# COUNTING POINTS ON CURVES: THE GENERAL CASE.

JAN TUITMAN

ABSTRACT. We introduce a new algorithm to compute the zeta function of a curve over a finite field. This method extends previous work of ours to all curves for which a lift to characteristic zero is known that satisfies certain conditions. 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 $\sigma \in \mathrm{Gal}(\mathbf{Q}_q/\mathbf{Q}_p)$ 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 $t \to q/t$. Moreover, by the Lefschetz formula for rigid cohomology, we have that

$$\chi(T) = \det\left(1 - T\,\mathrm{F}_p^n\,|H^1_{\mathrm{rig}}(X)\right),$$

where $\mathrm{F}_p$ denotes the $p$-th power Frobenius map.

In [13], 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 $\mathrm{F}_p$ on $H^1_{\mathrm{rig}}(X)$. His algorithm was then extended to characteristic 2 [6] and also to superelliptic curves [9], $C_{ab}$ curves [5] and nondegenerate curves [4]. 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 does not send $x$ to $x^p$ anymore. 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 low degree map to the projective line, as is the case for most nondegenerate curves, this is not fully exploited.

We proposed a very general and practical extension of Kedlaya's algorithm in [20]. Our approach there can be summarised as follows. We start with a finite separable map $x : X \to \mathbf{P}^1_{\mathbf{F}_q}$ of degree $d_x$ and a rational function $y : X \to \mathbf{P}^1_{\mathbf{F}_q}$ of degree $d_y$ 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 $y$. Note that the degrees of $Q$ in $x, y$ are $d_y, d_x$, respectively. The polynomial $Q$ is the natural input to our algorithm. After removing the ramification locus of $x$, 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 [16] to find the matrix of Frobenius and the zeta function of $X$.

The algorithm from [20] can be applied to generic, or in other words random, equations $Q$. However, there are equations to which it cannot be applied including some very interesting examples. For example, when $Q$ is (the reduction at some prime number $p$ of) one of the defining equations computed for modular curves in [19, 22], it turns out that the algorithm can almost never be applied. The reason is that in [20] we assume that $Q$, or rather its lift $\mathcal{Q}$ to characteristic zero, defines a smooth curve in the affine $(x, y)$-plane, i.e. that all the singularities of the plane curve defined by $\mathcal{Q}$ are at infinity. We expect that this condition is similar to nondegeneracy in the sense that a generic equation satisfies it, but a generic curve cannot be defined by an equation that satisfies it. Moreover, even when an equation satisfying the assumption does exist it is not clear how to find it.

In this paper we eliminate the assumption that $\mathcal{Q}$ does not have any singularities in the affine $(x, y)$-plane. As a consequence, our algorithm can now be applied to any curve for which we know a good lift to characteristic zero in the sense of Assumption 1. In particular, for any smooth curve $\mathbb{X}$ defined over the rational numbers, our algorithm can be applied to the reduction of $X$ at $p$ for almost all prime numbers $p$. The time complexity of the algorithm is $\tilde{O}(pd_x^6 d_y^4 n^3)$ by Theorem 4.12, and the space complexity $\tilde{O}(pd_x^4 d_y^3 n^3)$ by Theorem 4.13, under two rather harmless additional assumptions (numbered 2 and 3 below), as was the case in [20].

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 [4]. However, for hyperelliptic curves, the dependence on $p$ of the time and space complexities of Kedlaya's algorithm has been improved to $\tilde{O}(p^{1/2})$ [10] and average polynomial time [11] 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 have updated our Magma [3] implementation. The new code can be found at our webpage[1]. Some new examples that our algorithm could not handle before are included at the end of the paper.

## 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 $x : X \to \mathbf{P}^1_{\mathbf{F}_q}$ be a finite separable map of degree $d_x$ and $y : X \to \mathbf{P}^1_{\mathbf{F}_q}$ a rational function that generates the function

---

[1] https://perswww.kuleuven.be/jan_tuitman

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 $y$. Below in Assumption 1 we will impose an important condition on this lift.

**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 $[1, y, \ldots, y^{d_x-1}]$.*

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

**Definition 2.2.** *We let $\Delta(x) \in \mathbf{Z}_q[x]$ denote the discriminant of $\mathcal{Q}$ with respect to the variable $y$, define $r(x) \in \mathbf{Z}_q[x]$ as the squarefree polynomial $r = \Delta/(\gcd(\Delta, \frac{d\Delta}{dx}))$ and let $m \in \mathbf{N}$ be the least positive integer such that there exist a polynomial $g(x) \in \mathbf{Z}_q[x]$ that satisfies $r(x)^m = g(x)\Delta(x)$.*

**Remark 2.3.** *Note that $\Delta(x) \neq 0 \pmod{p}$ since $x : X \to \mathbf{P}^1_{\mathbf{F}_q}$ is separable.*

**Proposition 2.4.** *The element*

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

*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:

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

with respect to the bases $[1, y, \ldots, y^{d_x-2}]$, $[1, y, \ldots, y^{d_x-1}]$ and $[1, y, \ldots, y^{2d_x-2}]$. 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 $\mathcal{A}$. $\qquad\square$

**Definition 2.5.** *We denote*

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

*and write $\mathcal{V} = \operatorname{Spec}\mathcal{S}$, $\mathcal{U} = \operatorname{Spec}\mathcal{R}$, 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 $\mathbf{P}^1_{\mathbf{Z}_q}$ such that $\mathcal{V} = \mathbf{P}^1_{\mathbf{Z}_q} \setminus \mathcal{D}_{\mathbf{P}^1}$.*

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

**Remark 2.6.** *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 points on $\mathbf{P}^1_{\mathbf{Q}_q}$ in the support of $\mathcal{D}_{\mathbf{P}^1}$ and all points on $\mathbb{X}$ lying over these points are distinct modulo $p$.*

**Remark 2.7.** *The general problem of when there exists a lift $\mathcal{Q}$ of the polynomial $Q$ to characteristic zero such that Assumption 1 is satisfied, and if so how this lift can be constructed, is not easy and will not be answered here. We only make the following observations:*

(1) *For a generic $Q$ a random lift usually works.*
(2) *Starting from $\mathcal{Q} \in \mathbf{Z}[x, y]$, it is easy to show that for all but a finite number of prime numbers $p$, the polynomial $\mathcal{Q}$ is a good lift of its reduction modulo $p$ $Q$ in the sense of Assumption 1.*
(3) *We will see in Remark 3.14 that if Assumption 1 is satisfied, then the finite map $x : X \to \mathbf{P}_{\mathbf{F}_q}$ is tamely ramified. Therefore, a lift $\mathcal{Q}$ satisfying Assumption 1 does not always exist.*

**Definition 2.8.** *We denote the ring of overconvergent 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 $\mathcal{S}^\dagger = \mathbf{Z}_q\langle x, 1/r\rangle^\dagger$ and that a basis is given by $[y^0, \ldots, y^{d_x-1}]$. A Frobenius lift $\mathrm{F}_p : \mathcal{R}^\dagger \to \mathcal{R}^\dagger$ is defined as a $\sigma$-semilinear ring homomorphism that reduces modulo $p$ to the $p$-th power Frobenius map.*

**Theorem 2.9.** *There exists a Frobenius lift $\mathrm{F}_p : \mathcal{R}^\dagger \to \mathcal{R}^\dagger$ for which $\mathrm{F}_p(x) = x^p$.*

*Proof.* Let notation be as in Definition 2.2 and Proposition 2.4. 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:

$$\alpha_0 = \frac{1}{r^p},$$
$$\beta_0 = y^p,$$
$$\alpha_{i+1} = \alpha_i(2 - \alpha_i r^\sigma(x^p)) \qquad\qquad (\mathrm{mod}\ p^{2^{i+1}}),$$
$$\beta_{i+1} = \beta_i - \mathcal{Q}^\sigma(x^p, \beta_i)s^\sigma(x^p, \beta_i)g^\sigma(x^p)\alpha_i^m \qquad (\mathrm{mod}\ p^{2^{i+1}}).$$

Then one easily checks that the $\sigma$-semilinear ring homomorphism $\mathrm{F}_p : \mathcal{R}^\dagger \to \mathcal{R}^\dagger$ defined by

$$\mathrm{F}_p(x) = x^p, \qquad\quad \mathrm{F}_p(1/r) = \lim_{i\to\infty} \alpha_i, \qquad\quad \mathrm{F}_p(y) = \lim_{i\to\infty} \beta_i,$$

is a Frobenius lift. $\qquad\square$

**Assumption 2.** *We will assume that matrices $W^0 \in Gl_{d_x}(\mathbf{Z}_q[x, 1/r])$ and $W^\infty \in Gl_{d_x}(\mathbf{Z}_q[x, 1/x, 1/r])$ are known such that, if we denote $b_j^0 = \sum_{i=0}^{d_x-1} W^0_{i+1,j+1}y^i$ and $b_j^\infty = \sum_{i=0}^{d_x-1} W^\infty_{i+1,j+1}y^i$ for all $0 \leq j \leq d_x - 1$, then:*

