mathcal{D}_{\mathcal{X}}$  is smooth over  $\mathbf{Z}_q$ . □

We can effectively reduce any 1-form to one of low pole order using linear algebra following the work of Laufer [Lau06]. The procedure consists of two parts, reducing the pole order at the points not lying over  $x = \infty$  and at those lying over  $x = \infty$ , respectively. From now on we let  $r'$  denote the polynomial  $\frac{dx}{dx}$ . We start with the points not lying over  $x = \infty$ .

**Proposition 3.3.** For all  $\ell \in \mathbf{N}$  and every vector  $w \in \mathbf{Q}_q[x]^{\oplus d_x}$ , there exist vectors  $u, v \in \mathbf{Q}_q[x]^{\oplus d_x}$  with  $\deg(v) < \deg(r)$ , such that

$$\frac{\sum_{i=0}^{d_x-1} w_i y^i}{r^\ell} \frac{dx}{r} = d \left( \frac{\sum_{i=0}^{d_x-1} v_i y^i}{r^\ell} \right) + \frac{\sum_{i=0}^{d_x-1} u_i y^i}{r^{\ell-1}} \frac{dx}{r}.$$

*Proof.* Note that since  $r$  is separable,  $r'$  is invertible in the ring  $\mathbf{Q}_q[x]/(r)$ . One checks that  $v$  has to satisfy the  $d_x \times d_x$  linear system

$$\left( \frac{M}{r'} - \ell I \right) v \equiv \frac{w}{r'} \pmod{r}$$

over  $\mathbf{Q}_q[x]/(r)$ . However, since  $\ell \geq 1$  is not an exponent of  $Gdx$  by Proposition 2.10, we have that  $\det(\ell I - M/r')$  is invertible in  $\mathbf{Q}_q[x]/(r)$ , so that this system has a unique solution  $v$ . We now take

$$u = \frac{w - (M - \ell r' I) v}{r} - \frac{dv}{dx}. \quad \square$$

We now move on to the points lying over  $x = \infty$ .

**Proposition 3.4.** For every vector  $w \in \mathbf{Q}_q[x, x^{-1}]^{\oplus d_x}$  with

$$\text{ord}_\infty(w) \leq -\deg(r),$$

there exist vectors  $u, v \in \mathbf{Q}_q[x, x^{-1}]^{\oplus d_x}$  with  $\text{ord}_\infty(u) > \text{ord}_\infty(w)$  such that

$$\left( \sum_{i=0}^{d_x-1} w_i b_i^\infty \right) \frac{dx}{r} = d \left( \sum_{i=0}^{d_x-1} v_i b_i^\infty \right) + \left( \sum_{i=0}^{d_x-1} u_i b_i^\infty \right) \frac{dx}{r}.$$

*Proof.* We still denote  $t = 1/x$ . By Proposition 2.8, we can expand

$$G^\infty dx = \left( \frac{G_{-1}^\infty}{t} + G_0^\infty + \dots \right) dt,$$

where  $G_i^\infty \in M_{d_x \times d_x}(\mathbf{Q}_q)$  for all  $i \geq -1$ . Writing  $m = -\text{ord}_\infty(w) - \text{deg}(r) + 1$ , we can also expand

$$w \frac{dx}{r} = \sum_{j=-(m+1)}^\infty \bar{w}_j t^j dt,$$

where  $\bar{w}_j \in \mathbf{Q}_q^{\oplus d_x}$  for all  $j \geq -(m+1)$ . Note that  $m \geq 1$ . By Proposition 2.10, we have that  $\det(mI - G_{-1}^\infty)$  is nonzero, so that the linear system

$$(G_{-1}^\infty - mI)\bar{v} = \bar{w}_{-(m+1)}$$

has a unique solution  $\bar{v} \in \mathbf{Q}_q^{\oplus d_x}$ . We can now take

$$v = \bar{v}x^m, \quad u = w - r \left( G^\infty v + \frac{dv}{dx} \right). \quad \square$$

*Remark 3.5.* Note that when  $\text{ord}_\infty(w) \leq \text{ord}_0(W^\infty) - \text{deg}(r) + 1$ , we have that  $\text{ord}_0(v) \geq -\text{ord}_0(W^\infty)$ , so that the function  $\sum_{i=0}^{d_x-1} v_i b_i^\infty$  only has poles at points lying over  $x = \infty$ .

We now give an explicit description of the cohomology space  $H_{\text{rig}}^1(U)$ .

**Theorem 3.6.** *Define the following  $\mathbf{Q}_q$ -vector spaces:*

$$\begin{aligned} E_0 &= \left\{ \left( \sum_{i=0}^{d_x-1} u_i(x)y^i \right) \frac{dx}{r} : u \in \mathbf{Q}_q[x]^{\oplus d_x} \right\}, \\ E_\infty &= \left\{ \left( \sum_{i=0}^{d_x-1} u_i(x, x^{-1})b_i^\infty \right) \frac{dx}{r} : u \in \mathbf{Q}_q[x, x^{-1}]^{\oplus d_x}, \right. \\ &\quad \left. \text{ord}_\infty(u) > \text{ord}_0(W^\infty) - \text{deg}(r) + 1 \right\}, \\ B_0 &= \left\{ \sum_{i=0}^{d_x-1} v_i(x)y^i : v \in \mathbf{Q}_q[x]^{\oplus d_x} \right\}, \\ B_\infty &= \left\{ \sum_{i=0}^{d_x-1} v_i(x, x^{-1})b_i^\infty : v \in \mathbf{Q}_q[x, x^{-1}]^{\oplus d_x}, \text{ord}_\infty(v) > \text{ord}_0(W^\infty) \right\}. \end{aligned}$$

Then  $E_0 \cap E_\infty$  and  $d(B_0 \cap B_\infty)$  are finite dimensional  $\mathbf{Q}_q$ -vector spaces and

$$H_{\text{rig}}^1(U) \cong (E_0 \cap E_\infty) / d(B_0 \cap B_\infty).$$

*Proof.* First, note that elements of  $E_0, B_0$  have bounded poles everywhere but at the points lying over  $x = \infty$  and elements of  $E_\infty, B_\infty$  everywhere but at the points lying over  $x = 0$ . So elements of  $E_0 \cap E_\infty$  and  $d(B_0 \cap B_\infty)$  have bounded poles everywhere on  $\mathbb{X}$ . Hence these vector spaces are contained in the space of global sections of some line bundle on  $\mathbb{X}$  and are therefore finite dimensional.

Next, we show that every class in  $H_{\text{rig}}^1(U)$  can be represented by a 1-form in  $E_0 \cap E_\infty$ . Note that by Theorem 3.2 we can restrict to classes in  $H_{\text{dR}}^1(\mathbb{U})$ . Now every such class can be represented by a 1-form in  $E_0$  by (repeatedly) applying Proposition 3.3. Then we change basis by the matrix  $W^\infty$  from Assumption 3. Observe that this change of basis might introduce a pole at  $x = 0$ . Now our

cohomology class can be represented by 1-form in  $E_0 \cap E_\infty$  by (repeatedly) applying Proposition 3.4 and Remark 3.5.

Finally, we have to prove that if a 1-form  $\omega \in E_0 \cap E_\infty$  is exact, then it lies in  $d(B_0 \cap B_\infty)$ . So let  $\omega \in E_0 \cap E_\infty$  denote such an exact 1-form. From Assumption 2 and the definition of  $[b_0^\infty, \dots, b_{d_x-1}^\infty]$ , it follows that  $\text{ord}_P(\omega) \geq -1$  all points  $P$  not lying over  $x = \infty$  and  $\text{ord}_P(\omega) \geq (\text{ord}_0(W^\infty) + 1)e_P - 1$  at all points  $P$  lying over  $x = \infty$ . Note that the exterior derivative lowers the order by at most 1. So if  $\omega = df$  for some  $f \in \mathcal{O}(U)$ , then  $\text{ord}_P(f) \geq 0$  at all points  $P$  not lying over  $x = \infty$  and  $\text{ord}_P(f) \geq (\text{ord}_0(W^\infty) + 1)e_P$  at all points  $P$  lying over  $x = \infty$ . Using Assumption 2 and the definition of  $[b_0^\infty, \dots, b_{d_x-1}^\infty]$  again, it follows that  $f$  is an element of  $B_0 \cap B_\infty$ .  $\square$

Note that by the proof of Theorem 3.6, we can effectively reduce any 1-form to one in  $E_0 \cap E_\infty$  with the same cohomology class. However, the reduction procedure will introduce  $p$ -adic denominators and therefore suffer from loss of  $p$ -adic precision. In the following two propositions we bound these denominators. Our bounds and their proofs generalise the ones from [Ked01].

