# Quantitative Siegel's theorem for Galois coverings


**YURI BILU**

Max-Planck-Institut
für Mathematik
Gottfried-Claren-Str. 26
53225 Bonn
GERMANY



MPI 95-105

# Quantitative Siegel's theorem for Galois coverings

Yuri Bilu
Max-Planck-Institut für Mathematik
Gottfried-Claren-Strasse 26
53225 Bonn GERMANY.

**Abstract**

It is known that Siegels theorem on integral points is effective for Galois coverings of the projective line. In this paper we obtain a quantitative version of this result, giving an explicit upper bound for the heights of $S$-integral $\mathbf{K}$-rational points in terms of the number field $\mathbf{K}$, the set of places $S$ and the defining equation of the curve. Our main tools are Baker's theory of linear forms in the logarithms and the quantitative Eisenstein theorem due to Schmidt, Dwork and van der Poorten.

# 1 Introduction

## 1.1 The main result

Let $C$ be a projective curve defined over a number field $\mathbf{K}$ and $x \in \mathbf{K}(C)$ non-constant. For any finite set $S$ of places of $\mathbf{K}$ containing the set $S_\infty$ of archimedean places define the set of $S$-integral points of the curve $C$ (with respect to $x$) as follows:

$$C(x, \mathbf{K}, S) = \{P \in C(\mathbf{K}) : x(P) \in \mathcal{O}_{\mathbf{K},S}\} ,$$

where $\mathcal{O}_{\mathbf{K},S}$ is the ring of $S$-integers of the field $\mathbf{K}$. The classical theorem of Siegel [31, 23] states that $|C(x, \mathbf{K}, S)| < \infty$ as soon as the genus $\mathbf{g}(C) \geq 1$. For curves of genus 2 or more this is covered by a result of Faltings [18], who proved that $|C(\mathbf{K})| < \infty$ when $\mathbf{g}(C) \geq 2$, as was originally conjectured by Mordell.

Both the theorems of Siegel and Faltings are, in general, non-effective. However, Siegel's theorem is effective in some particular cases, for instance, for curves of genus 1 (Baker and Coates [5]). See [22, 29, 9] for quantitative improvements of the result of Baker and Coates.

One more general case of effectivity of Siegel's theorem is when $x : C \to \mathbf{P}^1$ is a *geometrically Galois covering* of the projective line (that is $\overline{\mathbf{Q}}(C)/\overline{\mathbf{Q}}(x)$ is a Galois extension, where $\overline{\mathbf{Q}}$ is the argebraic closure of $\mathbf{Q}$). This was proved by the author [8, Sec. 7], and, independently, by Dvornicich and Zannier [15]. Partial results were obtained by H. Kleiman [21, Cor. (3) of Th. 3] and Poulakis [27, Sec. 2].

In all cases the method of Gelfond–Baker [19, 2] was used, so far the single general effective method in Diophantine analysis. In [35, 30, 9, 11] one can find further information on the effective study of Diophantine equations by Baker's method, including extensive bibliography.

Here we obtain a quantitative version of the effective Siegel's theorem for Galois coverings. Introduce some notation. Given a projective vector $\underline{\alpha} = (\alpha_0 : \ldots : \alpha_k) \in \mathbf{P}^k\left(\overline{\mathbf{Q}}\right)$, we denote by $h(\underline{\alpha})$ its *absolute logarithmic heihgt* (further *height*, we recall the definition in Subsection 1.4). The height of a polynomial is the height of the projective vector composed from its coeffitients. Also, we define the height function $h_x : C\left(\overline{\mathbf{Q}}\right) \to \mathbf{R}^{\geq 0}$ by $h_x(P) = h_{\mathbf{P}^1}(x(P))$, where $h_{\mathbf{P}^1} : \mathbf{P}^1\left(\overline{\mathbf{Q}}\right) \to \mathbf{R}^{\geq 0}$ is the height on $\mathbf{P}^1$.

Let $y \in \mathbf{K}(C)$ be such that $\mathbf{K}(C) = \mathbf{K}(x,y)$ and $f(X,Y) \in \mathbf{K}[X,Y]$ a non-zero separable polynomial such that $f(x,y) = 0$. (We use lowercase letters $x$, $y$, ... for rational functions on $C$ and uppercase letters $X$, $Y$, ... for indeterminants.) For some flexibility, we do not assume $f(x,Y)$ to be the minimal polynomial of $y$ over the ring $\mathbf{K}[x]$; in particular, it can be reducible.

Put

$$\begin{array}{l} m = \deg_X f(X,Y)\,, \quad n = \deg_Y f(X,Y)\,, \quad N = \max(m,n,3)\,, \quad s = |S|\,, \\ d = d_{\mathbf{K}} = [\mathbf{K} : \mathbf{Q}]\,, \quad \mathcal{D} = \mathcal{D}_{\mathbf{K}} \text{ — the absolute discriminant of } \mathbf{K}\,. \end{array} \tag{1}$$

We denote by $\mathcal{N} = \mathcal{N}_{\mathbf{K}} : \mathbf{K} \to \mathbf{Q}$ the norm map. The norm of a fractional ideal is well-defined as a non-negative rational number. For any place $v$ of the field $\mathbf{K}$ we define $\mathcal{N}v$ as the norm of the corresponding prime ideal if $v$ is non-archimedean, and put $\mathcal{N}v = 1$ if $v$ is archimedean. Also, we denote by $p(v)$ the underlying rational prime (which is assumed to be $\infty$ for archimedean $v$), and put

$$\widehat{p}(v) = \begin{cases} p(v)\,, & p(v) < \infty\,, \\ 1\,, & p(v) = \infty\,, \end{cases} \qquad \widehat{p}(S) = \max_{v \in S} \widehat{p}(v)\,. \tag{2}$$

Finally, throughout the paper the symbols $O(\ldots)$, $\ll$ and $\gg$ imply *absolute effective constants.*

**Theorem 1.1** *Suppose that* $\mathbf{g}(C) \geq 1$ *and* $x : C \to \mathbf{P}^1$ *is a Galois covering. Then for any* $P \in C(x, \mathbf{K}, S)$ *we have*

$$h_x(P) \leq \widehat{p}(S)^{dN_1} \left(\mathcal{D} \prod_{v \in S} \mathcal{N}v\right)^{4N_2} \exp\left(400sN_2\left(\log(NS) + O(1)\right) + 600dN_3\left(h(f) + O(N)\right)\right), \tag{3}$$

*where*

$$N_1 = \max\left(n^5\,, 16n^2m^2\,, 256m^3\right), \quad N_2 = \max\left(n^4\,, 10m^2n\right), \quad N_3 = \max\left(mn^7\,, 500m^2n^4\right).$$

## 1.2 An application: the superelliptic Diophantine equation

The Diophantine equation

$$y^n = F(x) \tag{4}$$

is called *superelliptic* if the pair $(n\,, F)$ satisfies the following "LeVeque condition": write $F(x) = a(x - \alpha_1)^{r_1} \cdots (x - \alpha_k)^{r_k}$ with pairwise distinct $\alpha_1, \ldots, \alpha_k$; then

$k \geq 2$ and the $k$-tuple $\left(\frac{n}{(n,r_1)}, \ldots, \frac{n}{(n,r_k)}\right)$ is not a permutation of $(\nu, 1, \ldots, 1)$ or $(2, 2, 1, \ldots, 1)$. An equivalent condition: the (non-singular model of the) plain curve (4) has positive genus.

As follows from Siegel's theorem (see also [24]), the equation (4) has finitely many $S$-integral solutions $(x, y)$ in the field $\mathbf{K}$. A. Baker [4] was the first to obtain an effective bound for the size of the solutions. Though he considered only the case $\mathbf{K} = \mathbf{Q}$ and $S = \{\infty\}$, and his condition on $(n, F)$ was stronger than stated above, it was clear that his method, suitably modified, can be applied in the general situation. Indeed, Baker's result was sharpened and extended to arbitrary number fields and/or $S$-integral solutions in [34, 36, 6, 26].

Recently P. Voutier obtained a new effective bound for the integral solutions of (4), having considerably improved the previous results (in the case $S = S_\infty$). He proved that any solution $(x, y) \in \mathcal{O}_\mathbf{K} \times \mathbf{K}$ of (4) satisfies

$$\max\left(h(x), h(y)\right) \leq c(N, d) \left(\mathcal{D} \exp(dh(f))\right)^{n^8 m^2/3} \left(h(f) + \log \mathcal{D} + 1\right)^{n^6 m^2 d}, \qquad (5)$$

the constant $c(N, d)$ being effective. Here $f(X, Y) = F(X) - Y^n$ and we use the notation (1). (The reader should be warned that we express Voutier's result in our notation, which is different from his. He uses the relative exponetial height $H_\mathbf{K}(\ldots)$ (instead of the absolute logarithmic height $h(\ldots)$, as in the present paper), and his $m$ and $n$ correspond to our $n$ and $m$, respectively.)

Since the curve (4) has positive genus and $\overline{\mathbf{Q}}\left(x, \sqrt[n]{F(x)}\right)$ is a Galois extension of $\overline{\mathbf{Q}}(x)$, Theorem 1.1 is applicable to the superelliptic equation. Therefore we can evaluate the quality of the estimate (3), looking at what it gives for the superelliptic equation in comparison with the result of Voutier. For any solution $(x, y) \in \mathcal{O}_\mathbf{K} \times \mathbf{K}$ we have

$$\max\left(h(x), h(y)\right) \leq c(N, d) \mathcal{D}^{4N_2} \exp\left(600 d N_3 h(f)\right).$$

which is better than (5) when $n$ is suficiienly large.

Of course, the superelliptic equation is a very particular case of Theorem 1.1. Thus, we obtain an asymptotically better result in a more general setting.

## 1.3 Ramification indices

We identify set-theoretically $\mathbf{P}^1\left(\overline{\mathbf{Q}}\right)$ and $\overline{\mathbf{Q}} \cup \{\infty\}$ in the obvious way. For any $\alpha \in \overline{\mathbf{Q}} \cup \{\infty\}$ we denote by $e_1 = e_1(\alpha), \ldots, e_\rho = e_{\rho(\alpha)}(\alpha)$ the ramification indices of the covering $x : C \to \mathbf{P}^1$ over the point $\alpha$. Put

$$e_\alpha = \gcd(e_1, \ldots, e_\rho). \qquad (6)$$

(Sometimes we write $e_\alpha(x)$, when several coverings of the projective line are considered.)

Actually, we shall prove a more general result.

**Theorem 1.2** *Suppose that*

$$\sum_{\alpha \in \overline{\mathbf{Q}}} \left(1 - e_\alpha^{-1}\right) > 1. \qquad (7)$$

*Then any $P \in C(x, \mathbf{K}, S)$ satisfies (3).*

When $\mathbf{g} \geq 1$ and the covering is Galois, the relation (7) holds. Indeed, in this case all ramification indices over a point $\alpha$ are equal to $e_\alpha$, and we write Hurwitz formula as

$$2\mathrm{g} - 2 + 2n = \sum_{\alpha \in \overline{\mathbf{Q}} \cup \{\infty\}} \frac{n}{e_\alpha} (e_\alpha - 1) .$$

Then

$$\sum_{\alpha \in \overline{\mathbf{Q}}} \left(1 - e_\alpha^{-1}\right) = 1 + \frac{2\mathrm{g} - 2}{n} + e_\infty^{-1} > 1 ,$$

which is (7). Thus, Theorem 1.1 follows from Theorem 1.2.

## 1.4 Additional notation and conventions

For any place $v$ of the field $\mathbf{K}$ (and any number field to occur) the corresponding (multiplicative) valuation $| \dots |_v$ is normalized so that its restriction to $\mathbf{Q}$ is a standard infinite or $p$-adic valuation. In addition, for a non-archimedean $v$ we shall use an additive valuation $\mathrm{Ord}_v : \mathbf{K}^* \to \mathbf{Z}$ normalizing it so that 1 belongs to the image of $\mathrm{Ord}_v$. In explicit terms $\mathrm{Ord}_v(\alpha) = d_v \log |\alpha|_v / \log \mathcal{N} v$, where $d_v = d_v(\mathbf{K}) = \left[\mathbf{K}_v : \mathbf{Q}_{p(v)}\right]$ is the local degree of $v$.

Recall the definition of the *absolute logarithmic heght* of a projective vector $\underline{\alpha} = (\alpha_0 : \dots : \alpha_k) \in \mathbf{P}^k \left(\overline{\mathbf{Q}}\right)$:

$$h(\underline{\alpha}) = d_{\mathbf{L}}^{-1} \sum_v \max_{0 \leq i \leq k} d_v(\mathbf{L}) \log |\alpha_i|_v , \tag{8}$$

the sum being over all places of the field $\mathbf{L} = \mathbf{Q}(\alpha_0, \dots, \alpha_k)$ (by the product formula, it does not depend on the choice of the homogenious coordinates).

With an abuse of notation, for $\alpha \in \overline{\mathbf{Q}}$ we write $h(\alpha)$ instead of $h(1 : \alpha)$. As follows from the definition of the absolute logarithmic height, for any $\nu \in \mathbf{Z}$ and $\alpha_1, \dots, \alpha_k, \alpha \in \overline{\mathbf{Q}}$ we have

$$\begin{aligned} h(\alpha_1 + \dots + \alpha_k) &\leq h(\alpha_1) + \dots + h(\alpha_k) + \log k , & (9) \\ h(\alpha_1 \cdots \alpha_k) &\leq h(\alpha_1) + \dots + h(\alpha_k) , & (10) \\ h(\alpha^\nu) &= |\nu| h(\alpha) . & (11) \end{aligned}$$

We write

$$f(X, Y) = g_0(X) Y^n + \text{ terms of lower degree in } Y . \tag{12}$$

Denote by $R(X)$ the resultant of $f(X,Y)$ and $\frac{\partial f}{\partial Y}(X,Y)$ with respect to $Y$ and by $D(X)$ the discriminant of $f(X,Y)$ with respect to $Y$. Then we have

$$\begin{aligned} R(X) &= g_0(X) D(X) , & (13) \\ \deg R(X) &\leq (2n-1)m , & (14) \\ h(R) &\leq (2n-1)h(f) + O(n \log N) , & (15) \\ \deg D(X) &\leq (2n-2)m , & (16) \\ h(D) &\leq (2n-2)h(f) + O(n \log N) , & (17) \end{aligned}$$

as follows from the standard determinant representations of the resultant and discriminant.