(1) *$[b_0^0, \ldots, b_{d_x-1}^0]$ is an integral basis for $\mathbf{Q}_q(x, y)$ over $\mathbf{Q}_q[x]$,*
(2) *$[b_0^\infty, \ldots, b_{d_x-1}^\infty]$ is an integral basis for $\mathbf{Q}_q(x, y)$ over $\mathbf{Q}_q[1/x]$.*

*Let $W \in Gl_{d_x}(\mathbf{Z}_q[x, 1/x])$ be the matrix defined by $W = (W^0)^{-1}W^\infty$.*

**Remark 2.10.** *There are good algorithms available to compute integral bases in function fields. In our implementation, we use the one that comes with Magma. In practice, the time required to compute the integral bases always seems to be neglible compared to the other steps in our algorithm.*

**Proposition 2.11.** *Let $G^0 \in M_{d_x \times d_x}(\mathbf{Z}_q[x, 1/r])$ and $G^\infty \in M_{d_x \times d_x}(\mathbf{Z}_q[x, 1/x, 1/r])$ denote the matrices such that*

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

*for all $0 \le j \le d_x - 1$. Let $x_0 \neq \infty$ be a geometric point of $\mathbf{P}^1(\bar{\mathbf{Q}}_q)$. Then the matrix $G^0 dx$ has at most a simple pole at $x_0$. Similarly, the matrix $G^\infty dx$ has at most a simple pole at $x = \infty$.*

*Proof.* Note that $\mathrm{ord}_P(dx/(x - x_0)) = -1$ at every $P \in \mathcal{X} \setminus \mathcal{U}$ lying over $x_0$. At every such $P$ and for all $0 \le i \le d_x - 1$ we clearly have $\mathrm{ord}_P(db_i^0) \ge 0$, so that $\mathrm{ord}_P((x - x_0)db_i^0) - \mathrm{ord}_P(dx) \ge 1$. Since $[b_0^0, \ldots, b_{d_x-1}^0]$ is an integral basis for $\mathbf{Q}_q(x, y)$ over $\mathbf{Q}_q[x]$, we conclude that $(x - x_0)G^0$ does not have a pole at $x_0$, so that $G^0 dx$ has at most a simple pole there. To show that $G^\infty dx$ has at most a simple pole at $x = \infty$, replace the geometric point $x_0$ by $\infty$, the basis $[b_0^0 \ldots, b_{d_x-1}^0]$ by $[b_0^\infty \ldots, b_{d_x-1}^\infty]$, and the local parameter $(x - x_0)$ by $t = 1/x$ in the argument. $\square$

**Remark 2.12.** *In particular, we have that $rG^0 \in M_{d_x \times d_x}(\mathbf{Z}_q[x])$.*

**Definition 2.13.** *Let $x_0 \in \mathbf{P}^1(\bar{\mathbf{Q}}_q) \backslash \infty$ be a geometric point. The exponents of $G^0 dx$ at $x_0$ are defined as the eigenvalues of the residue matrix $G_{-1}^{x_0} = (x - x_0)G^0|_{x=x_0}$. Moreover, the exponents of $G^\infty dx$ at $x = \infty$ are defined as its exponents at $t = 0$, after substituting $x = 1/t$.*

**Proposition 2.14.** *The exponents of $G^0 dx$ at any geometric point $x_0 \in \mathbf{P}^1(\bar{\mathbf{Q}}_q) \backslash \infty$ 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 $G^0 dx$ at $x_0 \neq \infty$. Then there exists $f = \sum_{i=0}^{d_x-1} a_i b_i^0$ with $a_0, \ldots, a_{d_x-1} \in \bar{\mathbf{Q}}_q$ such that

$$df = \left( \frac{\lambda f}{x - x_0} + g \right) dx \tag{2}$$

as 1-forms on $\mathbb{U} \otimes \bar{\mathbf{Q}}_q$, where $g \in \mathcal{O}(\mathbb{U} \otimes \bar{\mathbf{Q}}_q)$ satisfies $\mathrm{ord}_P(g) \ge 0$ at all points $P \in x^{-1}(x_0)$. Note that for at least one $P \in x^{-1}(x_0)$ we have $\mathrm{ord}_P(f) < \mathrm{ord}_P(x - x_0)$, 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 (2) and taking residues, we obtain

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

Since $0 \le \mathrm{ord}_P(f) < \mathrm{ord}_P(x - x_0)$, 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 $G_{-1}^{x_0} \in M_{d_x \times d_x}(\mathbf{Z}_q)$. 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 the geometric point $x_0$ by $\infty$, the basis $[b_0^0 \ldots, b_{d_x-1}^0]$ by $[b_0^\infty \ldots, b_{d_x-1}^\infty]$, and the local parameter $(x - x_0)$ by $t = 1/x$ in the argument. $\square$

**Definition 2.15.** *For a geometric point $x_0 \in \mathbf{P}^1(\bar{\mathbf{Q}}_q)$, we let $\mathrm{ord}_{x_0}(\cdot)$ denote the discrete valuation on $\bar{\mathbf{Q}}_q(x)$ corresponding to $x_0$. Moreover, we define*

$$\mathrm{ord}_{\neq\infty}(\cdot) = \min_{x_0 \in \mathbf{P}^1(\bar{\mathbf{Q}}_q) \backslash \infty} \{\mathrm{ord}_{x_0}(\cdot)\}.$$

*We extend these definitions to matrices over $\bar{\mathbf{Q}}_q(x)$ by taking the minimum over their entries.*

**Proposition 2.16.** *Let $N \in \mathbf{N}$ be a positive integer.*

(1) *The element $\mathrm{F}_p(1/r)$ 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 $p \le i \le pN$.*

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

$$\phi_{i,j} = \sum_{k=0}^{p(N-1)-\mathrm{ord}_{\neq\infty}(W^0)-p\,\mathrm{ord}_{\neq\infty}((W^0)^{-1})} \frac{\phi_{i,j,k}(x)}{r^k}$$

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

$$\deg(\phi_{i,j,0}) \le -\mathrm{ord}_\infty(W^\infty) - p\,\mathrm{ord}_\infty((W^\infty)^{-1}),$$
$$\deg(\phi_{i,j,k}) < \deg(r),$$

*for all $0 \le j \le d_x-1$ and $1 \le k \le p(N-1)-\mathrm{ord}_{\neq\infty}(W^0)-p\,\mathrm{ord}_{\neq\infty}((W^0)^{-1})$.*

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

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

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

$$\deg(\psi_{i,j,0}) \le -\mathrm{ord}_\infty(W) - p\,\mathrm{ord}_\infty(W^{-1}) - (p-1)\deg(r),$$
$$\deg(\psi_{i,j,k}) < \deg(r),$$

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

*Proof.*

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

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

(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 $\mathbf{R}^0 x_*(\mathcal{O}_\mathbb{U})$ with respect to the basis $[y^0, \ldots, y^{d_x-1}]$. From Proposition 2.11 we know that the matrix of the Gauss–Manin connection $\nabla$ on $\mathbf{R}^0 x_*(\mathcal{O}_\mathbb{U})$ has at most a simple pole at any $x_0 \ne \infty \in \mathbf{P}^1(\bar{\mathbf{Q}}_q)$ with respect to the basis $[b_0^0, \ldots, b_{d_x-1}^0]$ and at most a simple pole at $x = \infty$ with respect to the basis $[b_0^\infty, \ldots, b_{d_x-1}^\infty]$. Moreover, by Proposition 2.14, at every $x_0 \in \mathbf{P}^1(\bar{\mathbf{Q}}_q)$ the exponents of $\nabla$ with respect to these bases are elements of $\mathbf{Q} \cap \mathbf{Z}_p$ that are contained in the interval $[0, 1)$. Finally, from the definition of $\Phi$ it follows that $\mathrm{ord}_p(\Phi) \ge 0$ and by Poincaré duality we find that $\mathrm{ord}_p(\Phi^{-1}) \ge 0$ as well. The result is now a consequence of [15, Corollary 2.6].

