Quantitative Siegel's theorem for Galois coverings

YURI F. BILU
Max-Planck-Institut für Mathematik, Gottfried-Claren-Strasse 26, D-53225 Bonn, Germany

Abstract. It is known that Siegel's 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 **K**-rational points in terms of the number field **K**, the set of places S and the defining equation of the curve. Our main tools are Baker's theory of linear forms in 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 **K** and $x \in \mathbf{K}(C)$ non-constant. For any finite set S of places of **K** containing the set S_∞ 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 **K**. The classical theorem of Siegel [38, 26] states that $|C(x, \mathbf{K}, S)| < \infty$ as soon as the genus $\mathbf{g}(C) \geqslant 1$. For curves of genus 2 or more this is covered by a result of Faltings [19], who proved that $|C(\mathbf{K})| < \infty$ when $\mathbf{g}(C) \geqslant 2$, 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 [24, 34, 8] for quantitative improvements of the result of Baker and Coates.

One more general case of effectivity of Siegel's theorem is when $x \colon C \to \mathbf{P}^1$ 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 algebraic closure of **Q**). This was proved by the author [7, Sect. 7], and, independently, by Dvornicich and Zannier [16]. Partial results were obtained by H. Kleiman [23, Cor. (3) of Thm 3] and Poulakis [31, Sect. 2].

In all cases the method of Gelfond–Baker [20, 2] was used, so far the single general effective method in Diophantine analysis. In [42, 37, 36, 8, 10] 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 $\underline{\alpha} = (\alpha_0 : \ldots : \alpha_k) \in \mathbf{P}^k(\overline{\mathbf{Q}})$, we denote by $h(\underline{\alpha})$ its *absolute logarithmic height* (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 coefficient. Also, we define the height function $h_x \colon C(\overline{\mathbf{Q}}) \to \mathbf{R}^{\geqslant 0}$ by $h_x(P) = h_{\mathbf{P}^1}(x(P))$, where $h_{\mathbf{P}^1} \colon \mathbf{P}^1(\overline{\mathbf{Q}}) \to \mathbf{R}^{\geqslant 0}$ 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 nonzero separable polynomial such that $f(x, y) = 0$. (We use lowercase letters x, y, $\ldots$, for rational functions on C and uppercase letters X, Y, $\ldots$, 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

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

We denote by $\mathcal{N} = \mathcal{N}_\mathbf{K} \colon \mathbf{K} \to \mathbf{Q}$ 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 ∞ for archimedean v), and put

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

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

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

$$h_x(P) \leqslant \widehat{p}(S)^{dN_1} \left(\mathcal{D} \prod_{v \in S} \mathcal{N}v \right)^{4N_2} \exp(\Psi), \tag{3}$$

where Ψ *is* $400 s N_2 (\log(Ns) + O(1)) + 600 d N_3 (h(f) + O(N))$*,* N_1 *is* $\max(n^5, 16n^2m^2, 256m^3)$*,* N_2 *is* $\max(n^4, 10m^2n)$*,* N_3 *is* $\max(mn^7, 500m^2n^4)$*.*

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 $k \geqslant 2$ 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 [27]), 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 [41, 43, 11, 30].

Recently P. Voutier [44] 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

$$h(x) \leqslant c(N, d)(\mathcal{D} \exp(dh(f)))^{n^8 m^2/3}(h(f) + \log \mathcal{D} + 1)^{n^6 m^2 d}, \tag{5}$$

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 exponential 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 $\overline{\mathbf{Q}}(x, \sqrt[n]{F(x)})$ 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

$$\max(h(x), h(y)) \leqslant c(N, d)\mathcal{D}^{4N_2} \exp(600dN_3 h(f)),$$

which is better than (5) when n is sufficiently 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 $\mathbf{P}^1(\overline{\mathbf{Q}})$ 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 $x \colon C \to \mathbf{P}^1$ over the point α. Put

$$e_\alpha = \gcd(e_1, \ldots, e_\rho). \tag{6}$$

(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*

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

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

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

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

Then

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

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

Condition (7) first appeared in [6], where it was expressed in a slightly different (but equivalent) form. The simple argument above deducing Theorem 1.1 from Theorem 1.2 goes back to [7]. See also [8, 10].

After having submitted this paper I received a preprint of Poulakis [32] where he also estimates heights of integral points on curves subject to condition (7). He considers only the case $S = S_\infty$, and obtains an estimate

$$h_x(P) \leqslant d^{-1} N^{10^5 d^2 N^{35}} \mathcal{D}^{9N^{13}} \exp(16 \cdot 10^3 d^2 N^{35} h(f)).$$

As one can see, our estimate is considerably sharper.

The approach of Poulakis is different from ours and from the approach of Dvornicich and Zannier [16]. In the last section of his paper Poulakis gives several concrete examples of curves satisfying the condition (7).

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 $|\ldots|_v$ 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 $\mathrm{Ord}_v \colon \mathbf{K}^* \to \mathbf{Z}$ normalizing it so that 1 belongs to the image of Ord_v. In explicit terms $\mathrm{Ord}_v(\alpha) = d_v \log |\alpha|_v / \log \mathcal{N} v$, where $d_v = d_v(\mathbf{K}) = [\mathbf{K}_v \colon \mathbf{Q}_{p(v)}]$ is the local degree of v.

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

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

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

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

$$h(\alpha_1 + \cdots + \alpha_k) \leqslant h(\alpha_1) + \cdots + h(\alpha_k) + \log k, \tag{9}$$

$$h(\alpha_1 \cdots \alpha_k) \leqslant h(\alpha_1) + \cdots + h(\alpha_k), \tag{10}$$

$$h(\alpha^\nu) = |\nu| h(\alpha). \tag{11}$$

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

$$R(X) = g_0(X) D(X), \tag{13}$$

$$\deg R(X) \leqslant (2n - 1)m, \tag{14}$$

$$h(R) \leqslant (2n - 1)h(f) + O(n \log N), \tag{15}$$

$$\deg D(X) \leqslant (2n - 2)m, \tag{16}$$

$$h(D) \leqslant (2n - 2)h(f) + O(n \log N), \tag{17}$$

as follows from the standard determinant representations of the resultant and discriminant. (Of course, the constants in (15) and (17) can be easily made explicit. For example, the $O(\ldots)$-term in (15) can be replaced by $(2n - 1) \log\big((m + 1)(n + 1)\sqrt{n}\big)$, as shown by Schmidt [33, Lemma 4].)

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

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

where Ord_α is the order of vanishing at α. Then $f_0(X, Y) = f(X, Y)$, and we similarly write u and μ instead of u_0 and μ_0, respectively. We have trivially

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

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

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

This, 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, [33, Lemma 3]).

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

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

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

1.5. PLAN OF THE PAPER

In Section 2 we summarize the necessary properties of algebraic power series, in particular, the quantitative Eisenstein theorem due to Schmidt [33] and Dwork–van der Poorten [18].

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_αth 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 [39, 12, 8] and Baker's theory [45, 46].

