ACTA ARITHMETICA
145.4 (2010)

Quantitative Riemann existence theorem over a number field

by

Yuri F. Bilu (Talence) and Marco Strambi (Pisa)

1. Introduction. The *Riemann Existence Theorem* asserts that every compact Riemann surface is (analytically isomorphic to) a complex algebraic curve. In other words, if f is a non-constant meromorphic function on a compact Riemann surface S, then the field of all meromorphic functions on S is a finite extension of $\mathbb{C}(f)$.

One of the most common ways of defining Riemann surfaces is realizing them as finite ramified coverings of the Riemann sphere $\mathbb{P}^1(\mathbb{C})$. Moreover, even if the covering is purely topological, the $\mathbb{C}$-analytic structure on the Riemann sphere lifts, in a unique way, to the covering surface. Thus, the Riemann Existence Theorem can be restated as follows.

Theorem 1.1. *Let* M *be a finite subset of* $\mathbb{P}^1(\mathbb{C})$*. Then for any finite covering of* $\mathbb{P}^1(\mathbb{C})$ *by a closed oriented surface, unramified outside* M*, there exists a complex algebraic curve* $\mathcal{C}$ *and a rational function* $x \in \mathbb{C}(\mathcal{C})$ *such that our covering is isomorphic* ([1]) *to* $\mathcal{C}(\mathbb{C}) \xrightarrow{x} \mathbb{P}^1(\mathbb{C})$*, the covering defined by* x*. Moreover, the couple* $(\mathcal{C}, x)$ *is unique up to a naturally defined isomorphism* ([2])*.*

We refer to [4] for several more precise statements.

The purpose of this article is to give an effective description of the curve $\mathcal{C}$, or, more precisely, of the couple $(\mathcal{C}, x)$, in terms of the degree of the initial topological covering and the set M of ramification points, provided those points are defined over the field $\bar{\mathbb{Q}}$ of all algebraic numbers. In this

2010 *Mathematics Subject Classification*: Primary 11G30; Secondary 14H25, 14H05, 14H55, 11G50.

Key words and phrases: Riemann existence theorem, algebraic functions, coverings.

([1]) Two morphisms $S_1 \xrightarrow{\pi_1} S$ and $S_2 \xrightarrow{\pi_2} S$ of topological spaces are *isomorphic* if there exists a homeomorphism $S_1 \xrightarrow{\varphi} S_2$ such that $\pi_1 = \pi_2 \circ \varphi$.

([2]) If $(\mathcal{C}', x')$ is another such couple, then the field isomorphism $\mathbb{C}(x) \to \mathbb{C}(x')$ given by $x \mapsto x'$ extends to a field isomorphism $\mathbb{C}(\mathcal{C}) \to \mathbb{C}(\mathcal{C}')$.

DOI: 10.4064/aa145-4-2

case the curve $\mathcal{C}$ is also defined over $\bar{\mathbb{Q}}$ (this is the "easy" direction of the Theorem of Belyi). We produce a plane model of $\mathcal{C}$ over $\bar{\mathbb{Q}}$ such that one of the coordinates is x, and we give explicit bounds for the degree and height of the defining equation of this model, and of the degree and discriminant of the number field over which this model is defined.

Notice that we do not produce a new proof of the Riemann Existence Theorem. In fact, we do use both the existence and uniqueness statements of Theorem 1.1.

Let us state our principal result. Everywhere in this article, by *height* we mean *logarithmic affine height*; see Section 2.

THEOREM 1.2. *Let $S \to \mathbb{P}^1(\mathbb{C})$ be a finite covering of degree $n \geq 2$ by a closed oriented surface S of genus $\mathbf{g}$, unramified outside a finite set $\mathrm{M} \subset \mathbb{P}^1(\bar{\mathbb{Q}})$. Put* ([3])

$$\mathbb{K} = \mathbb{Q}(\mathrm{M}), \qquad h = \max\{\mathrm{h}(\alpha) : \alpha \in \mathrm{M}\}, \qquad \Lambda = (2(\mathbf{g}+1)n^2)^{10\mathbf{g}n+12n}.$$

Then there exist a number field $\mathbb{L}$ containing $\mathbb{K}$, an algebraic curve $\mathcal{C}$ defined over $\mathbb{L}$ and rational functions $x, y \in \mathbb{L}(\mathcal{C})$ such that $\mathbb{L}(\mathcal{C}) = \mathbb{L}(x, y)$ and the following is true.

(a) *The covering $\mathcal{C}(\mathbb{C}) \xrightarrow{x} \mathbb{P}^1(\mathbb{C})$ defined by x is isomorphic to the given covering $S \to \mathbb{P}^1(\mathbb{C})$.*

(b) *The rational functions $x, y \in \mathbb{L}(\mathcal{C})$ satisfy the equation $f(x, y) = 0$, where $f(X, Y) \in \mathbb{L}[X, Y]$ is an absolutely irreducible polynomial and*

$$\deg_X f = \mathbf{g} + 1, \qquad \deg_Y f = n, \qquad \mathrm{h}(f) \leq \Lambda(h+1). \tag{1.1}$$

(c) *The degree and the discriminant of $\mathbb{L}$ over $\mathbb{K}$ satisfy*

$$[\mathbb{L} : \mathbb{K}] \leq \Lambda, \qquad \frac{\log \mathcal{N}_{\mathbb{K}/\mathbb{Q}} \mathcal{D}_{\mathbb{L}/\mathbb{K}}}{[\mathbb{L} : \mathbb{Q}]} \leq \Lambda(h+1), \tag{1.2}$$

where $\mathcal{N}_{\mathbb{K}/\mathbb{Q}}$ is the norm map.

The principal motivation of this theorem lies in the field of effective Diophantine analysis, where the covering technique is widely used. It happens quite often that only the degree of the covering and the ramification points are known, and to work with the covering curve, one needs to have an effective description of it. In particular, in [2] we use Theorem 1.2 to get a user-friendly version of the Chevalley–Weil theorem, one of the main tools of Diophantine analysis.

In brief, our method of proof is as follows. First, we use the existence part of Theorem 1.1 to show the existence of $\mathcal{C}$ and x. Next, we define "quasi-canonically" a generator y of $\bar{\mathbb{Q}}(\mathcal{C})$ over $\bar{\mathbb{Q}}(x)$, and denote by $f(X, Y)$ the

([3]) A pedantic reader may complain that the definition of h below is formally incorrect, because $\mathrm{h}(\cdot)$ is the *affine* height, and M is a subset of the *projective* line. Of course, this can be easily overcome, for instance by writing $\mathbb{P}^1 = \mathbb{A}^1 \cup \{\infty\}$ and defining $\mathrm{h}(\infty) = 0$.

irreducible polynomial satisfying $f(x, y) = 0$. Further, we show that the coefficients of this polynomial satisfy a certain system of algebraic equations and inequalities, and we use the uniqueness part of Theorem 1.1 to show that the system has finitely many solutions. (To be more precise, the coefficients of f form only a part of the variables involved in the equations and inequalities.) Using this, we estimate the height of the polynomial, and the degree and discriminant of the number field generated by its coefficients.

This argument is inspired by the work of Zverovich [13], who applies a rather similar approach, though he works only in the complex domain. The system of equations considered by Zverovich is simpler than ours, but we could not understand one key point in his proof of the finiteness of the number of solutions. See more on this in Section 16.

Our result is sensitive only to the set M of ramification points, and the degree n of the covering. It would be interesting to obtain a more precise result, which depends on the more subtle elements of the "covering data", like the monodromy permutations associated to every ramification point. Probably, the "correct" statement of Theorem 1.2 must involve the notion of the Hurwitz space associated to the given topological covering (see [5]). Another interesting problem is to characterize our curve not in terms of the defining equation, but in more invariant terms, for instance, to estimate its Faltings height.

In our result, the quantity Λ depends exponentially on n. This improves on Theorem 3A from [1], where the dependence is doubly exponential. There are strong reasons to believe that the "correct" estimate is polynomial in n. Indeed, this is the case for a similar problem over a function field, in the recent work of Edixhoven et al. [7].

In Sections 2–4 we collect various auxiliary facts needed for the proof of Theorem 1.2. The proof itself occupies Sections 5–15. In Section 16 we very briefly discuss the work of Zverovich.

Notation and conventions. If $F(X)$ is a polynomial in X over some field (or integral domain), and β is an element of this field (or domain), then we denote by $\operatorname{ord}_{X=\beta} F$ the order of vanishing of F at β. Sometimes we write simply ord_β or even ord, when this does not lead to confusion. We employ the same notation not only for polynomials, but also for formal power series in $X - \beta$.

