A quantitative Chevalley-Weil theorem for curves

Bilu, Yuri and Strambi, Marco and Surroca, Andrea

Quantitative Chevalley-Weil Theorem for Curves

Yuri Bilu[1] (Bordeaux),
Marco Strambi (Livorno),
Andrea Surroca[2] (Basel)

December 23, 2011

Abstract

The classical Chevalley-Weil theorem asserts that for an étale covering of projective varieties over a number field $\mathbb{K}$, the discriminant of the field of definition of the fiber over a $\mathbb{K}$-rational point is uniformly bounded. We obtain a fully explicit version of this theorem in dimension 1.

Contents

1 Introduction

The Chevalley-Weil theorem is one of the most basic principles of the Diophantine analysis. Already Diophantus of Alexandria routinely used reasoning of the kind "if a and b are 'almost' co-prime integers and ab is a square, then each of a and b is 'almost' a square". The Chevalley-Weil [5, 20] theorem provides a general set-up for this kind of arguments.

Theorem 1.1 (Chevalley-Weil) *Let $\widetilde{V} \stackrel{\phi}{\to} V$ be a finite étale covering of normal projective varieties, defined over a number field $\mathbb{K}$. Then there exists a non-zero integer T such that for any $P \in V(K)$ and $\widetilde{P} \in \widetilde{V}(\bar{K})$ such that $\phi(\widetilde{P}) = P$, the relative discriminant of $\mathbb{K}(\widetilde{P})/\mathbb{K}(P)$ divides T.*

There is also a similar statement for coverings of affine varieties and integral points. See [13, Section 2.8] or [18, Section 4.2] for more details.

The Chevalley-Weil theorem is indispensable in the Diophantine analysis, because it reduces a Diophantine problem on the variety V to that on the covering variety $\widetilde{V}$, which can often be simpler to deal. In particular, the Chevalley-Weil theorem is used, albeit implicitly, in the proofs of the great finiteness theorems of Mordell-Weil, Siegel and Faltings.

In view of all this, a quantitative version of the Chevalley-Weil theorem, at least in dimension 1, would be useful to have. One such version appears in Chapter 4 of [1], but it is not explicit in

[1]Supported by the ANR project HAMOT, and *Ambizione* fund PZ00P2_121962 of the Swiss National Foundation
[2]Supported by the Marie Curie IEF 025499 of the European Community and the Ambizione Fund PZ00P2_121962 of the Swiss National Science Foundation

all parameters; neither is the version recently suggested by Draziotis and Poulakis [7, 8], who also make some other restrictive assumptions (see Remark 1.4 below).

In the present article we present a version of the Chevalley-Weil theorem in dimension 1, which is explicit in all parameters and considerably sharper than the previous versions. Our approach is different from that of [7, 8] and goes back to [1, 2].

To state our principal results, we have to introduce some notation. Let $\mathbb{K}$ be a number field, $\mathcal{C}$ an absolutely irreducible smooth projective curve $\mathcal{C}$ defined over $\mathbb{K}$, and $x \in \mathbb{K}(\mathcal{C})$ a non-constant $\mathbb{K}$-rational function on $\mathcal{C}$. We also fix a covering $\widetilde{\mathcal{C}} \xrightarrow{\phi} \mathcal{C}$ of $\mathcal{C}$ by another smooth irreducible projective curve $\widetilde{\mathcal{C}}$; we assume that both $\widetilde{\mathcal{C}}$ and the covering ϕ are defined over $\mathbb{K}$. We consider $\mathbb{K}(\mathcal{C})$ as a subfield of $\mathbb{K}(\widetilde{\mathcal{C}})$; in particular, we identify the functions $x \in \mathbb{K}(\mathcal{C})$ and $x \circ \phi \in \mathbb{K}(\widetilde{\mathcal{C}})$.

We also fix one more rational function $y \in \mathbb{K}(\mathcal{C})$ such that $\mathbb{K}(\mathcal{C}) = \mathbb{K}(x, y)$ (existence of such y follows from the primitive element theorem). Let $f(X, Y) \in \mathbb{K}[X, Y]$ be the $\mathbb{K}$-irreducible polynomial such that $f(x, y) = 0$ (it is well-defined up to a constant factor). Since $\mathcal{C}$ is absolutely irreducible, so is the polynomial $f(X, Y)$. We put $m = \deg_X f$ and $n = \deg_Y f$.

Similarly, we fix a function $\widetilde{y} \in \mathbb{K}(\widetilde{\mathcal{C}})$ such that $K(\widetilde{\mathcal{C}}) = \mathbb{K}(x, \widetilde{y})$. We let $\tilde{f}(X, \widetilde{Y}) \in \mathbb{K}[X, \widetilde{Y}]$ be an irreducible polynomial such that $\tilde{f}(x, \widetilde{y}) = 0$. We put $\widetilde{m} = \deg_X \tilde{f}$ and $\widetilde{n} = \deg_Y \tilde{f}$. We denote by ν the degree of the covering ϕ, so that $\widetilde{n} = n\nu$.

Remark 1.2 Equations $f(X, Y) = 0$ and $\tilde{f}(X, \widetilde{Y}) = 0$ define affine plane models of our curves $\mathcal{C}$ and $\widetilde{\mathcal{C}}$; *we do not assume these models non-singular*.

In the sequel, $\mathrm{h_p}(\cdot)$ and $\mathrm{h_a}(\cdot)$ denote the projective and the affine absolute logarithmic heights, respectively, see Section 2 for the definitions. We also define normalized logarithmic discriminant $\partial_{\mathbb{L}/\mathbb{K}}$ and the height $\mathrm{h}(S)$ of a finite set of places S as

$$\partial_{\mathbb{L}/\mathbb{K}} = \frac{\log \mathcal{N}_{\mathbb{K}/\mathbb{Q}} \mathcal{D}_{\mathbb{L}/\mathbb{K}}}{[\mathbb{L} : \mathbb{Q}]}, \qquad \mathrm{h}(S) = \frac{\sum_{v \in S} \log \mathcal{N}_{\mathbb{K}/\mathbb{Q}}(v)}{[\mathbb{K} : \mathbb{Q}]};$$

see Section 2 for the details.

Put

$$\begin{gathered} \Omega = mn^2\big(\mathrm{h_p}(f) + 2m + 2n\big), \qquad \widetilde{\Omega} = \widetilde{m}\widetilde{n}^2\big(\mathrm{h_p}(\tilde{f}) + 2\widetilde{m} + 2\widetilde{n}\big), \\ \Upsilon = 2\widetilde{n}\big(\widetilde{m}\mathrm{h_p}(f) + m\mathrm{h_p}(\tilde{f})\big). \end{gathered} \tag{1}$$

Theorem 1.3 ("projective" Chevalley-Weil theorem) *In the above set-up, assume that the covering $\widetilde{\mathcal{C}} \xrightarrow{\phi} \mathcal{C}$ is unramified. Then for every $P \in \mathcal{C}(\bar{\mathbb{K}})$ and $\widetilde{P} \in \widetilde{\mathcal{C}}(\bar{\mathbb{K}})$ such that $\phi(\widetilde{P}) = P$ we have*

$$\partial_{\mathbb{K}(\widetilde{P})/\mathbb{K}(P)} \le 400(\Omega + \widetilde{\Omega}) + 2\Upsilon + 6m\widetilde{n}^2.$$

Remark 1.4 Draziotis and Poulakis [8, Theorem 1.1], assume that $\mathcal{C}$ is a non-singular plane curve (which is quite restrictive) and that $P \in \mathcal{C}(\mathbb{K})$. Their set-up is slightly different, and the two estimates cannot be compared directly. But it would be safe to say that their estimate is not sharper than

$$\partial_{\mathbb{K}(\widetilde{P})/\mathbb{K}(P)} \le cN^{30}\widetilde{N}^{13}\left(\mathrm{h_p}(f) + \mathrm{h_p}(\tilde{f})\right) + C,$$

where $N = \deg f$, $\widetilde{N} = \deg \tilde{f}$, the constant c is absolute and C depends of N, $\widetilde{N}$ and the degree $[\mathbb{K} : \mathbb{Q}]$.

Now let S be a finite set of places of $\mathbb{K}$, including all the archimedean places. A point $P \in \mathcal{C}(\bar{\mathbb{K}})$ will be called *S-integral* if for any $v \in M_{\mathbb{K}} \smallsetminus S$ and any extension $\bar{v}$ of v to $\bar{\mathbb{K}}$ we have $|x(P)|_{\bar{v}} \le 1$.

Theorem 1.5 ("affine" Chevalley-Weil theorem) *In the above set-up, assume that the covering $\widetilde{\mathcal{C}} \xrightarrow{\phi} \mathcal{C}$ is unramified outside the poles of x. Then for every S-integral point $P \in \mathcal{C}(\bar{\mathbb{K}})$ and $\widetilde{P} \in \widetilde{\mathcal{C}}(\bar{\mathbb{K}})$ such that $\phi(\widetilde{P}) = P$ we have*

$$\partial_{\mathbb{K}(\widetilde{P})/\mathbb{K}(P)} \le 300(\Omega + \widetilde{\Omega}) + \Upsilon + 3m\widetilde{n}^2 + \mathrm{h}(S). \tag{2}$$

Again, Draziotis and Poulakis [7, Theorem 1.1] obtain a less sharp result under more restrictive assumptions.

It might be also useful to have a statement free of the defining equations of the curves $\mathcal{C}$ and $\widetilde{\mathcal{C}}$. Using the result of [4], we obtain versions of Theorems 1.3 and 1.5, which depend only on the degrees and the ramification points of our curves over $\mathbb{P}^1$. For a finite set $A \subset \mathbb{P}^1(\bar{\mathbb{K}})$ we define $\mathrm{h_a}(A)$ as the affine height of the vector whose coordinates are the finite elements of A.

Theorem 1.6 *Let A be a finite subset of $\mathbb{P}^1(\bar{\mathbb{K}})$ such that the covering $\mathcal{C} \xrightarrow{x} \mathbb{P}^1$ is unramified outside A. Put*

$$\delta = [\mathbb{K}(A) : \mathbb{K}], \qquad \widetilde{\mathbf{g}} = \mathbf{g}(\widetilde{\mathcal{C}}), \qquad \Lambda = \big((\widetilde{\mathbf{g}}+1)\widetilde{n}\big)^{25(\widetilde{\mathbf{g}}+1)\widetilde{n}} + 2(\delta-1).$$

1. *Assume that the covering $\phi : \widetilde{\mathcal{C}} \to \mathcal{C}$ is unramified. Then for every $P \in \mathcal{C}(\bar{\mathbb{K}})$ and $\widetilde{P} \in \widetilde{\mathcal{C}}(\bar{\mathbb{K}})$ such that $\phi(\widetilde{P}) = P$ we have*

$$\partial_{\mathbb{K}(\widetilde{P})/\mathbb{K}(P)} \le \Lambda\big(\mathrm{h_a}(A)+1\big).$$

2. *Assume that the covering $\phi : \widetilde{\mathcal{C}} \to \mathcal{C}$ is unramified outside the poles of x, and let S be as above. Then for every S-integral point $P \in \mathcal{C}(\bar{\mathbb{K}})$ and $\widetilde{P} \in \widetilde{\mathcal{C}}(\bar{\mathbb{K}})$ such that $\phi(\widetilde{P}) = P$ we have*

$$\partial_{\mathbb{K}(\widetilde{P})/\mathbb{K}(P)} \le \mathrm{h}(S) + \Lambda\big(\mathrm{h_a}(A)+1\big).$$

Acknowledgments The authors thank Carlo Gasbarri for useful discussions. Yuri Bilu thanks the University of Basel for hospitality in late 2011, when a substantial part of this work was done.

2 Preliminaries

Let $\mathbb{K}$ be any number field and let $M_{\mathbb{K}} = M_{\mathbb{K}}^0 \cup M_{\mathbb{K}}^\infty$ be the set of its places, with $M_{\mathbb{K}}^0$ and $M_{\mathbb{K}}^\infty$ denoting the sets of finite and infinite places, respectively. For every place $v \in M_{\mathbb{K}}$ we normalize the corresponding valuation $|\cdot|_v$ so that its restriction to $\mathbb{Q}$ is the standard infinite or p-adic valuation. Also, we let $\mathbb{K}_v$ be the v-adic completion of $\mathbb{K}$, (in particular, $\mathbb{K}_v$ is $\mathbb{R}$ or $\mathbb{C}$ when v is infinite).

Heights For a vector $\underline{\alpha} = (\alpha_1, \dots, \alpha_N) \in \bar{\mathbb{Q}}^N$ we define, as usual, the *absolute logarithmic projective height* and *absolute logarithmic affine height* (in the sequel simply *projective* and *affine heights*) by[3]

$$\mathrm{h_p}(\underline{\alpha}) = \frac{1}{[\mathbb{K}:\mathbb{Q}]} \sum_{v \in M_{\mathbb{K}}} [\mathbb{K}_v : \mathbb{Q}_v] \log \|\underline{\alpha}\|_v, \qquad \mathrm{h_a}(\underline{\alpha}) = \frac{1}{[\mathbb{K}:\mathbb{Q}]} \sum_{v \in M_{\mathbb{K}}} [\mathbb{K}_v : \mathbb{Q}_v] \log^+ \|\underline{\alpha}\|_v, \quad (3)$$

where $\mathbb{K}$ is any number field containing the coordinates of $\underline{\alpha}$,

$$\|\underline{\alpha}\|_v = \max\{|\alpha_0|_v, \dots, |\alpha_N|_v\}$$

and $\log^+ = \max\{\log, 0\}$. With our choice of normalizations, the right-hand sides in (3) are independent of the choice of the field $\mathbb{K}$. For a polynomial f with algebraic coefficients we denote by $\mathrm{h_p}(f)$ and by $\mathrm{h_a}(f)$ the projective height and the affine height of the vector of its coefficients respectively, ordered somehow.

[3] In the definition of the projective height we assume that at least one coordinate of $\underline{\alpha}$ is non-zero.