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.

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 always assume $k_0 \geqslant 0$ 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, $k_0/e \leqslant u \leqslant m$.

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

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

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

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

and for any place $w|v$ of the field **L** *we have*

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

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

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

This theorem is a combination of results of Schmidt [33] and Dwork–van der Poorten [18]. 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 details 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 coefficients in the number field **K** and $\alpha_1, \ldots, \alpha_t$ its roots. For any place v of the field **K** fix a prolongation to $\mathbf{K}(\alpha_1, \ldots, \alpha_t)$ and put

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

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 coefficients β of $f(X, Y)$.

THEOREM 2.2 (Dwork–Robba–Schmidt–van der Poorten). *For a given place v of the field* **K** *put*

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

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

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

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

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 **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* $n < p(v) \leqslant \infty$ *then*

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

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

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

Proof. For the case $n < p(v) \leqslant \infty$ see Schmidt [33]. (As indicated by Schmidt, the case $n < p(v) < \infty$ is a direct consequence of a result of Dwork and Robba [17].)

The case $p(v) \leqslant n$ is due to Dwork and van der Poorten [18]. Let α be a root of $R^*(X)$ with the property $|\alpha|_v = \sigma_v$. Then by [18, Thm 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 [33, 18] only the case $k_0 = 0$ is treated. However, the general case can be easily reduced to the case $k_0 = 0$. Indeed, put

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

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 $\widetilde{f}(X, Y) = X^{k_0 n} f(X, X^{-k_0} Y)$. 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 $k_0 \geqslant 0$ to the case $k_0 = 0$.

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

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).

Proof. This is a result of Schmidt [33, Lemma 2]. Though he considers only the case $k_0 = 0$, his argument plainly works for arbitrary $k_0 \geqslant 0$. 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

$$\widetilde{f}(X,Y) = f(X^e, Y), \qquad \widetilde{y} = \sum_{k=-k_0}^{\infty} a_k x^k, \tag{31}$$

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

$$\widetilde{R}^*(X) = R^*(X^e), \tag{32}$$

$$\widetilde{u} = eu. \tag{33}$$

As follows from (32),

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

Now put

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

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 $\widetilde{y}$, we have (23). Further, Schmidt [33, Lemma 5] showed that

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

Therefore

$$d^{-1} \sum_v d_v \log A_v \leqslant d^{-1} \sum_v d_v \log \sigma_v^{-1} + n d^{-1} \sum_{p(v)\leqslant n} \log c(v,n) + O(n)$$
$$\leqslant (2n-1)h(f) + O(n(n + \log N)),$$

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-archimedean 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) $p(v) \leqslant n$;

(b) there is a root α 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-archimedean v belonging to the cases (a) and (b) above. For (a) the estimate is straightforward:

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

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 ν_i. Therefore

$$d^{-1} \sum_{(b)} \log \mathcal{N} v \leqslant \sum_{i=1}^{s} \nu_i h(\alpha_i)$$

$$\leqslant h(R^*) + \log(1 + \deg R^*)$$

$$\leqslant (2n - 1)h(f) + O(n \log N),$$

where the second inequality is by Proposition 1.4. 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 $\mathcal{K}_1 \subseteq \mathcal{L} = \mathbf{L}((x^{1/e}))$, the field $\mathbf{K}_1$ is a subfield of $\mathbf{L}$. It remains to prove that $[\mathbf{L} : \mathbf{K}] \leqslant e$.

By [25, Prop. 2.12], there exists a primitive element t of the field $\mathcal{K}_1$, such that t^e is a primitive element of $\mathcal{K}_1((x))$. Write

$$t^e = \lambda_1 x + \lambda_2 x^2 + \cdots,$$

where $\lambda_1, \lambda_2, \ldots \in \mathbf{K}_1$ and $\lambda_1 \neq 0$. By Hensel's lemma, there exists $t_1 \in \mathcal{K}_1$ such that $t_1^e = \lambda_1 x$. Since t_1 is a primitive element of $\mathcal{K}_1$, we have

$$y = \sum_{k=-k_0}^{\infty} b_k t_1^k = \sum_{k=-k_0}^{\infty} b_k (\lambda_1^{1/e})^k x^{k/e}$$

with $b_k \in \mathbf{K}_1$. Therefore $\mathbf{L} \subseteq \mathbf{K}_1(\lambda_1^{1/e})$. The proof is complete.

Put

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

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*

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

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

   Proof. Put

$$\begin{aligned}
\mathbf{L}_0 &= \mathbf{K}(a_{-k_0}, \ldots, a_\kappa), \qquad \mathcal{L}_0 = \mathbf{L}_0((x^{1/e})), \\
\delta &= [\mathbf{L} \colon \mathbf{L}_0] = [\mathcal{L} \colon \mathcal{L}_0].
\end{aligned} \tag{40}$$

Clearly, $\mathcal{L}_0(y) = \mathcal{L}$. Let $\varphi(Y) = \varphi_0 Y^\delta + \cdots$ be a minimal polynomial of y over the ring $\mathcal{R} = \mathbf{L}_0[[x^{1/e}]]$ (at least one of its coefficients 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 $f(x, Y)^{[\mathbf{L} \colon \mathbf{K}_1]}$ in the ring $\mathbf{K}[[x]](Y)$.

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

$$\mathcal{N}_{\mathcal{L}_0/\mathbf{K}((x))}\Delta(x) \mid D(x)^{[\mathbf{L} \colon \mathbf{K}_1]} \tag{41}$$

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

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

Comparing Ord_x of the both sides of (41), we obtain

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

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

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

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 [40, Thm 2].

PROPOSITION 2.4.3 (Silverman). *Let $\underline{\alpha} = (\alpha_0 \colon \ldots \colon \alpha_k)$ be a point in $\mathbf{P}^k(\overline{\mathbf{Q}})$, and $[\mathbf{K}(\underline{\alpha}) : \mathbf{K}] = \nu$. Then*

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

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

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

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*

$$k(q) \leqslant \kappa/(q - 1) - 1. \tag{44}$$

Proof. It is very similar to the proof of the fist part of Lemma 2.4, with the field $\mathcal{L}_0$ replaced by $\mathcal{L}_1 = \mathcal{L}((x^{1/e_1}))$, where $e_1 = e/q$.

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

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

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

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

$$\mathcal{N}_{\mathcal{L}_1/\mathbf{K}((x))}\Delta(x) \, | \, D(x)^{[\mathbf{L} \colon \mathbf{K}_1]} \tag{45}$$

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

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

we have

$$e_1 \, [\mathbf{L} \colon \mathbf{K}] \, q\,(q - 1)\,(k(q) + 1)/e \leqslant [\mathbf{L} \colon \mathbf{K}_1]\,\mu, \tag{46}$$

which yields (44) at once.

