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 1
1. Absolute values and norms of polynomials 5
2. Height functions 10
3. A height function on the algebraic closure of $\mathbb{Q}(X)$ 26
4. Restricted power series 32
5. Hermann’s method for restricted power series 46
6. Criteria for ideal membership 56
References 60

INTRODUCTION

Let $f_1, \ldots, f_n \in \mathbb{Z}[X]$, where $X = (X_1, \ldots, X_N)$ 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 $f_1, \ldots, f_n$ (implicit in [37]): there exist positive integers $\delta$ and $e$ such that for every polynomial $f_0 \in \mathbb{Z}[X]$:

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

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 $\delta$ and $e$ can be easily seen as follows. Put $I := (f_1, \ldots, f_n)\mathbb{Z}[X]$, 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
positive integer $\delta$ 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

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

With this $e$ it is easy to check that

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

hence $\delta$ and $e$ have the required properties.

It is well-known that an integer $\delta$ with the property that $J = I : (\delta)$ can be effectively computed from $f_1, \ldots, f_n$. In fact, we may take $\delta = P(c)$ for some polynomial $P(C) \in \mathbb{Z}[C]$ in the coefficient tuple $c$ of $f_1, \ldots, f_n$. 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 $\delta_1 = \delta$ 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 $\delta_1$ was given by Clivio [12] (based on earlier work of Ayoub [6]).

The computability of $\delta$ 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 $(g_1, \ldots, g_n)$ 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 $e = 1, 2, \ldots$, whether $f_0 \in (f_1, \ldots, f_n)(\mathbb{Z}/\delta^e \mathbb{Z})[X]$. 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 $\delta$), where Hermann’s method applies. The first procedure stops if $f_0 \notin I$, 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 $\delta$ 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 $f_1, \ldots, f_n$ are polynomials in $\mathbb{Z}[X] = \mathbb{Z}[X_1, \ldots, X_N]$ with $\deg f_1, \ldots, \deg f_n \leq \beta$ and $||f_1||_\infty, \ldots, ||f_n||_\infty \leq \beta$, 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 $\beta$. 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 $p_1, \ldots, p_K$ be the distinct prime factors of $\delta$. A short argument using the Euclidean Algorithm (see [8], pp. 409–410) shows that then for $f_0 \in \mathbb{Z}[X]$:

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

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 $\mathbb{Z}_p\langle X \rangle_{\text{alg}}$ be the subring of $\mathbb{Z}_p\langle X \rangle$ consisting of the restricted power series that are algebraic over the rational function field $\mathbb{Q}\langle X \rangle$. (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

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

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 $\mathbb{Z}_p\langle X \rangle_{\text{alg}}$, in (0.3) we may further replace $\mathbb{Z}_p\langle X \rangle$ by its subring $\mathbb{Z}_p\langle X \rangle_{\text{alg}}$. Thus we can improve (0.2) as follows:

$$f_0 \in (f_1, \ldots, f_n)\mathbb{Z}[X] \iff f_0 \in (f_1, \ldots, f_n)\mathbb{Q}[X] \text{ and } f_0 \in (f_1, \ldots, f_n)\mathbb{Z}_p\langle X \rangle_{\text{alg}} \text{ for all } k. \quad (0.4)$$

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 $\mathbb{Z}_p\langle X \rangle_{\text{alg}}$; for example, the ring $\mathbb{Z}_p\langle X \rangle_{\text{alg}}$ is closed under Weierstraß Division in $\mathbb{Z}_p\langle X \rangle$. Moreover, in $\mathbb{Z}_p\langle X \rangle_{\text{alg}}$ 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

$$h : \mathbb{Q}(X)_{\text{alg}} \to \mathbb{R} \geq 0$$

on the algebraic closure $\mathbb{Q}(X)_{\text{alg}}$ 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 $\mathbb{Q}(X)_{\text{alg}}$; 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

$$h(f) = \deg_X f + \cdots + \deg_{X_N} f + \int_0^1 \cdots \int_0^1 \max \{ \log |f(e^{2\pi i \theta_1}, \ldots, e^{2\pi i \theta_N})|, 0 \} d\theta_1 \cdots d\theta_N,$$
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 $\mathbb{Q}(X)_{\text{alg}}$, and also has the following finiteness property: there are only finitely many $f \in \mathbb{Q}(X)_{\text{alg}}$ with given bounds on the height $h(f)$ and the degree $\Delta(f) := \max \{ \deg f, \deg g \}$ 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 $g \in \mathbb{Z}_p(X)$ is said to be regular in $\mathbb{Z}[X]$ of degree $s \in \mathbb{N}$ if there exist a monic polynomial $g_0 \in \mathbb{Z}_p[X]/[X_N]$ of degree $s$ and a unit $u$ of $\mathbb{Z}_p$ such that $g - u g_0 \in \mathbb{Z}_p(X)$. Here and below $X' := (X_1, \ldots, X_{N-1})$.

**Theorem 0.1.** Let $f, g \in \mathbb{Z}_p(X)_{\text{alg}}$, and suppose that $g$ is regular of degree $s > 0$ in $X_N$. Let $(q, r)$ be the unique pair consisting of $q \in \mathbb{Z}_p(X)_{\text{alg}}$ and $r \in \mathbb{Z}_p(X')_{\text{alg}}[X_N]$ such that $f = q g + r$, and bounds for the complexity of quotient and remainder in the Weierstraß Division Theorem above can be used to adapt Hermann’s method to treat ideal membership problems for ideals in $\mathbb{Z}_p(X)_{\text{alg}}$. Here, the role played by the Euclidean Algorithm (for polynomials) in Hermann’s method is taken over by the effective Weierstraß Division Theorem above can be used to adapt Hermann’s method to treat ideal membership problems for ideals in $\mathbb{Z}_p(X)_{\text{alg}}$. Here, the role played by the Euclidean Algorithm (for polynomials) in Hermann’s method is taken over by the

where $\beta > 0$ allows us to show the following “effective” $p$-adic analogue of (0.1). Here $\overline{f}$ denotes the image of $f \in \mathbb{Z}_p(X)$ under the canonical homomorphism $\mathbb{Z}_p(X) \to \mathbb{Z}_p(X)/p^E \mathbb{Z}_p(X) \cong (\mathbb{Z}/p^E \mathbb{Z})[X]$. 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

