AN EFFECTIVE WEIERSTRASS DIVISION THEOREM

MATTHIAS ASCHENBRENNER

ABSTRACT. We prove an effective Weierstrass Division Theorem for algebraic restricted power series with p-adic coefficients. The complexity of such power series is measured using a certain height function on the algebraic closure of the field of rational functions over $\mathbb{Q}$. The paper includes a construction of this height function, following an idea of Kani. We apply the effective Weierstrass Division Theorem to obtain a number-theoretic criterion for membership in ideals of polynomial rings with integer coefficients.

CONTENTS

INTRODUCTION

Let $f_1, \dots, f_n \in \mathbb{Z}[X]$, where $X = (X_1, \dots, X_N)$ is an N-tuple of indeterminates. The starting point for this paper was the following criterion for membership in the ideal of $\mathbb{Z}[X]$ generated by $f_1, \dots, f_n$ (implicit in [37]): there exist positive integers δ and e such that for every polynomial $f_0 \in \mathbb{Z}[X]$:

$$f_0 \in (f_1, \dots, f_n)\mathbb{Z}[X] \quad \Longleftrightarrow \quad f_0 \in (f_1, \dots, f_n)\mathbb{Q}[X] \text{ and } \overline{f_0} \in (\overline{f_1}, \dots, \overline{f_n})(\mathbb{Z}/\delta^e\mathbb{Z})[X]. \quad (0.1)$$

Here and below, for $a \in \mathbb{Z}$ we denote by $\overline{f}$ the polynomial in $(\mathbb{Z}/a\mathbb{Z})[X]$ obtained by applying the canonical surjection $\mathbb{Z} \to \mathbb{Z}/a\mathbb{Z}$ to the coefficients of $f \in \mathbb{Z}[X]$. The existence of such δ and e can be easily seen as follows. Put $I := (f_1, \dots, f_n)\mathbb{Z}[X]$, and for $a \in \mathbb{Z}$ denote by $I : (a)$ the ideal of $\mathbb{Z}[X]$ consisting of all $f \in \mathbb{Z}[X]$ such that $af \in I$. The ring $\mathbb{Z}[X]$ being noetherian, the ideal $J := I\mathbb{Q}[X] \cap \mathbb{Z}[X]$ is finitely generated, so there exists a

Date: October 2005.
2000 *Mathematics Subject Classification.* Primary 13J05; Secondary 11G50, 13P10.
Key words and phrases. Restricted p-adic power series; ideal membership; height function.
Partially supported by NSF grant DMS 03-03618.

positive integer δ such that $\delta f \in I$ for all $f \in J$, or in other words, such that $J = I : (\delta)$. Consider now the ascending chain of ideals

$$I \subseteq I : (\delta) \subseteq \cdots \subseteq I : (\delta^i) \subseteq \cdots$$

in $\mathbb{Z}[X]$. Again, since $\mathbb{Z}[X]$ is noetherian, there exists an integer $e > 0$ such that

$$I : (\delta^e) = I : (\delta^{e+i}) \qquad \text{for all } i \in \mathbb{N}.$$

With this e it is easy to check that

$$I = \big(I : (\delta)\big) \cap \big(I + (\delta^e)\big),$$

hence δ and e have the required properties.

It is well-known that an integer δ with the property that $J = I : (\delta)$ can be effectively computed from $f_1, \dots, f_n$. In fact, we may take $\delta = P(c)$ for some polynomial $P(C) \in \mathbb{Z}[C]$ in the coefficient tuple c of $f_1, \dots, f_n$. This is a byproduct of Hermann's classical algorithm [18] for deciding ideal membership in polynomial rings over fields. (See [8], Section 3.) The smallest positive integer $\delta_I = \delta$ such that $I : (\delta) = J$ is the exponent of the torsion subgroup of the abelian group $\mathbb{Z}[X]/I$. (An abelian group G is said to have finite exponent if there exists a positive integer m with $mG = \{0\}$, and the smallest such m is called the exponent of G.) An algorithm to compute δ_I was given by Clivio [12] (based on earlier work of Ayoub [6]).

The computability of δ can be used to turn the criterion (0.1) into a (hopelessly inefficient) procedure, due to Simmons [37], for deciding whether a given polynomial $f_0 \in \mathbb{Z}[X]$ lies in the ideal I: First we check whether $f_0 \in I\mathbb{Q}[X]$, say using Hermann's method; if the answer is negative, we already know that $f_0 \notin I$. Otherwise, we begin running two effective procedures simultaneously. In the first one, we enumerate all n-tuples of polynomials in $\mathbb{Z}[X]$, and for each such n-tuple $(g_1, \dots, g_n)$ we compute $f_1 g_1 + \cdots + f_n g_n$. We stop when we find that $f_0 = f_1 g_1 + \cdots + f_n g_n$. In the second procedure, we successively check, for each $e = 1, 2, \dots$, whether $\overline{f_0} \in (\overline{f_1}, \dots, \overline{f_n})(\mathbb{Z}/\delta^e\mathbb{Z})[X]$. As shown in [37] (see also Section 6 below), the ideal membership problem for $(\mathbb{Z}/\delta^e\mathbb{Z})[X]$ can be easily, and elementary recursively in the data, reduced to solving systems of linear equations with coefficients in $\mathbb{F}_p[X]$ (with p ranging over the prime divisors of δ), where Hermann's method applies. The first procedure stops if $f_0 \in I$, and by (0.1), the second algorithm terminates if $f_0 \notin I$.

All this raises the obvious question:

> *Can one compute an exponent e such that* (0.1) *holds for every polynomial $f_0 \in \mathbb{Z}[X]$ and every positive integer δ with the property that $I : (\delta) = J$?*

One of the aims of this paper is to answer this question positively. In fact, we will exhibit an explicit number-theoretic function $(N, \beta) \mapsto e(N, \beta)$ such that if $f_1, \dots, f_n$ are polynomials in $\mathbb{Z}[X] = \mathbb{Z}[X_1, \dots, X_N]$ with $\deg f_1, \dots, \deg f_n \leq \beta$ and $||f_1||_\infty, \dots, ||f_n||_\infty \leq \beta$, then the equivalence (0.1) holds for every $f_0 \in \mathbb{Z}[X]$ and every positive $\delta \in \mathbb{Z}$ such that $I : (\delta) = J$, if one takes $e = e(N, \beta)$. Here, $\deg f$ denotes the (total) degree of $f \in \mathbb{Z}[X]$, and $||f||_\infty$ denotes the maximum of the absolute values of the coefficients of f.

Our exponent $e(N, \beta)$ is most likely highly excessive: it takes the form of an N-times iterated exponential in N and β. However, it still seems worth discussing, for two reasons: First, because of the inherent *non-primitive recursive* features of Hilbert's Basis Theorem exemplified in [16], [28]. (This theorem, in the form of noetherianity of $\mathbb{Z}[X]$, was used above to establish the existence of e.) The class of *primitive recursive* functions forms a proper subclass of all algorithmically computable (also called recursive) functions. The function $(N, \beta) \mapsto e(N, \beta)$ is primitive recursive (and in fact, *for fixed* N, belongs to the

smaller class of functions which are elementary recursive in the sense of Kalmár [19]). Second, because some of the results obtained *en route* might be of independent interest and useful for further investigations of the properties of finitely generated commutative rings (such as $\mathbb{Z}[X]/I$). With this in mind, we now explain our method to construct the exponent e.

Let $p_1, \dots, p_K$ be the distinct prime factors of δ. A short argument using the Euclidean Algorithm (see [8], pp. 409–410) shows that then for $f_0 \in \mathbb{Z}[X]$:

$$f_0 \in (f_1, \dots, f_n)\mathbb{Z}[X] \quad \Longleftrightarrow \\ f_0 \in (f_1, \dots, f_n)\mathbb{Q}[X] \text{ and } f_0 \in (f_1, \dots, f_n)\mathbb{Z}_{(p_k)}[X] \text{ for } k = 1, \dots, K. \quad (0.2)$$

Here $\mathbb{Z}_{(p)}$ denotes the localization of $\mathbb{Z}$ at the prime ideal $(p) = p\mathbb{Z}$, where p is a prime number. We write $\mathbb{Z}_p\langle X\rangle$ for the ring of *restricted power series* with p-adic integer coefficients, that is, the completion of $\mathbb{Z}[X]$ at the ideal $p\mathbb{Z}[X]$ generated by p. (See [10] for basic facts about $\mathbb{Z}_p\langle X\rangle$.) We also let $\mathbb{Z}_p\langle X\rangle_{\mathrm{alg}}$ be the subring of $\mathbb{Z}_p\langle X\rangle$ consisting of the restricted power series that are *algebraic* over the rational function field $\mathbb{Q}(X)$. (This is the henselization of $\mathbb{Z}[X]$ with respect to the ideal $p\mathbb{Z}[X]$, see [35], p. 126.) A faithful flatness argument ([8], Lemma 2.6) yields

$$f_0 \in (f_1, \dots, f_n)\mathbb{Z}_{(p)}[X] \quad \Longleftrightarrow \\ f_0 \in (f_1, \dots, f_n)\mathbb{Q}[X] \text{ and } f_0 \in (f_1, \dots, f_n)\mathbb{Z}_p\langle X\rangle. \quad (0.3)$$

Hence in (0.2) we may replace each ring $\mathbb{Z}_{(p_k)}[X]$ by $\mathbb{Z}_{(p_k)}\langle X\rangle$. Moreover, since $\mathbb{Z}_p\langle X\rangle$ is faithfully flat over $\mathbb{Z}_p\langle X\rangle_{\mathrm{alg}}$, in (0.3) we may further replace $\mathbb{Z}_p\langle X\rangle$ by its subring $\mathbb{Z}_p\langle X\rangle_{\mathrm{alg}}$. Thus we can improve (0.2) as follows:

$$f_0 \in (f_1, \dots, f_n)\mathbb{Z}[X] \quad \Longleftrightarrow \\ f_0 \in (f_1, \dots, f_n)\mathbb{Q}[X] \text{ and } f_0 \in (f_1, \dots, f_n)\mathbb{Z}_{(p_k)}\langle X\rangle_{\mathrm{alg}} \text{ for all } k.$$

What we have gained is that the rings $\mathbb{Z}_p\langle X\rangle$ have some very nice properties: besides being noetherian and henselian, they also satisfy Weierstraß Division and Preparation Theorems. These properties continue to hold for $\mathbb{Z}_p\langle X\rangle_{\mathrm{alg}}$; for example, the ring $\mathbb{Z}_p\langle X\rangle_{\mathrm{alg}}$ is closed under Weierstraß Division in $\mathbb{Z}_p\langle X\rangle$. Moreover, in $\mathbb{Z}_p\langle X\rangle_{\mathrm{alg}}$ the complexity of quotient and remainder obtained through Weierstraß Division can be explicitly bounded.

In order to formulate this fact precisely, we employ a certain height function

$$h\colon \mathbb{Q}(X)_{\mathrm{alg}} \to \mathbb{R}^{\geq 0}$$

on the algebraic closure $\mathbb{Q}(X)_{\mathrm{alg}}$ of $\mathbb{Q}(X)$, a variant of which was constructed by Kani in his Ph.D. thesis [20]. We have not found a suitable analogue in the literature (for example, in [32], Section B, one only finds something like local height functions on $\mathbb{Q}(X)_{\mathrm{alg}}$); since it is fundamental for our work, this paper includes an account of Kani's (unpublished) construction. Based on ideas of Arakelov [4], it is nevertheless quite elementary, in contrast to the constructions of height functions in (higher-dimensional) Arakelov theory, which usually heavily rely on algebraic geometry (Chow forms, intersection theory etc.).

The height function h extends the usual (absolute logarithmic) height on the algebraic closure of $\mathbb{Q}$ (as defined in [23]). For non-zero $f \in \mathbb{Z}[X]$ we have

$$h(f) = \deg_{X_1} f + \dots + \deg_{X_N} f + \\ \int_0^1 \cdots \int_0^1 \max\big\{\log|f(e^{2\pi i\theta_1}, \dots, e^{2\pi i\theta_N})|, 0\big\} d\theta_1 \cdots d\theta_N,$$

where $\deg_{X_i}$ is the degree in the variable X_i. The restriction of h to $\mathbb{Z}[X]$ is related to Mahler's measure (see Section 1.2 below), which has been used before in the context of ideal membership problems in $\mathbb{Z}[X]$, see, e.g., [30], [31]. The function h behaves well with respect to algebraic operations in $\mathbb{Q}(X)_{\mathrm{alg}}$, and also has the following *finiteness property:* there are only finitely many $f \in \mathbb{Q}(X)_{\mathrm{alg}}$ with given bounds on the height $h(f)$ and the degree $\Delta(f) := \big[\mathbb{Q}(X, f) : \mathbb{Q}(X)\big]$ of f. Here now is the "effective" Weierstraß Division Theorem alluded to in the title of this paper, which will be established in Section 4. Recall that $g \in \mathbb{Z}_p\langle X\rangle$ is said to be regular in X_N of degree $s \in \mathbb{N}$ if there exist a monic polynomial $g_0 \in \mathbb{Z}_p\langle X'\rangle[X_N]$ of degree s and a unit u of $\mathbb{Z}_p$ such that $g - ug_0 \in p\mathbb{Z}_p\langle X\rangle$. Here and below $X' := (X_1, \ldots, X_{N-1})$.

Theorem 0.1. *Let $f, g \in \mathbb{Z}_p\langle X\rangle_{\mathrm{alg}}$, and suppose that g is regular of degree $s > 0$ in X_N. Let (q, r) be the unique pair consisting of $q \in \mathbb{Z}_p\langle X\rangle_{\mathrm{alg}}$ and $r \in \mathbb{Z}_p\langle X'\rangle_{\mathrm{alg}}[X_N]$ such that $f = qg + r$ and $\deg_{X_N} r < s$. Then, writing*

$$r = r_0 + r_1 X_N + \cdots + r_{s-1} X_N^{s-1} \qquad (r_0, \ldots, r_{s-1} \in \mathbb{Z}_p\langle X'\rangle_{\mathrm{alg}}),$$

we have

$$\begin{aligned} \Delta(r_i) &\leq \big(\Delta(f)\Delta(g)\big(h(g) + \log 2\big)\big)^s, \\ h(r_i) &\leq O(1)^s\big(\Delta(f)\big)^s \Delta(g)\big(h(f) + \log 2\big)\big(h(g) + \log 2\big). \end{aligned}$$

Similar results are known for the ring $K[[X]]$ of formal power series with coefficients in a field K. See [22], [25] for a Weierstraß Division Theorem for the subring of $K[[X]]$ consisting of those power series which are algebraic over $K(X)$. In [33], [34] a complexity measure for Nash functions is defined and used to prove an effective version of Bézout's Theorem for these functions, and, in [13], an effectivization of the Artin-Mazur Theorem. In [1], [2], related measures for the complexity of algebraic formal power series are introduced, and bounds for the complexity of quotient and remainder in the Weierstraß Division Theorem in terms of the complexity of dividend and divisor are deduced. The bounds obtained there (by very different methods) are also of a single exponential nature, like ours.

The effective Weierstraß Division Theorem above can be used to adapt Hermann's method to treat ideal membership problems for ideals in $\mathbb{Z}_p\langle X\rangle_{\mathrm{alg}}$. Here, the role played by the Euclidean Algorithm (for polynomials) in Hermann's method is taken over by the Weierstraß Division Theorem in $\mathbb{Z}_p\langle X\rangle_{\mathrm{alg}}$. Using the height function on $\mathbb{Q}(X)_{\mathrm{alg}}$, bounds for the complexity of the power series occurring in the individual steps can be found. This allows us to show the following "effective" p-adic analogue of (0.1). Here $\overline{f}$ denotes the image of $f \in \mathbb{Z}_p\langle X\rangle$ under the canonical homomorphism $\mathbb{Z}_p\langle X\rangle \to \mathbb{Z}_p\langle X\rangle/p^E\mathbb{Z}_p\langle X\rangle \cong (\mathbb{Z}/p^E\mathbb{Z})[X]$. As before, let $\beta \in \mathbb{N}$ be such that $||f_i||_\infty \leq \beta$ and $\deg f_i \leq \beta$ for all i, and suppose $\beta > 0$.

Theorem 0.2. *There exists a positive integer E with*

$$E \leq 2^{2^{\cdot^{\cdot^{\cdot^{2^{O(1)^N(2N\beta + \log\beta + 1)^{N+1}}}}}}} \qquad (N \text{ many } 2\text{'s}),$$

such that for all $f_0 \in \mathbb{Z}_p\langle X\rangle$:

$$f_0 \in (f_1, \ldots, f_n)\mathbb{Z}_p\langle X\rangle \iff f_0 \in (f_1, \ldots, f_n)\mathbb{Q}_p\langle X\rangle \text{ and } \overline{f_0} \in (\overline{f_1}, \ldots, \overline{f_n})(\mathbb{Z}/p^E\mathbb{Z})[X].$$

Together with the discussion above, this theorem leads to the computation of the exponent $e(N, \beta)$. The upper bound on the exponent E in the last theorem is probably far from optimal. It would be interesting to obtain a qualitatively better (say, doubly-exponential) bound on E, and hence improve $e(N, \beta)$. (The bottleneck here is Theorem 0.1.) Note however that $e(N, \beta)$ has the advantage of being *independent* of f_0. If we are willing to also bound the complexity of f_0 and the degrees of the solutions of the reduced equation $\overline{f_0} = \overline{f_1}y_1 + \cdots + \overline{f_n}y_n$, then using the results of [8] we obtain the following membership criterion:

Theorem 0.3. *There exists a positive integer $D = (2\beta)^{2^{O(N \log(N+1))}}$ with the following property: let $f_0, \ldots, f_n \in \mathbb{Z}[X]$ with $\deg f_i, ||f_i||_\infty \leq \beta$ for $i = 0, \ldots, n$ and $f_0 \in (f_1, \ldots, f_n)\mathbb{Q}[X]$; then $f_0 \in (f_1, \ldots, f_n)\mathbb{Z}[X]$ if and only if and there exist $g_1, \ldots, g_n \in \mathbb{Z}[X]$ of degree at most D such that*

$$f_0 = f_1 g_1 + \cdots + f_n g_n \qquad \mod \delta^D.$$

For ideal membership in polynomial rings over residue class rings of $\mathbb{Z}$ one shows rather easily:

Proposition 0.4. *Let $f_0, \ldots, f_n \in \mathbb{Z}[X]$ such that $\deg f_i \leq d$ for $i = 0, \ldots, n$, and let $a \in \mathbb{N}$, $a > 0$. If $\overline{f_0} \in (\overline{f_1}, \ldots, \overline{f_n})(\mathbb{Z}/a\mathbb{Z})[X]$, then there exist $g_1, \ldots, g_n \in \mathbb{Z}[X]$ of degree at most $(\log_2 a)(2d)^{a^{N+1}}$ such that*

$$f_0 = f_1 g_1 + \cdots + f_n g_n \qquad \mod a.$$

Unfortunately, however, the bound in this proposition does not enable us to straight away remove the bounds on the degrees of the g_i in Theorem 0.3.

Organization of the paper. Section 1 has preliminary character and collects some definitions and basic results concerning absolute values and norms of polynomials. In Section 2 we introduce height functions, in the general context of Kani's theory of arithmetic fields [20], and in Section 3 we construct a height function on the algebraic closure of $\mathbb{Q}(X)$. We also go beyond [20] and discuss height on projective and affine space, and height of matrices. In Section 4 we prove Theorem 0.1. In Section 5 we adapt Hermann's method to $\mathbb{Z}_p\langle X\rangle$, leading to a proof of Theorem 0.2. Finally, Section 6 contains the proofs of Theorem 0.3 and Proposition 0.4.

Conventions and notations. Throughout this paper, N, m and n range over the set $\mathbb{N} = \{0, 1, 2, \ldots\}$ of natural numbers.

Acknowledgments. This paper derives from a part of my Ph.D. thesis [7], written under the guidance of Lou van den Dries, whose advice I gratefully acknowledge. In particular, the idea of using Kani's height function to compute the exponent e is due to him. I also thank Ernst Kani for the permission to include the construction of his height function.

1. Absolute values and norms of polynomials

For the convenience of the reader, and to fix notations, we first collect a few definitions and basic results concerning absolute values. (See [11], Chapitre VI for more facts about absolute values.) We then discuss several measures for the complexity of polynomials, most notably the (logarithmic) Mahler measure [26] from transcendental number theory.

1.1. **Absolute values.** Let K be a field and $|\cdot|$ be an **absolute value** on K, that is, a function $x \mapsto |x|\colon K \to \mathbb{R}^{\geq 0}$ with $|0| = 0$, whose restriction to $K^\times = K \setminus \{0\}$ is a homomorphism $K^\times \to \mathbb{R}^{>0}$ of (multiplicative) groups which satisfies the **triangle inequality**

$$|x + y| \leq |x| + |y| \qquad \text{for all } x, y \in K.$$

We always assume that absolute values are non-trivial, i.e., $|x| \neq 1$ for some $x \in K^\times$. If instead of the triangle inequality the stronger **ultrametric triangle inequality**

$$|x + y| \leq \max\big\{|x|, |y|\big\} \qquad \text{for all } x, y \in K$$

holds, then $|\cdot|$ is called **ultrametric**, and otherwise, **archimedean.** The absolute value $|\cdot|$ is ultrametric if and only if $|n \cdot 1| \leq 1$ for all n. (In particular, if $\mathrm{char}(K) > 0$, then each absolute value on K is ultrametric.) The map

$$(x, y) \mapsto |x - y|\colon K \times K \to \mathbb{R}^{\geq 0} \tag{1.1}$$

is a metric on K and makes K into a topological field. Two absolute values $|\cdot|_1$ and $|\cdot|_2$ on K induce the same topology on K if and only if $|\cdot|_1 = (|\cdot|_2)^r$ for some $r \in \mathbb{R}^{>0}$. In this case, $|\cdot|_1$ and $|\cdot|_2$ are called **equivalent.** For example, the only absolute values on $\mathbb{Q}$, up to equivalence, are

(1) the usual (archimedean) absolute value, given by $|x|_\infty := \max\{x, -x\}$ for $x \in \mathbb{Q}$, and
(2) for each prime number p, the (ultrametric) p-adic absolute value $|x|_p := p^{-v_p(x)}$, where $k = v_p(x) \in \mathbb{Z}$ is such that $x = p^k \frac{a}{b}$ with $a, b \in \mathbb{Z} \setminus \{0\}$ not divisible by p, and $|0|_p := 0$.

We associate to $|\cdot|$ its corresponding **additive absolute value**

$$v = v_{|\cdot|}\colon K^\times \to \mathbb{R}, \quad v(x) := -\log|x|,$$

a homomorphism of the multiplicative group of K into the additive group of $\mathbb{R}$. (Here and below, "log" will always stand for the natural logarithm.) We put $\Gamma_v := v(K^\times)$, an additive subgroup of $\mathbb{R}$. From v we can reconstruct $|\cdot| = |\cdot|_v$, since

$$|x|_v = \exp\big(-v(x)\big) \text{ for } x \in K^\times, \qquad |0|_v = 0. \tag{1.2}$$

This justifies the usual practice (which we will also adopt) of speaking of an "absolute value v", if one in fact means that "v is the additive absolute value corresponding to the absolute value $|\cdot|_v$." Note that if $|\cdot|$ is ultrametric, then v is a (real) valuation on K with

$$\mathcal{O}_v := \big\{x \in K : |x| \leq 1\big\} \quad \text{and} \quad \mathfrak{m}_v := \big\{x \in K : |x| < 1\big\} \tag{1.3}$$

as the corresponding valuation ring and its maximal ideal, respectively. Conversely, if $v\colon K^\times \to \mathbb{R}$ is a valuation on K, then $|\cdot|_v$ as given by (1.2) is an ultrametric absolute value on K. This gives a bijection $v \mapsto |\cdot|_v$ (with inverse $|\cdot| \mapsto v_{|\cdot|}$) between the set of all real valuations on K and the set of all ultrametric absolute values on K. An important class of absolute values arises from unique factorization:

Example 1.1. Suppose that $K = \mathrm{Frac}(R)$ is the fraction field of a unique factorization domain R. Every irreducible element $p \in R$ determines the (discrete) p-adic valuation $v_p\colon K^\times \to \mathbb{Z}$ by the rule

$$v_p(p^n a) = n \quad \text{for } 0 \neq a \in R, \, p \text{ does not divide } a.$$

Clearly, for irreducible $p, q \in R$ we have $v_p = v_q$ if and only if $p = uq$ for some unit u of R. The p-adic valuations v_p are called the **essential valuations of** R**.** Each essential valuation v_p determines an **essential absolute value** $|\cdot|_{v_p}$ on K by $|x|_{v_p} := \exp\big(-v_p(x)\big)$ for $x \in K^\times$. For instance, the essential valuations of $R = \mathbb{Z}$ are the p-adic valuations on

$\mathbb{Q}$, p a prime number. Note that $|\cdot|_p = |\cdot|_{v_p}^{\log p}$. We also write v_∞ for the absolute value $v_{|\cdot|_\infty}$ on $\mathbb{Q}$.

Addition and multiplication of K extend continuously to the completion $\widehat{K}$ of K with respect to the metric (1.1), making $\widehat{K}$ a field. The absolute value $|\cdot|$ on K extends continuously to an absolute value on $\widehat{K}$, also denoted by $|\cdot|$. Ostrowski's Theorem says that if $|\cdot|$ is archimedean, then either $(\widehat{K}, |\cdot|) \cong (\mathbb{R}, |\cdot|_{\mathbb{R}}^r)$ for some real number $r \in (0,1]$, or $(\widehat{K}, |\cdot|) \cong (\mathbb{C}, |\cdot|_{\mathbb{C}}^r)$ for some real number $r \in (0,1]$. Here $|\cdot|_{\mathbb{R}}$ and $|\cdot|_{\mathbb{C}}$ are the usual absolute values on $\mathbb{R}$ and $\mathbb{C}$, respectively.

For convenience, we put

$$\varepsilon_v := \begin{cases} |1|_v & \text{if } v \text{ is ultrametric,} \\ |2|_v & \text{if } v \text{ is archimedean.} \end{cases}$$

Then $1 \le \varepsilon_v \le 2$, with $\varepsilon_v = 1$ if and only if v is ultrametric, and for all $x, y \in K$ we have $|x+y|_v \le \varepsilon_v \max\{|x|_v, |y|_v\}$. (All this is clear if v is ultrametric; if v is archimedean, use Ostrowski's Theorem and the triangle inequality for $|\cdot|_{\mathbb{C}}$.)

The next lemma contains some inequalities that will become useful for later estimates.

Lemma 1.2. *Let $A = (a_{ij}) \in K^{n\times n}$ be an $n \times n$-square matrix over K, where $n > 0$. Suppose $r_1, \dots, r_n \in \mathbb{R}$ are such that $|a_{ij}|_v \le r_j$ for all $i, j = 1, \dots, n$. Then:*

(1) *if v is ultrametric, then $|\det A|_v \le r_1 \cdots r_n$;*
(2) *if v is archimedean, then $|\det A|_v \le |n|_v^{n/2} r_1 \cdots r_n$.*

Proof. Part (1) is clear from the ultrametric triangle inequality. For (2), by Ostrowski's Theorem, we may assume that $K = \mathbb{C}$ and $|\cdot|_v = |\cdot|^r$ for some $r \in (0,1]$, where $|\cdot|$ denotes the usual absolute value on $\mathbb{C}$. Let A have rows $a_1, \dots, a_n$. By Hadamard's inequality (see [17], §2.12),

$$|\det A|_v = |\det A|^r \le ||a_1||_2^r \cdots ||a_n||_2^r,$$

where $||b||_2 = \left(\sum_j |b_j|^2\right)^{1/2}$ for $b = (b_1, \dots, b_n) \in \mathbb{C}^n$. We have

$$||b||_2 \le \sqrt{n} \max\{|b_1|, \dots, |b_n|\} \qquad \text{for all } b \in \mathbb{C}^n.$$

This implies (2). □

Finally, let us recall that the absolute value $|\cdot|$ on K always extends to an absolute value on the algebraic closure K_{alg} of K, and any two such extensions $|\cdot|_1$ and $|\cdot|_2$ are conjugate, that is: there exists $\sigma \in \mathrm{Gal}(K_{\mathrm{alg}}|K)$ such that $|x|_1 = |\sigma(x)|_2$ for all $x \in K_{\mathrm{alg}}$. This fact will turn out to be crucial for the construction of a height function on K_{alg} in the next section.

1.2. **Norms of polynomials.** We now discuss several different measures for the size of polynomials, the most important among them the so-called logarithmic Mahler measure. We first introduce some notations used throughout the paper. Let R be a commutative ring, $X = (X_1, \dots, X_N)$ an N-tuple of indeterminates over R, and

$$f(X) = \sum_\nu a_\nu X^\nu \tag{1.4}$$

a polynomial in $R[X]$. Here the multiindex $\nu = (\nu_1, \dots, \nu_N)$ ranges over $\mathbb{N}^N$, $X^\nu = X_1^{\nu_1} \cdots X_N^{\nu_N}$, and $a_\nu \in R$. For $i = 1, \dots, N$, the degree of f in X_i will be denoted by $\deg_{X_i} f$, and the (total) degree of f by $\deg f$ or $\deg_X f$, with $\deg_{X_i} 0 = \deg 0 = -\infty <$

$\mathbb{N}$. We also set $\deg_{(X)} f := \sum_{i=1}^{N} \deg_{X_i} f$ for non-zero f, and $\deg_{(X)} 0 := -\infty$. Note that with $m = \max\{\deg_{X_1} f, \dots, \deg_{X_N} f\}$ we have

$$m \leq \deg_X f \leq \deg_{(X)} f \leq Nm. \tag{1.5}$$

Below, we let $R = K$ be a field and v be an absolute value on K, and we let f as in (1.4) range over $K[X]$.

First recall that the absolute value $|\cdot|_v$ extends to a norm on the polynomial ring $K[X]$, again denoted by $|\cdot|_v$, by $|f|_v := \max_\nu |a_\nu|_v$. The norm $|\cdot|_v$ is called the **Gauß norm** on $K[X]$. We also write $v(f) := -\log |f|_v$ for non-zero f. If the absolute value $|\cdot|_v$ of K is ultrametric, then by *Gauß's Lemma,* the norm $|\cdot|_v$ on $K[X]$ is multiplicative:

$$|f \cdot g|_v = |f|_v \cdot |g|_v \quad \text{for all } f, g \in K[X]. \tag{1.6}$$

So $|\cdot|_v$ extends uniquely to an absolute value on the fraction field $K(X)$ of $K[X]$, and v extends to a valuation on $K(X)$. If $|\cdot|_v$ on K is archimedean, then the Gauß norm on $K[X]$ is no longer multiplicative. However, for $f_1, \dots, f_n \in K[X]$, we do have the following fundamental inequality (known as *Gelfond's Lemma*):

$$|2^{-d} f_1 \cdots f_n|_v \leq |f_1|_v \cdots |f_n|_v \leq |2^d f_1 \cdots f_n|_v, \quad d = \deg_{(X)} f_1 \cdots f_n \tag{1.7}$$

(See [23], Chapter 3, §2.) Note that (1.6) and (1.7) may be combined to the inequality

$$\varepsilon_v^{-d} \prod_{i=1}^{n} |f_i|_v \leq |f_1 \cdots f_n|_v \leq \varepsilon_v^{d} \prod_{i=1}^{n} |f_i|_v. \tag{1.8}$$

The proof of the next lemma is straightforward from the definitions:

Lemma 1.3. *Suppose that f is non-zero, let $d = \deg_X f$ and $\operatorname{mon} f =$ the number of monomials occurring in f with a non-zero coefficient, so $\operatorname{mon} f \leq \binom{N+d}{d}$. Then, for all $x \in K^N$ with $\max_i |x_i|_v \geq 1$:*

(1) $|f(x)|_v \leq \max_i |x_i|_v^d \cdot |f|_v$ *if v is ultrametric,*
(2) $|f(x)|_v \leq \max_i |x_i|_v^d \cdot |f|_v \cdot |\operatorname{mon} f|_v$ *if v is archimedean.*

For archimedean $|\cdot|_v$ there are other measures for the size of a polynomial in $K[X]$, which fit different purposes. By Ostrowski's Theorem, we may restrict our attention to the field $K = \mathbb{C}$ of complex numbers with its usual absolute value $|z| = \sqrt{z\overline{z}}$, for $z \in \mathbb{C}$. The maps

$$f \mapsto ||f||_1 := \sum_\nu |a_\nu|, \qquad f \mapsto ||f||_2 := \left(\sum_\nu |a_\nu|^2 \right)^{1/2}$$

are norms on the $\mathbb{C}$-vector space $\mathbb{C}[X]$ extending the absolute value on $\mathbb{C}$, called the **l_1-norm** (or **length**) and the **l_2-norm** (or **euclidean norm**) on $\mathbb{C}[X]$, respectively. We also denote the Gauß norm (sometimes called the **l_∞-norm**) on $\mathbb{C}[X]$ by $|\cdot| = ||\cdot||_\infty$. These norms are connected by the following relations, valid for $f \neq 0$:

$$\begin{cases} ||f||_\infty \leq ||f||_1 \leq \big(1 + \deg_{X_1} f\big) \cdots \big(1 + \deg_{X_N} f\big) ||f||_\infty, \\ ||f||_\infty \leq ||f||_2 \leq \big(1 + \deg_{X_1} f\big)^{1/2} \cdots \big(1 + \deg_{X_N} f\big)^{1/2} ||f||_\infty, \\ ||f||_2 \leq ||f||_1 \leq \big(1 + \deg_{X_1} f\big)^{1/2} \cdots \big(1 + \deg_{X_N} f\big)^{1/2} ||f||_2. \end{cases} \tag{1.9}$$

A more subtle measure for the complexity of a non-zero polynomial f is its **Mahler measure**

$$m(f) = \int_0^1 \cdots \int_0^1 \log |f(e^{2\pi i \theta_1}, \dots, e^{2\pi i \theta_N})| \, d\theta_1 \cdots d\theta_N, \tag{1.10}$$

where the integral in (1.10) is taken in the sense of Lebesgue. See [15], §3.3 for a proof that $m(f) \in \mathbb{R}$. Note that $m(\,\cdot\,)$ is additive, i.e., $m(fg) = m(f) + m(g)$ for non-zero $f, g \in \mathbb{C}[X]$. In the next lemma, we collect a few other well-known properties of Mahler measure. (For a proof see [15].) We write $\log^+ x = \max\{0, \log x\}$, for any positive real number x.