For $\alpha \in \overline{\mathbf{Q}}$ put

$$u_\alpha = \mathrm{Ord}_\alpha g_0(X), \qquad \mu_\alpha = \mathrm{Ord}_\alpha D(X), \qquad f_\alpha(X,Y) = f(\alpha + X, Y), \tag{18}$$

where $\mathrm{Ord}_\alpha$ is the order of vanishing at $\alpha$. Then $f_0(X,Y) = f(X,Y)$, and we similarly write $u$ and $\mu$ instead of $u_0$ and $\mu_0$, respectively. We have trivially

$$h(f_\alpha) \le h(f) + mh(\alpha) + O(\log N). \tag{19}$$

The relation (7) is false when $m = 1$ or $n = 1$. Therefore we suppose further that

$$n, m \ge 2. \tag{20}$$

which, together with (9)–(17) and (19) will be frequently used in our estimates, mostly without special referring.

We also need the following well-known fact (see, for example, [28, Lemma 3]).

**Proposition 1.4.1** *Let $F(X)$ be a polynomial of degree $\rho$ with algebraic coefficients and $\alpha_1, \dots, \alpha_\rho$ its roots counted with multiplicities. Then*

$$h(\alpha_1) + \cdots + h(\alpha_\rho) \le h(F) + \log(\rho + 1).$$

**Warning** The letter $e$ is reserved here exclusively for ramification indices, being never used for $2.718\ldots$ (for the latter we write $\exp(1)$).

## 1.5 Plan of the paper

In Section 2 we summarize neccessary properties of algebraic power series, including the quantitative Eisenstein theorem due to Schmidt [28] and Dwork–van der Poorten [17].

In Section 3 we prove that, given $P \in C(x, \mathbf{K}, S)$ and $\alpha \in \mathbf{K}$, the principal ideal $(x(P) - \alpha)$ is "almost a $e_\alpha$-th power". The qualitative part (Proposition 3.2) is self-contained, while the quantitative part (how "almost"?) depends on the estimates of Section 2.

Section 4 is a summary of the auxiliary material needed for the proof of Theorem 1.2, in particular, Siegel's construction of convenient units [32, 7, 9] and Baker's theory [38, 39].

In Section 5 we give a detailed proof of a particular case of Theorem 1.2. The argument is based on the results obtained or quoted in Sections 3 and 4. In Section 6 we prove Theorem 1.2 in its full generality, reducing it to the result of Section 5.

**Acknoledgements.** I would like to thank P.M. Voutier for putting at my disposal his paper [37]. I was inspired by his result and some of his ideas. In fact, it was he who pointed me that some ideas from my own unpublished paper [10] can be used for sharpening bounds for integral points.

This work was done during the author's post-doctoral stage at the Max-Planck-Institute für Mathematik. I am grateful to Prof. F. Hirzebruch and MPI für Mathematik for their kind invitation and excellent working conditions.

# 2 Eisenstein theorem and further properties of algebraic power series

## 2.1 Preliminaries

Let $y = \sum_{k=-k_0}^{\infty} a_k x^{k/e}$ be an algebraic power series, where we alwayse assume $k_0 \geq 0$ and $a_{-k_0} \neq 0$ when $k_0 > 0$. Also, we suppose that $y$ cannot be presented as a power series in $x^{1/e'}$ with $e' < e$.

Let $y$ satisfy an algebraic equation $f(x, y) = 0$ with $f(X, Y) \in \mathbf{K}(X, Y)$. We use the notation $m$, $n$, $u$, $\mu$,, etc., introduced in Subsections 1.1 and 1.4. Clearly, $k_0/e \leq u \leq m$.

Let $\mathbf{L}$ be the extension of $\mathbf{K}$ generated by all the coeffitients $a_k$ of the series $y$. It is well-known that $[\mathbf{L} : \mathbf{K}] \leq n$.

**Theorem 2.1** *For any place $v$ of the field $\mathbf{K}$ there exist real numbers $A_v$, $A'_v \geq 1$ such that $A_v = A'_v = 1$ for all but finitely many $v$,*

$$d^{-1} \sum_v d_v \log A_v \leq (2n-1)\, h(f) + O\Big(n(n + \log N)\Big), \tag{21}$$

$$d^{-1} \sum_v d_v \log A'_v \leq h(f) + O(\log n), \tag{22}$$

*and for any place $w|v$ of the field $\mathbf{L}$ we have*

$$|a_k|_w \leq A'_v A_v^{u+k/e} \quad (k \geq -k_0). \tag{23}$$

*Furthermore, for any non-archimedean place $v$ we have $\frac{d_v \log A'_v}{\log \mathcal{N}v} \in \mathbf{Z}$ and*

$$d^{-1} \sum_{\frac{d_v \log A_v}{\log \mathcal{N}v} \notin \mathbf{Z}} \log \mathcal{N}v \leq (2n-1)h(f) + O(n \log N). \tag{24}$$

This theorem is a combination of results of Schmidt [28] and Dwork–van der Poorten [17]. Formally, they considered only the case $e = 1$. Though the general case requires no new ideas, it cannot be reduced to the case $e = 1$ just by the substitution $x = x_1^e$. Therefore we include some detailes for the sake of completeness (see Subsections 2.2 and 2.3).

In Subsection 2.4 we obtain additional auxiliary properties of algebraic power series.

## 2.2 Eisenstein theorem: the unramified case

In this subsection we assume that $e = 1$. Then $y = \sum_{k=-k_0}^{\infty} a_k x^k$. We need one more definition. Let $F(X)$ be a polynomial with coeffitients in the number field $\mathbf{K}$ and $\alpha_1, \ldots, \alpha_t$ its roots. For any place $v$ of the field $\mathbf{K}$ fix a prolongation to $\mathbf{K}(\alpha_1, \ldots, \alpha_t)$ and put

$$\sigma_v(F) = \min\left(1, |\alpha_1|_v, \ldots, |\alpha_t|_v\right).$$

Clearly, $\sigma_v(F)$ does not depend on the fixed prolongation.

Recall that $R(X)$ is the resultant of $f(X, Y)$ and $\frac{\partial f}{\partial Y}(X, Y)$ with respect to $Y$. We write $R(X) = A x^{u+\mu} R^*(X)$ where $R^*(0) = 1$.

Normalize the polynomial $f(X,Y) = g_0(X)Y^n + \ldots$ so that $g_0(X) = X^u g_0^*(X)$ with $g_0^*(0) = 1$. As usual, denote by $|f|_v$ the maximum of $|\beta|_v$ over all the coeffitients $\beta$ of $f(X,Y)$.

**Theorem 2.2 (Dwork–Robba–Schmidt–van der Poorten)** *For any valuation $v$ of the field* $\mathbf{K}$ *with put*

$$A'_v = \begin{cases} 2n|f|_v\,, & p(v) = \infty, \\ |f|_v\,, & p(v) < \infty\,. \end{cases} \tag{25}$$

$$A_v = \begin{cases} 2/\sigma_v\,(R^*)\,, & p(v) = \infty\,, \\ 1/\sigma_v\,(R^*)\,, & n < p(v) < \infty\,, \\ c(v,n)/\sigma_v\,(R^*)\,, & p(v) \le n\,, \end{cases} \tag{26}$$

*where* $c(v,n) = np(v)^{\frac{1}{p(v)-1}}$. *Then for any place $w|v$ of the field* $\mathbf{L}$ *we have*

$$|a_k|_w \le A'_v A_v^{u+k} \quad (k \ge -k_0)\,. \tag{27}$$

We indicate the main steps of the proof. Until the end of this subsection we write $\sigma_v = \sigma_v(R^*)$. Given a place $w$ of the field $\mathbf{L}$, denote by $r_w$ the $w$-adic radius of convergence of the series $y = \sum_{k=-k_0}^{\infty} a_k x^k$.

The heart of the proof is the following

**Lemma 2.2.1** *If $w|v$ with $n < p(v) \le \infty$ then*

$$r_w \ge \sigma_v \tag{28}$$

*If $w|v$ with $p(v) \le n$ then*

$$r_w \ge c(v,n)^{-1}\sigma_v\,. \tag{29}$$

For the case $n < p(v) \le \infty$ see Schmidt [28]. (As indicated by Schmidt, the case $n < p(v) < \infty$ is a direct consequence of a result of Dwork and Robba [16].)

The case $p(v) \le n$ is due to Dwork and van der Poorten [17]. Let $\alpha$ be a root of $R^*(X)$ with the property $|\alpha|_v = \sigma_v$. Then by [17, Th. 3], the series $\widehat{y} = \sum_{k=-k_0}^{\infty} a_k\alpha^{-k}x^k = \sum_{k=-k_0}^{\infty} \widehat{a}_k x^k$ converges for $|x|_v < c(v,n)$, whence the result.

It should be mentioned that in [28, 17] only the case $k_0 = 0$ is treated. However, the general case can be easily reduced to the case $k_0 = 0$. Indeed, put

$$\widetilde{y} = x^{k_0} y = \sum_{k=0} \widetilde{a}_k x^k\,, \qquad \widetilde{a}_k = a_{k-k_0}\,.$$

Clearly, the radii $r_w$ and $\widetilde{r}_w$ of $w$-adic convergence of respectively $y$ and $\widetilde{y}$ are equal. Further, $\widetilde{y}$ satisfies the equation $\widetilde{f}(x,\widetilde{y}) = 0$, where $\widetilde{f}(X,Y) = X^{k_0 n} f\left(X, X^{-k_0}Y\right)$. Defining $\widetilde{R}$ and $\widetilde{R}^*$ for $\widetilde{f}$ as $R$ and $R^*$ were defined for $f$, we see that $\widetilde{R}^* = R^*$. Thus, $\widetilde{r}_w = r_w$ and $\widetilde{\sigma}_v = \sigma_v$. This reduces the case of arbitrary $k_0 \ge 0$ to the case $k_0 = 0$.

Put $\tau_v = \min\left(\sigma_v\,, \min_{w|v} r_w\right)$.

**Lemma 2.2.2** *The inequality (27) holds with*

$$A_v = \begin{cases} 2/\sigma_v\,, & p(v) = \infty\,, \\ 1/\tau_v\,, & p(v) < \infty\,, \end{cases} \tag{30}$$

*and $A'_v$ defined as in (25).*

This is a result of Schmidt [28, Lemma 2]. Though he considers only the case $k_0 = 0$, his argument plainly works for arbitrary $k_0 \geq 0$. Also, what he proves is exactly the inequality (27), but he formulates his result in a slightly weaker form, with $m$ instead of $u$ in (27).

Now Theorem 2.2 follows as a direct consequence of Lemmas 2.2.1 and 2.2.2.

## 2.3 Eisenstein theorem: the general case

Put

$$\tilde{f}(X,Y) = f\left(X^e, Y\right), \quad \tilde{y} = \sum_{k=-k_0}^{\infty} a_k x^k, \tag{31}$$

so that $\tilde{f}(x, \tilde{y}) = 0$. Define $\tilde{R}$, $\tilde{R}^*$ and $\tilde{u}$ for $\tilde{f}$ as $R$, $R^*$ and $u$ were defined for $f$. Then

$$\begin{aligned} \tilde{R}^*(X) &= R^*\left(X^e\right), & (32)\\ \tilde{u} &= eu. & (33) \end{aligned}$$

As follows from (32),

$$\sigma_v\left(\tilde{R}^*\right) = \sigma_v\left(R^*\right)^{1/e}. \tag{34}$$

Now put

$$A_v = \begin{cases} 2^n/\sigma_v\left(R^*\right), & p(v) = \infty, \\ 1/\sigma_v\left(R^*\right), n < p(v) < \infty, & \\ c(v,n)^n/\sigma_v\left(R^*\right), & p(v) \leq n, \end{cases} \tag{35}$$

and define $A'_v$ as in (25) (provided $f(X,Y)$ is normalized as described above). By (33), (34) and Theorem 2.2, applied to the series $\tilde{y}$, we have (23). Further, Schmidt [28, Lemma 5] showed that

$$d^{-1}\sum_v d_v \log\left(1/\sigma_v\right) \leq (2n-1)h(f) + O(n\log N). \tag{36}$$

Therefore

$$\begin{aligned} d^{-1}\sum_v d_v \log A_v &\leq d^{-1}\sum_v d_v \log(1/\sigma_v) + nd^{-1}\sum_{p(v)\leq n} \log c(v,n) + O(n) \\ &\leq (2n-1)h(f) + O(n(n+\log N)), \end{aligned}$$

which is (21).

The inequality (22) is obvious and, as follows from (25), for any non-archimedean place $v$ the quotient $\frac{d_v \log A'_v}{\log \mathcal{N}v}$ is an integer. It remains to establish (24). In view of (35), for a non-arcimedean $v$, the quotient $\frac{d_v \log A_v}{\log \mathcal{N}v}$ can be not an integer only in one of the following cases:

(a) $p(v) \leq n$;

(b) there is a root $\alpha$ of $R^*(X)$ such that $\frac{d_v \log|\alpha|_w}{\log \mathcal{N}v} \notin \mathbf{Z}$ for some place $w$ of $\mathbf{K}(\alpha)$ lying above $v$.