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

such that for all $f_0 \in \mathbb{Z}_p(X)$:

$$ f_0 \in (f_1, \ldots, f_n)\mathbb{Z}_p(X) \iff f_0 \in (f_1, \ldots, f_n)\mathbb{Q}_p(X) \text{ and } \overline{f}_0 \in (\overline{f}_1, \ldots, \overline{f}_n)(\mathbb{Z}/p^E \mathbb{Z})[X]. $$
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 $D = (2\beta)^{2O(N \log(N+1))}$ 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

$$f_0 = f_1g_1 + \cdots + f_ng_n \mod \delta^D.$$ 

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

$$f_0 = f_1g_1 + \cdots + f_ng_n \mod a.$$ 

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 $\mathbb{Z}_p(X)$, 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. \textbf{Absolute values.} Let $K$ be a field and $\| \cdot \|$ be an \textbf{absolute value} on $K$, that is, a function $x \mapsto |x|: K \to \mathbb{R}_{\geq 0}$ 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 \textbf{triangle inequality}

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

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 \textbf{ultrametric triangle inequality}

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

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

$$(x, y) \mapsto |x - y|: K \times K \to \mathbb{R}_{\geq 0}$$

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 \textbf{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 \textbf{additive absolute value}

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

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

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

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

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

as the corresponding valuation ring and its maximal ideal, respectively. Conversely, if $v: K^\times \to \mathbb{R}$ 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:

\textbf{Example 1.1.} Suppose that $K = \text{Frac}(R)$ is the fraction field of a unique factorization domain $R$. Every irreducible element $p \in R$ determines the (discrete) $p$-adic valuation $v_p: K^\times \to \mathbb{Z}$ 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 \textbf{essential valuations of $R$}. Each essential valuation $v_p$ determines an \textbf{essential absolute value} $| \cdot |_{v_p}$ on $K$ by $|x|_{v_p} := \exp(-v_p(x))$ for $x \in K^\times$. For instance, the essential valuations of $R = \mathbb{Z}$ are the $p$-adic valuations on
where inequality (see [17], § value on the algebraic closure \( K \))
denotes the usual absolute value on \( \mathbb{Q} \).

We first introduce some notations used throughout the paper. Let polynomials, the most important among them the so-called logarithmic Mahler measure.

Theorem, we may assume that \( K \) is archimedean, then either \((\hat{K}, \cdot | \cdot) \cong (\mathbb{R}, \cdot | \cdot_p^p)\) for some real number \( r \in (0, 1) \), or \((\hat{K}, \cdot | \cdot) \cong (\mathbb{C}, \cdot | \cdot)\) for some real number \( r \in (0, 1) \). Here \( \cdot | \cdot_p^p \) and \( \cdot | \cdot ) \) 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 \( 1 \leq \varepsilon_v \leq 2 \), with \( \varepsilon_v = 1 \) if and only if \( v \) is ultrametric, and for all \( x, y \in K \) we have \( |x + y|_v \leq \varepsilon_v \max\{ |x|_v, |y|_v \} \). (All this is clear if \( v \) is ultrametric; if \( v \) is archimedean, use Ostrowski’s Theorem and the triangle inequality for \( \cdot | \cdot ) \).

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 \( r_1, \ldots, r_n \in \mathbb{R} \) are such that \( |a_{ij}|_v \leq r_j \) for all \( i, j = 1, \ldots, n \). Then:

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

**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 \( a_1, \ldots, a_n \). By Hadamard’s inequality (see [17], §2.12),

\[
|\det A|_v = |\det A|^r \leq \|a_1\|^r_2 \cdots \|a_n\|^r_2,
\]

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

\[
\|b\|_2 \leq \sqrt{n} \max\{ |b_1|, \ldots, |b_n| \}
\]

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 \( K_{\text{alg}} \) of \( K \), and any two such extensions \( | \cdot |_1 \) and \( | \cdot |_2 \) are conjugate, that is: there exists \( \sigma \in \text{Gal}(K_{\text{alg}}|K) \) such that \( |x|_1 = |\sigma(x)|_2 \) for all \( x \in K_{\text{alg}} \). This fact will turn out to be crucial for the construction of a height function on \( K_{\text{alg}} \) 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, \( X = (X_1, \ldots, X_N) \) an \( N \)-tuple of indeterminates over \( R \), and

\[
f(X) = \sum_\nu a_\nu X^\nu
\]