(3) The matrix $\Psi = (\psi_{i,j}) \in M_{d_x \times d_x}(\mathcal{S}^\dagger)$ defines a $p$-th power Frobenius structure on the (higher) direct image $\mathbf{R}^0 x_*(\mathcal{O}_\mathbb{U})$ with respect to the basis $[b_0^0/r, \ldots, b_{d_x-1}^0/r]$. With respect to this basis, the matrix of the Gauss–Manin connection $\nabla$ on $\mathbf{R}^0 x_*(\mathcal{O}_\mathbb{U})$ has at most a simple pole at every $x_0 \neq \infty \in \mathbf{P}^1(\bar{\mathbf{Q}}_q)$ and the exponents of $\nabla$ are elements of $\mathbf{Q} \cap \mathbf{Z}_p$ that are contained in the interval $[-1, 0)$. We still have that $\mathrm{ord}_p(\Psi), \mathrm{ord}_p(\Psi^{-1}) \geq 0$. The result is now again a consequence of [15, Corollary 2.6]. $\qquad\square$

## 3. Computing (in) the cohomology

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

$$H^1_{rig}(U) = \mathrm{coker}(d : \mathcal{R}^\dagger \to \Omega^1(\mathbb{U}) \otimes \mathcal{R}^\dagger).$$

**Theorem 3.2.**

$$H^1_{rig}(U) \cong H^1_{dR}(\mathbb{U})$$

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

We can effectively reduce any 1-form to one of low pole order using linear algebra following work of Lauder [16]. 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 $\frac{dr}{dx}$. 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*

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

*Proof.* Recall from Remark 2.12 that $rG^0 \in M_{d_x \times d_x}(\mathbf{Z}_q[x])$. 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

$$\left( \frac{rG^0}{r'} - \ell I \right) v \equiv \frac{w}{r'} \pmod{r}$$

over $\mathbf{Q}_q[x]/(r)$. However, since $\ell \geq 1$ is not an exponent of $G^0 dx$ by Proposition 2.14, we have that $\det(\ell I - rG^0/r')$ is invertible in $\mathbf{Q}_q[x]/(r)$, so that this system has a unique solution $v$. We take

$$u = \frac{w - \left(rG^0 - \ell r' I\right) v}{r} - \frac{dv}{dx}. \qquad\square$$

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

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

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

*there exist vectors $u, v \in \mathbf{Q}_q[x, 1/x]^{\oplus d_x}$ with $\mathrm{ord}_\infty(u) > \mathrm{ord}_\infty(w)$ 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 $t = 1/x$. 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 $i \geq -1$. Putting $m = -\mathrm{ord}_\infty(w) - \deg(r) + 1$, we can also write

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

where $\bar{w}_j \in \mathbf{Q}_q^{\oplus d_x}$ for all $j \geq -(m+1)$. Note that $m \geq 1$. By Proposition 2.14, 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 take

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

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

Next we give an explicit description of the cohomology space $H^1_{\mathrm{rig}}(U)$.

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

$$E_0 = \left\{ \left( \sum_{i=0}^{d_x-1} u_i(x)b_i^0 \right) \frac{dx}{r} \qquad : u \in \mathbf{Q}_q[x]^{\oplus d_x} \right\},$$

$$E_\infty = \left\{ \left( \sum_{i=0}^{d_x-1} u_i(x, 1/x)b_i^\infty \right) \frac{dx}{r} \quad : u \in \mathbf{Q}_q[x, 1/x]^{\oplus d_x}, \mathrm{ord}_\infty(u) > \mathrm{ord}_0(W) - \deg(r) + 1 \right\},$$

$$B_0 = \left\{ \sum_{i=0}^{d_x-1} v_i(x)b_i^0 \qquad : v \in \mathbf{Q}_q[x]^{\oplus d_x} \right\},$$

$$B_\infty = \left\{ \sum_{i=0}^{d_x-1} v_i(x, 1/x)b_i^\infty \qquad : v \in \mathbf{Q}_q[x, 1/x]^{\oplus d_x}, \mathrm{ord}_\infty(v) > \mathrm{ord}_0(W) \right\}.$$

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

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

*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 $H^1_{\mathrm{rig}}(U)$ 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 $H^1_{\mathrm{dR}}(\mathbb{U})$. Now every such class can be represented by a 1-form in $E_0$ by (repeatedly) applying Proposition 3.3. Then we change basis from $[b^0_0, \ldots, b^0_{d_x-1}]$ to $[b^{\infty_0}, \ldots, b^\infty_{d_x-1}]$ by the matrix $W$. 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 the definitions of $E_0$ and $[b^0_0, \ldots, b^0_{d_x-1}]$ it follows that $\mathrm{ord}_P(\omega) \geq -1$ at all points $P$ not lying over $x = \infty$ and from the definitions of $E_\infty$ and $[b^\infty_0, \ldots, b^\infty_{d_x-1}]$ it follows that $\mathrm{ord}_P(\omega) \geq (\mathrm{ord}_0(W) + 1)e_P - 1$ 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 $f \in \mathcal{O}(\mathbb{U})$, then $\mathrm{ord}_P(f) \geq 0$ at all points $P$ not lying over $x = \infty$ and $\mathrm{ord}_P(f) \geq (\mathrm{ord}_0(W) + 1)e_P$ at all points $P$ lying over $x = \infty$. Using the definitions of $[b^0_0, \ldots, b^0_{d_x-1}]$ and $[b^\infty_0, \ldots, b^\infty_{d_x-1}]$ again, it follows that $f$ is an element of $B_0 \cap B_\infty$. $\qquad\square$

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 [13].

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

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

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

$$e_0 = \max\{e_P | P \in \mathcal{X} \setminus \mathcal{U}, x(P) \neq \infty\}.$$

*If we represent the class of $\omega$ in $H^1_{rig}(U)$ by*

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

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

$$p^{\lfloor \log_p(\ell e_0) \rfloor} u \in \mathbf{Z}_q[x]^{\oplus d_x}.$$

*Proof.* We have

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

with $f = \sum_{j=1}^{\ell} (\sum_{i=0}^{d_x-1} (v_j)_i b^0_i)/r^j$, where $v_j \in \mathbf{Q}_q[x]^{\oplus d_x}$ satisfies $\deg(f_j) < \deg(r)$ for all $1 \leq j \leq \ell$. Note that it is sufficient to show that $p^{\lfloor \log_p(\ell e_0) \rfloor} f \in \mathcal{R}$. 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, by definition $[b_0^0, \ldots, b_{d_x-1}^0]$ is a basis for $\mathcal{O}(\mathbb{X} - x^{-1}(\infty))$ over $\mathbf{Q}_q[x]$. To show that $p^{\lfloor \log_p(\ell e_0) \rfloor} f$ 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

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

of $p^{\lfloor \log_p(\ell e_0) \rfloor} f$ is integral. However, the differential $df$ has a pole of order at most $\ell e_P + 1$ at $P$, and its Laurent series expansion

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

is integral since $\omega$ is integral. The worst denominator we get by integrating this series is therefore $p^{\lfloor \log_p(\ell e_0) \rfloor}$ and the result follows. $\qquad\square$

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

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

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

$$e_\infty = \max\{e_P | P \in \mathcal{X} \setminus \mathcal{U}, x(P) = \infty\}.$$

*If we represent the class of $\omega$ in $H^1_{rig}(U)$ by*

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

*with $u \in \mathbf{Q}_q[x, x^{-1}]^{\oplus d_x}$ such that $\mathrm{ord}_\infty(u) > \mathrm{ord}_0(W^\infty) - \deg(r) + 1$ 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 $f = \sum_{j=-\mathrm{ord}_0(W^\infty)}^{m} (\sum_{i=0}^{d_x-1} (v_j)_i b_i^\infty) x^j$, where $v_j \in \mathbf{Q}_q^{\oplus d_x}$ for all $j$ such that $-\mathrm{ord}_0(W^\infty) \leq j \leq m$. Note that it is sufficient to show that $p^{\lfloor \log_p(\ell e_\infty) \rfloor} f \in \mathcal{R}$. By Assumption 1, we have that

$$\mathcal{O}(\mathcal{X} - x^{-1}(0))/(t)^k \cong \prod_{P \in \mathcal{X} \setminus \mathcal{U}, x(P) = \infty} \mathcal{O}_{\mathcal{X}, P}/(z_P^{e_P})^k, \tag{3}$$

for all $k \in \mathbf{N}$. Moreover, by definition $[b_0^\infty, \ldots, b_{d_x-1}^\infty]$ 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 $x(P) = \infty$, the Laurent series expansion

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

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

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

is integral since $\omega$ is integral. The worst denominator we get by integrating this series is therefore $p^{\lfloor \log_p(me_\infty) \rfloor}$ 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 $H^1_{\mathrm{rig}}(U)$ was reduced to a finite dimensional linear algebra problem. However, the dimension of $H^1_{\mathrm{rig}}(U)$ is generally much higher than the dimension of $H^1_{\mathrm{rig}}(X)$, 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*

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

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

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

*Moreover, we denote*

$$res_0 = \bigoplus_{P \in \mathcal{X} \setminus \mathcal{U}: \, x(P) \neq \infty} res_P, \qquad res_\infty = \bigoplus_{P \in \mathcal{X} \setminus \mathcal{U}: \, x(P) = \infty} res_P.$$

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

$$0 \longrightarrow H^1_{rig}(X) \longrightarrow H^1_{rig}(U) \xrightarrow{(res_0 \oplus res_\infty) \otimes \mathbf{Q}_q} \bigoplus_{P \in \mathcal{X} \setminus \mathcal{U}} \mathcal{O}_{\mathcal{X},P}/(z_P) \otimes \mathbf{Q}_q.$$

*Proof.* This is well known. $\square$

The kernels of $res_0$ and $res_\infty$ 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 lying over $x = \infty$.