We denote by α the finite point $(\alpha : 1)$ of the projective line $\mathbb{P}^1$, and by ∞ the infinite point $(1 : 0)$.

More specific notation will be introduced at appropriate places.

2. Heights and algebraic equations. Let $\underline{\alpha} = (\alpha_1, \dots, \alpha_N) \in \bar{\mathbb{Q}}^N$ be a point with algebraic coordinates in the affine space of dimension N. Let $\mathbb{K}$

be a number field containing $\alpha_1, \dots, \alpha_N$ and $M_{\mathbb{K}}$ the set of its valuations. We assume that every valuation $v \in M_{\mathbb{K}}$ is normalized so that its restriction to $\mathbb{Q}$ is the standard infinite or p-adic valuation. Also, we let $\mathbb{K}_v$ be the v-adic completion of $\mathbb{K}$ (in the case of an infinite v, the field $\mathbb{K}_v$ is either $\mathbb{R}$ or $\mathbb{C}$). For $v \in M_{\mathbb{K}}$ we put

$$|\underline{\alpha}|_v = \max\{|\alpha_1|_v, \dots, |\alpha_N|_v\}.$$

We now define the *absolute logarithmic affine height* (or simply *height*) of the point $\underline{\alpha}$ as

$$\mathrm{h}(\underline{\alpha}) = \frac{1}{[\mathbb{K} : \mathbb{Q}]} \sum_{v \in M_{\mathbb{K}}} [\mathbb{K}_v : \mathbb{Q}_v] \log^+ |\underline{\alpha}|_v, \tag{2.1}$$

where $\log^+ x := \log \max\{1, x\}$. It is well-known and easy to verify that the expression on the right is independent of the choice of the field $\mathbb{K}$. The height of $\alpha \in \bar{\mathbb{Q}}$ is, by definition, the height of the one-dimensional vector (α).

For a polynomial f with algebraic coefficients we denote by $\mathrm{h}(f)$ the height of the vector of its coefficients, ordered somehow. More generally, the height $\mathrm{h}(f_1, \dots, f_s)$ of a finite system of polynomials is, by definition, the height of the vector formed by all the non-zero coefficients of all these polynomials.

2.1. Estimates for sums and products of polynomials. The following is an immediate consequence of Lemma 1.2 from [9].

LEMMA 2.1. *Let $f_1, \dots, f_s$ be polynomials in $\bar{\mathbb{Q}}[X_1, \dots, X_n]$ and put*

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

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

(a) $\mathrm{h}(\prod_{i=1}^s f_i) \le \sum_{i=1}^s \mathrm{h}(f_i) + \log(n+1) \sum_{i=1}^{s-1} \deg f_i$,
(b) $\mathrm{h}(g(f_1, \dots, f_s)) \le \mathrm{h}(g) + (h + \log(s+1) + d \log(n+1)) \deg g$. ∎

REMARK 2.2. Item (b) of Lemma 2.1 extends to a slightly more general situation when the polynomial g depends, besides $Y_1, \dots, Y_s$, on some other indeterminates $T_1, \dots, T_r$, but one substitutes new polynomials only for the Y_i's, leaving the T_j's intact. In this case we again have the estimate

$$\mathrm{h}(g(f_1, \dots, f_s, T_1, \dots, T_r)) \le \mathrm{h}(g) + (h + \log(s+1) + d \log(n+1)) \deg_{\underline{Y}} g$$

(independently of r and $\deg_{\underline{T}} g$). Indeed, we can write $g = \sum_k g_k(\underline{Y}) h_k(\underline{T})$, where $h_k(\underline{T})$ are pairwise distinct monomials in $\underline{T} = (T_1, \dots, T_r)$, and apply Lemma 2.1(b) to each g_k.

Here is a particular case of Lemma 2.1, where a slightly sharper estimate holds (see [9, end of Section 1.1.1]).

LEMMA 2.3. *Let $(f_{ij})_{ij}$ be an $s\times s$ matrix of polynomials in $\bar{\mathbb{Q}}[X_1,\dots,X_n]$ of degrees and heights bounded by d and h, respectively. Then*

$$\mathrm{h}(\det(f_{ij})_{ij}) \le s(h+\log s + d\log(n+1)). \ \blacksquare$$

We need one more technical lemma.

LEMMA 2.4. *Let $g(X,Y)\in\bar{\mathbb{Q}}[X,Y]$ be of X-degree m, and fix $\rho\in\bar{\mathbb{Q}}$. Put*

$$f(X,Y) := (X-\rho)^m g((X-\rho)^{-1},Y).$$

Then

$$\mathrm{h}(f) \le \mathrm{h}(g) + m\mathrm{h}(\rho) + 2m\log 2.$$

Proof. The polynomials $g(X,Y)$ and $\widetilde{g}(X,Y) := X^m g(X^{-1},Y)$ have the same coefficients and thereby the same height. Applying Lemma 2.1 and Remark 2.2, we obtain the result. ∎

2.2. Bounds for solutions of algebraic equations. By an *algebraic set* we mean a subset of $\bar{\mathbb{Q}}^N$ defined by a system of polynomial equations. We treat algebraic sets as in [12, 16. Kapitel] (where they are called *algebraische Mannigfaltigkeiten*), that is, purely set-theoretically, without counting multiplicities. By a *component* of an algebraic set we mean an irreducible component.

Let $p_1(\underline{X}),\dots,p_k(\underline{X})$ be polynomials in $\underline{X}=(X_1,\dots,X_N)$ with algebraic coefficients. By an *isolated solution* of the system of polynomial equations

$$p_1(\underline{X}) = \cdots = p_k(\underline{X}) = 0 \tag{2.2}$$

we mean a zero-dimensional component of the algebraic set in $\bar{\mathbb{Q}}^N$ defined by (2.2). (Existence of such a component implies that $k\ge N$.) Our aim is to bound the height of an isolated solution in terms of the degrees and heights of the polynomials $p_1,\dots,p_k$.

Such a bound follows from the arithmetical Bézout inequality due to Bost, Gillet and Soulé [3] and Philippon [10]. Krick, Pardo and Sombra [9] did a great job of producing a user-friendly version of this fundamental result. We very briefly recall some facts from [9] which will be used here. For an affine algebraic set $V\subset\mathbb{A}^N$ defined over $\bar{\mathbb{Q}}$, Krick, Pardo and Sombra [9, Section 1.2] define the *height* of V, to be denoted by $\mathrm{h}_{\mathrm{KPS}}(V)$. We do not reproduce here the full definition of this height function, but only list four of its properties.

PROPOSITION 2.5. *The Krick–Pardo–Sombra height function has the following properties.*

- (Positivity) *For any V we have* $\mathrm{h}_{\mathrm{KPS}}(V)\ge 0$.

- (Additivity) *The height function is "additive" in the following sense: for any V_1 and V_2 without common components we have*
$$\mathrm{h}_{\mathrm{KPS}}(V_1 \cup V_2) = \mathrm{h}_{\mathrm{KPS}}(V_1) + \mathrm{h}_{\mathrm{KPS}}(V_2).$$
- (One-point set) *If $\{\underline{\alpha}\}$ is an algebraic set, then $\mathrm{h}(\underline{\alpha}) \le \mathrm{h}_{\mathrm{KPS}}(\{\underline{\alpha}\})$.*
- (Bézout inequality) *Let V be the algebraic set defined by*
$$p_1(\underline{X}) = \cdots = p_N(\underline{X}) = 0,$$
where $p_i(\underline{X}) \in \bar{\mathbb{Q}}(\underline{X})$ for $i = 1, \dots, N$. Put
$$\nabla = \deg p_1 \cdots \deg p_N, \quad \Sigma = \frac{1}{\deg p_1} + \cdots + \frac{1}{\deg p_N}, \quad h = \max\{\mathrm{h}(p_1), \dots, \mathrm{h}(p_N)\}. \tag{2.3}$$
Then
$$\mathrm{h}_{\mathrm{KPS}}(V) \le \nabla \Sigma h + 2\nabla N \log(N+1). \tag{2.4}$$

Proof. The positivity and additivity follow immediately from the definition. For the height of a one-point set see [9, end of Section 1.2.3]; in fact, $\mathrm{h}_{\mathrm{KPS}}(\{\underline{\alpha}\})$ is defined as the right-hand side of (2.1) but with $\log^+ |\underline{\alpha}|_v$ replaced by $\log(1 + |\alpha_1|_v^2 + \cdots + |\alpha_N|_v^2)^{1/2}$ for archimedean v. Finally, for the Bézout inequality see Corollary 2.11 from [9], or, more precisely, the displayed inequality just before the beginning of Section 2.2.3 on page 555 there. ∎

We adapt the work of Krick, Pardo and Sombra as follows.