3. Study of a fixed α

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

$$e_\alpha = \gcd(e_1, \ldots, e_\rho), \tag{47}$$

where $e_1, \ldots, e_\rho$ are the ramification indices of the covering $x\colon C \to \mathbf{P}^1$ over the point α. 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*

$$d^{-1} \sum_{\substack{v \ \text{is ir-} \\ \text{regular,} \\ v \notin S}} \log \mathcal{N}v \quad \leqslant 12n^2(\mu_\alpha + nu_\alpha)$$

$$\times (h(f) + mh(\alpha) + O(n + \log N)). \tag{48}$$

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

$$\mu_\alpha \geqslant (e_1 - 1) + \cdots + (e_\rho - 1) = n - \rho \geqslant n/2, \tag{49}$$

which will be used in our estimates.

Let

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

be the Puiseux expansions of y at α. Actually, for any i we have e_i equivalent expansions

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

where ξ_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 coefficients a_{ik} of the series y_i. Put $\nu_i = [\mathbf{L}_i\colon \mathbf{K}]$. 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, for any (archimedean and non-archimedean) place v of the field $\mathbf{K}$ there exist A_v, $A_v' \geqslant 1$ such that

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

and satisfying (21), (22) and (24) with $h(f)$, u and μ replaced by $h(f_\alpha)$, u_α and μ_α respectively.

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

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

where

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

PROPOSITION 3.2. *Any $v \notin S \cup \mathbf{M}$ is regular.*

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

$$\Sigma_1 \ll n, \tag{53}$$

$$\Sigma_2 \leqslant h(\alpha), \tag{54}$$

$$\Sigma_3 \leqslant 5n^2(\mu_\alpha + u_\alpha n)(h(f_\alpha) + O(n + \log N)), \tag{55}$$

$$\Sigma_4 \leqslant 4n^2(\mu_\alpha + u_\alpha \log_2 n)(h(f_\alpha) + O(n + \log N)), \tag{56}$$

$$\Sigma_5 \leqslant 4n(h(f_\alpha) + O(n + \log N)). \tag{57}$$

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 $r \leqslant s$ 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

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

which yields

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

We estimate the relative discriminant $D_{\mathbf{L}_i/\mathbf{K}}$ by Lemma 2.4:

$$d^{-1} \log \mathcal{N} \mathcal{D}_{\mathbf{L}_i/\mathbf{K}} \leqslant 2(2n\nu_i(\mu_\alpha + u_\alpha \nu_i) + \nu_i^2)h(f_\alpha)$$
$$+ O(n\nu_i(\mu_\alpha + u_\alpha \nu_i)(n + \log N) + \nu_i \log \nu_i).$$

Using (58) and (59), we obtain

$$\Sigma_3 \leqslant d^{-1} \sum_{i=1}^{r} \log \mathcal{N}(\mathcal{D}_{\mathbf{L}_i/\mathbf{K}})$$
$$\leqslant 2(2n^2(\mu_\alpha + u_\alpha n) + n^2)h(f_\alpha)$$
$$+ O(n^2(\mu_\alpha + u_\alpha n)(n + \log N) + n \log n),$$

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

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

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

By Lemma 2.4 we have $k_i(q) \leqslant e_i \mu_\alpha / \nu_i$. Using Theorem 2.1, we obtain

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

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

$$\Sigma_4(i, q) := d^{-1} \sum_{v \in \mathbf{M}_4(i,q)} \log \mathcal{N}v$$
$$\leqslant \nu_i \Sigma_4'(i, q)$$
$$\leqslant 4n(\nu_i u_\alpha + \mu_\alpha)(h(f_\alpha) + O(n + \log N)).$$

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

$$\log_2 e_1 + \cdots + \log_2 e_r \leqslant e_1 + \cdots + e_r \leqslant n,$$
$$\nu_1 \log_2 e_1 + \cdots + \nu_r \log_2 e_r \leqslant (\nu_1 + \cdots + \nu_r) \log_2 n \leqslant n \log_2 n,$$

we have

$$\Sigma_4 \leqslant \sum_{i=1}^{r} \sum_{q|e_i} \Sigma_4(i, q)$$
$$\leqslant 4n(u_\alpha n \log_2 n + \mu_\alpha n)(h(f_\alpha) + O(n + \log N)),$$

which is (56).

Proof of (57). We write $\mathrm{M}_5 = \mathrm{M}_5' \cup \mathrm{M}_5''$, where M_5' consists of those $v \in \mathrm{M}_5$ for which $\frac{d_v \log A_v}{\log \mathcal{N} v} \in \mathbf{Z}$, and $\mathrm{M}_5'' = \mathrm{M}_5 \setminus \mathrm{M}_5'$. 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 Σ_5' is estimated using (21):

$$
\begin{aligned}
\Sigma_5' &\leqslant d^{-1} \sum_{v \in \mathrm{M}_5'} d_v (\log A_v + \log A_v') \\
&\leqslant d^{-1} \sum_v d_v (\log A_v + \log A_v') \\
&\leqslant 2n(h(f_\alpha) + O(n + \log N)).
\end{aligned}
$$

The sum Σ_5'' is estimated using (24):

$$
\Sigma_5'' \leqslant 2nh(f_\alpha) + O(\log N).
$$

This proves (57).

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

PROPOSITION 3.3. *Let K be a field of characteristic 0, complete with respect to a discrete valuation with residue field of characteristic p. Let π 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 $K(\eta^{1/e})$ over K is $e/\gcd(e, \mathrm{Ord}_\pi(\eta))$.*

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

Put $e' = e/\gcd(e, \tau)$ and $\tau' = \tau/\gcd(e, \tau)$ Then $\eta^{1/e} = (\pi^{1/e'})^{\tau'}$ for some choice of the root $\pi^{1/e'}$. Therefore $K(\eta^{1/e}) \subseteq K(\pi^{1/e'})$. On the other hand, $\gcd(e', \tau') = 1$, therefore exists $a \in \mathbf{Z}$ such that $\tau' a \equiv 1 \bmod e'$. Then $K(\eta^{1/e}) \supseteq K((\eta^{1/e})^a) = K(\pi^{1/e'})$. Thus, $K(\eta^{1/e}) = K(\pi^{1/e'})$, 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 $v \notin \mathrm{M} \cup S$. Then $\tau = \mathrm{Ord}_v(x_0) \geqslant 0$, because $v \notin \mathrm{M}_2 \cup S$. 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

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

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 τ, the proof is finished. Therefore we may suppose that e does not divide τ, that is $e' = e / \gcd(e, \tau) > 1$.

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

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

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

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

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

On the other hand

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

where $e'' = e/q$. Since $v \notin \mathrm{M}_1$, we may apply Proposition 3.3. It implies that the ramification 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 proof of Lemma 3.1.

4. Auxiliary material

4.1. SIEGEL'S CONSTRUCTION OF CONVENIENT UNITS IN NUMBER FIELDS

Propositions 4.1.1 and 4.1.3 of this subsection go back to Siegel [39].