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

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

*where $u \in \mathbf{Q}_q[x, x^{-1}]^{\oplus d_x}$ satisfies $\mathrm{ord}_\infty(u) > -\deg(r)$, and let a vector $v \in \mathbf{Q}_q^{\oplus d_x}$ be defined by $v = \left( x^{1-\deg(r)}u \right)|_{x=\infty}$. 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 $\mathcal{E}_\lambda^\infty$ denote the (generalised) eigenspace of $G_{-1}^\infty$ with eigenvalue $\lambda$, so that $\mathbf{Q}_q^{\oplus d_x}$ decomposes as $\bigoplus \mathcal{E}_\lambda^\infty$. Then*

$$res_\infty(\omega) = 0 \quad \Leftrightarrow \quad \text{the projection of } v \text{ onto } \mathcal{E}_0^\infty \text{ vanishes.}$$

*Proof.* Let $P$ run over all points in $\mathcal{X} \setminus \mathcal{U}$ such that $x(P) = \infty$. One checks that $\mathrm{ord}_P(\frac{dx}{r}) = -1 + (\deg(r)-1)e_P$ and $\mathrm{ord}_P(\omega) \geq -1$. Since $\mathrm{ord}_P(x) = -e_P$, we have that $res_P(\omega) = 0$ if and only if $\mathrm{ord}_P(\sum_{i=0}^{d_x-1} v_i b_i^\infty) \geq 1$. We still denote $t = 1/x$.

Note that $[b_0^\infty, \ldots, b_{d_x-1}^\infty]$ is a $\mathbf{Q}_q$-basis for $\mathcal{O}(\mathbb{X} - x^{-1}(0))/(t)$ and that

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

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

We know from Proposition 2.14 that the eigenvalues of $G^\infty_{-1}$ 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 $\mathrm{ord}_P(f) < e_P$ for some $P$, then we have that $\mathrm{ord}_P(f) = \lambda e_P$. We claim that the eigenvalues of $G^\infty_{-1}$ on the factor corresponding to the point $P$ in (4) are $[0, 1/e_P, \ldots, (e_P - 1)/e_P]$. In particular they are all different, so that $G^\infty_{-1}$ 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 $G^\infty_{-1}$ on the factor corresponding to the point $P$ in (4).

Now, if we decompose $v$ onto a basis of eigenvectors compatible with the decomposition (4), then we see that $\mathrm{ord}_P(\sum_{i=0}^{d_x-1} v_i b_i^\infty) \geq 1$ 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. □

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*

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

*with* $u \in \mathbf{Q}_q[x]^{\oplus d_x}$. *For every geometric point* $x_0 \in \mathcal{D}_{\mathbf{P}^1}(\bar{\mathbf{Q}}_q) \setminus \infty$, *let the vector* $v_{x_0} \in \bar{\mathbf{Q}}_q^{\oplus d_x}$ *be defined by* $v_{x_0} = u|_{x=x_0}$. *Moreover, let the residue matrix* $G^{x_0}_{-1} \in M_{d_x \times d_x}(\bar{\mathbf{Q}}_q)$ *be defined as* $G^{x_0}_{-1} = (x - x_0)G^0|_{x=x_0}$, *and let* $\mathcal{E}^{x_0}_\lambda$ *denote the (generalised) eigenspace of* $G^{x_0}_{-1}$ *with eigenvalue* $\lambda$, *so that* $\bar{\mathbf{Q}}_q^{\oplus d_x}$ *decomposes as* $\bigoplus \mathcal{E}^{x_0}_\lambda$. *Then*

$$res_0(\omega) = 0 \quad \Leftrightarrow \quad \text{the projection of } v_{x_0} \text{ onto } \mathcal{E}^{x_0}_0 \text{ vanishes}$$

$$\text{for all } x_0 \in \mathcal{D}_{\mathbf{P}^1}(\bar{\mathbf{Q}}_q) \setminus \infty.$$

*Proof.* Let $P$ be a point of $\mathcal{X} \setminus \mathcal{U}$ with $x(P) \neq \infty$. One checks that $\mathrm{ord}_P(\frac{dx}{r}) = -1$ and $\mathrm{ord}_P(\omega) \geq -1$. Therefore, we have that $res_P(\omega) = 0$ if and only if $\mathrm{ord}_P(\sum_{i=0}^{d_x-1} v_i b_i^0) \geq 1$. This can clearly be checked on geometric points lying over $P$.

Let us denote $\mathcal{X}_{\bar{\mathbf{Q}}_q} = \mathcal{X} \otimes \bar{\mathbf{Q}}_q$. Note that for all $x_0 \in \mathcal{D}_{\mathbf{P}^1}(\bar{\mathbf{Q}}_q) \setminus \infty$, we have that $[b_0^0, \ldots, b_{d_x-1}^0]$ is a $\bar{\mathbf{Q}}_q$-basis for $\mathcal{O}(\mathcal{X}_{\bar{\mathbf{Q}}_q} - x^{-1}(\infty))/(x - x_0)$ and that

$$\mathcal{O}(\mathcal{X}_{\bar{\mathbf{Q}}_q} - x^{-1}(\infty))/(x - x_0) \cong \prod_{\mathbb{P} \in \mathcal{X}(\bar{\mathbf{Q}}_q), x(\mathbb{P})=x_0} \mathcal{O}_{\mathcal{X}_{\bar{\mathbf{Q}}_q}, \mathbb{P}}/(z_{\mathbb{P}}^{e_{\mathbb{P}}}), \tag{5}$$

where $z_{\mathbb{P}}$ and $e_{\mathbb{P}}$ are defined as $z_P$ and $e_P$. Under this isomorphism every factor on the right-hand side is an invariant subspace for $G^{x_0}_{-1}$ since $\mathrm{ord}_{\mathbb{P}}(f) \geq e_{\mathbb{P}}$ implies that $\mathrm{ord}_{\mathbb{P}}((x - x_0)df/dx) \geq e_{\mathbb{P}}$.

We know from Proposition 2.14 that the eigenvalues of $G^{x_0}_{-1}$ are elements of $\mathbf{Q} \cap \mathbf{Z}_p$ contained in the interval $[0, 1)$ and if $f \in \mathcal{O}(\mathcal{X}_{\bar{\mathbf{Q}}_q} - x^{-1}(\infty))/(x - x_0)$ is an eigenvector with eigenvalue $\lambda$ and $\mathrm{ord}_{\mathbb{P}}(f) < e_{\mathbb{P}}$ for some $\mathbb{P}$, then we have that $\mathrm{ord}_{\mathbb{P}}(f) = \lambda e_{\mathbb{P}}$. We claim that the eigenvalues of $G^{x_0}_{-1}$ on the factor corresponding to the geometric point $\mathbb{P}$ in (5) are $[0, 1/e_{\mathbb{P}}, \ldots, (e_{\mathbb{P}} - 1)/e_{\mathbb{P}}]$. In particular they are all different, so that $G^{x_0}_{-1}$ is diagonalisable. This follows since locally around the geometric point

$\mathbb{P}$ the map $(x - x_0)$ is the $e_\mathbb{P}$-th power map, so the eigenvalues of its monodromy are all the $e_\mathbb{P}$-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 $G^{x_0}_{-1}$ on the factor corresponding to the geometric point $\mathbb{P}$ in (5).