Proposition 2.6. *Let K be a number field and let $p_1(\underline{X}), \dots, p_k(\underline{X}) \in \mathbb{K}[\underline{X}]$ be polynomials in $\underline{X} = (X_1, \dots, X_N)$. Let $\underline{\alpha}$ be an isolated solution of (2.2) and $\mathbb{L} = \mathbb{K}(\underline{\alpha})$ the number field generated by the coordinates of $\underline{\alpha}$. Then $k \ge N$. Further, assume that*
$$\deg p_1 \ge \cdots \ge \deg p_k.$$
Let also ∇, Σ be defined as in (2.3) and $h = \max\{\mathrm{h}(p_1), \dots, \mathrm{h}(p_k)\}$. Then
$$[\mathbb{L} : \mathbb{K}] \le \nabla, \tag{2.5}$$
$$[\mathbb{L} : \mathbb{K}]\mathrm{h}(\underline{\alpha}) \le \nabla \Sigma h + 2\nabla N \log(N+1), \tag{2.6}$$
$$\frac{\log \mathcal{N}_{\mathbb{K}/\mathbb{Q}} \mathcal{D}_{\mathbb{L}/\mathbb{K}}}{[\mathbb{L} : \mathbb{Q}]} \le 2\nabla \Sigma h + 5\nabla N \log(N+1), \tag{2.7}$$
where $\mathcal{D}_{\mathbb{L}/\mathbb{K}}$ is the discriminant of $\mathbb{L}$ over $\mathbb{K}$ and $\mathcal{N}_{\mathbb{K}/\mathbb{Q}}$ is the norm map.

The following consequence is immediate.

Corollary 2.7. *In the set-up of Proposition 2.6, denote by V the algebraic subset of $\bar{\mathbb{Q}}^N$ defined by (2.2), and let W be another algebraic subset of $\bar{\mathbb{Q}}^N$ such that the difference set $V \setminus W$ is finite. Then every $\underline{\alpha} \in V \setminus W$ satisfies (2.5)–(2.7).* ∎

For the proof of Proposition 2.6 we shall use the following lemma, due to Silverman [11, Theorem 2].

LEMMA 2.8. *Let $\mathbb{K}$ be a number field and let $\underline{\alpha}$ be a point in $\bar{\mathbb{Q}}^N$. Let $\mathbb{L} = \mathbb{K}(\underline{\alpha})$. Then*

$$\frac{\log \mathcal{N}_{\mathbb{K}/\mathbb{Q}} \mathcal{D}_{\mathbb{L}/\mathbb{K}}}{[\mathbb{L} : \mathbb{Q}]} \le 2([\mathbb{L} : \mathbb{K}] - 1)\mathrm{h}(\underline{\alpha}) + \log[\mathbb{L} : \mathbb{K}]. \ \blacksquare$$

Proof of Proposition 2.6. We denote by V the algebraic set defined by (2.2). Since it has a 0-dimensional component $\underline{\alpha}$, we have $k \ge N$. Among the k polynomials $p_1, \dots, p_k$ one can select N polynomials $q_1, \dots, q_N$ such that $\underline{\alpha}$ is an isolated solution of the system $q_1(\underline{X}) = \dots = q_N(\underline{X})$. The algebraic set defined by this system has at most $\deg q_1 \cdots \deg q_N \le \nabla$ irreducible (over $\bar{\mathbb{Q}}$) components: this follows from the geometric Bézout inequality. In particular, there are at most ∇ isolated solutions. Since a $\mathbb{K}$-conjugate of an isolated solution is again an isolated solution, we must have (2.5).

Further, the four properties in Proposition 2.5 imply that

$$\sum_{\{\underline{\alpha}\} \text{ component of } V} \mathrm{h}(\underline{\alpha}) \le \mathrm{h}_{\mathrm{KPS}}(V) \le \nabla \Sigma h + 2\nabla N \log(N+1), \tag{2.8}$$

where the sum is over the 0-dimensional components of $V(\bar{\mathbb{Q}})$. Since all conjugates of α have the same height, the left side of (2.8) exceeds $[\mathbb{L}:\mathbb{K}]\mathrm{h}(\underline{\alpha})$, which proves (2.6). Combining it with Lemma 2.8, we obtain (2.7). $\blacksquare$

3. Power series. In this section K is a field of characteristic 0 and $f(X, Y) \in K[[X]][Y]$ is a polynomial in Y with coefficients in the ring $K[[X]]$ of formal power series. We denote by ord the order of vanishing at 0. By the *initial segment of length κ* (or simply *κ-initial segment*) of a power series $y = \sum_{k=0}^{\infty} \gamma_k X^k$ we mean $y = \sum_{k=0}^{\kappa} \gamma_k X^k$.

LEMMA 3.1. *Let $\widetilde{y} = \sum_{k=0}^{\kappa} \gamma_k X^k \in K[X]$ be a polynomial in X of degree at most κ. Assume that*

$$\operatorname{ord} f(X, \widetilde{y}) > 2\kappa, \qquad \operatorname{ord} f'_Y(X, \widetilde{y}) = \kappa.$$

Then there exists a unique formal power series $y = \sum_{k=0}^{\infty} \gamma_k X^k \in K[[X]]$ such that $f(X, y) = 0$ and $\widetilde{y}$ is the initial segment of y of length κ.

Proof. By Hensel's Lemma, there exists a unique power series y such that $f(X, y) = 0$ and $\operatorname{ord}(y - \widetilde{y}) > \kappa$. The last inequality implies that $\widetilde{y}$ is the initial segment of y of length κ. $\blacksquare$

LEMMA 3.2.

(a) *Let $y \in K[[X]]$ be a formal power series such that $f(X, y) = 0$. We define $\kappa = \operatorname{ord} f'_Y(X, y)$ and we let $\widetilde{y}$ be the initial segment of y of length κ. Then $\operatorname{ord} f(X, \widetilde{y}) > 2\kappa$ and $\operatorname{ord} f'_Y(X, \widetilde{y}) = \kappa$.*

(b) *Let $y_1, y_2 \in K[[X]]$ be distinct formal power series such that*

$$f(X, y_1) = f(X, y_2) = 0,$$

and let κ_1, κ_2 be defined as κ in (a)*. Then the kth coefficients of y_1 and y_2 are distinct for some $k \le \min\{\kappa_1, \kappa_2\}$.*

Proof. Since $\widetilde{y}$ is the κ-initial segment of y, we have $\operatorname{ord}(y - \widetilde{y}) > \kappa$. Hence

$$f(X, \widetilde{y}) = f(X, y) + f'_Y(X, y)(y - \widetilde{y}) + \text{terms of order} > 2\kappa.$$

Since $f(X, y) = 0$ and $\operatorname{ord} f'_Y(X, y) = \kappa$, the right-hand side is of order $> 2\kappa$. Similarly,

$$f'_Y(X, \widetilde{y}) = f'_Y(X, y) + \text{terms of order} > \kappa,$$

which implies that the right-hand side is of order κ. We have proved part (a).

For (b), Lemma 3.1 implies that y_j is the single power series satisfying $f(X, y_j) = 0$ and having $\widetilde{y}_j$ as an initial segment. Since the series y_1 and y_2 are distinct, none of $\widetilde{y}_j$ can be an initial segment of the other ([4]), whence the result. ∎

LEMMA 3.3. *Suppose K is algebraically closed and let $y_1, \dots, y_\ell \in K[[X]]$ be pairwise distinct formal power series such that*

$$f(X, y_1) = \cdots = f(X, y_\ell) = 0.$$

Assume that the polynomial f is monic in Y (that is, $f = Y^n +$ terms of lower degree in Y) and

$$\sum_{j=1}^{\ell} \operatorname{ord} f'_Y(y_j) = \operatorname{ord} d(X), \tag{3.1}$$

where $d(X)$ is the Y-discriminant of f. Then f splits into linear factors over the ring $K[[X]]$:

$$f(X, Y) = (Y - y_1) \cdots (Y - y_n),$$

where $y_1, \dots, y_n \in K[[X]]$.

Proof. Since f is monic, it splits, by the Puiseux theorem, into linear factors over the ring $K[[X^{1/e}]]$ for some e:

$$f(X, Y) = (Y - y_1) \cdots (Y - y_n),$$

where $y_{\ell+1}, \dots, y_n \in K[[X^{1/e}]]$. Further, $d(X) = \prod_{j=1}^{n} f'_Y(y_j)$, which, together with (3.1) implies that

$$\operatorname{ord} f'_Y(y_j) = 0 \quad (j = \ell + 1, \dots, n). \tag{3.2}$$

([4]) If, say, $\widetilde{y}_1$ is an initial segment of $\widetilde{y}_2$ then the same argument as above shows that $\operatorname{ord} f'_Y(X, \widetilde{y}_2) = \operatorname{ord} f'_Y(X, \widetilde{y}_1)$, that is, $\kappa_1 = \kappa_2$, whence $\widetilde{y}_1 = \widetilde{y}_2$. Lemma 3.1 now implies that $y_1 = y_2$, a contradiction.