**Proposition 3.7.** *Let  $\omega \in \Omega^1(\mathcal{U})$  be of the form*

$$\omega = \frac{\sum_{i=0}^{d_x-1} w_i y^i dx}{r^\ell},$$

where  $\ell \in \mathbf{N}$  and  $w \in \mathbf{Z}_q[x]^{\oplus d_x}$  satisfies  $\text{deg}(w) < \text{deg}(r)$ . We define

$$e = \max\{e_P \mid P \in \mathcal{X} \setminus \mathcal{U}, x(P) \neq \infty\}.$$

If we represent the class of  $\omega$  in  $H_{\text{rig}}^1(U)$  by

$$\left( \sum_{i=0}^{d_x-1} u_i y^i \right) \frac{dx}{r},$$

with  $u \in \mathbf{Q}_q[x]^{\oplus d_x}$  as in the proof of Theorem 3.6, then

$$p^{\lfloor \log_p(\ell e) \rfloor} u \in \mathbf{Z}_q[x]^{\oplus d_x}.$$

*Proof.* We have

$$\omega = df + \left( \sum_{i=0}^{d_x-1} u_i y^i \right) \frac{dx}{r}$$

with  $f = \sum_{j=1}^\ell (\sum_{i=0}^{d_x-1} (v_j)_i y^i) / r^j$ , where  $v_j \in \mathbf{Q}_q[x]^{\oplus d_x}$  satisfies  $\text{deg}(f_j) < \text{deg}(r)$  for all  $1 \leq j \leq \ell$ . Note that it is sufficient to show that  $p^{\lfloor \log_p(\ell e) \rfloor} f \in \mathcal{R}$ . By Assumption 1, we have that

$$\mathcal{O}(\mathcal{X} - x^{-1}(\infty)) / (r)^k \cong \prod_{P \in \mathcal{X} \setminus \mathcal{U}, x(P) \neq \infty} \mathcal{O}_{\mathcal{X},P} / (z_P^{e_P})^k,$$

for all  $k \in \mathbf{N}$ . Moreover, we have that  $\mathcal{O}(\mathbb{X} - x^{-1}(\infty)) \cong \mathcal{A} \otimes \mathbf{Q}_q$  by Assumption 2. To show that  $p^{\lfloor \log_p(\ell e) \rfloor} f$  is integral, it is therefore enough to show that for every  $P \in \mathcal{X} \setminus \mathcal{U}$  with  $x(P) \neq 0$ , the Laurent series expansion

$$a_{-\ell e_P} z_P^{-\ell e_P} + \dots + a_{-e_P-1} z_P^{-e_P-1} + \mathcal{O}(z_P^{-e_P})$$

of  $p^{\lfloor \log_p(\ell e) \rfloor} f$  is integral. However, the differential  $df$  has a pole of order at most  $\ell e_P + 1$  at  $P$ , and its Laurent series expansion

$$\left( b_{-\ell e_P-1} z_P^{-\ell e_P-1} + \dots + b_{-e_P-2} z_P^{-e_P-2} + \mathcal{O}(z_P^{-e_P-1}) \right) dz_P$$

is integral since  $\omega$  is integral. The worst denominator we get by integrating this series is therefore  $p^{\lfloor \log_p(\ell e) \rfloor}$  and the result follows.  $\square$

**Proposition 3.8.** *Let  $\omega \in \Omega^1(\mathcal{U})$  be of the form*

$$\omega = \left( \sum_{i=0}^{d_x-1} w_i(x, x^{-1}) b_i^\infty \right) \frac{dx}{r},$$

where  $w \in \mathbf{Z}_q[x, x^{-1}]^{\oplus d_x}$  satisfies  $\text{ord}_\infty(w) \leq \text{ord}_0(W^\infty) - \text{deg}(r) + 1$ . We define

$$m = -\text{ord}_\infty(w) - \text{deg}(r) + 1, \\ e_\infty = \max\{e_P \mid P \in \mathcal{X} \setminus \mathcal{U}, x(P) = \infty\}.$$

If we represent the class of  $\omega$  in  $H_{\text{rig}}^1(U)$  by

$$\left( \sum_{i=0}^{d_x-1} u_i y^i \right) \frac{dx}{r},$$

with  $u \in \mathbf{Q}_q[x, x^{-1}]^{\oplus d_x}$  such that  $\text{ord}_\infty(u) > \text{ord}_0(W^\infty) - \text{deg}(r) + 1$  as in the proof of Theorem 3.6, then

$$p^{\lfloor \log_p(m e_\infty) \rfloor} u \in \mathbf{Z}_q[x, x^{-1}]^{\oplus d_x}.$$

*Proof.* We have

$$\omega = df + \left( \sum_{i=0}^{d_x-1} u_i b_i^\infty \right) \frac{dx}{r}$$

with  $f = \sum_{j=-\text{ord}_0(W^\infty)}^m (\sum_{i=0}^{d_x-1} (v_j)_i b_i^\infty) x^j$ , where  $v_j \in \mathbf{Q}_q^{\oplus d_x}$  for all  $-\text{ord}_0(W^\infty) \leq j \leq m$ . Note that it is sufficient to show that  $p^{\lfloor \log_p(\ell e) \rfloor} f \in \mathcal{R}$ . By Assumption 1, we have that

$$(4) \quad \mathcal{O}(\mathcal{X} - x^{-1}(0))/(t)^k \cong \prod_{P \in \mathcal{X} \setminus \mathcal{U}, x(P)=\infty} \mathcal{O}_{\mathcal{X},P}/(z_P^{e_P})^k,$$

for all  $k \in \mathbf{N}$ . Moreover, by definition  $[b_0^\infty, \dots, b_{d_x-1}^\infty]$  is a basis for  $\mathcal{O}(\mathbb{X} - x^{-1}(0))$  over  $\mathbf{Q}_q[x^{-1}]$ . To show that  $p^{\lfloor \log_p(\ell e_\infty) \rfloor} f$  is integral, it is therefore enough to show that for every  $P \in \mathcal{X} \setminus \mathcal{U}$  with  $x(P) = 0$ , the Laurent series expansion

$$a_{-m e_P} z_P^{-m e_P} + \dots + a_{(\text{ord}_0(W^\infty)+1)e_P-1} z_P^{(\text{ord}_0(W^\infty)+1)e_P-1} + \mathcal{O}(z_P^{(\text{ord}_0(W^\infty)+1)e_P})$$

of  $p^{\lfloor \log_p(\ell e_\infty) \rfloor} f$  is integral. However, the differential  $df$  has a pole of order at most  $m e_P + 1$  at  $P$ , and its Laurent series expansion

$$\left( b_{-m e_P-1} z_P^{-m e_P-1} + \dots + b_{(\text{ord}_0(W^\infty)+1)e_P} z_P^{(\text{ord}_0(W^\infty)+1)e_P} \right. \\ \left. + \mathcal{O}\left( z_P^{(\text{ord}_0(W^\infty)+1)e_P-1} \right) \right) dz_P$$

is integral since  $\omega$  is integral. The worst denominator we get by integrating this series is therefore  $p^{\lfloor \log_p(m e_\infty) \rfloor}$  and the result follows.  $\square$

*Remark 3.9.* Note that Propositions 3.3, 3.4, 3.7 and 3.8 can be used to give an alternative effective proof of Theorem 3.2.