Logarithmic discriminant Given an extension $\mathbb{L}/\mathbb{K}$ of number fields, we denote by $\partial_{\mathbb{L}/\mathbb{K}}$ the *normalized logarithmic relative discriminant*:

$$\partial_{\mathbb{L}/\mathbb{K}} = \frac{\log \mathcal{N}_{\mathbb{K}/\mathbb{Q}} \mathcal{D}_{\mathbb{L}/\mathbb{K}}}{[\mathbb{L}:\mathbb{Q}]},$$

where $\mathcal{D}_{\mathbb{L}/\mathbb{K}}$ is the usual relative discriminant and $\mathcal{N}_{\mathbb{K}/\mathbb{Q}}$ is the norm map. The properties of this quantity are summarized in the following proposition.

Proposition 2.1 *1. (additivity in towers) If $\mathbb{K} \subset \mathbb{L} \subset \mathbb{M}$ is a tower of number fields, then $\partial_{\mathbb{M}/\mathbb{K}} = \partial_{\mathbb{L}/\mathbb{K}} + \partial_{\mathbb{M}/\mathbb{L}}$.*

2. (base extension) If $\mathbb{K}'$ is a finite extension of $\mathbb{K}$ and $\mathbb{L}' = \mathbb{L}\mathbb{K}'$ then $\partial_{\mathbb{L}'/\mathbb{K}'} \leq \partial_{\mathbb{L}/\mathbb{K}}$.

3. (triangle inequality) If $\mathbb{L}_1$ and $\mathbb{L}_2$ are two extensions of $\mathbb{K}$, then $\partial_{\mathbb{L}_1\mathbb{L}_2/\mathbb{K}} \leq \partial_{\mathbb{L}_1/\mathbb{K}} + \partial_{\mathbb{L}_2/\mathbb{K}}$.

These properties will be used without special reference.

Height of a set of places Given a number field $\mathbb{K}$ and finite set of places $S \subset M_{\mathbb{K}}$, we define the *absolute logarithmic height* of this set as

$$\mathrm{h}(S) = \frac{\sum_{v \in S} \log \mathcal{N}_{\mathbb{K}/\mathbb{Q}}(v)}{[\mathbb{K}:\mathbb{Q}]},$$

where the norm $\mathcal{N}_{\mathbb{K}/\mathbb{Q}}(v)$ of the place v is the norm of the corresponding prime ideal if v is finite, and is set to be 1 when v is infinite. The properties of this height are summarized in the following proposition.

Proposition 2.2 *1. (field extension) Let $\mathbb{L}$ be an extension of $\mathbb{K}$ and $S_{\mathbb{L}}$ the set of extensions of the places from S to $\mathbb{L}$. Assume that no place from S ramifies in $\mathbb{L}$. Then $\mathrm{h}(S) = \mathrm{h}(S_{\mathbb{L}})$. Without this assumption we have the inequalities $\mathrm{h}(S_{\mathbb{L}}) \leq \mathrm{h}(S) \leq [\mathbb{L}:\mathbb{K}]\mathrm{h}(S_{\mathbb{L}})$.*

2. (denominators and numerators) For $\underline{\alpha} \in \mathbb{K}^N$ let the sets $\mathrm{Den}(\underline{\alpha})$ and $\mathrm{Num}(\underline{\alpha})$ consist of all $v \in M_{\mathbb{K}}$ such that $\|\underline{\alpha}\|_v > 1$, respectively, $\|\underline{\alpha}\|_v < 1$. Then

$$\mathrm{h}\big(\mathrm{Den}_{\mathbb{K}}(\underline{\alpha})\big) \leq \mathrm{h}_{\mathrm{a}}(\underline{\alpha}),$$
$$\mathrm{h}\big(\mathrm{Num}_{\mathbb{K}}(\underline{\alpha})\big) \leq \big(\mathrm{h}_{\mathrm{a}}(\underline{\alpha}) - \mathrm{h}_{\mathrm{p}}(\underline{\alpha})\big) \qquad (\underline{\alpha} \neq \underline{0}).$$

In particular, for $\alpha \in \mathbb{K}^*$ we have $\mathrm{h}\big(\mathrm{Num}_{\mathbb{K}}(\alpha)\big) \leq \mathrm{h}_{\mathrm{a}}(\alpha)$. □

This will also be used without special reference.

Sums over primes We shall systematically use the following estimates from [14]:

$$\sum_{p \leq x} 1 \leq 1.26 \frac{x}{\log x}, \tag{4}$$

$$\sum_{p \leq x} \log p \leq 1.02x. \tag{5}$$

See [14], Corollary 1 of Theorem 2 for (4) and Theorem 9 for (5).

3 Auxiliary Material

In this section we collect miscellaneous facts, mostly elementary and/or well-known, to be used in the article.

3.1 Integral Elements

In this subsection R is an integrally closed integral domain and $\mathbb{K}$ its quotient field.

Lemma 3.1 *Let $\mathbb{L}$ be a finite separable extension of $\mathbb{K}$ of degree n and $\bar{R}$ the integral closure of R in $\mathbb{L}$. Let $\omega_1, \dots, \omega_n \in \bar{R}$ form a base of $\mathbb{L}$ over $\mathbb{K}$. We denote by Δ the discriminant of this basis: $\Delta = \left(\det\left[\sigma_i(\omega_j)\right]_{ij}\right)^2$, where $\sigma_1, \dots, \sigma_n : \mathbb{L} \hookrightarrow \bar{\mathbb{K}}$ are the distinct embeddings of $\mathbb{L}$ into $\bar{\mathbb{K}}$. Then $\bar{R} \subset \Delta^{-1}(R\omega_1 + \cdots + R\omega_n)$.*

Proof This is standard. Write $\beta \in \bar{R}$ as $\beta = a_1\omega_1 + \cdots + a_n\omega_n$ with $a_i \in \mathbb{K}$. Solving the system of linear equations

$$\sigma_i(\beta) = a_1\sigma_i(\omega_1) + \cdots + a_n\sigma_i(\omega_n) \qquad (i = 1, \dots, n)$$

using the Kramer rule, we find that the numbers Δa_i are integral over R. Since R is integrally closed, we have $\Delta a_i \in R$. □

Corollary 3.2 *Let $f(T) = f_0T^n + f_1T^{n-1} + \cdots + f_n \in R[T]$ be a $\mathbb{K}$-irreducible polynomial, and $\alpha \in \bar{\mathbb{K}}$ one of its roots. Let $\bar{R}$ be the integral closure of R in $\mathbb{K}(\alpha)$. Then $\bar{R} \subset \Delta(f)^{-1}R[\alpha]$. where $\Delta(f)$ is the discriminant of f.*

Proof It is well-known that the quantities

$$\omega_1 = 1, \quad \omega_2 = f_0\alpha, \quad \omega_3 = f_0\alpha^2 + f_1\alpha, \quad \dots \qquad \omega_n = f_0\alpha^{n-1} + f_1\alpha^{n-2} + \cdots + f_{n-2}\alpha$$

are integral over R; see, for example, [16, page 183]. Applying Lemma 3.1 to the basis $\omega_1, \dots, \omega_n$, we complete the proof. □

3.2 Local Lemmas

In this subsection $\mathbb{K}$ is a field of characteristic 0 supplied with a discrete valuation v. We denote by $\mathcal{O}_v$ the local ring of v.

The proof of the following lemma is a simple exercise left to the reader.

Lemma 3.3 *Assume that K is complete, and let π be a primitive element of $\mathbb{K}$.*

1. *Let $\alpha \in \mathbb{K}^\times$. For a positive integer e not divisible by the characteristic of the residue field and any choice of the root $\alpha^{1/e}$ the ramification index of $\mathbb{K}(\alpha^{1/e})/\mathbb{K}$ is $e/\gcd(e, \operatorname{ord}_\pi\alpha)$.*
2. *Let $\mathbb{L}_1$ and $\mathbb{L}_2$ be finite extensions of $\mathbb{K}$ (inside some algebraic closure of $\mathbb{K}$) of ramification e_1 and e_2 respectively. Assume that none of e_1, e_2 is divisible by the characteristic of the residue field. Then the ramification of $\mathbb{L}_1\mathbb{L}_2/\mathbb{K}$ is $\operatorname{lcm}(e_1, e_2)$.* □

We say that a polynomial $F(X) \in \mathbb{K}[X]$ is *v-monic* if its leading coefficient is a v-adic unit[4].

Lemma 3.4 *Let $F(X) \in \mathcal{O}_v[X]$ be a v-monic polynomial, and let $\eta \in \bar{\mathbb{K}}$ be a root of F. (We do not assume F to be the minimal polynomial of η over $\mathbb{K}$, because we do not assume it $\mathbb{K}$-irreducible.) Assume that v ramifies in the field $\mathbb{K}(\eta)$. Then $|F'(\eta)|_v < 1$ (for any extension of v to $\mathbb{K}(\eta)$).*

[4] We say that α is a v-adic unit if $|\alpha|_v = 1$.

Proof We may assume that $\mathbb{K}$ is v-complete, and we let $\mathfrak{d} = \mathfrak{d}_{\mathbb{K}(\eta)/\mathbb{K}}$ be the different of the extension $\mathbb{K}(\eta)/\mathbb{K}$. Since v ramifies in $\mathbb{K}(\eta)$, the different is a non-trivial ideal of $\mathcal{O}_v$.

Since η is a root of a v-monic polynomial, it is integral over $\mathcal{O}_v$. Let $G(X) \in \mathcal{O}_v[X]$ be the minimal polynomial of η. Then the different $\mathfrak{d}$ divides $G'(\eta)$, which implies that $|G'(\eta)|_v < 1$.

Write $F(X) = G(X)H(X)$. By the Gauss lemma, $H(X) \in \mathcal{O}_v[X]$. Since $F'(\eta) = G'(\eta)H(\eta)$, we obtain $|F'(\eta)|_v \le |G'(\eta)|_v < 1$, as wanted. □

Given a polynomial $F(X)$ over some field of characteristic 0, we define by $\widehat{F}(X)$ the *radical* of F, that is, the separable polynomial, having the same roots and the same leading coefficient as F:

$$\widehat{F}(X) = f_0 \prod_{F(\alpha)=0} (X - \alpha),$$

where f_0 is the leading coefficient of F and the product runs over the **distinct** roots of F (in an algebraic closure of the base field).

Lemma 3.5 *Assume that $F(X) \in \mathcal{O}_v[X]$. Then the radical $\widehat{F}(X)$ is in $\mathcal{O}_v[X]$ as well. Also, if $|F(\xi)|_v < 1$ for some $\xi \in \mathcal{O}_v$, then we have $|\widehat{F}(\xi)|_v < 1$ as well.*

Proof Let $F(X) = P_1(X)^{\alpha_1} \cdots P_k(X)^{\alpha_k}$ be the irreducible factorization of F in $\mathbb{K}[X]$. The Gauss Lemma implies that we can choose $P_i(X) \in \mathcal{O}_v[X]$ for $i = 1, \dots, k$. Since the characteristic of $\mathbb{K}$ is 0, every P_i is separable. Obviously, the leading coefficient of the separable polynomial $P_1(X) \cdots P_k(X)$ divides that of $F(X)$ in the ring $\mathcal{O}_v$. Hence $\widehat{F}(X) = \gamma P_1(X) \cdots P_k(X)$ with some $\gamma \in \mathcal{O}_v$, which proves the first part of the lemma. The second part is obvious: if $|F(\xi)|_v < 1$ then $|P_i(\xi)|_v < 1$ for some i, which implies $|\widehat{F}(\xi)|_v < 1$. □

Lemma 3.6 *Let $F(X) \in \mathcal{O}_v[X]$ and $\xi \in \mathcal{O}_v$ satisfy $|F(\xi)|_v < 1$ and $|F'(\xi)|_v = 1$. Let $\bar{v}$ be an extension of v to $\bar{\mathbb{K}}$. Then there exists exactly one root $\alpha \in \bar{\mathbb{K}}$ of F such that $|\xi - \alpha|_{\bar{v}} < 1$.*

Proof This is a consequence of Hensel's lemma. Extending $\mathbb{K}$, we may assume that it contains all the roots of F. Hensel's lemma implies that there is exactly one root α in the v-adic completion of $\mathbb{K}$ with the required property. This root must belong to $\mathbb{K}$. □

Lemma 3.7 *Let $F(X), G(X) \in \mathcal{O}_v[X]$ and $\alpha, \xi \in \mathcal{O}_v$ satisfy*

$$F(X) = (X - \alpha)^m G(X), \qquad G(\alpha) \neq 0, \qquad |\xi - \alpha|_v < |G(\alpha)|_v$$

with some non-negative integer m. Expand the rational function $F(X)^{-1}$ into the Laurent series at α. Then this series converges at $X = \xi$.

Proof Substituting $X \mapsto \alpha + X$, we may assume $\alpha = 0$, in which case the statement becomes obvious. □

3.3 Heights

Recall that, for a polynomial f with algebraic coefficients, we denote by $\mathrm{h_p}(f)$ and by $\mathrm{h_a}(f)$, respectively, the projective height and the affine height of the vector of its coefficients ordered somehow. More generally, the height $\mathrm{h_a}(f_1, \dots, f_s)$ of a finite system of polynomials is, by definition, the affine height of the vector formed of all the non-zero coefficients of all these polynomials.

Lemma 3.8 *Let $f_1, \dots, f_s$ be polynomials in $\bar{\mathbb{Q}}[X_1, \dots, X_r]$ and put*

$$N = \max\{\deg f_1, \dots, \deg f_s\}, \qquad h = \mathrm{h_a}(f_1, \dots, f_s).$$

Let also g be a polynomial in $\bar{\mathbb{Q}}[Y_1, \dots, Y_s]$. Then

1. $\mathrm{h_a}\left(\prod_{i=1}^{s} f_i\right) \le \sum_{i=1}^{s} \mathrm{h_a}\left(f_i\right) + \log(r+1)\sum_{i=1}^{s-1} \deg f_i$,
2. $\mathrm{h_p}\left(\prod_{i=1}^{s} f_i\right) \ge \sum_{i=1}^{s} \mathrm{h_p}\left(f_i\right) - \sum_{i=1}^{s} \deg f_i$,
3. $\mathrm{h_a}\big(g\left(f_1,\dots,f_s\right)\big) \le \mathrm{h_a}(g) + \big(h + \log(s+1) + N\log(r+1)\big)\deg g$.

Notice that we use the projective height in item 2, and the affine height in the other items.