If we now write $y_j = a_{j0} + a_{j1}X^{1/e} + \cdots$, then (3.2) implies that

$$\operatorname{ord} f'_Y(X, a_{j0}) = 0 \quad (j = \ell + 1, \ldots, n).$$

Lemma 3.1 now implies that in each of the rings $K[[X]]$ and $K[[X^{1/e}]]$, the polynomial f has exactly one root with initial term a_{j0}. Hence $y_j \in K[[X]]$ for $j = \ell + 1, \ldots, n$, as desired. ∎

4. Miscellaneous lemmas

LEMMA 4.1. *Let $\mathcal{C}$ be a smooth projective curve defined over an algebraically closed field K of characteristic* 0. *Let $x \in K(\mathcal{C})$ have only simple poles, and let $y \in K(\mathcal{C})$ have a single (possibly, multiple) pole which is a pole of x as well. Then $K(\mathcal{C}) = K(x, y)$.*

Proof. Since x has only simple poles in $K(\mathcal{C})$, the place at ∞ of the field $K(x)$ splits completely in $K(\mathcal{C})$. Let P be the pole of y, and let $\widetilde{P}$ be the place of $K(x, y)$ below P. Then $\widetilde{P}$ is above the place at ∞ of $K(x)$. Hence $\widetilde{P}$ also splits completely in $K(\mathcal{C})$.

Now assume that $K(x, y)$ is a proper subfield of $K(\mathcal{C})$. Then there are at least two places of $K(\mathcal{C})$ above $\widetilde{P}$. In particular, there is a place $P' \neq P$ above $\widetilde{P}$. This P' must be a pole of y, a contradiction. ∎

LEMMA 4.2. *Let K be an algebraically closed field of characteristic* 0, *and V a non-empty quasiprojective variety over K. Let $\{(\mathcal{C}_t, D_t) : t \in V\}$ be an algebraic family of curves supplied with an effective divisor. Also, let s be a positive integer. Assume that there exists $\tau \in V$ such that $\mathcal{C}_\tau$ is irreducible and $h^0(D_\tau) = s$. Then the set*

$$\left\{ t \in V : \begin{array}{l} \textit{either } \mathcal{C}_t \textit{ is reducible} \\ \textit{or } \mathcal{C}_t \textit{ is irreducible and } h^0(D_t) > s \end{array} \right\}$$

is not Zariski dense in V.

Proof. This is a consequence of the theorems of Bertini and semi-continuity (see, for instance, Theorem 12.8 in [8, Chapter III]). ∎

LEMMA 4.3. *Given a positive integer n and a finite set* $\mathrm{M} \subset \mathbb{C}$, *there exist only finitely many extensions of the rational function field $\mathbb{C}(x)$ of degree n, unramified outside* M.

Proof. This lemma (which may be viewed as an analogue of the Hermite theorem for function fields) is an immediate consequence of the uniqueness statement of Theorem 1.1. Alternatively, it is a direct consequence of the fact that the fundamental group of a compact Riemann surface is finitely generated. ∎

5. Launching the proof of Theorem 1.2. Let $S \to \mathbb{P}^1(\mathbb{C})$ be a covering as in the statement of Theorem 1.2. According to Theorem 1.1, our covering is isomorphic to $\mathcal{C}(\mathbb{C}) \xrightarrow{x} \mathbb{P}^1(\mathbb{C})$, where $\mathcal{C}$ is a complex algebraic curve and x is a rational function on $\mathcal{C}$. Since all ramification points of the latter covering are algebraic, the curve $\mathcal{C}$ and the function x are definable over $\bar{\mathbb{Q}}$.

We are going to find a number field $\mathbb{L} \supset \mathbb{K}$, a rational function $y \in \mathbb{L}(\mathcal{C})$ such that $\bar{\mathbb{Q}}(\mathcal{C}) = \bar{\mathbb{Q}}(x, y)$, and an absolutely irreducible polynomial $f(X, Y) \in \mathbb{L}[X, Y]$ such that $f(x, y) = 0$ and the degrees $\deg_X f$, $\deg_Y f$, the height $\mathrm{h}(f)$, as well as the degree $[\mathbb{L} : \mathbb{K}]$ and the relative discriminant of $\mathbb{L}/\mathbb{K}$ satisfy the required (in)equalities. To achieve this, we define algebraic sets V and W in a high-dimensional affine space such that the set $V \setminus W$ contains a point having the coefficients of f as part of its coordinates. We then show that the set $V \setminus W$ is finite and use Corollary 2.7 to bound its elements (and thereby the coefficients of f). As a by-product, we will also bound the degree and the discriminant of the field generated by the coefficients.

We write

$$\mathrm{M} = \{\alpha_1, \dots, \alpha_\mu\}.$$

For the main part of the proof we shall assume that the curve $\mathcal{C}$ is unramified over ∞ (that is, ∞ is not one of the points $\alpha_1, \dots, \alpha_\mu$), and that $\mathcal{C}$ has no Weierstrass point above ∞. In other words, the poles of x are neither ramified nor Weierstrass. The general case easily reduces to this one (see Section 15).

Now we start the detailed proof. Since it is going to be long and involved, we divide it into short logically complete steps.

6. The function y and polynomial $f(X, Y)$. Fix a pole P of x. Since P is not a Weierstrass point of $\mathcal{C}$, we have

$$h^0(mP) = 2, \quad h^0((m-1)P) = 1$$

with $m = \mathbf{g}(\mathcal{C}) + 1$.

Since x is unramified above the infinity, x^{-1} can serve as a local parameter at P. If y belongs to $H^0(mP)$, but not to $H^0((m-1)P)$, then y has the Puiseux expansion at P of the form $\sum_{k=-m}^{\infty} c_k x^{-k}$ with $c_{-m} \neq 0$. Since $h^0(mP) = 2$, there exists a unique $y \in H^0(mP)$ with the properties

$$c_{-m} = 1, \quad c_0 = 0. \tag{6.1}$$

In the sequel y will be the function satisfying these conditions.

The function y has a single pole P which is a pole of x as well. Lemma 4.1 implies now that $\bar{\mathbb{Q}}(\mathcal{C}) = \bar{\mathbb{Q}}(x, y)$ (here we use the assumption that x is unramified above ∞). Also, since y has no poles outside the poles of x,

it is integral over the ring $\bar{\mathbb{Q}}[x]$. Hence, there exists a unique absolutely irreducible polynomial $f(X,Y) \in \bar{\mathbb{Q}}[X,Y]$, monic in Y, such that $f(x,y)=0$ and

$$\deg_Y f = [\bar{\mathbb{Q}}(\mathcal{C}) : \bar{\mathbb{Q}}(x)] = n.$$

We also have

$$\deg_X f = [\bar{\mathbb{Q}}(\mathcal{C}) : \bar{\mathbb{Q}}(y)] = \deg (y)_\infty = m,$$

where $(y)_\infty = mP$ is the divisor of poles of y. We write

$$f(X,Y) = Y^n + \sum_{j=0}^{n-1} \sum_{i=0}^{m} \theta_{ij} X^i Y^j. \tag{6.2}$$

7. Discriminant, its roots, and Puiseux expansions. Let $d(X)$ be the discriminant of $f(X,Y)$ with respect to Y. Every α_i is a root of $d(X)$. Besides the α_i's, the polynomial $d(X)$ may have other roots; we denote them $\beta_1, \dots, \beta_\nu$. Thus, we have

$$d(X) = \delta \prod_{i=1}^{\mu} (X-\alpha_i)^{\sigma_i} \prod_{i=1}^{\nu} (X-\beta_i)^{\tau_i}, \tag{7.1}$$

where $\delta \in \bar{\mathbb{Q}}^*$ and where σ_i and τ_i are positive integers.