Now, if we decompose $v_{x_0}$ onto a basis of eigenvectors compatible with the decomposition (5), then we see that $\mathrm{ord}_\mathbb{P}(\sum_{i=0}^{d_x-1} v_i b_i^\infty) \geq 1$ for all $\mathbb{P} \in \mathcal{X}(\bar{\mathbf{Q}}_q)$ such that $x(\mathbb{P}) = x_0$ if and only if the components of $v_{x_0}$ along the eigenvectors of $G^{x_0}_{-1}$ with eigenvalue 0 all vanish. $\qquad\square$

**Remark 3.14.** *From Proposition 2.14 we know that all the exponents of $G^0 dx$ and $G^\infty dx$ are elements of $\mathbf{Q} \cap \mathbf{Z}_p$ under Assumption 1. However, in the proofs of Proposition 3.12 and Proposition 3.13, we have seen that every $e_P$ appears as the denominator of one of these exponents. Therefore, if Assumption 1 is satisfied, then $x : X \to \mathbf{P}^1_{\mathbf{F}_q}$ must be tamely ramified.*

## 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 $\theta \leq 2.3729$ [21]. 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 $e_0, e_\infty$ as in Section 3. We have:*

$$\deg(\Delta), \deg(r), \deg(s) \leq 2(d_x - 1)d_y \qquad \in O(d_x d_y), \qquad (6a)$$

$$e_0, e_\infty \leq d_x \qquad \in O(d_x), \qquad (6b)$$

$$g \leq (d_x - 1)(d_y - 1) \in O(d_x d_y). \qquad (6c)$$

*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 [2] that $g$ is at most the number of interior points of the Newton polygon of $Q$, which is clearly bounded by $(d_x - 1)(d_y - 1)$. $\qquad\square$

For our complexity analysis, we will need one additional assumption. Note that this assumption is not required for the algorithm to work.

**Assumption 3.** *In all complexity statements, we will assume that*

$$- \mathrm{ord}_{\neq\infty}((W^0)^\pm), - \mathrm{ord}_\infty((W^\infty)^\pm), - \mathrm{ord}_0(W), - \mathrm{ord}_\infty(W^\pm) \in O(d_x d_y).$$

**Remark 4.2.** *It is not hard to show that*

$$-\operatorname{ord}_{\neq\infty}((W^0)^{-1}), -\operatorname{ord}_\infty((W^\infty)^{-1}) \leq (d_x - 1)d_y \in O(d_x d_y).$$

*For all the other* ord*'s that appear in Assumption 3, the same bound holds experimentally in all examples that we have tried. Perhaps it is known that $W^0$ and $W^\infty$ can be chosen this way, but have not found a reference.*

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

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

(1) $[\omega_1, \ldots, \omega_\kappa]$ is a basis for $H^1_{\mathrm{rig}}(U) \cong (E_0 \cap E_\infty)/d(B_0 \cap B_\infty)$,

(2) the class of every element of $(E_0 \cap E_\infty) \cap \Omega^1(\mathcal{U})$ in $H^1_{\mathrm{rig}}(U)$ has $p$-adically integral coordinates with respect to $[\omega_1, \ldots, \omega_\kappa]$,

(3) $[\omega_1, \ldots, \omega_{2g}]$ is a basis for the kernel of $res_0 \oplus res_\infty$ and hence for the subspace $H^1_{\mathrm{rig}}(X)$ of $H^1_{\mathrm{rig}}(U)$.

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

$$d_x \left( \deg(r) - 1 - \operatorname{ord}_0(W) - \operatorname{ord}_\infty(W) \right) \subset O(d_x^2 d_y),$$

where we have used Proposition 4.1 and Assumption 3. Therefore, we need $O((d_x^2 d_y)^\theta)$ ring operations in $\mathbf{Z}_q$ by [18, 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

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

### 4.2. **Step II: Compute the map $\mathrm{F}_p$.**

We use Theorem 2.9 to compute approximations:

$$\mathrm{F}_p(1/r) = \alpha_i + \mathcal{O}(p^{2^i}),$$

$$\mathrm{F}_p(y) = \beta_i + \mathcal{O}(p^{2^i}),$$

for $i = 1, \ldots, \nu = \lceil \log_2(N) \rceil$. 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:

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

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 2.16 and Assumption 3, for all elements of $\mathcal{S}$ and $\mathcal{R}$ that we encounter in the algorithm, we have that

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

Hence, the time required for a ring operation in $\mathcal{S}$ is

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

Similarly, a ring operation in $\mathcal{R}$ requires time

$$\tilde{O}\Big(pd_x^2 d_y\big(N + d_x\big)nN\Big).$$

By [12], he image of an element of $\mathbf{Q}_q$ under the map $\sigma$ can be computed in time $\tilde{O}(\log^2(p)n + \log(p)nN)$. 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

$$\tilde{O}\left(pd_x^3 d_y\big(N + d_x\big)nN\right).$$

Recall the definitions of the matrices $\Phi, \Psi \in M_{d_x \times d_x}(\mathcal{S}^\dagger)$ from the proof of Theorem 2.16. The matrix $\Phi$ can clearly be computed from $\beta_\nu$ using $O(d_x)$ ring operations in $\mathcal{R}$. It now follows from the formula

$$\Psi = (W_0/r)\Phi(rW_0^{-1})^{F_p},$$

and Assumption 3 that $\Psi$ can be computed from $\Phi$ and $\alpha_\nu$ using $O(d_x^\theta + d_x d_y)$ ring operations in $\mathcal{S}$ and

$$O\left((d_x d_y)\deg(r)d_x^2\right) \subset O(d_x^4 d_y^2)$$

applications of $\sigma$. Therefore, the matrix $\Psi$ can be computed from $(\alpha_\nu, \beta_\nu)$ in time

$$\tilde{O}\left(pd_x^{\theta+1}d_y^2(N + d_x)nN\right).$$

Finally, for each $\omega_i = \left(\sum_{k=0}^{d-1} u_k(x)b_k^0\right)dx/r$ with $1 \le i \le 2g$, we compute

$$\mathrm{F}_p(\omega_i) = \sum_{k=0}^{d_x-1} px^{p-1}u_k^\sigma(x^p)\,\mathrm{F}_p\left(\frac{b_k^0}{r}\right)dx$$

$$= \sum_{j=0}^{d_x-1}\left(\sum_{k=0}^{d_x-1} px^{p-1}u_k^\sigma(x^p)\psi_{j,k}\right)b_j^0\frac{dx}{r} + O\left(p^N\right). \tag{7}$$

For a single $\omega_i$ this takes $O(d_x^2)$ ring operations in $\mathcal{S}$ and

$$O\left(d_x\left(\deg(r) - 2 - \mathrm{ord}_0(W) - \mathrm{ord}_\infty(W)\right)\right) \subset O(d_x^3 d_y)$$

applications of $\sigma$. Hence the complete set of $\mathrm{F}_p(\omega_i)$ can be computed in time

$$\tilde{O}\left(gpd_x^3 d_y\left(N + d_x\right)nN\right) \subset \tilde{O}\left(pd_x^4 d_y^2\left(N + d_x\right)nN\right),$$

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 $\mathcal{F} \in M_{2g \times 2g}(\mathbf{Q}_q)$ such that

$$\mathrm{F}_p(\omega_i) = \sum_{j=1}^{2g}\mathcal{F}_{j,i}\omega_j$$

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

$$\mathrm{F}_p(\omega_i) = \sum_{j \in J}\left(\sum_{k=0}^{d_x-1}\frac{w_{i,j,k}(x)}{r^j}b_k^0\right)\frac{dx}{r} + O\left(p^N\right), \tag{8}$$

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

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

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 $\mathrm{F}_p(\omega_i)$. Every single finite reduction step takes time $\tilde{O}(\log(p)d_x^2\deg(r)nN)$, so all $\mathrm{F}_p(\omega_i)$ can be reduced in time

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

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

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

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

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

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

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

### 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 we have

$$\chi(T) = \det\left(1 - \mathrm{F}_p^n\, T | H^1_{\mathrm{rig}}(X)\right).$$

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

$$\mathcal{F}^{(n)} = \mathcal{F}^{\sigma^{(n-1)}}\mathcal{F}^{\sigma^{(n-2)}}\cdots\mathcal{F}.$$

Note that $\chi(T)$ is the reverse characteristic polynomial of $\mathcal{F}^{(n)}$. It is known (see for example [17]) that $\mathcal{F}^{(n)}$ can be computed from $\mathcal{F}$ in time $\tilde{O}(\log^2(p)g^\theta nN)$ and that $\chi(T)$ can be computed from $\mathcal{F}^{(n)}$ in time $\tilde{O}(\log(p)g^\theta nN)$. Therefore, the total time complexity of this step is