Proof Item 2 is the famous Gelfond inequality, see, for instance, Proposition B.7.3 in [11]. The rest is an immediate consequence of Lemma 1.2 from [12]. □

Remark 3.9 If in item 3 we make substitution $Y_i = f_i$ only for a part of the indeterminates Y_i, say, for t of them, where $t \le s$, then we may replace $\log(s+1)$ by $\log(t+1)$, and $\deg g$ by the degree with respect to these indeterminates:

$$\mathrm{h_a}\big(g\left(f_1,\dots,f_t,Y_{t+1},\dots,Y_s\right)\big) \le \mathrm{h_a}(g) + \big(h + \log(t+1) + N\log(r+1)\big)\deg_{Y_1,\dots,Y_t} g.$$

Remark 3.10 When all the f_i are just linear polynomials in one variable, item 2 can be refined as follows: let $F(X)$ be a polynomial of degree ρ, and $\beta_1,\dots,\beta_\rho$ are its roots (counted with multiplicities); then

$$\mathrm{h_a}(\beta_1) + \dots + \mathrm{h_a}(\beta_\rho) \le \mathrm{h_p}(F) + \log(\rho+1).$$

This is a classical result of Mahler, see, for instance, [15, Lemma 3].

Corollary 3.11 *Let f and g be polynomials with algebraic coefficients such that f divides g. Let also a be a non-zero coefficient of f. Then*

1. $\mathrm{h_p}(f) \le \mathrm{h_p}(g) + \deg g$,
2. $\mathrm{h_a}(f) \le \mathrm{h_p}(g) + \mathrm{h_a}(a) + \deg g$.

Proof Item 1 is a direct consequence of item 2 of Lemma 3.8. For item 2 remark that one of the coefficients of f/a is 1, which implies that

$$\mathrm{h_a}(f/a) = \mathrm{h_p}(f/a) = \mathrm{h_p}(f) \le \mathrm{h_p}(g) + \deg g.$$

Since $\mathrm{h_a}(f) \le \mathrm{h_a}(a) + \mathrm{h_a}(f/a)$, the result follows. □

Corollary 3.12 *Let α be an algebraic number and $f \in \bar{\mathbb{Q}}[X,Y]$ be a polynomial with algebraic coefficients, let also $f^{(\alpha)}(X,Y) = f(X+\alpha,Y)$ then*

$$\mathrm{h_a}(f^{(\alpha)}) \le \mathrm{h_a}(f) + m\mathrm{h_a}(\alpha) + 2m\log 2,$$

where $m = \deg_X f$.

Proof This is a direct application of item 3 of Lemma 3.8, together with Remark 3.9. □

In one special case item 3 of Lemma 3.8 can be refined.

Lemma 3.13 *Let*

$$F_{ij}(X) \in \bar{\mathbb{Q}}[X] \qquad (i,j = 1,\dots,s)$$

be polynomials of degree bounded by μ and of affine height bounded by h; then

$$\mathrm{h_a}\left(\det(F_{ij})\right) \le sh + s(\log s + \mu\log 2).$$

For the proof see [12], end of Section 1.1.1.

We also need an estimate for both the affine and the projective height of the Y-resultant $R_f(X)$ of a polynomial $f(X,Y) \in \bar{\mathbb{Q}}[X,Y]$ and its Y-derivative f'_Y, in terms of the affine (respectively, projective) height of f.

Lemma 3.14 *Let $f(X,Y) \in \bar{\mathbb{Q}}[X,Y]$ be of X-degree m and Y-degree n. Then*

$$\mathrm{h_a}(R_f) \le (2n-1)\mathrm{h_a}(f) + (2n-1)\left(\log(2n^2) + m\log 2\right), \tag{6}$$

$$\mathrm{h_p}(R_f) \le (2n-1)\mathrm{h_p}(f) + (2n-1)\log\left((m+1)(n+1)\sqrt{n}\right), \tag{7}$$

Proof Estimate (7) is due to Schmidt [15, Lemma 4]. To prove (6), we invoke Lemma 3.13. Since $R_f(X)$ can be presented as a determinant of dimension $2n-1$, whose entries are polynomials of degree at most m and of affine height at most $\mathrm{h_a}(f)+\log n$, the result follows after an obvious calculation. □

Remark 3.15 Estimate (6) holds true also when $m=0$. We obtain the following statement: the resultant R_f of a polynomial $f(X)$ and its derivative $f'(X)$ satisfies

$$\mathrm{h_a}(R_f) \le (2\deg f - 1)\mathrm{h_a}(f) + (2\deg f - 1)\log\left(2(\deg f)^2\right).$$

3.4 Number fields and Discriminants

We need some estimates for the discriminant of a number field in terms of the heights of its generators. In this subsection $\mathbb{K}$ is a number field, $d=[\mathbb{K}:\mathbb{Q}]$ and $\mathcal{N}(\cdot)=\mathcal{N}_{\mathbb{K}/\mathbb{Q}}(\cdot)$. The following result is due to Silverman [19, Theorem 2].

Lemma 3.16 *Let $\underline{a}=(a_1,\dots,a_k)$ be a point in $\bar{\mathbb{K}}^k$. Put $\nu=[\mathbb{K}(\underline{a}):\mathbb{K}]$. Then*

$$\partial_{\mathbb{K}(\underline{a})/\mathbb{K}} \le 2(\nu-1)\mathrm{h_a}(\underline{a}) + \log\nu.$$ □

This has the following consequence.

Corollary 3.17 *Let $F(X)\in\mathbb{K}[X]$ be a polynomial of degree N. Then*

$$\sum_{F(\alpha)=0} \partial_{\mathbb{K}(\alpha)/\mathbb{K}} \le 2(N-1)\mathrm{h_p}(F) + 3N\log N, \tag{8}$$

the sum being over the roots of F.

Proof Since for any root α we have $[\mathbb{K}(\alpha):\mathbb{K}]\le N$, we estimate the left-hand side of (8) as

$$2(N-1)\sum_{F(\alpha)=0}\mathrm{h_a}(\alpha) + N\log N$$

Remark 3.10 allows us to bound the sum on the right by $\mathrm{h_p}(F)+\log(N+1)$. Now, to complete the proof, just remark that $(N-1)\log(N+1)\le N\log N$. □

We shall also need a bound for the discriminant of a different nature, known as the *Dedekind-Hensel inequality* (see [6, page 397] for historical comments and further references). This inequality gives an estimate of the relative discriminant of a number field extension in terms of the ramified places.

Lemma 3.18 *Let $\mathbb{K}$ be a number field of degree d over $\mathbb{Q}$, and $\mathbb{L}$ an extension of $\mathbb{K}$ of finite degree ν, and let $\mathrm{Ram}(\mathbb{L}/\mathbb{K})$ be the set of places of $\mathbb{K}$ ramified in $\mathbb{L}$. Then*

$$\partial_{\mathbb{L}/\mathbb{K}} \le \frac{\nu-1}{\nu}\mathrm{h}\big(\mathrm{Ram}(\mathbb{L}/\mathbb{K})\big) + 1.26\nu. \tag{9}$$

This is Proposition 4.2.1 from [2] (though the notation in [2] is different, and the quantity estimated therein is $\nu\partial\mathbb{L}/\mathbb{K}$ in our notation), the only difference being that the error term is now explicit. The proof is the same as in [2], but in the very last line one should use the estimate $\sum_{p\le\nu}1 \le 1.26\nu/\log\nu$, which is (4).

A similar estimate was obtained by Serre [17, Proposition 4]. However, (9) is more suitable for our purposes.

It is useful to have an opposite estimate as well. The following lemma is obvious.

Lemma 3.19 *In the set-up of Lemma 3.18 we have* $\mathrm{h}\big(\mathrm{Ram}(\mathbb{L}/\mathbb{K})\big) \le \nu\partial_{\mathbb{L}/\mathbb{K}}$.

This has the following consequence.

Corollary 3.20 *Let $\mathbb{L}_1, \dots, \mathbb{L}_n$ be a family of finite extensions of $\mathbb{K}$ closed under the Galois conjugation over $\mathbb{K}$. Then*

$$\mathrm{h}\left(\bigcup_{i=1}^{n} \mathrm{Ram}(\mathbb{L}_i/\mathbb{K})\right) \le \sum_{i=1}^{n} \partial_{\mathbb{L}_i/\mathbb{K}}.$$

Proof We may assume that the Galois action over K is transitive on $\mathbb{L}_1, \dots, \mathbb{L}_n$ (otherwise, one obtains the estimate for every orbit of the Galois action and then sums the resulting inequalities up). In other words, the fields $\mathbb{L}_1, \dots, \mathbb{L}_n$ form a full system of conjugates over K, which means that

$$[\mathbb{L}_1 : \mathbb{K}] = \ldots = [\mathbb{L}_n : \mathbb{K}] = n, \quad \mathrm{Ram}(\mathbb{L}_1/\mathbb{K}) = \ldots = \mathrm{Ram}(\mathbb{L}_n/\mathbb{K}), \quad \partial_{\mathbb{L}_1/\mathbb{K}} = \ldots = \partial_{\mathbb{L}_n/\mathbb{K}}.$$

Hence

$$\mathrm{h}\left(\bigcup_{i=1}^{n} \mathrm{Ram}(\mathbb{L}_i/\mathbb{K})\right) = \mathrm{h}\big(\mathrm{Ram}(\mathbb{L}_1/\mathbb{K})\big) \le n\partial_{\mathbb{L}_1/\mathbb{K}} = \sum_{i=1}^{n} \partial_{\mathbb{L}_i/\mathbb{K}}. \qquad \square$$

4 Power Series

Our main technical tool is the quantitative Eisenstein theorem, based on the work of Dwork, Robba, Schmidt and van der Poorten [9, 10, 15], in the form presented in [3]. Let

$$y = \sum_{k=-k_0}^{\infty} a_k x^{k/e} \tag{10}$$

be an algebraic power series with coefficients in $\bar{\mathbb{Q}}$, where we assume $k_0 \ge 0$ and $a_{-k_0} \ne 0$ when $k_0 > 0$. The classical Eisenstein theorem tells that the coefficients of this series belong to some number field, that for every valuation v of this field $|a_k|_v$ grows at most exponentially in k, and for all but finitely many v we have $|a_k|_v \le 1$ for all k. We need a quantitative form of this statement, in terms of an algebraic equation $f(x, y) = 0$ satisfied by y.

4.1 Eisenstein Theorem

Thus, let $f(X, Y) \in \mathbb{K}(X, Y)$ be a polynomial over a number field $\mathbb{K}$. We put

$$d = [\mathbb{K} : \mathbb{Q}], \qquad m = \deg_X f, \qquad n = \deg_Y f. \tag{11}$$

Write

$$f(X, Y) = f_0(X)Y^n + f_1(X)Y^{n-1} + \ldots \tag{12}$$

$[\mathbb{L} : \mathbb{K}] \le n$. Finally, for $v \in M_{\mathbb{K}}$ we denote by d_v its local degree over $\mathbb{Q}$, and by $\mathcal{N}v$ its absolute norm:

$$d_v = [\mathbb{K}_v : \mathbb{Q}_v], \qquad \mathcal{N}v = \mathcal{N}_{\mathbb{K}/\mathbb{Q}}(v). \tag{13}$$

With this notation, the height $\mathrm{h}(S)$ of a finite set of places $S \subset M_{\mathbb{K}}$ is given by $d^{-1}\sum_{v\in S} d_v \log \mathcal{N}v$.

The following is Theorem 6.3 from [3].

Theorem 4.1 *Let $\mathbb{K}$ be a number field and $f(X, Y) \in \mathbb{K}(X, Y)$ a separable polynomial. We use notation* (11) *and* (12)*. Let y be an algebraic power series, written as in (10), and satisfying $f(x, y) = 0$. For every $v \in M_{\mathbb{K}}$ there exist real numbers $A_v, B_v \ge 1$, with $A_v = B_v = 1$ for all but finitely many v, such that*

$$d^{-1} \sum_{v\in M_{\mathbb{K}}} d_v \log A_v \le 3n\big(\mathrm{h_p}(f) + \log(mn) + 3e\big), \tag{14}$$

$$d^{-1} \sum_{v\in M_{\mathbb{K}}} d_v \log B_v \le \mathrm{h_p}(f) + 2. \tag{15}$$

and for any extension $\bar v$ of v to $\bar{\mathbb{K}}$ we have

$$|a_k|_{\bar v} \le B_v A_v^{k/e - \lfloor -k_0/e \rfloor} \qquad (k \ge -k_0). \tag{16}$$

Remark 4.2 We shall use this theorem only in the "integral case" $k_0 = 0$, when (16) becomes

$$|a_k|_{\bar v} \le B_v A_v^{k/e} \qquad (k \ge 0), \tag{17}$$

but we prefer to state the theorem in full generality.

We will also use two consequences of this theorem, obtained in [3] as well. To state them, recall that the Puiseux theorem implies existence of $n = \deg_Y f$ distinct series $y_1, \dots, y_n$, which can be written as

$$y_i(x) = \sum_{k=-k_0(i)}^{\infty} a_{ik} x^{k/e_i} \qquad (i = 1, \dots, n), \tag{18}$$

and which satisfy $f\big(x, y_i(x)\big) = 0$.

We denote by $D(X) = D_f(X)$ the Y-discriminant of the polynomial $f(X, Y)$. Given a polynomial $P(X)$, we denote by $\operatorname{ord}_\alpha P(X)$ the order of α as the root of $P(X)$.

The following proposition is composed from Theorems 6.4 and 8.5 from [3].

Proposition 4.3 *Let $f(X, Y) \in \mathbb{K}[X, Y]$ be as above and let $y_1, \dots, y_n$ be the n distinct series, written as in* (18) *and satisfying $f\big(x, y_i(x)\big) = 0$.*

1. *Let T be the (finite) set of $v \in M_{\mathbb{K}}$ such that $|a_{ik}|_{\bar v} > 1$ for some coefficient a_{ik} and some extension $\bar v$ of v to $\bar{\mathbb{K}}$. Then*

$$\mathrm{h}(T) \le 3n\big(\mathrm{h}_{\mathrm{p}}(f) + \log(mn) + 1\big). \tag{19}$$

2. *The number fields $\mathbb{L}_1, \dots, \mathbb{L}_n$, generated over $\mathbb{K}$ by the coefficients of $y_1, \dots, y_n$, respectively, satisfy*

$$\sum_{i=1}^{n} \partial_{\mathbb{L}_i/\mathbb{K}} \le 8n\big(\operatorname{ord}_0 D(X) + 1\big)\big(\mathrm{h}_{\mathrm{p}}(f) + 5n + \log m\big). \tag{20}$$

4.2 The "Essential" Coefficients

Let $y \in \bar{\mathbb{Q}}((x^{1/e}))$ be a an algebraic power series written as in (10). We assume that e is smallest possible: $y \notin \bar{\mathbb{Q}}((x^{1/e'}))$ for $e' < e$.