Lemma 1.4. *Let $f \neq 0$, $d_j = \deg_{X_j} f$ for $j = 1, \ldots, N$, and $d = \sum_j d_j = \deg_{(X)} f$.*

(1) *If $N = 1$ and $\alpha_1, \ldots, \alpha_d \in \mathbb{C}$ are the zeros of f, with multiplicities, then*

$$m(f) = \log|a_d| + \sum_{j=1}^{d} \log^+ |\alpha_j|.$$

(2) $m(f) \geq \log|a_\nu|$ *for some non-zero coefficient a_ν of f. (In particular, $m(f) \geq 0$ if all $a_\nu \in \mathbb{Z}$.)*
(3) $m(f) \leq \log||f||_2$. *(Landau's Inequality.)*
(4) $m(f) - \frac{1}{2}\sum_j \log(d_j + 1) \leq \log||f||_\infty \leq m(f) + d\log 2$.
(5) $m(f) \leq \log||f||_1 \leq m(f) + d\log 2$.
(6) *If $j \in \{1, \ldots, N\}$ and $d_j > 0$, then $m(\partial f/\partial X_j) \leq m(f) + \log d_j$.*

Let us also introduce the notation

$$m^+(f) = \int_0^1 \cdots \int_0^1 \log^+ |f(e^{2\pi i\theta_1}, \ldots, e^{2\pi i\theta_N})|\, d\theta_1 \cdots d\theta_N$$

for non-zero f. Note that by Jensen's Formula

$$\int_0^1 \cdots \int_0^1 \log^+ |f(e^{2\pi i\theta_1}, \ldots, e^{2\pi i\theta_N})|\, d\theta_1 \cdots d\theta_N = \int_0^1 \cdots \int_0^1 \log|f(e^{2\pi i\theta_1}, \ldots, e^{2\pi i\theta_N}) + e^{2\pi i\theta}|\, d\theta d\theta_1 \cdots d\theta_N,$$

so that $m^+(f) = m\big(f(X) + Y\big)$, where Y is a new indeterminate, different from each of $X_1, \ldots, X_N$. It is convenient to define $m^+(0) := 0$. For $N = 1$, the function m^+ as a complexity measure for polynomials has been introduced in [27], and was further studied in [3].

For non-zero polynomials $f, f_1, \ldots, f_n \in \mathbb{C}[X]$, $n > 0$ and $k \in \mathbb{N}$, we have:

$$m^+(f_1 \cdots f_n) \leq m^+(f_1) + \cdots + m^+(f_n), \tag{1.11}$$

$$m^+(f_1 + \cdots + f_n) \leq m^+(f_1) + \cdots + m^+(f_n) + \log n, \tag{1.12}$$

$$m^+(f^k) = k m^+(f). \tag{1.13}$$

The next lemma compares $m^+(\,\cdot\,)$ with $m(\,\cdot\,)$, $||\cdot||_1$ and $||\cdot||_\infty$:

Lemma 1.5. *Let $f(X) \in \mathbb{Z}[X]$, $f \neq 0$, $d = \deg_{(X)} f$. Then*

$$m(f) \leq m^+(f) \leq \log||f||_1 \leq m(f) + d\log 2 \tag{1.14}$$

and

$$\log||f||_\infty - d\log 2 \leq m^+(f) \leq \log||f||_\infty + d. \tag{1.15}$$

In the proof, we use the following estimate. Here, $Y = (Y_1, \ldots, Y_M)$ is a tuple of distinct indeterminates over $\mathbb{Z}$.

Lemma 1.6. *Suppose $f(X) \in \mathbb{Z}[X]$ and $g_1(Y), \dots, g_N(Y) \in \mathbb{Z}[Y]$. Then*

$$m^+\big(f(g_1, \dots, g_N)\big) \leq \log ||f||_1 + d_1 m^+(g_1) + \cdots + d_N m^+(g_N),$$

where $d_1 = \deg_{X_1} f, \dots, d_N = \deg_{X_N} f$.

Proof. Write f as in (1.4), with all $a_\nu \in \mathbb{Z}$. Then for all $z = (z_1, \dots, z_N) \in \mathbb{C}^N$

$$\big|f\big(g_1(z), \dots, g_N(z)\big)\big| \leq \sum_\nu |a_\nu| \max\big\{|g_1(z)|^{d_1}, 1\big\} \cdots \max\big\{|g_N(z)|^{d_N}, 1\big\},$$

for $1 \leq j \leq N$. Thus

$$\log^+ \big|f\big(g_1(z), \dots, g_N(z)\big)\big| \leq \log ||f||_1 + d_1 \log^+|g_1(z)| + \cdots + d_N \log^+|g_N(z)|.$$

Taking $z_j = e^{2\pi i \theta_j}$ with $0 \leq \theta_j \leq 1$ for $j = 1, \dots, N$ and integrating both sides over $[0,1]^N$ with respect to $(\theta_1, \dots, \theta_N)$ gives the desired inequality. □

The first inequality in (1.14) is clear since $\log x \leq \log^+ x$ for all $x \in \mathbb{R}^{>0}$, and the last one is part of Lemma 1.4, (5). For the second one, apply the lemma above, taking $g_j = X_j$ for $j = 1, \dots, N$. The inequalities in (1.15) now follow from (1.14), Lemma 1.4, (4), and (1.9).

2. Height functions

In this section we first give a brief treatment of the relevant facts from Kani's theory of arithmetic fields, divisors, degree functions, and height functions, in a less general setting than in [20]. Its purpose is to abstract the main features of the theory of height functions on algebraic number fields (as presented in [23], say) to "non-classical" arithmetic fields, such as the function field $\mathbb{Q}(X_1, \dots, X_N)$. (See Definition 2.2 below.) We give a self-contained proof of Kani's theorem about the extension of a degree function on an arithmetic field to its algebraic closure. Using this theorem, in the next section we obtain a height function on the algebraic closure of $\mathbb{Q}(X_1, \dots, X_N)$, which will be the main tool for the computations in the sections to follow. In the later part of this section we discuss height on projective and affine space, and establish a few basic estimates used later.

2.1. **Arithmetic fields and divisors.** An **arithmetic field** is a pair (K, M) consisting of a field K and a collection M of mutually non-equivalent absolute values on K. (This is less general than the notion of arithmetic field in [20], Kapitel I, Definition 1.6.)

Examples 2.1.

(1) The collection $M = \{v_p : p \text{ prime}\} \cup \{v_\infty\}$ makes $(\mathbb{Q}, M)$ an arithmetic field, and we refer to this also as **the arithmetic field $\mathbb{Q}$.**
(2) Let R be a unique factorization domain, $K = \mathrm{Frac}(R)$, and

$$M_R := \{v_p : p \text{ irreducible element of } R\}.$$

The pair (K, M_R) is an arithmetic field.
(3) Given an arithmetic field (K, M), let M_{alg} be the set of absolute values on K_{alg} extending an absolute value in M. Then $(K_{\mathrm{alg}}, M_{\mathrm{alg}})$ is an arithmetic field.

We call an arithmetic field (K_1, M_1) an extension of the arithmetic field (K, M), in symbols $(K, M) \subseteq (K_1, M_1)$, if K is a subfield of K_1 and M_1 consists of all extensions of the absolute values in M to absolute values on K_1. Given any intermediate field $K \subseteq L \subseteq K_{\mathrm{alg}}$, we can make L into an arithmetic field (L, M_L) with $(K, M) \subseteq (L, M_L)$ in a unique way by setting $M_L := \{v \restriction L : v \in M_{\mathrm{alg}}\}$. (So $(K, M) \subseteq (K_{\mathrm{alg}}, M_{\mathrm{alg}})$ and $M_{\mathrm{alg}} = M_{K_{\mathrm{alg}}}$.)

Definition 2.2. An arithmetic field (K, M) is called **classical** if for each $a \in K^\times$, the **support of** a,

$$\operatorname{supp}(a) := \big\{v \in M : v(a) \neq 0\big\},$$

is a finite set.

Examples 2.3. The arithmetic field $\mathbb{Q}$ is classical. If R is a unique factorization domain, $K = \operatorname{Frac}(R)$, then (K, M_R) is a classical arithmetic field. If (K, M) is classical, so is (L, M_L) for each finite extension $L|K$. (Hence in particular, every algebraic number field, as an arithmetic field extending the arithmetic field $\mathbb{Q}$, is classical.) In the next section we will encounter examples of non-classical arithmetic fields.

Let (K, M) be an arithmetic field. We let

$$\mathcal{D}^{\mathrm{f}}(K, M) := \prod_{v \in M} \Gamma_v,$$

a product of additive subgroups of $\mathbb{R}$. Its elements $D = (v(D))_{v \in M}$ are called **formal divisors** on (K, M). Since each Γ_f is a linearly ordered additive group, the (additively written) group $\mathcal{D}^{\mathrm{f}}(K, M)$ is a lattice-ordered group: for $D_1, D_2 \in \mathcal{D}^{\mathrm{f}}(K, M)$

$$D_1 \leq D_2 \quad \iff \quad v(D_1) \leq v(D_2) \text{ for all } v \in M,$$

and for all $v \in M$:

$$v(D_1 \wedge D_2) = \min\big\{v(D_1), v(D_2)\big\}, \qquad v(D_1 \vee D_2) = \max\big\{v(D_1), v(D_2)\big\}.$$

In particular, every formal divisor D is of the form

$$D = D_0 - D_\infty, \qquad \text{where } D_0 = D \vee 0 \text{ and } D_\infty = (-D) \vee 0.$$

We also put

$$|D| := D \vee (-D) = D_0 + D_\infty \qquad \text{for } D \in \mathcal{D}^{\mathrm{f}}(K, M).$$

Every element $a \in K^\times$ determines a formal divisor, written $\operatorname{div}(a) = \operatorname{div}_{(K,M)}(a)$ or just (a), by

$$v\big(\operatorname{div}(a)\big) = v(a) \qquad \text{for all } v \in M.$$

We call $(a)_0$ and $(a)_\infty$ the **divisor of zeros of** a and the **divisor of poles of** a**,** respectively.

Example 2.4. We have $(2)_\infty = (\log \varepsilon_v)_{v \in M}$, using the notation introduced in the previous section. Thus $(2)_\infty = 0$ if M contains only ultrametric absolute values.

The formal divisors of the form (a) for $a \in K^\times$ are called **principal divisors.** Note that $(ab) = (a) + (b)$ and $(a^{-1}) = -(a)$ for $a, b \in K^\times$, so we have a group homomorphism

$$a \mapsto \operatorname{div}(a) \colon K^\times \to \mathcal{D}^{\mathrm{f}}(K, M)$$

whose image we denote by $\mathcal{P}(K, M)$. We let $\mathcal{D}(K, M)$ be the smallest subgroup of $\mathcal{D}^{\mathrm{f}}(K, M)$ closed under the operation $\wedge$ and containing $\mathcal{P}(K, M)$. (Then $\mathcal{D}(K, M)$ is also closed under $\vee$.) We call $\mathcal{D}(K, M)$ the **divisor group of** (K, M) and its elements **divisors.** The quotient group

$$\operatorname{Cl}(K, M) = \mathcal{D}(K, M)/\mathcal{P}(K, M)$$

is called the **divisor class group of** (K, M), and its elements **divisor classes.**

Lemma 2.5. $\mathcal{D}(K, M) = \Big\{\bigwedge_{1 \leq i \leq m} \bigvee_{1 \leq j \leq n} (a_{ij}) : m, n \geq 1, a_{ij} \in K^\times\Big\}$.

Proof. Let $\mathcal{S}$ be the set on the right hand side. Then $\mathcal{S}$ is closed under $\wedge$ and contains $\mathcal{P}(K,M)$; in order to show $\mathcal{S}=\mathcal{D}(K,M)$, it therefore suffices to prove that $\mathcal{S}$ is a subgroup of $\mathcal{D}^{\mathrm{f}}(K,M)$. This follows immediately from the following two easily verified facts: if

$$D=\bigwedge_{1\le i\le m}\bigvee_{1\le j\le n}(a_{ij}) \text{ and } D'=\bigwedge_{1\le k\le m'}\bigvee_{1\le l\le n'}(a'_{kl})$$

with $m,n,m',n'\ge 1$ and $a_{ij},a'_{kl}\in K^\times$, then

$$D+D'=\bigwedge_{\substack{1\le i\le m,\\ 1\le k\le m'}}\bigvee_{\substack{1\le j\le n,\\ 1\le l\le n'}}(a_{ij}a'_{kl}) \tag{2.1}$$

and

$$-D=\bigwedge_{\varphi}\bigvee_{1\le i\le m}(1/a_{i,\varphi(i)}),$$

where φ runs through all maps $\{1,\dots,m\}\to\{1,\dots,n\}$. □

In the classical case, $\mathcal{D}(K,M)$ admits an even simpler representation:

Lemma 2.6. *If (K,M) is a classical arithmetic field, then*

$$\mathcal{D}(K,M)=\bigoplus_{v\in M}\Gamma_v,$$

where $\bigoplus_v\Gamma_v$ is viewed as a subgroup of $\mathcal{D}^{\mathrm{f}}(K,M)=\prod_v\Gamma_v$.

Proof. Since $\operatorname{supp} D$ is finite for each $D\in\mathcal{D}(K,M)$, we have $\mathcal{D}(K,M)\subseteq\bigoplus_v\Gamma_v$. For the reverse inclusion, it suffices to show: for every $v\in M$ and $\gamma\in\Gamma_v$ there exists a divisor $D\in\mathcal{D}(K,M)$ with $v(D)=\gamma$ and $w(D)=0$ for $w\ne v$ in M. For this, we may assume $\gamma=v(x)>0$ with $x\in K^\times$. By the Approximation Theorem for absolute values (see [11], Chapitre VI, §7, Théorème 2), we find $y\in K^\times$ with $v(y)>0$ and $w(y)<0$ for all $w\ne v$ in M with $w(x)>0$. Set $D'=(x)_0\wedge(y)_0$. If $v(y)\ge v(x)$, then $D:=D'$ is the desired divisor; otherwise, replace y by a suitable power y^n, $n>0$. □

2.2. **Extensions of arithmetic fields.** Suppose (K,M), (K_1,M_1) are arithmetic fields with $(K,M)\subseteq(K_1,M_1)$. The restriction map

$$v\mapsto\varrho(v):=v\restriction K\colon M_1\to M$$

induces an embedding of ordered groups

$$\varrho^*\colon\mathcal{D}^{\mathrm{f}}(K,M)\to\mathcal{D}^{\mathrm{f}}(K_1,M_1)$$

between the formal divisor groups via the rule

$$v\big(\varrho^*(D)\big)=\big(\varrho(v)\big)(D)\qquad\text{for all } D\in\mathcal{D}^{\mathrm{f}}(K,M),\ v\in M_1.$$

We have

$$\varrho^*\big(\operatorname{div}_{(K,M)}(a)\big)=\operatorname{div}_{(K_1,M_1)}(a)\qquad\text{for } a\in K^\times,$$

hence $\varrho^*\big(\mathcal{P}(K,M)\big)\subseteq\mathcal{P}(K_1,M_1)$; moreover, for all $D_1,D_2\in\mathcal{D}^{\mathrm{f}}(K,M)$ we have

$$\varrho^*(D_1\vee D_2)=\varrho^*(D_1)\vee\varrho^*(D_2),\quad \varrho^*(D_1\wedge D_2)=\varrho^*(D_1)\wedge\varrho^*(D_2),$$

therefore $\varrho^*\big(\mathcal{D}(K,M)\big)\subseteq\mathcal{D}(K_1,M_1)$. Below, we will always identify the ordered groups $\mathcal{D}^{\mathrm{f}}(K,M)$, $\mathcal{P}(K,M)$ and $\mathcal{D}(K,M)$ with their respective images under the embedding ϱ^*.

Let now (K,M) be an arithmetic field and $L|K$ be a normal field extension. The group $\mathrm{Gal}(L|K)$ acts on M_L in a natural way from the right: for $v\in M_L$ and $\sigma\in\mathrm{Gal}(K_L|K)$

put $v^\sigma := v \circ \sigma \in M_L$. This action induces a left action $(\sigma, D) \mapsto \sigma D$ of $\mathrm{Gal}(L|K)$ on $\mathcal{D}^{\mathrm{f}}(L, M_L)$:

$$(\sigma D)(v) := v^\sigma(D) \quad \text{for } D \in \mathcal{D}^{\mathrm{f}}(L, M_L), \sigma \in \mathrm{Gal}(L|K), v \in M_L.$$

The map $D \mapsto \sigma D$ is an automorphism of the ordered group $\mathcal{D}^{\mathrm{f}}(L, M_L)$, with inverse $D \mapsto \sigma^{-1}D$. Every divisor $D \in \mathcal{D}^{\mathrm{f}}(K, M)$ is invariant under $\mathrm{Gal}(L|K)$: $\sigma D = D$ for all $\sigma \in \mathrm{Gal}(L|K)$. Note that for each $\sigma \in \mathrm{Gal}(L|K)$ we have

$$\sigma \operatorname{div}(a) = \operatorname{div}(\sigma(a)) \qquad \text{for } a \in L^\times$$

and, for all $D_1, D_2 \in \mathcal{D}^{\mathrm{f}}(L, M_L)$:

$$\sigma(D_1 \vee D_2) = \sigma D_1 \vee \sigma D_2, \quad \sigma(D_1 \wedge D_2) = \sigma D_1 \wedge \sigma D_2.$$

Hence for every intermediate field $K \subseteq F \subseteq L$:

$$\sigma\mathcal{P}(F, M_F) = \mathcal{P}\big(\sigma(F), M_{\sigma(F)}\big), \quad \sigma\mathcal{D}(F, M_F) = \mathcal{D}\big(\sigma(F), M_{\sigma(F)}\big).$$

In particular, $\sigma\mathcal{D}(L, M_L) = \mathcal{D}(L, M_L)$.

2.3. **Degree functions and height functions.** A **degree function** on an arithmetic field (K, M) is a function $\deg\colon \mathcal{D}(K, M) \to \mathbb{R}$ satisfying

(D1) $\deg(D + E) = \deg(D) + \deg(E)$ (additivity)
(D2) $D \geq 0 \Rightarrow \deg(D) \geq 0$ (monotonicity)
(D3) $\deg(\operatorname{div} x) = 0$ for all $x \in K^\times$ (product formula).

Because of (D3), $\deg$ induces a homomorphism $\mathrm{Cl}(K, M) \to \mathbb{R}$, which we also denote by $\deg$. A **global field** $(K, M, \deg)$ is an arithmetic field (K, M) equipped with a degree function $\deg$ on $\mathcal{D}(K, M)$.

Remark 2.7. Suppose (K, M) is classical. Then by Lemma 2.6, every divisor $D \in \mathcal{D}(K, M)$ is completely determined by $v \mapsto v(D)\colon M \to \mathbb{R}$, a function of finite support. So every real-valued function $\deg$ on $\mathcal{D}(K, M)$ satisfying (D1) and (D2) is of the form

$$\deg(D) = \sum_{v \in M} v(D)\lambda_v \qquad \text{for } D \in \mathcal{D}(K, M), \tag{2.2}$$

where the $\lambda_v \in \mathbb{R}^{\geq 0}$ are independent of D and uniquely determined by (2.2). Conversely, given any $\lambda_v \in \mathbb{R}^{\geq 0}$ for $v \in M$, we obtain a function $\deg\colon \mathcal{D}(K, M) \to \mathbb{R}$ satisfying (D1) and (D2) by defining $\deg$ as in (2.2). So in this case, (D3) is equivalent to the **product formula**

$$\prod_{v \in M} |a|_v^{\lambda_v} = 1 \qquad \text{for all } a \in K^\times.$$

The construction of a degree function via (2.2) will be generalized below to the case of non-classical (K, M).

Example 2.8. For the classical arithmetic field $\mathbb{Q}$ we have the product formula

$$\prod_{v \in M} |a|_v^{\lambda_v} = 1 \qquad (a \in \mathbb{Q}^\times),$$

with $\lambda_{v_\infty} = 1$ and $\lambda_{v_p} = \log p$ for all primes p. By (2.2), we obtain a degree function on the arithmetic field $\mathbb{Q}$, and we call $(\mathbb{Q}, M, \deg)$ the **global field** $\mathbb{Q}$**.** (By [5] and Remark 2.7, $\deg$ is the *only* degree function on $\mathbb{Q}$, up to multiplication by a non-negative real number.) More generally, for an algebraic number field K, we have the product formula

$$\prod_{v \in M_K} |a|_v^{\lambda_v} = 1 \qquad (a \in K^\times),$$

where the λ_v are determined as follows: For $v \in M_K$, we let $n_v = [K_v : \mathbb{Q}_v]$, where K_v denotes the v-adic completion of K and $\mathbb{Q}_v \subseteq K_v$ is the $(v \restriction \mathbb{Q})$-adic completion of $\mathbb{Q}$. (So either $\mathbb{Q}_v = \mathbb{R}$ or $\mathbb{Q}_v = \mathbb{Q}_p$ for a prime p.) Now

$$\lambda_v = \begin{cases} n_v & \text{if } v \text{ is archimedean,} \\ n_v \log \kappa_v & \text{if } v \text{ is non-archimedean,} \end{cases}$$

where κ_v is the cardinality of the (finite) residue field $\mathcal{O}_v/\mathfrak{m}_v$ of v.

To a global field $(K, M, \deg)$ we associate a **height function** $h\colon K \to \mathbb{R}$ by

$$h(x) := \deg\big((x)_\infty\big) \quad \text{for } x \in K.$$

Here and below, $(0)_\infty := 0$ by convention. One verifies easily the following properties, for elements $x, y, x_1, \dots, x_n \in K$ and $n > 0$:

$$h(x) \geq 0,\ h(0) = h(1) = 0, \tag{2.3}$$

$$h(x) = h(-x) \tag{2.4}$$

$$h(x^k) = |k|h(x) \quad \text{for all } k \in \mathbb{Z},\ x \neq 0, \tag{2.5}$$

$$h(xy) \leq h(x) + h(y), \tag{2.6}$$

$$h(x_1 + \cdots + x_n) \leq \max\big\{h(x_1), \dots, h(x_n)\big\}, \tag{2.7}$$

if M consists only of ultrametric absolute values,

$$h(x_1 + \cdots + x_n) \leq h(x_1) + \cdots + h(x_n) + h(n), \tag{2.8}$$

if M contains an archimedean absolute value.

Remark 2.9. By (2.3), (2.5) and (2.6), the subset of $K^\times$ consisting of the elements of height 0 is a subgroup of $K^\times$ containing each root of unity of K. If M consists only of ultrametric absolute values, then, using also (2.7), one can show that the elements of K of height 0 form a subfield of K which is algebraically closed in K. (See [21], (2.17).)

Example 2.10. The height function of the global field $\mathbb{Q}$ is given by

$$h(a/b) = \max\big\{\log|a|_\infty, \log|b|_\infty\big\} \quad \text{for relatively prime } a, b \in \mathbb{Z} \setminus \{0\}.$$

By Example 2.8, for every global field $(K, M, \deg)$ with K of characteristic zero there exists $c \in \mathbb{R}^{\geq 0}$ such that

$$h(a/b) = c \cdot \max\big\{\log|a|_\infty, \log|b|_\infty\big\} \quad \text{for relatively prime } a, b \in \mathbb{Z} \setminus \{0\}.$$

Let (K, M) be an arithmetic field and $L|K$ be a normal field extension. We say that a degree function $\deg$ on (L, M_L) is $\mathrm{Gal}(L|K)$**-invariant** if $\deg(\sigma D) = \deg(D)$ for all $\sigma \in \mathrm{Gal}(L|K)$ and $D \in \mathcal{D}(L, M_L)$. In this case, $h\big(\sigma(x)\big) = h(x)$ for all $x \in L$. To construct degree and height functions, the following is essential:

Theorem 2.11. (Kani, [20].) *Let $(K, M, \deg)$ be a global field. Then $\deg$ extends uniquely to a $\mathrm{Gal}(K_{\mathrm{alg}}|K)$-invariant degree function on $(K_{\mathrm{alg}}, M_{\mathrm{alg}})$.*

We give a direct proof of this theorem, avoiding the non-standard methods used in the original proof (in order to obtain the auxiliary Proposition 2.12 below). We also close a small gap in [20], see remark (1) at the end of the proof. First, we have to introduce another basic concept in the theory of divisors.

Let (K, M) be an arithmetic field, $L|K$ a finite normal field extension. By

$$\mathrm{N}_{L|K}(D) := [L : K]_{\mathrm{ins}} \sum_{\sigma \in \mathrm{Gal}(L|K)} \sigma D,$$

where $[L:K]_{\text{ins}}$ = degree of inseparability of $L|K$, we define a homomorphism

$$\mathrm{N}_{L|K}\colon \mathcal{D}(L,M_L) \to \mathcal{D}(L,M_L),$$

called the **norm map.** It is clear that for all $D \in \mathcal{D}(L,M_L)$

(N1) $D \geq 0 \;\Rightarrow\; \mathrm{N}_{L|K}(D) \geq 0$,
(N2) $D \in \mathcal{D}(K,M) \;\Rightarrow\; \mathrm{N}_{L|K}(D) = [L:K]\,D$.

By (N1) we have, for $D, D' \in \mathcal{D}(L,M_L)$:

(N3) $\mathrm{N}_{L|K}(D \wedge D') \leq \mathrm{N}_{L|K}(D) \wedge \mathrm{N}_{L|K}(D')$,
(N4) $\mathrm{N}_{L|K}(D \vee D') \geq \mathrm{N}_{L|K}(D) \vee \mathrm{N}_{L|K}(D')$.

Moreover, for all $x \in L^\times$

(N5) $\mathrm{N}_{L|K}\big(\operatorname{div}(x)\big) = \operatorname{div}\big(\mathrm{N}_{L|K}(x)\big)$,

where $\mathrm{N}_{L|K}(x)$ is the field-theoretic norm of the element x. In particular, we have

$$\mathrm{N}_{L|K}\big(\mathcal{P}(L,M_L)\big) \subseteq \mathcal{P}(K,M).$$

This raises the question whether we also have $\mathrm{N}_{L|K}\big(\mathcal{D}(L,M_L)\big) \subseteq \mathcal{D}(K,M)$. This turns out to be false (unless M consists of ultrametric absolute values only), but we can show a weaker statement, sufficient for our purposes. To formulate it, let $\mathcal{D}^{(N)}(L,M_L)$, for $N \geq 1$, denote the set of all divisors $D \in \mathcal{D}(L,M_L)$ admitting a representation of the form $D = \bigwedge_{i=1}^N \bigvee_{j=1}^N (a_{ij})$ with $a_{ij} \in L^\times$ for $1 \leq i,j \leq N$. For $D, D' \in \mathcal{D}(L,M_L)$ we also set $[D,D'] := \big\{E \in \mathcal{D}(L,M_L) : D \leq E \leq D'\big\}$.

Proposition 2.12. *For all $N \geq 1$ there exists a divisor $D_N \in \mathcal{D}(K,M)$ such that $D_N \geq 0$ and*

$$\mathrm{N}_{L|K}\big(\mathcal{D}^{(N)}(L,M_L)\big) \subseteq \mathcal{D}(K,M) + [-D_N, D_N].$$

If all absolute values in M are ultrametric, then

$$\mathrm{N}_{L|K}\big(\mathcal{D}(L,M_L)\big) \subseteq \mathcal{D}(K,M).$$

The last statement (for which also see [21], Satz 1.9) immediately implies:

Corollary 2.13. *Suppose that all absolute values in M are ultrametric. If $L'|L$ is another finite normal field extension, then $\mathrm{N}_{L'|K} = \mathrm{N}_{L|K} \circ \mathrm{N}_{L'|L}$.* □

For the proof of Proposition 2.12, we let $\mathcal{D}_\vee^{(N)}(L,M_L)$ denote the set of all divisors $D \in \mathcal{D}(L,M_L)$ of the form $D = \bigvee_{i=1}^N (a_i)$, where $a_i \in L^\times$. Similarly we define $\mathcal{D}_\wedge^{(N)}(L,M_L)$. We also define

$$\mathcal{D}_\vee(L,M_L) = \bigcup_{N\geq 1} \mathcal{D}_\vee^{(N)}(L,M_L) \quad \text{and} \quad \mathcal{D}_\wedge(L,M_L) = \bigcup_{N\geq 1} \mathcal{D}_\wedge^{(N)}(L,M_L).$$

Both sets are sub-semigroups of $\mathcal{D}(L,M_L)$, by the identity (2.1) in the proof of Lemma 2.5.

Lemma 2.14. *Let $N \geq 1$ and put $D'_N = (2^{N^2})_\infty$. Then for all $D \in \mathcal{D}^{(N)}(L,M_L)$ we have*

$$-D'_N + D_1 - D_2 \leq D \leq D'_N + D_1 - D_2$$

for certain $D_1, D_2 \in \mathcal{D}_\vee^{(N^2)}(L,M_L)$. If all absolute values in M are ultrametric, then $D'_N = 0$, so in this case every $D \in \mathcal{D}(L,M_L)$ is a difference of divisors from $\mathcal{D}_\vee(L,M_L)$.