Now fix $i \in \{1, \dots, \nu\}$. Since x is unramified over β_i, the function y has n Puiseux expansions at β_i of the form

$$y_{ij} = \sum_{k=0}^{\infty} \gamma_{ijk} (x-\beta_i)^k \quad (j = 1, \dots, n).$$

We put

$$\kappa_{ij} = \operatorname{ord}_{\beta_i} f'_Y(x, y_{ij}).$$

Then

$$\kappa_{i1} + \cdots + \kappa_{in} = \tau_i. \tag{7.2}$$

We may assume that $\kappa_{i1} \geq \cdots \geq \kappa_{in}$ and we define ℓ_i from the condition

$$\kappa_{i\ell_i} > 0, \quad \kappa_{ij} = 0 \quad \text{for } j > \ell_i. \tag{7.3}$$

Then (7.2) reads

$$\sum_{j=1}^{\ell_i} \kappa_{ij} = \tau_i, \tag{7.4}$$

which implies that

$$\sum_{\substack{1\leq i\leq \nu \\ 1\leq j\leq \ell_i}} (\kappa_{ij}+1) \leq \sum_{\substack{1\leq i\leq \nu \\ 1\leq j\leq \ell_i}} 2\kappa_{ij} = 2(\tau_1 + \cdots + \tau_\nu) \leq 2 \deg d(X). \tag{7.5}$$

This inequality will be used in Section 9.

We also let $\widetilde{y}_{ij}$ be the initial segment of the series y_{ij} of length κ_{ij}:

$$\widetilde{y}_{ij} = \sum_{k=0}^{\kappa_{ij}} \gamma_{ijk}(x-\beta_i)^k. \tag{7.6}$$

By Lemma 3.2 we have

$$\operatorname{ord}_{\beta_i} f(x, \widetilde{y}_{ij}) > 2\kappa_{ij}, \qquad \operatorname{ord}_{\beta_i} f'_Y(x, \widetilde{y}_{ij}) = \kappa_{ij}.$$

Lemma 3.2 also implies that, for every fixed i, none of $\widetilde{y}_{i1}, \dots, \widetilde{y}_{in}$ is an initial segment of any other. In other words, for any distinct $j_1, j_2 \in \{1, \dots, n\}$ there exists a non-negative integer $\lambda(i, j_1, j_2) \le \min\{\kappa_{ij_1}, \kappa_{ij_2}\}$ such that

$$\gamma_{ij_1\lambda(i,j_1,j_2)} \neq \gamma_{ij_2\lambda(i,j_1,j_2)}.$$

8. Expansions at infinity. We also have the Puiseux expansions of y at infinity:

$$\begin{aligned} y_{\infty j} &= \sum_{k=0}^{\infty} \gamma_{\infty jk} x^{-k} \qquad (j = 2, \dots, n), \\ y_{\infty 1} &= \sum_{k=-m}^{\infty} \gamma_{\infty 1k} x^{-k}. \end{aligned} \tag{8.1}$$

We define the polynomials

$$g(T, Y) = T^m f(T^{-1}, Y), \qquad h(T, Y) = T^{m(n+1)} f(T^{-1}, T^{-m}Y)$$

and put $t = x^{-1}$, so that the expansions (8.1) can be written in powers of t. Now we define the numbers

$$\begin{aligned} \kappa_{\infty j} &= \operatorname{ord}_{t=0} g'_Y(t, y_{\infty j}) \qquad (j = 2, \dots, n), \\ \kappa_{\infty 1} &= \operatorname{ord}_{t=0} h'_Y(t, t^m y_{\infty 1}). \end{aligned}$$

We have $h(T, T^m Y) = T^{mn} g(T, Y)$, whence

$$\kappa_{\infty 1} = mn + \operatorname{ord}_{t=0} g'_Y(t, y_{\infty 1}).$$

Hence the sum $\kappa_{\infty 1} + \kappa_{\infty 2} + \cdots + \kappa_{\infty n}$ is bounded by mn plus the order at $T = 0$ of the Y-discriminant of $g(T, Y)$. Bounding the latter order by the degree of this discriminant, we obtain

$$\kappa_{\infty 1} + \kappa_{\infty 2} + \cdots + \kappa_{\infty n} \le mn + \deg d(X). \tag{8.2}$$

Putting

$$\ell_\infty = n, \tag{8.3}$$

we rewrite (8.2) as

$$\sum_{1 \le j \le \ell_\infty} (\kappa_{\infty j} + 1) \le (m+1)n + \deg d(X). \tag{8.4}$$

This will be used in Section 9.

Further, for $j = 2, \dots, n$ let $\widetilde{y}_{\infty j}$ be the initial segment of the series $y_{\infty j}$ of length $\kappa_{\infty j}$, and let $\widetilde{y}_{\infty 1}$ be the initial segment of $y_{\infty 1}$ of length $\kappa_{\infty 1}$:

$$\widetilde{y}_{\infty j} = \sum_{k=0}^{\kappa_{\infty j}} \gamma_{\infty jk} t^k \quad (j = 2, \dots, n), \tag{8.5}$$

$$\widetilde{y}_{\infty 1} = \sum_{k=-m}^{\kappa_{\infty 1}-m} \gamma_{\infty 1k} t^k. \tag{8.6}$$

Then we have

$$\operatorname{ord}_{t=0} g(t, \widetilde{y}_{\infty j}) > 2\kappa_{\infty j}, \qquad \operatorname{ord}_{t=0} g'_Y(t, \widetilde{y}_{\infty j}) = \kappa_{\infty j} \quad (j = 2, \dots, n),$$
$$\operatorname{ord}_{t=0} h(t, t^m \widetilde{y}_{\infty 1}) > 2\kappa_{\infty 1}, \quad \operatorname{ord}_{t=0} h'_Y(t, t^m \widetilde{y}_{\infty 1}) = \kappa_{\infty 1}.$$

Identities (6.1) now become

$$\gamma_{\infty 1,-m} = 1, \qquad \gamma_{\infty 10} = 0.$$

As in the finite case, for any distinct $j_1, j_2 \in \{2, \dots, n\}$ there exists a non-negative integer $\lambda(\infty, j_1, j_2) \le \min\{\kappa_{\infty j_1}, \kappa_{\infty j_2}\}$ such that

$$\gamma_{\infty j_1 \lambda(\infty, j_1, j_2)} \ne \gamma_{\infty j_2 \lambda(\infty, j_1, j_2)}.$$

9. Indeterminates. We consider the vector

$$\varphi = (\underline{\theta}, \underline{\alpha}, \underline{\beta}, \underline{\gamma}, \delta) \in \bar{\mathbb{Q}}^{\Omega},$$

where the dimension Ω is defined below in (9.1). Here:

- $\underline{\theta} = (\theta_{ij})_{\substack{0 \le i \le m \\ 0 \le j \le n-1}}$ is the vector of coefficients of f (see (6.2));
- $\underline{\alpha} = (\alpha_i)_{1 \le i \le \mu}$ and $\underline{\beta} = (\beta_i)_{1 \le i \le \nu}$ are the vectors of roots of the discriminant $d(X)$, and δ is its leading coefficient (see (7.1));
- $\underline{\gamma} = (\underline{\gamma}_{ij})_{\substack{i \in \{1, \dots, \nu, \infty\} \\ 1 \le j \le \ell_i}}$, where ℓ_i are defined in (7.3) and (8.3), and $\underline{\gamma}_{ij}$ is the vector of coefficients of the initial segment $\widetilde{y}_{ij}$ of the Puiseux expansion y_{ij} (see (7.6), (8.5) and (8.6)); that is, $\underline{\gamma}_{ij} = (\gamma_{ijk})_{0 \le k \le \kappa_{ij}}$ for $(i, j) \ne (\infty, 1)$ and $\underline{\gamma}_{\infty 1} = (\gamma_{\infty 1k})_{-m \le k \le \kappa_{\infty 1} - m}$.

We are only interested in the vectors $\underline{\theta}$ and $\underline{\alpha}$, but we cannot study them separately from the other vectors defined above.

The dimension Ω is defined by

$$\Omega = (m+1)n + \mu + \nu + \sum_{\substack{1 \le i \le \nu \\ 1 \le j \le \ell_i}} (\kappa_{ij} + 1) + \sum_{1 \le j \le \ell_\infty} (\kappa_{\infty j} + 1) + 1. \tag{9.1}$$

We have

$$\Omega \le 2(m+1)n + 4 \deg d(X) + 1 \le 10mn + 2n - 8m + 1, \tag{9.2}$$

where we use (7.5), (8.4) and the estimates $\mu + \nu \le \deg d(X) \le 2m(n-1)$.

We shall define algebraic sets V and W in $\bar{\mathbb{Q}}^{\Omega}$ such that $\varphi \in V \setminus W$ and $V \setminus W$ is finite. This will allow us to use Corollary 2.7 to bound the height of φ. This would imply a bound on the height of $\underline{\theta}$, which is the height of the polynomial f.