a polynomial in \( R[X] \). Here the multiindex \( \nu = (\nu_1, \ldots, \nu_N) \) ranges over \( \mathbb{N}^N \), \( X^\nu = X_1^{\nu_1} \cdots X_N^{\nu_N} \), and \( a_\nu \in R \). For \( i = 1, \ldots, N \), 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 \( \deg_X 0 = \deg 0 = -\infty < \)
\[ \mathbb{N}. \] 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 \( m = \max \{ \deg_{X_{1}} f, \ldots, \deg_{X_{N}} f \} \) we have
\[
m \leq \deg_{X} f \leq \deg_{(X)} f \leq Nm.
\] (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 \( |f|_{v} := \max_{\nu \in \nu} |a_{\nu}|_{v} \). 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:
\[
|f \cdot g|_{v} = |f|_{v} \cdot |g|_{v} \quad \text{for all } f, g \in K[X].
\] (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 \( f_{1}, \ldots, f_{n} \in K[X] \), we do have the following fundamental inequality (known as Gelfond's Lemma):
\[
|2^{-d} f_{1} \cdots f_{n}|_{v} \leq |f_{1}|_{v} \cdots |f_{n}|_{v} \leq |2^{d} f_{1} \cdots f_{n}|_{v}, \quad d = \deg_{(X)} f_{1} \cdots f_{n}
\] (1.7)
(See [23], Chapter 3, §2.) Note that (1.6) and (1.7) may be combined to the inequality
\[
\varepsilon v^{-d} \prod_{i=1}^{n} |f_{i}|_{v} \leq |f_{1} \cdots f_{n}|_{v} \leq \varepsilon v^{d} \prod_{i=1}^{n} |f_{i}|_{v}.
\] (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 \( \text{mon } f = \text{the number of monomials occurring in } f \) with a non-zero coefficient, so \( \text{mon } f \leq (N+d) \). 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. \( |f(x)|_{v} \leq \max_{i} |x_{i}|_{v}^{d} \cdot |f|_{v} \cdot \text{mon } f \) 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 \( |z| = \sqrt{\Re z} \), for \( z \in \mathbb{C} \). The maps
\[
f \mapsto ||f||_{1} := \left( \sum_{\nu} |a_{\nu}| \right)^{1/2}, \quad f \mapsto \|f\|_{2} := \left( \sum_{\nu} |a_{\nu}|^{2} \right)^{1/2}
\]
are norms on the \( \mathbb{C} \)-vector space \( \mathbb{C}[X] \) extending the absolute value on \( \mathbb{C} \), called the \( l_{1} \)-\textbf{norm} (or length) and the \( l_{2} \)-\textbf{norm} (or euclidean norm) on \( \mathbb{C}[X] \), respectively. We also denote the Gauß norm (sometimes called the \( l_{\infty} \)-\textbf{norm}) on \( \mathbb{C}[X] \) by \( | \cdot | = || \cdot ||_{\infty} \). These norms are connected by the following relations, valid for \( f \neq 0 \):
\[
\begin{align*}
||f||_{\infty} &\leq ||f||_{1} \leq (1 + \deg_{X_{1}} f) \cdots (1 + \deg_{X_{N}} f) ||f||_{1}, \\
||f||_{\infty} &\leq ||f||_{2} \leq (1 + \deg_{X_{1}} f)^{1/2} \cdots (1 + \deg_{X_{N}} f)^{1/2} ||f||_{1}, \\
||f||_{2} &\leq ||f||_{1} \leq (1 + \deg_{X_{1}} f)^{1/2} \cdots (1 + \deg_{X_{N}} f)^{1/2} ||f||_{2}.
\end{align*}
\] (1.9)

A more subtle measure for the complexity of a non-zero polynomial \( f \) is its \textbf{Mahler measure}
\[
m(f) = \int_{0}^{1} \cdots \int_{0}^{1} \log |f(e^{2\pi i \theta_{1}}, \ldots, e^{2\pi i \theta_{N}})|
\]
\[
d\theta_{1} \cdots d\theta_{N}. \quad (1.10)
\]

\[2\]
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 $a_{\nu}$ 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 $m^+(f) = m(f(X) + Y)$, 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:

$$m^+(f_1 \cdots f_n) \leq m^+(f_1) + \cdots + m^+(f_n), \quad (1.11)$$

$$m^+(f_1 + \cdots + f_n) \leq m^+(f_1) + \cdots + m^+(f_n) + \log n, \quad (1.12)$$

$$m^+(f^k) = km^+(f). \quad (1.13)$$

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

$$m(f) \leq m^+(f) \leq \log ||f||_1 \leq m(f) + d \log 2 \quad (1.14)$$

and

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

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 \( g_1(Y), \ldots, g_N(Y) \in \mathbb{Z}[Y] \). Then
\[
m^+ \left( f(g_1, \ldots, g_N) \right) \leq \log \| f \|_1 + d_1 m^+ (g_1) + \cdots + d_N m^+ (g_N),
\]
where \( d_1 = \deg_X f, \ldots, d_N = \deg_X f \).

Proof. Write \( f \) as in (1.4), with all \( a_\nu \in \mathbb{Z} \). Then for all \( z = (z_1, \ldots, z_N) \in \mathbb{C}^N \)
\[
| f(g_1(z), \ldots, g_N(z)) | \leq \sum_\nu |a_\nu| \max \{ |g_1(z)|^{d_1}, 1 \} \cdots \max \{ |g_N(z)|^{d_N}, 1 \},
\]
for \( 1 \leq j \leq N \). Thus
\[
\log^+ | f(g_1(z), \ldots, g_N(z)) | \leq \log \| f \|_1 + d_1 \log^+ | g_1(z) | + \cdots + d_N \log^+ | g_N(z) |.
\]
Taking \( z_j = e^{2\pi i \theta_j} \) with \( 0 \leq \theta_j \leq 1 \) for \( j = 1, \ldots, N \) and integrating both sides over \( [0, 1]^N \) with respect to \((\theta_1, \ldots, \theta_N)\) gives the desired inequality.

The first inequality in (1.14) is clear since \( \log x \leq \log^+ x \) for all \( x \in \mathbb{R}^+ \), 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 \( j = 1, \ldots, N \). 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 \( \mathbb{Q}(X_1, \ldots, X_N) \). (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 \( \mathbb{Q}(X_1, \ldots, X_N) \), 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, \( K = \text{Frac}(R) \), 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 \( M_{\text{alg}} \) be the set of absolute values on \( K_{\text{alg}} \) extending an absolute value in \( M \). Then \((K_{\text{alg}}, M_{\text{alg}})\) 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 \( K \subseteq L \subseteq K_{\text{alg}} \), we can make \( L \) into an arithmetic field \((L, M_L)\) with \((K, M) \subseteq (L, M_L)\) in a unique way by setting \( M_L := \{ v_L : v \in M \} \). (So \((K, M) \subseteq (K_{\text{alg}}, M_{\text{alg}})\) and \( M_{\text{alg}} = M_{K_{\text{alg}}} \).)
Definition 2.2. An arithmetic field $(K, M)$ is called classical if for each $a \in K^\times$, the support of $a$,
\[
\text{supp}(a) := \{ v \in M : v(a) \neq 0 \},
\]
is a finite set.