We estimate separately $d^{-1}\sum \log \mathcal{N}v$ over non-arcimedean $v$ belonging to the cases (a) and (b) above. For (a) the estimate is straightforward:

$$d^{-1}\sum_{(a)} \log \mathcal{N}v = \sum_{p\leq n} \log p \ll n\,. \tag{37}$$

For (b), let $\alpha_1, \ldots, \alpha_s$ be a maximal selection of roots of $R^*(X)$ pairwise non-conjugate over $\mathbf{K}$. Put $\nu_i = [\mathbf{K}(\alpha_i):\mathbf{K}]$. If $w$ is a prolongation of $v$ to $\mathbf{K}(\alpha)$, then the denominator of the rational number $\frac{d_v \log|\alpha_i|_w}{\log \mathcal{N}v}$ is at most $\nu_i$. Therefore

$$\begin{aligned} d^{-1}\sum_{(b)} \log \mathcal{N}v &\leq \sum_{i=1}^{s} \nu_i h(\alpha_i) \\ &\leq h(R^*) + \log(1+\deg R^*) \\ &\leq (2n-1)h(f) + O(n\log N)\,, \end{aligned}$$

where the second inequality is by Proposition 1.4.1. Together with (37) this proves (24). Theorem 2.1 is proved.

**Remark 2.3.1** Given a polynomial $f(X,Y)$, separable in $Y$, there exist $n\,(=\deg_Y f)$ distinct power series $y_i = \sum_{k=-k_0(i)}^{\infty} a_{ik}x^{k/e_i}$ such that $f(x,y_i)=0$. As follows from definitions (35) and (25), the values of $A_v$ and $A'_v$ depend only on the polynomial $f(X,Y)$ and are common for all the series $y_i$. This observation will not help us to improve the final result, but will simplify our notation in Section 3.

## 2.4 Field generated by the coefficients, etc.

Let $\mathbf{K}_1$ be the subfield of constants of the field $\mathcal{K}_1 = \mathbf{K}((x))(y)$. We begin with the following standard fact.

**Proposition 2.4.1** *The field $\mathbf{L}$ is an extension of $\mathbf{K}_1$ of degree at most $e$.*

**Proof** Since $\mathcal{K}_1 \subseteq \mathcal{L} = \mathbf{L}\left(\left(x^{1/e}\right)\right)$, the field $\mathbf{K}_1$ is a subfield of $\mathbf{L}$. It remains to prove that $[\mathbf{L}:\mathbf{K}]\leq e$.

Using Hensel's lemma, one can easily show that $\mathcal{K}_1 = \mathbf{K}_1\left(\left(\widehat{x}^{1/e}\right)\right)$, where $\widehat{x} = \alpha x$ with $\alpha\in\mathbf{K}_1$. Hence $y = \sum_{k=-k_0}^{\infty} b_k \widehat{x}^{k/e}$ with $b_k\in\mathbf{K}_1$. Therefore $\mathbf{L}\subseteq \mathbf{K}_1\left(\alpha^{1/e}\right)$. The proof is complete.

Put

$$\kappa = \left\lfloor \frac{\mu}{[\mathbf{K}_1:\mathbf{K}]} \right\rfloor \leq \frac{e\mu}{[\mathbf{L}:\mathbf{K}]}\,, \tag{38}$$

where $\lfloor\gamma\rfloor$ is the maximal integer not exceeding $\gamma\in\mathbf{R}$.

**Lemma 2.4.2** *The field $\mathbf{L}$ is generated over $\mathbf{K}$ by $a_{-k_0}, \ldots, a_\kappa$. The relative discriminant $\mathcal{D}_{\mathbf{L}/\mathbf{K}}$ satisfies*

$$\begin{aligned} d^{-1}\log\mathcal{N}\left(\mathcal{D}_{\mathbf{L}/\mathbf{K}}\right) &\leq 2\left(2n\nu(\mu+u\nu)+\nu^2\right)h(f) + \\ & \qquad O\left(n\nu(\mu+u\nu)(n+\log N)+\nu\log\nu\right)\,, \end{aligned} \tag{39}$$

*where $\nu = [\mathbf{L}:\mathbf{K}]$.*

**Proof** Put

$$\mathbf{L}_0 = \mathbf{K}\left(a_{-k_0}, \dots, a_\kappa\right), \quad \mathcal{L}_0 = \mathbf{L}_0\left(\left(x^{1/e}\right)\right), \quad \delta = [\mathbf{L}:\mathbf{L}_0] = [\mathcal{L}:\mathcal{L}_0]. \tag{40}$$

Clearly, $\mathcal{L}_0(y) = \mathcal{L}$. Let $\varphi(Y) = \varphi_0 Y^\delta + \dots$ be a minimal polynomial of $y$ over the ring $\mathcal{R} = \mathbf{L}_0\left[\left[x^{1/e}\right]\right]$ (at least one of its coeffitients is invertible in $\mathcal{R}$). By the Gauss Lemma, the polynomial

$$\mathcal{N}_{\mathcal{L}_0/\mathbf{K}((x))}(\varphi(Y)) \in \mathbf{K}[[x]](Y)$$

divides $f(x,Y)^{[\mathbf{L}:\mathbf{K}_1]}$ in the ring $\mathbf{K}[[x]](Y)$.

Denote by $\Delta(x)$ the discriminant of $\varphi(Y)$. Then

$$\mathcal{N}_{\mathcal{L}_0/\mathbf{K}((x))}\Delta(x)\ \Big|\, D(x)^{[\mathbf{L}:\mathbf{K}_1]} \tag{41}$$

in the ring $\mathbf{K}[[x]]$. Obviously,

$$\mathrm{Ord}_x\Delta(x) \geq \delta(\delta-1)(\kappa+1)/e.$$

Comparing $\mathrm{Ord}_x$ of the both sides of (41), we obtain

$$e[\mathbf{L}_0:\mathbf{K}]\,\delta(\delta-1)(\kappa+1)/e \leq [\mathbf{L}:\mathbf{K}_1]\,\mu. \tag{42}$$

Since $\delta = [\mathbf{L}:\mathbf{L}_0]$, we can rewrite (42) as

$$\delta - 1 \leq \frac{\mu}{[\mathbf{K}_1:\mathbf{K}]}\Big/(\kappa+1) < 1.$$

Thus, $\delta = 1$. This proves the first assertion. (See [13, Lemma 3] for a similar result.)

For the second assertion we need a result of Silverman [33, Th. 2].

**Proposition 2.4.3 (Silverman)** *Let* $\underline{\alpha} = (\alpha_0 : \dots : \alpha_k) \in \mathbf{P}^k\left(\overline{\mathbf{Q}}\right)$, *and* $[\mathbf{K}(\alpha) : \mathbf{K}] = \nu$. *Then*

$$d^{-1}\log\mathcal{N}\left(\mathcal{D}_{\mathbf{K}(\alpha)/\mathbf{K}}\right) \leq 2\nu(\nu-1)h(\underline{\alpha}) + \nu\log\nu. \tag{43}$$

In our case $\underline{\alpha} = (1 : a_{-k_0} : \dots : a_\kappa)$. We obtain an upper estimate for $h(\underline{\alpha})$ from Theorem 2.1:

$$\begin{aligned} h(\underline{\alpha}) &\leq h(f) + \left((2n-1)\,h(f) + O\left(n\,(n+\log N)\right)\right)(u+\kappa/e) \\ &\leq \left(2n\left(\frac{\mu}{\nu}+u\right)+1\right)h(f) + O\left(n\,(n+\log N)\left(\frac{\mu}{\nu}+u\right)\right). \end{aligned}$$

Together with (43) this gives the desired estimate for the relative discriminant. The lemma is proved.

Recall that $y$ cannot be written as a power series in $x^{1/e'}$ with $e' < e$. Hence for any prime $q|e$ there exists $k \not\equiv 0 \bmod q$ such that $a_k \neq 0$. Denote by $k(q)$ the minimal among such $k$.

**Lemma 2.4.4** *For any prime $q|e$ we have*

$$k(q) \le \kappa/(q-1) - 1 . \qquad (44)$$

**Proof** It is very similar to the proof of the fist part of Lemma 2.4.2, with the field $\mathcal{L}_0$ replaced by $\mathcal{L}_1 = \mathcal{L}\left(\left(x^{1/e_1}\right)\right)$, where $e_1 = e/q$.

Clearly, $\mathcal{L}_1(y) = \mathcal{L}$. Let $\varphi(Y) = \varphi_0 Y^q + \ldots$ be a minimal polynomial of $y$ over the ring $\mathcal{R} = \mathbf{L}\left[\left[x^{1/e_1}\right]\right]$. Then

$$\mathcal{N}_{\mathcal{L}_1/\mathbf{K}((x))}\left(\varphi(Y)\right) \Big| f(x,Y)^{[\mathbf{L}:\mathbf{K}_1]}$$

in the ring $\mathbf{K}\left[[x]\right](Y)$.

Denote by $\Delta(x)$ the discriminant of $\varphi(Y)$. Then

$$\mathcal{N}_{\mathcal{L}_1/\mathbf{K}((x))}\Delta(x) \ \Big| D(x)^{[\mathbf{L}:\mathbf{K}_1]} \qquad (45)$$

in the ring $\mathbf{K}\left[[x]\right]$. Since

$$\mathrm{Ord}_x\Delta(x) \ge q\,(q-1)\,(k(q)+1)/e ,$$

we have

$$e_1\,[\mathbf{L}:\mathbf{K}]\,q\,(q-1)\,(k(q)+1)/e \le [\mathbf{L}:\mathbf{K}_1]\,\mu , \qquad (46)$$

which yields (44) at once.

# 3 Study of a fixed $\alpha$

*In this section we consider only non-archimedean places, unless the contrary is stated explicitly.*

Until the end of this section we fix $\alpha \in \mathbf{K}$ and $P \in C\,(x\,,\mathbf{K}\,,S)$. Recall that

$$e_\alpha = \gcd(e_1\,,\ldots,\,e_\rho) , \qquad (47)$$

where $e_1\,,\ldots,\,e_\rho$'are the ramification indices of the covering $x: C \to \mathbf{P}^1$ over the point $\alpha$. We say that a (non-archimedean) place $v$ is *regular* if $e_\alpha|\mathrm{Ord}_v\,(x(P)-\alpha)$, and *irregular* otherwise.

In this section we prove

**Lemma 3.1** *We have*

$$d^{-1} \sum_{\substack{v \text{ is irregular,} \\ v \notin S}} \log \mathcal{N}v \le 12n^2\,(\mu_\alpha + nu_\alpha)\left(h\,(f) + mh(\alpha) + O(n+\log N)\right) \qquad (48)$$

If $e_\alpha = 1$, the lemma is trivial. Therefore we may suppose that $e_\alpha \ge 2$. In this case all the ramification indices $e_1\,,\ldots,\,e_\rho$ over $\alpha$ are greater or equal to 2, whence $\rho \le n/2$. Consequently

$$\mu_\alpha \ge (e_1 - 1) + \cdots + (e_\rho - 1) = n - \rho \ge n/2 , \qquad (49)$$

which will be used in our estimates.

Let

$$y_i = \sum_{k=-k_0(i)}^{\infty} a_{ik}(x-\alpha)^{k/e_i} \quad (1 \le i \le \rho) \tag{50}$$

be the Puiseaux expansions of $y$ at $\alpha$. Actually, for any $i$ we have $e_i$ equivalent expansions

$$y_{ij} = \sum_{k=-k_0(i)}^{\infty} a_{ik}\xi_i^{jk}(x-\alpha)^{k/e_i} \quad (0 \le j \le e_i - 1), \tag{51}$$

where $\xi_i$ is a fixed primitive root of unity of degree $e_i$. We have

$$f(x,Y) = g_0(x)\prod_{i=1}^{\rho}\prod_{j=0}^{e_i-1}(Y - y_{ij}) .$$

We denote by $\mathbf{L}_i$ the field generated by all the coeffitients $a_{ik}$ of the series $y_i$. Put $\nu_i = [\mathbf{L}_i : \mathbf{K}]$. Further, for any prime $q|e_i$ let $k_i(q)$ be the minimal $k \not\equiv 0 \bmod q$ such that $a_{ik} \neq 0$. By Theorem 2.1 together with Remark 2.3.1, for any (archimedean and non-archimedean) place $v$ of the field $\mathbf{K}$ there exist $A_v$, $A'_v \ge 1$ such that

$$|a_{ik}|_v \le A'_v A_v^{k/e_i} \tag{52}$$

and satisfying (21), (22) and (24) with $h(f)$, $u$ and $\mu$ replaced by $h(f_\alpha)$, $u_\alpha$ and $\mu_\alpha$ respectively.

Let M be a finite set of non-archimedean places of the field $K$ defined as follows:

$$\mathrm{M} = \mathrm{M}_1 \cup \mathrm{M}_2 \cup \mathrm{M}_3 \cup \mathrm{M}_4 \cup \mathrm{M}_5 ,$$

where

$$\begin{aligned}
\mathrm{M}_1 &= \{v : p(v) \le n\} , \\
\mathrm{M}_2 &= \{v : |\alpha|_v < 1\} , \\
\mathrm{M}_3 &= \{v : v \text{ is ramified in one of the fields } \mathbf{L}_1, \ldots, \mathbf{L}_\rho\} , \\
\mathrm{M}_4 &= \left\{v : \left|a_{ik_i(q)}\right|_v < 1 \text{ for some } i \in \{i, \ldots, \rho\} \text{ and prime } q|e_i\right\} , \\
\mathrm{M}_5 &= \{v : A_v A'_v > 1\} .
\end{aligned}$$

**Proposition 3.2** *Any $v \notin S \cup \mathrm{M}$ is regular.*

In view of this proposition, Lemma 3.1 is a direct consequence of the following estimates (we put $\Sigma_i = d^{-1}\sum_{v\in \mathrm{M}_i} \log \mathcal{N}v$):

$$\begin{aligned}
\Sigma_1 &\ll n , && (53)\\
\Sigma_2 &\le h(\alpha) , && (54)\\
\Sigma_3 &\le 5n^2(\mu_\alpha + u_\alpha n)\Big(h(f_\alpha) + O(n + \log N)\Big) , && (55)\\
\Sigma_4 &\le 4n^2(\mu_\alpha + u_\alpha \log_2 n)\Big(h(f_\alpha) + O(n + \log N)\Big) , && (56)\\
\Sigma_5 &\le 4n\Big(h(f_\alpha) + O(n + \log N)\Big) , && (57)
\end{aligned}$$

Here (53) and (54) are obvious. It remains to establish (55)–(57) and to prove Proposition 3.2.