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

### 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 4.3.** *In order to recover $\chi(T) \in \mathbf{Z}[T]$ exactly, it is sufficient to know it to $p$-adic precision*

$$\max_{1 \leq i \leq g} \left\{ \left\lfloor \log_p\left(\frac{4g}{i}\right) + \left(\frac{ni}{2}\right) \right\rfloor + 1 \right\} \qquad \in O(d_x d_y n).$$

*Proof.* The expression for the precision is a straightforward consequence of a result of Kedlaya, which can be found in [14, Lemma 1.2.3]. That this precision is $O(d_x d_y n)$ follows from the bound on $g$ from Proposition 4.1. $\qquad \square$

**Definition 4.4.** *Let $H^1_{cris}(\mathcal{X}, \mathcal{D}_{\mathcal{X}})$ denote the log-crystalline cohomology of $\mathcal{X}$ along the divisor $\mathcal{D}_{\mathcal{X}}$. We define the following $\mathbf{Z}_q$-lattices in $H^1_{rig}(U)$:*

$$\Lambda_{E_0 \cap E_\infty} = \mathrm{im}\left( (E_0 \cap E_\infty) \cap \Omega^1(\mathcal{U}) \to H^1_{rig}(U) \right),$$
$$\Lambda_{cris} = \mathrm{im}\left( H^1_{cris}(\mathcal{X}, \mathcal{D}_{\mathcal{X}}) \to H^1_{rig}(U) \right).$$

**Definition 4.5.** *Let us denote*

$$\delta_1 = \left\lfloor \log_p\left( -(\mathrm{ord}_0(W)+1)e_\infty \right) \right\rfloor,$$
$$\delta_2 = \left\lfloor \log_p\left( (\lfloor (2g-2)/d_x \rfloor + 1)e_\infty \right) \right\rfloor,$$
$$\delta = \delta_1 + \delta_2.$$

**Proposition 4.6.** *We have the following inclusions of lattices:*

$$p^{\delta_1} \Lambda_{E_0 \cap E_\infty} \subset \Lambda_{cris} \subset p^{-\delta_2} \Lambda_{E_0 \cap E_\infty}.$$

*Proof.* Our proof generalises that of [8]. We define the effective divisor

$$\mathcal{D}_\infty = \sum_{P \in \mathcal{X} \setminus \mathcal{U} : x(P) = \infty} e_P P$$

on the curve $\mathcal{X}$. For any integer $m \geq 0$, we let $\mathcal{C}^\bullet(m)$ denote the complex

$$\mathcal{O}(m\mathcal{D}_\infty) \longrightarrow \Omega^1(\log(\mathcal{D}_{\mathcal{X}})) \otimes \mathcal{O}(m\mathcal{D}_\infty),$$

i.e. the De Rham complex on $\mathcal{X}$ with logarithmic poles along $\mathcal{D}_{\mathcal{X}}$ twisted by the line bundle $\mathcal{O}(m\mathcal{D}_\infty)$. Note that $\mathcal{C}^\bullet(l)$ is a subcomplex of $\mathcal{C}^\bullet(m)$ whenever $l \leq m$. From the comparison theorem between log-De Rham and log-crystalline cohomology, we know that $\mathbb{H}^1(\mathcal{C}^\bullet(0)) = H^1_{cris}(\mathcal{X}, \mathcal{D}_{\mathcal{X}})$.

For any integer $m \geq 0$, we have the following diagram:

$$
\begin{array}{ccccc}
H^0(\Omega^1(\log(\mathcal{D}_{\mathcal{X}}))) & \longrightarrow & \mathbb{H}^1(\mathcal{C}(0)) & \longrightarrow & H^1(\mathcal{O}) \\
\downarrow & & \downarrow & & \downarrow \\
H^0(\Omega^1(\log(\mathcal{D}_{\mathcal{X}})) \otimes \mathcal{O}(m\mathcal{D}_{\infty})) & \longrightarrow & \mathbb{H}^1(\mathcal{C}(m)) & \longrightarrow & H^1(\mathcal{O}(m\mathcal{D}_{\infty})) \\
\downarrow & & \downarrow & & \\
\displaystyle\bigoplus_{P \in \mathcal{X} \backslash \mathcal{U} : x(P) = \infty} \dfrac{z_P^{-me_P} \mathbf{Z}_q[[z_P]] \frac{dz_P}{z_P}}{\mathbf{Z}_q[[z_P]] \frac{dz_P}{z_P}} & \longrightarrow & \displaystyle\bigoplus_{\substack{P \in \mathcal{X} \backslash \mathcal{U} : x(P) = \infty \\ -me_P \leq i < 0}} (\mathbf{Z}_q/i\mathbf{Z}_q) z_P^i \frac{dz_P}{z_P} & & \\
\downarrow & & \downarrow & & \\
0 & & 0 & &
\end{array}
$$

where the first two rows and columns are exact and all (hyper)cohomology is taken with respect to global sections on $\mathcal{X}$. Hence the cokernel of the map $\mathbb{H}^1(\mathcal{C}(0)) \rightarrow \mathbb{H}^1(\mathcal{C}(m))$ is annihilated by $p^{\lfloor \log_p(me_\infty) \rfloor}$. For $m = -(\mathrm{ord}_0(W) + 1)$, we have that $H^0(\Omega^1(\log(\mathcal{D}_{\mathcal{X}})) \otimes \mathcal{O}(m\mathcal{D}_{\infty})) = (E_0 \cap E_\infty) \cap \Omega^1(\mathcal{U})$. Therefore, it follows that $p^{\delta_1} \Lambda_{E_0 \cap E_\infty} \subset \Lambda_{cris}$.

We now prove the other inclusion. For $m = \lfloor (2g-2)/d_x \rfloor + 1$, it follows from Serre duality that $H^1(\mathcal{O}(m\mathcal{D}_{\infty})) = H^0(\mathcal{O}(\omega_{\mathcal{X}} - m\mathcal{D}_{\infty})) = 0$, since we have that $\deg(m\mathcal{D}_{\infty}) > 2g - 2 = \deg(\omega_{\mathcal{X}})$. So the map $H^0(\Omega^1(\log(\mathcal{D}_{\mathcal{X}})) \otimes \mathcal{O}(m_1\mathcal{D}_{\infty})) \rightarrow \Lambda_{cris}$ is surjective. However, by Proposition 3.4, the class in $H^1_{\mathrm{rig}}(U)$ of an element of $H^0(\Omega^1(\log(\mathcal{D}_{\mathcal{X}})) \otimes \mathcal{O}(m_1\mathcal{D}_{\infty}))$ can be represented by an element of $p^{-\delta_2} \Lambda_{E_0 \cap E_\infty}$. This finishes the proof. $\qquad\square$

**Corollary 4.7.** *We have that* $\mathrm{ord}_p(\mathcal{F}) \geq -\delta$.

*Proof.* Note that $\Lambda_{cris}$ is mapped into itself by $F_p$ and that the basis $[\omega_1, \ldots, \omega_\kappa]$ for $H^1_{\mathrm{rig}}(U)$ is by construction a basis for $\Lambda_{E_0 \cap E_\infty}$. Therefore, the result follows from Proposition 4.6. $\qquad\square$

**Proposition 4.8.** *In order to recover* $\chi(T) \in \mathbf{Z}[T]$ *exactly, it is sufficient to know the matrix* $\mathcal{F}$ *to $p$-adic precision*

$$
\max_{1 \leq i \leq g} \left\{ \left\lfloor \log_p\left(\frac{4g}{i}\right) + \left(\frac{ni}{2}\right) \right\rfloor + 1 \right\} + \delta \qquad\qquad \in O(d_x d_y n).
$$

*Proof.* We have to compute $\mathcal{F}^{(n)}$ and its reverse characteristic polynomial $\chi$. The basis $[\omega_1, \ldots, \omega_\kappa]$ for $H^1_{\mathrm{rig}}(U)$ that we constructed is a basis for $\Lambda_{E_0 \cap E_\infty}$. Note that with respect to a basis for $\Lambda_{cris}$ there would be no loss of precision in the computation. Therefore, the result follows from Proposition 4.6 by changing basis from $[\omega_1, \ldots, \omega_\kappa]$ to a basis for $\Lambda_{cris}$, computing $\chi(T)$ with respect to this basis, and changing basis back to $[\omega_1, \ldots, \omega_\kappa]$. $\qquad\square$

**Definition 4.9.** *We define* $f : \mathbf{N} \rightarrow \mathbf{Z}_{\geq 0}$ *and* $g \in \mathbf{Z}_{\geq 0}$ *by*

$$
f(N) = \left\lfloor \log_p\big(p(N-1)e_0\big) \right\rfloor + \left\lfloor \log_p\big(-(\mathrm{ord}_\infty(W^{-1}) + 1)e_\infty\big) \right\rfloor,
$$
$$
g = \left\lfloor \log_p\big(-p(\mathrm{ord}_0(W) + 1)e_\infty\big) \right\rfloor.
$$

**Proposition 4.10.** *In order to recover $\chi(T) \in \mathbf{Z}[T]$ exactly, it is sufficient to choose the p-adic precision $N$ such that*

$$N - \max\{f(N), g\} \geq \max_{1 \leq i \leq g} \left\{ \left\lfloor \log_p\left(\frac{4g}{i}\right) + \left(\frac{ni}{2}\right) \right\rfloor + 1 \right\} + \delta,$$

*so we may take $N \in \tilde{O}(d_x d_y n)$.*

*Proof.* We divide the sum in (8) up into two parts, consisting of the terms with $j > 0$ and the ones with $j \leq 0$.