Examples 2.3. The arithmetic field $\mathbb{Q}$ is classical. If $R$ is a unique factorization domain, $K = \text{Frac}(R)$, 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
\[
\mathcal{D}^f(K, M) := \prod_{v \in M} \Gamma_v,
\]
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 $\Gamma_f$ is a linearly ordered additive group, the (additively written) group $\mathcal{D}^f(K, M)$ is a lattice-ordered group: for $D_1, D_2, D \in \mathcal{D}^f(K, M)$
\[
D_1 \leq D_2 \iff v(D_1) \leq v(D_2) \quad \text{for all } v \in M,
\]
and for all $v \in M$:
\[
v(D_1 \wedge D_2) = \min\{v(D_1), v(D_2)\}, \quad v(D_1 \vee D_2) = \max\{v(D_1), v(D_2)\}.
\]
In particular, every formal divisor $D$ is of the form
\[
D = D_0 - D_\infty, \quad \text{where } D_0 = D \lor 0 \text{ and } D_\infty = (-D) \lor 0.
\]
We also put
\[
|D| := D \lor (-D) = D_0 + D_\infty \quad \text{for } D \in \mathcal{D}^f(K, M).
\]
Every element $a \in K^\times$ determines a formal divisor, written $\text{div}(a) = \text{div}_{(K, M)}(a)$ or just $(a)$, by
\[
v(\text{div}(a)) = v(a) \quad \text{for all } v \in M.
\]
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
\[
a \mapsto \text{div}(a) : K^\times \to \mathcal{D}^f(K, M)
\]
whose image we denote by $\mathcal{P}(K, M)$. We let $\mathcal{D}(K, M)$ be the smallest subgroup of $\mathcal{D}^f(K, M)$ closed under the operation $\wedge$ and containing $\mathcal{P}(K, M)$. (Then $\mathcal{D}(K, M)$ is also closed under $\lor$.) We call $\mathcal{D}(K, M)$ the divisor group of $(K, M)$ and its elements divisors. The quotient group
\[
\text{Cl}(K, M) = \mathcal{D}(K, M) / \mathcal{P}(K, M)
\]
is called the divisor class group of $(K, M)$, and its elements divisor classes.

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

AN EFFECTIVE WEIERSTRASS DIVISION THEOREM 11
Proof. Let \( S \) be the set on the right hand side. Then \( S \) is closed under \( \wedge \) and contains \( P(K, M) \); in order to show \( S = D(K, M) \), it therefore suffices to prove that \( S \) is a subgroup of \( D^f(K, M) \). This follows immediately from the following two easily verified facts: if

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

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

\[
D + D' = \bigwedge_{1 \leq i \leq m} \bigvee_{1 \leq j \leq n} (a_{ij}a'_{kl})
\]

and

\[
-D = \bigvee_{1 \leq i \leq m} (1/a_{ij}),
\]

where \( \varphi \) runs through all maps \( \{1, \ldots, m\} \to \{1, \ldots, n\} \).

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

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

\[
D(K, M) = \bigoplus_{v \in M} \Gamma_v,
\]

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

Proof. Since \( \text{supp} \ D \) is finite for each \( D \in D(K, M) \), we have \( D(K, M) \subseteq \bigoplus_v \Gamma_v \). For the reverse inclusion, it suffices to show: for every \( v \in M \) and \( \gamma \in \Gamma_v \) there exists a divisor \( D \in D(K, M) \) with \( v(D) = \gamma \) and \( w(D) = 0 \) for \( w \neq v \) 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 \( w \neq v \) in \( M \) with \( w(x) > 0 \). Set \( D' = (x) \circ (y) \). If \( v(y) \geq v(x) \), 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

\[
v \mapsto \varrho(v) := v \upharpoonright K_1: M_1 \to M
\]

induces an embedding of ordered groups

\[
\varrho^*: D^f(K, M) \to D^f(K_1, M_1)
\]

between the formal divisor groups via the rule

\[
v(\varrho^*(D)) = (\varrho(v))(D) \quad \text{for all} \ D \in D^f(K, M), v \in M_1.
\]

We have

\[
\varrho^*(\text{div}(K, M)(a)) = \text{div}(K_1, M_1)(a) \quad \text{for} \ a \in K^\times,
\]

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

\[
\varrho^*(D_1 \lor D_2) = \varrho^*(D_1) \lor \varrho^*(D_2), \quad \varrho^*(D_1 \land D_2) = \varrho^*(D_1) \land \varrho^*(D_2),
\]

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

Let now \( (K, M) \) be an arithmetic field and \( L/K \) be a normal field extension. The group \( \text{Gal}(L/K) \) acts on \( M_L \) in a natural way from the right: for \( v \in M_L \) and \( \sigma \in \text{Gal}(L/K) \).
put \( v^\sigma := v \circ \sigma \in M_L \). This action induces a left action \((\sigma, D) \mapsto \sigma D\) of \( \text{Gal}(L|K) \) on \( \mathcal{D}^f(L, M_L) \):

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

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

\[
\sigma \text{div}(a) = \text{div}(\sigma(a)) \quad \text{for } a \in L^\times
\]

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

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

Hence for every intermediate field \( K \subseteq F \subseteq L \):

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

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 \( \deg : \mathcal{D}(K, M) \rightarrow \mathbb{R} \) satisfying

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

Because of (D3), \( \deg \) induces a homomorphism \( \text{Cl}(K, M) \rightarrow \mathbb{R} \), 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 \( v \mapsto v(D) : M \rightarrow \mathbb{R} \), a function of finite support. So every real-valued function \( \deg \) on \( \mathcal{D}(K, M) \) satisfying (D1) and (D2) is of the form

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

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 \( \deg : \mathcal{D}(K, M) \rightarrow \mathbb{R} \) satisfying (D1) and (D2) by defining \( \deg \) as in (2.2). So in this case, (D3) is equivalent to the **product formula**

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

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

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

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

\[
\prod_{v \in M_K} |a|_v^{\lambda_v} = 1 \quad (a \in K^\times),
\]
where the \( \lambda_v \) 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 \((v \upharpoonright \mathbb{Q})\)-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 \( \kappa_v \) is the cardinality of the (finite) residue field \( \mathcal{O}_v / m_v \) of \( v \).