Recall that in Theorem 3.6 the computation of a basis for  $H_{\text{rig}}^1(U)$  was reduced to a (small) finite dimensional linear algebra problem. However, the dimension of  $H_{\text{rig}}^1(U)$  is generally about  $d_x$  times the dimension of  $H_{\text{rig}}^1(X)$ , so that we would like to compute a basis for this last space. For this we will need to compute the kernel of a cohomological residue map.

**Definition 3.10.** For a 1-form  $\omega \in \Omega^1(\mathcal{U})$  and a point  $P \in \mathcal{X} \setminus \mathcal{U}$ , we let

$$\text{res}_P(\omega) \in \mathcal{O}_{\mathcal{X},P}/(z_P)$$

denote the coefficient  $a_{-1}$  in the Laurent series expansion

$$\omega = (a_{-k}z_P^k + \dots + a_{-1}z_P^{-1} + \dots)dz_P.$$

Moreover, we denote

$$\text{res} = \bigoplus_{P \in \mathcal{X} \setminus \mathcal{U}: x(P) \neq \infty} \text{res}_P, \quad \text{res}_\infty = \bigoplus_{P \in \mathcal{X} \setminus \mathcal{U}: x(P) = \infty} \text{res}_P.$$

**Theorem 3.11.** We have an exact sequence

$$0 \longrightarrow H_{\text{rig}}^1(X) \longrightarrow H_{\text{rig}}^1(U) \xrightarrow{(\text{res} \oplus \text{res}_\infty) \otimes \mathbf{Q}_q} \bigoplus_{P \in \mathcal{X} \setminus \mathcal{U}} \mathcal{O}_{\mathcal{X},P}/(z_P) \otimes \mathbf{Q}_q.$$

*Proof.* This is well known. □

The kernels of  $\text{res}$  and  $\text{res}_\infty$  can be computed without having to compute the Laurent series expansions at all  $P \in \mathcal{X} \setminus \mathcal{U}$  using the following two propositions. We start with the residues at the points not lying over  $x = \infty$ .

**Proposition 3.12.** Let  $\omega \in \Omega^1(\mathbb{U})$  be a 1-form of the form

$$\omega = \left( \sum_{i=0}^{d_x-1} u_i(x)y^i \right) \frac{dx}{r},$$

with  $u \in \mathbf{Q}_q[x]^{\oplus d_x}$ . Then

$$\text{res}(\omega) = 0 \iff \frac{\partial Q}{\partial y} \sum_{i=0}^{d_x-1} u_i y^i = 0 \text{ in } \mathcal{O}(\mathbb{X} - x^{-1}(\infty))/(r).$$

*Proof.* Let  $P$  run over all points in  $\mathcal{X} \setminus \mathcal{U}$  such that  $x(P) \neq \infty$ . One checks that  $\text{ord}_P(\frac{dx}{r}) = -1$  and  $\text{ord}_P(\omega) \geq -1$ . Hence  $\text{res}_P(\omega) = 0$  if and only if  $\text{ord}_P(\sum_{i=0}^{d_x-1} u_i y^i) \geq 1$ . However, since  $\text{ord}_P(\frac{\partial Q}{\partial y}) = e_P - 1$  by Assumption 2, this is the case if and only if  $\text{ord}_P(\frac{\partial Q}{\partial y} \sum_{i=0}^{d_x-1} u_i y^i) \geq e_P$ . Finally, we have that  $\text{ord}_P(\frac{\partial Q}{\partial y} \sum_{i=0}^{d_x-1} u_i y^i) \geq e_P$  at all  $P$  in  $\mathcal{X} \setminus \mathcal{U}$  such that  $x(P) \neq \infty$  if and only if  $\frac{\partial Q}{\partial y} \sum_{i=0}^{d_x-1} u_i y^i$  maps to 0 in  $\mathcal{O}(\mathbb{X} - x^{-1}(\infty))/(r)$ . □

We now move on to the residues at the points lying over  $x = \infty$ .

**Proposition 3.13.** Let  $\omega \in \Omega^1(\mathbb{U})$  be a 1-form of the form

$$\omega = \left( \sum_{i=0}^{d_x-1} u_i(x, x^{-1})b_i^\infty \right) \frac{dx}{r},$$

where  $u \in \mathbf{Q}_q[x, x^{-1}]^{\oplus d_x}$  satisfies  $\text{ord}_\infty(u) > -\text{deg}(r)$ , and let  $v \in \mathbf{Q}_q^{\oplus d_x}$  be defined by  $v = (x^{1-\text{deg}(r)}u)|_{x=\infty}$ . Moreover, let the residue matrix  $G_{-1}^\infty \in M_{d_x \times d_x}(\mathbf{Q}_q)$

be defined as in the proof of Proposition 3.4, and let  $V_\lambda$  denote the generalised eigenspace of  $G_{-1}^\infty$  with eigenvalue  $\lambda$ , so that  $\mathbf{Q}_q^{\oplus d_x}$  decomposes as  $\bigoplus V_\lambda$ . Then

$$\text{res}_\infty(\omega) = 0 \iff \text{the projection of } v \text{ onto } V_0 = 0.$$

*Proof.* Let  $P$  run over all points in  $\mathcal{X} \setminus \mathcal{U}$  such that  $x(P) = \infty$ . One checks that  $\text{ord}_P(\frac{dx}{r}) = -1 + (\deg(r) - 1)e_P$  and  $\text{ord}_P(\omega) \geq -1$ . Since  $\text{ord}_P(x) = -e_P$ , we have that  $\text{res}_P(\omega) = 0$  if and only if  $\text{ord}_P(\sum_{i=0}^{d_x-1} v_i b_i^\infty) \geq 1$ . We still denote  $t = 1/x$ . Note that  $[b_0^\infty, \dots, b_{d_x-1}^\infty]$  is a  $\mathbf{Q}_q$ -basis for  $\mathcal{O}(\mathbb{X} - x^{-1}(0))/(t)$  and that

$$(5) \quad \mathcal{O}(\mathbb{X} - x^{-1}(0))/(t) \cong \prod_{P \in \mathcal{X} \setminus \mathcal{U}, x(P)=\infty} \mathcal{O}_{\mathbb{X},P}/(z_P^{e_P}).$$

Under this isomorphism every factor on the right-hand side is an invariant subspace for  $G_{-1}^\infty$  since  $\text{ord}_P(f) \geq e_P$  implies that  $\text{ord}_P(tdf/dt) \geq e_P$ .

We know from Proposition 2.10 that the eigenvalues of  $G_{-1}^\infty$  are elements of  $\mathbf{Q} \cap \mathbf{Z}_p$  contained in the interval  $[0, 1)$  and that if  $f \in \mathcal{O}(\mathbb{X} - x^{-1}(0))/(t)$  is an eigenvector with eigenvalue  $\lambda$  and  $\text{ord}_P(f) < e_P$  for some  $P$ , then we have that  $\text{ord}_P(f) = \lambda e_P$ . We claim that the eigenvalues of  $G_{-1}^\infty$  on the factor corresponding to the point  $P$  in (5) are  $[0, 1/e_P, \dots, (e_P - 1)/e_P]$ . In particular they are all different, so that  $G_{-1}^\infty$  is diagonalisable. This follows since locally around the point  $P$  the map  $t$  is the  $e_P$ -th power map, so the eigenvalues of its monodromy are all the  $e_P$ -th roots of unity, but these eigenvalues of monodromy are of the form  $e^{2\pi i \lambda}$  where  $\lambda$  runs over the eigenvalues of  $G_{-1}^\infty$  on the factor corresponding to the point  $P$  in (5).