**Proof of (55)** We may suppose that for some $r \le s$ the fields $\mathbf{L}_1, \ldots, \mathbf{L}_r$ are pairwise non-conjugate over $\mathbf{K}$ and any $\mathbf{L}_i$ is conjugate to one of $\mathbf{L}_1, \ldots, \mathbf{L}_r$. Then

$$\nu_i + \cdots + \nu_r \le n, \tag{58}$$

which yields

$$\nu_i^2 + \cdots + \nu_r^2 \le n^2. \tag{59}$$

We estimate the relative discriminant $D_{\mathbf{L}_i/\mathbf{K}}$ by Lemma 2.4.2:

$$\begin{aligned} d^{-1} \log \mathcal{N}\mathcal{D}_{\mathbf{L}_i/\mathbf{K}} &\le 2\left(2n\nu_i\left(\mu_\alpha + u_\alpha\nu_i\right) + \nu_i^2\right) h\left(f_\alpha\right) + \\ &\quad O\left(n\nu_i\left(\mu_\alpha + u_\alpha\nu_i\right)(n + \log N) + \nu_i \log \nu_i\right). \end{aligned}$$

Using (58) and (59), we obtain

$$\begin{aligned} \Sigma_3 &\le d^{-1} \sum_{i=1}^{r} \log \mathcal{N}\left(\mathcal{D}_{\mathbf{L}_i/\mathbf{K}}\right) \\ &\le 2\left(2n^2\left(\mu_\alpha + u_\alpha n\right) + n^2\right) h\left(f_\alpha\right) + \\ &\quad O\left(n^2\left(\mu_\alpha + u_\alpha n\right)(n + \log N) + n \log n\right), \end{aligned}$$

which proves (55) (recall that $\mu_\alpha \ge n/2 \ge 1$).

**Proof of (56)** Let $\mathrm{M}_4'(i,q)$ be the set of all non-archimedean places $w$ of the field $\mathbf{L}_i$ such that $\left|a_{ik_i(q)}\right|_w < 1$, and $\mathrm{M}_4(i,q)$ the set of all points $\mathbf{K}$ below $\mathrm{M}_4'(i,q)$. Then

$$\mathrm{M}_4 = \bigcup_{i=1}^{r} \bigcup_{q|e_i} \mathrm{M}_4(i,q).$$

By Lemma 2.4.4 we have $k_i(q) \le e_i\mu_\alpha/\nu_i$. Using Theorem 2.1, we obtain

$$\begin{aligned} \Sigma_4'(i,q) &:= [d_{\mathbf{L}_i}]^{-1} \sum_{w \in \mathrm{M}_4'(i,q)} \log \mathcal{N}_{\mathbf{L}_i}(w) \\ &\le h\left(a_{ik_i(q)}\right) \\ &\le h\left(f_\alpha\right) + \left((2n-1)h\left(f_\alpha\right) + O\left(n(n + \log N)\right)\right)\left(u_\alpha + k_i(q)/e_i\right) \\ &\le 4n\left(u_\alpha + \mu_\alpha/\nu_i\right)\left(h\left(f_\alpha\right) + O(n + \log N)\right) \end{aligned}$$

(we again use $\mu_\alpha \ge n/2$). Further,

$$\begin{aligned} \Sigma_4(i,q) &:= d^{-1} \sum_{v \in \mathrm{M}_4(i,q)} \log \mathcal{N}v \\ &\le \nu_i \Sigma_4'(i,q) \\ &\le 4n\left(\nu_i u_\alpha + \mu_\alpha\right)\left(h\left(f_\alpha\right) + O(n + \log N)\right). \end{aligned}$$

There are at most $\log_2 e_i$ distinct prime divisors of $e_i$. Since

$$\begin{aligned} \log_2 e_1 + \cdots + \log_2 e_r &\le e_1 + \cdots + e_r &&\le n, \\ \nu_1 \log_2 e_1 + \cdots + \nu_r \log_2 e_r &\le (\nu_1 + \cdots + \nu_r) \log_2 n &&\le n \log_2 n, \end{aligned}$$

we have

$$\begin{aligned}\Sigma_4 &\leq \sum_{i=1}^{r}\sum_{q|e_i}\Sigma_4(i,q)\\ &\leq 4n\left(u_\alpha n\log_2 n+\mu_\alpha n\right)\left(h\left(f_\alpha\right)+O(n+\log N)\right),\end{aligned}$$

which is (56).

**Proof of (57)** We write $M_5 = M_5' \cup M_5''$, where $M_5'$ consists of those $v \in M_5$ for which $\frac{d_v \log A_v}{\log \mathcal{N}v} \in \mathbf{Z}$, and $M_5'' = M_5 \setminus M_5'$. In accordance with this partition of the set$M_5$, we write $\Sigma_5 = \Sigma_5' + \Sigma_5''$. Recall that $\frac{d_v \log A_v'}{\log \mathcal{N}v}$ is always in $\mathbf{Z}$.

The sum $\Sigma_5'$ is estimated using (21):

$$\begin{aligned}\Sigma_5' &\leq d^{-1}\sum_{v\in M_5'} d_v\left(\log A_v+\log A_v'\right)\\ &\leq d^{-1}\sum_{v} d_v\left(\log A_v+\log A_v'\right)\\ &\leq 2n\left(h\left(f_\alpha\right)+O(n+\log N)\right).\end{aligned}$$

The sum $\Sigma_5''$ is estimated using (24):

$$\Sigma_5'' \leq 2nh\left(f_\alpha\right)+O\left(\log N\right).$$

This proves (57).

The proof of Proposition 3.2 is based on the following almost trivial fact.

**Proposition 3.3** *Let $K$ be a local field of characteristic 0, with residue field of characteristic $p$. Let $\pi$ be a primitive element of $K$ and $\eta \in K$. For any $e \in \mathbf{Z}$ not divisible by $p$ and for any choice of the root $\eta^{1/e}$, the ramification index of $K\left(\eta^{1/e}\right)$ over $K$ is $e/\gcd\left(e, \mathrm{Ord}_\pi(\eta)\right)$.*

**Proof** Write $\eta = \pi^\tau \theta$, where $\tau = \mathrm{Ord}_\pi(\eta)$ and $\theta$ is a unit of $K$. Fix a root $\theta^{1/e}$. Since $p$ is not a divisor of $e$, the field $K\left(\theta^{1/e}\right)$ is unramified over $K$. Replacing $K$ by $K\left(\theta^{1/e}\right)$ and $\pi$ by $\pi\left(\theta^{1/e}\right)^{-1}$, we may suppose that $\eta = \pi^\tau$.

Put $e' = e/\gcd(e,\tau)$ and $\tau' = \tau/\gcd(e,\tau)$ Then $\eta^{1/e} = \left(\pi^{1/e'}\right)^{\tau'}$ for some choice of the root $\pi^{1/e'}$. Therefore $K\left(\eta^{1/e}\right) \subseteq K\left(\pi^{1/e'}\right)$. On the other side, $\gcd(e',\tau') = 1$, therefore exists $a \in \mathbf{Z}$ such that $\tau' a \equiv 1 \bmod e'$. Then $K\left(\eta^{1/e}\right) \supseteq K\left(\left(\eta^{1/e}\right)^a\right) = K\left(\pi^{1/e'}\right)$. Thus, $K\left(\eta^{1/e}\right) = K\left(\pi^{1/e'}\right)$, the latter field being a totally ramified extension of $K$ of degree $e'$. The proposition is proved.

**Proof of Proposition 3.2** Put $x_0 = x(P) - \alpha$ and fix $v \notin M \cup S$. Then $\tau = \mathrm{Ord}_v(x_0) \geq 0$, because $v \notin M_2 \cup S$. If $\tau = 0$ then there is nothing to prove. Thus, assume that $\tau > 0$. Fix a prolongation of $v$ to $\overline{\mathbf{Q}}$. Then all the series

$$y_i(P) = \sum_{k=-k_0(i)}^{\infty} a_{ik}\left(x_0^{1/e_i}\right)^k \quad (1 \leq i \leq s)$$

converge in $v$-metric, because $|x_0|_v < 1$ and $v \notin \mathrm{M}_5$. For some $i$ and some choice of the root $x_0^{1/e_i}$ we have $y_i(P) = y(P)$. Fix this $i$ and this choice of the root until the end of the proof, and omit the index $i$ in the further reasoning.

Since $v \notin \mathrm{M}_3$, it is not ramified in the field $\mathbf{L} = \mathbf{L}_i$.

Denote by $\mathbf{K}_v$ and $\mathbf{L}_v$ the completions with respect to (the fixed prolongation of) $v$. If $e = e_i$ divides $\tau$, the proof is finished. Therefore we may suppose that $e$ does not divide $\tau$, that is $e' = e/\gcd(e, \tau) > 1$.

Let $q$ be a prime divisor of $e'$. Then $q$ does not divide $\tau' = e/\gcd(e, \tau)$. Put

$$\omega = \sum_{k=k(q)}^{\infty} a_k \left(x_0^{1/e}\right)^k .$$

Since $v \notin \mathrm{M}_4 \cup \mathrm{M}_5$, we have

$$\mathrm{Ord}_v(\omega) = \mathrm{Ord}_v\left(\left(x_0^{1/e}\right)^{k(q)}\right) = \frac{\tau' k(q)}{e'}$$

(there is a unique prolongation of $\mathrm{Ord}_v$ to the algebraic closure of $\mathbf{L}_v$).

On the other hand

$$\omega = \sum_{k=-k_0}^{k(q)-1} a_k \left(x_0^{1/e}\right)^k \in \widetilde{\mathbf{L}}_v := \mathbf{L}_v\left(x_0^{1/e''}\right) ,$$

where $e'' = e/q$. Since $v \notin \mathrm{M}_1$, we may apply Proposition 3.3. It implies that the ramication index of $\widetilde{\mathbf{L}}_v$ over $\mathbf{L}_v$ is $e''/\gcd(e'', \tau) = e'/q$. Therefore $\mathrm{Ord}_v(\omega)$ is $q/e'$ times an integer (recall that $\mathbf{L}_v$ is unramified over $\mathbf{K}_v$). Thus, $q$ divides the product $\tau' k(q)$ — a contradiction. The proposition is proved, which completes the prooof of Lemma 3.1.

# 4 Auxiliary material

## 4.1 Siegel's construction of convenient units of number fields

Propositions 4.1.1 and 4.1.3 of this subsection go back to Siegel's famous paper [32].

Let $S = (v_0, \ldots, v_{s-1})$ be a finite set of places of the number field $\mathbf{K}$ and $\eta_1, \ldots, \eta_{s-1}$ a fundamental system of $S$-units. The *S-regulator* $R(S) = R_{\mathbf{K}}(S)$ is, by definition, the absolute value of the determinant of the matrix

$$[d_{v_i} \log |\eta_j|_{v_i}]_{1 \le i,j \le s-1} . \tag{60}$$

It is well-defined and equal to the usual regulator $R = R_{\mathbf{K}}$ when $S = S_\infty$.

**Proposition 4.1.1** *There exists a fundemental system of $S$-units $\eta_1, \ldots, \eta_{s-1}$ satisfying*

$$h(\eta_1) \cdots h(\eta_{s-1}) \le s^{2s-2} d^{1-s} R(S) , \tag{61}$$

$$h^*(\eta_1) \cdots h^*(\eta_{s-1}) \le s^{2s-2} \zeta^{s-1} R(S) , \tag{62}$$

$$(\zeta d)^{-1} \le h(\eta_i) \le s^{2s-2} \zeta^{s-2} R(S) . \tag{63}$$

*Here* $h^*(\eta) = \max(1, h(\eta))$ *and* $\zeta = 1201 \left(\frac{\log d'}{\log\log d'}\right)^3$ *with* $d' = \max(d, 3)$. *Furthermore, let* $[a_{ij}]_{1\le i,j\le s-1}$ *be the matrix inverse to (60). Then*

$$|a_{ij}| \le s^{2s-2}\zeta \quad (1 \le i, j \le s-1). \tag{64}$$

**Proof** See Bugeaud and Győry [7, Lemma 1]. Note that the left-hand inequality in (63) is the well-known result of Dobrowolski [14].