First, consider the terms with $j > 0$. Because of Proposition 2.16 (3), and the factor $p$ appearing in (7), we have that $j \leq p(N-1)$. Therefore, it follows from Proposition 3.7 that the loss of precision during the reductions at the points not lying over $\infty$ is at most $\left\lfloor \log_p\left(p(N-1)e_0\right) \right\rfloor$. However, the reductions at the points not lying over $\infty$ can introduce a (small) pole at the points lying over $\infty$, which still has to be reduced. The matrix of the change of basis from $[b_0^0, \ldots, b_{d_x-1}^0]$ to $[b_0^\infty, \ldots, b_{d_x-1}^\infty]$ is $W^{-1}$ and $\mathrm{ord}_\infty(v_i/r^\ell) \geq 1$ for all $0 \leq i \leq d_x - 1$ and $\ell > 0$ in Proposition 3.3. Therefore, it follows from Proposition 3.8 that the loss of precision during these final reductions at the points lying over $\infty$ is at most $\left\lfloor \log_p\left(-(\mathrm{ord}_\infty(W^{-1})+1)e_\infty\right) \right\rfloor$. We conclude that the total loss of precision during the reductions of the terms in (8) with $j > 0$ is at most $f(N)$.

Second, consider the terms with $j \leq 0$. By the definition of $E_\infty$, the coefficients of the 1-forms $[\omega_0, \ldots, \omega_{2g}]$ with respect to the basis $[b_0^\infty, \ldots, b_{d_x-1}^\infty]dx/r$ have order at $\infty$ bounded below by $\mathrm{ord}_\infty(u_i) \geq \mathrm{ord}_0(W) - \deg(r) + 2$ for all $0 \leq i \leq d_x - 1$. Therefore, with respect to the basis $[b_0^\infty, \ldots, b_{d_x-1}^\infty]dx/x$ the coefficients have order at $\infty$ bounded below by $\mathrm{ord}_0(W) + 1$. By (the proof of) Proposition 2.16, the Frobenius structure on $\mathbf{R}^0 x_*(\mathcal{O}_\mathbb{U})$ does not have a pole at $\infty$ with respect to the basis $[b_0^\infty, \ldots, b_{d_x-1}^\infty]$. Moreover, note that $\mathrm{F}_p$ sends the 1-form $dx/x$ to $pdx/x$. Hence the coefficients of $\mathrm{F}_p(\omega_i)$ in (8) with respect to the basis $[b_0^\infty, \ldots, b_{d_x-1}^\infty]dx/x$ have order at $\infty$ bounded below by $p(\mathrm{ord}_0(W) + 1)$. So the coefficients of $\mathrm{F}_p(\omega_i)$ with respect to the basis $[b_0^\infty, \ldots, b_{d_x-1}^\infty]dx/r$ have order at $\infty$ bounded below by $p(\mathrm{ord}_0(W) + 1) - \deg(r) + 1$. Therefore, it follows from Proposition 3.8 that the loss of precision during the reductions of the terms in (8) with $j < 0$ is at most $\left\lfloor \log_p\left(-p(\mathrm{ord}_0(W) + 1)e_\infty)\right) \right\rfloor$. □

**Remark 4.11.** *To be entirely precise, we should mention that at two points in the algorithm a p-adic precision slightly higher than $N$ is required:*

(1) *the cohomological reductions are to be carried out to absolute p-adic precision $N$. However, sometimes in a reduction step the differential that has to be reduced already has a p-adic denominator, so that the reduction matrices computed in Step III have to be known to precision $N + f(N)$ and $N + f(N) + g$ for the points away from infinity and the ones lying over infinity, respectively,*

(2) *while in Step IV it is sufficient to know $\mathcal{F}$ to the p-adic precision given by Proposition 4.8, the working precision used for intermediate rounding in the computation of $\mathcal{F}^{(n)}$ and its reverse characteristic polynomial has to be*

$$\max_{1 \leq i \leq g} \left\{ \left\lfloor \log_p\left(\frac{4g}{i}\right) + \left(\frac{ni}{2}\right) \right\rfloor + 1 \right\} + (n - 1 + 2g - 1)\delta.$$

*In both cases this higher p-adic precision is still $\tilde{O}(d_x d_y n)$ and does not have a significant effect on running times either, since it only affects the p-adic precision as such and not the number of terms that have to be computed in the different expansions as is the case for $N$.*

**Theorem 4.12.** *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 4.10, 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, one can prove the following theorem.

**Theorem 4.13.** *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 $\mathrm{F}_p(\omega_i)$, 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 4.10. □

**Remark 4.14.** *To be entirely precise, we should mention that two things have been excluded from our complexity estimates:*

(1) *the computation of the matrices $W^0$ and $W^\infty$,*
(2) *the computation of the matrices of the maps $res_0$ and $res_\infty$, or rather the computation of the eigenspaces of the matrices $G_{-1}^{x_0}$ and $G_{-1}^\infty$.*

*We treat both of these as black boxes. Analyzing the available algorithms would take us too far, as they involve for example factorizing polynomials. In all examples that we have computed with our implementation, the time and space that Magma requires for these computations is neglible compared to those for the rest of our algorithm.*

## 5. Implementation

We have updated our Magma [3] implementation from [20] in the case of prime-fields (i.e. the package `pcc_p`). The code can be found at our webpage[2]. Work on updating the implementation in the case of non-primefields (i.e. the package `pcc_q`) is in progress and should be finished soon. We now provide two examples that our implementation from [20] was not able to handle. 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 and `pcc_p-2.7` on a 3.0GHz Intel Core i7-3540M processor.

*Example 1.* The modular curve $X_1(23)$.

Sutherland [19] gives an equation $\mathcal{Q}$ for a plane model of the modular curve $X_1(23)$. This equation can be loaded into our code in the following way:

```
load "pcc_q.m";
Q:=y^7+(x^5-x^4+x^3+4*x^2+3)*y^6+(x^7+3*x^5+x^4+5*x^3+7*x^2-4*x+3)*y^5+(2*x^7+3*x^5-x^4-2*x^3-x^2-8*x+1)*y^4+
(x^7-4*x^6-5*x^5-6*x^4-6*x^3-2*x^2-3*x)*y^3-(3*x^6-5*x^4-3*x^3-3*x^2-2*x)*y^2+(3*x^5+4*x^4+x)*y-x^2*(x+1)^2;
```

---

[2]`http://perswww.kuleuven.be/jan_tuitman`

Note that $d_x = 7$, which is known to be optimal [7]. It turns out that $\mathcal{Q}$ satisfies Assumption 1 for all prime numbers

$$p \notin \{2, 3, 23, 41, 73, 83, 2039\}.$$

To compute the numerator of the zeta function of $X_1(23)$ modulo 11, we enter the following commands:

```
p:=11;
N:=0;
chi:=num_zeta(Q,p,N:verbose:=true);
```

Here the $p$-adic precision $N$ is set to 0, which means that we let the code determine a provably correct precision, using the results in this paper. In this case this provably correct precision is 11, while experimentally it turns out that precision 9 is still sufficient to obtain the correct answer. In the tabel below we collect some data on the provably correct precision $N$, running time and memory usage for various prime numbers $p$.

| p | N | time | memory |
|---|---|---|---|
| $2^2 + 1$ | 18 | 58.2 s | 73 Mb |
| $2^3 + 3$ | 11 | 70.4 s | 73 Mb |
| $2^4 + 1$ | 10 | 111.8 s | 93 Mb |
| $2^5 + 5$ | 9 | 255.4 s | 146 Mb |
| $2^6 + 3$ | 8 | 441.6 s | 214 Mb |
| $2^7 + 3$ | 8 | 1026.9 s | 371 Mb |

Actually, there is a more efficient way to compute the zeta function of a modular curve modulo a prime number $p$, using modular symbols. However, our algorithm works and has about the same running time and memory usage for any curve of genus 12 with a map of degree 7 to the projective line!