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

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

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 fundamental system of S-units $\eta_1, \ldots, \eta_{s-1}$ satisfying*

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

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

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

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

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

Proof. See Bugeaud and Győry [12, Lemma 1]. Note that the left-hand inequality in (63) is a well-known result of Dobrowolski [15]. ∎

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. Then

$$h(\eta) \leqslant s^{2s-1}\zeta^{s-2}R(S)B, \tag{65}$$

$$B \leqslant s^{2s}\zeta h(\eta), \tag{66}$$

where $B = \max(|b_1|, \ldots, |b_{s-1}|)$.
Proof. Straightforward from (63) and (64). ∎

PROPOSITION 4.1.3. *For any $\alpha \in \mathbf{K}$ there exists an S-unit η such that $\beta = \alpha\eta^{-1}$ satisfies*

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

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

$$d^{-1} \sum_{\substack{v \in S \\ v \neq v_0}} d_v \left|\log |\beta|_v\right| = d^{-1} \sum_{\substack{v \in S \\ v \neq v_0}} d_v \left|\sum_{i=1}^{s-1} (\theta_i - b_i) \log |\eta_i|_v\right|$$

$$\leqslant d^{-1} \sum_{i=1}^{s-1} \sum_{v \in S} |\theta_i - b_i| d_v \left|\log |\eta_i|_v\right|$$

$$\leqslant (2d)^{-1} \sum_{i=1}^{s-1} \sum_{v \in S} d_v \left|\log |\eta_i|_v\right|$$

$$= h(\eta_1) + \cdots + h(\eta_{s-1})$$

$$\leqslant s^{2s-1}\zeta^{s-2}R(S),$$

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 [12, Lemma 3] and by the author [8, Prop. 1.4.8]. (See Pethő [29] and Hajdu [21] for similar results.)

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

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

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

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

Proof. The first inequality follows from Siegel's estimate [39, Satz 1]

$$hR \ll d^{-d} \sqrt{\mathcal{D}} (\log \mathcal{D})^{d-1}.$$

Further, note that $\log \mathcal{D} \leqslant c_1 \mathcal{D}^{0.01}$ and $\log \mathcal{N} v \leqslant c_1 (\mathcal{N} v)^{0.01}$, where $c_1, c_2, \ldots$ are absolute effective constants. Also, $\log \mathcal{N} v \leqslant (\mathcal{N} v)^{0.01}$ as soon as $\mathcal{N} v \geqslant c_2$, and there exist at most $c_2 d$ non-archimedean places v with $\mathcal{N} v \leqslant c_2$. We obtain

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

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. We also need an estimate of a different type, in terms of ramified places. We obtain it as a consequence of Hensel's inequality (69).

PROPOSITION 4.2.1. *Let* $\mathbf{L}/\mathbf{K}$ *be a finite extension of number fields and* $\nu = [\mathbf{L} : \mathbf{K}]$. *Then*

$$d^{-1} \log \mathcal{N}(\mathcal{D}_{\mathbf{L}/\mathbf{K}}) \leqslant d^{-1}(\nu - 1) \sum_{v \text{ is ramified in } \mathbf{L}} \log \mathcal{N} v + O(\nu^2). \quad (68)$$

Proof. Given a non-archimedean place w of $\mathbf{L}$, we denote by $e_w = e_w(\mathbf{L}/\mathbf{K})$, the relative ramification index, and by $f_w = f_w(\mathbf{L}/\mathbf{K})$ the relative degree of the residue fields. We put $s_w = \mathrm{Ord}_w(e_w)$.

By Proposition 6.3 from [28], for any non-archimedean place v of $\mathbf{K}$

$$\mathrm{Ord}_v(\mathcal{D}_{\mathbf{L}/\mathbf{K}}) \leqslant \sum_{w|v}(e_w - 1 + s_w)f_w. \tag{69}$$

(This goes back to Hensel [22], who confirmed a conjecture of Dedekind [14]. See [14, p. 397] and [28] for historical comments and further references.) Since $\sum_{w|v} e_w f_w = \nu$, we have $\sum_{w|v}(e_w - 1)f_w \leqslant \nu - 1$. Further,

$$s_w \log \mathcal{N}v = e_w d_v \mathrm{Ord}_p(e_w) \log p \leqslant e_w d_v \log e_w \leqslant e_w d_v \log \nu,$$

where $p = p(v)$. Therefore

$$\log \mathcal{N}v \sum_{w|v} s_w f_w \leqslant d_v \log \nu \sum_{w|v} e_w f_w = d_v \nu \log \nu.$$

When $p(v) > \nu$ we have $s_w = 0$ for all $w|v$. We obtain:

$$\mathrm{Ord}_v(\mathcal{D}_{\mathbf{L}/\mathbf{K}}) \log \mathcal{N}v \leqslant \begin{cases} (\nu - 1) \log \mathcal{N}v, & p(v) > \nu, \\ (\nu - 1) \log \mathcal{N}v + = d_v \nu \log \nu, & p(v) \leqslant \nu. \end{cases}$$

Therefore

$$d^{-1} \log \mathcal{N}(\mathcal{D}_{\mathbf{L}/\mathbf{K}}) \leqslant d^{-1}(\nu - 1) \sum_v \mathrm{Ord}_v(\mathcal{D}_{\mathbf{L}/\mathbf{K}}) \log \mathcal{N}v$$

$$\leqslant d^{-1}(\nu - 1) \sum_{v \text{ is ramified in } \mathbf{L}} \log \mathcal{N}v$$

$$+ d^{-1}\nu \log \nu \sum_{p(v) \leqslant \nu} d_v.$$

For the last sum we have $\sum_{p(v) \leqslant \nu} d_v = d \sum_{p \leqslant \nu} 1 \ll d\nu / \log \nu$, which completes the proof.

Remark 4.2.2. A similar estimate was obtained by Serre [35, Prop. 4]. However, (68) is more suitable for our purposes.

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 $\alpha_0, \ldots, \alpha_r$ nonzero elements of $\mathbf{K}$. Also, let v be a place of $\mathbf{K}$ and $0 < \varepsilon \leqslant 1$. Suppose that*

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

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

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

where

$$h' = \max(h(\alpha_1), \ldots, h(\alpha_r), 3),$$

$$c(r, d) = \exp(3r\log(rd) + O(r + \log d)).$$

and $h^*(\ldots)$ is defined as in Proposition 4.1. (In the archimedean case the factor $\log h'$ can be skipped.)

Proof. The archimedean case is due to Waldschmidt [45]. Define the parameters in [45, p. 215] as follows:

$$n = r + 1, \quad A_i = \exp(h^*(\alpha_i)) \ \ (0 \leqslant i \leqslant r),$$
$$E = \exp(1), \ f = \exp(-1), \qquad Z_0 = 7 + 3\log(r + 1) + \log d.$$

Applying Corollary 10.2 from [45] in this set-up, we obtain

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

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

The non-archimedean case is due to Yu [46]. Define the parameters in [46, p. 241–242] as follows:

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

where e_v is the ramification index of v over $\mathbf{Q}$ (in particular, $e_v \leqslant d$). In this set-up the third displayed formula on [46, p. 242] would turn into

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

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_α and e_β have a common divisor $e \geqslant 3$.
(B) There exist distinct $\alpha, \beta, \gamma \in \overline{\mathbf{Q}}$ such that e_α, e_β and e_γ have a common divisor $e = 2$.