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

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

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

\[
\begin{align*}
h(x) &\geq 0, \quad h(0) = h(1) = 0, \quad (2.3) \\
h(x) &= h(-x) \quad (2.4) \\
h(x^k) &= |k|h(x) \quad \text{for all } k \in \mathbb{Z}, \quad x \neq 0, \quad (2.5) \\
h(xy) &\leq h(x) + h(y), \quad (2.6) \\
h(x_1 + \cdots + x_n) &\leq \max \{h(x_1), \ldots, h(x_n)\}, \quad (2.7) \\
&\quad \text{if } M \text{ consists only of ultrametric absolute values,} \\
h(x_1 + \cdots + x_n) &\leq h(x_1) + \cdots + h(x_n) + h(n), \quad (2.8) \\
&\quad \text{if } M \text{ contains an archimedean absolute value.}
\end{align*}
\]

**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

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

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

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

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 Gal\((L/K)\)-invariant if \( \deg(\sigma D) = \deg(D) \) for all \( \sigma \in \text{Gal}(L/K) \) and \( D \in D(L, M_L) \). In this case, \( h(\sigma(x)) = h(x) \) 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 Gal\((K_{\text{alg}}/K)\)-invariant degree function on \((K_{\text{alg}}, M_{\text{alg}})\).

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

\[
N_{L/K}(D) := [L : K]_{\text{ins}} \sum_{\sigma \in \text{Gal}(L/K)} \sigma D,
\]
where $[L : K]_{\text{ins}}$ = degree of inseparability of $L|K$, we define a homomorphism
\[ N_{L|K} : D(L, M_L) \to D(L, M_L), \]
called the norm map. It is clear that for all $D \in D(L, M_L)$
\begin{align*}
(N1) \quad & D \geq 0 \Rightarrow N_{L|K}(D) \geq 0, \\
(N2) \quad & D \in D(K, M) \Rightarrow N_{L|K}(D) = [L : K] D.
\end{align*}
By (N1) we have, for $D, D' \in D(L, M_L)$:
\begin{align*}
(N3) \quad & N_{L|K}(D \wedge D') \leq N_{L|K}(D) \wedge N_{L|K}(D'), \\
(N4) \quad & N_{L|K}(D \vee D') \geq N_{L|K}(D) \vee N_{L|K}(D').
\end{align*}
Moreover, for all $x \in L^\times$
\begin{align*}
(N5) \quad & N_{L|K} \left( \text{div}(x) \right) = \text{div} \left( N_{L|K}(x) \right),
\end{align*}
where $N_{L|K}(x)$ is the field-theoretic norm of the element $x$. In particular, we have
\[ N_{L|K} \left( \mathcal{P}(L, M_L) \right) \subseteq \mathcal{P}(K, M). \]
This raises the question whether we also have $N_{L|K}(D(L, M_L)) \subseteq D(K, M)$. 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 $D^{(N)}(L, M_L)$, for $N \geq 1$, denote the set of all divisors $D \in D(L, M_L)$ 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 $D, D' \in D(L, M_L)$ we also set $[D, D'] := \{ E \in D(L, M_L) : D \leq E \leq D' \}$.

**Proposition 2.12.** For all $N \geq 1$ there exists a divisor $D_N \in D(K, M)$ such that $D_N \geq 0$ and
\[ N_{L|K}(D^{(N)}(L, M_L)) \subseteq D(K, M) + [-D_N, D_N]. \]
If all absolute values in $M$ are ultrametric, then
\[ N_{L|K}(D(L, M_L)) \subseteq D(K, M). \]
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 $N_{L'|K} = N_{L|K} \circ N_{L'|L}$.

For the proof of Proposition 2.12, we let $D^{(N)}_\vee(L, M_L)$ denote the set of all divisors $D \in D(L, M_L)$ of the form $D = \bigvee_{i=1}^N \{a_i\}$, where $a_i \in L^\times$. Similarly we define $D^{(N)}_\wedge(L, M_L)$. We also define
\[ D_\vee(L, M_L) = \bigcup_{N \geq 1} D^{(N)}_\vee(L, M_L) \quad \text{and} \quad D_\wedge(L, M_L) = \bigcup_{N \geq 1} D^{(N)}_\wedge(L, M_L). \]
Both sets are sub-semigroups of $D(L, M_L)$, by the identity (2.1) in the proof of Lemma 2.5.