*Example 2.* The intersection of three quadrics in $\mathbf{P}^4$.

We consider the curve $\mathcal{X}$ over $\mathbf{Z}$ that is the intersection of the following three quadrics in $\mathbf{P}^4$:

```
g1:=-2*x0^2-x0*x1+2*x0*x2-x0*x3+3*x0*x4+2*x1^2+2*x1*x2-3*x1*x3+2*x1*x4+3*x2*x3-2*x2*x4-2*x3^2+3*x3*x4-x4^2;
g2:=x0^2+3*x0*x1+x0*x2+2*x0*x3+2*x0*x4-x1^2-2*x1*x2+2*x1*x3+2*x1*x4-2*x2^2+2*x2*x3+x3^2-2*x3*x4;
g3:=x0^2-3*x0*x1-3*x0*x3+3*x0*x4-x1*x2-x1*x3+3*x1*x4-3*x2^2+2*x2*x3+3*x2*x4-x3^2+x3*x4+2*x4^2;
```

It is known that the intersection of three such quadrics is generically a genus 5 curve that does not have a nondegenerate model in the sense of [4]. Using the Magma function `Genus5PlaneCurveModel` we can compute an equation $\mathcal{Q}$ for a plane model of the curve. This equation can be loaded into our code in the following way:

```
load "pcc_p.m";
Q:=y^6-1676934*y^5*x-4686402670719*y^4*x^2+15696362774510389026*y^3*x^3-55124559862209734538679587*y^2*x^4
-12426954155876087436514613041719*y*x^5+12387480414774175009296660783451616802*x^6+1238738*y^5+4311969288924*y^4*x
+4465251377444310750*y^3*x^2-409626113445348805406666148*y^2*x^3+23782828369628730861153991263231*y*x^4
+1155759032566287728149204151297589545*x^5+621429270264*y^4+4142281363533276678*y^3*x+1747680928896108814456899 6*y^2*x^2
+15441330427942847249621800244376*y*x^3+702696445506259805789719429312894788*x^4+1687186698521956434*y^3
+5474941829683717195663056*y^2*x-1105459121160795888720225718862 4*y*x^2+935772151623461779074550065875834 3070*x^3
+18370437585555563687791 26*y^2+598747551487309181249973752058 9*y*x+5798417136543799255353703675081070079*x^2
+10287283997203637915813597689 59*y-16229692654258133152315647321928560 63*x+5030974936216271474383179756160414 98;
```

It turns out that $\mathcal{Q}$ satisfies Assumption 1 for all prime numbers

$$p \notin \{2, 3, 5, 7, 11, 13, 37, 59, 157, 251, 269, 349, 521, 839, 883, 2213, 3167, 5023,$$
$$9817, 11549, 76487, 79631, 814531, 857977, 1215521, 1811179, 3409999\}.$$

To compute the numerator of the zeta function of $\mathcal{X}$ modulo 17, we enter the following commands:

```
p:=17;
N:=0;
chi:=num_zeta(Q,p,N:verbose:=true);
```

In this case the provably correct precision computed by the code is 4, which is experimentally found to be sharp.

| p | N | time | memory |
|---|---|---|---|
| $2^4 + 1$ | 4 | 125.6 s | 64Mb |
| $2^5 - 1$ | 4 | 144.6 s | 64Mb |
| $2^6 + 3$ | 4 | 189.5 s | 64Mb |
| $2^7 + 3$ | 4 | 270.8 s | 86Mb |
| $2^8 + 1$ | 4 | 441.6 s | 146Mb |
| $2^9 - 3$ | 4 | 792.9 s | 209Mb |
| $2^{10} + 7$ | 4 | 1637.7 s | 404Mb |

**Remark 5.1.** *Note that for our plane model of $\mathcal{X}$, we have that $d_x = d_y = 6$. However, such a model does not always exist. Wouter Castryck has Magma code that given three quadrics in $\mathbf{P}^4$ over a finite field, finds a lift to characteristic zero with $d_x = 4$ and $d_y = 10$ as long as $p$ is not too small. More details on this are to appear elsewhere.*

## References

1. Francesco Baldassarri and Bruno 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. MR 1307391 (96f:14024)
2. Peter Beelen and Ruud Pellikaan, *The Newton polygon of plane curves with many rational points*, Des. Codes Cryptogr. **21** (2000), no. 1-3, 41–67, Special issue dedicated to Dr. Jaap Seidel on the occasion of his 80th birthday (Oisterwijk, 1999). MR 1801161 (2003c:14024)
3. Wieb Bosma, John Cannon, and Catherine Playoust, *The Magma algebra system. I. The user language*, J. Symbolic Comput. **24** (1997), no. 3-4, 235–265, Computational algebra and number theory (London, 1993). MR MR1484478
4. W. Castryck, J. Denef, and F. Vercauteren, *Computing zeta functions of nondegenerate curves*, IMRP Int. Math. Res. Pap. (2006), Art. ID 72017, 57. MR 2268492 (2007h:14026)
5. Jan Denef and Frederik Vercauteren, *Counting points on $C_{ab}$ curves using Monsky-Washnitzer cohomology*, Finite Fields Appl. **12** (2006), no. 1, 78–102. MR 2190188 (2007c:11075)
6. ———, *An extension of Kedlaya's algorithm to hyperelliptic curves in characteristic 2*, J. Cryptology **19** (2006), no. 1, 1–25. MR 2210897 (2007d:11069)
7. Maarten Derickx and Mark Van Hoeij, *Gonality of the modular curve $X_1(N)$*, preprint (2014), http://arxiv.org/pdf/1307.5719v3.pdf.
8. Bas Edixhoven, *Point counting after Kedlaya*, lecture notes (2006), http://www.math.leidenuniv.nl/~edix/oww/mathofcrypt/carls_edixhoven/kedlaya.pdf.
9. Pierrick Gaudry and Nicolas 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. MR 1934859 (2003h:11159)
10. David Harvey, *Kedlaya's algorithm in larger characteristic*, Int. Math. Res. Not. IMRN (2007), no. 22, Art. ID rnm095, 29. MR 2376210 (2009d:11096)
11. ———, *Counting points on hyperelliptic curves in average polynomial time*, Ann. of Math. (2) **179** (2014), no. 2, 783–803. MR 3152945

12. Hendrik Hubrechts, *Fast arithmetic in unramified p-adic fields*, Finite Fields Appl. **16** (2010), no. 3, 155–162. MR 2610706 (2011d:11277)
13. Kiran S. Kedlaya, *Counting points on hyperelliptic curves using Monsky-Washnitzer cohomology*, J. Ramanujan Math. Soc. **16** (2001), no. 4, 323–338. MR 1877805 (2002m:14019)
14. _____, *Effective p-adic cohomology for cyclic cubic threefolds*, Computational algebraic and analytic geometry, Contemp. Math., vol. 572, Amer. Math. Soc., Providence, RI, 2012, pp. 127–171. MR 2953828
15. Kiran S. Kedlaya and Jan. Tuitman, *Effective convergence bounds for Frobenius structures on connections*, Rend. Semin. Mat. Univ. Padova. (2012), 7–16.
16. Alan G. B. Lauder, *A recursive method for computing zeta functions of varieties*, LMS J. Comput. Math. **9** (2006), 222–269. MR 2261044 (2007g:14022)
17. Sebastian Pancratz and Jan Tuitman, *Improvements to the deformation method for counting points on smooth projective hypersurfaces*, preprint (2013), http://arxiv.org/abs/1307.1250.
18. Arne Storjohann, *Algorithms for matrix canonical forms*, PhD thesis, Swiss Federal Institute of Technology – ETH, 2000.
19. Andrew V. Sutherland, *Constructing elliptic curves over finite fields with prescribed torsion*, Mathematics of Computation **81** (2012), 1131–1147.
20. Jan Tuitman, *Counting points on curves using a map to $\mathbf{P}^1$*, Mathematics of Computation (to appear) (2014).
21. Virginia Vassilevska 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. MR 2961552
22. Yifan Yang, *Defining equations of modular curves*, Advances in Mathematics **204** (2006), 481–508.

KU Leuven, Departement Wiskunde, Celestijnenlaan 200B, 3001 Leuven, Belgium
*E-mail address*: jan.tuitman@wis.kuleuven.be