We define the k-th ramification index $\epsilon_k = \epsilon_k(y)$ as the smallest natural e' such that the k-th partial sum $y^{(k)} = \sum_{\ell=-k_0}^{k} a_\ell x^{\ell/e}$ belongs to $\bar{\mathbb{Q}}((x^{1/e'}))$. By the definition,

$$\epsilon_{-k_0} = 1, \qquad \epsilon_k \mid \epsilon_{k+1},$$

and since e is smallest possible, we have $\epsilon_k = e$ for all sufficiently large k.

We call an index $k > -k_0$ *essential* if $\epsilon_k > \epsilon_{k-1}$ (that is, we "gain new ramification" with the term $a_k x^{k/e}$). The corresponding coefficient a_k is called an *essential coefficient*. Clearly, an essential coefficient cannot be 0.

The series y can have only finitely many essential indices. We want to estimate the sum of the heights of the essential coefficients. We denote by ord_0 the discrete valuation on the local ring $\bar{\mathbb{Q}}[[x^{1/e}]]$ normalized to have $\operatorname{ord}_0(x) = 1$.

Proposition 4.4 *Let $f(X, Y) \in \bar{\mathbb{Q}}(X, Y)$ be a separable polynomial. We use notation* (11) *and* (12). *Let y be an algebraic power series satisfying $f(x, y) = 0$. Assume that $f_0(0) \ne 0$. Then*

$$\sum_{k \text{ essential}} \mathrm{h}(a_k) \le \big(\mathrm{h}_{\mathrm{p}}(f) + 2\big)\log_2 e + 3n\big(\mathrm{h}_{\mathrm{p}}(f) + \log(mn) + 3e\big)\operatorname{ord}_0\big(f'_Y(x, y)\big) \tag{21}$$

If $f_0(0) \neq 0$ then the the series $y(x)$ is integral over the ring $\bar{\mathbb{Q}}[[x]]$ and can be written as

$$y(x) = \sum_{k=0}^{\infty} a_k x^{k/e} \tag{22}$$

The assumption $f_0(0) \neq 0$ is purely technical; a similar result holds in general as well. However, without this assumption estimate (21) gets weaker than we need, while assuming $f_0(0) \neq 0$ does not hurt generality: see Section 7.

The proof of Proposition 4.4 relies on the following lemma (which is an analog of Lemma 7.2 in [3]).

Lemma 4.5 *Assume that $f_0(0) \neq 0$. Then there is at most $\log_2 e$ essential indices, and their sum does not exceed $e \operatorname{ord}_0\bigl(f'_Y(x,y)\bigr)$.*

Proof Since $\epsilon_{k-1} \mid \epsilon_k$, we have $\epsilon_k \geq 2\epsilon_{k-1}$ whenever k is essential, which means that there can be at most $\log_2 e$ essential indices.

Now let us prove the statement about the sum. Together with the series y we consider the "twisted series"

$$\sum_{k=0}^{\infty} a_k \zeta^{(j-1)k} x^{k/e} \in \bar{\mathbb{Q}}[[x^{1/e}]] \qquad (j = 1, \dots, e),$$

where ζ is a primitive e-th root of unity. These e series are among the n distinct series $y_1, \dots, y_n$, which satisfy $f(x, y_i) = 0$, and after re-numbering we may assume that

$$y_j = \sum_{k=0}^{\infty} a_k \zeta^{(j-1)k} x^{k/e} \in \bar{\mathbb{Q}}[[x^{1/e}]] \qquad (j = 1, \dots, e).$$

In particular, $y = y_1$.

By the definition of ϵ_k we have $y_j^{(k)} = y_{j'}^{(k)}$ if and only if $j \equiv j' \mod \epsilon_k$. In particular, $y_j^{(k)} = y^{(k)}$ if and only if $\epsilon_k \mid (j-1)$. We partition the set $J = \{2, 3, \dots, e\}$ as

$$J = J_1 \cup J_2 \cup J_3 \dots, \qquad J_k \cap J_\ell = \varnothing \quad (k \neq \ell)$$

where

$$J_k = \bigl\{ j \in J : \epsilon_{k-1} \mid (j-1),\ \epsilon_k \nmid (j-1) \bigr\}.$$

The following two observations are now crucial:

- for $j \in J$ we have $\operatorname{ord}_0(y - y_j) = k/e$ if and only if $j \in J_k$;
- the set J_k is not empty if and only if k is an essential index for y.

Using this, we find

$$\sum_{k \text{ essential}} \frac{k}{e} \leq \sum_{k=0}^{\infty} \frac{k}{e} |J_k| = \operatorname{ord}_0 \left(\prod_{j=2}^{e} (y - y_j) \right) \tag{23}$$

Since $f_n(0) \neq 0$, all the series $y_1, \dots, y_n$ are integral over $\bar{\mathbb{Q}}[[x]]$. Hence the product in the right-hand side of (23) divides $f'_Y(x,y) = f_0(x) \prod_{j=2}^{n} (y - y_j)$. It follows that the right-hand side of (23) does not exceed $\operatorname{ord}_0\bigl(f'_Y(x,y)\bigr)$, which proves the lemma. $\square$

Proof of Proposition 4.4 By Theorem 4.1 we have

$$\mathrm{h}(a_k) \le \mathrm{h_p}(f) + 2 + \frac{k}{e} \cdot 3n\big(\mathrm{h_p}(f) + \log(mn) + 3e\big).$$

Hence

$$\sum_{k \text{ essential}} \mathrm{h}(a_k) \le \big(\mathrm{h_p}(f) + 2\big) \sum_{k \text{ essential}} 1 + 3n\big(\mathrm{h_p}(f) + \log(mn) + 3e\big) \sum_{k \text{ essential}} \frac{k}{e}.$$

We conclude, applying the lemma. □

Now assume that $\mathbb{K}$ is a number field and y a series with coefficients in $\mathbb{K}$. We denote by $\mathrm{Ess}\,(y)$ the set of places $v \in M_{\mathbb{K}}$ such that $|a_k|_v < 1$ for some essential coefficient a_k of y:

$$\mathrm{Ess}\,(y) = \{v \in M_K : \text{ there exists an essential index } k \text{ such that } |a_k|_v < 1\}$$

Proposition 4.6 *Let $f(X,Y) \in \mathbb{K}[X,Y]$ be a separable polynomial. We use notation* (11) *and* (12)*. Assume that $f_0(0) \neq 0$. Let $y_1, \ldots, y_n$ be the n distinct series, satisfying $f\big(x, y_i(x)\big) = 0$. Assume that the coefficients of all these series belong to $\mathbb{K}$. Then*

$$\mathrm{h}\left(\bigcup_{i=1}^{n} \mathrm{Ess}\,(y_i)\right) \le n\big(\mathrm{h_p}(f) + 2\big) + 3n\big(\mathrm{h_p}(f) + 4n + \log m\big)\mathrm{ord}_0 D(X). \tag{24}$$

where $D(X)$ is the Y-discriminant of $f(X,Y)$.

Proof Recall that the series y_i has e_i "twists" among $y_1, \ldots, y_n$, as defined in the proof of Lemma 4.5. If y_j is a twist of y_i then each coefficient of y_j is equal to the corresponding coefficient of y_i times an e_i-th root of unity, which implies that $\mathrm{Ess}\,(y_i) = \mathrm{Ess}\,(y_j)$.

Select a maximal subset from $\{y_1, \ldots, y_n\}$ such that none of its elements is a twist of the other. After re-numbering, we may assume that this subset is $\{y_1, \ldots, y_s\}$ (this is NOT the numbering adopted in the proof of Lemma 4.5). Then each of $y_1, \ldots, y_n$ is a twist of one of $y_1, \ldots, y_s$, which implies that $\bigcup_{i=1}^{n} \mathrm{Ess}\,(y_i) = \bigcup_{i=1}^{s} \mathrm{Ess}\,(y_i)$ and $e_1 + \ldots + e_s = n$.

Item 2 of Proposition 2.2 implies that $\mathrm{h}\big(\mathrm{Ess}\,(y_i)\big)$ is bounded by the sum of the heights of the essential coefficients of y_i. Now, using Proposition 4.4 we obtain

$$\begin{aligned}
\mathrm{h}\left(\bigcup_{i=1}^{n} \mathrm{Ess}\,(y_i)\right) &= \mathrm{h}\left(\bigcup_{i=1}^{s} \mathrm{Ess}\,(y_i)\right) \\
&\le \sum_{i=1}^{s} \Big(\big(\mathrm{h_p}(f) + 2\big)\log_2 e_i + 3n\big(\mathrm{h_p}(f) + \log(mn) + 3e_i\big)\mathrm{ord}_0\big(f'_Y(x, y_i)\big)\Big) \\
&\le \big(\mathrm{h_p}(f) + 2\big)\sum_{i=1}^{s} \log_2 e_i + 3n\big(\mathrm{h_p}(f) + 4n + \log m\big) \sum_{i=1}^{s} \mathrm{ord}_0\big(f'_Y(x, y_i)\big) \\
&\le \big(\mathrm{h_p}(f) + 2\big)\sum_{i=1}^{s} e_i + 3n\big(\mathrm{h_p}(f) + 4n + \log m\big) \sum_{i=1}^{n} \mathrm{ord}_0\big(f'_Y(x, y_i)\big) \\
&= n\big(\mathrm{h_p}(f) + 2\big) + 3n\big(\mathrm{h_p}(f) + 4n + \log m\big)\mathrm{ord}_0 D(X),
\end{aligned}$$

as wanted. □

5 Proximity and Ramification

This section is the technical heart of the article. We consider a covering $\mathcal{C} \xrightarrow{x} \mathbb{P}^1$, defined over a number field $\mathbb{K}$, and call a point $P \in \mathcal{C}(\bar{\mathbb{K}})$ *semi-defined* over $\mathbb{K}$ if $x(P) \in \mathbb{P}^1(\mathbb{K})$. We define a finite set $\mathcal{Q}$ of points from $\mathcal{C}(\bar{\mathbb{K}})$ (which include the finite ramified points of the covering x, but may contain some other points as well) and prove two statements (Propositions 5.2 and 5.3 below) which, informally, assert the following.

- If a finite place $v \in M_{\mathbb{K}}$ ramifies in the field $\mathbb{K}(P)$ (where P is semi-defined over $\mathbb{K}$) then (unless v is "bad" in certain sense) the point P must be "v-adically close" to a point from the set $\mathcal{Q}$ (Propositions 5.2).
- Given a point Q on $\mathcal{C}$ and a finite place v (again, it should not be "bad" in some sense), for the points P (semi-defined over $\mathbb{K}$) in a "v-adic neighborhood" of Q, the v-ramification in the field $\mathbb{K}(P)$ is determined by the "v-adic distance" between P and Q and the ramification of the point Q over $\mathbb{P}^1$. Roughly speaking, "geometric ramification determines arithmetic ramification" (Propositions 5.3).

It is not difficult to make qualitative statements of this kind, but it is a rather delicate task to make everything explicit. In particular, we will explicitly estimate (Proposition 5.4) the set of the "bad" places.

5.1 Proximity

Now let us be precise. In this section we fix, once and for all:

- a number field $\mathbb{K}$;
- an absolutely irreducible smooth projective curve $\mathcal{C}$ defined over $\mathbb{K}$;
- a non-constant rational function $x \in \mathbb{K}(\mathcal{C})$;
- one more rational function $y \in \mathbb{K}(\mathcal{C})$ such that $\mathbb{K}(\mathcal{C}) = \mathbb{K}(x, y)$ (existence of such y follows from the primitive element theorem).

Let $f(X, Y) \in \mathbb{K}[X, Y]$ be the $\mathbb{K}$-irreducible polynomial such that $f(x, y) = 0$ (it is well-defined up to a constant factor). Since $\mathcal{C}$ is absolutely irreducible, so is the polynomial $f(X, Y)$.

We put $m = \deg_X f$, $n = \deg_Y f$, and write

$$f(X, Y) = f_0(X)Y^n + f_1(X)Y^{n-1} + \cdots + f_n(X). \tag{25}$$

Let $Q \in \mathcal{C}(\bar{\mathbb{K}})$ be a finite $\bar{\mathbb{K}}$-point of $\mathcal{C}$ ("finite" means that Q is not a pole of x). We set $\alpha = x(Q)$ and we denote by e_Q the ramification index of x at Q (that is, $e_Q = \mathrm{ord}_Q(x - \alpha)$). When it does not cause a confusion we write e instead of e_Q. Fix a primitive e-th root of unity $\zeta = \zeta_e$. Then there exist e equivalent Puiseux expansions of y at Q:

$$y_j^{(Q)} = \sum_{k=-k^{(Q)}}^{\infty} a_k^{(Q)} \zeta^{(j-1)k} (x - \alpha)^{k/e} \qquad (j = 1, \ldots, e), \tag{26}$$

where $k^{(Q)} = \max\{0, -\mathrm{ord}_Q(y)\}$.

Let $\bar{v}$ be a place of $\bar{\mathbb{K}}$. We say that the series (26) converge $\bar{v}$-adically at $\xi \in \bar{\mathbb{K}}$, if, for a fixed e-th root $\sqrt[e]{\xi - \alpha}$, the e numerical series

$$\sum_{k=-k^{(Q)}}^{\infty} a_k^{(Q)} \left(\zeta^{j-1} \sqrt[e]{\xi - \alpha}\right)^k \qquad (j = 1, \ldots, e)$$

converge in the $\bar{v}$-adic topology. We denote by $y_j^{(Q)}(\xi)$, with $j = 1, \ldots, e$, the corresponding sums. While the individual sums depend on the particular choice of the root $\sqrt[e]{\xi - \alpha}$, the very fact of convergence, as well as the set $\left\{y_1^{(Q)}(\xi), \ldots, y_e^{(Q)}(\xi)\right\}$ of the sums, are independent of the choice of the root.

Now we are ready to introduce the principal notion of this section, that is of *proximity of a point to a different point* with respect to a given place $\bar{v} \in M_{\bar{\mathbb{K}}}$.