Proof. Suppose $D = \bigwedge_{i=1}^N E_i$ with $E_i = \bigvee_{j=1}^N (a_{ij})$. We let, for $1 \leq i \leq N$,

$$f_i(T) := \sum_{j=1}^N a_{ij}^{-1} T^j \in L[T],$$

and $f := f_1 \cdots f_N$. Then for each $v \in M_L$, we have $v(E_i) = -v(f_i)$, so $v(D) = \min_i -v(f_i)$. By the inequality (1.8) in the last section applied to the polynomials f/f_i and f_i in $L[T]$, we have

$$-N^2 \log \varepsilon_v - v(f) \leq -v(f/f_i) - v(f_i) \leq N^2 \log \varepsilon_v - v(f)$$

and thus

$$-N^2 \log \varepsilon_v - v(f) + v(f/f_i) \leq -v(f_i) \leq N^2 \log \varepsilon_v - v(f) + v(f/f_i),$$

for $i = 1, \ldots, N$. Hence

$$\begin{aligned} -N^2 \log \varepsilon_v - v(f) + \min_i v(f/f_i) \leq \min_i -v(f_i) \leq \\ N^2 \log \varepsilon_v - v(f) + \min_i v(f/f_i). \end{aligned}$$

Since this holds for all $v \in M_L$, we get

$$-D'_N + D_1 - D_2 \leq D \leq D'_N + D_1 - D_2,$$

where $D'_N = (2^{N^2})_\infty$ and $D_1, D_2 \in \mathcal{D}_\vee^{(N^2)}(L, M_L)$. If M contains only ultrametric absolute values, then $D'_N = 0$. (See Example 2.4.) □

Lemma 2.15. *Let $N \geq 1$ and put $D''_N = (2^{N \cdot [L:K]})_\infty$. Then for each divisor $D \in \mathcal{D}_\vee^{(N)}(L, M_L)$ we have*

$$-D''_N + E \leq \mathrm{N}_{L|K}(D) \leq D''_N + E$$

for some $E \in \mathcal{D}_\vee(K, M)$, and for each $D \in \mathcal{D}_\wedge^{(N)}(L, M_L)$ we have

$$-D''_N + E' \leq \mathrm{N}_{L|K}(D) \leq D''_N + E'$$

for some $E' \in \mathcal{D}_\wedge(K, M)$. If all absolute values in M are ultrametric, then

$$\mathrm{N}_{L|K}\big(\mathcal{D}_\vee(L, M_L)\big) \subseteq \mathcal{D}_\vee(K, M) \quad \textit{and} \quad \mathrm{N}_{L|K}\big(\mathcal{D}_\wedge(L, M_L)\big) \subseteq \mathcal{D}_\wedge(K, M).$$

Proof. Since $D \in \mathcal{D}_\vee^{(N)}(L, M_L) \Leftrightarrow -D \in \mathcal{D}_\wedge^{(N)}(L, M_L)$ for all $N \geq 1$ and $D \in \mathcal{D}(L, M_L)$, it obviously suffices to prove the statement about $\mathcal{D}_\wedge^{(N)}(L, M_L)$. Let $D = \bigwedge_{i=1}^N (a_i)$ with $a_1, \ldots, a_N \in L^\times$ be an element of $\mathcal{D}_\wedge^{(N)}(L, M_L)$. Put

$$f(T) = \sum_{i=1}^N a_i T^i, \qquad g(T) = \left(\prod_{\sigma \in \mathrm{Gal}(L|K)} (\sigma f)(T) \right)^{[L:K]_{\mathrm{ins}}},$$

where for $\sigma \in \mathrm{Gal}(L|K)$ we let $(\sigma f)(T) := \sum_{i=1}^N \sigma(a_i) T^i$. Then $g(T) \in K[T]$, and for all $v \in M_L$

$$v\big(\mathrm{N}_{L|K}(D)\big) = [L:K]_{\mathrm{ins}} \sum_\sigma v(\sigma D) = [L:K]_{\mathrm{ins}} \sum_\sigma v(\sigma f).$$

By (1.8) we have, for all $v \in M_L$:

$$-N' \log \varepsilon_v + v(g) \leq [L:K]_{\mathrm{ins}} \sum_\sigma v(\sigma f) \leq N' \log \varepsilon_v + v(g),$$

where $N' = \deg g = N \cdot [L:K]$. Hence

$$-D''_N + E \le \mathrm{N}_{L|K}(D) \le D''_N + E$$

where $E \in \mathcal{D}_\wedge(K,M)$ and $D''_N = (2^{N'})_\infty$, with $D''_N = 0$ if M contains only ultrametric absolute values. □

Proof of Proposition 2.12. Let $D \in \mathcal{D}(L, M_L)$. By Lemma 2.14,

$$-D'_N + D_1 - D_2 \le D \le D'_N + D_1 - D_2$$

with $D'_N = (2^{N^2})_\infty$ and certain $D_1, D_2 \in \mathcal{D}_\vee^{(N^2)}(L, M_L)$. By Lemma 2.15, there are $E_1, E_2 \in \mathcal{D}_\vee(K,M)$ such that

$$-D''_{N^2} + E_i \le \mathrm{N}_{L|K}(D_i) \le D''_{N^2} + E_i \qquad \text{for } i = 1, 2,$$

with $D''_{N^2} = (2^{N^2[L:K]})_\infty$. It now follows easily that

$$-D_N + (E_1 - E_2) \le \mathrm{N}_{L|K}(D) \le D_N + (E_1 - E_2)$$

where $D_N = [L:K]D'_N + 2D''_{N^2} \in \mathcal{D}(K,M)$, with $D_N \ge 0$, and $D_N = 0$ if M contains no archimedean absolute value. □

Now we turn to the proof of Theorem 2.11. It clearly suffices to show that for each finite normal extension $L|K$, the degree function on (K,M) extends uniquely to a $\mathrm{Gal}(L|K)$-invariant degree function $\deg_L$ on (L, M_L). Hence suppose that $L|K$ is a finite normal extension. Put, for $D \in \mathcal{D}(L, M_L)$:

$$\overline{\deg}(D) := \inf\big\{\deg(E) : E \in \mathcal{D}(K,M), E \ge \mathrm{N}_{L|K}(D)\big\}, \tag{2.9}$$

$$\underline{\deg}(D) := \sup\big\{\deg(E) : E \in \mathcal{D}(K,M), E \le \mathrm{N}_{L|K}(D)\big\}. \tag{2.10}$$

Note that for each $D \in \mathcal{D}(L, M_L)$ we have $D \in \mathcal{D}^{(N)}(L, M_L)$ for some $N \ge 1$, and thus $-D_N \le \mathrm{N}_{L|K}(D) - E \le D_N$ for some $E \in \mathcal{D}(K,M)$, by Proposition 2.12. Hence the infimum (2.9) and the supremum (2.10) exist and are finite. Note that we have

$$D \in \mathcal{D}(K,M) \;\Rightarrow\; \overline{\deg}(D) = \underline{\deg}(D) = d\deg(D) \tag{2.11}$$

by (N2), where $d = [L:K]$. Moreover, for $D, D' \in \mathcal{D}(L, M_L)$,

$$\underline{\deg}(D) \le \overline{\deg}(D), \tag{2.12}$$

$$\underline{\deg}(D + D') \ge \underline{\deg}(D) + \underline{\deg}(D'), \tag{2.13}$$

$$\overline{\deg}(D + D') \le \overline{\deg}(D) + \overline{\deg}(D'), \tag{2.14}$$

and hence, by induction,

$$\underline{\deg}(D) \le \frac{1}{n}\underline{\deg}(nD) \le \frac{1}{n}\overline{\deg}(nD) \le \overline{\deg}(D) \qquad \text{for all } n \ge 1.$$

Let $D \in \mathcal{D}^{(N)}(L, M_L)$. Since $2^n D \in \mathcal{D}^{(N)}(L, M_L)$ for all n, we obtain in particular

$$\frac{1}{2^n}\underline{\deg}(2^n D) \le \frac{1}{2^{n+1}}\underline{\deg}(2^{n+1}D) \quad \text{and} \quad \frac{1}{2^n}\overline{\deg}(2^n D) \ge \frac{1}{2^{n+1}}\overline{\deg}(2^{n+1}D)$$

for all n, and, by Proposition 2.12,

$$0 \le \overline{\deg}(2^n D) - \underline{\deg}(2^n D) \le 2\deg(D_N) \qquad \text{for all } n.$$

Hence the sequences $\big\{\overline{\deg}(2^n D)/2^n\big\}_n$ and $\big\{\underline{\deg}(2^n D)/2^n\big\}_n$ of real numbers converge and have the same limit. Put

$$\deg_L(D) := \frac{1}{d}\lim_{n\to\infty}\frac{\overline{\deg}(2^n D)}{2^n} = \frac{1}{d}\lim_{n\to\infty}\frac{\underline{\deg}(2^n D)}{2^n}.$$

We claim that $\deg_L$ is the required degree function on (L, M_L): from (2.13) and (2.14) we see that $\deg_L$ is additive, from (N1) we get that $\overline{\deg}(D) \geq 0$ and $\underline{\deg}(D) \geq 0$ if $D \geq 0$, and by (N5) we have for a principal divisor $D = (x)$, $x \in L^\times$, that $\overline{\deg}(nD) = \deg\big(\operatorname{div}\big(\mathrm{N}_{L|K}(x^n)\big)\big) = 0$. So $\deg_L$ is a degree function on (L, M_L) which extends $\deg$, by (2.11).

To prove uniqueness of $\deg_L$, let $\deg'_L$ be another $\mathrm{Gal}(L|K)$-invariant degree function on (L, M_L) extending $\deg$. Then for all $D \in \mathcal{D}(L, M_L)$

$$\deg'_L(D) = \frac{1}{d}\deg'_L\big(\mathrm{N}_{L|K}(D)\big),$$

hence $\deg'_L(D) \leq \deg(E)/d$ whenever $E \in \mathcal{D}(K, M)$ is such that $E \geq \mathrm{N}_{L|K}(D)$, and thus $\deg'_L(D) \leq \overline{\deg}(D)/d$. Similarly we get $\deg'_L(D) \geq \underline{\deg}(D)/d$ and hence $\deg_L = \deg'_L$. This finishes the proof of Theorem 2.11.

Remarks 2.16.

(1) In the proof of Theorem 2.11 given in [20], Kani defines $\deg_L(D)$ as the limit $\lim_{n\to\infty} \frac{1}{d}\frac{\overline{\deg}(nD)}{n}$. Note however that it is not clear *a priori* that the sequence $\big\{\overline{\deg}(nD)/n\big\}_{n>0}$ has a limit in $\mathbb{R}$.

(2) Suppose that all absolute values in M are ultrametric. Then the extension of $\deg$ to a $\mathrm{Gal}(K_{\mathrm{alg}}|K)$-invariant degree function on $(K_{\mathrm{alg}}, M_{\mathrm{alg}})$ admits a simple description involving the norm map: For a divisor $D \in \mathcal{D}(K_{\mathrm{alg}}, M_{\mathrm{alg}})$ choose a finite normal extension L of K such that $D \in \mathcal{D}(L, M_L)$. By Proposition 2.12, we have $\mathrm{N}_{L|K}(D) \in \mathcal{D}(K, M)$, so we may set

$$\deg(D) := \frac{1}{d}\deg\big(\mathrm{N}_{L|K}(D)\big) \qquad \text{where } d := [L : K].$$

Using (N2) and Corollary 2.13 one shows that this definition does not depend on the choice of L. Clearly the map $\deg\colon \mathcal{D}(K_{\mathrm{alg}}, M_{\mathrm{alg}}) \to \mathbb{R}$ so defined is a degree function which is $\mathrm{Gal}(K_{\mathrm{alg}}|K)$-invariant and extends the one on $\mathcal{D}(K, M)$.

We denote the unique extension of $\deg$ to a $\mathrm{Gal}(K_{\mathrm{alg}}|K)$-invariant degree function on $(K_{\mathrm{alg}}, M_{\mathrm{alg}})$ also by $\deg$, and the corresponding height function by h.

Example 2.17. The height function which is associated to the extension of the degree function on the global field $\mathbb{Q}$ to a $\mathrm{Gal}(\mathbb{Q}_{\mathrm{alg}}|\mathbb{Q})$-invariant degree function on $(\mathbb{Q}_{\mathrm{alg}}, M_{\mathrm{alg}})$ is the usual (logarithmic) height function on $\mathbb{Q}_{\mathrm{alg}}$ used in diophantine geometry [23]. For an element α of an algebraic number field K we have

$$h(\alpha) = \frac{1}{[K : \mathbb{Q}]}\sum_{v\in M_K} \lambda_v \log^+ |\alpha|_v, \tag{2.15}$$

where the λ_v are as in Example 2.8. It is also well-known that if $f(T) \in \mathbb{Z}[T]$ is a non-zero irreducible polynomial and $\alpha \in \mathbb{Q}_{\mathrm{alg}}$ a zero of f, then $h(\alpha) = m(f)/\deg f$. (See, e.g., [23], Chapter 3, §1.)

2.4. **Height on projective and affine space.** Until the end of this section, let $(K, M, \deg)$ be a global field. Given a point $P = [x_0 : \cdots : x_n]$ of projective space $\mathbb{P}^n = \mathbb{P}^n(K_{\mathrm{alg}})$ over K_{alg}, we have $\bigwedge_{x_i\neq 0}(\lambda x_i) = \bigwedge_{x_i\neq 0}(x_i) + (\lambda)$ for all $\lambda \in K_{\mathrm{alg}}^\times$. Hence to P we may associate a divisor class

$$\operatorname{div}[x_0 : \cdots : x_n] := \bigwedge_{x_i\neq 0}(x_i) \qquad \text{(modulo principal divisors).}$$

We define the **projective height function** $h\colon \mathbb{P}^n \to \mathbb{R}$ associated to $(K, M, \deg)$ by

$$h\big([x_0 : \cdots : x_n]\big) := -\deg\big(\operatorname{div}[x_0 : \cdots : x_n]\big).$$

Note that $h\big([x_0 : \cdots : x_n]\big) \geq 0$. (Since one can always assume that one of the x_i is 1.) We record a few more basic properties:

Lemma 2.18. *Let $P = [x_0 : \cdots : x_n]$ and Q be points in $\mathbb{P}^n$. Then:*

(1) *For every permutation π of the set $\{0, \ldots, n\}$,*

$$h\big([x_{\pi(0)} : \cdots : x_{\pi(n)}]\big) = h\big([x_0 : \cdots : x_n]\big).$$

(2) *If $\sigma \in \operatorname{Gal}(K_{\mathrm{alg}}|K)$, then $h(\sigma P) = h(P)$, where*

$$\sigma P := \big[\sigma(x_0) : \cdots : \sigma(x_n)\big] \in \mathbb{P}^n.$$

(3) *If $\operatorname{div} P \leq \operatorname{div} Q$, then $h(P) \geq h(Q)$. In particular: If $x_{n+1} \in K_{\mathrm{alg}}$, then*

$$h\big([x_0 : \cdots : x_n]\big) \leq h\big([x_0 : \cdots : x_{n+1}]\big).$$

(4) *For all $x \in K$, $h(x) = h\big([1 : x]\big)$.*

(5) $h\big([x_0 : \cdots : x_n]\big) \geq \max\big\{h(x_i/x_j) : x_j \neq 0\big\}$.

Proof. Self-evident, except maybe for (5). Note that both sides of the inequality remain unchanged if we multiply each x_i with a non-zero constant. Hence, by (1) we may assume $x_0 = 1$ and $h(x_1) = h(x_1/x_0) \geq h(x_i/x_j)$ for all $0 \leq i, j \leq n$ with $x_j \neq 0$, so $h\big([1 : x_1 : \cdots : x_n]\big) \geq h\big([1 : x_1]\big) = h(x_1) \geq h(x_i/x_j)$ for all such i, j, by (3) and (4). □

Example 2.19. For the global field $\mathbb{Q}$, the height of a point $P = [x_0 : \cdots : x_n] \in \mathbb{P}^n(\mathbb{Q}_{\mathrm{alg}})$ is the so-called *absolute logarithmic height* of P (see [23]), given by

$$h(P) = \frac{1}{[\mathbb{Q}(P) : \mathbb{Q}]} \sum_{v \in M_{\mathbb{Q}(P)}} \lambda_v \max_{x_i \neq 0} \log |x_i|_v.$$

We now want to prove some results that are analogous to the ones proved for the height function on $\mathbb{P}^n(\mathbb{Q}_{\mathrm{alg}})$ in [23]. The first one (Corollary 2.22) connects the height of the coefficient tuple of a univariate polynomial (considered as a point in projective space) and the height of its zeros (as elements of K_{alg}). We will deduce it from a global version of Gelfond's Lemma. First we define a height function for polynomials:

Definition 2.20. Let $f(X) = \sum_\nu a_\nu X^\nu$ be a non-zero polynomial in the indeterminates $X = (X_1, \ldots, X_N)$ with coefficients $a_\nu \in K_{\mathrm{alg}}$. We put

$$\operatorname{coeff}(f) = \bigwedge_{a_\nu \neq 0} (a_\nu),$$

a divisor in $\mathcal{D}(K_{\mathrm{alg}}, M_{\mathrm{alg}})$, and

$$h_{\mathrm{coeff}}(f) := -\deg \operatorname{coeff}(f).$$

Note that $h_{\mathrm{coeff}}(f)$ is nothing but the height of the tuple consisting of the non-zero coefficients of f (in any order), considered as a point in a projective space over K_{alg}. Similarly to $h_{\mathrm{coeff}}(f)$ for a single polynomial f, we define $h_{\mathrm{coeff}}(f_1, \ldots, f_n)$ for a finite sequence of polynomials $(f_1, \ldots, f_n)$ in $K_{\mathrm{alg}}[X]$, where $n > 0$ and not all $f_i = 0$.

Lemma 2.21. (Gelfond's Lemma, global version.) *Let $f_1, \dots, f_n \in K_{\mathrm{alg}}[X]$ be non-zero, where $n > 0$, and put $f = f_1 \cdots f_n \neq 0$ and $d = \deg_{(X)} f$. Then*

$$-dh(2) + \sum_{i=1}^{n} h_{\mathrm{coeff}}(f_i) \leq h_{\mathrm{coeff}}(f) \leq dh(2) + \sum_{i=1}^{n} h_{\mathrm{coeff}}(f_i).$$

Proof. By (1.8), in $\mathcal{D}(K_{\mathrm{alg}}, M_{\mathrm{alg}})$ we have

$$-d(2)_\infty + \sum_{i=1}^{n} \mathrm{coeff}(f_i) \leq \mathrm{coeff}(f) \leq d(2)_\infty + \sum_{i=1}^{n} \mathrm{coeff}(f_i).$$

Applying $-\deg$ to these inequalities gives the desired result. □

Corollary 2.22. *Let $f(T) = a_0 + a_1 T + \cdots + a_d T^d \in K_{\mathrm{alg}}[T]$, $a_d \neq 0$, and $\alpha_1, \dots, \alpha_d \in K_{\mathrm{alg}}$ with $f(T) = a_d \prod_{j=1}^{d} (T - \alpha_j)$. Then*

$$-dh(2) + \sum_{j=1}^{d} h(\alpha_j) \leq h_{\mathrm{coeff}}(f) \leq dh(2) + \sum_{j=1}^{d} h(\alpha_j).$$

So if all $v \in M$ are ultrametric, then

$$h_{\mathrm{coeff}}(f) = h\big([a_0 : \cdots : a_d]\big) = \sum_{j=1}^{d} h(\alpha_j),$$

and if $a_0, \dots, a_d \in K$, then for every $\alpha = \alpha_j$:

$$d\big(h(\alpha) - h(2)\big) \leq h_{\mathrm{coeff}}(f) \leq d\big(h(\alpha) + h(2)\big).$$

Proof. Note that the inequalities to be proved remain unchanged if $f(T)$ is replaced by $(1/a_d)f(T)$. Therefore we may assume $a_d = 1$. Now the desired result follows from the previous lemma, applied to the polynomials $f_1, \dots, f_d$ with $f_j(T) = T - \alpha_j$ for $j = 1, \dots, d$. □

We say that $(K, M, \deg)$ **has the finiteness property** if for every real number C, there are only finitely many $\alpha \in K$ with $h(\alpha) \leq C$. Note that if $(K, M, \deg)$ has the finiteness property, then for every real number C there are only finitely many $P \in \mathbb{P}^n(K)$ such that $h(P) \leq C$. (By (5) in Lemma 2.18.) We now use the corollary above to show:

Proposition 2.23. *If $(K, M, \deg)$ has the finiteness property, then for all real numbers C and d, there are only finitely many $\alpha \in K_{\mathrm{alg}}$ such that $h(\alpha) \leq C$ and $[K(\alpha) : K] \leq d$.*

Proof. Suppose $\alpha \in K_{\mathrm{alg}}$ is such that $h(\alpha) \leq C$. Let $\alpha_1, \dots, \alpha_d$ be the conjugates of α in K_{alg}, where $d := [K(\alpha) : K]$. The minimal polynomial of α over K is given by

$$(T - \alpha_1) \cdots (T - \alpha_d) = T^d + a_{d-1} T^{d-1} + \cdots + a_0$$

for certain $a_0, \dots, a_{d-1} \in K$. Now by Corollary 2.22:

$$h\big([1 : a_{d-1} : \cdots : a_0]\big) \leq dh(2) + \sum_{j=1}^{d} h(\alpha_j) = dh(2) + dh(\alpha) \leq d\big(h(2) + C\big).$$

So there are only finitely many possibilities for choosing $a_0, \dots, a_{d-1} \in K$, and for given $a_0, \dots, a_{d-1} \in K$, the polynomial $T^d + a_{d-1} T^{d-1} + \cdots + a_0$ has at most d distinct zeros in K_{alg}. Hence there are only finitely many $\alpha \in K_{\mathrm{alg}}$ with $h(\alpha) \leq C$ and $[K(\alpha) : K] \leq d$. □

Example 2.24. The global field $\mathbb{Q}$ has the finiteness property. In this case, Proposition 2.23 is due to Northcott [29].

For a point $P = [x_0 : \cdots : x_n] \in \mathbb{P}^n$, we let

$$K(P) := K(x_0/x_i, \ldots, x_n/x_i) \quad \text{for any } i \text{ with } x_i \neq 0.$$

If K is perfect, then $K(P)$ is the fixed field in K_{alg} of the subgroup of $\mathrm{Gal}(K_{\mathrm{alg}}|K)$ consisting of all σ such that $\sigma P = P$.

Corollary 2.25. *Suppose that $(K, M, \deg)$ has the finiteness property.*

(1) *For all real numbers C and d, there are only finitely many $P \in \mathbb{P}^n(K_{\mathrm{alg}})$ with $h(P) \leq C$ and $[K(P) : K] \leq d$.*
(2) *A non-zero element of K_{alg} has height 0 if and only if it is a root of unity.*

Proof. Part (1) follows from the last proposition and part (5) of Lemma 2.18. If $x \in K_{\mathrm{alg}}^{\times}$ is a root of unity, then $h(x) = 0$ by (2.5). For the converse, let $x \in K_{\mathrm{alg}}^{\times}$ with $h(x) = 0$, and $L := K(x)$, $d := [L : K]$. By the proposition, the subgroup of $L^{\times}$ consisting of all elements of height 0 is finite. Therefore $x^m = 1$ for some positive m. □

The embedding of affine space $\mathbb{A}^n = \mathbb{A}^n(K_{\mathrm{alg}})$ over K_{alg} into projective space

$$\mathbb{A}^n(K_{\mathrm{alg}}) \hookrightarrow \mathbb{P}^n(K_{\mathrm{alg}}), \qquad (x_1, \ldots, x_n) \mapsto [1 : x_1 : \cdots : x_n],$$

can be used to define the **affine height** of a point $P = (x_1, \ldots, x_n) \in \mathbb{A}^n$: if $x_i \neq 0$ for some i, then

$$h(P) := h\big([1 : x_1 : \cdots : x_n]\big) = \deg \bigvee_{x_i \neq 0} (x_i)_{\infty},$$

and $h(P) := 0$ if all $x_i = 0$. It is also convenient to introduce the abbreviation

$$h_{\max}(P) := \max_i h(x_i) \qquad \text{for } P = (x_1, \ldots, x_n) \in \mathbb{A}^n,$$

with the convention that the maximum is 0 for $n = 0$. Here are some properties of h (affine height) and $h_{\max}$ on $\mathbb{A}^n$:

Lemma 2.26. *Let $P = (x_1, \ldots, x_n), Q \in \mathbb{A}^n$.*

(1) *For every permutation π of the set $\{1, \ldots, n\}$, we have*

$$h\big(x_{\pi(1)}, \ldots, x_{\pi(n)}\big) = h(x_1, \ldots, x_n),$$

and similarly for $h_{\max}$ in place of h.
(2) *If $\sigma \in \mathrm{Gal}(K_{\mathrm{alg}}|K)$, then $h(\sigma P) = h(P)$ and $h_{\max}(\sigma P) = h_{\max}(P)$, where $\sigma P := \big(\sigma(x_1), \ldots, \sigma(x_n)\big)$.*
(3) *$h(x_1, \ldots, x_n) \geq h\big([x_1 : \cdots : x_n]\big)$, if $P \neq 0$.*
(4) *$h_{\max}(x_1, \ldots, x_n) \leq h(x_1, \ldots, x_n) \leq h(x_1) + \cdots + h(x_n)$.*
(5) *$h_{\max}(P + Q) \leq h_{\max}(P) + h_{\max}(Q) + h(2)$.*
(6) *For $\lambda \in K$, we have $h_{\max}(\lambda P) \leq h(\lambda) + h_{\max}(P)$.*

Proof. Parts (1)–(3) are clear. For (4), note that $\bigvee_{x_i \neq 0}(x_i)_{\infty} \leq \sum_{x_i \neq 0}(x_i)_{\infty}$. Now take $\deg$ on both sides of this inequality to get the second inequality in (4). The first inequality follows from (3) and Lemma 2.18, (5). Items (5) and (6) follow from (2.6) and (2.8), respectively. □

Let $f = (f_0, \dots, f_m)$ be a sequence of homogeneous polynomials of degree $d \in \mathbb{N}$ in $K_{\mathrm{alg}}[Y_0, \dots, Y_n]$, not all $f_0, \dots, f_m$ equal to zero. Then f defines a rational map

$$f\colon \mathbb{P}^n \setminus Z \to \mathbb{P}^m, \quad P \mapsto f(P),$$

where $Z = \big\{P \in \mathbb{P}^n : f_0(P) = \cdots = f_m(P) = 0\big\}$. For a point $P \in \mathbb{P}^n \setminus Z$ we have the inequality

$$h\big(f(P)\big) \leq d\, h(P) + h_{\mathrm{coeff}}(f) + h\big(\mathrm{mon}(f)\big), \tag{2.16}$$

where $\mathrm{mon}(f) := \max_i \mathrm{mon}(f_i)$ and $\mathrm{mon}(f_i)$ is as defined in Lemma 1.3. If M contains only ultrametric absolute values, then we have

$$h\big(f(P)\big) \leq d\, h(P) + h_{\mathrm{coeff}}(f).$$

All this follows easily from Lemma 1.3.

A sequence $f = (f_1, \dots, f_m)$ of polynomials of degree $\leq d$ in $K_{\mathrm{alg}}[Y_1, \dots, Y_n]$ gives rise to a rational map

$$f\colon \mathbb{A}^n \to \mathbb{A}^m, \quad P \mapsto f(P).$$

Homogenizing, we obtain from (2.16) an upper bound on the affine height of $f(P)$:

$$h\big(f(P)\big) \leq d\, h(P) + h_{\mathrm{coeff}}(f) + h\big(\mathrm{mon}(f)\big), \tag{2.17}$$

where $\mathrm{mon}(f) := \max_i \mathrm{mon}(f_i)$. If M contains no archimedean absolute value, then $h\big(f(P)\big) \leq d\, h(P) + h_{\mathrm{coeff}}(f)$.

Until the end of this section we assume that K has characteristic zero. This has the pleasant consequence (by Example 2.10) that the height function h is increasing on $\mathbb{N}$ (for all $a, b \in \mathbb{N}$, if $a \leq b$, then $h(a) \leq h(b)$), which we use tacitly below.

Corollary 2.27. *Let $f, g \in K[T]$, $m = \deg f$, $n = \deg g$, and let $r = \mathrm{res}(f, g) \in K$ be the resultant of f and g. Then*

$$h(r) \leq n\, h_{\mathrm{coeff}}(f) + m\, h_{\mathrm{coeff}}(g) + m\, h(1+n) + mn \log 2.$$

Proof. We may assume that f is monic. Let $\alpha_1, \dots, \alpha_m$ be the zeros of f in K_{alg}. Then $r = (-1)^{mn} g(\alpha_1) \cdots g(\alpha_m)$, hence by (2.17) and Corollary 2.22

$$\begin{aligned} h(r) \leq \sum_{i=1}^m h\big(g(\alpha_i)\big) &\leq \sum_{i=1}^m \big(n\, h(\alpha_i) + h_{\mathrm{coeff}}(g) + h(1+n)\big) \\ &\leq n\, h_{\mathrm{coeff}}(f) + m\, h_{\mathrm{coeff}}(g) + m\, h(1+n) + mn \log 2 \end{aligned}$$

as claimed. □

2.5. **Height of matrices.** We identify each square matrix $A \in K^{n\times n}$ over K with a point in $\mathbb{A}^{n^2}(K)$. The next lemma contains two inequalities which allow us to estimate the height of the determinant of A in terms of $h_{\max}(A)$ and $h(A)$, respectively.

Lemma 2.28. *Let $A \in K^{n\times n}$. Then*

(1) $h(\det A) \leq n\big(h(n) + h_{\max}(A)\big)$;
(2) $h(\det A) \leq n\left(\frac{1}{2}h(n) + h(A)\right)$.

If M contains no archimedean absolute value, then $h(\det A) \leq n h_{\max}(A)$.

Proof. Write $A = (a_{ij})$. By estimating $h(\det A)$ using (2.6) and (2.8), we obtain the bound

$$h(\det A) \leq n h_{\max}(A) + h(n!) \leq n h_{\max}(A) + n h(n),$$

showing (1). In fact, if M consists only of ultrametric absolute values, one gets similarly $h(\det A) \leq nh_{\max}(A)$ by (2.7), which shows the last statement. For each $v \in M$, by Lemma 1.2, (2)

$$\max\{|\det A|_v, 1\} \leq \max\{|n|_v^{n/2}, 1\} \prod_j \max_i \max\{|a_{ij}|_v, 1\}.$$

Since this is valid for all $v \in M$, we have

$$(\det A)_\infty \leq \frac{n}{2}(n)_\infty + \sum_j \bigvee_i (a_{ij})_\infty.$$

Taking $\deg$ on both sides of the inequality yields (2). □

Suppose f is an invertible linear transformation

$$\mathbb{P}^n \to \mathbb{P}^n,\ P = [x_0 : \cdots : x_n] \mapsto P' = [x_0' : \cdots : x_n'],$$

given by

$$x_i' = \sum_{j=0}^n a_{ij} x_j \qquad \text{for } i = 0, \ldots, n, \tag{2.18}$$

with $A = (a_{ij}) \in \mathrm{GL}_{n+1}(K_{\mathrm{alg}})$. Note that $h_{\mathrm{coeff}}(f)$ is the height of A considered as a point in the projective space $\mathbb{P}^{n(n+2)}$.

Lemma 2.29. *For the height of the inverse transformation f^{-1} of f we have*

$$h_{\mathrm{coeff}}(f^{-1}) \leq n\left(\frac{1}{2}h(n) + h_{\mathrm{coeff}}(f)\right),$$

and the heights of the corresponding points P and $P' = f(P)$ satisfy

$$h(P) - h_{\mathrm{coeff}}(f^{-1}) - h(n+1) \leq h(P') \leq h(P) + h_{\mathrm{coeff}}(f) + h(n+1).$$

Proof. By Cramer's Rule the inverse f^{-1} of f is given by the adjoint matrix A^{ad} of A, whose entries are the signed $n \times n$-minors of A. Let B be range over the $n \times n$-submatrices of A. By Lemma 1.2, (2), for each $v \in M$ we have

$$|\det B|_v \leq \max\left\{|n|_v^{n/2}, 1\right\} \prod_j \max_i |a_{ij}|_v,$$

and hence in $\mathcal{D}(K_{\mathrm{alg}}, M_{\mathrm{alg}})$:

$$2\,(\det B) \geq n(n)_\infty + 2\sum_j \bigwedge_i (a_{ij}) \geq n\left((n)_\infty + 2\bigwedge_{i,j}(a_{ij})\right)$$

Therefore, construing A^{ad} as a point in $\mathbb{P}^{n(n+2)}$:

$$2\ \mathrm{div}\, A^{\mathrm{ad}} \geq n\left((n)_\infty + 2\bigwedge_{i,j}(a_{ij})\right).$$

The first inequality now follows by taking $-\deg$ on both sides of this inequality, and the rest is a consequence of (2.16). □

The following lemma gives an estimate for the height of the product of two matrices in terms of the heights of the two factors.

Lemma 2.30. *Let $A \in K^{m\times n}$, $B \in K^{n\times p}$, with $m, n, p \geq 1$. Then*

(1) $h(A \cdot B) \leq h(A) + h(B) + h(n)$;

(2) $h_{\max}(A \cdot B) \leq n\big(h_{\max}(A) + h_{\max}(B)\big) + h(n)$.