**Corollary 4.1.2** *Suppose that*

$$\eta = \eta_1^{b_1} \cdots \eta_{s-1}^{b_{s-1}},$$

*where* $\eta_1, \ldots, \eta_{s-1}$ *are from Proposition 4.1.1 and* $B = \max(|b_1|, \ldots, |b_{s-1}|)$. *Then*

$$\begin{aligned} h(\eta) &\le s^{2s-1}\zeta^{s-2}R(S)B, &(65)\\ B &\le s^{2s}\zeta h(\eta). &(66)\end{aligned}$$

**Proof** Straightforward from (63) and (64).

**Proposition 4.1.3** *For any* $\alpha \in \mathbf{K}$ *there exists an S-unit* $\eta = \eta_1^{b_1} \cdots \eta_{s-1}^{b_{s-1}}$ *such that* $\beta = \alpha\eta^{-1}$ *satisfies*

$$d^{-1} \sum_{\substack{v\in S\\ v\ne v_0}} d_v \left|\log|\beta|_v\right| \le s^{2s-1}\zeta^{s-2}R(S).$$

**Proof** Put $\eta = \eta_1^{b_1} \cdots \eta_{s-1}^{b_{s-1}}$, where $b_i$ is the nearest integer to $\theta_i = \sum_{j=1}^{s-1} a_{ij} d_{v_j} \log|\alpha|_{v_j}$. Then $\beta = \alpha\eta^{-1}$ satisfies

$$\begin{aligned} d^{-1} \sum_{\substack{v\in S\\ v\ne v_0}} d_v \left|\log|\beta|_v\right| &= d^{-1} \sum_{\substack{v\in S\\ v\ne v_0}} d_v \left|\sum_{i=1}^{s-1} (\theta_i - b_i)\log|\eta_i|_v\right| \\ &\le d^{-1} \sum_{i=1}^{s-1}\sum_{v\in S} |\theta_i - b_i| d_v \left|\log|\eta_i|_v\right| \\ &\le (2d)^{-1} \sum_{i=1}^{s-1}\sum_{v\in S} d_v \left|\log|\eta_i|_v\right| \\ &= h(\eta_1) + \cdots + h(\eta_{s-1}) \\ &\le s^{2s-1}\zeta^{s-2}R(S), \end{aligned}$$

as desired.

Let $h = h_{\mathbf{K}}$ be the class number of the field $\mathbf{K}$. (The letter $h$ will denote the class number only in the remaining part of this subsection, and nowhere more in this paper. Therefore there is no danger of confusing it with $h$ used for heights.) The following result was obtained independently by Bugeaud and Győry [7, Lemma 3] and by the author [9, Proposition 1.4.8]. (See Pethő [25] and Hajdu [20] for similar results.)

**Lemma 4.1.4** *Assume that* $S \supseteq S_\infty$. *Then*

$$R(S) \le hR \prod_{v\in S\setminus S_\infty} \log \mathcal{N}v.$$

**Corollary 4.1.5** *Suppose that* $\mathbf{K} \neq \mathbf{Q}$. *Then*

$$R(S) \ll d^{-d}\sqrt{D}(\log D)^{d-1} \prod_{v\in S\setminus S_\infty} \log \mathcal{N}v \ll \mathcal{D}^{0.51}\left(\prod_{v\in S}\mathcal{N}v\right)^{0.01}. \qquad (67)$$

**Proof** The first inequality follows from Siegel's estimate $hR \ll d^{-d}\sqrt{\mathcal{D}}(\log \mathcal{D})^{d-1}$ (see [32, Satz 1]). Further, note that $\log \mathcal{D} \leq c_1\mathcal{D}^{0.01}$ and $\log \mathcal{N}v \leq c_1(\mathcal{N}v)^{0.01}$, where $c_1, c_2, \ldots$ are absolute effective constants. Also, $\log \mathcal{N}v \leq (\mathcal{N}v)^{0.01}$ as soon as $\mathcal{N}v \geq c_2$, and there exist at most $c_2 d$ non-archimedean places $v$ with $\mathcal{N}v \leq c_2$. We obtain

$$(\log D)^{d-1} \prod_{v\in S\setminus S_\infty} \log \mathcal{N}v \leq c_3^d\left(\mathcal{D}\prod_{v\in S}\mathcal{N}v\right)^{0.01}$$

with $c_3 = c_1^{1+c_2}$. Since $c_3^d d^{-d} \ll 1$, this proves the second inequality.

## 4.2 One more estimate for the relative discriminant

In Subsection 2.4 we quoted Silverman's estimate for the relative discriminant in terms of generating elements. Below we obtain an estimate of a different type, in terms of ramified places. The results of this subsection are certainly not new, but we did not find a suitable reference.

Let $\mathbf{L}/\mathbf{K}$ be a finite extension of number fields. In this subsection $v$ (respectively $w$) is always a non-archimedean point of the field $\mathbf{K}$ (respectively $\mathbf{L}$). We denote by $e(w) = e_{\mathbf{L}/\mathbf{K}}(w)$, the relative ramification index, and by $f(w) = f_{\mathbf{L}/\mathbf{K}}(w)$ the relative degree of the residue fields. For any $v$, the function $|\ldots|_v$ is well-defined on the set of fractionl ideals of $\mathbf{K}$, as well as on the set of fractionl ideals of the $v$-adic completion $\mathbf{K}_v$. In particular, Proposition 4.5 (ii) in [12, Ch. 1] can be written as

$$\log\left|\mathcal{D}_{\mathbf{L}/\mathbf{K}}\right|_v = \sum_{w|v}\log\left|\mathcal{D}_{\mathbf{L}_w/\mathbf{K}_v}\right|_v. \qquad (68)$$

We begin with a local estimate.

**Proposition 4.2.1** *Suppose that* $w|v$. *Then*

$$-\log\left|\mathcal{D}_{\mathbf{L}_w/\mathbf{K}_v}\right|_v \leq \left(-e(w)\log|e(w)|_v + (e(w)-1)\, d_v^{-1}\log \mathcal{N}v\right) f(w). \qquad (69)$$

**Proof** Replacing $\mathbf{K}_v$ by its unramified closure in $\mathbf{L}_w$, we may assume that $f(w) = 1$. Let $\Pi$ be a primitive element of $\mathbf{L}_w$. Then $g(\Pi) = 0$, where $g(X) = a_eX^e + a_{e-1}X^{e-1} + \ldots + a_0 \in \mathbf{K}_v[X]$ is a polynomial of degree $e = e(w)$ with $a_e = 1$,

$$\begin{aligned} |a_i|_v &\leq 1 \quad (1 \leq i \leq e-1), \\ |a_0|_v &= 1. \end{aligned}$$

For $1 \leq i < j \leq e$ we have

$$\left|ia_i\Pi^{\,i}\right|_w \neq \left|ja_j\Pi^{\,j}\right|_w.$$

Hence

$$|g'(\Pi)|_w = \max_{1\leq i\leq e}\left|ia_i\Pi^{\,i}\right|_w \geq \left|e\Pi^{\,e-1}\right|_w.$$

Therefore $\varphi = \mathcal{N}_{\mathbf{L}_w/\mathbf{K}_v}(g'(\Pi))$ satisfies

$$|\varphi|_v \geq |e|_v^e \cdot |\pi|_v^{e-1},$$

where $\pi = \mathcal{N}_{\mathbf{L}_w/\mathbf{K}_v}(\Pi)$ is a primitive element of $\mathbf{K}_v$. Since $\mathcal{D}_{\mathbf{L}_w/\mathbf{K}_v}$ divides $\varphi$, we have

$$\begin{aligned} -\log\left|\mathcal{D}_{\mathbf{L}_w/\mathbf{K}_v}\right|_v &\leq -\log|\varphi|_v \\ &\leq -e\log|e|_v - (e-1)\log|\pi|_v \\ &= -e\log|e|_v + (e-1)d_v^{-1}\log\mathcal{N}v\,, \end{aligned}$$

as desired.

**Remark 4.2.2** It is well-known that (69) turns to equality when $|e|_v = 1$.

**Proposition 4.2.3** *Put* $\nu = [\mathbf{L}:\mathbf{K}]$. *Then*

$$d^{-1}\log\mathcal{N}\left(\mathcal{D}_{\mathbf{L}/\mathbf{K}}\right) \leq (\nu-1)d^{-1}\sum_{v \text{ is ramified in } \mathbf{L}}\log\mathcal{N}v + \nu^2\log\nu\,.$$

**Proof** By (68) and (69)

$$\begin{aligned} -\log\left|\mathcal{D}_{\mathbf{L}/\mathbf{K}}\right|_v &\leq \left(d_v^{-1}\log\mathcal{N}v\right)\sum_{w|v}(e(w)-1)\,f(w) - \sum_{w|v}e(w)f(w)\log|e(w)|_v \\ &\leq (\nu-1)d_v^{-1}\log\mathcal{N}v - \nu\log|\nu!|_v\,. \end{aligned}$$

Hence

$$\begin{aligned} \log\mathcal{N}\left(\mathcal{D}_{\mathbf{L}/\mathbf{K}}\right) &= \sum_{v \text{ is ramified in } \mathbf{L}} d_v\left(-\log\left|\mathcal{D}_{\mathbf{L}/\mathbf{K}}\right|_v\right) \\ &\leq (\nu-1)\sum_{v \text{ is ramified in } \mathbf{L}}\log\mathcal{N}v + d\,\nu\log\nu!\,, \end{aligned}$$

which completes the proof.

## 4.3 Baker's theory

We summarize necessary facts from Baker's theory of linear forms in the logarithms in the following proposition.

**Proposition 4.3.1 (Waldschmidt, Yu)** *Let* $\mathbf{K}$ *be a number field of degree* $d$ *and* $\alpha_0, \ldots, \alpha_r$ *non-zero elements of* $\mathbf{K}$. *Also, let* $v$ *be a place of* $\mathbf{K}$ *and* $0 < \varepsilon \leq 1$. *Suppose that*

$$0 < \left|\alpha_0\alpha_1^{b_1}\cdots\alpha_r^{b_r} - 1\right|_v \leq \exp(-\varepsilon B)\,, \tag{70}$$

*where* $b_1, \ldots, b_r \in \mathbf{Z}$ *and* $B = \max(b_1, \ldots, b_r, 3)$. *Then*

$$B \leq c(r,d)\varepsilon^{-1}\widehat{p}(v)^{d-0.5}h^*(\alpha_0)\cdots h^*(\alpha_r)(\log h')\log\left(\varepsilon^{-1}h'\right)\,, \tag{71}$$

*where* $h^*(\ldots)$ *is defined as in Proposition 4.1.1,* $h' = \max(h(\alpha_1), \ldots, h(\alpha_r), 3)$ *and* $c(r,d) = \exp(3r\log(rd) + O(r + \log d))$. *(In the archimedean case the multiple* $\log h'$ *can be skipped.)*

**Proof** The archmedean case is due to Waldschmidt [38]. Define the parameters in [38, p. 215] as follows:

$$\begin{array}{lll} n = r+1\,, & A_i = \exp\left(h^*(\alpha_i)\right) & (0 \le i \le r)\,, \\ E = \exp(1)\,, & f = \exp(-1)\,, & Z_0 = 7 + 3\log(r+1) + \log d\,. \end{array}$$

Applying Corollary 10.2 from [38] in this set-up, we obtain

$$\varepsilon B \le c(r,d) h^*(\alpha_0) \cdots h^*(\alpha_r) \log\left(3 + \frac{B}{h^*(\alpha_0)}\right),$$

which yields (71) (without the multiple $\log h'$) after obvious calculations.

The non-archmedean case is due to Yu [39]. Define the parameters in [39, p. 241–242] as follows:

$$n = r+1\,, \qquad \delta = \min\left(1\,, \frac{e_v \varepsilon}{2\log p}\right), \qquad h_i = 2h^*(\alpha_i)\log p \quad (0 \le i \le r)\,,$$

where $e_v$ is the ramification index of $v$ over $\mathbf{Q}$ (in particular, $e_v \le d$). In this set-up the third displayed formula on [39, p. 242] would turn to

$$\frac{\varepsilon e_v B}{\log p} \le c(r,d)\widehat{p}(v)^{d-1} h^*(\alpha_1) \cdots h^*(\alpha_r)(\log h') \log\left(\varepsilon^{-1} h' p\right),$$

which yields (71) at once.

# 5 The main argument

In this section we suppose that one of the following conditions holds:

(A) There exist distinct $\alpha, \beta \in \overline{\mathbf{Q}}$ such that $e_\alpha$ and $e_\beta$ have a common divisor $e \ge 3$.

(B) There exist distinct $\alpha, \beta, \gamma \in \overline{\mathbf{Q}}$ such that $e_\alpha$, $e_\beta$ and $e_\gamma$ have a common divisor $e \ge 2$.

Put

$$\nu_\alpha = [\mathbf{K}(\alpha) : \mathbf{K}]\,, \qquad \kappa_\alpha = n^2\left(\mu_\alpha + n u_\alpha\right)\left(h\left(f\right) + m h(\alpha) + O(n + \log N)\right),$$

and define $\nu_\beta$, $\nu_\gamma$, $\kappa_\beta$, $\kappa_\gamma$ similarly. Also, put

$$\Upsilon = \begin{cases} e\nu_\alpha \nu_\beta & \text{in the case (A)}, \\ e\nu_\beta \max(\nu_\alpha\,, \nu_\gamma) & \text{in the case (B)}, \end{cases}$$

$$\delta = \begin{cases} de^2\nu_\alpha \nu_\beta & \text{in the case (A)}, \\ de^2\nu_\alpha \nu_\beta \nu_\gamma & \text{in the case (B)}, \end{cases} \qquad \Theta = s\Upsilon\left(\log(Ns) + O(1)\right).$$

**Theorem 5.1** *Suppose that either (A) or (B) holds. Then for any $P \in C(x\,, \mathbf{K}\,, S)$ we have*

$$h_x(P) \le \widehat{p}(S)^\delta \left(\mathcal{D} \prod_{v \in S} \mathcal{N}v\right)^{2.1\Upsilon} \exp\left(130\Theta + ed\left(22\nu_\alpha\nu_\beta(\kappa_\alpha + \kappa_\beta) + 15\nu_\gamma\nu_\beta(\kappa_\gamma + \kappa_\beta)\right)\right), \tag{72}$$

*where in the case (A) the terms $\nu_\gamma$ and $\kappa_\gamma$ should be replaced by $\nu_\alpha$ and $\kappa_\alpha$.*

**Proof** In a few words, the proof is organized as follows. For a given $P \in C(x, \mathbf{K}, S)$ we construct algebraic numbers $\varphi$ and $\varphi'$ with the following three properties:

(i) the heights of $\varphi$ and $\varphi'$ are of the same magnitude as $h_x(P)$;

(ii) each of $\varphi$ and $\varphi'$ is "almost an $S$-unit" (an $S$-unit times an algebraic number of bounded height);

(iii) for some place $v_0$ the ratio $\varphi/\varphi'$ (slightly modified) is "very close to 1" with respect to the $v_0$-metric.