Definition 5.1 *Let $P \in \mathcal{C}(\bar{\mathbb{K}})$ be a finite $\bar{\mathbb{K}}$-point of $\mathcal{C}$, and put $\xi = x(P)$. We say that P is $\bar{v}$-**adically close** to Q if the following conditions are satisfied:*

- $|\xi - \alpha|_{\bar{v}} < 1$;
- *the e series (26) $\bar{v}$-adically converge at ξ, and one of the sums $y_j^{(Q)}(\xi)$ is equal to $y(P)$.*

An important warning: the notion of proximity just introduced is not symmetric in P and Q: the proximity of P to Q does not imply, in general, the proximity of Q to P. Intuitively, one should think of Q as a "constant" point, and of P as a "variable" point.

To state the main results of this section, we have to define a finite set $\mathcal{Q}$ of $\bar{\mathbb{K}}$-points of the curve $\mathcal{C}$, and certain finite sets of "bad" places of the field $\mathbb{K}$. Let $R(X) = R_f(X) \in \mathbb{K}[X]$ be the Y-resultant of $f(X,Y)$ and $f'_Y(X,Y)$, and let $\mathcal{A}$ be the set of the roots of $R(X)$:

$$\mathcal{A} = \{\alpha \in \bar{\mathbb{K}} : R(\alpha) = 0\}.$$

We define $\mathcal{Q}$ as follows:

$$\mathcal{Q} = \left\{Q \in \mathcal{C}(\bar{\mathbb{K}}) : x(Q) \in \mathcal{A}\right\}.$$

It is important to notice that $\mathcal{Q}$ contains all the finite ramification points of x (and may contain some other points as well). Also, the set $\mathcal{Q}$ is Galois-invariant over $\mathbb{K}$: every point belongs to it together with its Galois orbit over $\mathbb{K}$.

Now let us define the finite sets of "bad" places of $\mathbb{K}$ mentioned above. First of all we assume (as we may, without loss of generality) that

$$\text{the polynomial } f_0(X), \text{ defined in (25), is monic.} \tag{27}$$

In particular, f has a coefficient equal to 1, which implies equality of the affine and the projective heights of f:

$$\mathrm{h_a}(f) = \mathrm{h_p}(f). \tag{28}$$

Now, we define

$$T_1 = \left\{v \in M_{\mathbb{K}}^0 : \text{the prime below } v \text{ is} \le n\right\},$$
$$T_2 = \left\{v \in M_{\mathbb{K}}^0 : |f|_v > 1\right\}.$$

Further, let r_0 be the leading coefficient of $R(X)$. We define

$$T_3 = \left\{v \in M_{\mathbb{K}}^0 : |r_0|_v < 1\right\}.$$

Next, we let Δ be the resultant of $\widehat{R}(X)$ and $\widehat{R}'(X)$, where $\widehat{R}$ is the radical of R, see Subsection 3.2. Since the polynomial $\widehat{R}(X)$ is separable, we have $\Delta \in \mathbb{K}^*$. Now we define the set T_4 as follows:

$$T_4 = \left\{v \in M_{\mathbb{K}}^0 : |\Delta|_v < 1\right\}.$$

The sets T_5 and T_6 will be defined under the assumptions

$$\mathcal{Q} \subset \mathcal{C}(\mathbb{K}), \tag{29}$$
$$\mathbb{K} \text{ contains } e_Q\text{-th roots of unity for all } Q \in \mathcal{Q}. \tag{30}$$

Notice that (29) implies that

$$\mathcal{A} \subset \mathbb{K}. \tag{31}$$

Now fix $Q \in \mathcal{C}(\mathbb{K})$ and define the sets $T_5^{(Q)}$ and $T_6^{(Q)}$ using the Puiseux expansions of y at $Q \in \mathcal{Q}$. As in (26), we denote by $a_k^{(Q)}$ the coefficients of these expansions; by (30) we may assume that these coefficients are in $\mathbb{K}$. Now define

$$T_5^{(Q)} = \left\{v \in M_{\mathbb{K}}^0 : \left|a_k^{(Q)}\right|_v > 1 \text{ for some } k\right\}, \qquad T_5 = \bigcup_{Q \in \mathcal{Q}} T_5^{(Q)}.$$

The Eisenstein theorem implies that the set $T_5^{(Q)}$ is finite.

Finally, put

$$T_6^{(Q)} = \mathrm{Ess}\,(y_1^{(Q)}), \qquad T_6 = \bigcup_{Q\in\mathcal{Q}} T_6^{(Q)}.$$

where $\mathrm{Ess}\,(y)$ is defined in Subsection 4.2 (just before Proposition 4.6) as the set of places $v \in M_{\mathbb{K}}$ such that $|a_k|_{\bar v} < 1$ for some essential coefficient a_k.

For $P, Q \in \mathcal{C}(\bar{\mathbb{K}})$ and a finite place $\bar v \in M_{\bar{\mathbb{K}}}$ we let v be the restriction of $\bar v$ to $\mathbb{K}$ and π a primitive element of the local ring $\mathcal{O}_v$. Define

$$\ell(P, Q, \bar v) = \frac{\log|\xi - \alpha|_{\bar v}}{\log|\pi|_v} = \mathrm{ord}_\pi(\xi - \alpha), \tag{32}$$

where, as above, $\xi = x(P)$ and $\alpha = x(Q)$.

Now we are ready to state the principal results of this section. Recall that a point $P \in \mathcal{C}(\bar{\mathbb{K}})$ is *semi-defined over* $\mathbb{K}$ if $\xi = x(P) \in \mathbb{P}^1(\mathbb{K})$. We also call a point P *finite* if it is not a pole of x.

Proposition 5.2 *Let $\mathcal{Q}$ be the set defined above, and assume* (29), (30). *Then for any finite point $P \in \mathcal{C}(\bar{\mathbb{K}}) \smallsetminus \mathcal{Q}$ semi-defined over $\mathbb{K}$, and for any finite place $v \in M_{\mathbb{K}}$, at least one of the following conditions is satisfied (we again put $\xi = x(P)$).*

- $|\xi|_v > 1$.
- $v \in T_2 \cup T_3 \cup T_4 \cup T_5$.
- *v is not ramified in the field $\mathbb{K}(P)$.*
- *For any $\bar v \in M_{\bar{\mathbb{K}}}$, extending v, our point P is $\bar v$-adically close to some $Q \in \mathcal{Q}$, which depends only on v, but not on the particular extension $\bar v$. In particular, the integers e_Q and $\ell(P, Q, \bar v)$ are independent of the particular choice of $\bar v$.*

Proposition 5.3 *Let $P \in \mathcal{C}(\bar{\mathbb{K}})$ be a finite point semi-defined over $\mathbb{K}$, and assume that P is $\bar v$-adically close to some $Q \in \mathcal{C}(\mathbb{K})$ for some finite place $\bar v \in M_{\bar{\mathbb{K}}}$. Let v and w be the restrictions of $\bar v$ to $\mathbb{K}$ and $\mathbb{K}(P)$, respectively. Assume that v does not belong to $T_1 \cup T_5^{(Q)} \cup T_6^{(Q)}$, and that $\mathbb{K}$ contains the e_Q-th roots of unity. Then the ramification index of w over v is equal to $e_Q/(\gcd(e_Q, \ell)$, where $\ell = \ell(P, Q, \bar v)$ is defined in (32).*

Intuitively, the last condition means that the "arithmetic ramification is determined by the geometric ramification".

Proposition 5.4 *Assume* (29) *and* (30). *Put $T = T_1 \cup T_2 \cup \ldots \cup T_6$. Assume further that*

$$\text{the covering } \mathcal{C} \xrightarrow{x} \mathbb{P}^1 \text{ does not ramify over the roots of } f_0(X). \tag{33}$$

Then

$$\mathrm{h}(T) \le 52mn^2\big(\mathrm{h_p}(f) + 2m + 2n\big).$$

Finally, if we do *not* assume (29) and (30), then we have to estimate the smallest extension of K satisfying (29) and (30).

Proposition 5.5 *Let $\mathbb{L}$ be the compositum of the fields $\mathbb{K}(Q)$ and the fields generated over $\mathbb{K}$ by e_Q-th roots of unity, for all $Q \in \mathcal{Q}$. Then*

$$\partial_{\mathbb{L}/\mathbb{K}} \le 105mn^2\big(\mathrm{h_p}(f) + 2m + 2n\big). \tag{34}$$

5.2 Proof of Proposition 5.2

We fix, once and for all, a finite place $v \in M_{\mathbb{K}}$, its extension $\bar{v} \in M_{\bar{\mathbb{K}}}$, and a point $P \in \mathcal{C}(\bar{\mathbb{K}})$ semi-defined over $\mathbb{K}$ and such that $\xi = x(P) \notin \mathcal{A}$. We assume that $|\xi|_v \le 1$, that $v \notin T_2 \cup \ldots \cup T_5$ and that v is ramified in $\mathbb{K}(P)$, and we shall prove that P is $\bar{v}$-adically close to a unique $Q \in \mathcal{Q}$, and that the numbers e_Q and $\ell(P, Q, \bar{v})$ are independent of the selected $\bar{v}$.

Since $v \notin T_2 \cup T_3$, the polynomial $R(X)$ belongs to $\mathcal{O}_v[X]$ and is v-monic. Lemma 3.5 implies that so is its radical $\widehat{R}(X)$. Also, every root α of R is a v-adic integer.

Put $\eta = y(P)$. Since $\xi \notin \mathcal{A}$, the point (ξ, η) of the plane curve $f(X, Y) = 0$ is non-singular, which implies that $\mathbb{K}(P) = \mathbb{K}(\xi, \eta) = \mathbb{K}(\eta)$ (recall that $\xi \in \mathbb{K}$). Now Lemma 3.4 implies that $|f'_Y(\xi, \eta)|_{\bar{v}} < 1$. It follows that $|R(\xi)|_v < 1$, which implies that $|\widehat{R}(\xi)|_v < 1$ by Lemma 3.5.

Next, since $v \notin T_4$, we have $|\widehat{R}'(\xi)|_v = 1$. Lemma 3.6 implies now that there exists a unique $\alpha \in \mathcal{A}$ such that $|\xi - \alpha|_{\bar{v}} < 1$. Since $\alpha \in \mathbb{K}$ by (31), $|\xi - \alpha|_{\bar{v}} = |\xi - \alpha|_v$ depends only on v, but not on the extension $\bar{v}$; hence so does the quantity $\ell(P, Q, \bar{v})$.

Fix this α from now on. There is $\sum_{x(Q)=\alpha} e_Q = n$ Puiseux expansions of y at the points Q above α, and they satisfy

$$f(x, Y) = f_0(x) \prod_{x(Q)=\alpha} \prod_{j=1}^{e_Q} \left(Y - y_j^{(Q)} \right).$$

Since $v \notin T_5$, each of the series $y_j^{(Q)}$ has v-adic convergence radius at least 1. Since $|\xi - \alpha|_{\bar{v}} < 1$, all them $\bar{v}$-adically converge at ξ. Moreover, the convergence is absolute, because $\bar{v}$ is non-archimedean. Hence

$$f(\xi, Y) = f_0(\xi) \prod_{x(Q)=\alpha} \prod_{j=1}^{e_Q} \left(Y - y_j^{(Q)}(\xi) \right).$$

Since $R(\xi) \neq 0$, we have $f_0(\xi) \neq 0$ as well. Hence we have on the left and on the right polynomials of degree n in Y, the polynomial on the left having $\eta = y(P)$ as a simple root (here we again use that $R(\xi) \neq 0$). Hence exactly one of the sums $y_j^{(Q)}(\xi)$ is equal to η. We have proved that P is $\bar{v}$-adically close to exactly one $Q \in \mathcal{Q}$. □

5.3 Proof of Proposition 5.3

We may assume, by re-defining the root $\sqrt[e]{\xi - \alpha}$ that $\eta = y(P)$ is the sum of $y_1^{(Q)}$ at ξ. In the sequel we omit reference to Q (when it does not lead to confusion) and write e for e_Q, a_k for $a_k^{(Q)}$, etc. Thus, we have, in the sense of $\bar{v}$-adic convergence,

$$\eta = \sum_{k=-k^{(Q)}}^{\infty} a_k \left(\sqrt[e]{\xi - \alpha} \right)^k. \tag{35}$$

Let v and w be the restrictions of $\bar{v}$ to $\mathbb{K}$ and $\mathbb{K}(P)$, respectively. We assume that v does not belong to $T_1 \cup T_5^{(Q)} \cup T_6^{(Q)}$. Put

$$e' = \frac{e}{\gcd(e, \ell)}, \qquad \ell' = \frac{\ell}{\gcd(e, \ell)},$$

where $e = e_Q$ and $\ell = \ell(P, Q, \bar{v})$ is defined in (32). We have to show that the ramification index of w over v is equal to e'.

Recall that by the assumption $Q \in \mathcal{C}(\mathbb{K})$ and $\mathbb{K}$ contains e-th roots of unity. It follows that $\alpha = x(Q) \in \mathbb{K}$ and that $\mathbb{K}$ contains the coefficients of the Puiseux expansions of y at Q.

Let $\mathbb{K}_v$ be a v-adic completion of $\mathbb{K}$. We consider $\bar{\mathbb{K}}_{\bar{v}}$ as its algebraic closure, and the fields $\mathbb{K}_v(P) = \mathbb{K}_v(\eta)$ and $\mathbb{K}_v\left(\sqrt[e]{\xi - \alpha}\right)$ as subfields of the latter. According to (35), we have $\mathbb{K}_v(\eta) \subset \mathbb{K}_v\left(\sqrt[e]{\xi - \alpha}\right)$. The latter field has ramification e' over $\mathbb{K}_v$ by item 1 of Lemma 3.3. (The assumption $v \notin T_1$ implies that e is not divisible by the characteristic of the residue field.)

Assume that the ramification of $\mathbb{K}_v(\eta)/\mathbb{K}_v$ is not e'. Then there exists a prime divisor q of e' such that the ramification index of $\mathbb{K}_v(\eta)/\mathbb{K}_v$ divides e'/q. We want to show that this is impossible.