Proof. Write $A = (a_{ij})$, and consider A as a point in $\mathbb{A}^{mn}$ and B as a point in $\mathbb{A}^{np}$. Apply (2.17) to the map $f\colon \mathbb{A}^{np} \to \mathbb{A}^{mp}$ given by $f(x_{jk}) = \sum_j a_{ij}x_{jk}$, and $P = B$, and use that $h_{\mathrm{coeff}}(f) \leq h(A)$ by Lemma 2.26, (3) to obtain (1). For (2) just compute $h_{\max}(A \cdot B)$ using (2.6), (2.8). □

Now consider $A = (a_{ij}) \in \mathrm{GL}_n(K_{\mathrm{alg}})$, $n > 0$, and the invertible linear transformation

$$\mathbb{A}^n \to \mathbb{A}^n,\ P = (x_1, \ldots, x_n) \mapsto P' = (x_1', \ldots, x_n')$$

given by

$$x_i' = \sum_{j=1}^{n} a_{ij}x_j \qquad \text{for } i = 1, \ldots, n.$$

Corollary 2.31. *We have* $h(A^{-1}) \leq n\left(\frac{1}{2}h(n) + h(A)\right)$ *and*

$$h(P) - h(A^{-1}) - h(n) \leq h(P') \leq h(P) + h(A) + h(n). \tag{2.19}$$

Proof. Put $A^* := \left[\begin{smallmatrix} 1 & \\ & A \end{smallmatrix}\right] \in \mathrm{GL}_{n+1}(K_{\mathrm{alg}})$ and let f^* be the linear automorphism of $\mathbb{P}^n$ with coefficient matrix A^* as in (2.18). Note that $(f^*)^{-1}$ is given by the matrix $(A^{-1})^*$, and $h_{\mathrm{coeff}}(f^*) = h(A)$, $h_{\mathrm{coeff}}((f^*)^{-1}) = h(A^{-1})$. The first inequality now follows from Lemma 2.29, and (2.19) from the previous lemma. □

2.6. **Generalized Vandermonde matrices.** For $\xi \in K$ and $\mu, s \in \mathbb{N}$ with $1 \leq \mu \leq s$ we define the $\mu \times s$-matrix

$$A(\xi, \mu, s) := \begin{bmatrix} 1 & \xi & \xi^2 & \cdots & \cdots & \xi^i & \cdots & \xi^{s-1} \\ 0 & 1 & 2\xi & \cdots & \cdots & i\xi^{i-1} & \cdots & (s-1)\xi^{s-2} \\ \vdots & & \ddots & & & \vdots & & \vdots \\ 0 & \cdots & 0 & (\mu-1)! & \cdots & \frac{(i-1)!}{(i-\mu)!}\,\xi^{i-\mu} & \cdots & \frac{(s-1)!}{(s-\mu)!}\,\xi^{s-\mu} \end{bmatrix},$$

and given $\xi = (\xi_0, \ldots, \xi_n) \in K^{n+1}$, $\mu = (\mu_0, \ldots, \mu_n) \in \mathbb{N}^{n+1}$ and $s \in \mathbb{N}$ with $1 \leq \mu_i \leq s$ for all i, we define the generalized Vandermonde matrix $A(\xi, \mu, s)$ corresponding to (ξ, μ, s) as follows:

$$A(\xi, \mu, s) := \begin{bmatrix} A(\xi_0, \mu_0, s) \\ A(\xi_1, \mu_1, s) \\ \vdots \\ A(\xi_n, \mu_n, s) \end{bmatrix} \in K^{s \times s}.$$

These matrices will play an important role in the proof of Theorem 0.1. If all $\mu_i = 1$, then

$$A(\xi, \mu, s) = \begin{bmatrix} 1 & \xi_0 & \ldots & \xi_0^{s-1} \\ 1 & \xi_1 & \ldots & \xi_1^{s-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \xi_n & \ldots & \xi_n^{s-1} \end{bmatrix}.$$

So if in addition $n = s - 1$, then $A(\xi, \mu, s)$ is just the usual Vandermonde matrix for $\xi_0, \ldots, \xi_{s-1}$. At any rate, in this case we have

$$h\big(A(\xi, \mu, s)\big) = (s-1)h(\xi_0, \ldots, \xi_n),$$

since

$$\bigvee_{\substack{0\le i\le n\\ \xi_i\neq 0}} \bigvee_{1\le j<s} (\xi_i^j)_\infty = (s-1) \bigvee_{\substack{0\le i\le n\\ \xi_i\neq 0}} (\xi_i)_\infty. \tag{2.20}$$

In general, we have the following upper bound on the height of matrices of this type:

Proposition 2.32. *With $A = A(\xi,\mu,s)$ as above, we have*

$$h(A) \le (s-1)h(\xi_0,\dots,\xi_n) + (\mu_{\max}-1)h\big([1:2:\cdots:s-1]\big),$$

where $\mu_{\max} := \max\{\mu_0,\dots,\mu_n\}$.

In the proof we use the lemma below. Here $f(T)\in K[T]$ is a polynomial of degree $d>0$, and for $j\in\mathbb{N}$ we write $f^{(j)}$ for the j-th formal derivative of f, as usual.

Lemma 2.33. *For $\xi\in K^\times$ and $0\le j\le d$:*

$$\operatorname{coeff}\big(f^{(j)}(\xi T)\big) \ge \operatorname{coeff}\big(f(\xi T)\big) + j\cdot\big(\operatorname{div}[1:2:\cdots:d] - (\xi)\big).$$

In particular, we have $h_{\operatorname{coeff}}(f^{(j)}) \le h_{\operatorname{coeff}}(f) + j\cdot h\big([1:2:\cdots:d]\big)$.

Proof. By induction on j, it only suffices to consider the case $j=1$. We write $f(T) = a_0 + a_1T + \cdots + a_dT^d$ with $a_0,\dots,a_d\in K$ and compute:

$$\begin{aligned}\operatorname{coeff}\big(f'(\xi T)\big) &= \bigwedge_{\substack{1\le i\le d\\ a_i\neq 0}} (ia_i\xi^{i-1}) \ge \bigwedge_{\substack{1\le i\le d\\ a_i\neq 0}} (a_i\xi^{i-1}) + \bigwedge_{i=1}^{d}(i)\\ &= \bigwedge_{\substack{1\le i\le d\\ a_i\neq 0}} (a_i\xi^{i}) + \bigwedge_{i=1}^{d}(i) - (\xi) = \operatorname{coeff}\big(f(\xi T)\big) + \operatorname{div}[1:2:\cdots:d] - (\xi)\end{aligned}$$

as required. □

Let $s\in\mathbb{N}$ and $f(T) := 1+T+\cdots+T^{s-1}\in K[T]$. Note that for $\xi\in K^\times$ and $\mu\in\mathbb{N}$ with $1\le\mu\le s$ we have

$$\operatorname{div}\big(A(\xi,\mu,s)\big) = \bigwedge_{j=0}^{\mu-1} \operatorname{coeff}\big(f^{(j)}(\xi T)\big).$$

(Here we construe $A(\xi,\mu,s)$ as a point in $\mathbb{A}^{s^2} \hookrightarrow \mathbb{P}^{s^2-1}$.) Hence by the lemma, in $\operatorname{Cl}(K,M)$ we get:

$$\begin{aligned}\operatorname{div}\big(A(\xi,\mu,s)\big) &\ge \bigwedge_{j=0}^{\mu-1}\Big(\operatorname{coeff}\big(f(\xi T)\big) + j\cdot\big(\operatorname{div}[1:2:\cdots:s-1] - (\xi)\big)\Big)\\ &= \operatorname{coeff}\big(f(\xi T)\big) + \bigwedge_{j=0}^{\mu-1} j\cdot\operatorname{div}[1:2:\cdots:s-1]\\ &\ge \operatorname{coeff}\big(f(\xi T)\big) + (\mu-1)\cdot\operatorname{div}[1:2:\cdots:s-1].\end{aligned} \tag{2.21}$$

We now prove Proposition 2.32. For this, let $\xi = (\xi_0, \dots, \xi_n) \in K^{n+1}$ and $\mu = (\mu_0, \dots, \mu_n) \in \mathbb{N}^{n+1}$, and let s be an integer with $1 \leq \mu_i \leq s$ for all i. We may assume that $\xi_i \neq 0$ for all i. Then, by (2.21):

$$\begin{aligned}\operatorname{div}\big(A(\xi,\mu,s)\big) &= \bigwedge_{i=0}^{n} \operatorname{div}\big(A(\xi_i,\mu_i,s)\big)\\ &\geq \bigwedge_{i=0}^{n} \operatorname{coeff}\big(f(\xi_i T)\big) + (\mu_i - 1)\cdot \operatorname{div}[1:2:\cdots:s-1]\\ &\geq \bigwedge_{i=0}^{n}\bigwedge_{j=0}^{s-1} (\xi_i^j) + (\mu_{\max}-1)\cdot \operatorname{div}[1:2:\cdots:s-1].\end{aligned}$$

Applying $-\deg$ on both sides of this inequality and using (2.20) yields the claim.

3. A HEIGHT FUNCTION ON THE ALGEBRAIC CLOSURE OF $\mathbb{Q}(X)$

Throughout this section $X = (X_1, \dots, X_N)$ is a tuple of indeterminates, where $N \geq 1$. In this section we construct a height function on the algebraic closure $\mathbb{Q}(X)_{\mathrm{alg}}$ of the rational function field $\mathbb{Q}(X)$, following [20]. We begin by giving $\mathbb{Q}(X)$ the structure of a global field with the finiteness property: given a bound $C \in \mathbb{R}$, there are only finitely many elements of $\mathbb{Q}(X)$ with height at most C. Using Theorem 2.11, we then extend the degree function on $\mathbb{Q}(X)$ to a $\mathrm{Gal}\big(\mathbb{Q}(X)_{\mathrm{alg}}|\mathbb{Q}(X)\big)$-invariant degree function on $\mathbb{Q}(X)_{\mathrm{alg}}$, and establish a few of its properties needed later on.

3.1. **The arithmetic field $\big(\mathbb{Q}(X), M_{\mathbf{arch}}\big)$.** Let M_{arch} be the set of all archimedean absolute values on $\mathbb{Q}(X)$ which extend the usual archimedean absolute value v_∞ on $\mathbb{Q}$. There is a bijective correspondence between M_{arch} and the subset

$$U := \big\{(z_1,\dots,z_N) \in \mathbb{C}^N : \operatorname{trdeg}\big(\mathbb{Q}(z_1,\dots,z_N)|\mathbb{Q}\big) = N\big\}$$

of $\mathbb{C}^N$ as follows: each $z \in U$ defines an archimedean absolute value $v_z \in M_{\mathrm{arch}}$ on $\mathbb{Q}(X)$ by

$$|f/g|_z = |f/g|_{v_z} := |f(z)/g(z)| \qquad \text{for } f, g \in \mathbb{Z}[X],\ g \neq 0,$$

and conversely, by Ostrowski's Theorem, each $v \in M_{\mathrm{arch}}$ is of the form $v = v_z$ for a uniquely determined $z \in U$. Let $\mathcal{C}(U, \mathbb{R})$ be the lattice-ordered group of continuous functions $U \to \mathbb{R}$ (pointwise operations and ordering). The map $z \mapsto \varphi(z) := v_z \colon U \to M_{\mathrm{arch}}$ induces a homomorphism of lattice-ordered groups

$$\begin{aligned}\varphi^* \colon \mathcal{D}\big(\mathbb{Q}(X), M_{\mathrm{arch}}\big) &\to \mathcal{C}(U,\mathbb{R}),\\ &\qquad \varphi^*(D)(z) := \varphi(z)(D) = v_z(D) \text{ for all } z \in U.\end{aligned}$$

The homomorphism φ^* maps the principal divisor (f/g), for non-zero $f, g \in \mathbb{Z}[X]$, to the function

$$z \mapsto -\log|f(z)/g(z)| \colon U \to \mathbb{R}.$$

The arithmetic field $\big(\mathbb{Q}(X), M_{\mathrm{arch}}\big)$ is not classical.

3.2. **Integration of divisors on $\big(\mathbb{Q}(X), M_{\text{arch}}\big)$.** Our next goal is to define a height function on the field $\mathbb{Q}(X)$, as a measure for the complexity of elements of $\mathbb{Q}(X)$. The height of an element of $\mathbb{Q}(X)$ should take into account its values under the various (archimedean and non-archimedean) absolute values on $\mathbb{Q}(X)$. It turns out that when it comes to archimedean absolute values, it suffices to restrict our attention to those absolute values $v_z \in M_{\text{arch}}$ that come from elements $z \in U \cap \mathbb{T}^N$, where $\mathbb{T}^N = (S^1)^N$. We let μ_1 be the (non-atomic, positive) measure on $S^1 = \{z \in \mathbb{C}^N : ||z||_1 = 1\}$ with density function (in polar coordinates)

$$d\mu_1 = \frac{d\theta}{2\pi},$$

and we let $\mu_N := \mu_1 \otimes \cdots \otimes \mu_1$ be the N-fold product measure of μ_1 on $\mathbb{T}^N$. Then μ_N is a non-atomic probability measure on $\mathbb{T}^N$. Let $\mu = \varphi(\mu_N \restriction U \cap \mathbb{T}^N)$ be the image measure of the restriction of μ_N to $U \cap \mathbb{T}^N$ under the map φ, that is,

$$\mu(S) = \mu_N\big(\varphi^{-1}(S) \cap \mathbb{T}^N\big) \quad \text{for all } S \subseteq M_{\text{arch}}. \tag{3.1}$$

Each $D \in \mathcal{D}\big(\mathbb{Q}(X), M_{\text{arch}}\big)$ is μ-measurable and $\varphi^*(D) \restriction \mathbb{T}^N$ is defined μ_N-almost everywhere, since the set

$$\mathbb{T}^N \setminus U = \bigcup_{f \in \mathbb{Z}[X] \setminus \{0\}} \big\{z \in \mathbb{T}^N : f(z) = 0\big\}$$

has measure zero. Therefore,

$$\int_{M_{\text{arch}}} D\, d\mu = \int_{U \cap \mathbb{T}^N} \varphi^*(D) \restriction U \cap \mathbb{T}^N \, d\mu_N = \int_{\mathbb{T}^N} \varphi^*(D) \restriction \mathbb{T}^N \, d\mu_N.$$

For non-zero $f(X) \in \mathbb{Z}[X]$, we have

$$m(f) = \int_{\mathbb{T}^N} \log|f|\, d\mu_N = -\int_{M_{\text{arch}}} (f)\, d\mu, \tag{3.2}$$

where $m(f)$ is the Mahler measure of f. We also have

$$\int_{M_{\text{arch}}} (f)_\infty \, d\mu = \int_{\mathbb{T}^N} \big(-\varphi^*\big((f)\big) \vee 0\big)\, d\mu_N = \int_{\mathbb{T}^N} \log^+|f|\, d\mu_N = m^+(f), \tag{3.3}$$

using the notation m^+ introduced in Section 1.2. In fact:

Lemma 3.1. $\mathcal{D}\big(\mathbb{Q}(X), M_{\text{arch}}\big) \subseteq L^1(\mu)$.

Proof. Since $L^1(\mu)$ is closed under $\wedge$ and $\vee$, it is enough to verify that $\operatorname{div}(f) \in L^1(\mu)$ for all $f(X) \in \mathbb{Z}[X]$, $f \neq 0$. But $|D| = D_0 + D_\infty = D + 2D_\infty$ for all divisors D, so $\int_{M_{\text{arch}}} |\operatorname{div}(f)|\, d\mu = -m(f) + 2m^+(f) \in \mathbb{R}$ by (3.2) and (3.3), and hence $\operatorname{div}(f) \in L^1(\mu)$ as claimed. □

Example 3.2. For $N = 1$ and $f(X) = X$, we get

$$\int_{M_{\text{arch}}} (X)_\infty \, d\mu = \int_{S^1} \log^+|z|\, d\mu_1(z) = 0. \tag{3.4}$$

Therefore, for any monomial $X^\nu = X_1^{\nu_1} \cdots X_N^{\nu_N}$ with $\nu = (\nu_1, \ldots, \nu_N) \in \mathbb{N}^N$,

$$\int_{M_{\text{arch}}} (X^\nu)_\infty \, d\mu = 0. \tag{3.5}$$

(By (3.4), (1.11) and (1.13).)

3.3. **A method to construct height functions.** The following is a general procedure for constructing degree functions on the fraction field K of a unique factorization domain R. Let M_R be the set of essential absolute values of R, as defined in Example 2.1, and suppose that M_∞ is a set of pairwise non-equivalent absolute values on K, each $v \in M_\infty$ non-equivalent to all absolute values in M_R. We write $\operatorname{div}_\infty(a)$ for the principal divisor of $a \in K^\times$ in $\mathcal{D}(K, M_\infty)$. Assume that $\deg_\infty \colon \mathcal{D}(K, M_\infty) \to \mathbb{R}$ is a map satisfying, for each $D, E \in \mathcal{D}(K, M_\infty)$:

$$\deg_\infty(D + E) = \deg_\infty(D) + \deg_\infty(E), \tag{3.6}$$

$$D \geq 0 \Rightarrow \deg_\infty(D) \geq 0, \tag{3.7}$$

$$\deg_\infty\big(\operatorname{div}_\infty(a)\big) \leq 0 \quad \text{for all } 0 \neq a \in R. \tag{3.8}$$

Note that then $\deg_\infty\big(\operatorname{div}_\infty(u)\big) = 0$ for all units u of R. We set $M := M_R \cup M_\infty$ and define a degree function on the arithmetic field (K, M) as follows: for $v = v_p \in M_R$, where $p \in R$ is irreducible, put

$$\lambda_v := -\deg_\infty\big(\operatorname{div}_\infty(p)\big).$$

Clearly this definition does not depend on the choice of p, and $\lambda_v \geq 0$. Let

$$\pi_\infty \colon \mathcal{D}(K, M) \to \mathcal{D}(K, M_\infty), \qquad \pi_\infty(D) = D \restriction M_\infty$$

be the canonical homomorphism, and define $\deg \colon \mathcal{D}(K, M) \to \mathbb{R}$ by

$$\deg(D) := \sum_{v \in M_R} v(D)\lambda_v + \deg_\infty\big(\pi_\infty(D)\big). \tag{3.9}$$

Then $\deg$ is a degree function on (K, M), as one easily verifies. (For (D3), note that this holds trivially if x is a unit or an irreducible element of R, and these elements generate $K^\times$.) Its associated height function h is given by

$$h(x/y) = \deg_\infty\big(\big(-\operatorname{div}_\infty(x)\big) \vee \big(-\operatorname{div}_\infty(y)\big)\big), \tag{3.10}$$

if x and y are relatively prime elements of R, since

$$\begin{aligned} h(x/y) &= \deg\big((x/y)_\infty\big) \\ &= \sum_{v \in M_R} v(y)\lambda_v + \deg_\infty\big(\big(-\operatorname{div}_\infty(x/y)\big) \vee 0\big) \\ &= -\deg_\infty\big(\operatorname{div}_\infty(y)\big) + \deg_\infty\big(\big(-\operatorname{div}_\infty(x/y)\big) \vee 0\big) \\ &= \deg_\infty\big(\big(-\operatorname{div}_\infty(x)\big) \vee \big(-\operatorname{div}_\infty(y)\big)\big). \end{aligned}$$

In particular, for $f \in K^\times$:

$$h(f) \geq \sum_{v \in M_R} \lambda_{v_p} \log^+ |f|_{v_p}. \tag{3.11}$$

For $a_0, \ldots, a_n \in R$, not all zero, the projective height of the point $[a_0 : \cdots : a_n]$ of $\mathbb{P}^n(K)$ is given by

$$\begin{aligned} &h\big([a_0 : \cdots : a_n]\big) = \\ &\qquad -\sum_{v \in M_R} \min_i v(a_i)\lambda_v + \deg_\infty\big(\big(-\operatorname{div}_\infty(a_0)\big) \vee \cdots \vee \big(-\operatorname{div}_\infty(a_d)\big)\big), \end{aligned}$$

hence

$$h\big([a_0 : \cdots : a_n]\big) \leq \deg_\infty\big(\big(-\operatorname{div}_\infty(a_0)\big) \vee \cdots \vee \big(-\operatorname{div}_\infty(a_d)\big)\big),$$

with equality if the $a_0, \ldots, a_n$ have no common factor.

Here are some applications of this construction:

Example 3.3. Let $R = \mathbb{Z}$, $K = \mathbb{Q}$, $M_\infty = \{v_\infty\}$, and

$$\deg_\infty \colon \mathcal{D}(K, M_\infty) \to \mathbb{R}, \quad \deg_\infty(D) := v_\infty(D).$$

Then $\deg_\infty$ satisfies (3.6)–(3.8) from above, so (3.9) defines a degree function on the arithmetic field $\mathbb{Q}$, agreeing with the one defined in Example 2.8.

In the next two examples, we let $R = F[X]$ be the polynomial ring in the indeterminates $X = (X_1, \ldots, X_N)$ over a field F, and $K = F(X)$.

Example 3.4. Put $M_\infty := \{v_X\}$, where $v_X \colon K^\times \to \mathbb{Z}$ is the **total degree valuation** on K, given by

$$v_X(f/g) := \deg_X g - \deg_X f \qquad \text{for } f, g \in R \setminus \{0\}.$$

We let

$$\deg_\infty \colon \mathcal{D}(K, M_\infty) \to \mathbb{R}, \quad \deg_\infty(D) := v_X(D).$$

Then $\deg_\infty$ satisfies the conditions (3.6)–(3.8) above, so (3.9) defines a degree function on (K, M), which we denote by $\deg_X$. The height function h_X corresponding to $\deg_X$ is given by

$$h_X(f/g) = \max\bigl\{\deg_X f, \deg_X g\bigr\},$$

if $f, g \in R \setminus \{0\}$ are relatively prime. By (2.3) and (2.5)–(2.7), the extension of h_X to a $\mathrm{Gal}(K_{\mathrm{alg}}|K)$-invariant height function on K_{alg} is an "$F[X]$-stable degree function on K_{alg}" (as defined in [36]) satisfying $h_X(f) = \deg(f)$ for all non-zero $f \in F[X]$. (The existence of such a "degree function" is shown in [36], Corollary 3.2, by different methods, and used to compute degree bounds in the real Nullstellensatz.)

Example 3.5. Let $M_\infty := \{v_{X_1}, \ldots, v_{X_N}\}$, where $v_{X_i} \colon K^\times \to \mathbb{Z}$ is the X_i**-degree valuation** on K, given by

$$v_{X_i}(f/g) := \deg_{X_i}(g) - \deg_{X_i}(f) \qquad \text{for } f, g \in R \setminus \{0\}.$$

Then (K, M_∞) is a classical arithmetic field, so we can define the function

$$\deg_\infty \colon \mathcal{D}(K, M_\infty) \to \mathbb{R}, \quad \deg_\infty(D) := \sum_{i=1}^{N} v_{X_i}(D).$$

Then $\deg_\infty$ satisfies (3.6)–(3.8), so we have a degree function $\deg_{(X)}$ on (K, M) whose height function $h_{(X)}$ is given by

$$h_{(X)}(f/g) = \sum_{i=1}^{N} \max\bigl\{\deg_{X_i} f, \deg_{X_i} g\bigr\},$$

where $f, g \in R$ are non-zero and relatively prime.

In the previous example, note that $h_{(X)}(f) = \deg_{(X)}(f)$ for all $f \in R \setminus \{0\}$; in particular, the global field $\bigl(\mathbb{Q}(X), M, \deg_{(X)}\bigr)$ does not have the finiteness property. Next we show how, following the general pattern above, one can turn $\mathbb{Q}(X)$ into a global field which does have the finiteness property.

3.4. **Turning $\mathbb{Q}(X)$ into a global field with the finiteness property.** As in the last example, let $R = \mathbb{Z}[X]$, $K = \mathbb{Q}(X)$, but now put

$$M_\infty := M_{\text{arch}} \cup \{v_{X_1}, \dots, v_{X_N}\},$$

where M_{arch} as above is the set of all archimedean absolute values of K which extend the usual archimedean absolute value v_∞ on $\mathbb{Q}$. Let μ be a positive measure on the set M_{arch} satisfying

$$\mathcal{D}(K, M_{\text{arch}}) \subseteq L^1(\mu), \tag{3.12}$$

$$\int_{M_{\text{arch}}} (f)\, d\mu \leq 0 \text{ for all } 0 \neq f \in R. \tag{3.13}$$

The measure $\mu = \varphi(\mu_N \restriction U \cap \mathbb{T}^N)$ defined in (3.1), which is concentrated on the subset $\varphi(U \cap \mathbb{T}^N)$ of M_{arch}, has these properties, by Lemma 3.1, part (2) of Lemma 1.4, and (3.2). (Here, as before, $\varphi(z) = v_z$ for $z \in U$.) Kani in [20] uses instead the measure $\varphi(\mu_{\text{sp},N})$, where $\mu_{\text{sp},N}$ is the N-fold product of the so-called *spherical measure* μ_{sp} on $\mathbb{C}$, given by the density function (in polar coordinates)

$$d\mu_{\text{sp}} = \frac{r\,dr\,d\theta}{\pi(1+r^2)^2}.$$

Then $\mu_{\text{sp},N}$ is a non-atomic probability measure on $\mathbb{C}^N$, and analogues of Lemmas 3.1 and 1.4 hold for $\mu_{\text{sp},N}$ in place of μ_N. Note that $\mu_N = 2^N(\mu_{\text{sp},N} \restriction \mathbb{T}^N)$. We prefer to use μ_N, since this naturally gives a connection to Mahler's measure of transcendental number theory and simplifies some calculations.

For any positive measure μ on M_{arch} satisfying (3.12) and (3.13), we define $\deg_\infty = \deg_{\infty,\mu}$ by

$$\deg_\infty(D) := \sum_{i=1}^N v_{X_i}(D) + \int_{M_{\text{arch}}} D\, d\mu$$

for all $D \in \mathcal{D}(K, M_\infty)$. Any such $\deg_{\infty,\mu}$ satisfies (3.6)–(3.8) from above and hence gives rise to a degree function $\deg_\mu$ on $M = M_R \cup M_\infty$. According to (3.10), its associated height function h_μ is given by

$$h_\mu(f/g) = h_{(X)}(f/g) - \int_{M_{\text{arch}}} (\operatorname{div} f \wedge \operatorname{div} g)\, d\mu,$$

for relatively prime $f, g \in R \setminus \{0\}$. Here, $h_{(X)}$ is as in Example 3.5.

From now on, we fix $\mu = \varphi(\mu_N \restriction U \cap \mathbb{T}^N)$ and write $\deg = \deg_\mu$, $h = h_\mu$. In this case we have, for relatively prime $f, g \in R \setminus \{0\}$:

$$h(f/g) = h_{(X)}(f/g) + \int_{\mathbb{T}^N} \max\big\{\log|f|, \log|g|\big\}\, d\mu_N. \tag{3.14}$$

In particular, for non-zero $f \in R$ we have

$$h(f) = \deg_{(X)} f + \int_{\mathbb{T}^N} \log^+|f(z)|\, d\mu_N(z) = \deg_{(X)} f + m^+(f). \tag{3.15}$$

So for example, using (3.5), we see that

$$h(X^\nu) = \nu_1 + \cdots + \nu_N \qquad \text{for every } \nu = (\nu_1, \dots, \nu_N) \in \mathbb{N}^N.$$

We have the following useful bounds for h on $\mathbb{Q}(X)$, immediate from Lemma 1.5: if $f, g \in \mathbb{Z}[X]$ are non-zero and relatively prime, then

$$h(f/g) \geq (1 - \log 2) \max\big\{h(f), h(g)\big\}, \tag{3.16}$$

$$(1 - \log 2) \deg_{(X)} f + \log ||f||_1 \leq h(f) \leq \deg_{(X)} f + \log ||f||_1, \tag{3.17}$$

$$(1 - \log 2) \deg_{(X)} f + \log ||f||_\infty \leq h(f) \leq 2 \deg_{(X)} f + \log ||f||_\infty. \tag{3.18}$$

These estimates imply that the global field $\big(\mathbb{Q}(X), M, \deg\big)$ has the finiteness property. By Theorem 2.11, $\deg$ extends uniquely to a $\mathrm{Gal}(K_{\mathrm{alg}}|K)$-invariant degree function on $(K_{\mathrm{alg}}, M_{\mathrm{alg}})$, which we also denote by $\deg$. Its corresponding height function, also denoted by h, extends the absolute logarithmic height on $\mathbb{Q}_{\mathrm{alg}}$. Moreover, if $N > 1$, then the height function on $\mathbb{Q}(X_1, \ldots, X_N)_{\mathrm{alg}}$ so defined extends the height function on $\mathbb{Q}(X_1, \ldots, X_{N-1})_{\mathrm{alg}}$ obtained in the same way but with N replaced by $N - 1$. Proposition 2.23 implies that for all real numbers C and d, there are only finitely many $\alpha \in \mathbb{Q}(X)_{\mathrm{alg}}$ with $h(\alpha) \leq C$ and $\big[\mathbb{Q}(X, \alpha) : \mathbb{Q}\big] \leq d$.

3.5. **Height of polynomials with integer coefficients.** For later use, we collect some facts about the behavior of h on $R = \mathbb{Z}[X]$. First a version of the inequality in Lemma 1.4, (6), for the height:

Lemma 3.6. *Let $f(X) \in \mathbb{Z}[X]$ be a polynomial in which the variable X_j occurs, where $j \in \{1, \ldots, N\}$. Then $h(\partial f/\partial X_j) \leq h(f) + \log \deg_{X_j} f$.*

Proof. Let Y be an indeterminate different from $X_1, \ldots, X_N$. Then

$$\begin{aligned} h(\partial f/\partial X_j) &= \deg_{(X)}(\partial f/\partial X_j) + m(\partial f/\partial X_j + Y) \\ &\leq \big(\deg_{(X)}(f) - 1\big) + m\big(f(X) + X_j Y\big) + \log \deg_{X_j}(f) \end{aligned}$$

by Lemma 1.4, (6). Using Jensen's Formula, one sees easily that

$$m\big(f(X) + X_j Y\big) = 1 + m^+\big(f(X)\big),$$

proving the claim. □

By the remarks preceding Example 3.3:

Lemma 3.7. *For $a_0, \ldots, a_n \in \mathbb{Z}[X]$, not all zero, we have*

$$h\big([a_0 : \cdots : a_n]\big) \leq \max_i \deg_{(X)} a_i + \int_{\mathbb{T}^N} \max_i \log |a_i(z)| \, d\mu_N(z),$$

with equality if the $a_0, \ldots, a_n$ have no common factor. □

Example 3.8. Let $s \in \mathbb{N}$, $s > 0$. Then

$$h\big([1 : X_N : \cdots : X_N^{s-1}]\big) = s - 1$$

by the previous lemma and (3.4).

Example 3.9. Let $e \in \mathbb{N}$, $N > 0$. Then

$$h\big(\big[X_1 + X_N^{e^{N-1}} : \cdots : X_{N-1} + X_N^e : X_N : 1\big]\big) \leq e^{N-1} + 1 + \log 2.$$

To see this note that

$$\log^+ \Big| z_i + z_N^{e^{N-i}} \Big| \leq \log 2 \qquad \text{for } (z_1, \ldots, z_N) \in \mathbb{T}^N, \, 1 \leq i \leq N - 1,$$

and

$$\deg_{(X)}\big(X_i + X_N^{e^{N-i}}\big) = e^{N-i} + 1,$$

and apply Lemma 3.7.

Let $f(X,Y) \in \mathbb{Z}[X,Y]$, $Y = (Y_1,\dots,Y_M)$, be non-zero, and write $f(X,Y) = \sum_\nu a_\nu(X)Y^\nu$ with $a_\nu \in \mathbb{Z}[X]$. (Here ν ranges over $\mathbb{N}^M$.) We have two ways of measuring the "height" of f: first, the height $h(f)$ of f as an element of the global field $\mathbb{Q}(X,Y)$ as defined earlier in this section, and second, the height $h_{\mathrm{coeff}}(f)$ of f considered as a polynomial in the variables Y with coefficients $a_\nu \in \mathbb{Z}[X]$. The next proposition compares these two:

Proposition 3.10. *With $d = \deg_Y f$, we have*

$$h_{\mathrm{coeff}}(f) \le h(f) \le h_{\mathrm{coeff}}(f) + d + \log\binom{M+d}{d}.$$

Proof. To see the first inequality, note that for $z \in \mathbb{T}^N$, we have, by (1.15) in Lemma 1.5:

$$\max_\nu \log|a_\nu(z)| = \log||f(z,Y)||_\infty \le m^+\big(f(z,Y)\big) + \deg_{(Y)} f.$$

Integrating over $\mathbb{T}^N$ with respect to z gives

$$\int_{\mathbb{T}^N} \max_\nu \log|a_\nu(z)|\, d\mu_N(z) \le \int_{\mathbb{T}^{N+M}} \log^+|f(z,w)|\, d\mu_{N+M}(z,w) + \deg_{(Y)} f.$$

Since

$$\max_\nu \deg_{(X)} a_\nu \le \deg_{(X)} f,$$

the inequality now follows from Lemma 3.7 and (3.15). The second inequality is immediate from the estimate (2.17) and $h(Y_1,\dots,Y_M) = 1$. □

Remark 3.11. In general $h_{\mathrm{coeff}}(f) < h(f)$, as the example $f(T) = 2 + T \in \mathbb{Z}[T]$ shows. (Here $h(f) = 1 + \log 2 > \log 2 = h_{\mathrm{coeff}}(f)$.)