Using (ii), (iii) and Baker's theory, we estimate the heights of $\varphi$ and $\varphi'$. In view of (i), this would give a bound for $h_x(P)$.

### 1 The choice of $v_0$

Fix $P \in C(x, \mathbf{K}, S)$ and put $x_0 = x(P)$. We have

$$h_x(P) = d^{-1} \sum_{v \in S} d_v \log |x_0|_v ,$$

whence

$$h_x(P) = h(x_0) \leq s \log |x_0|_{v_0} \tag{73}$$

for some $v_0 \in S$. Prolong somehow $v_0$ to $\overline{\mathbf{Q}}$ and fix this prolongation until the end of the proof.

We put

$$\begin{aligned} \sigma_0 &= \begin{cases} \max(|\alpha|_{v_0}, |\beta|_{v_0}, 1) & \text{in the case (A)}, \\ \max(|\alpha|_{v_0}, |\beta|_{v_0}, |\gamma|_{v_0}, 1) & \text{in the case (B)}, \end{cases} \\ \tilde{h} &= \begin{cases} \max(h(\alpha), h(\beta), 1) & \text{in the case (A)}, \\ \max(h(\alpha), h(\beta), h(\gamma), 1) & \text{in the case (B)}. \end{cases} \end{aligned}$$

### 2 Construction of $\varphi$ and field $\mathbf{L}$

When $|x_0|_{v_0} \leq 10e\sigma_0$ then (73) implies an upper bound for $h_x(P)$ much better than (72). Hence we may suppose that $|x_0|_{v_0} \geq 10e\sigma_0$, whence the series $1 + \sum_{k=1}^{\infty} \binom{1/e}{k} \left(\frac{\beta-\alpha}{x_0-\beta}\right)^k$ converges in $v_0$-metric, and its sum, denoted by $\sqrt[e]{\frac{x_0-\alpha}{x_0-\beta}}$, satisfies

$$\left| \sqrt[e]{\frac{x_0 - \alpha}{x_0 - \beta}} - 1 \right|_{v_0} \ll \frac{e\sigma_0}{|x_0|_{v_0}} . \tag{74}$$

Fix a primitive $e$-th root of unity $\xi$ (in paricular $\xi = -1$ in the case (B)) and put

$$\theta = \xi \sqrt[e]{\frac{x_0 - \alpha}{x_0 - \beta}}, \qquad \mathbf{L} = \mathbf{K}(\alpha, \beta, \theta), \qquad \varphi = (x_0 - \beta)(\theta - 1)^e .$$

If $|x_0|_{v_0} \geq ce^6\sigma_0^2$, where $c$ is a sufficiently large absolute effective constant, then $\varphi \neq 0$ and

$$\left|(\xi - 1)^{-e}(x_0 - \beta)^{-1}\varphi - 1\right|_{v_0} \ll e^3\sigma_0|x_0|_{v_0}^{-1} \ll |x_0|_{v_0}^{-1/2} \tag{75}$$

as follows immedeately from (74) and the trivial estimate $|\xi - 1|_{v_0} \gg e^{-1}$.

It is worth mentioning that

$$\varphi = \left(\sqrt[e]{x_0 - \alpha} - \sqrt[e]{x_0 - \beta}\right)^e \tag{76}$$

when the roots $\sqrt[e]{x_0 - \alpha}$ and $\sqrt[e]{x_0 - \beta}$ are appropriately defined.

## 3 Estimating $h(\varphi)$ and $\mathcal{D}_{\mathbf{L}}$

We have either

$$h(\varphi) \ll h(x_0) \ll eh(\varphi) \tag{77}$$

or $h(x_0) \ll e\left(\tilde{h} + e\right)$, which is much better than (72). Indeed, by the definition of $\varphi$ we have $h(\varphi) \ll h(x_0) + \tilde{h} + e$ whence either $h(x_0) \leq \tilde{h} + e$ or $h(\varphi) \ll h(x_0)$. Further, rewrite (76) as $t_0^e + \beta - \alpha = \left(\sqrt[e]{\varphi} + t_0\right)^e$, where $t_0 = \sqrt[e]{x_0 - \beta}$. Thus, $t_0$ is a root of a polynomial of degree $e-1$ and height $O(h(\varphi) + \tilde{h} + e)$. By Proposition 1.4.1 we have $h(t_0) \ll h(\varphi) + \tilde{h} + e$, whence $h(x_0) \ll eh(t_0) + \tilde{h} \ll e\left(h(\varphi) + \tilde{h} + e\right)$. If $h(\varphi) \geq \tilde{h} + e$ then $h(x_0) \ll eh(\varphi)$, and if $h(\varphi) \leq \tilde{h} + e$ then we have $h(x_0) \ll e\left(\tilde{h} + e\right)$.

We also have to estimate the absolute discriminant $\mathcal{D}_{\mathbf{L}}$. By Proposition 2.4.3

$$d^{-1} \log \mathcal{N}\left(\mathcal{D}_{\mathbf{K}(\alpha)/\mathbf{K}}\right) \leq 2\nu_\alpha^2 h(\alpha) + \nu_\alpha \log \nu_\alpha \leq 2\nu_\alpha \kappa_\alpha\,, \tag{78}$$

$$(d\nu_\alpha)^{-1} \log \mathcal{N}_{\mathbf{K}(\alpha)}\left(\mathcal{D}_{\mathbf{K}(\alpha,\beta)/\mathbf{K}(\alpha)}\right) \leq 2\nu_\beta^2 h(\beta) + \nu_\beta \log \nu_\beta \leq 2\nu_\beta \kappa_\beta\,. \tag{79}$$

Indeed, $\mu_\alpha \geq n/2$, as we have seen in Section 3, and on the other hand $\mu_\alpha \nu_\alpha \leq \deg D(X) \leq (2n-2)m$. Therefore $\nu_\alpha \leq 4m$, whence $\nu_\alpha h(\alpha) + \log \nu_\alpha \leq \kappa_\alpha$. This proves (78); in the same manner one obtains (79).

Further, if a non-archimedean place $v$ of the field $\mathbf{L}_0 = \mathbf{K}(\alpha, \beta)$ is ramified in $\mathbf{L}$ then either $p(v) \leq n$, or $e$ does not divide one of the numbers $\mathrm{Ord}_v(x_0 - \alpha)$ and $\mathrm{Ord}_v(x_0 - \beta)$. By Lemma 3.1

$$\Sigma(\alpha) := d_{\mathbf{L}_0}^{-1} \sum_{\substack{v \notin S \text{ and } e \text{ does not} \\ \text{divide } \mathrm{Ord}_v(x_0 - \alpha)}} \log \mathcal{N}_{\mathbf{L}_0} v \leq 12 e \kappa_\alpha\,,$$

and similarly one defines and estimates $\Sigma(\beta)$. Denote by $S_0$ the set of places of $\mathbf{L}_0$ above $S$. Then

$$\begin{aligned} d_{\mathbf{L}_0}^{-1} \log \mathcal{N}_{\mathbf{L}_0}\left(\mathcal{D}_{\mathbf{L}/\mathbf{L}_0}\right) &\leq (e-1) d_{\mathbf{L}_0}^{-1} \sum_{v \text{ is ramified in } \mathbf{L}} \log \mathcal{N}_{\mathbf{L}_0} v + e^2 \log e \\ &\leq e\left(d_{\mathbf{L}_0}^{-1} \sum_{v \in S_0} \log \mathcal{N}_{\mathbf{L}_0} v + \Sigma(\alpha) + \Sigma(\beta)\right) + e^2 \log e \\ &\leq e\left(d^{-1} \sum_{v \in S} \log \mathcal{N} v + 12\kappa_\alpha + 12\kappa_\beta\right). \end{aligned}$$

Finally

$$\begin{aligned} \mathcal{D}_{\mathbf{L}} &\leq \mathcal{D}^{e\nu_\alpha \nu_\beta} \mathcal{N}\left(\mathcal{D}_{\mathbf{K}(\alpha)/\mathbf{K}}\right)^{e\nu_\beta} \mathcal{N}_{\mathbf{K}(\alpha)}\left(\mathcal{D}_{\mathbf{L}_0/\mathbf{K}(\alpha)}\right)^e \mathcal{N}_{\mathbf{L}_0}\left(\mathcal{D}_{\mathbf{L}/\mathbf{L}_0}\right) \\ &\leq \left(\mathcal{D}\left(\prod_{v \in S} \mathcal{N} v\right) \exp\left(14 d(\kappa_\alpha + \kappa_\beta)\right)\right)^{e\nu_\alpha \nu_\beta}. \end{aligned} \tag{80}$$

## 4 $\varphi$ is almost a unit

For any place $v$ of the field $\mathbf{L}$ put

$$\sigma_v = \max(1, |\alpha|_v, |\beta|_v), \qquad \varrho_v = |\alpha - \beta|_v.$$

Let $S_1$ be the set of places of $\mathbf{L}$ above $S$. Then $s_1 := |S_1| \le se\nu_\alpha\nu_\beta$ and for any $v \notin S_1$ we have

$$\varrho_v^e \sigma_v^{1-e} \le |\varphi|_v \le \sigma_v.$$

Indeed, $|\varphi|_v \le \sigma_v$ by (76). Further, we have $\varphi_0 \cdots \varphi_{e-1} = (\beta - \alpha)^e$, where $\varphi_k = (x_0 - \beta)(\xi^k\theta - 1)^e$. Since $\varphi = \varphi_0$ and for any $k$ we have $|\varphi_k|_v \le \sigma_v$, we obtain $|\varphi|_v \ge \varrho_v^e \sigma_v^{1-e}$.

Let $\eta_1, \ldots, \eta_{s_1-1}$ be a fundamental system of $S_1$-units of the field $\mathbf{L}$ constructed in Proposition 4.1.1. Then

$$h^*(\eta_1) \cdots h^*(\eta_{s_1-1}) \le s_1^{2.1s_1} R(S_1). \tag{81}$$

By Proposition 4.1.3 there exists a unit $\eta = \eta_1^{b_1} \cdots \eta_{s_1-1}^{b_{s_1-1}}$ such that

$$d_{\mathbf{L}}^{-1} \sum_{\substack{v \in S_1 \\ v \ne v_0}} d_v(\mathbf{L}) \left|\log |\psi|_v\right| \ll s_1^{2.1s_1} R(S_1),$$

where $\psi = \varphi\eta^{-1}$. We shall show that $\psi$ has a bounded heght.

Obviously, $|\psi|_v = |\varphi|_v$ for any $v \notin S_1$. Therefore

$$\begin{aligned} h(\psi) &= (2d_{\mathrm{L}})^{-1} \sum_v d_v(\mathbf{L}) \left|\log |\psi|_v\right| \\ &\le d_{\mathbf{L}}^{-1} \sum_{v \ne v_0} d_v(\mathbf{L}) \left|\log |\psi|_v\right| \\ &\le d_{\mathbf{L}}^{-1} \sum_{\substack{v \in S_1 \\ v \ne v_0}} d_v(\mathbf{L}) \left|\log |\psi|_v\right| + d_{\mathbf{L}}^{-1} \sum_{v \notin S_1} d_v(\mathbf{L}) \left|\log |\varphi|_v\right| \\ &\ll s_1^{2.1s_1} R(S_1) + d_{\mathbf{L}}^{-1} \sum_v e(\log \sigma_v + |\log \varrho_v|) \\ &\ll s_1^{2.1s_1} R(S_1) + e\tilde{h}. \end{aligned} \tag{82}$$

In addition, estimate $R(S_1)$. Corollary 4.1.5 and (80) yield

$$R(S_1) \ll \left( \mathcal{D}^{0.51} \left( \prod_{v \in S} \mathcal{N}v \right)^{0.52} \exp\left(7.2d(\kappa_\alpha + \kappa_\beta)\right) \right)^{e\nu_\alpha\nu_\beta}. \tag{83}$$

## 5 Construction of $\varphi'$ and $\mathbf{L}'$

In the case (A) let $\xi'$ be a primitive $e$-th root of unity distinct from $\xi$ (here we use the assumption $e \ge 3$). In the case (B) put $\xi' = -1$. Put

$$\theta' = \begin{cases} \xi' \sqrt[e]{\frac{x_0 - \alpha}{x_0 - \beta}} & \text{in the case (A)}, \\ \xi' \sqrt[e]{\frac{x_0 - \gamma}{x_0 - \beta}} & \text{in the case (B)}, \end{cases} \qquad \mathbf{L}' = \begin{cases} \mathbf{K}(\alpha, \beta, \theta') & \text{in the case (A)}, \\ \mathbf{K}(\gamma, \beta, \theta') & \text{in the case (B)}, \end{cases}$$

where the root $\sqrt[e]{\frac{x_0-\gamma}{x_0-\beta}}$ is defined as the sum of the series $1+\sum_{k=1}^{\infty}\binom{1/e}{k}\left(\frac{\beta-\gamma}{x_0-\beta}\right)^k$. Defining in the obvious manner $\varphi'$, $\psi'$, $\eta'$, etc., we refere to the analogues of (75)–(83) as (75′)–(83′). For example, in the case (B)

$$\mathcal{D}_{\mathbf{L}'} \le \left(\mathcal{D}\left(\prod_{v\in S}\mathcal{N}v\right)\exp\left(14d(\kappa_\gamma+\kappa_\beta)\right)\right)^{e\nu_\gamma\nu_\beta}. \tag{80'}$$

and in the case (A) $\kappa_\gamma$ and $\nu_\gamma$ should be replaced here (and everywhere below) by $\kappa_\alpha$ and $\nu_\alpha$.