To define our algebraic sets, we introduce the vector of indeterminates Φ whose coordinates correspond to the coordinates of φ:

$$\Phi = (\underline{\Theta}, \underline{\mathrm{A}}, \underline{\mathrm{B}}, \underline{\Gamma}, \Delta),$$

where

$$\underline{\Theta} = (\Theta_{ij})_{\substack{0\le i\le m\\ 0\le j\le n-1}}, \quad \underline{\mathrm{A}} = (\mathrm{A}_i)_{1\le i\le \mu}, \quad \underline{\mathrm{B}} = (\mathrm{B}_i)_{1\le i\le \nu}, \quad \underline{\Gamma} = (\underline{\Gamma}_{ij})_{\substack{i\in\{1,\dots,\nu,\infty\}\\ 1\le j\le \ell_i}}$$

with

$$\underline{\Gamma}_{ij} = (\Gamma_{ijk})_{0\le k\le \kappa_{ij}} \quad \text{for } (i,j)\neq(\infty,1), \qquad \underline{\Gamma}_{\infty 1} = (\Gamma_{\infty 1k})_{-m\le k\le \kappa_{\infty 1}-m}.$$

10. The algebraic set V. The first series of equations defining the algebraic set V is

$$\mathrm{A}_i = \alpha_i \quad (i = 1, \dots, \mu). \tag{10.1}$$

To write down the rest of the equations we introduce the polynomials $F(X,Y)$, $D(X)$, $G(T,Y)$ and $H(T,Y)$ with coefficients in $\mathbb{Z}[\underline{\Theta}]$, which correspond to the polynomials $d(X)$, $g(T,Y)$ and $h(T,Y)$ from Sections 7 and 8. More specifically, we put

$$F(X,Y) = Y^n + \sum_{j=0}^{n-1}\sum_{i=0}^{m} \Theta_{ij} X^i Y^j \in \mathbb{Z}[\underline{\Theta}][X,Y],$$

we define $D(X)$ to be the Y-discriminant of $F(X,Y)$ and we put

$$G(T,Y) = T^m F(T^{-1},Y), \qquad H(T,Y) = T^{m(n+1)} F(T^{-1}, T^{-m}Y).$$

The second series of equations comes from the equality

$$D(X) = \Delta \prod_{i=1}^{\mu} (X - \mathrm{A}_i)^{\sigma_i} \prod_{i=1}^{\nu} (X - \mathrm{B}_i)^{\tau_i}, \tag{10.2}$$

where the quantities σ_i and τ_i are defined in (7.1). In order to define the third set of equations we introduce the polynomials

$$\widetilde{Y}_{ij} = \sum_{k=0}^{\kappa_{ij}} \Gamma_{ijk} (X - \mathrm{B}_i)^k \qquad (1 \le i \le \nu,\ 1 \le j \le \ell_i),$$

$$\widetilde{Y}_{\infty j} = \sum_{k=0}^{\kappa_{\infty j}} \Gamma_{\infty jk} T^k \qquad (2 \le j \le \ell_\infty)$$

and the Laurent polynomial

$$\widetilde{Y}_{\infty 1} = \sum_{k=-m}^{\kappa_{\infty 1}-m} \Gamma_{\infty 1 k} T^k.$$

The equations come from the relations

$$\begin{aligned} \operatorname{ord}_{X=\mathrm{B}_i} F(X, \widetilde{Y}_{ij}) &> 2\kappa_{ij} \\ \operatorname{ord}_{X=\mathrm{B}_i} F'_Y(X, \widetilde{Y}_{ij}) &\geq \kappa_{ij} \end{aligned} \qquad (1 \leq i \leq \nu,\ 1 \leq j \leq \ell_i), \tag{10.3}$$

$$\begin{aligned} \operatorname{ord}_{T=0} G(T, \widetilde{Y}_{\infty j}) &> 2\kappa_{\infty j} \\ \operatorname{ord}_{T=0} G'_Y(T, \widetilde{Y}_{\infty j}) &\geq \kappa_{\infty j} \end{aligned} \qquad (2 \leq j \leq \ell_\infty), \tag{10.4}$$

$$\begin{aligned} \operatorname{ord}_{T=0} H(T, T^m \widetilde{Y}_{\infty 1}) &> 2\kappa_{\infty 1} \\ \operatorname{ord}_{T=0} H'_Y(T, T^m \widetilde{Y}_{\infty 1}) &\geq \kappa_{\infty 1}. \end{aligned} \tag{10.5}$$

The final two equations are

$$\Gamma_{\infty 1,-m} = 1, \quad \Gamma_{\infty 10} = 0. \tag{10.6}$$

The following statement is immediate in view of the definitions and properties from Sections 7 and 8.

PROPOSITION 10.1. *The vector φ belongs to the set V.* ∎

11. The algebraic set W. We write

$$W = W_1 \cup W_2 \cup W_3 \cup W_4 \cup W_5 \cup W_6,$$

where the sets $W_1, \ldots, W_6$ are defined below.

The set W_1 is defined by $\Delta = 0$. Next, put

$$W_2 = \bigcup_{\substack{1 \leq i \leq \mu \\ 1 \leq j \leq \nu}} W_2^{(ij)}, \quad W_3 = \bigcup_{1 \leq i < j \leq \nu} W_3^{(ij)},$$

where $W_2^{(ij)}$ is defined by $\mathrm{A}_i = \mathrm{B}_j$ and $W_3^{(ij)}$ is defined by $\mathrm{B}_i = \mathrm{B}_j$.

Further, we put

$$W_4 = \bigcup_{\substack{i \in \{1,\ldots,\nu,\infty\} \\ 1 \leq j \leq \ell_i}} W_4^{(ij)},$$

where the set $W_4^{(ij)}$ is defined by the relations

$$\operatorname{ord}_{X=\mathrm{B}_i} F'_Y(X, \widetilde{Y}_{ij}) > \kappa_{ij} \quad \text{when } i \neq \infty, \tag{11.1}$$

$$\operatorname{ord}_{T=0} G'_Y(T, \widetilde{Y}_{\infty j}) > \kappa_{\infty j} \quad \text{when } i = \infty \text{ and } j \neq 1, \tag{11.2}$$

$$\operatorname{ord}_{T=0} H'_Y(T, T^m \widetilde{Y}_{\infty 1}) > \kappa_{\infty 1} \quad \text{when } (i,j) = (\infty, 1). \tag{11.3}$$

Further, we put

$$W_5 = \Big(\bigcup_{\substack{1\le i\le \nu \\ 1\le j_1<j_2\le \ell_i}} W_5^{(ij_1j_2)} \Big) \cup \Big(\bigcup_{2\le j_1<j_2\le \ell_\infty} W_5^{(\infty j_1 j_2)} \Big),$$

where $W_5^{(ij_1j_2)}$ is defined by $\Gamma_{ij_1\lambda(i,j_1,j_2)} = \Gamma_{ij_2\lambda(i,j_1,j_2)}$ and $W_5^{(\infty j_1 j_2)}$ is defined by $\Gamma_{\infty j_1\lambda(\infty,j_1,j_2)} = \Gamma_{\infty j_2\lambda(\infty,j_1,j_2)}$, the numbers $\lambda(i,j_1,j_2)$ being defined at the end of Sections 7 and 8.

Finally, Lemma 4.2 implies that there is a proper Zariski-closed subset W_6 of V such that $\varphi \notin W_6$ and for any $\widehat{\varphi} = (\underline{\widehat{\theta}}, \underline{\widehat{\alpha}}, \underline{\widehat{\beta}}, \underline{\widehat{\gamma}}, \widehat{\delta}) \in V \setminus W_6$ the polynomial

$$Y^n + \sum_{j=0}^{n-1} \sum_{i=0}^{m} \widehat{\theta}_{ij} X^i Y^j \tag{11.4}$$

is irreducible and has the following property. Let $\widehat{x}$ and $\widehat{y}$ be the coordinate functions on the curve $\widehat{\mathcal{C}}$ defined by (11.4). Then the effective divisor $(\widehat{y})_\infty$ satisfies $h^0((\widehat{y})_\infty) = 2$.

The following statement is again immediate.

PROPOSITION 11.1. *The vector φ does not belong to the set W.* ∎

12. Finiteness of $V \setminus W$. Here we prove that the set $V \setminus W$ is finite. Let $\widehat{\varphi} = (\underline{\widehat{\theta}}, \underline{\widehat{\alpha}}, \underline{\widehat{\beta}}, \underline{\widehat{\gamma}}, \widehat{\delta})$ be a point in $V \setminus W$. Then $\underline{\widehat{\alpha}} = \underline{\alpha}$ because of (10.1).

Put

$$\widehat{f}(X,Y) = Y^n + \sum_{j=0}^{n-1} \sum_{i=0}^{m} \widehat{\theta}_{ij} X^i Y^j.$$

It is a $\bar{\mathbb{Q}}$-irreducible polynomial (because $\widehat{\varphi} \notin W_6$) and defines an algebraic curve $\widehat{\mathcal{C}}$ together with rational functions $\widehat{x}, \widehat{y} \in \bar{\mathbb{Q}}(\widehat{\mathcal{C}})$ satisfying $\widehat{f}(\widehat{x}, \widehat{y}) = 0$. Notice that this implies that $\widehat{y}$ is integral over $\bar{\mathbb{Q}}[\widehat{x}]$.