Let κ be the smallest k with the properties $a_k \neq 0$ and $q \nmid \kappa$. Then κ is an *essential index* of the series y_1 as defined in Subsection 4.2, and a_κ is an essential coefficient. Put

$$\theta = \eta - \sum_{k=k^{(Q)}}^{\kappa-1} a_k \left(\sqrt[e]{\xi-\alpha}\right)^k = a_\kappa \left(\sqrt[e]{\xi-\alpha}\right)^\kappa + \sum_{k=\kappa+1}^{\infty} a_k \left(\sqrt[e]{\xi-\alpha}\right)^k.$$

By the definition of κ, we have $\theta \in \mathbb{K}_v\left(\eta, \sqrt[e/q]{\xi-\alpha}\right)$. The ramification of $\mathbb{K}_v\left(\sqrt[e/q]{\xi-\alpha}\right)/\mathbb{K}_v$ is $(e/q)/\gcd(e/q,\ell)$ (we again use item 1 of Lemma 3.3). Since q divides e', it cannot divide ℓ', and we have $\gcd(e/q,\ell) = \gcd(e,\ell)$, which implies that $(e/q)/\gcd(e/q,\ell) = e'/q$.

Thus, the ramification of $\mathbb{K}_v\left(\sqrt[e/q]{\xi-\alpha}\right)/\mathbb{K}_v$ is e'/q, and the ramification of $\mathbb{K}_v(\eta)/\mathbb{K}_v$ divides e'/q. Item 2 of Lemma 3.3 now implies that the ramification of $\mathbb{K}_v\left(\eta, \sqrt[e/q]{\xi-\alpha}\right)/\mathbb{K}_v$ is e'/q. Hence the ramification of $\mathbb{K}_v(\theta)/\mathbb{K}_v$ divides e'/q, which implies that $\operatorname{ord}_\pi\theta \in (q/e')\mathbb{Z}$.

But, since $v \notin T_5^{(Q)} \cup T_6^{(Q)}$, we have $|a_k|_v \le 1$ for all k and $|a_\kappa|_v = 1$, which implies that $|\theta|_v = \left|\left(\sqrt[e]{\xi-\alpha}\right)^\kappa\right|_v$. It follows that

$$\operatorname{ord}_\pi\theta = \frac{\kappa}{e}\operatorname{ord}_\pi(\xi-\alpha) = \frac{\kappa\ell}{e} = \frac{\kappa\ell'}{e'}.$$

We have proved that $\kappa\ell/e' \in (q/e')\mathbb{Z}$. But q does not divide any of the numbers κ and ℓ', a contradiction. □

5.4 Proof of Proposition 5.4

The proposition is a direct consequence of the estimates

$$\mathrm{h}(T_1) \le 1.02n, \tag{36}$$
$$\mathrm{h}(T_2) \le \mathrm{h_p}(f), \tag{37}$$
$$\mathrm{h}(T_3) \le (2n-1)\big(\mathrm{h_p}(f) + m\log 2 + \log(2n^2)\big), \tag{38}$$
$$\mathrm{h}(T_4) \le 16mn^2\big(\mathrm{h_p}(f) + 2m + 2\log n\big), \tag{39}$$
$$\mathrm{h}(T_5) \le 14mn^2\big(\mathrm{h_p}(f) + 2m + 2n\big), \tag{40}$$
$$\mathrm{h}\big(T_6\big) \le 18mn^2\big(\mathrm{h_p}(f) + 2m + \log n\big). \tag{41}$$

Remark 5.6 Assumption (33) is used only in the proof of (41).

Proof of (36) Obviously, $\mathrm{h}(T_1) \le \sum_{p\le n}\log p$, which is bounded by $1.02n$ according to (5). □

Proof of (37) Item 2 of Proposition 2.2 implies that $\mathrm{h}(T_2) \le \mathrm{h_a}(f)$. Since $\mathrm{h_a}(f) = \mathrm{h_p}(f)$ by (28), the result follows. □

Proof of (38) Item 2 of Proposition 2.2 and Lemma 3.14 imply that

$$\mathrm{h}(T_3) \le \mathrm{h_a}(r_0) \le \mathrm{h_a}(R) \le (2n-1)\mathrm{h_a}(f) + (2n-1)\left(\log(2n^2) + m\log 2\right). \tag{42}$$

Again using (28), we have the result. □

Proof of (39) We have $\deg\widehat{R} \le \deg R \le (2n-1)m$. Further, using Corollary 3.11 and inequalities (42), we find

$$\mathrm{h_a}(\widehat{R}) \le \mathrm{h_p}(R) + \mathrm{h_a}(r_0) + \deg R \le (4n-2)\mathrm{h_a}(f) + (8n-4)\left(\log n + m\right).$$

Finally, using Remark 3.15 and the previous estimates, we obtain

$$\mathrm{h}(T_4) \le \mathrm{h_a}(\Delta) \le (2\deg\widehat{R} - 1)\left(\mathrm{h_a}(\widehat{R}) + \log(2(\deg\widehat{R})^2)\right) \le 16mn^2\mathrm{h_a}(f) + 32mn^2\left(\log n + m\right).$$

Using (28), we obtain the result. □

Preparation for the proofs of (40) and (41) Recall that we denote by $R(X)$ the Y-resultant of $f(X,Y)$ and $f'_Y(X,Y)$ and by $\mathcal{A}$ the set of the roots of $R(X)$. Then

$$|\mathcal{A}| \le \sum_{\alpha\in\mathcal{A}} \operatorname{ord}_\alpha R(X) \le \deg R(X) \le m(2n-1), \tag{43}$$

$$\sum_{\alpha\in\mathcal{A}} \mathrm{h_a}(\alpha) \le \sum_{\alpha\in\mathcal{A}} \operatorname{ord}_\alpha R(X)\mathrm{h_a}(\alpha) \le \mathrm{h_p}(R) + \log(2mn) \le (2n-1)\mathrm{h_p}(f) + 3n\log(4mn), \tag{44}$$

where for (44) we use Remark 3.10 and Lemma 3.14. Using the notation $f^{(\alpha)}(X,Y) = f(X+\alpha,Y)$ and Corollary 3.12, we obtain the inequality

$$\sum_{\alpha\in\mathcal{A}} \mathrm{h_a}(f^{(\alpha)}) \le \sum_{\alpha\in\mathcal{A}} \operatorname{ord}_\alpha R(X)\mathrm{h_a}(f^{(\alpha)}) \le 4mn\mathrm{h_p}(f) + 7m^2n + 3nm\log n. \tag{45}$$

Proof of (40) Fix $\alpha\in\mathbb{K}$. The height of the set $T_5^{(\alpha)} = \bigcup_{x(Q)=\alpha} T_5^{(Q)}$ can be estimated using item 1 of Proposition 4.3 with polynomial $f^{(\alpha)}$ instead of f. We obtain

$$\mathrm{h}\big(T_5^{(\alpha)}\big) \le 3n\big(\mathrm{h_p}(f^{(\alpha)}) + \log(mn) + 1\big). \tag{46}$$

The set T_5 is contained in the union of all $T_5^{(\alpha)}$ with $\alpha\in\mathcal{A}$. Hence combining (43), (45) and (46), we obtain

$$\mathrm{h}(T_5) \le 3n\left(\sum_{\alpha\in\mathcal{A}} \mathrm{h_p}(f^{(\alpha)}) + (\log(mn)+1)|\mathcal{A}|\right) \le 14mn^2\big(\mathrm{h_p}(f) + 2m + 2n\big),$$

as wanted. □

Proof of (41) It is totally analogous to the proof of (40). We define $T_6^{(\alpha)} = \bigcup_{x(Q)=\alpha} \operatorname{Ess}(y^{(Q)})$. If the set $T_6^{(\alpha)}$ is non-empty then the covering $\mathcal{C}\xrightarrow{x}\mathbb{P}^1$ ramifies over α, and condition (33) implies that $f_0^{(\alpha)}(0) = f_0(\alpha) \ne 0$. Hence we may apply (24) with $f^{(\alpha)}$ instead of f. We obtain

$$\mathrm{h}\big(T_6^{(\alpha)}\big) \le n\big(\mathrm{h_p}(f^{(\alpha)}) + 2\big) + 3n\big(\mathrm{h_p}(f^{(\alpha)}) + 4n + \log m\big)\operatorname{ord}_\alpha D(X)$$

Next, we use (43) and (45) to obtain

$$\begin{aligned}\mathrm{h}(T_6) &\le n\big(\sum_{\alpha\in\mathcal{A}} \mathrm{h_p}(f^{(\alpha)}) + 2|\mathcal{A}|\big) + 3n\left(\sum_{\alpha\in\mathcal{A}} \operatorname{ord}_\alpha R(X)\mathrm{h_p}(f^{(\alpha)}) + (4n+\log m)\sum_{\alpha\in\mathcal{A}} \operatorname{ord}_\alpha R(X)\right)\\ &\le 18mn^2\big(\mathrm{h_p}(f) + 2m + \log n\big),\end{aligned}$$

as wanted. This completes the proof of Proposition 5.4. □

5.5 Proof of Proposition 5.5

We have

$$\partial_{\mathbb{L}/\mathbb{K}} \le \sum_{\alpha\in\mathcal{A}} \partial_{\mathbb{K}(\alpha)/\mathbb{K}} + \sum_{\alpha\in\mathcal{A}}\sum_{x(Q)=\alpha} \partial_{\mathbb{K}(\alpha)(Q)/\mathbb{K}(\alpha)} + \sum_{r=1}^{n} \partial_{\mathbb{K}(\zeta_r)/\mathbb{K}}, \tag{47}$$

where ζ_r is a primitive r-th root of unity.

Each $\alpha\in\mathcal{A}$ generates over $\mathbb{K}$ a field of degree at most $\deg R(X) \le 2mn$. Lemma 3.16 and estimate (44) imply that

$$\sum_{\alpha\in\mathcal{A}} \partial_{\mathbb{K}(\alpha)/\mathbb{K}} \le 4mn\sum_{\alpha\in\mathcal{A}} \mathrm{h_a}(\alpha) + 2mn\log(2mn) \le 8mn^2\mathrm{h_p}(f) + 14mn^2\log(4mn).$$

The field $\mathbb{K}(\alpha)(Q)$ is contained in the field generated over $\mathbb{K}(\alpha)$ by the coefficients of the Puiseux expansions of y at Q. Using item 2 of Proposition 4.3, but with polynomial[5] $f^{(\alpha)}$ instead of f, we obtain

$$\sum_{x(Q)=\alpha} \partial_{\mathbb{K}(\alpha)(Q)/\mathbb{K}(\alpha)} \le 8n\big(\mathrm{ord}_\alpha D(X)+1\big)\big(\mathrm{h_p}(f^{(\alpha)})+5n+\log m\big).$$

Hence, applying (45), we obtain

$$\begin{aligned}\sum_{\alpha\in\mathcal{A}}\sum_{x(Q)=\alpha} \partial_{\mathbb{K}(\alpha)(Q)/\mathbb{K}(\alpha)} &\le 8n\left(\sum_{\alpha\in\mathcal{A}} \mathrm{ord}_\alpha R(X)\mathrm{h_p}(f^{(\alpha)})+\sum_{\alpha\in\mathcal{A}}\mathrm{h_p}(f^{(\alpha)})\right)\\ &\quad + 8n\left(\sum_{\alpha\in\mathcal{A}}\mathrm{ord}_\alpha R(X)+|\mathcal{A}|\right)\big(5n+\log m\big)\\ &\le 64mn^2\mathrm{h_p}(f)+144m^2n^2+208mn^3\end{aligned}$$

Finally, Lemma 3.16 implies that

$$\sum_{r=1}^{n}\partial_{\mathbb{K}(\zeta_r)/\mathbb{K}} \le \sum_{r=1}^{n}\log r \le n\log n.$$

Combining all this, we obtain (34). □

6 A Tower of $\bar{\mathbb{K}}$-Points

In this section we retain the set-up of Section 5; that is, we fix a number field $\mathbb{K}$, a curve $\mathcal{C}$ defined over $\mathbb{K}$ and rational functions $x, y \in \mathbb{K}(\mathcal{C})$ such that $\mathbb{K}(\mathcal{C}) = \mathbb{K}(x, y)$. Again, let $f(X, Y) \in \mathbb{K}[X, Y]$ be the $\mathbb{K}$-irreducible polynomial of X-degree m and Y-degree n such that $f(x, y) = 0$, and we again assume that $f_0(X)$ in (25) is monic. We again define the polynomial $R(X)$, the sets $\mathcal{A} \subset \bar{\mathbb{K}}$, $\mathcal{Q} \subset \mathcal{C}(\bar{\mathbb{K}})$ and $T_1, \dots, T_6 \subset M_{\mathbb{K}}$, etc.

We also fix a covering $\widetilde{\mathcal{C}} \xrightarrow{\phi} \mathcal{C}$ of $\mathcal{C}$ by another smooth irreducible projective curve $\widetilde{\mathcal{C}}$; we assume that both $\widetilde{\mathcal{C}}$ and the covering ϕ are defined over $\mathbb{K}$. We consider $\mathbb{K}(\mathcal{C})$ as a subfield of $\mathbb{K}(\widetilde{\mathcal{C}})$; in particular, we identify the functions $x \in \mathbb{K}(\mathcal{C})$ and $x \circ \phi \in \mathbb{K}(\widetilde{\mathcal{C}})$. We fix a function $\widetilde{y} \in \mathbb{K}(\widetilde{\mathcal{C}})$ such that $K(\widetilde{\mathcal{C}}) = \mathbb{K}(x, \widetilde{y})$. We let $\tilde{f}(X, \widetilde{Y}) \in \mathbb{K}[X, \widetilde{Y}]$ be an irreducible polynomial of X-degree $\widetilde{m}$ and $\widetilde{Y}$-degree $\widetilde{n}$ such that $\tilde{f}(x, \widetilde{y}) = 0$; we write

$$\tilde{f}(X, \widetilde{Y}) = \tilde{f}_0(X)\widetilde{Y}^{\widetilde{n}} + \tilde{f}_1(X)\widetilde{Y}^{\widetilde{n}-1} + \cdots + \tilde{f}_{\widetilde{n}}(X)$$

and assume that the polynomial $\tilde{f}_0(X)$ is monic. We define in the similar way the polynomial $\widetilde{R}(X)$, the sets $\widetilde{\mathcal{A}} \subset \bar{\mathbb{K}}$, $\widetilde{\mathcal{Q}} \subset \widetilde{\mathcal{C}}(\bar{\mathbb{K}})$ and $\widetilde{T}_1, \dots, \widetilde{T}_6 \subset M_{\mathbb{K}}$, etc. For defining $\widetilde{T}_5$ and $\widetilde{T}_6$ we need to assume that

$$\widetilde{\mathcal{Q}} \subset \widetilde{\mathcal{C}}(\mathbb{K}), \tag{48}$$

$$\mathbb{K} \text{ contains } e_{\widetilde{Q}}\text{-th roots of unity for all } \widetilde{Q} \in \widetilde{\mathcal{Q}}. \tag{49}$$

We also define the notion of proximity on the curve $\widetilde{\mathcal{C}}$ exactly in the same way as we did it for $\mathcal{C}$ in Definition 5.1, and we have the analogues of Propositions 5.2, 5.3 and 5.4.