Put

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

and define ν_β, ν_γ, κ_β, κ_γ similarly. Also, put

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

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

$$\Omega = 130\Theta + ed(22\nu_\alpha \nu_\beta(\kappa_\alpha + \kappa_\beta) + 15\nu_\gamma \nu_\beta(\kappa_\gamma + \kappa_\beta)).$$

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

$$h_x(P) \leqslant \widehat{p}(S)^\delta (\mathcal{D} \prod_{v \in S} \mathcal{N}v)^{2.1\Upsilon} \exp(\Omega), \tag{72}$$

where in the case (A) *the terms* ν_γ *and* κ_γ *should be replaced by* ν_α *and* κ_α.

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 φ and φ' with the following three properties:

  (i) the heights of φ and φ' are of the same magnitude as $h_x(P)$;
 (ii) each of φ and φ' is 'almost an S-unit' (an S-unit times an algebraic number of bounded height);
(iii) for some place v_0 the ratio φ/φ' (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 φ and φ'. 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

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

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

We put

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

$$
\widetilde{h} = \begin{cases} \max(h(\alpha),\ h(\beta),\ 1) & \text{in the case (A)}, \\ \max(h(\alpha),\ h(\beta),\ h(\gamma),\ 1) & \text{in the case (B)}. \end{cases}
$$

2. *Definition of φ and field $\mathbf{L}$*

If $|x_0|_{v_0} \leqslant 10e\sigma_0$ then (73) implies an upper bound for $h_x(P)$ much better than (72). Hence we may suppose that $|x_0|_{v_0} \geqslant 10e\sigma_0$, 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 eth root of unity ξ (in particular $\xi = -1$ 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 $|x_0|_{v_0} \geqslant ce^6\sigma_0^2$, where c is a sufficiently large absolute effective constant, then $\varphi \neq 0$ and

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

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

It is worth mentioning that

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

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 $h(x_0) \ll e(\widetilde{h} + e)$, which is much better than (72). Indeed, by the definition of φ we have $h(\varphi) \ll h(x_0) + \widetilde{h} + e$ whence either $h(x_0) \leqslant \widetilde{h} + e$ or $h(\varphi) \ll h(x_0)$. Further, rewrite (76) as $t_0^e + \beta - \alpha = (\sqrt[e]{\varphi} + t_0)^e$, where $t_0 = \sqrt[e]{x_0 - \beta}$. Thus, t_0 is a root of a polynomial of degree $e - 1$ and height $O(h(\varphi) + \widetilde{h} + e)$. By Proposition 1.4 we have $h(t_0) \ll h(\varphi) + \widetilde{h} + e$, whence $h(x_0) \ll eh(t_0) + \widetilde{h} \ll e(h(\varphi) + \widetilde{h} + e)$.

If $h(\varphi) \geqslant \widetilde{h} + e$ then $h(x_0) \ll eh(\varphi)$, and if $h(\varphi) \leqslant \widetilde{h} + e$ then we have $h(x_0) \ll e(\widetilde{h} + e)$.

We also have to estimate the discriminant $\mathcal{D}_{\mathbf{L}}$. By Proposition 2.4

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

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

Indeed, $\mu_\alpha \geqslant n/2$, as we have seen in Section 3, and on the other hand $\mu_\alpha \nu_\alpha \leqslant \deg D(X) \leqslant (2n-2)m$. Therefore $\nu_\alpha \leqslant 4m$, whence $\nu_\alpha h(\alpha) + \log \nu_\alpha \leqslant \kappa_\alpha$. 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 $p(v) \leqslant n$, 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

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

and similarly one defines and estimates $\Sigma(\beta)$. Denote by S_0 the set of places of $\mathbf{L}_0$ above S. By Proposition 4.2

$$d_{\mathbf{L}_0}^{-1} \log \mathcal{N}_{\mathbf{L}_0}(\mathcal{D}_{\mathbf{L}/\mathbf{L}_0}) \leqslant (e-1)d_{\mathbf{L}_0}^{-1} \sum_{v \text{ is ramified in } \mathbf{L}} \log \mathcal{N}_{\mathbf{L}_0} v + O(e^2)$$

$$\leqslant e\big(d_{\mathbf{L}_0}^{-1} \sum_{v \in S_0} \log \mathcal{N}_{\mathbf{L}_0} v + \Sigma(\alpha) + \Sigma(\beta)\big) + O(e^2)$$

$$\leqslant e\big(d^{-1} \sum_{v \in S} \log \mathcal{N} v + 12\kappa_\alpha + 12\kappa_\beta\big).$$

Finally

$$\mathcal{D}_{\mathbf{L}} \leqslant \mathcal{D}^{e\nu_\alpha \nu_\beta} \mathcal{N}(\mathcal{D}_{\mathbf{K}(\alpha)/\mathbf{K}})^{e\nu_\beta} \mathcal{N}_{\mathbf{K}(\alpha)}(\mathcal{D}_{\mathbf{L}_0/\mathbf{K}(\alpha)})^e \mathcal{N}_{\mathbf{L}_0}(\mathcal{D}_{\mathbf{L}/\mathbf{L}_0})$$

$$\leqslant \left(\mathcal{D}\left(\prod_{v \in S} \mathcal{N} v\right) \exp(14d(\kappa_\alpha + \kappa_\beta))\right)^{e\nu_\alpha \nu_\beta}. \tag{80}$$

4. φ 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 $s_1 := |S_1| \leqslant se\nu_\alpha \nu_\beta$ and for any $v \notin S_1$ we have

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

Indeed, $|\varphi|_v \leqslant \sigma_v$ 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 $|\varphi_k|_v \leqslant \sigma_v$, we obtain $|\varphi|_v \geqslant \varrho_v^e \sigma_v^{1-e}$.

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. Then

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

By Proposition 4.1 there exists a unit $\eta = \eta_1^{b_1} \cdots \eta_{s_1-1}^{b_{s_1-1}}$ such that

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

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

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

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

In addition, estimate $R(S_1)$. Corollary 4.1 and (80) yield

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

5. *Definition of φ' and $\mathbf{L}'$*

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

$$
\theta' = \begin{cases} \xi' \sqrt[e]{\dfrac{x_0 - \alpha}{x_0 - \beta}} & \text{in the case (A)}, \\[2ex] \xi' \sqrt[e]{\dfrac{x_0 - \gamma}{x_0 - \beta}} & \text{in the case (B)}, \end{cases}
$$

$$
\mathbf{L}' = \begin{cases} \mathbf{K}(\alpha, \, \beta, \, \theta') & \text{in the case (A)}, \\[1ex] \mathbf{K}(\gamma, \, \beta, \, \theta') & \text{in the case (B)}, \end{cases}
$$

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 φ', ψ', η', etc., we refer to the analogues of (75)–(83) as $(75')$–$(83')$. For example, in the case (B)

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

and in the case (A) κ_γ and ν_γ should be replaced here (and everywhere below) by κ_α and ν_α.