If the indeterminate X_j (where $j \in \{1,\dots,N\}$) occurs in the polynomial f, then by Lemma 3.6 and Proposition 3.10:

$$h_{\mathrm{coeff}}(\partial f/\partial X_j) \le h_{\mathrm{coeff}}(f) + d + \log(d+1) + \log\deg_{X_j}(f). \tag{3.19}$$

Corollary 2.27 and Proposition 3.10 yield:

Corollary 3.12. *Let $f, g \in \mathbb{Z}[X,T]$ be non-zero of degrees $m = \deg_T f$ and $n = \deg_T g$ in the indeterminate T, respectively. Let $r = \mathrm{res}_T(f,g) \in \mathbb{Z}[X]$ be the resultant of f and g with respect to the indeterminate T. Then*

$$h(r) \le n\,h(f) + m\,h(g) + mn\log 4.$$

□

4. Restricted power series

Let K be a field of characteristic zero and let $|\cdot| = |\cdot|_v$ be an ultrametric absolute value on K. We assume that K is complete with respect to the metric (1.1) induced by $|\cdot|$. In the following we write $\mathcal{O} = \mathcal{O}_v$ and $\mathfrak{m} = \mathfrak{m}_v$ for the valuation ring of v and its maximal ideal, respectively, as defined in (1.3). The residue field of $\mathcal{O}$ is denoted by $\overline{\mathcal{O}} = \mathcal{O}/\mathfrak{m}$, with residue homomorphism $a \mapsto \overline{a}\colon \mathcal{O} \to \overline{\mathcal{O}}$. The subset

$$K\langle X\rangle := \left\{\sum_\nu a_\nu X^\nu \in K[[X]] : a_\nu \in K,\ |a_\nu| \to 0 \text{ as } |\nu| \to \infty\right\}$$

of the ring $K[[X]] = K[[X_1,\dots,X_N]]$ of all formal power series with coefficients in K is a subring of $K[[X]]$, called the ring of **restricted power series** with coefficients in K. Here, as earlier, $\nu = (\nu_1,\dots,\nu_N)$ ranges over all multiindices in $\mathbb{N}^N$, and $|\nu| =$

$\nu_1 + \cdots + \nu_N$. The Gauß norm on $K[X]$ extends to an ultrametric absolute value on the domain $K\langle X\rangle$ (called the **Gauß norm** on $K\langle X\rangle$) by setting

$$|f| := \max_{\nu} |a_\nu| \qquad \text{for } f = \sum_{\nu} a_\nu X^\nu \in K\langle X\rangle,\ f \neq 0.$$

(See [10], p. 44, Corollary 2.) The set of all $f \in K\langle X\rangle$ with $|f| \leq 1$ forms a subring $\mathcal{O}\langle X\rangle$ of $K\langle X\rangle$ (the ring of restricted power series with coefficients in $\mathcal{O}$). We identify $\mathcal{O}\langle X\rangle/\mathfrak{m}\mathcal{O}\langle X\rangle$ with $\overline{\mathcal{O}}[X]$ in the natural way, and we denote the image of $f \in \mathcal{O}\langle X\rangle$ under the canonical surjection $\mathcal{O}\langle X\rangle \to \mathcal{O}\langle X\rangle/\mathfrak{m}\mathcal{O}\langle X\rangle = \overline{\mathcal{O}}[X]$ by $\overline{f}$.

Suppose from now on that $N \geq 1$, and let $X' := (X_1, \ldots, X_{N-1})$. We have $K\langle X'\rangle \subseteq K\langle X\rangle$ and $\mathcal{O}\langle X'\rangle \subseteq \mathcal{O}\langle X\rangle$ in a natural way. Every element $f \in \mathcal{O}\langle X\rangle$ can be written uniquely in the form

$$f = \sum_{i=0}^{\infty} f_i X_N^i \quad \text{with } f_i(X') \in \mathcal{O}\langle X'\rangle \text{ for all } i \in \mathbb{N}, \tag{4.1}$$

where the infinite sum converges with respect to the Gauß norm on $K\langle X\rangle$. An element f of $\mathcal{O}\langle X\rangle$, written as in (4.1), is called **regular in X_N of degree** $s \in \mathbb{N}$ if its reduction $\overline{f} \in \overline{\mathcal{O}}[X]$ is a unit-monic polynomial of degree s in X_N, that is,

(1) $\overline{f_s}$ is a unit in $\overline{\mathcal{O}}[X']$, and
(2) $\overline{f_i} = 0$ for all $i > s$.

If $f \in \mathcal{O}\langle X'\rangle[X_N]$ is monic of X_N-degree s (so that in particular f is regular in X_N of degree s, as an element of $\mathcal{O}\langle X\rangle$), then f is called a **Weierstraß polynomial (in X_N) of degree s.** For proofs of the following standard facts see, e.g., [10].

Theorem 4.1. (Weierstraß Division Theorem for $\mathcal{O}\langle X\rangle$.) *Let $g \in \mathcal{O}\langle X\rangle$ be regular in X_N of degree s. Then for each $f \in \mathcal{O}\langle X\rangle$ there are uniquely determined elements $q \in \mathcal{O}\langle X\rangle$ and $r \in \mathcal{O}\langle X'\rangle[X_N]$ with $\deg_{X_N} r < s$ such that $f = qg + r$.*

Applying Weierstraß Division with $f = X_N^s$, we obtain:

Corollary 4.2. (Weierstraß Preparation Theorem for $\mathcal{O}\langle X\rangle$.) *Let $g \in \mathcal{O}\langle X\rangle$ be regular in X_N of degree s. There are a unique Weierstraß polynomial $w \in \mathcal{O}\langle X'\rangle[X_N]$ of degree s and a unique unit $u \in \mathcal{O}\langle X\rangle$ such that $g = u \cdot w$.*

Regularity can achieved by a change of variables:

Lemma 4.3. (Noether Normalization.) *Let $e > 1$ and $f \in \mathcal{O}\langle X\rangle$, and suppose that $\overline{f} \in \overline{\mathcal{O}}[X]$ is non-zero of degree $< e$. Let $T_e\colon \mathcal{O}\langle X\rangle \to \mathcal{O}\langle X\rangle$ be the $\mathcal{O}$-automorphism defined by*

$$\begin{aligned} X_i &\mapsto X_i + X_N^{e^{N-i}} \qquad (\textit{for } 1 \leq i < N)\\ X_N &\mapsto X_N. \end{aligned}$$

Then $T_e(f)$ is regular in X_N of degree $< e^N$.

A non-zero element $f \in K\langle X\rangle$ is called **regular in X_N of degree** $s \in \mathbb{N}$ if there exists $b \in \mathcal{O}$ such that $bf \in \mathcal{O}\langle X\rangle$ and bf is regular in X_N of degree s (as defined above). From Theorem 4.1 and Corollary 4.2 we get:

Corollary 4.4. (Weierstraß Division and Preparation Theorems for $K\langle X\rangle$.) *Let $g \in K\langle X\rangle$ be regular in X_N of degree s. Then every $f \in K\langle X\rangle$ can be uniquely written as $f = qg + r$ with $q \in K\langle X\rangle$ and $r \in K\langle X'\rangle[X_N]$, $\deg_{X_N} r < s$. In particular, there are a unique*

Weierstraß polynomial $w \in \mathcal{O}\langle X'\rangle[X_N]$ of degree s and a unique unit $u \in K\langle X\rangle$ such that $g = u \cdot w$.

The rings $K\langle X\rangle$ and $\mathcal{O}\langle X\rangle$ are local, with maximal ideal $(X_1, \dots, X_N)K\langle X\rangle$ and $(\mathfrak{m}, X_1, \dots, X_N)\mathcal{O}\langle X\rangle$, respectively. Corollary 4.4 and Lemma 4.3 imply that $K\langle X\rangle$ is noetherian. (If $\mathcal{O}$ is a discrete valuation ring, then $\mathcal{O}\langle X\rangle$ is also noetherian.)

In the rest of this section we fix a subring D of $\mathcal{O}$, with fraction field F (a subfield of K).

Notation. For elements $\alpha_1, \dots, \alpha_n$ of an algebraic closure $F(X)_{\mathrm{alg}}$ of $F(X)$, we write

$$\Delta_X(\alpha_1, \dots, \alpha_n) := \big[F(X, \alpha_1, \dots, \alpha_n) : F(X)\big].$$

If X is understood, then we also write Δ for Δ_X, and we abbreviate $\Delta_{X'}$ to Δ'.

We have the following simple rules, for all $\alpha, \beta, \alpha_1, \dots, \alpha_n \in F(X)_{\mathrm{alg}}$:

(1) $\Delta(\alpha_1), \dots, \Delta(\alpha_n) \le \Delta(\alpha_1, \dots, \alpha_n)$;
(2) $\Delta(\alpha_1, \dots, \alpha_n) \le \Delta(\alpha_1) \cdots \Delta(\alpha_n)$, with equality if and only if the fields
$$F(X, \alpha_1), \dots, F(X, \alpha_n)$$
are pairwise linearly disjoint over $F(X)$, in particular, if the $\Delta(\alpha_i)$ are pairwise relatively prime;
(3) if $\alpha_1, \dots, \alpha_n$ are zeros of a common polynomial $P(X', T) \in F(X)[T]$ of degree $d > 0$, then $\Delta(\alpha_1, \dots, \alpha_n)$ is bounded from above by the degree of the splitting field of P over $F(X')$; so $\Delta(\alpha_1, \dots, \alpha_n)$ divides $d!$;
(4) $\Delta(\alpha + \beta, \alpha_1, \dots, \alpha_n), \Delta(\alpha \cdot \beta, \alpha_1, \dots, \alpha_n) \le \Delta(\alpha, \beta, \alpha_1, \dots, \alpha_n)$;
(5) $\Delta(\alpha^{-1}, \alpha_1, \dots, \alpha_n) = \Delta(\alpha, \alpha_1, \dots, \alpha_n)$ if $\alpha \neq 0$;
(6) $\Delta(\lambda\alpha, \alpha_1, \dots, \alpha_n) = \Delta(\alpha, \alpha_1, \dots, \alpha_n)$ for all $\lambda \in F(X)$, $\lambda \neq 0$;
(7) if $\alpha_1, \dots, \alpha_n \in F(X')_{\mathrm{alg}}$, then $\Delta(\alpha_1, \dots, \alpha_n) = \Delta'(\alpha_1, \dots, \alpha_n)$.

Let $\alpha \in F(X)_{\mathrm{alg}}$. Then there exists a non-zero polynomial $P(T) \in F[X, T]$ such that $P(\alpha) = 0$. Here, T is a new indeterminate, distinct from each of $X_1, \dots, X_N$. Among the non-zero polynomials $P(T) = \sum_{i=0}^{d} a_i T^i$ with coefficients $a_0, \dots, a_d \in F[X]$, $a_d \neq 0$, such that $P(\alpha) = 0$, there exists one of lowest possible degree d in T whose non-zero coefficients a_i have no common non-trivial factor in $F[X]$. The polynomial P is irreducible in $F[X, T]$ and uniquely determined, up to multiplication by a non-zero element of F. We say that such a polynomial P is a **minimal polynomial for α.**

Lemma 4.5. *Suppose that $F = \mathbb{Q}$ (or more generally, a Hilbertian field), and let*

$$\alpha = \alpha_0 + \alpha_1 X_N + \cdots + \alpha_s X_N^s \quad \textit{with } \alpha_0, \dots, \alpha_s \in F(X')_{\mathrm{alg}},\ s \in \mathbb{N}.$$

Then $\Delta(\alpha) = \Delta'(\alpha_0, \dots, \alpha_s)$.

Proof. Let $P(X, T) \in \mathbb{Q}[X, T]$ be a minimal polynomial for α, and let $d := \deg_T P = \Delta(\alpha)$. By Hilbert's Irreducibility Theorem (see [23], Chapter XIX) there exist infinitely many elements x_N of $\mathbb{Q}$ such that the polynomial $Q(X', T) := P(X', x_N, T) \in \mathbb{Q}[X', T]$ is irreducible of degree d in T. Then $Q\big(X', \alpha(x_N)\big) = 0$, so $d = \Delta'\big(\alpha(x_N)\big)$. The standard proof of the Primitive Element Theorem (as given in [24], say) shows that all but finitely many x_N have the additional property that $\alpha(x_N)$ is a primitive element of the finite field extension $\mathbb{Q}(X', \alpha_0, \dots, \alpha_s) \supseteq \mathbb{Q}(X')$, so $\Delta'(\alpha_0, \dots, \alpha_s) = \Delta'\big(\alpha(x_N)\big)$. The lemma follows. □

We now turn our attention to the power series in $K\langle X\rangle$ which are algebraic over $F(X)$. The set $K\langle X\rangle_{\mathrm{alg}}$ of such power series forms a subring of $K\langle X\rangle$, the **ring of algebraic restricted power series** with coefficients in K. (Note the potential notational confusion: for $N = 0$, the ring $K\langle X\rangle_{\mathrm{alg}}$ denotes $F_{\mathrm{alg}} \cap K$, and not an algebraic closure K_{alg} of K.) We put

$$\mathcal{O}\langle X\rangle_{\mathrm{alg}} := K\langle X\rangle_{\mathrm{alg}} \cap \mathcal{O}\langle X\rangle.$$

Note that $\mathcal{O}\langle X\rangle_{\mathrm{alg}} = \mathrm{Frac}\big(\mathcal{O}\langle X\rangle_{\mathrm{alg}}\big) \cap \mathcal{O}\langle X\rangle$. The $\mathcal{O}$-automorphism T_e of $\mathcal{O}\langle X\rangle$ in Lemma 4.3 extends uniquely to a K-automorphism of $K\langle X\rangle$. This automorphism of $K\langle X\rangle$, which we again denote by T_e, maps $F[X]$ into itself, so by restriction we get an F-automorphism of $K\langle X\rangle_{\mathrm{alg}}$ (which in turn restricts to a D-automorphism of $\mathcal{O}\langle X\rangle_{\mathrm{alg}}$).

In view of the applications in the later sections, we now make the following simplifying assumption: *for every $a \in \mathcal{O}$ there exists $b \in \mathcal{O} \cap F_{\mathrm{alg}}$ with $|a| = |b|$.* (So for every $f \in K\langle X\rangle_{\mathrm{alg}}$ which is regular in X_N of degree $s \in \mathbb{N}$ there exists $b \in \mathcal{O} \cap F_{\mathrm{alg}}$ such that $bf \in \mathcal{O}\langle X\rangle_{\mathrm{alg}}$ is regular in X_N of degree s.) We want to prove the following version of Weierstraß Division for algebraic restricted power series:

Theorem 4.6. (Weierstraß Division Theorem for $K\langle X\rangle_{\mathrm{alg}}$.) *In Corollary* 4.4*, if in addition f and g are in $K\langle X\rangle_{\mathrm{alg}}$, then $q \in K\langle X\rangle_{\mathrm{alg}}$ and $r \in K\langle X'\rangle_{\mathrm{alg}}[X_N]$. If moreover we write*

$$r = r_0 + r_1X_N + \cdots + r_{s-1}X_N^{s-1} \qquad \text{with } r_0, \ldots, r_{s-1} \in K\langle X'\rangle_{\mathrm{alg}},$$

and $Q(T) \in F[X, T]$ is a minimal polynomial for g, then

$$\Delta'(r_0, \ldots, r_{s-1}) \leq \big(\Delta(f)\deg_{X_N} Q(0)\big)^s.$$

We deduce Theorem 4.6 from the usual Weierstraß Division Theorem for $K\langle X\rangle$ using an idea from [14], §3. Before we give the proof, we collect a few useful results about substitution, along the lines of [9], §8.2 (which deals with the case of algebraic formal power series).

4.1. **Substitution into restricted power series.** Let $g \in K\langle X\rangle$ be regular in X_N of degree s, and write $g = uw$ with $u \in K\langle X\rangle$ a unit and $w \in \mathcal{O}\langle X'\rangle[X_N]$ a Weierstraß polynomial of degree s. We denote by Ω an algebraic closure of $\mathrm{Frac}\big(K\langle X'\rangle\big)$, and we fix a zero $\xi \in \Omega$ of w. (We will sometimes just say that ξ is a **zero of** g.) For $f \in K\langle X\rangle$, let $r_f = r_f(X_N) \in K\langle X'\rangle[X_N]$ be the remainder of f obtained through Weierstraß Division by g. We put

$$f(X', \xi) := r_f(X', \xi) \in \Omega.$$

For $f \in K\langle X'\rangle[X_N]$, Euclidean division of f by the monic polynomial w in $K\langle X'\rangle[X_N]$ yields that $r_f \in K\langle X'\rangle[X_N]$; so in this case the element $f(X', \xi)$ of Ω is indeed just obtained by substitution of ξ for X_N into the polynomial f. For notational simplicity, from now on we just write $f(\xi)$ for $f(X', \xi)$. The map $f \mapsto f(\xi)\colon K\langle X\rangle \to \Omega$ clearly is $K\langle X'\rangle$-linear. In fact, we have:

Lemma 4.7. *The map $f \mapsto f(\xi)\colon K\langle X\rangle \to \Omega$ is a $K\langle X'\rangle$-algebra homomorphism. If there exists a non-zero $f \in K\langle X\rangle_{\mathrm{alg}}$ with $f(\xi) = 0$, then ξ is algebraic over $F(X')$. Conversely, if ξ is algebraic over $F(X')$, then for all $f \in K\langle X\rangle_{\mathrm{alg}}$, the element $f(\xi)$ of Ω is algebraic over $F(X', \xi)$ and hence over $F(X')$.*

Proof. We have to show that

$$(f_1f_2)(\xi) = f_1(\xi)f_2(\xi) \qquad \text{for all } f_1, f_2 \in K\langle X\rangle.$$

Let $q_1, q_2 \in K\langle X\rangle$ be the quotients obtained by Weierstraß Division of f_1, f_2, respectively, by g; thus $f_i = q_ig + r_{f_i}$ for $i = 1, 2$. Let also $q \in K\langle X\rangle$ be the quotient obtained by

dividing f_1f_2 by g, so $f_1f_2 = qg + r_{f_1f_2}$. We also have $f_1f_2 = (q_1g + r_{f_1})(q_2g + r_{f_2})$, and subtracting these two representations of f_1f_2 one sees that $r_{f_1f_2} - r_{f_1}r_{f_2} \in K\langle X'\rangle[X_N]$ is a multiple of g in $K\langle X\rangle$, hence

$$(f_1f_2)(\xi) - f_1(\xi)f_2(\xi) = r_{f_1f_2}(\xi) - r_{f_1}(\xi)r_{f_2}(\xi) = (r_{f_1f_2} - r_{f_1}r_{f_2})(\xi) = 0$$

as required.

Now let $f \in K\langle X\rangle_{\mathrm{alg}}$ be non-zero, and let $P(T) = \sum_{i=0}^{d} a_i T^i \in F[X, T]$, with $a_0, \dots, a_d \in F[X]$, $a_d \neq 0$, be a minimal polynomial for f. Note that $a_0 \neq 0$. Since $P(f) = 0$ and $f \mapsto f(\xi)$ is a $K\langle X'\rangle$-algebra homomorphism, we have

$$0 = \big(P(f)\big)(\xi) = \sum_{i=0}^{d} a_i(\xi)\big(f(\xi)\big)^i = \sum_{i=0}^{d} a_i(X', \xi)\big(f(\xi)\big)^i. \tag{4.2}$$

Suppose $f(\xi) = 0$; then (4.2) implies $a_0(X', \xi) = 0$, which shows that ξ is algebraic over $F(X')$. Assume conversely that ξ is algebraic over $F(X')$. Not all the $a_i(X', \xi)$ vanish: otherwise, ξ would be algebraic over $F(X')$ and all $a_i(X', X_N)$ would be $F[X]$-multiples of the minimal polynomial of ξ over $F(X')$, in contradiction to $a_0, \dots, a_d$ having no common factor in $F[X]$. So by (4.2), $f(\xi)$ is algebraic over $F\big(a_0(X', \xi), \dots, a_d(X', \xi), X'\big)$, and hence $f(\xi)$ is algebraic over $F(X', \xi)$. □

In particular, if $g \in K\langle X\rangle_{\mathrm{alg}}$, then we have $\xi \in F(X')_{\mathrm{alg}}$, since $g(\xi) = 0$. Inspection of the proof of Lemma 4.7 also gives the following degree bounds, with $0 \neq f \in K\langle X\rangle_{\mathrm{alg}}$ and $P(T) \in F[X, T]$ a minimal polynomial for f as in that proof, and $\xi_0, \dots, \xi_{s-1}$ the zeros of g in Ω.

(1) If $f(\xi_0) = \cdots = f(\xi_{s-1}) = 0$, then

$$\Delta'(\xi_0, \dots, \xi_{s-1}) \leq \big(\deg_{X_N} P(0)\big)^s. \tag{4.3}$$

(2) If $\xi_0, \dots, \xi_{s-1}$ are algebraic over $F(X')$, then

$$\big[F\big(X', \xi_0, \dots, \xi_{s-1}, f(\xi_0), \dots, f(\xi_{s-1})\big) : F(X', \xi_0, \dots, \xi_{s-1})\big] \leq \Delta(f)^s. \tag{4.4}$$

Moreover:

Lemma 4.8. *Let $f_1, \dots, f_m \in K\langle X\rangle_{\mathrm{alg}}$, $m \geq 1$. There exists $\alpha \in K\langle X\rangle_{\mathrm{alg}}$ such that $\Delta(\alpha) \leq \Delta(f_1, \dots, f_m)$ and*

$$F\big(X', f_1(\xi), \dots, f_m(\xi)\big) = F\big(X', \alpha(\xi)\big)$$

for every zero $\xi \in F(X')_{\mathrm{alg}}$ of g.

Proof. Put $\alpha = f_1 + c_2f_2 + \cdots + c_mf_m \in K\langle X\rangle_{\mathrm{alg}}$ with as yet undetermined $c_2, \dots, c_m \in F$. Then $\Delta(\alpha) \leq \Delta(f_1, \dots, f_m)$. Moreover, $\alpha(\xi)$ is a primitive element of the field extension $F\big(X', f_1(\xi), \dots, f_m(\xi)\big) \supseteq F(X')$ for all but finitely many $(c_2, \dots, c_m) \in F^{m-1}$. Since g only has finitely (namely, at most s) many zeros, we can choose $(c_2, \dots, c_m) \in F^{m-1}$ such that $\alpha(\xi)$ has the required property for every zero ξ of g. □

In the next lemma, we let $Y_1, \dots, Y_M$ be a new set of pairwise distinct indeterminates, each Y_i distinct from $X_1, \dots, X_N$, and $Y := (Y_1, \dots, Y_M)$. For $f = \sum_\nu a_\nu X^\nu \in K\langle X\rangle$ and $g_1, \dots, g_N \in \mathcal{O}\langle Y\rangle$, the infinite sum

$$\sum_\nu a_\nu g_1^{\nu_1} \cdots g_N^{\nu_N}$$

converges to an element $f(g_1, \dots, g_N)$ in $K\langle Y\rangle$. With these notations and remarks:

Lemma 4.9. *Let $f \in K\langle X\rangle_{\mathrm{alg}}$ and $g_1, \dots, g_N \in \mathcal{O}\langle Y\rangle_{\mathrm{alg}}$. Then $f(g_1, \dots, g_N) \in K\langle Y\rangle_{\mathrm{alg}}$ with $\Delta_Y\big(f(g_1, \dots, g_N)\big) \leq \Delta(f)\Delta_Y(g_1) \cdots \Delta_Y(g_N)$.*

Proof. We prove the following slightly more general statement, by induction on N: if $f \in K\langle X, Y\rangle_{\mathrm{alg}}$ and $g_1, \dots, g_N \in \mathcal{O}\langle Y\rangle_{\mathrm{alg}}$, then $f(g_1, \dots, g_N, Y) \in K\langle Y\rangle_{\mathrm{alg}}$ with $\Delta_Y\big(f(g_1, \dots, g_N, Y)\big) \leq \Delta_{(X,Y)}(f)\Delta_Y(g_1) \cdots \Delta_Y(g_N)$. For $N = 0$, this is trivial. Let $N > 0$ and assume the corresponding statement is true for $N - 1$ instead of N. Let $f \in K\langle X, Y\rangle_{\mathrm{alg}}$ and $g_1, \dots, g_N \in \mathcal{O}\langle Y\rangle_{\mathrm{alg}}$. We set $p(X_N) := X_N - g_N \in \mathcal{O}\langle X', Y\rangle[X_N]$, a Weierstraß polynomial of degree 1, having the unique zero g_N. By Weierstraß Division in $K\langle X, Y\rangle$ we obtain $q \in K\langle X, Y\rangle$ and $r \in K\langle X', Y\rangle$ with $f = qp + r$, thus $f(X', g_N, Y) = r \in K\langle X', Y\rangle$. By Lemma 4.7 and the remarks following it we get $f(X', g_N, Y) = r \in K\langle X', Y\rangle_{\mathrm{alg}}$ with $\Delta_{(X',Y)}\big(f(X', g_N, Y)\big) \leq \Delta_{(X,Y)}(f)\Delta_Y(g_N)$. Now apply the inductive hypothesis to $f(X', g_N, Y)$. □

Lemma 4.10. *If $f \in K\langle X\rangle_{\mathrm{alg}}$, then $\partial f/\partial X_j \in F(X, f)$ for $j = 1, \dots, N$. (So in particular, $\partial f/\partial X_j \in K\langle X\rangle_{\mathrm{alg}}$, with $\Delta(\partial f/\partial X_j) \leq \Delta(f)$.)*

Proof. Let $f \in K\langle X\rangle_{\mathrm{alg}}$ and $j \in \{1, \dots, N\}$. Let $P(T) = \sum_{i=0}^{d} a_i T^i \in F[X, T]$, $a_0, \dots, a_d \in F[X]$, be a minimal polynomial for f. We then have

$$0 = \frac{\partial P(f)}{\partial X_j} = \sum_{i=0}^{d} \frac{\partial a_i}{\partial X_j} f^i + \frac{\partial f}{\partial X_j} \sum_{i=1}^{d} i a_i f^{i-1}$$

Then $\sum_{i=1}^{d} i a_i f^{i-1} \neq 0$ (by virtue of $\operatorname{char} F = 0$), and since $\partial a_i/\partial X_j \in F[X]$ for $i = 0, \dots, d$, this shows that $\partial f/\partial X_j \in F(X, f)$. □

Corollary 4.11. *Let $f \in K\langle X\rangle_{\mathrm{alg}}$ and let $\xi \in F(X')_{\mathrm{alg}}$ be a zero of g. Then*

$$\big(\partial^i f/\partial X_N^i\big)(\xi) \in F\big(X', \xi, f(\xi)\big) \qquad \text{for all } i \in \mathbb{N}.$$

Proof. By Lemma 4.10 and induction, it is enough to consider the case $i = 1$. By Lemma 4.7, $f(\xi)$ is algebraic over $F(X', \xi)$. Let $P(X, T) = \sum_{i=0}^{d} a_i(X) T^i \in F[X, T]$, $a_0, \dots, a_d \in F[X]$, $a_d(X', \xi) \neq 0$, be a polynomial of minimal degree in T such that $P\big(X', \xi, f(\xi)\big) = 0$. We have

$$\frac{\partial P(X, f)}{\partial X_N} = \sum_{i=0}^{d} \frac{\partial a_i}{\partial X_N} f^i + \frac{\partial f}{\partial X_N} \sum_{i=1}^{d} i a_i f^{i-1},$$

hence after applying the F-algebra homomorphism "substitution of ξ for X_N":

$$\left(\frac{\partial P(X, f)}{\partial X_N}\right)(\xi) = \sum_{i=0}^{d} \left(\frac{\partial a_i}{\partial X_N}\right)(X', \xi) f(\xi)^i + \left(\frac{\partial f}{\partial X_N}\right)(\xi) \sum_{i=1}^{d} i a_i(X', \xi) f(\xi)^{i-1}.$$

Now $\varrho := \sum_{i=1}^{d} i a_i(X', \xi) f(\xi)^{i-1}$ is non-zero by choice of P, and hence

$$\left(\frac{\partial f}{\partial X_N}\right)(\xi) = \frac{1}{\varrho}\left(\frac{\partial P(X, f)}{\partial X_N} - \sum_{i=0}^{d} \left(\frac{\partial a_i}{\partial X_N}\right) f^i\right)(\xi) \in F\big(X', \xi, f(\xi)\big)$$

as required. □

Corollary 4.12. *If $f = \sum_{i=0}^{\infty} f_i X_N^i \in K\langle X\rangle_{\mathrm{alg}}$ with $f_i \in K\langle X'\rangle$ for all $i \in \mathbb{N}$, then $f_i \in K\langle X'\rangle_{\mathrm{alg}}$ with $\Delta'(f_i) \leq \Delta(f)$ for all $i \in \mathbb{N}$.*

Proof. This follows from $f_i = \frac{1}{i!}\frac{\partial^i f}{\partial X_N^i}(X', 0)$ and Lemmas 4.9 and 4.10. (If f is a polynomial in X_N and $F = \mathbb{Q}$, then degree estimate can be strengthened, see Lemma 4.5.) $\square$

The **multiplicity** of a zero $\xi \in \Omega$ of g is its multiplicity as a zero of the polynomial w. Equivalently, $\xi \in \Omega$ is a zero of g of multiplicity n if and only if

$$g(\xi) = \left(\frac{\partial g}{\partial X_N}\right)(\xi) = \cdots = \left(\frac{\partial^{n-1} g}{\partial X_N^{n-1}}\right)(\xi) = 0,\ \left(\frac{\partial^n g}{\partial X_N^n}\right)(\xi) \neq 0.$$

4.2. **Proof of the Weierstraß Division Theorem for $K\langle X\rangle_{\mathrm{alg}}$.** Let f and g be as in Theorem 4.6. By Weierstraß Division in $K\langle X\rangle$,

$$f = qg + r \quad \text{with } q \in K\langle X\rangle,\ r \in K\langle X'\rangle[X_N],\ \deg_{X_N} r < s. \tag{4.5}$$

Write $r = r(X_N) = \sum_{j=0}^{s-1} r_j X_N^j$ with $r_0, \ldots, r_{s-1} \in K\langle X'\rangle$. For the first statement in Theorem 4.6, it suffices to show $r \in K\langle X'\rangle_{\mathrm{alg}}[X_N]$, since then

$$q = (f - r)/g \in \mathrm{Frac}\big(K\langle X\rangle_{\mathrm{alg}}\big) \cap K\langle X\rangle = K\langle X\rangle_{\mathrm{alg}}.$$

Let now $Q(T) \in F[X, T]$ be a minimal polynomial for g. We need to prove that

$$r_0, \ldots, r_{s-1} \in F(X')_{\mathrm{alg}}$$

and

$$\Delta'(r_0, \ldots, r_{s-1}) \leq \big(\Delta(f)\deg_{X_N} Q(0)\big)^s.$$

To see this, let $\xi_0, \ldots, \xi_{s-1}$ be the zeros of g in $F(X')_{\mathrm{alg}}$. Using Lemma 4.7 and (4.5), we get

$$f(\xi_i) = r(\xi_i) = \sum_{j=0}^{s-1} r_j \xi_i^j \qquad \text{for } i = 0, \ldots, s-1.$$

We have $f(\xi_i) \in F(X')_{\mathrm{alg}}$ by Lemma 4.7, since $f \in K\langle X\rangle_{\mathrm{alg}}$ and $\xi_i \in F(X')_{\mathrm{alg}}$. Hence by (4.3), (4.4) it suffices to show that

$$r_0, \ldots, r_{s-1} \in F\big(f(\xi_0), \ldots, f(\xi_{s-1}), \xi_0, \ldots, \xi_{s-1}\big). \tag{4.6}$$

If $\xi_0, \ldots, \xi_{s-1}$ are *distinct,* then $r_0, \ldots, r_{s-1}$ are uniquely determined by the system of linear equations:

$$\begin{bmatrix} 1 & \xi_0 & \ldots & \xi_0^{s-1} \\ 1 & \xi_1 & \ldots & \xi_1^{s-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \xi_{s-1} & \ldots & \xi_{s-1}^{s-1} \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \\ \vdots \\ r_{s-1} \end{bmatrix} = \begin{bmatrix} f(\xi_0) \\ f(\xi_1) \\ \vdots \\ f(\xi_{s-1}) \end{bmatrix} \tag{4.7}$$

(The Vandermonde matrix in (4.7) is non-singular.) So (4.6) follows by Cramer's Rule.

Consider now the case that the zero ξ_0 of g has multiplicity 2, and $\xi_1, \ldots, \xi_{s-2}$ each has multiplicity 1. Then we may use the identity

$$\frac{\partial f}{\partial X_N} = q \cdot \frac{\partial g}{\partial X_N} + \frac{\partial q}{\partial X_N} \cdot g + \big(r_1 + 2r_2 X_N + \cdots + (s-1) r_{s-1} X_N^{s-2}\big)$$

to get the following system of linear equations:

$$\begin{bmatrix} 1 & \xi_0 & \xi_0^2 & \ldots & \xi_0^{s-1} \\ 0 & 1 & 2\xi_0 & \ldots & (s-1)\xi_0^{s-2} \\ 1 & \xi_1 & \xi_1^2 & \ldots & \xi_1^{s-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \xi_{s-2} & \xi_{s-2}^2 & \ldots & \xi_{s-2}^{s-1} \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \\ r_2 \\ \vdots \\ r_{s-1} \end{bmatrix} = \begin{bmatrix} f(\xi_0) \\ (\partial f/\partial X_N)(\xi_0) \\ f(\xi_1) \\ \vdots \\ f(\xi_{s-2}) \end{bmatrix} \tag{4.8}$$

The matrix here is again non-singular. (The homogeneous system with this matrix has only the trivial solution.) By Corollary 4.11 we again get (4.6).