**Lemma 2.14.** Let $N \geq 1$ and put $D'_N = (2^N)^\infty$. Then for all $D \in D^{(N)}(L, M_L)$ we have
\[ -D'_N + D_1 - D_2 \leq D \leq D'_N + D_1 - D_2 \]
for certain $D_1, D_2 \in D^{(N^2)}_\vee(L, M_L)$. If all absolute values in $M$ are ultrametric, then $D'_N = 0$, so in this case every $D \in D(L, M_L)$ is a difference of divisors from $D_\vee(L, M_L)$.
Proof. Suppose $D = \bigwedge_{i=1}^N E_i$ with $E_i = \bigvee_{j=1}^N (a_{ij})$. We let, for $1 \leq i \leq N$,

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

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 $\mathbb{L}[T]$, 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

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

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 $D'_N = (2N^2)^\infty$ 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 $D'_N = (2N^{[L:K]} \log 2)^\infty$. Then for each divisor $D \in \mathcal{D}_\vee^{(N)}(L, M_L)$ we have

$$-D'_N + E \leq \min_{L/K} \log \varepsilon_v - v(D) \leq D'_N + E$$

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

$$-D'_N + E' \leq \min_{L/K} \log \varepsilon_v - v(D) \leq D'_N + E'$$

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

$$\min_{L/K} \mathcal{D}_\vee(L, M_L) \subseteq \mathcal{D}_\vee(K, M) \quad \text{and} \quad \min_{L/K} \mathcal{D}_\vee(L, M_L) \subseteq \mathcal{D}_\vee(K, M).$$

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

$$f(T) = \sum_{i=1}^N a_i T^i, \quad g(T) = \prod_{\sigma \in \text{Gal}(L/K)} (\sigma f)(T)^{[L:K]_{\text{inv}}},$$

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

$$v(N_{L/K}(D)) = [L : K]_{\text{inv}} \sum_{\sigma} v(\sigma D) = [L : K]_{\text{inv}} \sum_{\sigma} v(\sigma f).$$

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

$$-N^i \log \varepsilon_v + v(g) \leq [L : K]_{\text{inv}} \sum_{\sigma} v(\sigma f) \leq N^i \log \varepsilon_v + v(g),$$
\[ N' = \deg g = N \cdot [L : K]. \] Hence

\[ -D_N'' + E \leq N_{L/K}(D) \leq D_N'' + E \]

where \( E \in \mathcal{D}_r(K, M) \) and \( D_N'' = (2^N)'', \) with \( D_N'' = 0 \) if \( M \) contains only ultrametric absolute values. \(\square\)

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

\[ -D_N'' + D_1 - D_2 \leq D \leq D_N'' + D_1 - D_2 \]

with \( D_N' = (2^{N_2})' \) and certain \( D_1, D_2 \in \mathcal{D}^{(N_2)}(L, M_L) \). By Lemma 2.15, there are \( E_1, E_2 \in \mathcal{D}_r(K, M) \) such that

\[ -D_N'' + E_i \leq N_{L/K}(D_i) \leq D_N'' + E_i \quad \text{for } i = 1, 2, \]

with \( D_N'' = (2^{N_2[L : K]})' \). It now follows easily that

\[ -D_N + (E_1 - E_2) \leq N_{L/K}(D) \leq D_N + (E_1 - E_2) \]

where \( D_N = [L : K]D_N' + 2D_N'' \in \mathcal{D}(K, M) \), with \( D_N \geq 0 \), and \( D_N = 0 \) if \( M \) contains no archimedean absolute value. \(\square\)

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 \( \text{Gal}(L/K) \)-invariant degree function \( \deg_L \) on \((L, M_L)\). Hence suppose that \( L/K \) is a finite normal extension. For, for \( D \in \mathcal{D}(L, M_L) \):

\[ \overline{\deg}(D) := \inf \{ \deg(E) : E \in \mathcal{D}(K, M), E \geq N_{L/K}(D) \}, \quad (2.9) \]

\[ \underline{\deg}(D) := \sup \{ \deg(E) : E \in \mathcal{D}(K, M), E \leq N_{L/K}(D) \}. \quad (2.10) \]

Note that for each \( D \in \mathcal{D}(L, M_L) \) we have \( D \in \mathcal{D}^{(N)}(L, M_L) \) for some \( N \geq 1 \), and thus \( -D_N \leq N_{L/K}(D) \) 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

\[ D \in \mathcal{D}(K, M) \Rightarrow \overline{\deg}(D) = \deg(D) = d \deg(D) \quad (2.11) \]

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

\[ \overline{\deg}(D) \leq \overline{\deg}(D), \quad (2.12) \]

\[ \underline{\deg}(D + D') \geq \underline{\deg}(D) + \underline{\deg}(D'), \quad (2.13) \]

\[ \overline{\deg}(D + D') \leq \overline{\deg}(D) + \overline{\deg}(D'), \quad (2.14) \]

and hence, by induction,

\[ \overline{\deg}(D) \leq \frac{1}{n} \deg(nD) \leq \frac{1}{n} \underline{\deg}(nD) \leq \underline{\deg}(D) \quad \text{for all } n \geq 1. \]

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

\[ \frac{1}{2^n} \deg(2^n D) \leq \frac{1}{2^{n+1}} \deg(2^{n+1} D) \quad \text{and} \quad \frac{1}{2^n} \deg(2^n D) \leq \frac{1}{2^{n+1}} \deg(2^{n+1} D) \]

for all \( n \), and, by Proposition 2.12,

\[ 0 \leq \overline{\deg}(2^n D) - \deg(2^n D) \leq 2 \deg(D_N) \quad \text{for all } n. \]

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

\[ \deg_L(D) := \frac{1}{d} \lim_{n \to \infty} \frac{\deg(2^n D)}{2^n} = \frac{1}{d} \lim_{n \to \infty} \frac{\deg(2^n D)}{2^n}. \]
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 \( \deg(D) \geq 0 \) and \( \deg(D) \geq 0 \) if \( D \geq 0 \), and by (N5) we have for a principal divisor \( D = (x) \), \( x \in L^\times \), that \( \deg(nD) = \deg(\text{div}(\mathcal{N}_{L/K}(x^n))) = 0 \). 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 \( \text{Gal}(L/K) \)-invariant degree function on \((L, M_L)\) extending \( \deg \). Then for all \( D \in \mathcal{D}(L, M_L) \)

\[
\deg'_L(D) = \frac{1}{d} \deg'_L(N_{L/K}(D)),
\]

hence \( \deg'_L(D) \leq \deg(E)/d \) whenever \( E \in \mathcal{D}(K, M) \) is such that \( E \geq N_{L/K}(D) \), and thus \( \deg'_L(D) \leq \deg(D)/d \). Similarly we get \( \deg'_L(D) \geq \deg(D)/d \) 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 \( \lim_{n \to \infty} \frac{\deg(nD)}{n} \). Note however that it is not clear \textit{a priori} that the sequence \( \{\deg(nD)/n\} \) has a limit in \( \mathbb{R} \).