In addition to all this, we define one more finite set of places of the field $\mathbb{K}$ as follows. Write $\widetilde{R}(X) = \widetilde{R}_1(X)\widetilde{R}_2(X)$, where the polynomials $\widetilde{R}_1(X), \widetilde{R}_2(X) \in \mathbb{K}(X)$ are uniquely defined by the following conditions:

- the roots of $\widetilde{R}_1(X)$ are contained in the set of the roots of $f_0(X)$;

[5] Recall that $f^{(\alpha)}(X, Y) = f(X + \alpha, Y)$.

- the polynomial $\widetilde{R}_2(X)$ has no common roots with $f_0(X)$ and is monic.

Now let Θ be the resultant of $f_0(X)$ and $\widetilde{R}_2(X)$. Then $\Theta \neq 0$ by the definition of $\widetilde{R}_2(X)$, and we set

$$U = \{v \in M_{\mathbb{K}} : |\Theta|_v < 1\}.$$

Proposition 6.1 *Assume* (29), (30), (48) *and* (49). *Let* $P \in \mathcal{C}(\bar{\mathbb{K}})$ *be semi-defined over* $\mathbb{K}$ *(that is,* $\xi = x(P) \in \mathbb{K}$*), and let* $\widetilde{P} \in \widetilde{\mathcal{C}}(\bar{\mathbb{K}})$ *be a point above* P *(that is,* $\phi(\widetilde{P}) = P$*). Let* v *be a finite place of* $\mathbb{K}$*, and* $\bar{v}$ *an extension of* v *to* $\bar{\mathbb{K}}$*. Assume that* $\widetilde{P}$ *is* $\bar{v}$*-close to some* $\widetilde{Q} \in \widetilde{\mathcal{Q}}$*. Then we have one of the following options.*

- $|\xi|_v > 1$.
- $v \in T \cup \widetilde{T} \cup U$.
- P *is* $\bar{v}$*-adically close to the* $Q \in \mathcal{C}(\bar{\mathbb{K}})$ *which lies below* $\widetilde{Q}$.

For the proof we shall need a simple lemma.

Lemma 6.2 *In the above set-up, there exists a polynomial* $\Phi(X, \widetilde{Y}) \in \mathbb{K}[X, \widetilde{Y}]$ *such that*

$$y = \frac{\Phi(x, \widetilde{y})}{f_0(x)\widetilde{R}(x)}$$

Proof Since $f_0(x)y$ is integral over $\mathbb{K}[x]$, Corollary 3.2 implies that $f_0(x)y \in \widetilde{R}(x)^{-1}\mathbb{K}[x, \widetilde{y}]$, whence the result. □

Proof of Proposition 6.1 We put $\alpha = x(\widetilde{Q})$. By the definition of the set $\widetilde{\mathcal{Q}}$, we have $\alpha \in \widetilde{\mathcal{A}}$. Assume that $|\xi|_v \leq 1$ and $v \notin T \cup \widetilde{T} \cup U$. Let $\widetilde{e}$ be the ramification of $\widetilde{Q}$ over $\mathbb{P}^1$, and let

$$\widetilde{y}_i^{(\widetilde{Q})} = \sum_{k=-k^{(\widetilde{Q})}}^{\infty} a_k^{(\widetilde{Q})} \tilde{\zeta}^{(j-1)k} (x-\alpha)^{k/\widetilde{e}} \qquad (j = 1, \dots, \widetilde{e}), \tag{50}$$

be the equivalent Puiseux expansions of $\widetilde{y}$ at $\widetilde{Q}$ (here $\tilde{\zeta}$ is a primitive $\widetilde{e}$-th root of unity). Since $\widetilde{P}$ is $\bar{v}$-close to $\widetilde{Q}$, we have $|\xi - \alpha|_{\bar{v}} < 1$ and the $\widetilde{e}$ series (50) converge at ξ, with one of the sums being $\widetilde{y}(\widetilde{P})$.

Now let $\Phi(X, \widetilde{Y})$ be the polynomial from Lemma 6.2. Then the $\widetilde{e}$ series

$$\frac{\Phi\big(x, \widetilde{y}_j^{(\widetilde{Q})}\big)}{f_0(x)\widetilde{R}(x)} \qquad (j = 1, \dots, \widetilde{e}) \tag{51}$$

contain all the equivalent Puiseux series of y at $Q = \phi(\widetilde{Q})$. More precisely, if the ramification of Q over $\mathbb{P}^1$ is e, then every of the latter series occurs in (51) exactly $\widetilde{e}/e$ times.

Write $f_0(X)\widetilde{R}(X) = (X-\alpha)^r g(X)$ with $g(\alpha) \neq 0$. The assumption $v \notin T_2 \cup \widetilde{T}_2 \cup \widetilde{T}_3 \cup \widetilde{T}_4 \cup U$ implies that $|g(\alpha)|_{\bar{v}} = 1$. Now Lemma 3.7 implies that the Laurent series at α of the rational function $\big(f_0(x)\widetilde{R}(x)\big)^{-1}$ converges at ξ. Hence all the series (51) converge at ξ, and among the sums we find

$$\frac{\Phi\big(x(\widetilde{P}), \widetilde{y}(\widetilde{P})\big)}{f_0\big(x(\widetilde{P})\big)\widetilde{R}\big(x(\widetilde{P})\big)} = y(P).$$

Hence P is $\bar{v}$-close to Q. □

We shall also need a bound for U similar to that of Proposition 5.4.

Proposition 6.3 *We have* $\mathrm{h}(U) \leq \Upsilon + \Xi$*, where* Υ *is defined in (1) and*

$$\Xi = 2m\widetilde{n}(2\widetilde{m} + 3\log\widetilde{n}) + (m + 2\widetilde{m}\widetilde{n})\log(m + 2\widetilde{m}\widetilde{n}). \tag{52}$$

Proof Item 2 of Proposition 2.2 implies that $\mathrm{h}(U) \le \mathrm{h_a}(\Theta)$, where Θ is the resultant of $f_0(X)$ and $\widetilde{R}_2(X)$. Expressing Θ as the familiar determinant, we find

$$\mathrm{h_a}(\Theta) \le \deg \widetilde{R}_2 \mathrm{h_a}(f_0) + \deg f_0 \mathrm{h_a}(\widetilde{R}_2) + (\deg f_0 + \deg \widetilde{R}_2)\log(\deg f_0 + \deg \widetilde{R}_2). \tag{53}$$

Since both f_0 and $\widetilde{R}_2$ are monic polynomials (by the convention (27) and the definition of $\widetilde{R}_2$), we may replace the affine heights by the projective heights. Further, we have the estimates

$$\deg f_0 \le m, \qquad \deg \widetilde{R}_2 \le \widetilde{m}(2\widetilde{n}-1), \qquad \mathrm{h_p}(f_0) \le \mathrm{h_p}(f),$$
$$\mathrm{h_p}(\widetilde{R}_2) \le (2\widetilde{n}-1)\mathrm{h_p}(\tilde{f}) + (2\widetilde{n}-1)\left(2\widetilde{m} + \log\bigl((\widetilde{n}+1)\sqrt{\widetilde{n}}\bigr)\right),$$

the latter estimate being a consequence of Corollary 3.11 and Lemma 3.14. Substituting all this to (53), we obtain the result. □

7 The Chevalley-Weil Theorem

Now we may to gather the fruits of our hard work. In this section we retain the set-up of Section 6. Here is our principal result, which will easily imply all the theorems stated in the introduction.

Theorem 7.1 *Assume* (29), (30), (48) *and* (49). *Assume that the covering ϕ is unramified outside the poles of x. Let $P \in \mathcal{C}(\bar{\mathbb{K}})$ be semi-defined over $\mathbb{K}$, and let $\widetilde{P} \in \widetilde{\mathcal{C}}(\bar{\mathbb{K}})$ be a point above P. As before, we put $\xi = x(P) = x(\widetilde{P})$. Then for every $v \in M^0_{\mathbb{K}}$ we have one of the following options.*

- $|\xi|_v > 1$.
- $v \in T \cup \widetilde{T} \cup U$.
- *Any extension of v to $\mathbb{K}(P)$ is unramified in $\mathbb{K}(\widetilde{P})$.*

Proof Let $v \in M_{\mathbb{K}}$ be a non-archimedean valuation such that $|\xi|_v \le 1$ and $v \notin T \cup \widetilde{T} \cup U$. Fix an extension $\bar{v}$ of v to $\bar{\mathbb{K}}$, and let $\widetilde{w}$ and w be the restrictions of $\bar{v}$ to $\mathbb{K}(\widetilde{P})$ and $\mathbb{K}(P)$, and $\widetilde{e}$ and e their ramification indexes over v, respectively. We want to show that $\widetilde{e} = e$.

We may assume that $\widetilde{P} \notin \widetilde{\mathcal{Q}}$; otherwise there is nothing to prove by (48). Proposition 5.2 applied to the covering $\widetilde{\mathcal{C}} \to \mathbb{P}^1$ implies that either $\widetilde{e} = 1$ and we are done, or $\widetilde{P}$ is $\bar{v}$-adically close to some $\widetilde{Q} \in \widetilde{\mathcal{Q}}$, which will be assumed in the sequel. Now Proposition 5.3 implies that $\widetilde{e} = e_{\widetilde{Q}}/\gcd(e_{\widetilde{Q}}, \ell)$. Let Q be the point of $\mathcal{C}$ lying under $\widetilde{Q}$. Put $\alpha = x(\widetilde{Q}) = x(Q)$. If $\alpha \notin \mathcal{A}$ then the covering $\mathcal{C} \mapsto \mathbb{P}^1$ does not ramify at Q. Since ϕ is unramified outside the poles of x, the covering $\widetilde{\mathcal{C}} \mapsto \mathbb{P}^1$ does not ramify at $\widetilde{Q}$, that is, $e_{\widetilde{Q}} = 1$. Hence $\widetilde{e} = 1$, which means that v is not ramified in $\mathbb{K}(\widetilde{P})$.

Now assume that $\alpha \in \mathcal{A}$. Proposition 6.1 implies that P is $\bar{v}$-adically close to Q. Now notice that $e_Q = e_{\widetilde{Q}}$, again because ϕ is unramified. Also, $\ell(P, Q, \bar{v}) = \ell(\widetilde{P}, \widetilde{Q}, \bar{v}) = \ell$, just by the definition of this quantity. Again using Proposition 5.3, we obtain that $e = e_Q/\gcd(e_Q, \ell) = \widetilde{e}$. This shows tha $\widetilde{w}$ is unramified over w, completing the proof. □

We also need an estimate for $\mathrm{h}(T \cup \widetilde{T} \cup U)$. Recall the notation

$$\Omega = mn^2\bigl(\mathrm{h_p}(f) + 2m + 2n\bigr), \qquad \widetilde{\Omega} = \widetilde{m}\widetilde{n}^2\bigl(\mathrm{h_p}(\tilde{f}) + 2\widetilde{m} + 2\widetilde{n}\bigr),$$
$$\Upsilon = 2\widetilde{n}\bigl(\widetilde{m}\mathrm{h_p}(f) + m\mathrm{h_p}(\tilde{f})\bigr).$$

Proposition 7.2 *Assume* (29), (30), (48) *and* (49), *and assume in addition that*

$$\text{the covering } \mathcal{C} \xrightarrow{x} \mathbb{P}^1 \text{ does not ramify over the roots of } f_0(X), \tag{54}$$
$$\text{the covering } \widetilde{\mathcal{C}} \xrightarrow{x} \mathbb{P}^1 \text{ does not ramify over the roots of } \tilde{f}_0(X). \tag{55}$$

Then

$$\mathrm{h}(T \cup \widetilde{T} \cup U) \le 60(\Omega + \widetilde{\Omega}) + \Upsilon. \tag{56}$$

Proof Combining Propositions 5.4 and 6.3, we obtain the estimate

$$\mathrm{h}(T \cup \widetilde{T} \cup U) \le 52(\Omega + \widetilde{\Omega}) + \Upsilon + \Xi,$$

where Ξ is defined in (52). A routine calculation show that $\Xi \le 6(\Omega + \widetilde{\Omega})$, which proves (56). $\square$

Now we can prove the theorems from the introduction.