Now, if we decompose  $v$  onto a basis of eigenvectors compatible with the decomposition (5), then we see that  $\text{ord}_P(\sum_{i=0}^{d_x-1} v_i b_i^\infty) \geq 1$  for all  $P$  in  $\mathcal{X} \setminus \mathcal{U}$  such that  $x(P) = \infty$  if and only if the components along the eigenvectors with eigenvalue 0 all vanish. □

*Remark 3.14.* For any  $\omega \in \Omega^1(\mathbb{U})$  we can first apply Propositions 3.3 and 3.4 to represent the class of  $\omega$  in  $H_{\text{rig}}^1(\mathbb{U})$  by 1-forms to which we can apply Propositions 3.12 and 3.13.

#### 4. THE COMPLETE ALGORITHM AND ITS COMPLEXITY

In this section we describe all the steps in the algorithm and determine bounds for the complexity. Recall that  $X$  is a curve of genus  $g$  over a finite field  $\mathbf{F}_q$  with  $q = p^n$  and that  $d_x$  and  $d_y$  denote the degrees of the defining polynomial  $Q$  in the variables  $y$  and  $x$ , respectively. All computations are carried out to  $p$ -adic precision  $N$  which will be specified later. We use the  $\tilde{O}(-)$  notation that ignores logarithmic factors, i.e.,  $\tilde{O}(f)$  denotes the class of functions that lie in  $O(f \log^k(f))$  for some  $k \in \mathbf{N}$ . For example, two elements of  $\mathbf{Z}_q$  can be multiplied in time  $\tilde{O}(\log(p)nN)$ . We let  $\theta$  denote an exponent for matrix multiplication, so that two  $k \times k$  matrices can be multiplied in  $O(k^\theta)$  ring operations. It is known that  $\theta \geq 2$  and that one can take  $\theta \leq 2.3729$  [Wil12]. We start with some bounds that will be useful later on.

**Proposition 4.1.** *Let  $\Delta, s, r$  be defined as in Section 2 and  $e, e_\infty$  as in Section 3. We have:*

$$\begin{aligned} (6a) \quad & \deg(\Delta), \deg(r), \deg(s) \leq 2(d_x - 1)d_y \quad \in O(d_x d_y), \\ (6b) \quad & e, e_\infty \leq d_x \quad \in O(d_x), \\ (6c) \quad & g \leq (d_x - 1)(d_y - 1) \in O(d_x d_y). \end{aligned}$$

*Proof.* (6a) Note that the matrix  $\Sigma$  from Proposition 2.4 is a  $(2d_x - 1) \times (2d_x - 1)$  matrix over  $\mathbf{Z}_q[x]$  of degree at most  $d_y$  and that the row corresponding to  $y^{2d_x-2}$  has degree 0. Since  $\Delta = \det(\Sigma)$ , this implies that  $\deg(\Delta) \leq (2d_x - 2)d_y$ . Writing  $s = \sum_{i=0}^{d_x-1} s_i(x)y^i$  with  $s_i \in \mathbf{Z}_q[x]$ , the  $s_i$  are in fact entries of  $r\Sigma^{-1}$ , so that  $\deg(s_i) \leq (2d_x - 2)d_y$  for all  $0 \leq i \leq d_x - 1$ .

(6b) All the ramification indices  $e_P$  are at most  $d_x$ .

(6c) It is known [BP00] that  $g$  is at most the number of interior points of the Newton polygon of  $Q$ , which is clearly bounded by  $(d_x - 1)(d_y - 1)$ .  $\square$

**Proposition 4.2.** *We have*

$$\begin{aligned} (7a) \quad & \text{ord}_\infty(W^\infty) \geq -(d_x - 1)d_x d_y \in -O(d_x^2 d_y), \\ (7b) \quad & \text{ord}_\infty((W^\infty)^{-1}) \geq -(d_x - 1)d_y \in -O(d_x d_y). \end{aligned}$$

*Moreover, we may assume that*

$$(7c) \quad \text{ord}_0(W^\infty) \geq -(d_x - 1)d_y \in -O(d_x d_y).$$

*Proof.* We still denote  $t = 1/x$ . One easily checks that the minimal polynomial  $\mathcal{Q}^\infty$  of  $y' = y/x^{d_y}$  over  $\mathbf{Q}_q[t]$  is monic. Hence the functions  $1, y', \dots, y'^{d_x-1}$  are  $\mathbf{Q}_q[t]$ -linear combinations of  $b_0^\infty, \dots, b_{d_x-1}^\infty$ , so that  $\text{ord}_\infty((W^\infty)^{-1}) \geq -(d_x - 1)d_y$ .

Since the degree of  $\mathcal{Q}^\infty$  in the variable  $t$  is at most  $d_x d_y$ , its discriminant  $\Delta^\infty \in \mathbf{Z}_q[t]$  with respect to the variable  $y'$  has degree  $\leq 2(d_x - 1)d_x d_y$  by the argument from Proposition 4.1. Defining the matrix  $W^{\infty'} \in \text{Gl}_{d_x}(\mathbf{Z}_q[x, x^{-1}])$  such that

$$b_j^\infty = \sum_{i=0}^{d_x-1} W_{i+1, j+1}^{\infty'} y'^i$$

for all  $0 \leq j \leq d_x - 1$ , it follows from basic properties of the discriminant that  $\text{ord}_\infty(W^{\infty'}) \geq -\deg(\Delta^\infty)/2$ . Clearly,  $\text{ord}_\infty(W^\infty) \geq \text{ord}_\infty(W^{\infty'})$ , so this implies that  $\text{ord}_\infty(W^\infty) \geq -(d_x - 1)d_x d_y$ .

We may assume that  $\text{ord}_0(W^{\infty'}) \geq 0$ . When this is not the case, we can proceed as in [vH94] to obtain another integral basis such that  $\text{ord}_0(W^{\infty'}) \geq 0$ . Note that this does not involve computing Puiseux expansions etc. as in [vH94], since we already have the integral basis  $[b_0^\infty, \dots, b_{d_x-1}^\infty]$  at our disposal. Finally, clearly  $\text{ord}_0(W^{\infty'}) \geq 0$  implies that  $\text{ord}_0(W^\infty) \geq -(d_x - 1)d_y$ .  $\square$

In general algorithms like the one from [vH94] are available for computing integral bases in function fields. In the following important special case we can write down  $[b_0^\infty, \dots, b_{d_x-1}^\infty]$  directly.

**Proposition 4.3.** *For positive integers  $a, b \in \mathbb{N}$ , let  $\Gamma$  denote the triangle in the plane with vertices  $(0, 0), (a, 0)$  and  $(0, b)$ . If  $Q$  is nondegenerate with respect to  $\Gamma$ , then we can take  $[b_0^\infty, \dots, b_{d_x-1}^\infty]$  to be*

$$\left[ 1, x^{\lfloor -a/b \rfloor} y, x^{\lfloor -2a/b \rfloor} y^2, \dots, x^{\lfloor -(b-1)(a/b) \rfloor} y^{b-1} \right].$$

*Proof.* Let  $\Gamma'$  be the translation of  $\Gamma$  defined by  $\Gamma' = \Gamma - (a, 0)$ . If  $Q$  is nondegenerate with respect to  $\Gamma$ , then so is  $Q$  by [CDV06, Corollary 6]. The toric surface  $Y_\Gamma$  associated to  $\Gamma$  contains 3 divisors at infinity, corresponding to the edges of  $\Gamma$ . Now, a Laurent polynomial is regular on  $\mathbb{X} - x^{-1}(0)$  if and only if it is regular at the points lying on the intersection of  $\mathbb{X}$  with the divisors at infinity of  $Y_\Gamma$  corresponding to the horizontal and the diagonal edges of  $\Gamma$ . Therefore, it follows from (the proof of) [CDV06, Lemma 2] that

$$\mathcal{O}(\mathbb{X} - x^{-1}(0)) \cong \mathbf{Q}_q[\Gamma'] / (x^{-a}Q),$$

where  $\mathbf{Q}_q[\Gamma']$  is the algebra over  $\mathbf{Q}_q$  generated by the monomials supported on the cone generated by  $\Gamma'$ . Note that we can subtract a multiple of  $x^{-a}Q$  from an arbitrary element of  $\mathbf{Q}_q[\Gamma']$  to eliminate all powers of  $y$  greater than  $b$ . Therefore,  $\mathcal{O}(\mathbb{X} - x^{-1}(0))$  is generated as a  $\mathbf{Q}_q[x^{-1}]$ -module by the set

$$\left\{ 1, x^{\lfloor -a/b \rfloor} y, x^{\lfloor -2a/b \rfloor} y^2, \dots, x^{\lfloor -(b-1)(a/b) \rfloor} y^{b-1} \right\}.$$