The other (finitely many) possible configurations of the multiplicities of the zeros of g are treated similarly. They are in one-to-one correspondence with the finite sequences $\mu = (\mu_0, \ldots, \mu_n) \in \mathbb{N}^{n+1}$, where $\mu_0 \geq \mu_1 \geq \cdots \geq \mu_n \geq 1$ and $\mu_0 + \cdots + \mu_n = s$. For each such μ, suppose that each ξ_i has multiplicity μ_i, and let $\xi = (\xi_0, \ldots, \xi_n)$. Then $[r_0, \ldots, r_{s-1}]^{\mathrm{tr}}$ is the unique solution to a system of linear equations with the invertible matrix $A(\xi, \mu, s)$ as its coefficient matrix; here $A(\xi, \mu, s)$ is as introduced in Section 2.5. (For example, the matrices in (4.7) and (4.8) are $A(\xi, (1, 1, \ldots, 1), s)$ and $A(\xi, (2, 1, \ldots, 1), s)$, respectively.) □

The following strengthening of Theorem 4.6 will become useful:

Corollary 4.13. (Simultaneous Weierstraß Division.) *Suppose that the restricted power series g is regular in X_N of degree s. Let $m \geq 1$, and for every $i = 1, \ldots, m$ let $f_i, q_i \in K\langle X\rangle_{\mathrm{alg}}$ and $r_{i0}, \ldots, r_{i,s-1} \in K\langle X'\rangle_{\mathrm{alg}}$ such that*

$$f_i = q_i g + r_i \qquad \text{with } r_i = r_{i0} + r_{i1}X_N + \cdots + r_{i,s-1}X_N^{s-1} \in K\langle X\rangle_{\mathrm{alg}}.$$

Then

$$\Delta'(r_{ij} : 1 \leq i \leq m,\ 0 \leq j < s) \leq \big(\Delta(f_1, \ldots, f_m) \deg_{X_N} Q(0)\big)^s,$$

where $Q(X, T) \in F[X, T]$ is a minimal polynomial for g.

Proof. By (4.6) we have, for every i:

$$r_{i0}, \ldots, r_{i,s-1} \in F\big(f_i(\xi_0), \ldots, f_i(\xi_{s-1}), \xi_0, \ldots, \xi_{s-1}\big).$$

Now choose $\alpha \in K\langle X\rangle_{\mathrm{alg}}$ with the properties stated in Lemma 4.8. Then

$$F\big(X', f_i(\xi_j) : 1 \leq i \leq m,\ 0 \leq j < s\big) = F\big(X', \alpha(\xi_j) : 0 \leq j < s\big)$$

and $\Delta(\alpha) \leq \Delta(f_1, \ldots, f_m)$. The claim now follows from (4.3), (4.4). □

As in the case of Weierstraß Division over $K\langle X\rangle$, from Theorem 4.6 we get:

Corollary 4.14. (Weierstraß Preparation Theorem for $K\langle X\rangle_{\mathrm{alg}}$) *Let $g \in K\langle X\rangle_{\mathrm{alg}}$ be regular in X_N of degree $s \in \mathbb{N}$, and let u be the unit of $K\langle X\rangle$ and $w \in \mathcal{O}\langle X'\rangle[X_N]$ be the Weierstraß polynomial of degree s such that $g = u \cdot w$. Then $u \in K\langle X\rangle_{\mathrm{alg}}$ and $w \in \mathcal{O}\langle X'\rangle_{\mathrm{alg}}[X_N]$, and if $w_0, \ldots, w_{s-1} \in \mathcal{O}\langle X'\rangle_{\mathrm{alg}}$ are such that*

$$w = w_0 + w_1 X_N + \cdots + w_{s-1}X_N^{s-1} + X_N^s$$

and $Q(T) \in F[X, T]$ is a minimal polynomial for g, then

$$\Delta'(w_0, \ldots, w_{s-1}) \leq \big(\deg_{X_N} Q(0)\big)^s.$$

As a consequence, the local ring $K\langle X\rangle_{\mathrm{alg}}$ is noetherian.

4.3. **Effective Weierstraß Division.** We now apply the discussion in the last subsection to the field $K = \mathbb{C}_p$ of p-adic complex numbers, that is, the completion of the algebraic closure $(\mathbb{Q}_p)_{\mathrm{alg}}$ of the field $\mathbb{Q}_p$ of p-adic numbers. We denote by v_p the unique extension of the p-adic valuation on $\mathbb{Q}$ to a valuation on $\mathbb{C}_p$, and we put $|\cdot|_p = |\cdot|_{v_p}^{\log p}$. We take $D = \mathbb{Z}$; then the ring $\mathbb{C}_p\langle X\rangle_{\mathrm{alg}}$ consists of all power series in $\mathbb{C}_p\langle X\rangle$ which are algebraic over $\mathbb{Q}(X)$. Let $f \in \mathbb{C}_p\langle X\rangle_{\mathrm{alg}}$ be non-zero. Put $d := \Delta(f)$ and let $P \in \mathbb{Q}[X, T]$ be a minimal polynomial for f. We may choose P of the form

$$P(T) = a_0 + a_1 T + \cdots + a_d T^d$$

with $a_0, \dots, a_d \in \mathbb{Z}[X]$ without common factor in $\mathbb{Z}[X]$. By Corollary 2.22

$$h_{\mathrm{coeff}}(P) = h\big([a_0 : \cdots : a_d]\big) \leq d\big(h(f) + \log 2\big). \tag{4.9}$$

On the other hand, for all $j = 0, \dots, d$ with $a_j \neq 0$,

$$\begin{aligned} h\big([a_0 : \cdots : a_d]\big) &\geq \deg_{(X)} a_j + m(a_j) \\ &\geq (1 - \log 2) \deg_{(X)} a_j + \log ||a_j||_1 \\ &\geq (1 - \log 2)\big(\deg_{(X)} a_j + \log ||a_j||_1\big), \end{aligned} \tag{4.10}$$

by Lemmas 3.7 and 1.4, (5). Combining (4.9) and (4.10) we get

$$\deg_{(X)} a_j \leq d\big(h(f) + \log 2\big), \qquad \log ||a_j||_1 \leq d\big(h(f) + \log 2\big). \tag{4.11}$$

Using (3.17) we also obtain from (4.10):

$$h(a_j) \leq \frac{d}{1 - \log 2}\big(h(f) + \log 2\big) \qquad \text{for all } j = 0, \dots, d.$$

Moreover, by Proposition 3.10

$$\begin{aligned} h(P) &\leq d\big(h(f) + \log 2 + 1\big) + \log(d+1) \\ &\leq 4d\big(h(f) + \log 2\big). \end{aligned} \tag{4.12}$$

The main goal for this subsection is to show the following generalization of Theorem 0.1 stated in the introduction:

Theorem 4.15. (Effective Weierstraß Division.) *Let $f_1, \dots, f_m, g \in \mathbb{C}_p\langle X\rangle_{\mathrm{alg}}$, $m \geq 1$, and suppose g is regular in X_N of degree $s \in \mathbb{N}$. For each $i = 1, \dots, m$ let $q_i \in \mathbb{C}_p\langle X\rangle_{\mathrm{alg}}$ and $r_i \in \mathbb{C}_p\langle X'\rangle_{\mathrm{alg}}[X_N]$ such that $f_i = q_i g + r_i$ and $\deg_{X_N} r_i < s$. Write*

$$r_i = r_{i0} + r_{i1}X_N + \cdots + r_{i,s-1}X_N^{s-1} \quad \textit{with } r_{ij} \in \mathbb{C}_p\langle X'\rangle_{\mathrm{alg}} \textit{ for all } i, j. \tag{4.13}$$

Then

$$\Delta'(r_{ij} : 1 \leq i \leq m,\ 0 \leq j < s) \leq \big(\Delta(f_1, \dots, f_m)\Delta(g)\big(h(g) + \log 2\big)\big)^s$$

and

$$\begin{aligned} &h_{\max}\big(r_{ij} : 1 \leq i \leq m,\ 0 \leq j < s\big) \leq \\ &\qquad O(1)^s\big(\Delta(f_1, \dots, f_m)\big)^s \Delta(g)\big(h_{\max}(f_1, \dots, f_m) + \log 2\big)\big(h(g) + \log 2\big). \end{aligned}$$

We need the following auxiliary fact, with f as above:

Lemma 4.16. *For each $j \in \mathbb{N}$, we have*

$$h\big(\partial^j f/\partial X_N^j\big) + \log 2 \leq (6d)^j\big(h(f) + \log 2\big).$$

Proof. We first consider the case $j = 1$. We have

$$\frac{\partial f}{\partial X_N} = -\frac{P^*(f)}{P'(f)},$$

where

$$P^*(T) = \sum_{i=0}^{d} \frac{\partial a_i}{\partial X_N} T^i = \frac{\partial P}{\partial X_N} \quad \text{and} \quad P'(T) = \sum_{i=1}^{d} i a_i T^{i-1} = \frac{\partial P}{\partial T},$$

see proof of Lemma 4.10. We may assume that $P^* \neq 0$. By Lemma 2.33 and (4.9)

$$h_{\mathrm{coeff}}(P') \leq h_{\mathrm{coeff}}(P) + \log d \leq d\big(h(f) + \log 2\big) + \log d.$$

Hence by (2.17)

$$h\big(P'(f)\big) \le (d-1)h(f) + h_{\mathrm{coeff}}(P') + \log d \le (2d-1)h(f) + \log(2^d d^2).$$

By (3.19) and (4.11) we have

$$h_{\mathrm{coeff}}(P^*) \le h_{\mathrm{coeff}}(P) + d + \log(d+1) + \log d\big(h(f) + \log 2\big)$$

and hence by (2.17), (4.9) and using that $\log(x+\log 2) \le x$ and $x \le \log 2^x$ for all $x \in \mathbb{R}^{\ge 0}$ we get

$$\begin{aligned} h\big(P^*(f)\big) &\le d\,h(f) + h_{\mathrm{coeff}}(P^*) + \log(d+1) \\ &\le d\,h(f) + h_{\mathrm{coeff}}(P) + d + 2\log(d+1) + \log d\big(h(f) + \log 2\big) \\ &\le 2d\,h(f) + d + \log\big(2^d(d+1)^2 d\big) + \log\big(h(f) + \log 2\big) \\ &\le (2d+1)\,h(f) + \log\big(2^{d+1}(d+1)^2 d\big). \end{aligned}$$

So we have

$$h(\partial f/\partial X_N) \le h\big(P^*(f)\big) + h\big(P'(f)\big) \le 4d\,h(f) + \log\big(2^{2d+1}d^3(d+1)^2\big).$$

A computation shows that $\log\big(2^{2d+1}d^3(d+1)^2\big) \le (6d-1)\log 2$, hence

$$h(\partial f/\partial X_N) + \log 2 \le 6d\big(h(f) + \log 2\big).$$

The lemma now follows by induction on $j \ge 1$. □

For the proof of Theorem 4.15, we let $g \in \mathbb{C}_p\langle X\rangle_{\mathrm{alg}}$ be regular in X_N of degree s. Let

$$Q(T) = b_0 + b_1 T + \cdots + b_e T^e$$

be a minimal polynomial for g, where $e := \Delta(g)$ and $b_0, \ldots, b_e \in \mathbb{Z}[X]$ are without common factor in $\mathbb{Z}[X]$. So we have

$$\deg_{(X)} b_j \le \Delta(g)\big(h(g) + \log 2\big), \qquad \log||b_j||_1 \le \Delta(g)\big(h(g) + \log 2\big) \tag{4.14}$$

and

$$h(b_j) \le \frac{\Delta(g)}{1 - \log 2}\big(h(g) + \log 2\big) \tag{4.15}$$

for all $j = 0, \ldots, e$. (By (4.11), applied to g in place of f.)

For each $i = 1, \ldots, m$ let $q_i \in \mathbb{C}_p\langle X\rangle_{\mathrm{alg}}$ and let $r_i \in \mathbb{C}_p\langle X'\rangle_{\mathrm{alg}}[X_N]$ be of degree $< s$ such that $f_i = q_i g + r_i$, and write r_i as in (4.13). By Corollary 4.13 and (4.14) we get

$$\Delta'(r_{ij} : 1 \le i \le m,\ 0 \le j < s) \le \big(\Delta(f_1, \ldots, f_m)\Delta(g)\big(h(g) + \log 2\big)\big)^s.$$

This is the first estimate in Theorem 4.15.

Let $\xi_0, \ldots, \xi_{s-1}$ be the zeros of g in $\mathbb{Q}(X')_{\mathrm{alg}}$. Recall that for $h \in \mathbb{C}_p\langle X\rangle$ we write $h(\xi_i)$ for $r_h(\xi_i)$, where $r_h \in \mathbb{C}_p\langle X'\rangle[X_N]$ is the remainder obtained through Weierstraß Division of h by g. For each ξ_i we have $b_0(X', \xi_i) = 0$. (See proof of Lemma 4.7.) So by Corollary 2.22, Proposition 3.10, and (4.14), (4.15)

$$\sum_{i=0}^{s-1} h(\xi_i) \le h_{\mathrm{coeff}}(b_0) + \deg_{X_N} b_0 \log 2 \le 4e\big(h(g) + \log 2\big).$$

(Here, $h_{\mathrm{coeff}}(b_0)$ is the height of the polynomial $b_0(X', X_N)$ in the variable X_N with coefficients in $\mathbb{Z}[X']$.) By Corollary 2.22, this already allows us to conclude:

Proposition 4.17. (Effective Weierstraß Preparation.) *Let u be a unit of $\mathbb{C}_p\langle X\rangle_{\mathrm{alg}}$ and*

$$w = w_0 + \cdots + w_{s-1}X_N^{s-1} + X_N^s \quad \textit{with } w_0, w_1, \ldots, w_{s-1} \in \mathbb{C}_p\langle X'\rangle_{\mathrm{alg}}$$

be a Weierstraß polynomial of degree s such that $g = u \cdot w$. Then

$$\Delta'(w_0, \ldots, w_{s-1}) \leq \big(\Delta(g)\big(h(g) + \log 2\big)\big)^s,$$
$$h(w_0, \ldots, w_{s-1}) \leq s\log 2 + 4\Delta(g)\big(h(g) + \log 2\big).$$

Remark 4.18. It is unclear whether the degree bound in Proposition 4.17 is sharp. (Often much better estimates are possible, e.g., if $u \in \mathbb{Q}[X]$, see Lemma 4.5.)

More work is necessary to obtain the upper bound on $h(r_{ij})$ claimed in Theorem 4.15.

Lemma 4.19. *There exists a real constant $C_0 > 0$ with the following property: if $\xi = \xi_i$ for some i, then*

$$h\big(f(\xi)\big) \leq C_0 \cdot \Delta(f)\Delta(g)\big(h(f) + \log 2\big)\big(h(g) + \log 2\big).$$

Hence for all $j \in \mathbb{N}$

$$h\big((\partial^j f/\partial X_N^j)(\xi)\big) \leq C_0 \cdot 6^j\Delta(f)^{j+1}\Delta(g)\big(h(f) + \log 2\big)\big(h(g) + \log 2\big).$$

Proof. We have $b_0(X', \xi) = P\big(X', \xi, f(\xi)\big) = 0$. So if $r(T) = \mathrm{res}_{X_N}(b_0, P) \in \mathbb{Z}[X', T]$ denotes the resultant of $b_0 \in \mathbb{Z}[X]$ and $P \in \mathbb{Z}[X, T]$ with respect to the indeterminate X_N, then $r\big(f(\xi)\big) = 0$. Note that $r(T) \neq 0$, since $a_0, \ldots, a_d$ are relatively prime. Hence $\deg_T(r) > 0$. By Corollary 3.12

$$h(r) \leq \deg_{X_N}(P)\, h(b_0) + \deg_{X_N}(b_0)\, h(P) + \deg_{X_N}(b_0)\deg_{X_N}(P)\log 4.$$

Construing r as a polynomial in the indeterminate T with coefficients in X' we therefore obtain, using Proposition 3.10 and (4.11), (4.12), (4.14), (4.15):

$$h_{\mathrm{coeff}}(r) \leq 9de\big(h(f) + \log 2\big)\big(h(g) + \log 2\big).$$

So by Corollary 2.22, for some constant $C_0 > 0$:

$$h\big(f(\xi)\big) \leq h_{\mathrm{coeff}}(r) + \log 2 \leq C_0 \cdot de\big(h(f) + \log 2\big)\big(h(g) + \log 2\big)$$

as desired. The second estimate follows from Lemma 4.16. □

We may assume that $\xi_0, \ldots, \xi_n$ $(0 \leq n < s)$ are the distinct zeros of g. For $0 \leq i \leq n$ put $\mu_i :=$ the multiplicity of ξ_i. (So $\mu_0 + \cdots + \mu_n = s$.) Set $\xi := (\xi_0, \ldots, \xi_n)$, $\mu := (\mu_0, \ldots, \mu_n)$, and $A := A(\xi, \mu, s)$ (an invertible $s \times s$-matrix with entries in $\mathbb{Q}(X')_{\mathrm{alg}}$). If $s = 1$ then $h(A) = 0$; otherwise, from Proposition 2.32, Lemma 2.26, (4), and $h\big([1 : 2 : \cdots : s-1]\big) = \log(s-1)$ we get that

$$\begin{aligned} h(A) &\leq (s-1)\big(h(\xi_0, \ldots, \xi_n) + \log(s-1)\big) \\ &\leq (s-1)\big(4e\big(h(g) + \log 2\big) + \log(s-1)\big). \end{aligned}$$

Hence from Corollary 2.31 we obtain

$$\begin{aligned} h(A^{-1}) &\leq s\left(\frac{1}{2}\log s + h(A)\right) \\ &\leq \frac{1}{2}s\log s + s(s-1)\log(s-1) + s(s-1)4e\big(h(g) + \log 2\big) \\ &\leq s(s-1/2)\log s + s(s-1)4e\big(h(g) + \log 2\big) \\ &\leq s(s-1/2)\big(\log s + 4e\big(h(g) + \log 2\big)\big), \end{aligned}$$

hence there exists a universal constant $C_1 > 0$ with

$$h(A^{-1}) + \log s \le C_1 s^3 e\big(h(g) + \log 2\big).$$

Consider the $s \times m$-matrices

$$R := \begin{bmatrix} r_{10} & \cdots & r_{m0} \\ \vdots & \ddots & \vdots \\ r_{1,s-1} & \cdots & r_{m,s-1} \end{bmatrix}$$

and $B := [b_1 \cdots b_m]$ where b_i is the transpose of the vector

$$\left[f_i(\xi_0), \frac{\partial f_i}{\partial X_N}(\xi_0), \ldots, \frac{\partial^{\mu_0 - 1} f_i}{\partial X_N^{\mu_0 - 1}}(\xi_0), \ldots, f_i(\xi_n), \frac{\partial f_i}{\partial X_N}(\xi_n), \ldots, \frac{\partial^{\mu_n - 1} f_i}{\partial X_N^{\mu_n - 1}}(\xi_n) \right].$$

From Lemma 4.19 we get

$$\begin{aligned} h_{\max}(B) &= \max_{\substack{1 \le i \le m,\ 0 \le j \le n \\ 0 \le k < \mu_j}} h\big((\partial^k f_i / \partial X_N^k)(\xi_j)\big) \\ &\le C_0 s \cdot (6d)^s e\big(h_{\max}(f_1, \ldots, f_m) + \log 2\big)\big(h(g) + \log 2\big). \end{aligned}$$

Since $R = A^{-1}B$, by Lemma 2.30 therefore

$$\begin{aligned} h_{\max}(R) &\le s\big(h_{\max}(A^{-1}) + h_{\max}(B)\big) + \log s \\ &\le (C_2 d)^s e\big(h_{\max}(f_1, \ldots, f_m) + \log 2\big)\big(h(g) + \log 2\big) \end{aligned}$$

for some universal constant $C_2 > 0$. This finishes the proof of Theorem 4.15. □

Remarks 4.20.

(1) Let $y = y_0 + y_1 X_N + \cdots + y_{s-1} X_N^{s-1}$ with $y_0, \ldots, y_{s-1} \in \mathbb{C}_p\langle X' \rangle_{\mathrm{alg}}$, $s > 0$. Then
$$h(y) \le h(y_0, \ldots, y_{s-1}) + (s - 1 + \log s)$$
by (2.17), Example 3.8 and Lemma 2.26, (3). Therefore the bound on the quantity $h(r_{i0}, \ldots, h_{i,s-1})$ in Theorem 4.15 also entails an upper bound on $h(r_i)$.

(2) Theorem 4.15 in conjunction with Lemma 2.26 immediately gives a bound on the affine heights in the Weierstraß Division Theorem:

$$\begin{aligned} &h(r_{ij} : 1 \le i \le m,\ 0 \le j < s) \le \\ &\qquad mO(1)^s\big(\Delta(f_1, \ldots, f_m)\big)^s \Delta(g)\big(h(f_1, \ldots, f_m) + \log 2\big)\big(h(g) + \log 2\big). \end{aligned}$$

4.4. **Height and degree bounds for algebraic restricted power series.** In the rest of this section we construct a few bounds which will be used in the next section. We fix a non-zero $f \in \mathbb{C}_p\langle X \rangle_{\mathrm{alg}}$, and as in the last subsection we choose a minimal polynomial

$$P(T) = a_0 + a_1 T + \cdots + a_d T^d$$

of f with relatively prime $a_0, \ldots, a_d \in \mathbb{Z}[X]$. For the next proposition suppose that $N > 0$, and let T_e ($e \in \mathbb{N}$, $e > 1$) be the $\mathbb{C}_p$-automorphism of $\mathbb{C}_p\langle X \rangle$ given by

$$X_i \mapsto X_i + X_N^{e^{N-i}} \text{ for } 1 \le i < N, \qquad X_N \mapsto X_N.$$

(See Lemma 4.3.) The power series $T_e(f)$ satisfies the equation

$$T_e(a_d) T_e(f)^d + T_e(a_{d-1}) T_e(f)^{d-1} + \cdots + T_e(a_0) = 0,$$

with $T_e(a_0), \ldots, T_e(a_d) \in \mathbb{Z}[X]$ and $T_e(a_d) \ne 0$. In particular, $\Delta\big(T_e(f)\big) = d$.

Proposition 4.21. $h\big(T_e(f)\big) = O\big(e^{N-1}\big(h(f) + \log 2\big)\big)$.

Proof. Let $\phi = (\phi_0, \dots, \phi_d)$ be the rational map $\mathbb{P}^N \to \mathbb{P}^d$ given by $\phi_i =$ homogenization of a_i, for $i = 0, \dots, d$. Then

$$\phi\big([X_1 + X_N^{e^{N-1}} : \cdots : X_{N-1} + X_N^e : X_N : 1]\big) = \big[T_e(a_0) : \cdots : T_e(a_d)\big].$$

Lemma 3.7 and (4.11) yield

$$h_{\mathrm{coeff}}(\phi_0, \dots, \phi_d) = \max_i \log ||a_i||_\infty \leq d\big(h(f) + \log 2\big).$$

Moreover $\deg \phi_i \leq d\big(h(f) + \log 2\big)$ for all i. Applying (2.16) and using Example 3.9 now gives

$$h\big(\big[T_e(a_0) : \cdots : T_e(a_d)\big]\big) = O\big(e^{N-1}\big(h(f) + \log 2\big)\big),$$

hence by Corollary 2.22 we get

$$h\big(T_e(f)\big) \leq \frac{1}{d} h\big(\big[T_e(a_0) : \cdots : T_e(a_d)\big]\big) + \log 2 = O\big(e^{N-1}\big(h(f) + \log 2\big)\big),$$

as claimed. □

Remark 4.22. Similarly, one shows that $h\big(T_e^{-1}(f)\big) = O\big(e^{N-1}\big(h(f) + \log 2\big)\big)$, where T_e^{-1} is the inverse automorphism to T_e.

Next we establish an upper bound on the p-adic valuation of f:

Lemma 4.23. $|v_p(f)| \leq \Delta(f)\big(h(f) + \log 2\big)/\log p$.

Let $j \in \{0, \dots, d\}$. For every non-zero coefficient $a \in \mathbb{Z}$ of a_j, we have $\log |a| \geq v_p(a_j) \log p$. So by (4.11) we get for non-zero a_j:

$$0 \leq v_p(a_j) \leq d\big(h(f) + \log 2\big)/\log p.$$

The claim now follows from a simple observation about valuations:

Lemma 4.24. *Let v be a valuation on a field K. Let $\alpha \in K^\times$ be a zero of a polynomial*

$$Q(T) = b_0 + b_1 T + \cdots + b_n T^n \in K[T] \qquad (b_0, \dots, b_n \in K, b_n \neq 0).$$

If all non-zero $v(b_i)$ have the same sign, then $|v(\alpha)| \leq \max\big\{|v(b_i)| : b_i \neq 0\big\}$.

Proof. We have

$$|v(\alpha)| \leq \max\left\{\left|\frac{v(b_j) - v(b_i)}{i - j}\right| : i \neq j,\ b_i, b_j \neq 0\right\} \leq \max\big\{|v(b_i)| : b_i \neq 0\big\},$$

as required. □

Remark 4.25. If f lies in $(\mathbb{C}_p)_{\mathrm{alg}}$ or in $\mathbb{Q}(X)$, then we can sharpen the estimate in Lemma 4.23 to:

$$|v_p(f)| \leq \frac{\Delta(f) h(f)}{\log p}.$$

Proof. By passing from f to $1/f$, if necessary, we may assume $v_p(f) \leq 0$, so $|f|_{v_p} \geq 1$. Now if $f \in (\mathbb{C}_p)_{\mathrm{alg}}$, then by (2.15)

$$dh(f) = \sum_{v \in M_{\mathbb{Q}(f)}} \lambda_v \log^+ |f|_v \geq -v_p(f) \log p,$$

so $|v_p(f)| \leq \frac{dh(f)}{\log p}$ as required. If $f \in \mathbb{Q}(X)$, then by (3.11) we see that $h(f) \geq -v_p(f) \log p$ as claimed. □

Lemma 4.26. *Suppose* $|f|_p = 1$. *Then*

$$\deg_{(X)} \overline{f} \le 3d\big(h(f) + \log 2\big), \qquad \deg_X \overline{f} \le d\big(h(f) + \log 2\big).$$

Proof. Let F be an algebraic closure of $\mathbb{F}_p$. We use the height function $h_{(X)}$ on $F(X)$ defined in Example 3.5. For non-zero $a \in F[X]$ we have $h_{(X)}(a) = \deg_{(X)} a$. Now the image $\overline{f}$ of f in $F[X]$ under the residue homomorphism is a zero of the non-zero polynomial

$$\overline{P} = \overline{a_0} + \overline{a_1}T + \cdots + \overline{a_d}T^d \in F[X, T].$$

Moreover, for all $j = 0, \dots, d$ with $\overline{a_j} \ne 0$,

$$h_{(X)}(a_j) = \deg_{(X)} \overline{a_j} \le \deg_{(X)} a_j \le d\big(h(f) + \log 2\big),$$

by (4.11). Therefore, by Corollary 2.22:

$$\begin{aligned}\deg_{(X)} \overline{f} = h_{(X)}(\overline{f}) &\le \frac{1}{d} h_{(X)}\big([\overline{a_0} : \cdots : \overline{a_d}]\big) + \log 2 \\ &\le \frac{1}{d}\sum_j \deg_{(X)}(\overline{a_j}) + \log 2 \le (d+1)\big(h(f) + \log 2\big) + \log 2.\end{aligned}$$

The first inequality follows. For the second inequality use Lemma 4.24, applied to the total degree valuation on $K = F(X)$, the non-zero polynomial $Q(T) = \overline{P}(T) \in K[T]$, and $\alpha = \overline{f}$. □

Now assume that $f \in \mathbb{Z}_p\langle X\rangle_{\mathrm{alg}}$. Then there are uniquely determined polynomials $f_{(0)}, f_{(1)}, \dots$ in $\mathbb{Z}[X]$ such that $||f_{(j)}||_\infty < p$ for all j and

$$f = f_{(0)} + pf_{(1)} + p^2 f_{(2)} + \cdots \qquad \text{in } \mathbb{Z}_p\langle X\rangle.$$

If $f \in \mathbb{Z}[X]$, then $f_{(j)} = 0$ for all $j > h(f)/\log p$, by (3.18). We want to produce a bound on $\deg_{(X)} f_{(j)}$ in terms of $\Delta(f)$, $h(f)$, p, N and j. (This is not used in the later sections.) By the previous lemma

$$\deg_{(X)} f_{(0)} = \deg_{(X)} \overline{f} \le 3\Delta(f)\big(h(f) + \log 2\big),$$

so by (3.18):

$$h\big(f_{(0)}\big) \le 2\deg_{(X)} f_{(0)} + \log||f_{(0)}||_\infty \le 6d\big(h(f) + \log 2\big) + \log p.$$

Now put $f_1 := \big(f - f_{(0)}\big)/p \in \mathbb{Z}_p\langle X\rangle_{\mathrm{alg}}$. Then

$$\begin{aligned}h(f_1) &\le h(f) + h\big(f_{(0)}\big) + \log 2 + \log p \\ &\le h(f) + 6d\big(h(f) + \log 2\big) + \log 2 + 2\log p.\end{aligned}$$

It follows that for some constant $C > 0$:

$$h(f_1) + \log 2 \le C \log p\, \Delta(f)\big(h(f) + \log 2\big)$$

and hence, inductively,

$$\deg_{(X)} f_{(j)} \le (C\log p)^j \Delta(f)^{j+1}\big(h(f) + \log 2\big) \quad \text{for all } j \in \mathbb{N}. \tag{4.16}$$

Now let $j \ge 1$ and put $f_j := f_{(0)} + pf_{(1)} + \cdots + p^{j-1}f_{(j-1)} \in \mathbb{Z}[X]$. Then $f \equiv f_j \bmod p^j$ and $||f_j|| < p^j$, and by (3.18), (4.16):

$$h(f_j) \le 2\deg_{(X)} f_j + \log||f_j||_\infty = O(\log p)^{j-1}\Delta(f)^j\big(h(f) + \log 2\big).$$

These bounds may be used to show that power series like $\sum_{i=0}^\infty p^i X^{2^i} \in \mathbb{Z}_p\langle X\rangle$ (where X is a single indeterminate) are not algebraic over $\mathbb{Q}(X)$.

5. Hermann's method for restricted power series

Consider a system of linear equations $Ay = b$, where the matrix

$$A = (a_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}}$$

and the vector $b = [b_1, \dots, b_m]^{\mathrm{tr}}$ have entries from the ring $\mathbb{Z}_p\langle X\rangle$ of restricted power series with coefficients in $\mathbb{Z}_p$. A necessary condition for solvability of the system $Ay = b$ in $\mathbb{Z}_p\langle X\rangle$ is certainly that this system is solvable in $\mathbb{Q}_p\langle X\rangle$, and that for all $e \ge 1$, the system $\overline{A}y = \overline{b}$, obtained from $Ay = b$ by applying the canonical homomorphism $a \mapsto \overline{a} = a \bmod p^e$ to the coefficients of A and b, is solvable in $\mathbb{Z}_p\langle X\rangle \bmod p^e = (\mathbb{Z}/p^e\mathbb{Z})[X]$. (For the latter state of affairs, we will from now on just say "the system $Ay = b$ is solvable mod p^e.") In this section, we want to prove a partial converse to this: if all entries a_{ij} of A are *algebraic* over $\mathbb{Q}(X)$, then solvability of $Ay = b$ in $\mathbb{Q}_p\langle X\rangle$ and in $\mathbb{Z}_p\langle X\rangle \bmod p^e$ for large enough e implies solvability of $Ay = b$ in $\mathbb{Z}_p\langle X\rangle$. The following theorem contains the precise statement. As before m, n and N range over $\mathbb{N}$; we also let d range over $\mathbb{N}$ and h over $\mathbb{R}$.

Theorem 5.1. *For every tuple (N, d, h, m) with $d, m \ge 1$, $h \ge 0$ there exists a positive integer $E = E(N, d, h, m)$ with*

$$E = 2^{\cdot^{\cdot^{2^{O(1)^N (m^2 d(h+1))^{N+1}}}}} \qquad (N \text{ many } 2\text{'s})$$

such that the following holds: for every $A \in (\mathbb{Z}_p\langle X\rangle_{\mathrm{alg}})^{m\times n}$ with $\Delta(A) \le d$ and $h(A) \le h$ and every $b \in \mathbb{Z}_p\langle X\rangle^m$, the system

$$Ay = b$$

is solvable in $\mathbb{Z}_p\langle X\rangle$ if and only if it is solvable in $\mathbb{Q}_p\langle X\rangle$ and solvable mod p^E.

Provided that $b \in \big(\mathbb{Z}_p\langle X\rangle_{\mathrm{alg}}\big)^n$, we will also show: if $Ay = b$ is solvable in $\mathbb{Z}_p\langle X\rangle$, then there exists a solution $y \in \big(\mathbb{Z}_p\langle X\rangle_{\mathrm{alg}}\big)^n$ whose height and degree can be explicitly bounded in terms of N, $\Delta(A, b)$, $h(A, b)$, m and n. (See (5.10) below.) From the theorem, we get as the special case $m = 1$:

Corollary 5.2. *For all $f_0, f_1, \dots, f_n \in \mathbb{Z}_p\langle X\rangle_{\mathrm{alg}}$ with $f_0 \in (f_1, \dots, f_n)\mathbb{Q}_p\langle X\rangle$ and $\Delta(f_1, \dots, f_n) \le d$, $h(f_1, \dots, f_n) \le h$, we have*

$$f_0 \in (f_1, \dots, f_n)\mathbb{Z}_p\langle X\rangle \quad\Longleftrightarrow\quad \overline{f_0} \in (\overline{f_1}, \dots, \overline{f_n})(\mathbb{Z}/p^E\mathbb{Z})[X],$$

where $E = E(N, d, h, 1)$. (Together with (3.18), this yields Theorem 0.2.)