6. *Estimating B*

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

$$h(x_0) \leqslant c_1 R(S_1)B, \tag{84}$$

$$B \leqslant c_1 \log|x_0|_{v_0}, \tag{85}$$

where $c_1 = \exp(6.7\Theta)$. Indeed, $h(\varphi) \leqslant h(\psi) + h(\eta) + O(1)$, where $h(\eta) \ll s_1^{2.1s_1}R(S_1)B$ by Corollary 4.1. Combining this with (77) and (82), we obtain $h(x_0) \ll c_1 R(S_1)B + e^2\widetilde{h}$. We may assume $c_1 R(S_1)B \geqslant e^2\widetilde{h}$ (otherwise it would be $h(x_0) \ll e^2\widetilde{h}$, better than (72)). Therefore $h(x_0) \leqslant c_1 R(S_1)B$.

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

Further, by Corollary 4.1

$$\max(b_1,\ldots,b_{s_1-1}) \ll s_1^{2.1s_1}h(\eta) \leqslant s_1^{2.1s_1}(h(\varphi)+h(\psi)+O(1))$$

and similarly for $\max(b'_1,\ldots,b'_{s'_1-1})$. Combining this with (73), (82), $(82')$, (77) and $(77')$, we obtain $B \leqslant c_1(\log|x_0|_{v_0}+c_2)$, where $c_2 = \widetilde{h}+c_1(R(S_1)+R(S'_1))$. If $\log|x_0|_{v_0} \geqslant c_2$ then $B \leqslant c_1\log|x_0|_{v_0}$, as desired. If $\log|x_0|_{v_0} \leqslant c_2$ then $h(x_0) \leqslant sd^{-1}c_2$; 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 $\varphi/\varphi' \neq (\xi-1)^e/(\xi'-1)^e$. 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 $\eta_0 = (\frac{\xi'-1}{\xi-1})^e \frac{\psi}{\psi'}$. Using (75) and (85), we obtain

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

Put $r = s_1 + s_1' - 2 \leqslant 2se\nu_\beta(\nu_\alpha + \nu_\gamma)$ and write $\eta_{s_1}, \ldots, \eta_r$ and $b_{s_1}, \ldots, b_r$ instead of $\eta_1', \ldots, \eta_{s_1'-1}'$ and $-b_1', \ldots, -b_{s_1'-1}'$, respectively. Then

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

$$[\mathbf{Q}(\eta_0, \eta_1, \ldots, \eta_r):\mathbf{Q}] \leqslant \delta \ll sN^5,$$

$$h^*(\eta_0) \leqslant c_1 \widetilde{R} + O(e\widetilde{h}),$$

$$h^*(\eta_1) \cdots h^*(\eta_r) \leqslant c_1^2 R(S_1) R(S_1'),$$

$$h' := \max(h(\eta_1), \ldots, h(\eta_r), 3) \leqslant c_1 \widetilde{R},$$

where $\widetilde{R} = \max(R(S_1), R(S_1'), 3)$ and $h^*(\ldots)$ is defined in Proposition 4.1. By Proposition 4.3

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

By (84), (83) and (83′)

$$\begin{aligned} h(x_0) &\leqslant \widehat{p}(S)^{\delta - 0.3} (\widetilde{R} + \widetilde{h}) R(S_1)^2 R(S_1')(\log^2 \widetilde{R}) \exp(130\Theta) \\ &\leqslant \widehat{p}(S)^\delta \left(\mathcal{D} \prod_{v \in S} \mathcal{N}v \right)^{2.1\Upsilon} \exp(\Omega). \end{aligned}$$

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 $e_\alpha \geqslant 3$ and $e_\beta \geqslant 2$;
(b) there exist $\alpha, \beta, \gamma \in \overline{\mathbf{Q}}$ such that $e_\alpha = e_\beta = e_\gamma = 2$.

The case (a) is contained in (b) and the following three cases:

(a1) there exist $\alpha, \beta \in \overline{\mathbf{Q}}$ such that $e_\alpha = e_\beta \geqslant 3$;
(a2) there exist $\alpha, \beta \in \mathbf{K}$ such that $e_\alpha \geqslant 3$ and $e_\beta \geqslant 2$;

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

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

$$\beta \notin \mathbf{K}, \qquad e_\beta \geqslant 3$$

can be reduced to (a1), and the case

$$[\mathbf{K}(\beta) \colon \mathbf{K}] \geqslant 3, \qquad e_\beta = 2$$

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 α over $\mathbf{K}$. All them are roots of $D(X)$ of order μ_α. Therefore

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

Since $\mu_\alpha \geqslant n/2$, as we have seen in Section 3, we obtain $\nu_\alpha \leqslant 4m$. (We have already used this in the previous section.) Similarly, $\nu_\beta,\ \nu_\gamma \leqslant 4m$. Furthermore, $h(\alpha_1) = \ldots = h(\alpha_{\nu_\alpha})$. Hence, by Proposition 1.4

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

Combining (87)–(88), we obtain

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

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

$$\begin{aligned}
\nu_\alpha \nu_\beta (\kappa_\alpha + \kappa_\beta) &\leqslant 4m(\nu_\alpha \kappa_\alpha + \nu_\beta \kappa_\beta) \\
&\leqslant 48 n^3 m^2 (h(f) + O(n + \log N)),
\end{aligned} \tag{90}$$

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

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

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

$$\Theta \leqslant 16 m^2 n s (\log(Ns) + O(1)). \tag{93}$$

Substituting the estimates (90)–(92) to (72), we obtain

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

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 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 $\widetilde{f}(t, y) = 0$, where $\widetilde{f}(T, Y) = f(\beta + T^e, Y)$. In particular,

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

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 $\widetilde{P} \in \varphi^{-1}(P)$ there exists an extension $\widetilde{\mathbf{K}}$ of $\mathbf{K}$ such that $[\widetilde{\mathbf{K}}: \mathbf{K}] \leqslant e$,

$$
\mathcal{D}_{\widetilde{\mathbf{K}}} \leqslant \left(\mathcal{D} \left(\prod_{v \in S} \mathcal{N}v \right) \exp(72dn^3m(h(f) + O(n + \log N))) \right)^e, \qquad (95)
$$

and $\widetilde{P} \in \widetilde{C}(t, \widetilde{\mathbf{K}}, \widetilde{S})$, where $\widetilde{S}$ is the set of places of $\widetilde{\mathbf{K}}$ above S.