Since the rank of  $\mathcal{O}(\mathbb{X} - x^{-1}(0))$  over  $\mathbf{Q}_q[x^{-1}]$  is  $b$ , this finishes the proof. □

**Assumption 4.** *In the complexity analysis we will assume a couple of times (with explicit mention) that  $\text{ord}_\infty(W^\infty) \in -O(d_x d_y)$ .*

**4.1. Step I: Determine a basis for the cohomology.**

We want to find  $\omega_1, \dots, \omega_\kappa \in (E_0 \cap E_\infty) \cap \Omega^1(\mathcal{U})$  such that:

- (1)  $[\omega_1, \dots, \omega_\kappa]$  is a basis for  $H_{\text{rig}}^1(U) \cong (E_0 \cap E_\infty) / d(B_0 \cap B_\infty)$ ,
- (2) the class of every element of  $(E_0 \cap E_\infty) \cap \Omega^1(\mathcal{U})$  in  $H_{\text{rig}}^1(U)$  has  $p$ -adically integral coordinates with respect to  $[\omega_1, \dots, \omega_\kappa]$ ,
- (3)  $[\omega_1, \dots, \omega_{2g}]$  is a basis for the kernel of  $\text{res} \oplus \text{res}_\infty$  and hence for the subspace  $H_{\text{rig}}^1(X)$  of  $H_{\text{rig}}^1(U)$ .

This can be done using standard linear algebra over  $\mathbf{Z}_q$ , i.e. by computing the Smith normal forms (including unimodular transformations) of two matrices. Note for an element

$$\left( \sum_{i=0}^{d_x-1} u_i(x) y^i \right) \frac{dx}{r} \in E_0 \cap E_\infty,$$

we have that  $\deg(u) \leq \deg(r) - 2 - \text{ord}_0(W^\infty) - \text{ord}_\infty(W^\infty)$ . Hence the dimensions of the matrices involved are at most

$$d_x (\deg(r) - 1 - \text{ord}_0(W^\infty) - \text{ord}_\infty(W^\infty)).$$

Therefore, (under Assumption 4) we need  $O((d_x^2 d_y)^\theta)$  ring operations in  $\mathbf{Z}_q$  by [Sto00, Chapter 7], each of which can be carried out in time  $\tilde{O}(\log(p)nN)$ , so that the time complexity of this step is

$$\tilde{O}(\log(p) d_x^{2\theta} d_y^\theta nN).$$

**4.2. Step II: Compute the map  $F_p$ .**

We use Theorem 2.6 to compute approximations:

$$\begin{aligned} F_p(1/r) &= \alpha_i + \mathcal{O}(p^{2^i}), \\ F_p(y) &= \beta_i + \mathcal{O}(p^{2^i}), \end{aligned}$$

for  $i = 1, \dots, \nu = \lceil \log_2(N) \rceil$ . We carry out all computations using  $r$ -adic expansions for the elements of  $\mathcal{R}$  and  $\mathcal{S}$ , e.g., we represent  $\alpha_i, \beta_i$  as:

$$\alpha_i = \sum_{j \in J} \frac{\alpha_{i,j}(x)}{r^j}, \quad \beta_i = \sum_{k=0}^{d_x-1} \left( \sum_{j \in J} \frac{\beta_{i,j,k}(x)}{r^j} \right) y^k,$$

where  $J \subset \mathbf{Z}$  is finite and  $\alpha_{i,j}, \beta_{i,j,k} \in \mathbf{Z}_q[x]$  satisfy  $\deg(\alpha_{i,j}), \deg(\beta_{i,j,k}) < \deg(r)$ , for all  $i, j, k$ . By Propositions 2.12 and 4.2, we have that

$$|\min J|, |\max J| \in O\left(p\left(N + d_x^2 d_y / \deg(r)\right)\right).$$

Hence, a single ring operation in  $\mathcal{R}$  takes time

$$\tilde{O}(\log(p)|\max J - \min J|) \deg(r) d_x n N \subset \tilde{O}\left(p d_x^2 d_y (N + d_x) n N\right).$$

Moreover, the image of an element of  $\mathbf{Q}_q$  under the map  $\sigma$  can be computed in time  $\tilde{O}(\log^2(p)n + \log(p)nN)$  by [Hub10]. We need  $O(d_x \log(N))$  ring operations in  $\mathcal{R}$  and  $O(d_x d_y)$  applications of  $\sigma$  in order to compute  $(\alpha_\nu, \beta_\nu)$ . Therefore, this can be done in time

$$\tilde{O}\left(p d_x^3 d_y (N + d_x) n N\right).$$

Now for each  $\omega_i = \left(\sum_{k=0}^{d_x-1} u_k(x) y^k\right) \frac{dx}{r}$  with  $1 \leq i \leq 2g$ , we compute

$$(8) \quad F_p(\omega_i) = \sum_{k=0}^{d_x-1} p x^{p-1} u_k^\sigma(x^p) F_p\left(\frac{y^k}{r}\right) dx = \sum_{k=0}^{d_x-1} p x^{p-1} u_k^\sigma(x^p) \alpha_\nu \beta_\nu^k dx + O(p^N).$$

For a single  $\omega_i$  this takes  $O(d_x)$  ring operations in  $\mathcal{R}$  and

$$O(d_x(\deg(r) - 1 - \text{ord}_0(W^\infty) - \text{ord}_\infty(W^\infty))) \subset O(d_x^3 d_y)$$

applications of  $\sigma$ . Hence the complete set of  $F_p(\omega_i)$  can be computed in time

$$\tilde{O}\left(g p d_x^3 d_y (N + d_x) n N\right) \subset \tilde{O}\left(p d_x^4 d_y^2 (N + d_x) n N\right),$$

which is also the total time complexity of this step.

**4.3. Step III: Reduce back to the basis.**

We want to find the matrix  $\Phi \in M_{2g \times 2g}(\mathbf{Q}_q)$  such that

$$F_p(\omega_i) = \sum_{j=1}^{2g} \Phi_{j,i} \omega_j$$

in  $H_{\text{rig}}^1(U)$ . In the previous step, we have obtained an approximation

$$(9) \quad F_p(\omega_i) = \sum_{j \in J} \left( \sum_{k=0}^{d_x-1} \frac{w_{i,j,k}(x)}{r^j} y^k \right) \frac{dx}{r} + O(p^N),$$

where  $J \subset \mathbf{Z}$  is finite and  $w_{i,j,k}(x) \in \mathbf{Z}_q[x]$  satisfies  $\deg(w_{i,j,k}(x)) < \deg(r)$  for all  $i, j, k$ . We now use Proposition 3.3 and Proposition 3.4 (repeatedly) to reduce this 1-form to an element of  $E_0 \cap E_\infty$  as in Theorem 3.6.