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

\[
\deg(D) := \frac{1}{d} \deg(N_{L/K}(D)) \quad \text{where } d := [L : K].
\]

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

We denote the unique extension of \( \deg \) to a \( \text{Gal}(K_{alg}|K) \)-invariant degree function on \((K_{alg}, M_{alg})\) 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 \( Q \) to a \( \text{Gal}(Q_{alg}|Q) \)-invariant degree function on \((Q_{alg}, M_{alg})\) is the usual (logarithmic) height function on \( Q_{alg} \) used in diophantine geometry [23]. For an element \( \alpha \) of an algebraic number field \( K \) we have

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

where the \( \lambda_v \) 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 \( \alpha \in Q_{alg} \) 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 \( \mathbb{P}^n = \mathbb{P}^n(K_{alg}) \) over \( K_{alg} \), we have \( \bigwedge_{x_i \neq 0} (\lambda x_i) = \bigwedge_{x_i \neq 0} (x_i) + (\lambda) \) for all \( \lambda \in K_{alg}^\times \). Hence to \( P \) we may associate a divisor class

\[
\text{div} [x_0 : \cdots : x_n] := \bigwedge_{x_i \neq 0} (x_i) \quad \text{(modulo principal divisors)}.
\]
We define the **projective height function** $h : \mathbb{P}^n \rightarrow \mathbb{R}$ associated to $(K, M, \deg)$ by

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

Note that $h([x_0 : \cdots : x_n]) \geq 0$. (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 $\pi$ of the set $\{0, \ldots, n\}$,
   $$h([x_{\pi(0)} : \cdots : x_{\pi(n)}]) = h([x_0 : \cdots : x_n]).$$
2. If $\sigma \in \text{Gal}(K_{\text{alg}}|K)$, then $h(\sigma P) = h(P)$, where
   $$\sigma P := [\sigma(x_0) : \cdots : \sigma(x_n)] \in \mathbb{P}^n.$$
3. If $\text{div} P \leq \text{div} Q$, then $h(P) \geq h(Q)$. In particular: If $x_{n+1} \in K_{\text{alg}}$, then
   $$h([x_0 : \cdots : x_n]) \leq h([x_0 : \cdots : x_{n+1}]).$$
4. For all $x \in K$, $h(x) = h([1 : x])$.
5. $h([x_0 : \cdots : x_n]) \geq \max\{h(x_i/x_j) : x_j \neq 0\}$.

**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 $h([1 : x_1 : \cdots : x_n]) \geq h([1 : x_1]) = h(x_1) \geq h(x_i/x_j)$ for all such $i, j$, by (3) and (4). \qed

**Example 2.19.** For the global field $\mathbb{Q}$, the height of a point $P = [x_0 : \cdots : x_n] \in \mathbb{P}^n(\mathbb{Q}_{\text{alg}})$ is the so-called **absolute logarithmic height** $h(P)$ (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 $\mathbb{P}^n(\mathbb{Q}_{\text{alg}})$ 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 $K_{\text{alg}}$). We will deduce it from a global version of Gelfond’s Lemma. First we define a height function for polynomials:

**Definition 2.20.** Let $f(X) = \sum \alpha_v X^v$ be a non-zero polynomial in the indeterminates $X = (X_1, \ldots, X_N)$ with coefficients $\alpha_v \in K_{\text{alg}}$. We put

$$\text{coeff}(f) = \prod_{\alpha_v \neq 0} (\alpha_v),$$

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

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

Note that $h_{\text{coeff}}(f)$ 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 $K_{\text{alg}}$. Similarly to $h_{\text{coeff}}(f)$ for a single polynomial $f$, we define $h_{\text{coeff}}(f_1, \ldots, f_n)$ for a finite sequence of polynomials $(f_1, \ldots, f_n)$ in $K_{\text{alg}}[X]$, where $n > 0$ and not all $f_i = 0$. 

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

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

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

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

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

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

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

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

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

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

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

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 $f_1, \ldots, f_d$ with $f_j(T) = T - \alpha_j$ for $j = 1, \ldots, d$. \hfill \square

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 $\alpha \in K_{\text{alg}}$ such that $h(\alpha) \leq C$ and $[K(\alpha) : K] \leq d$.

Proof. Suppose $\alpha \in K_{\text{alg}}$ is such that $h(\alpha) \leq C$. Let $\alpha_1, \ldots, \alpha_d$ be the conjugates of $\alpha$ in $K_{\text{alg}}$, where $d := [K(\alpha) : K]$. The minimal polynomial of $\alpha$ 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 $a_0, \ldots, a_{d-1} \in K$. Now by Corollary 2.22:

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

So there are only finitely many possibilities for choosing $a_0, \ldots, a_{d-1} \in K$, and for given $a_0, \ldots, a_{d-1} \in K$, the polynomial $T^d + a_{d-1} T^{d-1} + \cdots + a_0$ has at most $d$ distinct zeros in $K_{\text{alg}}$. Hence there are only finitely many $\alpha \in K_{\text{alg}}$ with $h(\alpha) \leq C$ and $[K(\alpha) : K] \leq d$. \hfill \square
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 $K_{\text{alg}}$ of the subgroup of $\text{Gal}(K_{\text{alg}}|K)$ consisting of all $\sigma$ 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 $P \in \mathbb{P}^n(K_{\text{alg}})$ with $h(P) \leq C$ and $|K(P) : K| \leq d$.