(ii) Put

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

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

Proof of (i). We have $\widetilde{P} \in \widetilde{C}(t, \widetilde{\mathbf{K}}, \widetilde{S})$ with $\widetilde{\mathbf{K}} = \mathbf{K}(\sqrt[e]{x(P) - \beta})$ for an appropriate definition of the root. By Proposition 4.2, Lemma 3.1 and (89) we have

$$
d^{-1} \log \mathcal{N}(\mathcal{D}_{\widetilde{\mathbf{K}}}) \leqslant ed^{-1} \sum_{\substack{v \text{ is ramified in } \widetilde{\mathbf{K}}}} \log \mathcal{N}v + O(e^2)
$$

$$
\leqslant ed^{-1} \left(\sum_{v \in S} + \sum_{p(v) \leqslant e} + \sum_{\substack{v \notin S \text{ and } e \\ \text{does not divide} \\ \text{Ord}_v(x(P) - \alpha)}} \right) \log \mathcal{N}v + O(e^2)
$$

$$
\leqslant e \left(d^{-1} \sum_{v \in S} \log \mathcal{N}v + 12\kappa_\alpha \right)
$$

$$\leqslant e \left(d^{-1} \sum_{v \in S} \log \mathcal{N} v + 72 n^3 m (h(f) + O(n + \log N)) \right),$$

which yields (95).

Proof of (ii). Below *divisor* means *divisor on* $\widetilde{C}$. 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 $\widetilde{e}$. (We say that a divisor D is *divisible* by an integer l if $D = lD'$ for some divisor D'.) On the other hand,

$$(t^e - \gamma)_0 = (t - \widetilde{\alpha})_0 + (t - \xi \widetilde{\alpha})_0 + \cdots + (t - \xi^{e-1} \widetilde{\alpha})_0. \tag{96}$$

The divisor in the left-hand side of (96) is divisible by $\widetilde{e}$. Since the divisors in the right-hand side have pairwise disjoint supports, each of them is divisible by $\widetilde{e}$. In particular, $\widetilde{e}$ divides $e_{\widetilde{\alpha}}(t)$ and $e_{\widetilde{\beta}}(t)$, as desired.

We have

$$h(\widetilde{\alpha}) \leqslant e^{-1}(h(\alpha) + h(\beta) + O(1)) \leqslant 4e^{-1} n (h(f) + O(\log N)),$$

$$\widetilde{\nu}_{\widetilde{\alpha}} := [\widetilde{\mathbf{K}}(\widetilde{\alpha}) : \widetilde{\mathbf{K}}] \leqslant e,$$

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

$$\widetilde{\Upsilon} := e \widetilde{\nu}_{\widetilde{\alpha}} \widetilde{\nu}_{\widetilde{\beta}} \leqslant e^2 \widetilde{e},$$

$$\widetilde{\delta} := d_{\widetilde{\mathbf{K}}} \widetilde{e}^2 \widetilde{\nu}_{\widetilde{\alpha}} \widetilde{\nu}_{\widetilde{\beta}} \leqslant de^3 \widetilde{e}^2,$$

$$\widetilde{s} := |\widetilde{S}| \leqslant se,$$

$$\widetilde{\Theta} := \widetilde{s} \widetilde{\Upsilon} (\log(\widetilde{N} \widetilde{s}) + O(1)) \leqslant 3se^3 \widetilde{e} (\log(NS) + O(1)).$$

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

$$\widetilde{u}_{\widetilde{\alpha}} = \widetilde{u}_{\widetilde{\beta}} = u_\alpha, \qquad \widetilde{\mu}_{\widetilde{\alpha}} = \widetilde{\mu}_{\widetilde{\beta}} = \mu_\alpha. \tag{97}$$

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

$$\mu_\alpha + n u_\alpha \leqslant \mu_\beta + n u_\beta, \tag{98}$$

interchanging α and β if necessary. Defining in the obvious manner $\widetilde{\kappa}_{\widetilde{\alpha}}$ and using (87), (88) and (98), we obtain

$$
\begin{aligned}
\widetilde{\kappa}_{\widetilde{\alpha}} &= \widetilde{\nu}_{\widetilde{\alpha}}\widetilde{n}^2(\widetilde{\mu}_{\widetilde{\alpha}} + \widetilde{n}\widetilde{u}_{\widetilde{\alpha}})(h(\widetilde{f}) + \widetilde{m}h(\widetilde{\alpha}) + O(\widetilde{n} + \log N)) \\
&\leqslant n^2(\mu_\alpha + nu_\alpha)(h(f) + m(h(\alpha) + 2h(\beta)) + O(N)) \qquad (99) \\
&\leqslant n^2 m(3nh(f) + (\mu_\alpha + nu_\alpha)h(\alpha) \\
&\quad + 2(\mu_\beta + nu_\beta)h(\beta) + O(nN)) \qquad\qquad (100) \\
&\leqslant 12n^3 m(h(f) + O(N)). \qquad\qquad\qquad\qquad (101)
\end{aligned}
$$

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

$$
\begin{aligned}
h_x(P) &\leqslant e(h_t(\widetilde{P}) + h(\beta) + O(1)) \\
&\leqslant \widehat{p}(\widetilde{S})^{\widetilde{\delta}}\left(\mathcal{D}_{\widetilde{\mathbf{K}}}\prod_{v\in\widetilde{S}}\mathcal{N}_{\widetilde{\mathbf{K}}}v\right)^{2.1\widetilde{\Upsilon}}\exp(130\widetilde{\Theta} + 37\widetilde{e}d_{\widetilde{\mathbf{K}}}\widetilde{\nu}_{\widetilde{\alpha}}\widetilde{\nu}_{\widetilde{\beta}}(\widetilde{\kappa}_{\widetilde{\alpha}} + \widetilde{\kappa}_{\widetilde{\beta}})) \\
&\leqslant (\widehat{p}(S)^{\widetilde{de}}\mathcal{D}^{2.1}\left(\prod\mathcal{N}v\right)^{3.1}\exp(400\Omega_2))^{e^3\widetilde{e}} \qquad (102) \\
&\leqslant (\widehat{p}(S)^{dn}\mathcal{D}^{2.1}\left(\prod\mathcal{N}v\right)^{3.1}\exp(400\Omega_2))^{n^4} \qquad (103)
\end{aligned}
$$

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

Now suppose that $e = e_\beta = 2$. Then we cannot assume (98) any more. Instead, we shall use the estimates

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

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

We still have (99), but instead of (100) we obtain

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

Therefore instead of (102) we have

$$
\begin{aligned}
h_x(P) &\leqslant \left(\widehat{p}(S)^{\widetilde{de}}\mathcal{D}^{2.1}\left(\prod\mathcal{N}v\right)^{3.1}\exp(399\Omega_3)\right)^{e^3\widetilde{e}} \\
&\leqslant \left(\widehat{p}(S)^{dn}\mathcal{D}^{2.1}\left(\prod\mathcal{N}v\right)^{3.1}\exp(399\Omega_3)\right)^{8n} \qquad (104)
\end{aligned}
$$

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

6.3. CASE (a3)

We have

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

Applying (104) with the field $\mathbf{K}(\beta)$ instead of $\mathbf{K}$, we obtain

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

again better than (3).

Theorem 1.2 is proved.