To carry out the reduction procedure, it is sufficient to solve a linear system with parameter ( $\ell$  or  $m$ , respectively) only once in Propositions 3.3 and 3.4. After that, every reduction step corresponds to a multiplication of a vector by a  $d_x \times d_x$  matrix

(over  $\mathbf{Q}_q[x]/(r)$  or  $\mathbf{Q}_q$ , respectively). First, the linear systems with parameter can be solved in time

$$\tilde{O}(\log(p)d_x^{\theta+1} \deg(r)nN) \subset \tilde{O}(\log(p)d_x^{\theta+2}d_y nN),$$

where one factor  $d_x$  is from the degree in the parameter. Then, the number of reduction steps at the points not lying over  $x = \infty$  is  $O(pN)$  for each  $F_p(\omega_i)$ . Every single finite reduction step takes time  $\tilde{O}(\log(p)d_x^2 \deg(r)nN)$ , so all  $F_p(\omega_i)$  can be reduced in time

$$\tilde{O}(g(pN)d_x^2 \log(p) \deg(r)nN) \subset \tilde{O}(pd_x^4 d_y^2 nN^2).$$

Finally, the number of reduction steps at the points lying over  $x = \infty$  is  $O(pd_x^2 d_y)$  for each  $F_p(\omega_i)$ . Every single infinite reduction step takes time  $\tilde{O}(\log(p)d_x^2 nN)$ , so all  $F_p(\omega_i)$  can be reduced in time

$$\tilde{O}(g(pd_x^2 d_y) \log(p)d_x^2 nN) \subset \tilde{O}(pd_x^5 d_y^2 nN).$$

After this reduction procedure, we project from  $E_0 \cap E_\infty$  onto the basis  $[\omega_1, \dots, \omega_{2g}]$  and read off the entries of  $\Phi$ . This involves computing  $O(g)$  products of a vector by a matrix of size  $O(d_x^2 d_y)$  (under Assumption 4). Therefore, it can be done in time

$$\tilde{O}(\log(p)g(d_x^2 d_y)^2 nN) \subset \tilde{O}(\log(p)d_x^5 d_y^3 nN).$$

Combining all of this, the total time complexity of this step is

$$\tilde{O}(pd_x^4 d_y^2 nN^2 + d_x^5 d_y^3 nN).$$

**4.4. Step IV: Determine  $Z(X, T)$ .**

It follows from the Lefschetz formula for rigid cohomology that

$$Z(X, T) = \frac{\chi(T)}{(1 - T)(1 - qT)},$$

where

$$\chi(T) = \det(1 - F_p^n T | H_{\text{rig}}^1(X)).$$

Since  $F_p$  is not linear but  $\sigma$ -semilinear, the matrix of  $F_p^n$  with respect to the basis  $[\omega_1, \dots, \omega_{2g}]$  is given by

$$\Phi^{(n)} = \Phi^{\sigma^{(n-1)}} \Phi^{\sigma^{(n-2)}} \dots \Phi.$$

Note that  $\chi(T)$  is the reverse characteristic polynomial of  $\Phi^{(n)}$ . It is known (see for example [PT13]) that  $\Phi^{(n)}$  can be computed from  $\Phi$  in time  $\tilde{O}(\log^2(p)g^\theta nN)$  and that  $\chi(T)$  can be computed from  $\Phi^{(n)}$  in time  $\tilde{O}(\log(p)g^\theta nN)$ . Therefore, the total time complexity of this step is

$$\tilde{O}(\log^2(p)g^\theta nN) \subset \tilde{O}(\log^2(p)(d_x d_y)^\theta nN).$$

4.5. **The  $p$ -adic precision.**

So far we have only obtained an approximation to  $\chi(T)$ , since we have computed to  $p$ -adic precision  $N$ . Moreover, because of loss of precision in the computation, in general  $\chi(T)$  will not even be correct to precision  $N$ . So what precision  $N$  is sufficient to determine  $\chi(T)$  exactly?

**Proposition 4.4.** *The least  $p$ -adic precision  $N$  that is sufficient to determine  $\chi(T)$  satisfies  $N \in \tilde{O}(d_x d_y n)$ .*

*Proof.* We assume for simplicity as in [Ked01] that  $\text{ord}_p(\Phi) \geq 0$ . After the proof we will say something more about the general case.

It follows from the Weil conjectures that  $\chi(T)$  is determined by the bottom half of its coefficients, all of which are bounded in absolute value by  $\binom{2g}{g} q^{\frac{g}{2}}$ . Therefore, if  $\chi(T)$  is known to  $p$ -adic precision at least  $\lceil \log_p(2 \binom{2g}{g} q^{\frac{g}{2}}) \rceil$ , then it is determined exactly. Since  $\text{ord}_p(\Phi) \geq 0$ , there will be no loss of precision in computing  $\Phi^{(n)}$  and  $\chi(T)$ , so that it is sufficient to compute  $\Phi$  to  $p$ -adic precision  $\lceil \log_p(2 \binom{2g}{g} q^{\frac{g}{2}}) \rceil$ .

From Proposition 2.12 and formula (8), it follows that in equation (9) we have  $\max J \leq p(N - 1) - 1$ . Therefore, the loss of precision during the reductions at the points not lying over  $x = \infty$  is at most  $\lceil \log_p(p(N - 1)e) \rceil$  by Proposition 3.7.

Similarly, the coefficients of  $F_p(y^i/r)$  with respect to the basis  $[b_0^\infty, \dots, b_{d_x-1}^\infty]$  have order at  $x = \infty$  at least  $p(\text{ord}_\infty((W^\infty)^{-1}) + \text{deg}(r))$  by the proof of Proposition 2.12. It follows from formula (8) and the definition of  $E_\infty$  that the coefficients of  $F_p(\omega_i)$  with respect to the basis  $[b_0^\infty, \dots, b_{d_x-1}^\infty]$ , which are elements of  $\Omega^1(\mathbb{V})$  now, have order at  $x = \infty$  at least

$$\begin{aligned} p\left(\text{ord}_\infty((W^\infty)^{-1}) + \text{deg}(r)\right) - (p - 1) + p\left(\text{ord}_0(W^\infty) - \text{deg}(r) + 2\right) - 2 \\ \geq p\left(\text{ord}_\infty((W^\infty)^{-1}) + \text{ord}_0(W^\infty)\right) - 1. \end{aligned}$$

Note that the reductions at the points not lying over  $x = \infty$  can introduce poles at  $x = \infty$ , but these can be ignored since they have order at  $x = \infty$  at least

$$\text{ord}_\infty((W^\infty)^{-1}) \geq p\left(\text{ord}_\infty((W^\infty)^{-1}) + \text{ord}_0(W^\infty)\right) - 1,$$

using that  $\text{ord}_\infty((W^\infty)^{-1})$ ,  $\text{ord}_0(W^\infty)$  are both negative. Hence, when applying Proposition 3.8 to the 1-form that remains after the reductions at the points not lying over  $x = \infty$ , we have that  $m \leq -p(\text{ord}_\infty((W^\infty)^{-1}) + \text{ord}_0(W^\infty))$ . Therefore, the loss of precision during the reductions at the points lying over  $x = \infty$  is at most

$$\lceil \log_p\left(-p(\text{ord}_\infty((W^\infty)^{-1}) + \text{ord}_0(W^\infty))e_\infty\right) \rceil.$$

By construction of our basis  $[\omega_1, \dots, \omega_{2g}]$ , there will be no further loss of precision computing the matrix  $\Phi$ . We conclude that it is sufficient for  $N$  to satisfy

$$\begin{aligned} N - \lceil \log_p(p(N - 1)e) \rceil - \lceil \log_p\left(-p(\text{ord}_\infty((W^\infty)^{-1}) + \text{ord}_0(W^\infty))e_\infty\right) \rceil \\ \geq \lceil \log_p\left(2 \binom{2g}{g} q^{\frac{g}{2}}\right) \rceil. \end{aligned}$$

From this it follows that  $N \in \tilde{O}(d_x d_y n)$  using Propositions 4.1 and 4.2. □

*Remark 4.5.* If we do not assume that  $\text{ord}_p(\Phi) \geq 0$ , then we can use Propositions 2.12, 3.7 and 3.8 to obtain a lower bound for  $\text{ord}_p(\Phi)$ . Taking into account the extra loss of precision  $(n - 1) \text{ord}_p(\Phi)$  for computing  $\Phi^{(n)}$  and  $(2g - 1)n \text{ord}_p(\Phi)$  for computing  $\chi(T)$ , we still have that  $N \in \tilde{O}(d_x d_y n)$ . However, a bound for  $N$  obtained this way will not be very good in practice. One can obtain a much sharper bound for  $\text{ord}_p(\Phi)$  and the loss of precision in computing  $\Phi^{(n)}$  and  $\chi(T)$ , using the existence of the  $F_p$ -invariant  $\mathbf{Z}_q$ -lattice coming from the (log)-crystalline cohomology inside the rigid cohomology.