Let $\widehat{d}(X)$ be the Y-discriminant of $\widehat{f}(X,Y)$. Then

$$\widehat{d}(X) = \widehat{\delta} \prod_{i=1}^{\mu} (X - \alpha_i)^{\sigma_i} \prod_{i=1}^{\nu} (X - \widehat{\beta}_i)^{\tau_i}$$

because $\widehat{\varphi}$ satisfies (10.2). Since $\widehat{\varphi} \notin W_2 \cup W_3$, the numbers $\widehat{\beta}_i$ are pairwise distinct and also distinct from every α_i.

The covering $\widehat{\mathcal{C}} \xrightarrow{\widehat{x}} \mathbb{P}^1$ can be ramified only over the roots of $\widehat{d}(X)$, and, perhaps, over infinity. We want to show that $\widehat{x}$ is unramified over the numbers $\widehat{\beta}_i$ and over infinity.

Fix a root $\widehat{\beta}_i$ and define

$$\widetilde{\widehat{y}}_{ij}(X) = \sum_{k=0}^{\kappa_{ij}} \widehat{\gamma}_{ijk}(X - \widehat{\beta}_i)^k \quad (j = 1, \dots, \ell_i). \tag{12.1}$$

Then

$$\operatorname{ord}_{\widehat{\beta}_i} \widehat{f}(X, \widetilde{\widehat{y}}_{ij}) > 2\kappa_{ij}, \quad \operatorname{ord}_{\widehat{\beta}_i} \widehat{f}'_Y(X, \widetilde{\widehat{y}}_{ij}) = \kappa_{ij},$$

because $\widehat{\varphi}$ satisfies (10.3) and does not satisfy (11.1). Also, none of $\widetilde{\widehat{y}}_{ij}$ is an initial segment of another, because $\widehat{\varphi} \notin W_5$.

Using Lemma 3.1, we find ℓ_i pairwise distinct Puiseux expansions

$$\widehat{y}_{i1}, \dots, \widehat{y}_{i\ell_i} \in \bar{\mathbb{Q}}[[X - \widehat{\beta}_i]]$$

of $\widehat{y}$ at $\widehat{\beta}_i$ satisfying $\operatorname{ord}_{\widehat{\beta}_i} \widehat{f}'_Y(X, \widehat{y}_{ij}) = \kappa_{ij}$. Since

$$\sum_{j=1}^{\ell_i} \operatorname{ord}_{\widehat{\beta}_i} \widehat{f}'_Y(X, \widehat{y}_{ij}) = \sum_{j=1}^{\ell_i} \kappa_{ij} = \tau_i = \operatorname{ord}_{\widehat{\beta}_i} \widehat{d}(X)$$

by (7.4), Lemma 3.3 implies that all n Puiseux expansions of $\widehat{x}$ at $\widehat{\beta}_i$ are in $\bar{\mathbb{Q}}[[X - \widehat{\beta}_i]]$, which means that $\widehat{x}$ is unramified over $\widehat{\beta}_i$.

In a similar way we prove that $\widehat{x}$ is unramified over infinity (here $\ell_\infty = n$ and we do not need Lemma 3.3). Moreover, $\widehat{y}$ has at infinity $n-1$ Puiseux expansions without negative powers and one expansion starting from degree $-m$. Since $\widehat{y}$ is integral over $\bar{\mathbb{Q}}[\widehat{x}]$, we have $(\widehat{y})_\infty = m\widehat{P}$, where $\widehat{P}$ is a pole of $\widehat{x}$. Since $\widehat{\varphi} \notin W_6$, we have $h^0(m\widehat{P}) = 2$.

Thus, each $\widehat{\varphi} \in V \setminus W$ gives rise to a pair $(\widehat{\mathcal{C}}, \widehat{x})$, where $\widehat{\mathcal{C}}$ is an algebraic curve and $\widehat{x}$ a rational function on $\widehat{\mathcal{C}}$ of degree n, unramified outside the points α_i. By Lemma 4.3, there are only finitely many possibilities for $(\widehat{\mathcal{C}}, \widehat{x})$. Fix one. Since $h^0(m\widehat{P}) = 2$, the function $\widehat{y}$ is uniquely defined by the equations (10.6). It follows that the polynomial $\widehat{f}$ is uniquely defined as well. Hence so is $\widehat{\delta}$, and the vector $\underline{\widehat{\beta}}$ is uniquely defined up to ordering its components. Having this order fixed, we find that $\underline{\widehat{\gamma}}$ is uniquely defined.

This proves that the set $V \setminus W$ is finite.

13. Estimating the equations defining V. In this section we estimate the degrees and heights of the equations defining the algebraic set V.

Since $\kappa_{ij} \le \deg d(X) \le 2m(n-1)$, the equations defined by (10.3) are of degree at most

$$n(2m(n-1)+1) + 1 \le 2mn^2.$$

Here the "1" inside the parentheses is the degree of $\widetilde{Y}_{ij}$ in $\underline{\Gamma}$, and the "1" outside the parentheses is the degree of F (and of F'_Y) in $\underline{\Theta}$.

A straightforward verification shows that the degrees of the other equations are bounded by $2mn^2$ as well.

Now let us estimate the heights of the equations. The heights of the μ equations (10.1) are obviously bounded by $h = \max\{\mathrm{h}(\alpha_1), \dots, \mathrm{h}(\alpha_\mu)\}$.

Estimating the heights of the remaining equations can be done with Lemma 2.1. All of the polynomials occurring below have rational integer coefficients. We define the *size* of a polynomial p with coefficients in $\mathbb{Z}$ (denoted by $\|p\|$) to be the sup-norm of the vector of its coefficients. For a non-zero polynomial p we have $\mathrm{h}(p) \le \log \|p\|$, with equality if the coefficients are coprime. In particular, $\mathrm{h}(p) = 0$ if p is of size 1, which is the case for many polynomials below.

The left-hand side of (10.2) is a determinant of order $2n-1$ whose entries are polynomials in the $n(m+1)+1$ variables X and $\underline{\Theta}$, each entry being of degree at most $m+1$ and of size at most n. Hence its height can be estimated using Lemma 2.3:

$$\mathrm{h}(D) \le (2n-1)\big(\log n + \log(2n-1) + (m+1)\log(n(m+1)+2)\big) \le 10(mn)^2.$$

The right-hand side of (10.2) is a product of at most $2m(n-1)$ polynomials of degree 1 and size 1 in $\mu+\nu+1$ variables $\underline{\mathrm{A}}$, $\underline{\mathrm{B}}$ and X. Lemma 2.1(a) allows us to estimate the height of the right-hand side by $2m(n-1)\log(\nu+\mu+1)$ $\le 5(mn)^2$. We thereby bound the heights of the equations coming from (10.2) by $10(mn)^2$.

Equations (10.6) are, obviously, of height 0. The heights of the equations coming from (10.3)–(10.5) can be estimated using Lemma 2.1(b). For $i \ne \infty$ the polynomial $\widetilde{Y}_{ij}$ is in the $\kappa_{ij}+2 \le 2mn$ variables X, B_j, $\underline{\Gamma}_{ij}$. It is of degree $\kappa_{ij}+1 \le 2mn-1$ and of size bounded by $2^{\kappa_{ij}} \le 4^{mn}$. Lemma 2.1(b) together with Remark 2.2 bounds the height of the polynomials $F(X, \widetilde{Y}_{ij})$ and $F'_Y(X, \widetilde{Y}_{ij})$ by the quantities

$$(mn\log 4 + \log 2 + 2mn\log(2mn+1))(m+n)$$

and

$$\log n + (mn\log 4 + \log 2 + 2mn\log(2mn+1))(m+n-1),$$

respectively. Both are at most $6(mn)^3$, which bounds the heights of the equations coming from (10.3). Similarly, one bounds by $12(mn)^3$ the heights of the equations coming from (10.4) and (10.5).

Finally, we summarize all these calculations in the following proposition.