Acknowledgements

References

1. Baker, A.: 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. Baker, A.: Contributions to the theory of Diophantine equations. I, II, *Phil. Trans. R. Soc. London Ser. A* 263 (1967–68), 173–208.
3. Baker, A.: The Diophantine equation $y^2 = ax^3 + bx^2 + cx + d$, *J. London Math. Soc.* 43 (1968), 1–9.
4. Baker, A.: Bounds for the solutions of the hyperelliptic equations, *Proc. Camb. Phil. Soc.* 65 (1969), 439–444.
5. Baker, A. and Coates, J.: Integer points on curves of genus 1, *Proc. Camb. Phil. Soc.* 67 (1970), 592–602.
   Baker, A. and Wüstholz, G.: Logarithmic forms and group varieties, *J. reine und angew. Math.* 442 (1993), 19–62.
6. Bilu, Yu. (Belotserkovski): Effective analysis of a class of Diophantine equations (Russian), *Vestsi Akad. Navuk BSSR, Ser. Fiz.-Math. Navuk* 3 (1988), 111–115.
7. Bilu, Yu. (Belotserkovski): Effective analysis of a new class of Diophantine equations (Russian), *Vestsi Akad. Navuk BSSR, Ser. Fiz.-Math. Navuk* 6 (1988), 34–39, 125
8. Bilu, Yu.: *Effective Analysis of Integral Points on Algebraic Curves*, thesis, Beer Sheva, 1993.
9. Bilu, Yu.: *Solving Superelliptic Diophantine Equations by the Method of Gelfond–Baker*, preprint 94-09, Mathématiques Stochastiques, Univ. Bordeaux 2, 1994.
10. Bilu, Yu.: Effective analysis of integral points on algebraic curves, *Israel J. Math* 90 (1995), 235–252.
11. Brindza, B.: On S-integral solutions of the equation $y^m = f(x)$, *Acta Math. Hungar.* 44 (1984), 133–139.

12. Bugead, Y. and Győry, K.: Bounds for the solutions of unit equations, *Acta Arith.*, to appear.
13. Coates, J.: Construction of rational functions on a curve, *Proc. Camb. Phil. Soc.* 68 (1970), 105–123.
14. Dedekind, R.: Über die Diskriminanten endlicher Körper, *Werke* I, 351–397; Vieweg, 1930.
15. Dobrowolski, E.: On a question of Lehmer and the number of irreducible factors of a polynomial, *Acta Arith.* 34 (1979), 391–401.
16. Dvornicich, R. and Zannier, U.: Fields containing values of algebraic functions, *Ann. Sc. Norm. Sup. Pisa Cl. Sci. Ser. IV* 21 (1994), 421–443.
17. Dwork, B. M. and Robba, B.: On natural radii of p-adic convergence, *Trans. Amer. Math. Soc.* 256 (1979), 199–213.
18. Dwork, B. M. and van der Poorten, A. J.: The Eisenstein constant, *Duke Math. J.* 65 (1992), 23–43; Corrections: 76 (1994), 669–672.
19. Faltings, G.: Endlichkeitssätze für abelche Varietäten über Zahlkörpern, *Invent. Math.* 73 (1983), 349–366; Erratum: 75 (1984), 381.
20. Gelfond, A. O.: *Transcendent and Algebraic Numbers* (Russian), Moscow 1952; English trans.: New York, Dover, 1960.
21. Hajdu, L.: A quantitative version of Dirichlet's S-unit theorem in algebraic number fields, *Publ. Math. Debrecen* 42 (1993), 239–246.
22. Hensel, K.: Ueber die Bestimmung der Diskriminante eines algebraischen Körpers, *Nachr. Gött. Akad.* 1897, 247–253.
23. Kleiman, H.: On the Diophantine equation $f(x, y) = 0$, *J. reine und angew. Math.* 286/287 (1976), 124–131.
24. Kotov, S. V. and Trelina, L. A.: S-ganze Punkte auf elliptischen Kurven, *J. reine und angew. Math.* 306 (1979), 28–41.
25. Lang, S.: *Algebraic Number Theory*, Addison-Wesley, 1970.
26. Lang, S.: *Fundamentals of Diophantine Geometry*, Springer, 1983.
27. LeVeque, W. J.: On the equation $y^m = f(x)$, *Acta Arith.* 9 (1964), 209–219.
28. Narkiewicz, W.: *Elementary and Analytic Theory of Algebraic Numbers*, Springer and PWN, 1990.
29. Pethő, A.: Beiträge zur Theorie der S-Ordnungen, *Acta Math. Acad. Sci. Hungar.* 37 (1981), 51–57.
30. Poulakis, D.: Solutions entières de l'équation $Y^m = f(x)$. *Sém. Thm Nom. Bordeaux* 3 (1991), 187–199.
31. Poulakis, D.: Points entiers et modèles des courbes algébriques, *Monatsh. Math.* 118 (1994), 111–143.
32. Poulakis, D.: Estimation effective des points entiers d'une famille de courbes algébriques, *Ann. Fac. Sci. Toulouse*, to appear.
33. Schmidt, W. M.: Eisenstein theorem on power series expansions of algebraic functions, *Acta Arith.* 56 (1990), 161–179.
34. Schmidt, W. M.: Integer points on curves of genus 1, *Compositio Math.* 81 (1992), 33–59.
35. Serre, J.-P.: Quelques applications du théorème de densité de Chebotarev, *Publ. Math. I.H.E.S.* 54 (1981), 123–201.
36. Serre, J.-P.: *Lectures on the Mordell-Weil Theorem*, Aspects of math. E15, Vieweg 1989.
37. Shorey, T. N. and Tijdeman, R.: *Exponential Diophantine Equations*, Cambridge University Press, Cambridge, 1986.
38. Siegel, C. L.: Über einige Anwendungen Diophantischer Approximationen, *Abh. Preuss Akad. Wiss. Phys.-Math. Kl.* 1 (1929).
39. Siegel, C. L.: Abschätzung von Einheiten, *Nachr. Akad. Wiss. Göttingen II. Math.-Phys. Kl.* 9 (1969), 71–86.
40. Silverman, J. H.: Lower bounds for height functions, *Duke Math. J.* 51 (1984), 395–403.
41. Sprindžuk, V. G.: 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.
42. Sprindžuk, V. G.: *Classical Diophantine Equations in Two Unknowns* (Russian), Nauka, Moscow, 1982; English transl.: Lecture Notes in Math. 1559, Springer, 1994.

43. Trelina, L. A.: On S-integral solutions of the hyperelliptic equation (Russian), *Dokl. Akad. Nauk BSSR* 22 (1978), 881–884.
44. Voutier, P. M.: An Upper Bound for the Size of Integral Solutions to $Y^m = f(X)$, *J. Number Theory* 53 (1995), 247–271.
45. Waldschmidt, M.: Minorations de combinaisons linéares de logarithmes de nombres algébriques, *Canadian J. Math.* 45 (1993), 176–224.
46. Yu, Kunrui: Linear forms in p-adic logarithms III, *Compositio Math.* 91 (1994), 241–276.