**Theorem 4.6.** *The time complexity of the algorithm presented in this section is  $\tilde{O}(pd_x^6 d_y^4 n^3)$ .*

*Proof.* We take the sum of the complexities of the different steps using Proposition 4.4, leaving out terms and factors that are absorbed by the  $\tilde{O}$ . □

For the analysis of the space complexity, we will not go into the same detail as for the time complexity. However, using Assumption 4 at the same two points as in the analysis of the time complexity, one can prove the following theorem.

**Theorem 4.7.** *The space complexity of the algorithm presented in this section is  $\tilde{O}(pd_x^4 d_y^3 n^3)$ .*

*Proof.* The space complexity of the algorithm turns out to be that of storing a single  $F_p(\omega_i)$ , or equivalently an element of  $\mathcal{R}$ , which is  $\tilde{O}(pd_x^2 d_y (N + d_x)nN)$ . The result now follows using Proposition 4.4. □

*Remark 4.8.* There are some standard ways to improve the algorithm from this section in practice:

- (1) We computed the Frobenius lift by working with  $p$ -adic precision  $N_i = 2^i$  in the  $i$ th step of the Hensel lift. Setting  $N_\nu = N$  and  $N_{i-1} = \lceil N_i/2 \rceil$  for all  $1 \leq i \leq \nu$ , we still obtain the correct Frobenius lift to precision  $N$ , while having to compute to lower precision in every step.
- (2) The bound  $\log_p(2 \binom{2g}{g} q^{\frac{g}{2}})$  for the  $p$ -adic precision of  $\chi(T)$  can be lowered using the Newton-Girard identities [Ked08].

These improvements do not affect the complexity of the algorithm, but are important in practice.

**4.6. Our assumptions.**

4.6.1. *Assumption 1.* Without this assumption, Theorem 3.2 does not hold and we cannot compute in  $H_{\text{rig}}^1(U)$  as in Section 3. Therefore, Assumption 1 is essential and cannot be lifted. It would be interesting to know under what conditions a lift satisfying this assumption can be found. Note that for a smooth curve and a map  $x$  to the projective line defined over a number field  $K$ , Assumption 1 is satisfied at all but finitely many prime ideals of  $\mathcal{O}_K$ .

4.6.2. *Assumption 2.* This assumption serves to simplify the exposition and can be weakened as follows. Note that Assumption 2 is equivalent to asking that  $[y^0, \dots, y^{d_x-1}]$  is an integral basis for  $\mathbf{Q}_q(x, y)$  over  $\mathbf{Q}_q[x]$ . Let us now assume instead that a matrix  $W^0 \in Gl_{d_x}(\mathbf{Z}_q[x, 1/r])$  is known such that if we denote  $b_j^0 = \sum_{i=0}^{d_x-1} W_{i+1, j+1}^0 y^i$  for all  $0 \leq j \leq d-1$ , then  $[b_0^0, \dots, b_{d_x-1}^0]$  is an integral basis for  $\mathbf{Q}_q(x, y)$  over  $\mathbf{Q}_q[x]$ . Then our algorithm should continue to work, extending it

to arbitrary curves for which we can find a lift that satisfies Assumption 1. However, since quite a lot of small changes are needed in the different steps of the algorithm and, more importantly, we have not implemented this more general algorithm yet, for now we limit ourselves to the less general case.

**4.6.3. Assumption 3.** Note that Assumptions 2 and 3 are in fact very similar: we need an integral basis for  $\mathbf{Q}_q(x, y)$  over both  $\mathbf{Q}_q[x]$  and  $\mathbf{Q}_q[x^{-1}]$ . In both cases, algorithms like the one from [vH94] are available for computing the integral bases.

**4.6.4. Assumption 4.** This assumption is the least important of all the assumptions. We have used it a couple of times in the complexity analysis, to bound the complexity of doing linear algebra in  $E_0 \cap E_\infty$ . Note that in Proposition 4.3, we have that  $\text{ord}_\infty(W^\infty) = 0$ . For more general Newton polygons this also seems to be the case experimentally. With large random searches we have not been able to find a single example satisfying  $\text{ord}_\infty(W^\infty) \leq -d_x d_y / 2$ . Therefore, we expect that Assumption 4 can be removed.

## 4.7. Implementation.

We have implemented our algorithm in the computer algebra system Magma [BCP97]. In examples where we can compare against either [CDV06] or [Wal10], our algorithm runs at least two orders of magnitude faster. The code can be found at [http://perswww.kuleuven.be/jan\\_tuitman](http://perswww.kuleuven.be/jan_tuitman) and comes in two different packages: `pcc_p` for primefields and `pcc_q` for nonprimefields. We give an example for each package below, mainly to demonstrate how to use the code. More examples and timings can be found in the example files that come with the packages. The computations were carried out with Magma v2.20-3 on a 3.0GHz Intel Core i7-3540M processor.

**Example 1.** A random curve over  $\mathbf{F}_{11}$  with  $d_x = 4$  and  $d_y = 5$  (genus 12).

```
load "pcc_p.m";
Q:=y^4+(6*x^5+10*x^4+8*x^3+5*x^2+7*x+5)*y^3+(4*x^5*x^4+8*x^3+6*x^2+6*x)*y^2+(3*x^5+5*x^4+9*x^3+2*x^2+10*x+4)*y
+6*x^5+3*x^4+7*x^3+10*x^2+4*x+3;
p:=11;
N:=9;
chi:=num_zeta(Q,p,N:verbose:=true);
```

The input consists of the polynomial  $Q \in \mathbb{Z}[x, y]$ , the prime  $p$  and the  $p$ -adic working precision  $N$ . The output is the numerator  $\chi(T)$  of the zeta function  $Z(X, T)$ . In this case it is

```
3138428376721*T^24-285311670611*T^23-233436821409*T^22+80170221494*T^21-20364093695*T^20+3799998345*T^19
+2657341500*T^18-754684986*T^17+182500065*T^16-37234725*T^15-9607037*T^14+6197609*T^13-939504*T^12+563419*T^11
-79397*T^10-27975*T^9+12465*T^8-4686*T^7+1500*T^6+195*T^5-95*T^4+34*T^3-9*T^2-T+1.
```

The computation took 27.9s and less than 32MB of memory.