Proof of Theorem 1.3 We may replace $\mathbb{K}$ by $\mathbb{K}(P)$ and assume that $P \in \mathcal{C}(\mathbb{K})$. Put $\xi = x(P)$. Assume first that (29), (30), (48) and (49) hold, and assume in addition that[6]

$$\text{the covering } \mathcal{C} \xrightarrow{x} \mathbb{P}^1 \text{ does not ramify over the roots of } f_0(X)X^m f_0(X^{-1}), \tag{57}$$

$$\text{the covering } \widetilde{\mathcal{C}} \xrightarrow{x} \mathbb{P}^1 \text{ does not ramify over the roots of } \tilde{f}_0(X)X^{\tilde{m}} \tilde{f}_0(X^{-1}). \tag{58}$$

Theorem 7.1 and estimate (56) imply that

$$\mathrm{h}\big(\{v \in \mathrm{Ram}(\mathbb{K}(\widetilde{P})/\mathbb{K}) : |\xi|_v \le 1\}\big) \le 60(\Omega + \widetilde{\Omega}) + \Upsilon.$$

Replacing x by x^{-1} and the polynomials f, $\tilde{f}$ by $X^m f(X^{-1}, Y)$ and $X^{\tilde{m}} \tilde{f}(X^{-1}, Y)$, respectively, we obtain the estimate

$$\mathrm{h}\big(\{v \in \mathrm{Ram}(\mathbb{K}(\widetilde{P})/\mathbb{K}) : |\xi|_v \ge 1\}\big) \le 60(\Omega + \widetilde{\Omega}) + \Upsilon.$$

Thus,

$$\mathrm{h}\big(\mathrm{Ram}(\mathbb{K}(\widetilde{P})/\mathbb{K})\big) \le 120(\Omega + \widetilde{\Omega}) + 2\Upsilon,$$

and Lemma 3.18 implies that

$$\partial_{\mathbb{K}(\widetilde{P})/\mathbb{K}} \le \frac{\nu - 1}{\nu}\big(120(\Omega + \widetilde{\Omega}) + 2\Upsilon\big) + 1.26\nu \le 120(\Omega + \widetilde{\Omega}) + 2\Upsilon. \tag{59}$$

Now let us relax our assumptions. Suppose that we no longer assume (29), (30), (48) and (49), but continue to assume (57) and (58). Then (59) should be replaced by

$$\partial_{\mathbb{L}(\widetilde{P})/\mathbb{L}} \le 120(\Omega + \widetilde{\Omega}) + 2\Upsilon, \tag{60}$$

where $\mathbb{L}$ is the compositum of the fields $\mathbb{K}(Q)$, $\mathbb{K}(\widetilde{Q})$ and the fields generated over $\mathbb{K}$ by e_Q-th and $e_{\widetilde{Q}}$-th roots of unity, for all $Q \in \mathcal{Q}$ and $\widetilde{Q} \in \widetilde{\mathcal{Q}}$. Proposition 5.5 implies that $\partial_{\mathbb{L}/\mathbb{K}} \le 110(\Omega + \widetilde{\Omega})$. Hence

$$\partial_{\mathbb{K}(\widetilde{P})/\mathbb{K}} \le \partial_{\mathbb{L}(\widetilde{P})/\mathbb{K}} \le \partial_{\mathbb{L}(\widetilde{P})/\mathbb{L}} + \partial_{\mathbb{L}/\mathbb{K}} \le 230(\Omega + \widetilde{\Omega}) + 2\Upsilon.$$

Finally, suppose that we no longer assume (57) and (58) either. All finite ramification points are contained in the set $\mathcal{A}$. Hence there is at most $|\mathcal{A}| \le (2n-1)m$ finite ramification points. It follows that there exists a root of unity ζ of order $4m^2n$ such that $f(X, \zeta)X^m f(X^{-1}, \zeta)|_{X=\alpha} \ne 0$ for any finite ramification point α. Now instead of the function y we consider the new function $z = (y - \zeta)^{-1} \in \bar{\mathbb{K}}(\mathcal{C})$. It satisfies the equation $g(x, z) = 0$, where the polynomial

$$g(X, Z) = Z^n f(X, \zeta + Z^{-1}) = g_0(X)Z^n + g_1(X)Z^{n-1} + \cdots + g_n(X) \in \mathbb{K}(\zeta)[X, Z]$$

satisfies

$$\deg_X g = m, \quad \deg_Z g = n, \quad \mathrm{h_p}(g) \le \mathrm{h_p}(f) + 2n \log 2 \tag{61}$$

(we use Corollary 3.12). Also, $\partial_{\mathbb{K}(\zeta)/\mathbb{K}} \le \log(4m^2n)$ by Lemma 3.16.

We have $g_0(X)X^m g_0(X^{-1}) = f(X, \zeta)X^m f(X^{-1}, \zeta)$, and by the choice of ζ the covering $\mathcal{C} \xrightarrow{x} \mathbb{P}^1$ does not ramify over the roots of $g_0(X)X^m g_0(X^{-1})$.

[6] We have to replace here (54) and (55) by more restrictive conditions (57) and (58) because in the proof we deal not only with the function x, but with x^{-1} as well.

In the same way we find a root of unity $\tilde\zeta$ of order $4\widetilde m^2\widetilde n$ such that the function $\widetilde z=(\widetilde y-\tilde\zeta)^{-1}$ satisfies $\widetilde g(x,\widetilde z)=0$ with $g(X,\widetilde Z)\in\mathbb{K}(\tilde\zeta)[X,\widetilde Z]$ satisfying

$$\deg_X\widetilde g=\widetilde m,\quad \deg_{\widetilde Z}\widetilde g=\widetilde n,\quad \mathrm{h_p}(\widetilde g)\le \mathrm{h_p}(\tilde f)+2\widetilde n\log 2 \tag{62}$$

and the covering $\widetilde{\mathcal C}\xrightarrow{x}\mathbb{P}^1$ is unramified over the roots of the polynomial $\widetilde g_0(X)X^{\widetilde m}\widetilde g_0(X^{-1})$. Also, $\partial_{\mathbb{K}(\tilde\zeta)/\mathbb{K}}\le\log(4\widetilde m^2\widetilde n)$.

Thus, (57) and (58) hold with f, $\tilde f$ replaced by g, $\widetilde g$. It follows that

$$\partial_{\mathbb{K}(\zeta,\tilde\zeta)(\widetilde P)/\mathbb{K}(\zeta,\tilde\zeta)}\le 230(\Omega'+\widetilde\Omega')+2\Upsilon',$$

where Ω', $\widetilde\Omega'$ and Υ' are defined like Ω, $\widetilde\Omega$ and Υ but with f, $\tilde f$ replaced by g, $\widetilde g$. Hence

$$\begin{aligned}\partial_{\mathbb{K}(\widetilde P)/\mathbb{K}}&\le 230(\Omega'+\widetilde\Omega')+2\Upsilon'+\partial_{\mathbb{K}(\zeta)/\mathbb{K}}+\partial_{\mathbb{K}(\tilde\zeta)/\mathbb{K}}\\&\le 230(\Omega'+\widetilde\Omega')+2\Upsilon'+\log(4m^2n)+\log(4\widetilde m^2\widetilde n).\end{aligned}\tag{63}$$

A messy calculation using (61) and (62) shows that the right-hand side of (63) does not exceed $400(\Omega+\widetilde\Omega)+2\Upsilon+6m\widetilde n^2$. Theorem 1.3 is proved. □

Proof of Theorem 1.5 Let S' be set of places of the field $\mathbb{K}(P)$ extending the places from S. The right-hand side of (2) will not increase (see item 1 of Proposition 2.2) if we replace $\mathbb{K}$ by $\mathbb{K}(P)$ and S by S'. Thus, we may assume that $P\in\mathcal{C}(\mathbb{K})$. As in the proof of Theorem 1.3 assume first that (29), (30), (48) and (49) hold, and in addition assume[7] (54) and (55). Again using Theorem 7.1 and (56), we obtain

$$\mathrm{h}\big(\mathrm{Ram}(\mathbb{K}(\widetilde P)/\mathbb{K})\smallsetminus S\big)\le 60(\Omega+\widetilde\Omega)+\Upsilon,$$

and applying Lemma 3.18, we obtain

$$\partial_{\mathbb{K}(\widetilde P)/\mathbb{K}}\le 60(\Omega+\widetilde\Omega)+\Upsilon+\mathrm{h}(S).$$

Now we get rid of the assumptions (29), (30), (48), (49), (54) and (55) in exactly the same manner as we did in the proof of Theorem 1.3. The details are routine, we leave them out. □

To prove Theorem 1.6, we need the following result from [4].

Theorem 7.3 *Let $x:\mathcal{C}\to\mathbb{P}^1$ be a finite covering of degree $n\ge 2$, defined over $\mathbb{K}$ and unramified outside a finite set $A\subset\mathbb{P}^1(\bar{\mathbb{K}})$. Put $h=\mathrm{h_a}(A)$ and $\Lambda'=\big(2(\mathbf{g}+1)n^2\big)^{10\mathbf{g}n+12n}$, where $\mathbf{g}=\mathbf{g}(\mathcal{C})$. Then there exists a rational function $y\in\bar{\mathbb{K}}(\mathcal{C})$ such that $\bar{\mathbb{K}}(\mathcal{C})=\bar{\mathbb{K}}(x,y)$ and the rational functions $x,y\in\bar{\mathbb{K}}(\mathcal{C})$ satisfy the equation $f(x,y)=0$, where $f(X,Y)\in\bar{\mathbb{K}}[X,Y]$ is an absolutely irreducible polynomial satisfying*

$$\deg_X f=\mathbf{g}+1,\qquad \deg_Y f=n,\qquad \mathrm{h_p}(f)\le\Lambda'(h+1).\tag{64}$$

Moreover, the number field $\mathbb{L}$, generated over $\mathbb{K}$ by the set A and by the coefficients of f satisfies $\partial_{\mathbb{L}/\mathbb{K}(A)}\le\Lambda'(h+1)$.

Proof of Theorem 1.6 We shall prove the "projective" case (that is, item 1) of this theorem. The "affine" case is proved similarly.

We define $\widetilde\Lambda'$ in the same way as Λ' in Theorem 7.3, but with n and $\mathbf{g}$ replaced by $\widetilde n$ and $\widetilde{\mathbf{g}}$. We use Theorem 7.3 to find functions $y\in\bar{\mathbb{K}}(\mathcal{C})$ and $\widetilde y\in\bar{\mathbb{K}}(\widetilde{\mathcal{C}})$, and polynomials $f(X,Y)\in\bar{\mathbb{K}}[X,Y]$ and $\tilde f(X,\widetilde Y)\in\bar{\mathbb{K}}[X,\widetilde Y]$. Denoting by $\mathbb{L}$ the field generated by the set A and the coefficients of

[7] In this proof we deal only with the function x, and do not need x^{-1}, as we did in the projective case. Therefore we may assume (54) and (55), and do not need more restrictive (57) and (58).

both the polynomials, we find $\partial_{\mathbb{L}/\mathbb{K}(A)} \leq (\Lambda' + \widetilde{\Lambda}')(h+1)$ with $h = \mathrm{h_a}(A)$. Using Lemma 3.16, we estimate $\partial_{\mathbb{K}(A)/\mathbb{K}} \leq 2(\delta-1)h + \log\delta$. Hence

$$\partial_{\mathbb{L}/\mathbb{K}} \leq \big(\Lambda' + \widetilde{\Lambda}' + 2(\delta-1)\big)(h+1).$$

We define the quantities Ω, $\widetilde{\Omega}$ and Υ as in the introduction. Then, applying Theorem 1.3, but over the field $\mathbb{L}$ rather than $\mathbb{K}$, we find $\partial_{\mathbb{L}(\widetilde{P})/\mathbb{L}(P)} \leq 400(\Omega + \widetilde{\Omega}) + 2\Upsilon + 6m\widetilde{n}^2$. We have

$$\partial_{\mathbb{K}(\widetilde{P})/\mathbb{K}(P)} \leq \partial_{\mathbb{L}(\widetilde{P})/\mathbb{K}(P)} = \partial_{\mathbb{L}(\widetilde{P})/\mathbb{L}(P)} + \partial_{\mathbb{L}(P)/\mathbb{K}(P)} \leq \partial_{\mathbb{L}(\widetilde{P})/\mathbb{L}(P)} + \partial_{\mathbb{L}/\mathbb{K}}.$$

The last sum is bounded by

$$400(\Omega + \widetilde{\Omega}) + 2\Upsilon + 6m\widetilde{n}^2 + \big(\Lambda' + \widetilde{\Lambda}' + 2(\delta-1)\big)(h+1),$$

which, obviously, does not exceed $\Lambda(h+1)$, as wanted. □

References

[1] Yu. Bilu, *Effective Analysis of Integral Points on Algebraic Curves*, Ph. D. Thesis, Beer Sheva, 1993.

[2] Yu. Bilu, Quantitative Siegel's Theorem for Galois Coverings, *Compositio Math.*, **106(2)** (1997), 125–158.

[3] Yu. Bilu, A. Borichev, Remarks on Eisenstein, submitted; arXiv:1112.2290.

[4] Yu. Bilu, M. Strambi, Quantitative Riemann Existence Theorem over a Number Field, *Acta Arith.* **145** (2010), 319–339.

[5] C. Chevalley, A. Weil, Un théorème d'arithmétique sur les courbes algébriques, *C. R. Acad. Sci. Paris* **195** (1932), 570–572.

[6] R. Dedekind, *Werke* I, Vieweg, 1930.

[7] K. Draziotis, D. Poulakis, Explicit Chevalley-Weil Theorem for Affine Plane Curves, *Rocky Mountain J. of Math.* **39** (2009), 49–70.

[8] K. Draziotis, D. Poulakis, An Effective Version of Chevalley-Weil Theorem for Projective Plane Curves, arXiv:0904.3845.

[9] B. Dwork and P. Robba, On natural radii of p-adic convergence, *Trans. Amer. Math. Soc.* **256** (1979), 199–213.

[10] B. M. Dwork and A. J. van der Poorten, The Eisenstein Constant, *Duke Math. J.* **65 (1)** (1992), 23–43.

[11] M. Hindry, J. H. Silverman, Diophantine Geometry: an Introduction, Graduate Texts in Math. **201**, Springer Verlag, 2000.

[12] T. Krick, L.M. Pardo, M. Sombra, Sharp estimates for the arithmetic Nullstellensatz, *Duke Math. J.* **109** (2001), 521–598.

[13] S. Lang, *Fundamentals of Diophantine Geometry*, Springer, New York, 1983.

[14] J. B. Rosser, L. Schoenfeld, Lowell Approximate formulas for some functions of prime numbers, *Illinois J. Math.* **6** (1962), 64–94.

[15] W. M. Schmidt, Eisenstein's theorem on power series expansions of algebraic functions. *Acta Arith.* **56(2)** (1990), 161–179.

[16] W. M. Schmidt, Construction and estimates of bases in function fields. *J. Number Theory* **39)** (1991), 181–224.

[17] J.-P. Serre, Quelques applications du théorème de densité de Chebotarev, *Publ. Math. IHES* **54** (1981), 323–401.

[18] J.-P. Serre, Lectures on Mordell-Weil Theorem, 3rd edition, Vieweg, Braunschweig, 1997.

[19] J. H. Silverman, Lower bounds for height functions. *Duke Math. J.* **51** (1984), 395–403.

[20] A. Weil, Arithmétique et géométrie sur les variétés algébriques, *Act. Sc. et Ind.* **206** (1935), 3–16.

Yuri Bilu
IMB, Université Bordeaux 1
351 cours de la Libération
33405 Talence CEDEX
France

Marco Strambi
via del Litorale 145
Antignano, Livorno
57128 Italy

Andrea Surroca
Mathematisches Institut
Universität Basel
Rheinsprung 21
CH-4051 Basel
Switzerland