PROPOSITION 13.1. *The algebraic set V is defined by equations of degree bounded by $2mn^2$ and height bounded by $h + 12(mn)^3$.*

14. The height of φ and the field $\mathbb{K}(\varphi)$. Now we may apply Proposition 2.6, or, more precisely, Corollary 2.7 to bound the height of the vector φ,

and the number field generated by its coordinates. Recall that φ belongs to $\bar{\mathbb{Q}}^{\Omega}$, where the dimension Ω satisfies

$$\Omega \le 10mn + 2n - 7$$

(see (9.2)). If we define ∇ and Σ as in Proposition 2.6, we will have

$$\mathrm{h}(f) \le \mathrm{h}(\varphi) \le \nabla\Sigma(h + 12(mn)^3) + 2\nabla\Omega\log(\Omega+1).$$

Furthermore, the field $\mathbb{L} = \mathbb{K}(\varphi)$ satisfies $[\mathbb{L}:\mathbb{K}] \le \nabla$ and

$$\frac{\mathcal{N}_{\mathbb{K}/\mathbb{Q}}\mathcal{D}_{\mathbb{L}/\mathbb{K}}}{[\mathbb{L}:\mathbb{Q}]} \le 2\nabla\Sigma(h + 12(mn)^3) + 5\nabla\Omega\log(\Omega+1).$$

Since the degrees of the equations defining V are bounded by $2mn^2$, we have

$$\nabla \le (2mn^2)^{\Omega} \le (2mn^2)^{10mn+2n-7}.$$

Obviously, $\Sigma \le \Omega \le 12mn$. After trivial calculations we obtain

$$\mathrm{h}(f) \le \Lambda'(h+1), \quad [\mathbb{L}:\mathbb{K}] \le \Lambda', \quad \frac{\log \mathcal{N}_{\mathbb{K}/\mathbb{Q}}\mathcal{D}_{\mathbb{L}/\mathbb{K}}}{[\mathbb{L}:\mathbb{Q}]} \le \Lambda'(h+1) \tag{14.1}$$

with $\Lambda' = (2mn^2)^{10mn+2n-3}$. Since $m = \mathbf{g}+1$, this proves Theorem 1.2 in the case when there are no ramified points and no Weierstrass points among the poles of x.

15. The general case. We no longer assume that the set of poles of x has no Weierstrass and no ramified points (called *bad* points in what follows). Since there exist at most $\mathbf{g}^3 - \mathbf{g}$ Weierstrass points and at most $2\mathbf{g}$ ramified points, there exists $\rho \in \mathbb{Z}$ satisfying

$$|\rho| \le \mathbf{g}^3 + \mathbf{g} \le m^3$$

(recall that $m = \mathbf{g}+1$) such that the fiber of x above ρ contains no bad points. It follows that the function $\check{x} = (x-\rho)^{-1}$ has no bad points among its poles, and the previous argument applies to it. We find a number field $\mathbb{L}$, a rational function $y \in \mathbb{L}(\mathcal{C})$ such that $\mathbb{L}(\mathcal{C}) = \mathbb{L}(\check{x}, y)$, and a polynomial $\check{f}(X,Y) \in \mathbb{L}[X,Y]$ such that $\check{f}(\check{x}, y) = 0$,

$$\deg_X \check{f} = m = \mathbf{g}+1, \quad \deg_Y \check{f} = n,$$

and (14.1) holds with f replaced by $\check{f}$ and h replaced by

$$\check{h} := \max\{\mathrm{h}((\alpha_1-\rho)^{-1}), \dots, \mathrm{h}((\alpha_\mu-\rho)^{-1})\}.$$

Obviously

$$\check{h} \le h + \log(2\max\{1,|\rho|\}) \le h + 3\log(2m),$$

which proves (1.2) after a short calculation. Further, the polynomial

$$f(X,Y) := (X-\rho)^m \check{f}((X-\rho)^{-1}, Y)$$

satisfies $f(x,y)=0$ and

$$\mathrm{h}(f) \le \mathrm{h}(\check{f}) + 3m\log(2m)$$

by Lemma 2.4. Again a trivial calculation implies (1.1). Theorem 1.2 is completely proved. ∎

16. On the work of Zverovich. As we already indicated in the introduction, the prototype of our proof is the work of Zverovich [13]. Given a covering $\mathcal{C} \xrightarrow{x} \mathbb{P}^1$ and a point $\alpha \in \mathbb{P}^1$, define the *total ramification* of x at α to be the quantity

$$e(\alpha) = e_x(\alpha) = (e_1 - 1) + \cdots + (e_s - 1),$$

where $e_1, \dots, e_s$ are the ramification indices of x over α. In particular, $e(\alpha) > 0$ if and only if x is ramified over α.

Loosely, Zverovich's argument is as follows. He defines x, y and the polynomial f in (almost) the same way as we do. Then, denoting by $d(X)$ the Y-discriminant of f, one has the equality

$$d(X) = \prod_{i=1}^{\mu} (X - \alpha_i)^{e(\alpha_i)} \psi(X)^2,$$

where ψ is a polynomial. Zverovich considers the equations which follow from the relation

$$D(X) = \prod_{i=1}^{\mu} (X - \alpha_i)^{e(\alpha_i)} \Psi(X)^2, \tag{16.1}$$

where the unknowns are the coefficients of the variable polynomials F and Ψ, and, as in our argument, $D(X)$ is the Y-discriminant of the variable polynomial F. He adds to this two equations similar to our normalization equations (10.6). He observes that (f, ψ) satisfies his system of equations, and wants to prove that the system has finitely many equations.

Unfortunately, Zverovich's proof of finiteness seems to be incomplete. In fact, he implicitly assumes that, for any solution $(\hat{f}, \widehat{\psi})$ of (16.1), the curve $\hat{\mathcal{C}}$ defined by $\hat{f}(X,Y)=0$ is ramified over the points $\alpha_1, \dots, \alpha_\mu$, and moreover the total ramification is the same as for our curve. If this were true, then Zverovich would have correctly proved that there is no other ramification, and Lemma 4.3 would imply finiteness. The problem is that a curve defined by a polynomial satisfying Zverovich's equations is not obliged *a priori* to have the same ramification at the points $\alpha_1, \dots, \alpha_\mu$ as our curve, and without this his argument does not seem to work.

We failed to repair Zverovich's argument and had to invent another system of equations defining our polynomial f, which is much more complicated than his. It would be of interest to reconsider his work and justify his ar-

gument. This would not only improve on the estimates of this article, but would also probably imply a relatively practical algorithm (see [6] for some indications) for actual calculation of the polynomial f. Evidently, our equations are too bulky for this purpose.

Acknowledgements. We thank Anna Cadoret, Pierre Dèbes, Bas Edixhoven, Carlo Gasbarri and Martin Sombra for helpful discussions.

References

[1] Yu. Bilu,*Effective analysis of integral points on algebraic curves*, Ph.D. Thesis, Beer Sheva, 2003.

[2] Yu. Bilu, M. Strambi and A. Surroca, *Quantitative Chevalley–Weil theorem for curves*, arXiv:0908.1233.

[3] J.-B. Bost, H. Gillet and C. Soulé, *Heights of projective varieties and positive Green forms*, J. Amer. Math. Soc. 7 (1994), 903–1027.

[4] P. Dèbes, *Méthodes topologiques et analytiques en théorie inverse de Galois: théorème d'existence de Riemann*, in [5], 27–41.

[5] B. Deschamps (ed.), *Arithmétique des revètements algébriques* (Saint-Étienne, 2000), Sém. Congrès 5, SMF, Paris, 2001.

[6] O. B. Dolgopolova and È. I. Zverovich, *Explicit construction of global uniformization of an algebraic correspondence*, Sibirsk. Mat. Zh. 41 (2000), 72–87, ii (in Russian); English transl.: Siberian Math. J. 41 (2000), 61–73.

[7] B. Edixhoven, R. de Jong and J. Schepers, *Covers of surfaces with fixed branch locus*, arXiv:0807.0184.

[8] R. Hartshorne, *Algebraic Geometry*, Grad. Texts in Math. 52, Springer, 1977.

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

[10] P. Philippon, *Sur des hauteurs alternatives*, I, Math. Ann. 289 (1991), 255–283; II, Ann. Inst. Fourier (Grenoble) 44 (1994), 1043–1065; III, J. Math. Pures Appl. 74 (1995), 345–365.

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

[12] B. L. van der Waerden, *Algebra II*, 6th German ed., Springer, 1993.

[13] È. I. Zverovich, *An algebraic method for constructing the basic functionals of a Riemann surface given in the form of a finite covering of a sphere*, Sibirsk. Mat. Zh. 28 (1987), 32–43, 217 (in Russian); English transl.: Siberian Math. J. 28 (1987), 889–898.

Yuri F. Bilu
Institut de Mathématiques
Université Bordeaux 1
351 cours de la Libération
33405 Talence, France
E-mail: yuri@math.u-bordeaux1.fr

Marco Strambi
Dipartimento di Matematica
Università di Pisa
Lago Bruno Pontecorvo 5
56127 Pisa, Italy
E-mail: strambi@mail.dm.unipi.it

Received on 17.9.2008
and in revised form on 2.8.2009 (5800)