**Example 2.** A random curve over  $\mathbf{F}_{710}$  with  $d_x = 3$  and  $d_y = 5$  (genus 8).

```
load "pcc_q.m";
Q:=y^3+((a^9+5*a^7+3*a^5+6*a^4+4*a^3+2*a^2+5*a+1)*x^5+(5*a^9+5*a^8+2*a^7+2*a^6+3*a^5+a^4+6*a^3+4*a+4)*x^4
+(2*a^9+6*a^7+6*a^6+2*a^5+6*a^4+5*a^3+6)*x^3+(3*a^9+2*a^8+3*a^7+3*a^6+a^5+4*a^4+5*a^3+4*a^2+3*a+3)*x^2
+(5*a^9+3*a^8+a^7+2*a^6+4*a^5+a^4+3*a^3+5*a^2+2)*x+(4*a^8+2*a^7+4*a^6+a^5+4*a^4+4*a^3+a^2+2*a+4))*y^2
+((2*a^9+3*a^8+3*a^7+6*a^6+6*a^5+6*a^4+4*a^3+5*a^2+6*a)*x^5+(5*a^9+3*a^8+2*a^6+2*a^5+4*a^4+2*a^3+4*a^2+3*a+6)*x^4
+(3*a^9+3*a^8+6*a^7+5*a^6+3*a^5+3*a^4+5*a^3+4*a^2+4*a+1)*x^3+(2*a^9+2*a^8+5*a^7+5*a^6+5*a^5+6*a^4+a^3+a^2+2*a+2)*x^2
+(3*a^8+3*a^6+3*a^5+5*a^3+4*a^2+4*a+2)*x+(4*a^9+2*a^8+5*a^6+2*a^5+5*a^4+6*a^3+2*a+4))*y
+(a^9+a^8+2*a^7+4*a^6+2*a^5+a^4+2*a^3+4*a^2+2*a+2)*x^5+(4*a^9+5*a^7+a^6+a^5+3*a^4+2*a^3+6*a+6)*x^4
+(4*a^9+4*a^8+4*a^7+a^6+a^5+5*a^4+2*a^3+a^2+2*a)*x^3+(5*a^9+5*a^7+6*a^6+3*a^5+6*a^4+4*a^3+3*a^2+6*a)*x^2
+(5*a^8+2*a^7+2*a^6+3*a^2+a)*x+a^9+6*a^8+6*a^7+2*a^6+6*a^5+4*a^4+3*a^3+5*a+2;
p:=7;
n:=10;
N:=45;
chi:=num_zeta(Q,p,n,N:verbose:=true);
```

The input consists of the polynomial  $\mathcal{Q} \in \mathbf{Z}[a][x, y]$ , the prime  $p$ , the extension degree  $n$  and the  $p$ -adic working precision  $N$ . Here  $a$  represents a standard generator for  $\mathbf{Z}_q/\mathbf{Z}_p$ , i.e., it is a root of a Conway polynomial. The output is the numerator  $\chi(T)$  of the zeta function  $Z(X, T)$ . In this case it is

```
40536215597144386832065866109016673800875222251012083746192454448001*T^16
+734594936640916515108002147869799216237456127361200615126315631*T^15
+37833822114992619972303659616442535094177702647200606500823*T^14
+2969545553762454604862263614126054405430871338256835484*T^13+323896800674094517822826810513267326953587001034849*T^12
+22636175881373275379227578482427791310493422448*T^11+146359712260050195498039226426210033108323*T^10
+66506665686156219471818560867075857462*T^9+3128031304748736252054098124793644*T^8+235442453530348846499533702038*T^7
+1834259371881387520432323*T^6+1004296292146625341552*T^5+50872731607858849*T^4+1651155559516*T^3+74472823*T^2+5119*T+1.
```

The computation took 2458s and about 350MB of memory.

## REFERENCES

- [BC94] F. Baldassarri and B. Chiarellotto, *Algebraic versus rigid cohomology with logarithmic coefficients*, Barsotti Symposium in Algebraic Geometry (Abano Terme, 1991), Perspect. Math., vol. 15, Academic Press, San Diego, CA, 1994, pp. 11–50. MR1307391 (96f:14024)
- [BP00] P. Beelen and R. Pellikaan, *The Newton polygon of plane curves with many rational points*, Des. Codes Cryptogr. **21** (2000), no. 1-3, 41–67, DOI 10.1023/A:1008323208670. Special issue dedicated to Dr. Jaap Seidel on the occasion of his 80th birthday (Oisterwijk, 1999). MR1801161 (2003c:14024)
- [BCP97] W. Bosma, J. Cannon, and C. Playoust, *The Magma algebra system. I. The user language*, J. Symbolic Comput. **24** (1997), no. 3-4, 235–265, DOI 10.1006/jSCO.1996.0125. Computational algebra and number theory (London, 1993). MR1484478
- [CDV06] W. Castryck, J. Denef, and F. Vercauteren, *Computing zeta functions of nondegenerate curves*, IMPR Int. Math. Res. Pap. (2006), Art. ID 72017, 57. MR2268492 (2007h:14026)
- [DV06a] J. Denef and F. Vercauteren, *Counting points on  $C_{ab}$  curves using Monsky-Washnitzer cohomology*, Finite Fields Appl. **12** (2006), no. 1, 78–102, DOI 10.1016/j.faa.2005.01.003. MR2190188 (2007c:11075)
- [DV06b] J. Denef and F. Vercauteren, *An extension of Kedlaya’s algorithm to hyperelliptic curves in characteristic 2*, J. Cryptology **19** (2006), no. 1, 1–25, DOI 10.1007/s00145-004-0231-y. MR2210897 (2007d:11069)
- [GG01] P. Gaudry and N. Gürel, *An extension of Kedlaya’s point-counting algorithm to superelliptic curves*, Advances in cryptology—ASIACRYPT 2001 (Gold Coast), Lecture Notes in Comput. Sci., vol. 2248, Springer, Berlin, 2001, pp. 480–494, DOI 10.1007/3-540-45682-1\_28. MR1934859 (2003h:11159)
- [Har07] D. Harvey, *Kedlaya’s algorithm in larger characteristic*, Int. Math. Res. Not. IMRN **22** (2007), Art. ID rnm095, 29, DOI 10.1093/imrn/rnm095. MR2376210 (2009d:11096)
- [Har14] D. Harvey, *Counting points on hyperelliptic curves in average polynomial time*, Ann. of Math. (2) **179** (2014), no. 2, 783–803, DOI 10.4007/annals.2014.179.2.7. MR3152945
- [Hub10] H. Hubrechts, *Fast arithmetic in unramified  $p$ -adic fields*, Finite Fields Appl. **16** (2010), no. 3, 155–162, DOI 10.1016/j.faa.2009.12.004. MR2610706 (2011d:11277)
- [Ked01] K. S. Kedlaya, *Counting points on hyperelliptic curves using Monsky-Washnitzer cohomology*, J. Ramanujan Math. Soc. **16** (2001), no. 4, 323–338. MR1877805 (2002m:14019)
- [Ked08] K. S. Kedlaya, *Search techniques for root-unitary polynomials*, Computational arithmetic geometry, Contemp. Math., vol. 463, Amer. Math. Soc., Providence, RI, 2008, pp. 71–81, DOI 10.1090/conm/463/09047. MR2459990 (2009h:26022)
- [KT12] K. S. Kedlaya and J. Tuitman, *Effective convergence bounds for Frobenius structures on connections*, Rend. Semin. Mat. Univ. Padova **128** (2012), 7–16 (2013), DOI 10.4171/RSMUP/128-2. MR3076829
- [Lau06] A. G. B. Lauder, *A recursive method for computing zeta functions of varieties*, LMS J. Comput. Math. **9** (2006), 222–269, DOI 10.1112/S146115700001261. MR2261044 (2007g:14022)
- [PT13] S. Pancratz and J. Tuitman, *Improvements to the deformation method for counting points on smooth projective hypersurfaces*, preprint (2013), <http://arxiv.org/abs/1307.1250>.
- [Sto00] A. Storjohann, *Algorithms for matrix canonical forms*, PhD thesis, Swiss Federal Institute of Technology – ETH, 2000.

- [vH94] M. van Hoeij, *An algorithm for computing an integral basis in an algebraic function field*, J. Symbolic Comput. **18** (1994), no. 4, 353–363, DOI 10.1006/jsco.1994.1051. MR1324494 (96e:11166)
- [Wal10] G. Walker, *Computing zeta functions of varieties via fibration*, PhD thesis, Oxford, 2010.
- [Wil12] V. V. Williams, *Multiplying matrices faster than Coppersmith-Winograd [extended abstract]*, STOC'12—Proceedings of the 2012 ACM Symposium on Theory of Computing, ACM, New York, 2012, pp. 887–898, DOI 10.1145/2213977.2214056. MR2961552

KU LEUVEN, DEPARTEMENT WISKUNDE, CELESTIJNENLAAN 200B, 3001 LEUVEN, BELGIUM  
*E-mail address*: `jan.tuitman@wis.kuleuven.be`