To discuss our strategy for the proof of Theorem 5.1, let D be an integral domain with fraction field F, and let $A = (a_{ij})$ be an $m \times n$-matrix with entries $a_{ij} \in D$ and $b = [b_1, \dots, b_m]^{\mathrm{tr}} \in D^m$. We are mostly interested in the case where $D = \mathbb{Z}_p\langle X\rangle$ or $D = \mathbb{Z}_p\langle X\rangle_{\mathrm{alg}}$; the ideas explained below were first used by Hermann [18] in the case where D is a polynomial ring over a field. We want to determine whether the system

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \tag{I}$$

(or: $Ay = b$) has a solution $y = [y_1, \dots, y_n]^{\mathrm{tr}} \in D^n$, and if it does, effectively find such a solution. Of course, we may assume $A \ne 0$, so the rank $r = \mathrm{rank}_F(A)$ of A (considered as a matrix over F) is ≥ 1. Let Δ be an $r \times r$-submatrix of A with $\delta = \det\Delta \ne 0$.

After rearranging the order of the equations and permuting the unknowns $y_1, \dots, y_n$ in (I) we may assume that $\Delta = (a_{ij})_{1 \le i,j \le r}$. A *necessary condition* for (I) to have a solution $y \in D^n$ is clearly that

$$\operatorname{rank}_F(A) = \operatorname{rank}_F(A, b). \tag{NC}$$

Assume (NC) holds. Then (I) has the same solutions in D^n as the system:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rn} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_r \end{bmatrix}$$

Changing notation, we assume from now on that $r = m$. Multiplying both sides of $Ay = b$ on the left by the adjoint Δ^{ad} of Δ, (I) turns into the system

$$\begin{bmatrix} \delta & & & c_{1,r+1} & \cdots & c_{1,n} \\ & \delta & & c_{2,r+1} & \cdots & c_{2,n} \\ & & \ddots & \vdots & \ddots & \vdots \\ & & \delta & c_{r,r+1} & \cdots & c_{rn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_r \end{bmatrix} \tag{S}$$

(with $c_{ij}, d_i \in D$ for $1 \le i \le r < j \le n$) which has the same solutions in every domain extending D as (I). Clearly, a sufficient condition for (S) to have a solution $y = [y_1, \dots, y_n]^{\mathrm{tr}} \in D^n$ is that $d_1, \dots, d_r$ are each divisible by δ. This will be the case if δ is a unit; then a solution to (S) (and hence to (I)) is given by

$$y_j = \begin{cases} d_j/\delta & \text{for } 1 \le j \le r, \\ 0 & \text{for } r < j \le n. \end{cases}$$

Suppose δ is not a unit, so $\overline{D} = D/\delta D \neq 0$. Then, reducing the coefficients in (S) modulo δ, the system (S) turns into

$$\begin{bmatrix} \overline{c_{1,r+1}} & \cdots & \overline{c_{1n}} \\ \vdots & \ddots & \vdots \\ \overline{c_{r,r+1}} & \cdots & \overline{c_{rn}} \end{bmatrix} \begin{bmatrix} y_{r+1} \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \overline{d_1} \\ \vdots \\ \overline{d_r} \end{bmatrix} \tag{$\overline{\mathrm{S}}$}$$

over $\overline{D}$. (Here $\overline{a}$ denotes the image of $a \in D$ in $\overline{D}$.) For any $y_{r+1}, \dots, y_n \in D$ with the property that $[\overline{y_{r+1}}, \dots, \overline{y_n}]^{\mathrm{tr}}$ is a solution of the reduced system ($\overline{\mathrm{S}}$) there are uniquely determined $y_1, \dots, y_r \in D$ such that

$$[y_1, \dots, y_r, y_{r+1}, \dots, y_n]^{\mathrm{tr}} \in D^n$$

is a solution of (S), and hence of (I). In particular, (I) is solvable in D if and only if ($\overline{\mathrm{S}}$) is solvable in $\overline{D}$.

Now we specialize to the case $D = \mathbb{Z}_p\langle X \rangle$ with $N > 0$. The procedure above can be used to reduce solving a system (I) over D to solving an equivalent system over $D' := \mathbb{Z}_p\langle X' \rangle$. For this, suppose that the system $Ay = b$ is solvable in $\mathbb{Q}_p\langle X \rangle$; in particular, the necessary condition (NC) holds. Assume moreover that A has an $r \times r$-minor δ with $\delta \bmod p \neq 0$. Then, after applying the $\mathbb{Z}_p$-automorphism T_d of D given by Lemma 4.3, where $d > \deg_X(\delta \bmod p)$, to all coefficients of (S), we may even assume that δ is regular in X_N of some degree s. By Weierstraß Division we now have

$$\overline{D} = D/\delta D \cong D' \oplus D'\overline{X_N} \oplus \cdots \oplus D'\overline{X_N}^{s-1}$$

as D'-algebras. This allows us to replace the system ($\overline{\mathrm{S}}$) with an equivalent system over D' as follows. Each $\overline{c_{ij}}$ can be written uniquely as

$$\overline{c_{ij}} = c_{ij,0} + c_{ij,1}\overline{X_N} + c_{ij,2}\overline{X_N}^2 + \cdots + c_{ij,s-1}\overline{X_N}^{s-1}$$

with $c_{ij,0}, \dots, c_{ij,s-1} \in D'$, $1 \le i \le r < j \le n$. We also write each $\overline{d_i}$ and each power $\overline{X_N}^t$ ($t \ge s$) in this way,

$$\overline{d_i} = d_{i0} + d_{i1}\overline{X_N} + d_{i2}\overline{X_N}^2 + \cdots + d_{i,s-1}\overline{X_N}^{s-1}$$

with $d_{ik} \in D'$, $1 \le i \le r$, $0 \le k < s$,

$$\overline{X_N}^t = \xi_{t0} + \xi_{t1}\overline{X_N} + \cdots + \xi_{t,s-1}\overline{X_N}^{s-1} \tag{5.1}$$

with $\xi_{tk} \in D'$, and also each unknown y_j as

$$y_j = y_{j0} + y_{j1}\overline{X_N} + \cdots + y_{j,s-1}\overline{X_N}^{s-1}$$

with new unknowns y_{jk} ($r < j \le n$, $0 \le k < s$) ranging over D'. The coefficient matrix of the system ($\overline{\mathrm{S}}$) above may then be written as

$$C(0) + C(1)\overline{X_N} + \cdots + C(s-1)\overline{X_N}^{s-1},$$

where

$$C(k) = (c_{ijk})_{\substack{1 \le i \le r \\ r < j \le n}} \in (D')^{r \times (n-r)} \qquad \text{for } k = 0, \dots, s-1.$$

Similarly we may write

$$\begin{bmatrix} \overline{d_1} \\ \vdots \\ \overline{d_r} \end{bmatrix} = d(0) + d(1)\overline{X_N} + \cdots + d(s-1)\overline{X_N}^{s-1}, \qquad d(k) = \begin{bmatrix} d_{1k} \\ \vdots \\ d_{rk} \end{bmatrix},$$

and also

$$\begin{bmatrix} y_{r+1} \\ \vdots \\ y_n \end{bmatrix} = y(0) + y(1)\overline{X_N} + \cdots + y(s-1)\overline{X_N}^{s-1}, \qquad y(k) = \begin{bmatrix} y_{r+1,k} \\ \vdots \\ y_{nk} \end{bmatrix}.$$

So our system may be rewritten as

$$\sum_{k=0}^{2(s-1)} \left(\sum_{l=0}^{k} C(k-l)y(l) \right) \overline{X_N}^k = d(0) + d(1)\overline{X_N} + \cdots + d(s-1)\overline{X_N}^{s-1}. \tag{5.2}$$

Using the identities (5.1), the left-hand side reduces to

$$\sum_{k=0}^{s-1} \left(\sum_{l=0}^{k} \left(C(k-l) + \sum_{t=s}^{2(s-1)} C(t-l)\xi_{tk} \right) y(l) + \sum_{l=k+1}^{s-1} \left(\sum_{t=s}^{2(s-1)} C(t-l)\xi_{tk} \right) y(l) \right) \overline{X_N}^k.$$

Comparing the coefficients of equal powers of $\overline{X_N}$ in (5.2) we thus obtain s systems of linear equations over D':

$$\sum_{l=0}^{s-1} \left(C(k-l) + \sum_{t=s}^{2(s-1)} C(t-l)\xi_{tk} \right) y(l) = d(k)$$

for $k = 0, \dots, s-1$, where we put $C(t) := 0$ for $t < 0$. Combining these systems into a single one, we obtain a system

$$A'y' = b', \tag{I'}$$

where

$$A' \in (D')^{m' \times n'},\ b' \in (D')^{m'},\ m' = rs,\ n' = s(n-r),$$

whose solutions in D' are in one-to-one correspondence with the solutions in D of $Ay = b$. Note that the new system is solvable in $\mathbb{Q}_p\langle X'\rangle$, by Weierstraß Division for $\mathbb{Q}_p\langle X\rangle$. (Here the right choice of rings is essential: if we replace $\mathbb{Q}_p\langle X\rangle$ by $\mathbb{Q}_p[X]$ or by the fraction field of $\mathbb{Z}_p\langle X\rangle$, say, there is no such "preservation of solvability.") Moreover, if the original system $Ay = b$ is solvable modulo some power of p, then $A'y' = b'$ is solvable modulo the same power of p. Setting $D_{\mathrm{alg}} := \mathbb{Z}_p\langle X\rangle_{\mathrm{alg}}$, $D'_{\mathrm{alg}} := \mathbb{Z}_p\langle X'\rangle_{\mathrm{alg}}$ we also have: if $A \in (D_{\mathrm{alg}})^{m\times n}$, then $A' \in (D'_{\mathrm{alg}})^{m'\times n'}$, and if in addition $b \in (D_{\mathrm{alg}})^m$, then also $b' \in (D'_{\mathrm{alg}})^{m'}$.

The reduction from $\mathbb{Z}_p\langle X\rangle$ to $\mathbb{Z}_p\langle X'\rangle$ described above breaks down if $\delta \bmod p = 0$ for all $r \times r$-minors δ of A, since then Weierstraß Division by δ is *inapplicable.* To overcome this obstacle, in this case we shall first transform the system (I) into an equivalent system for which $\delta \bmod p \neq 0$ for a suitable $r \times r$-minor δ of the new coefficient matrix, which is of rank r.

5.1. **Desingularization.** This process, which we call p**-desingularization**, can be formulated in the quite general context that D is an integral domain and $p \in D$ a non-zero generator of a prime ideal $(p) = pD$ such that $\bigcap_{e\in\mathbb{N}} p^e D = (0)$. We write v for the p-adic valuation on D, given by $v(a) = e \in \mathbb{N}$ if $a \in (p^e) \setminus (p^{e+1})$, for non-zero $a \in D$, and $v(0) := -\infty < \mathbb{N}$. We also put $F = \mathrm{Frac}(D)$ and $F(p) = \mathrm{Frac}\big(D/pD\big)$.

Let $A = (a_{ij})$ be an $m \times n$-matrix over D, of rank $r = \mathrm{rank}_F(A)$, and let $b = [b_1, \dots, b_m]^{\mathrm{tr}} \in D^m$. We shall show how to construct an $m \times n$-matrix B over D, depending only on A (and not on b), and a vector $c = [c_1, \dots, c_m]^{\mathrm{tr}} \in D^m$ with the following properties:

(1) $r = \mathrm{rank}_F(B) = \mathrm{rank}_{F(p)}(B \bmod p)$, and
(2) the systems $Ay = b$ and $By = c$ have the same solutions (in every domain containing D).

It may happen that one of the steps of the algorithm to construct (B, c) cannot be carried out, but then we will know that $Ay = b$ has no solution $y \in D^n$.

We may assume $A \neq 0$, since otherwise we may just take $(B, c) = (A, b)$. By removing superfluous rows from A we may of course assume that the rows of A are F-linearly independent, i.e., $m = r$. Let Δ be an $r \times r$-submatrix of A such that the value $v(\det \Delta)$ is *minimal* among all $r \times r$-submatrices of A. Without loss of generality, $\Delta = (a_{ij})_{1\le i,j\le r}$. As above, consider now the system

$$\begin{bmatrix} \delta & & & c_{1,r+1} & \cdots & c_{1,n} \\ & \delta & & c_{2,r+1} & \cdots & c_{2,n} \\ & & \ddots & \vdots & \ddots & \vdots \\ & & \delta & c_{r,r+1} & \cdots & c_{rn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_r \end{bmatrix} \tag{5.3}$$

which is obtained by multiplying both sides of $Ay = b$ from the left with the adjoint of Δ. It has the same solutions as $Ay = b$, in any domain extending D. Here, $\delta = \det \Delta$, the c_{ij} are certain signed $r \times r$-minors of A, and the d_i are certain signed $r \times r$-minors of the extended matrix (A, b). In particular, $v(c_{ij}) \ge v(\delta)$ for all i, j, by choice of Δ. Therefore,

if $v(d_i) < v(\delta)$ for some $i \in \{1, \dots, m\}$, then (5.3) and hence the original system $Ay = b$ are not solvable in D. Suppose $v(d_i) \geq v(\delta) = e$ for all $i = 1, \dots, m$. Dividing all coefficients δ, c_{ij} and d_i in (5.3) by p^e, we obtain a system $By = c$ as required.

Put $e(A) = \max_S v(\det S)$, with S ranging over all $r \times r$-submatrices of A. The p-desingularization process described above can be successfully performed on the system $Ay = b$ provided this system is solvable mod p^e, where $e = v(\delta) \leq e(A)$. If $y \in D^n$ is a solution to $Ay = b \bmod p^t$, where $t \geq e$, then clearly y is also a solution to $By = c \bmod p^{t-e}$. Hence:

Lemma 5.3. *If $Ay = b$ is solvable mod p^t, where $t \geq e(A)$, then the p-desingularization process may be successfully performed on (A, b), and the resulting system $By = c$ is solvable mod $p^{t-e(A)}$.* □

5.2. **Desingularization and height.** Let us now return to $D = \mathbb{Z}_p\langle X\rangle$. In the remainder of this section, $A = (a_{ij})$ will always denote an $m \times n$-matrix with coefficients in $D_{\mathrm{alg}} = \mathbb{Z}_p\langle X\rangle_{\mathrm{alg}}$, of rank r over $F = \mathrm{Frac}(D)$, and $b = [b_1, \dots, b_m]^{\mathrm{tr}}$ will denote a vector with $b_1, \dots, b_m \in D$. Note that the entries of the matrix B produced by p-desingularization lie in D_{alg}, and if $b \in (D_{\mathrm{alg}})^m$, then $c \in (D_{\mathrm{alg}})^m$ as well. We now want to estimate the height of B. For every $r \times r$-submatrix S of A we have

$$h(\det S) + \log 2 \leq r\big(\log r + h_{\max}(S)\big) + \log 2 \leq r^2\big(h_{\max}(A) + \log 2\big). \tag{5.4}$$

(By Lemma 2.28 and the crude estimate $r \log r + \log 2 \leq r^2 \log 2$.) Hence

$$e(A) \leq r^2\Delta(A)\big(h_{\max}(A) + \log 2\big)/\log p \tag{5.5}$$

by Lemma 4.23. Therefore, if $Ay = b$ is solvable mod p^t, with $t \geq s$, where s is the integral part

$$s = \big[m^2\Delta(A)\big(h_{\max}(A) + \log 2\big)/\log p\big],$$

then p-desingularization can be carried out, and the system $By = c$ resulting from this process is solvable mod p^{t-s}.

The heights of the entries of the coefficient matrix of the system (5.3) above are bounded from above by $m^2\big(h_{\max}(A) + \log 2\big) - \log 2$, by (5.4). So the matrix B produced by p-desingularization satisfies:

$$\begin{aligned} h_{\max}(B) + \log 2 &\leq e(A)\log p + m^2\big(h_{\max}(A) + \log 2\big) \\ &\leq 2m^2\Delta(A)\big(h_{\max}(A) + \log 2\big). \end{aligned}$$

Similarly, if all entries $b_1, \dots, b_m$ of the vector b are algebraic over $\mathbb{Q}(X)$, one gets

$$h_{\max}(B, c) + \log 2 \leq 2m^2\Delta(A, b)\big(h_{\max}(A, b) + \log 2\big).$$

We summarize this discussion:

Proposition 5.4. *If the system $Ay = b$ is solvable mod p^t, where*

$$t \geq s = \big[m^2\Delta(A)\big(h_{\max}(A) + \log 2\big)/\log p\big],$$

then the p-desingularization algorithm for (A, b) can be carried out, and the resulting system $By = c$ is solvable mod p^{t-s}. Moreover, we have

$$h_{\max}(B) \leq 2m^2\Delta(A)\big(h_{\max}(A) + \log 2\big) - \log 2$$

and if the entries of b are from D_{alg}, then so are the entries of c, with

$$h_{\max}(B, c) \leq 2m^2\Delta(A, b)\big(h_{\max}(A, b) + \log 2\big) - \log 2.$$

□

5.3. **Height and degree bounds in Hermann's method.** We now want to determine how m', $h_{\max}(A')$ and $\Delta'(A')$ depend on m, $h_{\max}(A)$ and $\Delta(A)$. In order to do this efficiently, we introduce the following ad-hoc notation:

Notation. We call the real number

$$\mathrm{c}(A) := \Delta(A)\big(h_{\max}(A)+1\big)$$

the **complexity of** A**.**

Note that always $\mathrm{c}(A) \geq 1$, and $\mathrm{c}(S) \leq \mathrm{c}(A)$ for every submatrix S of A (in particular $\Delta(a)\big(h(a)+1\big) \leq \mathrm{c}(A)$ for every entry a of A). For given real number β and integers $m, n > 0$, there exist only finitely many $m \times n$-matrices A over $\mathbb{Z}_p\langle X\rangle_{\mathrm{alg}}$ with $\mathrm{c}(A) \leq \beta$. We also abbreviate $m^2\,\mathrm{c}(A)$ by $\mathrm{c}^*(A)$.

Rather than studying how $\Delta(A)$ and $h_{\max}(A)$ change during the Hermann algorithm, we will now concentrate on estimating the *complexity* of the matrices constructed during this process.

5.3.1. *Complexity of p-desingularization.* By Proposition 5.4, the p-desingularization algorithm can be performed on (A, b) provided the system $Ay = b$ is solvable mod p^s, where $s = \big[\mathrm{c}^*(A)/\log p\big]$, and the matrix B obtained in this way satisfies $h_{\max}(B) \leq 2\,\mathrm{c}^*(A) - \log 2$. If $b \in (D_{\mathrm{alg}})^m$, then similarly $h_{\max}(B, c) \leq 2\,\mathrm{c}^*(A, b) - \log 2$.

5.3.2. *Complexity of $T_d(B)$.* We now want to bound $\mathrm{c}\,\big(T_d(B)\big)$. Recall that the system $By = c$ obtained from $Ay = b$ by p-desingularization has the shape (5.3), where $\delta \bmod p \neq 0$. By Lemma 4.26 we have

$$\deg_X(\delta \bmod p) \leq \Delta(\delta)\big(h(\delta)+\log 2\big) \leq \mathrm{c}(B) \leq 2\Delta(A)\,\mathrm{c}^*(A). \tag{5.6}$$

Let $d = \big[2\Delta(A)\,\mathrm{c}^*(A)\big] + 1$ (so that $d > \deg_X(\delta \bmod p)$ by (5.6)). Then $T_d(\delta)$ is X_N-distinguished of some degree s with $s < d^N \leq \big(4\Delta(A)\,\mathrm{c}^*(A)\big)^N$, and by Proposition 4.21,

$$h\big(T_d(f)\big) = O(1)^{N-1}\big(\Delta(A)\,\mathrm{c}^*(A)\big)^{N-1}\big(h(f)+\log 2\big) \qquad \text{for all } f \in D_{\mathrm{alg}}.$$

We conclude that

$$\mathrm{c}\,\big(T_d(B)\big) = O(1)^{N-1}\big(\Delta(A)\,\mathrm{c}^*(A)\big)^N. \tag{5.7}$$

Similarly, if $b \in (D_{\mathrm{alg}})^m$, then applying T_d to all coefficients of $By = c$ yields a system of linear equations over D_{alg} whose extended coefficient matrix has complexity $O(1)^{N-1}\big(\Delta(A, b)\,\mathrm{c}^*(A, b)\big)^N$.

5.3.3. *Complexity of A'.* From the system obtained by applying T_d to all coefficients of the desingularized system $By = c$, we now construct $A' \in (D'_{\mathrm{alg}})^{m' \times n'}$ and $b' \in (D')^{m'}$ as above. Let $f_1, \ldots, f_K \in D_{\mathrm{alg}} = \mathbb{Z}_p\langle X\rangle_{\mathrm{alg}}$, and write each $\overline{f_k}$ = canonical image of f_k in $D_{\mathrm{alg}}/T_d(\delta)D_{\mathrm{alg}}$ as

$$\overline{f_k} = f_{k0} + f_{k1}\overline{X_N} + \cdots + f_{k,s-1}\overline{X_N}^{s-1} \qquad \text{with } f_{kl} \in D'_{\mathrm{alg}}.$$

Then by Theorem 4.15 and the estimate (5.7):

$$\Delta'(f_{kl} : 1 \leq k \leq K,\, 0 \leq l < s) = O(1)^{(N-1)s}\Delta(f_1, \ldots, f_K)^s\big(\Delta(A)\,\mathrm{c}^*(A)\big)^{Ns}$$

and

$$h_{\max}(f_{kl} : 1 \leq k \leq K,\, 0 \leq l < s) = \\ O(1)^{N+s-1}\big(\Delta(f_1, \ldots, f_K)\big)^s\big(h_{\max}(f_1, \ldots, f_K) + \log 2\big)\big(\Delta(A)\,\mathrm{c}^*(A)\big)^N.$$

Hence

$$\Delta'\big(C(k), \xi_{tk} : 0 \le k < s \le t \le 2(s-1)\big) = O(1)^{Ns}\Delta(A)^{(N+1)s}\,\mathrm{c}^*(A)^{Ns}$$

and

$$h_{\max}\big(C(k) : 0 \le k < s\big) = O(1)^{Ns}\Delta(A)^{2N+s}\,\mathrm{c}^*(A)^{2N}, \tag{5.8}$$

and since $h_{\max}(X_N^s, \dots, X_N^{2(s-1)}) = 2(s-1)$, we also have

$$h_{\max}\big(\xi_{tk} : 0 \le k < s \le t \le 2(s-1)\big) = O(1)^{Ns}\big(\Delta(A)\,\mathrm{c}^*(A)\big)^N. \tag{5.9}$$

Moreover, for $0 \le k, l < s$ we get

$$h_{\max}\left(C(k-l) + \sum_{t=s}^{2(s-1)} C(t-l)\xi_{tk}\right) \le h_{\max}\big(C(k-l)\big) + \sum_{t=s}^{2(s-1)} h_{\max}\big(C(t-l)\big) + h(\xi_{tk}) + \log s.$$

The right-hand side in this inequality may be bounded from above by

$$s \cdot h_{\max}\big(C(0), \dots, C(s-1)\big) + (s-1) \cdot h_{\max}(\xi_{sk}, \dots, \xi_{2(s-1),k}) + \log s,$$

so with (5.8), (5.9):

$$h_{\max}(A') = O(1)^{Ns}\Delta(A)^{2N+s}\,\mathrm{c}^*(A)^{2N}.$$

This yields $\mathrm{c}(A') = O(1)^{Ns}\,\mathrm{c}^*(A)^{O(1)Ns}$, and therefore, since $m' \le ms$ and $s = O(1)^N\,\mathrm{c}^*(A)^N$:

$$\mathrm{c}^*(A') = O(1)^{Ns}\,\mathrm{c}^*(A)^{O(1)Ns}.$$

We can generously estimate

$$\mathrm{c}^*(A') \le 2^{O(1)^N\,\mathrm{c}^*(A)^{N+1}}.$$

Now suppose that $b \in (D_{\mathrm{alg}})^m$. Then similarly as above we obtain

$$\mathrm{c}(A', b') \le 2^{O(1)^N\,\mathrm{c}^*(A,b)^{N+1}}.$$

In addition, let $y' = (y_{jk}) \in (D'_{\mathrm{alg}})^{n'}$ be a solution to the system $A'y' = b'$, and $y \in (D_{\mathrm{alg}})^n$ the corresponding solution to the original system $Ay = b$. We have, for $j = r+1, \dots, n$,

$$T_d(y_j) = y_{j0} + y_{j1}X_N + \cdots + y_{j,s-1}X_N^{s-1},$$

and thus, by Remarks 4.20, (1)

$$h\big(T_d(y_j)\big) \le (s-1) + s \cdot h_{\max}(y') + \log s.$$

So we get

$$h(y_j) = O\big(d^{N-1}\big(h\big(T_d(y_j)\big) + \log 2\big)\big) = O(1)^N\,\mathrm{c}^*(A)^{3N-1}\big(h_{\max}(y') + 1\big).$$

For $i = 1, \dots, r$, we have

$$\delta y_i = c_{i,r+1}y_{r+1} + \cdots + c_{i,n}y_n - d_i,$$

where $h(c_{ij})$ and $h(d_i)$ are bounded from above by $\mathrm{c}^*(A,b)$, as is $h(\delta)$. (See (5.4).) Hence

$$\begin{aligned} h(y_i) &\le h(\delta) + \sum_{j=r+1}^{n} \big(h(c_{ij}) + h(y_j)\big) + h(d_i) + \log(n-r+1) \\ &\le (n-r+2)\,\mathrm{c}^*(A,b) + (n-r)h_{\max}(y_{r+1},\dots,y_n) + \log(n-r+1) \\ &\le O(1)^N n\,\mathrm{c}^*(A,b)^{3N-1}\big(h_{\max}(y')+1\big), \end{aligned}$$

and also $\Delta(y_i) \le \Delta(A,b)\Delta(y')$. We put everything together:

Proposition 5.5. *Let $A \in D_{\mathrm{alg}}^{m\times n}$, $b \in D^m$, and assume that $N > 0$. Suppose the system $Ay = b$ is solvable mod p^t, where $t \ge s = \big[\mathrm{c}^*(A)/\log p\big]$. Let $A'y' = b'$ with $A' \in (D'_{\mathrm{alg}})^{m'\times n'}$, $b' \in (D')^{m'}$ be the system obtained from $Ay = b$ by the procedure sketched above. Then $A'y' = b'$ is solvable mod p^{t-s}, and*

$$\mathrm{c}^*(A') = 2^{O(1)^N\,\mathrm{c}^*(A)^{N+1}}, \qquad n' = O(1)^N n\,\mathrm{c}^*(A)^N.$$

If in addition $b \in (D_{\mathrm{alg}})^m$, then

$$\mathrm{c}^*(A',b') = 2^{O(1)^N\,\mathrm{c}^*(A,b)^{N+1}},$$

and if $y' \in (D'_{\mathrm{alg}})^{n'}$ is a solution to $A'y' = b'$ and $y \in (D_{\mathrm{alg}})^n$ the corresponding solution to $Ay = b$, then

$$\begin{aligned} h_{\max}(y) &= O(1)^N n\,\mathrm{c}^*(A,b)^{3N-1}\big(h_{\max}(y')+1\big), \\ \Delta(y) &\le \Delta(A,b)\Delta(y'). \end{aligned}$$

□

5.4. **Towers of exponentials.** Before we go on, let us introduce the following useful notation: for a real number a, we define recursively

$$2 \uparrow^0 a := a, \qquad 2 \uparrow^{n+1} a := 2^{2\uparrow^n a}.$$

So $2 \uparrow^1 a = 2^a$, $2 \uparrow^2 a = 2^{2^a}$, $2 \uparrow^3 a = 2^{2^{2^a}}$, and so on. It is clear that the function $(n,a) \mapsto 2 \uparrow^n a \colon \mathbb{N}\times\mathbb{N} \to \mathbb{N}$ is primitive recursive. (However, it is *not* elementary recursive in the sense of Kalmár, see [19].) Note that for $a \le b$ in $\mathbb{R}$ and $n \le m$, we have $2 \uparrow^n a \le 2 \uparrow^m b$. We note a few other basic estimates:

Lemma 5.6. *Let $a, b \in \mathbb{R}$, $a, b \ge 1$. Then*

(1) $(2 \uparrow^n a) + b \le 2 \uparrow^n (a+b)$.
(2) $(2 \uparrow^n a)\cdot b \le 2 \uparrow^n (a+b)$, *for* $n \ge 1$.
(3) $(2 \uparrow^n a)^b \le 2 \uparrow^n (a+b)$, *for* $n \ge 2$.

Proof. We prove (1) by induction on n. The case $n = 0$ being trivial, suppose (1) holds for a certain n. We have $1 + b \le 2^b$, so that

$$(2 \uparrow^{n+1} a) + b = 2^{2\uparrow^n a} + b \le 2^{2\uparrow^n a}(1+b) \le 2^{2\uparrow^n a}2^b = 2^{(2\uparrow^n a)+b},$$

and by applying the inductive hypothesis, the claim follows. Property (2) is now a consequence of (1), since

$$(2 \uparrow^n a)\cdot b \le 2^{2\uparrow^{n-1} a}\cdot 2^b \le 2^{2\uparrow^{n-1}(a+b)} = 2 \uparrow^n (a+b),$$

for $n \ge 1$. Similarly, (3) follows from (2), since

$$(2 \uparrow^n a)^b = 2^{(2\uparrow^{n-1} a)\cdot b} \le 2^{2\uparrow^{n-1}(a+b)} = 2 \uparrow^n (a+b)$$

for all $n \ge 2$. □

5.5. **Proof of Theorem 5.1.** Suppose the system $Ay = b$ has a solution in $\mathbb{Q}_p\langle X\rangle$ (in particular, $Ay = b$ satisfies (NC) above) and is also solvable *modulo some very high power p^E of p.* (We will determine a suitable E in the process.) We now successively construct "equivalent" matrix equations

$$A^{(N)}y^{(N)} = b^{(N)} \tag{S_N}$$

$$\vdots$$

$$A^{(\nu)}y^{(\nu)} = b^{(\nu)} \tag{S_ν}$$

$$\vdots$$

$$A^{(0)}y^{(0)} = b^{(0)}, \tag{S_0}$$

where

(1) $0 \leq \nu \leq N$,
(2) $A^{(\nu)}$ is an $m(\nu) \times n(\nu)$-matrix with entries in the ring $\mathbb{Z}_p\langle X_1, \dots, X_\nu\rangle_{\mathrm{alg}}$,
(3) $y^{(\nu)} = \big[y_1^{(\nu)}, \dots, y_{n(\nu)}^{(\nu)}\big]^{\mathrm{tr}}$ is a vector of unknowns, and
(4) $b^{(\nu)} = \big[b_1^{(\nu)}, \dots, b_{m(\nu)}^{(\nu)}\big]^{\mathrm{tr}}$ is a vector with coordinates in $\mathbb{Z}_p\langle X_1, \dots, X_\nu\rangle$.