## 6 Estimating $B$

Put $B=\max\left(3,\, b_1,\dots,b_{s_1-1},\, b'_1,\dots,b'_{s'_1-1}\right)$. We shall see that either $h(x_0)$ can be estimated much better than in (72) or

$$\begin{aligned} h(x_0) &\le c_1R(S_1)B\,, && (84)\\ B &\le c_1\log|x_0|_{v_0}\,, && (85)\end{aligned}$$

where $c_1=\exp(6.7\Theta)$. Indeed, $h(\varphi)\le h(\psi)+h(\eta)+O(1)$, where $h(\eta)\ll s_1^{2.1s_1}R(S_1)B$ by Corollary 4.1.2. Combining this with (77) and (82), we obtain $h(x_0)\ll c_1R(S_1)B+e^2\tilde{h}$. We may assume $c_1R(S_1)B\ge e^2\tilde{h}$ (otherwise it would be $h(x_0)\ll e^2\tilde{h}$, better than (72)). Therefore $h(x_0)\le c_1R(S_1)B$.

**Remark 5.2** Here and below we may write $\le$ instead of $\ll$ because the implicit constant is absorbed by the $O(1)$-term of $\Theta$.

Further, by Corollary 4.1.2

$$\max(b_1,\dots,b_{s_1-1})\ll s_1^{2.1s_1}h(\eta)\le s_1^{2.1s_1}\left(h(\varphi)+h(\psi)+O(1)\right),$$

and similarly for $\max\left(b'_1,\dots,b'_{s'_1-1}\right)$. Combining this with (73), (82), (82′), (77) and (77′), we obtain $B\le c_1\left(\log|x_0|_{v_0}+c_2\right)$, where $c_2=\tilde{h}+c_1\left(R(S_1)+R(S'_1)\right)$. If $\log|x_0|_{v_0}\le c_2$ then $B\le c_1\log|x_0|_{v_0}$, as desired. If $\log|x_0|_{v_0}\ge c_2$ then $h(x_0)\le sc_2$; using (83) and (83′), we estimate $h(x_0)$ better than in (72).

## 7 Use of Baker's theory

In the sequel we can assume the inequality $\varphi/\varphi'\ne(\xi-1)^e/(\xi'-1)^e$. Indeed, the equality $\varphi/\varphi'=(\xi-1)^e/(\xi'-1)^e$ is a non-trivial algebraic relation involving $x_0$, which yields an estimate for $h_x(P)$ much better than (72).

Put $\eta_0=\left(\frac{\xi'-1}{\xi-1}\right)^e\frac{\psi}{\psi'}$. Using (75) and (85), we obtain

$$0<\left|\left(\frac{\xi'-1}{\xi-1}\right)^e\frac{\varphi}{\varphi'}-1\right|_{v_0}=\left|\eta_0\eta(\eta')^{-1}-1\right|_{v_0}\ll|x_0|_{v_0}^{-1/2}\le\exp(-c_1^{-1}B). \tag{86}$$

Put $r=s_1+s'_1-2\le 2se\nu_\beta(\nu_\alpha+\nu_\gamma)$ and write $\eta_{s_1},\dots,\eta_r$ and $b_{s_1},\dots,b_r$ instead of $\eta'_1,\dots,\eta'_{s'_1-1}$ and $-b'_1,\dots,-b'_{s'_1-1}$, respectively. Then

$$0<\left|\eta_0\eta_1^{b_1}\cdots\eta_r^{b_r}-1\right|_{v_0}\le\exp(-c_1^{-1}B),$$

$$
\begin{aligned}
[\mathbf{Q}(\eta_0, \eta_1, \ldots, \eta_r) : \mathbf{Q}] &\leq \delta \ll sN^5, \\
h^*(\eta_0) &\leq c_1\widetilde{R} + O(e\widetilde{h}), \\
h^*(\eta_1)\cdots h^*(\eta_r) &\leq c_1^2 R(S_1)R(S_1'), \\
h' := \max(h(\eta_1), \ldots, h(\eta_r), 3) &\leq c_1\widetilde{R}.
\end{aligned}
$$

where $h^*(\ldots)$ is defined in Proposition 4.1.1 and $\widetilde{R} = \max(R(S_1), R(S_1'), 3)$. By Proposition 4.3.1

$$
\begin{aligned}
B &\leq c_1\widehat{p}(v_0)^{\delta-0.5}h^*(\eta_0)h^*(\eta_1)\cdots h^*(\eta_r)(\log h')\log(c_1h')\exp(3r\log(r\delta) + O(r+\log\delta)) \\
&\leq \widehat{p}(S)^{\delta-0.3}\left(\widetilde{R}+\widetilde{h}\right)R(S_1)R(S_1')\left(\log^2\widetilde{R}\right)\exp(123\Theta).
\end{aligned}
$$

By (84), (83) and (83′)

$$
\begin{aligned}
h(x_0) &\leq \widehat{p}(S)^{\delta-0.3}\left(\widetilde{R}+\widetilde{h}\right)R(S_1)^2R(S_1')\left(\log^2\widetilde{R}\right)\exp(130\Theta) \\
&\leq \widehat{p}(S)^{\delta}\left(\mathcal{D}\prod_{v\in S}\mathcal{N}v\right)^{2.1\Upsilon}\exp\left(130\Theta + ed\left(22\nu_\alpha\nu_\beta(\kappa_\alpha+\kappa_\beta) + 15\nu_\gamma\nu_\beta(\kappa_\gamma+\kappa_\beta)\right)\right).
\end{aligned}
$$

Theorem 5.1 is proved.

# 6 Proof of Theorem 1.2

The relation (7) implies that one of the following conditions holds:

(a) there exist $\alpha, \beta \in \overline{\mathbf{Q}}$ such that $e_\alpha \geq 3$ and $e_\beta \geq 2$;

(b) there exist $\alpha, \beta, \gamma \in \overline{\mathbf{Q}}$ such that $e_\alpha = e_\beta = e_\gamma = 2$.

We can split (a) into three subcases:

(a1) there exist $\alpha, \beta \in \overline{\mathbf{Q}}$ such that $e_\alpha = e_\beta \geq 3$;

(a2) there exist $\alpha, \beta \in \mathbf{K}$ such that $e_\alpha \geq 3$ and $e_\beta \geq 2$;

(a3) there exist $\alpha \in \mathbf{K}$ and $\beta \in \overline{\mathbf{Q}}$ such that $[\mathbf{K}(\beta):\mathbf{K}] \leq 2$ and $e_\alpha \geq 3$, $e_\beta = 2$.

Indeed, suppose that $\alpha \notin \mathbf{K}$ and $\alpha_1 = \alpha, \alpha_2, \ldots, \alpha_\nu$ are the conjugates of $\alpha$ over $\mathbf{K}$. Then we redefine $\beta$, putting $\beta = \alpha_2$, and obtain the case (a1). In a similar manner the case

$$\beta \notin \mathbf{K}, \quad e_\beta \geq 3$$

can be reduced to (a1), and the case

$$[\mathbf{K}(\beta):\mathbf{K}] \geq 3, \quad e_\beta = 2$$

can be reduced to (b).

## 6.1 Cases (a1) and (b)

Let $\alpha_1 = \alpha$, $\alpha_2, \ldots, \alpha_{\nu_\alpha}$ be the conjugates of $\alpha$ over $\mathbf{K}$. All them are roots of $D(X)$ of order $\mu_\alpha$. Therefore

$$\begin{array}{rclcl} \nu_\alpha \mu_\alpha & \le & \deg D(X) & \le & 2mn\,, \\ \nu_\alpha u_\alpha & \le & \deg g_0(X) & \le & m\,. \end{array} \tag{87}$$

Since $\mu_\alpha \ge n/2$, as we have seen in Section 3, we obtain $\nu_\alpha \le 4m$. (We have already used this in the previous section.) Similarly, $\nu_\beta, \nu_\gamma \le 4m$. Furthermore, $h(\alpha_1) = \ldots = h(\alpha_{\nu_\alpha})$. Hence, by Proposition 1.4.1

$$\begin{array}{rclcl} \nu_\alpha \mu_\alpha h(\alpha) & \le & h(D) + \log(2mn) & \le & 2nh(f) + O(n \log N)\,, \\ \nu_\alpha u_\alpha h(\alpha) & \le & h(g_0) + \log(m+1) & \le & h(f) + O(\log N)\,. \end{array} \tag{88}$$

Combining (87)–(88), we obtain

$$\nu_\alpha \kappa_\alpha \le 6n^3 m \Big(h(f) + O(n + \log N)\Big)\,. \tag{89}$$

In the same manner one estimates $\nu_\beta \kappa_\beta$ and $\nu_\gamma \kappa_\gamma$. Hence

$$\nu_\alpha \nu_\beta (\kappa_\alpha + \kappa_\beta) \le 4m(\nu_\alpha \kappa_\alpha + \nu_\beta \kappa_\beta) \le 48n^3 m^2 \Big(h(f) + O(n + \log N)\Big)\,, \tag{90}$$

and similarly one estimates $\nu_\gamma \nu_\beta (\kappa_\gamma + \kappa_\beta)$. Further,

$$\Upsilon \le 16m^2 n\,, \tag{91}$$

$$\delta \le \begin{cases} 16dm^2n^2 & \text{in the case (A)}, \\ 256dm^3 & \text{in the case (B)}, \end{cases} \tag{92}$$

$$\Theta \le 16m^2 ns(\log(Ns) + O(1))\,. \tag{93}$$

Substituting the estimates (90)–(93) to (72), we obtain

$$h_x(P) \le \left( \widehat{p}(S)^{d \max(n^2, 8m)} \left( \mathcal{D} \prod_{v \in S} \mathcal{N}v \right)^{2.1n} \exp(130\Omega_1) \right)^{16m^2}\,, \tag{94}$$

where $\Omega_1 = sn \log(Ns) + O(sn) + 3dn^4(h(f) + O(n + \log N))$. As one can easily see, (94) is better than (3). This completes the proof in the cases (a1) and (b).

## 6.2 Case (a2)

Put $e = e_\beta$ and $t = \sqrt[e]{x - \beta}$. Let $\widetilde{C} \xrightarrow{\varphi} C$ be the covering corresponding to the embedding $\overline{\mathbf{Q}}(C) \hookrightarrow \overline{\mathbf{Q}}(C)(t)$. The curve $\widetilde{C}$ is defined over $\mathbf{K}$ and $\mathbf{K}(\widetilde{C}) = \mathbf{K}(y,t)$. We have $\tilde{f}(t,y) = 0$, where $\tilde{f}(T,Y) = f(\beta + T^e, Y)$. In particular,

$$\widetilde{m} := \deg_T \tilde{f} = me\,, \quad \tilde{n} := \deg_Y \tilde{f} = n\,, \quad h(\tilde{f}) \le h(f) + mh(\beta)\,.$$

The coverings $\widetilde{C} \xrightarrow{\varphi} C \xrightarrow{x} \mathbf{P}^1$ and $\widetilde{C} \xrightarrow{t} \mathbf{P}^1$ have the following two properties.

(i) For any $P \in C(x, \mathbf{K}, S)$ and $\tilde{P} \in \varphi^{-1}(P)$ there exists an extension $\widetilde{\mathbf{K}}$ of $\mathbf{K}$ such that $\left[\widetilde{\mathbf{K}}:\mathbf{K}\right] \leq e$ and

$$\begin{aligned}
\mathcal{D}_{\widetilde{\mathbf{K}}} &\leq \left(\mathcal{D}\left(\prod_{v\in S}\mathcal{N}v\right)\exp\left(72dn^3m\left(h(f)+O(n+\log N)\right)\right)\right)^e, && (95)\\
\tilde{P} &\in \tilde{C}(t, \widetilde{\mathbf{K}}, \tilde{S}), && (96)
\end{aligned}$$

where $\tilde{S}$ is the set of places of $\widetilde{\mathbf{K}}$ above $S$.

(ii) Put

$$\tilde{\alpha} = \sqrt[\tilde{e}]{\alpha - \beta}, \quad \tilde{\beta} = \xi\tilde{\alpha}, \quad \tilde{e} = e_\alpha(x)$$

where $\xi$ is a primitive $\tilde{e}$-th root of unity. Then $e_{\tilde{\alpha}}(t)$ and $e_{\tilde{\beta}}(t)$ are divisible by $\tilde{e}$.

**Proof of (i)** We have (96) with $\widetilde{\mathbf{K}} = \mathbf{K}\left(\sqrt[e]{x(P)-\beta}\right)$ for an appropriate definition of the root. By Proposition 4.2.3, Lemma 3.1 and (89).

$$\begin{aligned}
d^{-1}\log\mathcal{N}\left(\mathcal{D}_{\widetilde{\mathbf{K}}}\right) &\leq ed^{-1}\sum_{v \text{ is ramified in } \widetilde{\mathbf{K}}}\log\mathcal{N}v + e\log e\\
&\leq ed^{-1}\left(\sum_{v\in S} + \sum_{p(v)\leq e} + \sum_{\substack{v\notin S \text{ and } e \text{ does not}\\ \text{divide } \mathrm{ord}_v(x(P)-\alpha)}}\right)\log\mathcal{N}v + e\log e\\
&\leq e\left(d^{-1}\sum_{v\in S}\log\mathcal{N}v + 12\kappa_\alpha\right)\\
&\leq e\left(d^{-1}\sum_{v\in S}\log\mathcal{N}v + 72n^3m\left(h(f)+O(n+\log N)\right)\right),
\end{aligned}$$

which yields (95).