The *initial* equation (S_N) is just $Ay = b$, and if $\nu > 0$, the equation ($\mathrm{S}_{\nu-1}$) is obtained from (S_ν) by the procedure described earlier in this section: Weierstraß Division by a suitable minor, after a preliminary p-desingularization. The matrices $A^{(N-1)}, \dots, A^{(0)}$ obtained in this way only depend on the initial matrix $A^{(N)} = A$, and not on the initial vector $b^{(N)} = b$. Also, if the entries of $b^{(N)} = b$ are algebraic over $\mathbb{Q}(X)$, then the entries of $b^{(\nu)}$ are algebraic over $\mathbb{Q}(X_1, \dots, X_\nu)$, for $0 \leq \nu \leq N$, and if (S_ν) is solvable in $\mathbb{Q}_p\langle X_1, \dots, X_\nu\rangle$, then ($\mathrm{S}_{\nu-1}$) is solvable in $\mathbb{Q}_p\langle X_1, \dots, X_{\nu-1}\rangle$. Of course, we have to ensure that at each stage of this process, we are able to successfully carry out p-desingularization on $(A^{(\nu)}, b^{(\nu)})$. This can be achieved by choosing E large enough: Suppose that $E \geq e_N + e_{N-1} + \cdots + e_1$ with $e_\nu \geq e(A^{(\nu)})$ for $1 \leq \nu \leq N$; then p-desingularization is applicable to (S_N), and the system (S_{N-1}) will be solvable modulo p^{E-e_N}, hence p-desingularization is applicable to (S_{N-1}), and so on. By Proposition 5.4, it suffices to take for e_ν the integral part $\big[\mathrm{c}^*(A^{(\nu)})/\log p\big]$. By Proposition 5.5, the complexity of $A^{(\nu)}$ in turn can be bounded in terms of the complexity of A:

Lemma 5.7. *For all $\nu = 0, \dots, N$, we have*

$$\mathrm{c}^*\big(A^{(\nu)}\big) = 2 \uparrow^{N-\nu} O(1)^N\, \mathrm{c}^*(A)^{N+1},$$

and if $b \in (D_{\mathrm{alg}})^m$, then

$$\mathrm{c}^*\big(A^{(\nu)}, b^{(\nu)}\big) = 2 \uparrow^{N-\nu} O(1)^N\, \mathrm{c}^*(A, b)^{N+1}.$$

Proof. Let $C \geq 1$ be the universal constant such that $\mathrm{c}^*(A') \leq 2^{C^N\, \mathrm{c}^*(A)^{N+1}}$ from Proposition 5.5, and put $\mathrm{c}^{**}(A) := C\, \mathrm{c}^*(A)$. Since $Cx \leq x^C$ for all $x \in \mathbb{R}$, $x \geq 2$, we have the simple estimate $\mathrm{c}^{**}(A') \leq 2^{\mathrm{c}^{**}(A)^{N+1}}$. We now claim that

$$\mathrm{c}^{**}\big(A^{(\nu)}\big) \leq 2 \uparrow^{N-\nu} \big((\nu+2) + \cdots + (N-1) + N\, \mathrm{c}^{**}(A)^{N+1}\big)$$

for all $\nu = 0, \dots, N$. This is clear for $\nu = N$. If $\nu = N-1$, then

$$\mathrm{c}^{**}\big(A^{(N-1)}\big) \leq 2^{\mathrm{c}^{**}(A)^{N+1}} \leq 2^{N\, \mathrm{c}^{**}(A)^{N+1}},$$

and if $\nu = N-2$, then

$$\mathrm{c}^{**}\big(A^{(N-2)}\big) \le 2^{\mathrm{c}^{**}(A^{(N-1)})^N} \le 2^{2^{N\,\mathrm{c}^{**}(A)^{N+1}}}.$$

Suppose we have proved the inequalities in question for some ν with $0 < \nu \le N-2$. Then by inductive hypothesis and Lemma 5.6, (3), we have

$$\mathrm{c}^{**}\big(A^{(\nu-1)}\big) \le 2^{\mathrm{c}^{**}(A^{(\nu)})^{\nu+1}} \le 2^{(2\uparrow^{N-\nu}((\nu+2)+\cdots+(N-1)+N\,\mathrm{c}^{**}(A)^{N+1}))^{\nu+1}} \le$$
$$2\uparrow^{N-\nu+1}\big((\nu+1)+(\nu+2)+\cdots+(N-1)+N\,\mathrm{c}^{**}(A)^{N+1}\big).$$

This shows the claim, which in turn easily yields the desired bound on $\mathrm{c}^*\big(A^{(\nu)}\big)$. For $\mathrm{c}^*\big(A^{(\nu)}, b^{(\nu)}\big)$, in case $b \in (D_{\mathrm{alg}})^m$, one argues similarly. □

Proceeding in this way we ultimately arrive at the last equation (S_0) over $(\mathbb{Z}_p)_{\mathrm{alg}}$, which will be solvable mod $p^{E-(e_1+\cdots+e_N)}$. We shall consider this situation in more detail.

5.6. **Construction of e_0.** In the proposition below, we let $A = A^{(0)}$, $m = m(0)$, $n = n(0)$, $b = b(0)$, so $A = (a_{ij})$ is an $m \times n$-matrix with entries $a_{ij} \in \mathbb{Z}_p$ which are *algebraic* over $\mathbb{Q}$, $b = [b_1, \ldots, b_m]^{\mathrm{tr}} \in \mathbb{Z}_p^m$, and suppose $\mathrm{rank}_{\mathbb{Q}_p}(A) = \mathrm{rank}_{\mathbb{Q}_p}(A, b)$, that is, the system $Ay = b$ is solvable in $\mathbb{Q}_p$.

Proposition 5.8. *Let $e_0 := \big[\mathrm{c}^*(A)/\log p\big]$. Then the system $Ay = b$ is solvable in $\mathbb{Z}_p$ if and only if it is solvable mod p^{e_0}. In this case, if in addition $b_1, \ldots, b_m \in (\mathbb{Z}_p)_{\mathrm{alg}}$, then the system $Ay = b$ has a solution $y = [y_1, \ldots, y_n]^{\mathrm{tr}}$ with $y_1, \ldots, y_n \in (\mathbb{Z}_p)_{\mathrm{alg}}$ and*

$$\Delta(y) \le \Delta(A, b), \quad h_{\max}(y) \le 4\,\mathrm{c}^*(A, b).$$

Proof. We may assume $m = r = \mathrm{rank}_{\mathbb{Q}_p}(A)$. The discussion of Hermann's method earlier in this section shows that for solvability of $Ay = b$ in $\mathbb{Z}_p$ it suffices to have solvability of the reduced system $\overline{A}y = \overline{b}$ in $\mathbb{Z}_p/\delta\mathbb{Z}_p$, where δ is a non-zero $r \times r$-minor of A. Now $v_p(\delta) \le e(A) \le e_0$, so $\mathbb{Z}_p/\delta\mathbb{Z}_p = \mathbb{Z}/p^e\mathbb{Z}$ for some $e \le e_0$. This shows the first statement. Suppose now that $Ay = b$ is solvable in $\mathbb{Z}_p$, and $b_1, \ldots, b_m$ are all algebraic over $\mathbb{Q}$. By applying p-desingularization to the pair (A, b) we transform $Ay = b$ into a system

$$\begin{bmatrix} \delta & & & c_{1,r+1} & \cdots & c_{1,n} \\ & \delta & & c_{2,r+1} & \cdots & c_{2,n} \\ & & \ddots & \vdots & \ddots & \vdots \\ & & \delta & c_{r,r+1} & \cdots & c_{rn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_r \end{bmatrix}$$

with the same solutions in $\mathbb{Z}_p$, where $c_{ij}, d_i \in (\mathbb{Z}_p)_{\mathrm{alg}}$ for $1 \le i \le r < j \le n$ and $\delta \in (\mathbb{Z}_p)_{\mathrm{alg}}$ is a unit, whose heights are bounded from above by $2\,\mathrm{c}^*(A, b)$. This system has a solution $y = [y_1, \ldots, y_n]^{\mathrm{tr}}$ given by $y_j = d_j/\delta$ for $1 \le j \le r$ and $y_j = 0$ for $r < j \le n$, and we have

$$h(y_j) \le h(d_j) + h(\delta) \le 4\,\mathrm{c}^*(A, b)$$

for $1 \le j \le r$, as required. □

We return to the general case, i.e., $N \ge 0$. By the proposition above, our system $Ay = b$ is solvable if the system $A^{(0)}y^{(0)} = b^{(0)}$ is solvable mod p^{e_0} where $e_0 = \big[\mathrm{c}^*(A^{(0)})/\log p\big]$. So it suffices to take the exponent E such that

$$E \ge \sum_{\nu=0}^{N} \big[\mathrm{c}^*(A^{(\nu)})/\log p\big].$$

By Lemma 5.7, we have

$$\mathrm{c}^*\big(A^{(\nu)}\big) = 2\uparrow^{N-\nu} O(1)^N\,\mathrm{c}^*(A)^{N+1} \qquad \text{for all } \nu = 0,\dots,N,$$

and hence we see that for some universal constant $C > 0$,

$$E(N,d,h,m) := 2\uparrow^N \big[C^N\big(m^2d(h+1)\big)^{N+1}\big]$$

has the properties required in Theorem 5.1. Suppose now that $b \in (D_{\mathrm{alg}})^m$. Then

$$\mathrm{c}^*\big(A^{(\nu)}, b^{(\nu)}\big) = 2\uparrow^{N-\nu} O(1)^N\,\mathrm{c}^*(A,b)^{N+1} \qquad \text{for all } \nu = 0,\dots,N.$$

Let $y^{(0)} \in (\mathbb{Z}_p)^{n(0)}_{\mathrm{alg}}$ be a solution to (S_0) with

$$\Delta\big(y^{(0)}\big) \le \Delta\big(A^{(0)}, b^{(0)}\big), \quad h_{\max}\big(y^{(0)}\big) \le 4\,\mathrm{c}^*(A^{(0)}, b^{(0)})$$

as in Proposition 5.8, and $y^{(1)},\dots,y^{(N)} = y$ the corresponding solutions to the systems (S_1),…, (S_N), respectively. Using Proposition 5.5, we obtain:

$$\Delta(y) = 2\uparrow^N O(1)^N\,\mathrm{c}^*(A,b)^{N+1}. \tag{5.10}$$

We leave it to the reader to deduce a similar bound on $h_{\max}(y)$ (involving N, m, n, $\Delta(A)$ and $h_{\max}(A)$). This finishes the proof of Theorem 5.1.

6. Criteria for ideal membership

In this section we show how the results of the previous section give rise to the computation of a function $(N,\beta) \mapsto e(N,\beta)$ with the properties discussed in the introduction, and we prove Theorem 0.3, both in a slightly more general situation:

Theorem 6.1. *Let $A \in \mathbb{Z}[X]^{m\times n}$, $b \in \mathbb{Z}[X]^m$, with $\deg_{(X)} A \le d$, $\log||A||_\infty \le h$, where $d \in \mathbb{N}$, $h \in \mathbb{R}$, $d, h > 0$. There exist positive integers δ and E_1, E_2 with*

$$E_1 = 2\uparrow^N O(1)^N\big(m^2(2d+h+1)\big)^{N+1},$$
$$E_2 = (2md)^{2^{O(N\log(N+1))}}(h+1)^{2N+1},$$

having the following properties:

1. *the system $Ay = b$ has a solution in $\mathbb{Z}[X]$ if and only if $Ay = b$ has a solution in $\mathbb{Q}[X]$ and a solution modulo δ^{E_1};*
2. *if $\deg b \le d$ and $||b||_\infty \le h$, then the system $Ay = b$ has a solution in $\mathbb{Z}[X]$ if and only if $Ay = b$ has a solution in $\mathbb{Q}[X]$ and a solution modulo δ^{E_2} of degree at most $(2md)^{2^{O(N\log(N+1))}}(h+1)$.*

We give the proof of this theorem after some preliminary remarks. In the final subsection, we also show Proposition 0.4.

6.1. Preliminaries.

Fix a commutative ring R. Let $A = (a_{ij})$ be an $m \times n$-matrix with entries in $R[X]$ and $b = [b_1,\dots,b_m]^{\mathrm{tr}}$ with $b_1,\dots,b_m \in R[X]$, and suppose the a_{ij} and the b_i all have total degree $\le d$. Let λ,μ,ν range over $\mathbb{N}^N$ and write $a_{ij} = \sum_\mu a_{ij,\mu}X^\mu$ and $b_i = \sum_\lambda b_{i,\lambda}X^\lambda$ with $a_{ij,\mu}, b_{i,\lambda} \in R$. Fix $\gamma \in \mathbb{N}$ and let $y = [y_1,\dots,y_n]^{\mathrm{tr}}$, where

$$y_j = \sum_{|\nu|\le\gamma} y_{j,\nu}X^\nu$$

with new indeterminates $y_{j,\nu}$ ranging over R. A polynomial in $X_1, \dots, X_N$ of degree at most d has at most $M(N,d) = \binom{N+d}{N}$ monomials. Hence for every R-algebra S, the solutions in $S[X]$ of every equation

$$a_{i1}y_1 + \cdots + a_{in}y_n = b_i \qquad (1 \le i \le m)$$

such that $\deg y_j \le \gamma$ for all $j = 1, \dots, n$ are in one-to-one correspondence with the solutions in S of the system

$$\sum_{\mu+\nu=\lambda} \sum_j a_{ij,\mu} y_{j,\nu} = b_{i,\lambda} \qquad (|\lambda| \le \gamma + d)$$

of $M(N, \gamma + d)$ equations in the $n \cdot M(N,\gamma)$ many variables $y_{j,\nu}$, with coefficients $a_{ij,\mu}$, $b_{i,\lambda}$ in R. So the entire system $Ay = b$ over $R[X]$ may be replaced by a certain system

$$A^* y^* = b^*, \qquad y^* = (y_{j,\nu})_{1 \le j \le n,\ |\nu| \le \gamma}$$

of $m \cdot M(N, \gamma + d)$ equations over R whose solutions are in one-to-one correspondence with the solutions to $Ay = b$ of degree at most γ, uniformly for all R-algebras. Note that if R is a subring of $\mathbb{Q}$, then $||A^*||_\infty = ||A||_\infty$ and $||b^*||_\infty = ||b||_\infty$. We use this discussion to show:

Lemma 6.2. *Let $d, h \in \mathbb{N}$. There exist positive integers γ_0, γ_1 with*

$$\gamma_0 \le (2md)^{2^{O(N \log(N+1))}} (h+1)^{2N+1}, \quad \gamma_1 \le (2md)^{2^{O(N \log(N+1))}} (h+1)$$

and having the following properties: Let $A = (a_{ij})$ be a non-zero $m \times n$-matrix and $b = [b_1, \dots, b_m]^{\mathrm{tr}}$ be a vector with entries $a_{ij}, b_i \in \mathbb{Z}[X]$ of degree at most d and $\log ||a_{ij}||_\infty, \log ||b_i||_\infty \le h$. Then for every prime number p, the system $Ay = b$ has a solution in $\mathbb{Z}_{(p)}[X]$ if and only if $Ay = b$ has a solution modulo p^{γ_0} of degree at most γ_1,

Proof. By Theorem 8.6 in [8], the system $Ay = b$ has a solution in $\mathbb{Z}_{(p)}[X]$ if and only if it has such a solution of degree at most γ_1. As we've seen above, there exists a certain system $A^*y^* = b^*$ with coefficients in $\mathbb{Z}$, consisting of $m \cdot M(N, \gamma_1 + d)$ equations in $n \cdot M(N, \gamma_1)$ unknowns, with the following properties: the solutions to $Ay = b$ in $\mathbb{Z}_{(p)}[X]$ of degree $\le \gamma_1$ are in one-to-one correspondence with the solutions of $A^*y^* = b^*$ in $\mathbb{Z}_{(p)}$, and for any $e \ge 1$, the solutions to $\overline{A}y = \overline{b}$ in $(\mathbb{Z}/p^e\mathbb{Z})[X]$ of degree $\le \gamma_1$ are in one-to-one correspondence with the solutions of $\overline{A^*}y^* = \overline{b^*}$ in $\mathbb{Z}/p^e\mathbb{Z}$. By Lemma 5.8, the system $A^*y^* = b^*$ has a solution in $\mathbb{Z}_{(p)}$ provided it has a solution modulo p^e where

$$e = \Big[\big(m \cdot M(N, \gamma_1 + d)\big)^2 \big(h_{\max}(A^*) + 1\big) \Big].$$

Now $h_{\max}(A^*) = \log ||A, b||_\infty \le h$ and for $d > 0$

$$m^2 M(N, \gamma_1 + d)^2 \le m^2 (\gamma_1 + d + 1)^{2N} \le (2md)^{2^{O(N \log(N+1))}} (h+1)^{2N},$$

yielding the lemma. □

Proof of Theorem 6.1. Let A and b be as in the statement of Theorem 6.1. Let M be a submodule of the free $\mathbb{Z}[X]$-module $\mathbb{Z}[X]^m$. Given a ring extension R of $\mathbb{Z}[X]$ we denote by MR the submodule of the free R-module R^m generated by M.

Lemma 6.3. *From a given finite set of generators for M one can compute a positive integer δ such that*

$$M\mathbb{Q}[X] \cap \mathbb{Z}[X]^m = (M : \delta) := \big\{ y \in \mathbb{Z}[X]^m : \delta y \in M \big\}.$$

If $p_1, \dots, p_K$ are the distinct prime factors of a number δ with this property, then

$$\begin{aligned} M &= M\mathbb{Q}[X] \cap M\mathbb{Z}_{(p_1)}[X] \cap \cdots \cap M\mathbb{Z}_{(p_K)}[X] \\ &= M\mathbb{Q}[X] \cap M\mathbb{Z}_{p_1}\langle X\rangle \cap \cdots \cap M\mathbb{Z}_{p_K}\langle X\rangle. \end{aligned} \tag{6.1}$$

Proof. The existence of the integer δ is a consequence of the fact that the $\mathbb{Z}[X]$-module $M\mathbb{Q}[X]\cap\mathbb{Z}[X]^m$ is finitely generated; its computability is established in [8], Corollary 3.5. The first equation in (6.1) is shown using the argument on pp. 409–410 of [8]. The second equation is by [8], Lemma 2.6. □

Let δ and $p_1, \dots, p_K$ be as in the lemma, applied to $M =$ the submodule of $\mathbb{Z}[X]^m$ generated by the columns of A. Then $Ay = b$ has a solution in $\mathbb{Z}[X]$ if and only if $Ay = b$ has a solution in $\mathbb{Q}[X]$ and a solution in $\mathbb{Z}_{p_k}\langle X\rangle$ for all $k = 1, \dots, K$. For given positive integer E, by the Chinese Remainder Theorem we have a surjection

$$\begin{aligned} (\mathbb{Z}/\delta^E\mathbb{Z})[X] &\to (\mathbb{Z}/p_1^E\mathbb{Z})[X] \times \cdots \times (\mathbb{Z}/p_k^E\mathbb{Z})[X], \\ a \bmod \delta^E &\mapsto \big(a \bmod p_1^E, \dots, a \bmod p_k^E\big). \end{aligned}$$

Combining this with Theorem 5.1 and Lemma 6.2 yields Theorem 6.1.

6.2. **Linear algebra modulo prime powers.** In this final part of the paper we show how degree bounds for solving linear equations in polynomial rings over $\mathbb{Z}/p\mathbb{Z}$, where p is a prime, entail degree bounds for solving linear equations in polynomial rings over $\mathbb{Z}/p^e\mathbb{Z}$ where $e > 1$. For the moment, we let more generally D be a commutative ring, p be a non-zero divisor of D, and $e \geq 1$. Let $A = (a_{ij}) \in D^{m\times n}$ and $b = [b_1, \dots, b_m]^{\mathrm{tr}} \in D^m$. We want to consider the following two problems:

$H_e(A)$**:** *Find a finite set of generators for the submodule*

$$\mathrm{S}_e(A) := \big\{y \in D^n : Ay = 0 \bmod p^e\big\}$$

of the free D-module D^n.

$I_e(A, b)$**:** *Determine whether the (possibly inhomogeneous) system $Ay = b$ is solvable mod p^e, and if it is, find an element $y \in D^n$ with $Ay = b \bmod p^e$.*

We will show, by induction on e, that $H_e(A)$ and $I_e(A, b)$ can be reduced to numerous instances of the two simpler problems $H_1(A)$ and $I_1(A, b)$, for various matrices A and vectors b with entries from D. Suppose we can solve the problems $H_1(A)$ and $I_1(A, b)$ for all matrices $A \in D^{m\times n}$ and vectors $b \in D^m$, with $m, n \geq 1$. Suppose $e > 1$, and let

$$g_1 = \begin{bmatrix} g_{11} \\ \vdots \\ g_{n1} \end{bmatrix}, \dots, g_r = \begin{bmatrix} g_{1r} \\ \vdots \\ g_{nr} \end{bmatrix}$$

be vectors in D^n generating the submodule $\mathrm{S}_1(A)$ of D^n. Let $A^{(1)} \in D^{m\times r}$ be such that $AG = pA^{(1)}$, where $G = (g_{jk}) \in D^{n\times r}$. By induction hypothesis we can solve $H_{e-1}\big(A^{(1)}\big)$, that is, we can find

$$h_1 = \begin{bmatrix} h_{11} \\ \vdots \\ h_{r1} \end{bmatrix}, \dots, h_s = \begin{bmatrix} h_{1s} \\ \vdots \\ h_{rs} \end{bmatrix} \in D^r$$

generating $\mathrm{S}_{e-1}\big(A^{(1)}\big)$. Let $H = (h_{kl}) \in D^{r\times s}$. It is now easy to verify that the column vectors of the matrix GH form a set of generators of $\mathrm{S}_e(A)$. So we have solved $H_e(A)$.

For $I_e(A,b)$, we first determine whether $Ay = b$ is solvable mod p (using $I_1(A,b)$). If it is not, we are already done: then $Ay = b$ is not solvable mod p^e. Suppose $Ay = b$ *is* solvable mod p, and let $z \in D^n$, $b^{(1)} \in D^m$ be such that $Az = b - pb^{(1)}$. Let $G = (g_{jk}) \in D^{n\times r}$ as above be a matrix with entries from D whose column vectors generate $\mathrm{S}_1(A)$, and $A^{(1)} \in D^{m\times r}$ with $AG = pA^{(1)}$.

Lemma 6.4. *The system $Ay = b$ is solvable mod p^e if and only if the system $A^{(1)}y^{(1)} = b^{(1)}$ is solvable mod p^{e-1}.*

Proof. Given $y^{(1)} \in D^r$, we have the equivalence

$$A^{(1)}y^{(1)} = b^{(1)} \mod p^{e-1} \qquad \Longleftrightarrow \qquad A(Gy^{(1)} + z) = b \mod p^e.$$

So if $A^{(1)}y^{(1)} = b^{(1)} \bmod p^{e-1}$, then $y = Gy^{(1)} + z$ solves $Ay = b \bmod p^e$. Conversely, if $y \in D^n$ is such that $Ay = b \bmod p^e$, then $Ay = b \bmod p$, so $y - z \in \mathrm{S}_1(A)$ and thus $y = z + Gy^{(1)}$ for some $y^{(1)} \in D^r$; so $A^{(1)}y^{(1)} = b^{(1)} \bmod p^{e-1}$. □

Using the inductive hypothesis for $I_{e-1}\big(A^{(1)}, b^{(1)}\big)$, we can now determine whether $A^{(1)}y^{(1)} = b^{(1)}$ has a solution mod p^{e-1}. This solves $I_e(A,b)$. (Similar algorithms were used in [37] for deciding the universal theory of commutative rings of fixed positive characteristic.)

Suppose now that R is commutative ring, $D = R[X]$, and $p \in R$ is a non-zero divisor. Assume moreover that p generates a maximal ideal of R. By Hermann's method applied to the polynomial ring $(R/pR)[X]$ over the field R/pR (see [8], Section 3) we can choose the matrix G above of degree $\le (2md)^{2^N}$. Then $A^{(1)}$ has degree $\le d + (2md)^{2^N}$, and if the $h_1, \ldots, h_s$ are of degree $\le d'$, then GH is of degree $\le d' + (2md)^{2^N}$. Let $\gamma_e(N,d,m)$ be the smallest natural number such that $\mathrm{S}_e(A)$ for $A \in D^{m\times n}$ of degree $\le d$ is generated by elements of degree $\le \gamma_e(N,d,m)$. We have the recursive relation

$$\gamma_e(N,d,m) \le \gamma_{e-1}\big(N, d + (2md)^{2^N}, m\big) + (2md)^{2^N} \qquad \text{for } e > 1.$$

Using that γ_{e-1} is increasing in the second variable and that $d + (2md)^{2^N} \le (2md)^{2^N+1}$ for $d > 0$, we see that for $d, e > 0$,

$$\gamma_e(N,d,m) \le (2md)^{2^N} + (2md)^{2^N(2^N+2)} + \cdots + (2md)^{2^N(2^N+2)^{e-1}}.$$

Since $(2md)^{2^N(2^N+2)^i} \le (2md)^{2^N(2^N+2)^{e-1}}$ for $i = 0, \ldots, e-1$, we get for $N, d, e > 0$,

$$\gamma_e(N,d,m) \le e(2md)^{2^N(2^N+2)^{e-1}} \le e(2md)^{2^{N+1}(2^{N+1})^{e-1}} \le e(2md)^{2^{e(N+1)}}.$$

Similarly one deduces (using the proof of Lemma 6.4) that if $A \in D^{m\times n}$ and $b \in D^m$ are of degree $\le d$, and the system of congruence equations $Ay = b \bmod p^e$ is solvable in D, then it has such a solution $y \in D^n$ having degree at most $\gamma_e(N,d,m)$. Note that γ_e neither depends on R nor on p. We have proved:

Proposition 6.5. *Let R be a commutative ring, $p \in R$ be a non-zero divisor which generates a maximal ideal of R, and $D = R[X]$. Let $A \in D^{m\times n}$ and $b \in D^m$ be of degree $\le d$, and $e \ge 1$.*

(1) *There exist $y^{(1)}, \ldots, y^{(r)} \in D^n$ of degree $\le e(2md)^{2^{e(N+1)}}$ such that*

$$Ay^{(k)} = 0 \bmod p^e \qquad \text{for } k = 1, \ldots, r,$$

and every $y \in D^n$ with $Ay = 0 \bmod p^e$ is a D-linear combination of the $y^{(k)}$s.

(2) *If there exists $y \in D^n$ such that $Ay = b \bmod p^e$, then there exists such a y of degree $\le e(2md)^{2^{e(N+1)}}$.* □

Let now R be a unique factorization domain, $\delta \neq 0$ a non-unit of R, and $\delta = p_1^{e_1} \cdots p_k^{e_k}$ a decomposition of δ into powers of pairwise non-associated irreducibles $p_1, \ldots, p_k$, where $e_1, \ldots, e_k \geq 1$. We put $e(\delta) = \max\{e_1, \ldots, e_k\}$. Proposition 6.5 (applied to the $p_i^{e_i}$) and the Chinese Remainder Theorem imply:

Corollary 6.6. *Let $D = R[X]$ and $A \in D^{m\times n}$, $b \in D^m$ be of degree $\leq d$.*

(1) *There exist $y^{(1)}, \ldots, y^{(r)} \in D^n$ of degree $\leq e(\delta)(2md)^{2^{e(\delta)(N+1)}}$ such that*

$$Ay^{(k)} = 0 \bmod \delta \qquad \text{for } k = 1, \ldots, r,$$

and every $y \in D^n$ with $Ay = 0 \bmod \delta$ is a D-linear combination of the vectors $y^{(1)}, \ldots, y^{(r)}$.

(2) *If there exists $y \in D^n$ such that $Ay = b \bmod \delta$, then there exists such a y of degree $\leq e(\delta)(2md)^{2^{e(\delta)(N+1)}}$.* □

Taking $R = \mathbb{Z}$ in this proposition and noting that $e(\delta) \leq \log_2 |\delta|$ yields Proposition 0.4.

REFERENCES

1. M. E. Alonso, T. Mora, and M. Raimondo, *On the complexity of algebraic power series*, Applied Algebra, Algebraic Algorithms and Error-correcting Codes, Tokyo, 1990 (S. Sakata, ed.), Lecture Notes in Comput. Sci., vol. 508, Springer, Berlin, 1991, pp. 197–207.
2. ———, *A computational model for algebraic power series*, J. Pure Appl. Algebra **77** (1992), no. 1, 1–38.
3. F. Amoroso, *Algebraic numbers close to* 1 *and variants of Mahler's measure*, J. Number Theory **60** (1996), no. 1, 80–96.
4. S. Ju. Arakelov, *An intersection theory for divisors on an arithmetic surface*, Izv. Akad. Nauk SSSR Ser. Mat. **38** (1974), 1179–1192.
5. E. Artin and G. Whaples, *Axiomatic characterization of fields by the product formula for valuations*, Bull. Amer. Math. Soc. **51** (1945), 469–492.
6. C. Ayoub, *On constructing bases for ideals in polynomial rings over the integers*, J. Number Theory **17** (1983), no. 2, 204–225.
7. M. Aschenbrenner, *Ideal Membership in Polynomial Rings over the Integers*, Ph.D. thesis, University of Illinois at Urbana-Champaign, 2001.
8. ———, *Ideal membership in polynomial rings over the integers*, J. Amer. Math. Soc. **17** (2004), no. 2, 407–441 (electronic).
9. J. Bochnak, M. Coste, and M.-F. Roy, *Real Algebraic Geometry*, Ergeb. Math. Grenzgeb. (3), vol. 36, Springer-Verlag, Berlin, 1998.
10. S. Bosch, U. Güntzer, and R. Remmert, *Non-Archimedean Analysis. A Systematic Approach to Rigid Analytic Geometry*, Grundlehren Math. Wiss., vol. 261, Springer-Verlag, Berlin, 1984.
11. N. Bourbaki, *Éléments de Mathématique. Algèbre Commutative*, Hermann, Paris, 1964.
12. A. Clivio, *Algorithmic aspects of $\mathbf{Z}[x_1, \ldots, x_n]$ with applications to tiling problems*, Z. Math. Logik Grundlag. Math. **36** (1990), no. 6, 493–515.
13. M. Coste, J. Ruiz, and M. Shiota, *Uniform bounds on complexity and transfer of global properties of Nash functions*, J. Reine Angew. Math. **536** (2001), 209–235.
14. L. van den Dries, *On the elementary theory of restricted elementary functions*, J. Symbolic Logic **53** (1988), no. 3, 796–808.
15. G. Everest and T. Ward, *Heights of Polynomials and Entropy in Algebraic Dynamics*, Universitext, Springer-Verlag London Ltd., London, 1999.
16. H. Friedman, *The Ackermann function in elementary algebraic geometry*, manuscript, December 1999.
17. G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1988, Reprint of the 1952 edition.
18. G. Hermann, *Die Frage der endlich vielen Schritte in der Theorie der Polynomideale*, Math. Ann. **95** (1926), 736–788.
19. J. P. Jones, *Basis for the Kalmár elementary functions*, Number Theory and Applications, Banff, AB, 1988 (Richard A. Mollin, ed.), NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., vol. 265, Kluwer Acad. Publ., Dordrecht, 1989, pp. 435–444.
20. E. Kani, *Nonstandard Diophantische Geometrie, insbesondere Satz von Mordell-Weil*, Ph.D. thesis, Universität Heidelberg, 1978.

21. ———, *Eine Verallgemeinerung des Satzes von Castelnuovo-Severi*, J. Reine Angew. Math. **318** (1980), 179–220.
22. J.-P. Lafon, *Séries formelles algébriques*, C. R. Acad. Sci. Paris **260** (1965), 3238–3241.
23. S. Lang, *Fundamentals of Diophantine Geometry*, Springer-Verlag, Berlin-New York, 1983.
24. ———, *Algebra*, third ed., Graduate Texts in Mathematics, vol. 211, Springer-Verlag, New York, 2002.
25. F. Lazzeri and A. Tognoli, *Alcune proprietà degli spazi algebrici*, Ann. Scuola Norm. Sup. Pisa (3) **24** (1970), 597–632.
26. K. Mahler, *An application of Jensen's formula to polynomials*, Mathematika **7** (1960), 98–100.
27. M. Mignotte, *Approximation des nombres algébriques par des nombres algébriques de grand degré*, Ann. Fac. Sci. Toulouse Math. (5) **1** (1979), no. 2, 165–170.
28. G. Moreno Socías, *Length of polynomial ascending chains and primitive recursiveness*, Math. Scand. **71** (1992), no. 2, 181–205.
29. D. G. Northcott, *An inequality in the theory of arithmetic on algebraic varieties*, Proc. Cambridge Philos. Soc. **45** (1949), 502–509.
30. P. Philippon, *Dénominateurs dans le théorème des zéros de Hilbert,* Acta Arith. **58** (1991), no. 1, 1–25.
31. , ———, *Critères pour l'indépendance algébrique dans les anneaux diophantiens,* C. R. Acad. Sci. Paris Sér. I Math. **315** (1992), no. 5, 511–515.
32. ———, *Sur des hauteurs alternatives. II*, Ann. Inst. Fourier (Grenoble) **44** (1994), no. 4, 1043–1065.
33. R. Ramanakoraisina, *Bezout theorem for Nash functions*, J. Pure Appl. Algebra **61** (1989), no. 3, 295–301.
34. ———, *Complexité des fonctions de Nash*, Comm. Algebra **17** (1989), no. 6, 1395–1406.
35. M. Raynaud, *Anneaux Locaux Henséliens*, Lecture Notes in Mathematics, vol. 169, Springer-Verlag, Berlin, 1970.
36. J. Schmid, *On the Degree Complexity of Hilbert's 17th Problem and the Real Nullstellensatz*, Habilitationsschrift, Universität Dortmund, 1998.
37. H. Simmons, *The solution of a decision problem for several classes of rings*, Pacific J. Math. **34** (1970), 547–557.

E-mail address: maschenb@math.uic.edu
URL: http://www.math.uic.edu/∼maschenb

DEPARTMENT OF MATHEMATICS, STATISTICS, AND COMPUTER SCIENCE, UNIVERSITY OF ILLINOIS AT CHICAGO, 851 S. MORGAN ST. (M/C 249), CHICAGO, IL 60607-7045, U.S.A.