**Proof of (ii)** Below *divisor* means *divisor on* $\tilde{C}$. We have $t^e - \gamma = x - \alpha$, where $\gamma = \alpha - \beta$. Write the principal divisor $(t^e - \gamma)$ as the difference of two positive divisors with disjoint supports:

$$(t^e - \gamma) = (t^e - \gamma)_0 - (t^e - \gamma)_\infty .$$

Then $(t^e - \gamma)_0 = (x - \alpha)_0$, the latter divisor being divisible by $\tilde{e}$. (We say that a divisor $D$ is *divisible* by an integer $l$ if $D = lD'$ for some divisor $D'$.) On the other hand,

$$(t^e - \gamma)_0 = (t - \tilde{\alpha})_0 + (t - \xi\tilde{\alpha})_0 + \cdots + \left(t - \xi^{e-1}\tilde{\alpha}\right)_0 . \quad (97)$$

Since the divisor in the left-hand side of (97) is divisible by $\tilde{e}$ and the divisors in the right-hand side have pairwise disjoint supports, each of the latter is divisible by $\tilde{e}$. In particular, $\tilde{e}$ divides $e_{\tilde{\alpha}}(t)$ and $e_{\tilde{\beta}}(t)$, as desired.

We have

$$\begin{aligned}
h(\tilde{\alpha}) &\leq e^{-1}(h(\alpha) + h(\beta) + O(1)) &&\leq 4e^{-1}n(h(f) + O(\log N)),\\
\tilde{\nu}_{\tilde{\alpha}} &:= \left[\widetilde{\mathbf{K}}(\tilde{\alpha}):\widetilde{\mathbf{K}}\right] &&\leq e,
\end{aligned}$$

and similarly one estimates $h\left(\widetilde{\beta}\right)$ and $\widetilde{\nu}_{\widetilde{\beta}}$. Also,

$$\begin{array}{rcll}
\widetilde{\Upsilon} & := & \widetilde{e}\widetilde{\nu}_{\widetilde{\alpha}}\widetilde{\nu}_{\widetilde{\beta}} & \leq \; e^2\widetilde{e}\,, \\
\widetilde{\delta} & := & d_{\widetilde{\mathbf{K}}}\widetilde{e}^2\widetilde{\nu}_{\widetilde{\alpha}}\widetilde{\nu}_{\widetilde{\beta}} & \leq \; de^3\widetilde{e}^2\,, \\
\widetilde{s} & := & \left|\widetilde{S}\right| & \leq \; se\,, \\
\widetilde{\Theta} & := & \widetilde{s}\widetilde{\Upsilon}\left(\log\left(\widetilde{N}\widetilde{s}\right)+O(1)\right) & \leq \; 3se^3\widetilde{e}\left(\log(NS)+O(1)\right)\,.
\end{array}$$

Furthermore, defining $\widetilde{u}_{\widetilde{\alpha}}$, $\widetilde{\mu}_{\widetilde{\alpha}}$, $\widetilde{u}_{\widetilde{\beta}}$ and $\widetilde{\mu}_{\widetilde{\beta}}$ in the obvious manner, one easily finds that

$$\widetilde{u}_{\widetilde{\alpha}}=\widetilde{u}_{\widetilde{\beta}}=u_\alpha\,,\qquad \widetilde{\mu}_{\widetilde{\alpha}}=\widetilde{\mu}_{\widetilde{\beta}}=\mu_\alpha\,. \tag{98}$$

The rest of the argument splits into two cases: $e_\beta\geq 3$ and $e_\beta=2$. In the first case we may suppose that

$$\mu_\alpha+nu_\alpha\leq\mu_\beta+nu_\beta\,, \tag{99}$$

interchanging $\alpha$ and $\beta$ if necessary. Defining in the obvious manner $\widetilde{\kappa}_{\widetilde{\alpha}}$ and using (87), (88) and (99), we obtain

$$\begin{aligned}
\widetilde{\kappa}_{\widetilde{\alpha}} &= \widetilde{\nu}_{\widetilde{\alpha}}\widetilde{n}^2\left(\widetilde{\mu}_{\widetilde{\alpha}}+\widetilde{n}\widetilde{u}_{\widetilde{\alpha}}\right)\left(h\left(\widetilde{f}\right)+\widetilde{m}h(\widetilde{\alpha})+O(\widetilde{n}+\log N)\right) \\
&\leq n^2\left(\mu_\alpha+nu_\alpha\right)\left(h\left(f\right)+m\left(h(\alpha)+2h(\beta)\right)+O(N)\right) && (100)\\
&\leq n^2m\left(3nh(f)+\left(\mu_\alpha+nu_\alpha\right)h(\alpha)+2\left(\mu_\beta+nu_\beta\right)h(\beta)+O(nN)\right) && (101)\\
&\leq 12n^3m\left(h(f)+O(N)\right)\,. && (102)
\end{aligned}$$

In the similar way one defines and estimates $\widetilde{\kappa}_{\widetilde{\beta}}$. By Theorem 5.1

$$\begin{aligned}
h_x(P) &\leq e\left(h_t\left(\widetilde{P}\right)+h(\beta)+O(1)\right) \\
&\leq \widehat{p}\left(\widetilde{S}\right)^{\widetilde{\delta}}\left(\mathcal{D}_{\widetilde{\mathbf{K}}}\prod_{v\in\widetilde{S}}\mathcal{N}_{\widetilde{\mathbf{K}}}v\right)^{2.1\widetilde{\Upsilon}}\exp\left(130\widetilde{\Theta}+37\widetilde{e}d_{\widetilde{\mathbf{K}}}\widetilde{\nu}_{\widetilde{\alpha}}\widetilde{\nu}_{\widetilde{\beta}}(\widetilde{\kappa}_{\widetilde{\alpha}}+\widetilde{\kappa}_{\widetilde{\beta}})\right) \\
&\leq \left(\widehat{p}(S)^{d\widetilde{e}}\mathcal{D}^{2.1}\left(\prod\mathcal{N}v\right)^{3.1}\exp\left(400\Omega_2\right)\right)^{e^3\widetilde{e}} && (103)\\
&\leq \left(\widehat{p}(S)^{dn}\mathcal{D}^{2.1}\left(\prod\mathcal{N}v\right)^{3.1}\exp\left(400\Omega_2\right)\right)^{n^4} && (104)
\end{aligned}$$

with $\Omega_2=s\log(NS)+O(s)+1.5dn^3m\left(h(f)+O(N)\right)$. This is better than (3).

Now suppose that $e=e_\beta=2$. Then we cannot assume (99) anymore. Instead, we shall use the estimates

$$\mu_\alpha+nu_\alpha\leq 3mn\,,\qquad h(\alpha)\,,\,h(\beta)\leq 4h(f)+O(\log N)\,,$$

which can be deduced from (87), (88) and (49).

We still have (100), but instead of (101) we obtain

$$\widetilde{\kappa}_{\widetilde{\alpha}}\leq 40n^3m\left(mh(f)+O\left(n+m\log N\right)\right)\,.$$

Therefore instead of (103) we have

$$\begin{aligned}
h_x(P) &\leq \left(\widehat{p}(S)^{d\widetilde{e}}\mathcal{D}^{2.1}\left(\prod\mathcal{N}v\right)^{3.1}\exp\left(399\Omega_3\right)\right)^{e^3\widetilde{e}} \\
&\leq \left(\widehat{p}(S)^{dn}\mathcal{D}^{2.1}\left(\prod\mathcal{N}v\right)^{3.1}\exp\left(399\Omega_3\right)\right)^{8n} && (105)
\end{aligned}$$

with $\Omega_3 = s\log(NS) + O(s) + 4dn^3m\left(mh(f) + O(n + m\log N)\right)$. Again we obtain an estimate better than (3).

## 6.3 Case (a3)

We have

$$\mathcal{D}_{\mathbf{K}(\beta)} \le \mathcal{D}^2\exp(2d(h(\beta) + O(1)) \le \mathcal{D}^2\exp(8d(h(f) + O(1)).$$

Applying (105) with the field $\mathbf{K}(\beta)$ instead of $\mathbf{K}$, we obtain

$$h_x(P) \le \left(\hat{p}(S)^{dn}\mathcal{D}^{2.1}\left(\prod \mathcal{N}v\right)^{3.1}\exp(400\Omega_3)\right)^{16n},$$

again better than (3).

Theorem 1.2 is proved.

# References


[1] **A. Baker**, Linear forms in the logarithms of algebraic numbers I, *Mathematica* **13** (1966), 204–216; II, ibid. **14** (1967), 102–107; III, ibid. **14** (1967); 220–224; IV, ibid. **15** (1968), 204–216.

[2] **A. Baker**, Contributions to the theory of Diophantine equations. I, II, *Phil. Trans. R. Soc. London Ser.A* **263** (1967–68), 173–208.

[3] **A. Baker**, The Diophantine equation $y^2 = ax^3 + bx^2 + cx + d$, *J. London Math. Soc.* **43** (1968), 1-9.

[4] **A. Baker**, Bounds for the solutions of the hyperelliptic equations, *Proc. Camb. Phil. Soc.*, **65** (1969), 439–444.

[5] **A. Baker, J. Coates**, Integer points on curves of genus 1, *Proc. Camb. Phil. Soc.* **67** (1970), 592–602.

[6] **B. Brindza**, On $S$-integral solutions of the equation $y^m = f(x)$, *Acta Math. Hungar.* **44** (1984), 133–139.

[7] **Y. Bugead, K. Győry**, Bounds for the solutions of unit equations, *Acta Arithm.*, to appear.

[8] **Yu. Bilu** (Belotserkovski), Effective analysis of a new class of Diophantine equations (Russian), *Vestsi Akad. Navuk BSSR, Ser. Fiz.-Math. Navuk*, 1988, *no.* 6, 34–39, 125

[9] **Yu. Bilu**, *Effective analysis of integral points on algebraic curves*, thesis, Beer Sheva, 1993.

[10] **Yu. Bilu**, *Solving superelliptic Diophantine equations by the method of Gelfond–Baker*, preprint 94-09, Mathématiques Stochastiques, Univ. Bordeaux 2, 1994.

[11] **Yu. Bilu**, Effective analysis of integral points on algebraic curves, *Israel J. Math* **90** (1995), 235–252.

[12] **J.W.S. Cassels, A. Fröhlich** (eds.), *Algebraic Number Theory*, Academic Press, 1967.

[13] **J. Coates**, Construction of rational functions on a curve, *Proc. Camb. Phil. Soc.* **68** (1970), 105–123.

[14] **E. Dobrowolski**, On a question of Lehmer and the number of irreducible factors of a polynomial, *Acta Arithm.* **34** (1979), 391–401.

[15] **R. Dvornicich, U. Zannier**, Fields containing values of algebraic functions, Ann. Sc. Norm. Sup. Pisa Cl. Sci. Ser. IV **21** (1994), 421–443.

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

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

[18] **G. Faltings**, Eindlichkeitssätze für abelche Varietäten über Zahlkörpern, *Invent. Math.* **73** (1983), 349–366; Erratum: **75** (1984), p.381.

[19] **A.O. Gelfond**, *Transcendent and Algebraic Numbers* (Russian), Moscow 1952; English trans.: New York, Dover, 1960.

[20] **L. Hajdu**, A quantitative version of Dirichlet's $S$-unit theorem in algebraic number fields, *Publ. Math. Debrecen* **42** (1993), 239–246.

[21] **H. Kleiman**, On the Diophantine equation $f(x,y)=0$, *J. reine und angew. Math.* **286/287** (1976), 124–131.

[22] **S.V. Kotov, L.A. Trellina**, $S$-ganze Punkte auf elliptischen Kurven, *J. reine und angew. Math.* **306** (1979), 28–41.

[23] **S. Lang**, *Fundamentals of Diophantine Geometry*, Springer, 1983.

[24] **W.J. LeVeque**, On the equation $y^m = f(x)$, *Acta Arithm.* 9 (1964), 209–219.

[25] **A. Pethő**, Beiträge zur Theorie der $S$-Ordnungen, *Acta Math. Acad. Sci. Hungar.* **37** (1981), 51–57.

[26] **D. Poulakis**, Solutions entières de l'équation $Y^m = f(x)$. *Sém. Th. Nom. Bordeaux* **3** (1991), 187–199.

[27] **D. Poulakis**, Points entiers et modèles des courbes algébriques, *Monatsh. Math.* 118 (1994), 111–143.

[28] **W.M. Schmidt**, Eisenstein theorem on power series expansions of algebraic functions, *Acta Arithm.* 56(1990), 161–179.

[29] **W.M. Schmidt**, Integer points on curves of genus 1, *Compositio Math.* **81** (1992), 33–59.

[30] **T.N. Shorey, R. Tijdeman**, *Exponential Diophantine equations*, Cambridge Univ. Press, Cambridge, 1986.

[31] **C.L. Siegel**, Über einige Anwendungen Diophantischer Approximationen, *Abh. Preuss Akad. Wiss. Phys.-Math. Kl.*, 1929, Nr. 1.

[32] **C.L. Siegel**, Abschätzung von Einheiten, *Nachr. Akad. Wiss. Göttingen II. Math.-Phys. Kl.*, 1969, Nr. 9, 71–86.

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

[34] **V.G. Sprindžuk**, The arithmetic structure of integral polynomials and class numbers (Russian), *Trudy Mat. Inst. Steklov* **143** (1977), 152–174; English transl.: Proc. Steklov Inst. Math. 1980, issue 1, 163–186.

[35] **V.G. Sprindžuk**, *Classical Diophantine Equations in Two Unknowns* (Russian), Nauka, Moscow, 1982; English transl.: Lecture Notes in Math. **1559**, Springer, 1994.

[36] **L.A. Trelina**, On $S$-integral solutions of the hyperelliptic equation (Russian), *Dokl. Akad. Nauk BSSR* **22** (1978), 881–884.

[37] **P.M. Voutier**, An Upper Bound for the Size of Integral Solutions to $Y^m = f(X)$, *J. Number Theory* **53** (1995), 247–271.

[38] **M. Waldschmidt**, Minorations de combinaisons linéares de logarithmes de nombres algébriques, *Canadian J. Math.* **45** (1993), 176–224.

[39] **Kunrui Yu**, Linear forms in $p$-adic logarithms III, *Compositio Math.* **91** (1994), 241–276.