2. A non-zero element of $K_{\text{alg}}$ 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 $x \in K_{\text{alg}}^*$ is a root of unity, then $h(x) = 0$ by (2.5). For the converse, let $x \in K_{\text{alg}}^*$ with $h(x) = 0$, and $L := K(x)$, $d := [L : K]$. By the proposition, the subgroup of $L^n$ consisting of all elements of height 0 is finite. Therefore $x^m = 1$ for some positive $m$. □

The embedding of affine space $\mathbb{A}^n = \mathbb{A}^n(K_{\text{alg}})$ over $K_{\text{alg}}$ into projective space
\[ \mathbb{A}^n(K_{\text{alg}}) \hookrightarrow \mathbb{P}^n(K_{\text{alg}}), \quad (x_1, \ldots, x_n) \mapsto [1 : x_1 : \cdots : x_n], \]
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
\[ h(P) := h([1 : x_1 : \cdots : x_n]) = \deg \bigvee_{x_i \neq 0} (x_i)_{\infty}, \]
and $h(P) := 0$ if all $x_i = 0$. It is also convenient to introduce the abbreviation
\[ h_{\text{max}}(P) := \max_i h(x_i) \quad \text{for } P = (x_1, \ldots, x_n) \in \mathbb{A}^n, \]
with the convention that the maximum is 0 for $n = 0$. Here are some properties of $h$ (affine height) and $h_{\text{max}}$ on $\mathbb{A}^n$:

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

1. For every permutation $\pi$ of the set $\{1, \ldots, n\}$, we have
\[ h(x_{\pi(1)}, \ldots, x_{\pi(n)}) = h(x_1, \ldots, x_n), \]
and similarly for $h_{\text{max}}$ in place of $h$.

2. If $\sigma \in \text{Gal}(K_{\text{alg}}|K)$, then $h(\sigma P) = h(P)$ and $h_{\text{max}}(\sigma P) = h_{\text{max}}(P)$, where
\[ \sigma P := (\sigma(x_1), \ldots, \sigma(x_n)). \]

3. If $P = (x_1, \ldots, x_n) \in \mathbb{P}^n$, then $h(P) = h([1 : x_1 : \cdots : x_n])$.

4. $h_{\text{max}}(x_1, \ldots, x_n) \geq h([x_1 : \cdots : x_n])$. If $P \neq 0$.

5. $h_{\text{max}}(P + Q) \leq h_{\text{max}}(P) + h_{\text{max}}(Q) + h(2)$.

6. For $\lambda \in K$, we have $h_{\text{max}}(\lambda P) \leq h(\lambda) + h_{\text{max}}(P)$.

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

\[ f: \mathbb{P}^n \setminus Z \to \mathbb{P}^m, \quad P \mapsto f(P), \]

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

\[ h(f(P)) \leq d h(P) + h_{\text{coeff}}(f) + h(\text{mon}(f)), \quad (2.16) \]

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

\[ h(f(P)) \leq d h(P) + h_{\text{coeff}}(f). \]

All this follows easily from Lemma 1.3.

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

\[ f: \mathbb{A}^n \to \mathbb{A}^m, \quad P \mapsto f(P). \]

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

\[ h(f(P)) \leq d h(P) + h_{\text{coeff}}(f) + h(\text{mon}(f)), \quad (2.17) \]

where \( \text{mon}(f) := \max_i \text{mon}(f_i) \). If \( M \) contains no archimedean absolute value, then

\[ h(f(P)) \leq d h(P) + h_{\text{coeff}}(f). \]

**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 \( r = \text{res}(f, g) \in K \) be the resultant of \( f \) and \( g \). Then

\[ h(r) \leq n h_{\text{coeff}}(f) + m h_{\text{coeff}}(g) + m h(1 + n) + mn \log 2. \]

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

\[ h(r) \leq \sum_{i=1}^m h(g(\alpha_i)) \leq \sum_{i=1}^m \left( n h(\alpha_i) + h_{\text{coeff}}(g) + h(1 + n) \right) \leq n h_{\text{coeff}}(f) + m h_{\text{coeff}}(g) + m h(1 + n) + mn \log 2 \]

as claimed. \( \square \)

**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 \( h_{\text{max}}(A) \) and \( h(A) \), respectively.

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

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

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

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

\[ h(\det A) \leq nh_{\text{max}}(A) + h(n!) \leq nh_{\text{max}}(A) + nh(n), \]
AN EFFECTIVE WEIERSTRASS DIVISION THEOREM 23

showing (1). In fact, if $M$ consists only of ultrametric absolute values, one gets similarly

$$h(\det A) \leq n h_{\text{max}}(A)$$

by (2.7), which shows the last statement. For each $v \in M$, by Lemma 1.2, (2)

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

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

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

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

$$x'_i = \sum_{j=0}^n a_{ij} x_j \quad \text{for } i = 0, \ldots, n,$$

(2.18)

with $A = (a_{ij}) \in \text{GL}_{n+1}(\mathbb{K}_{\text{alg}})$. Note that $h_{\text{coeff}}(f)$ 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

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

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

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

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

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

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

$$2 (\det B) \geq n (n)_{\infty} + 2 \sum_j \prod_i |a_{ij}| \geq n \left( (n)_{\infty} + 2 \prod_i |a_{ij}| \right).$$

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

$$2 \div A^{\text{ad}} \geq n \left( (n)_{\infty} + 2 \prod_i |a_{ij}| \right).$$

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. $h(A \cdot B) \leq h(A) + h(B) + h(n)$;
These matrices will play an important role in the proof of Theorem 0.1. If all \( \mu \), \( \xi \), so if in addition \( 24 \) MATTHIAS ASCHENBRENNER
\[ 2.6. \text{ Generalized Vandermonde matrices.} \]

For \( \xi \in K \) and \( \mu, s \in \mathbb{N} \) with \( 1 \leq \mu \leq s \) we define the \( \mu \times s \)-matrix

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

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 \( (\xi, \mu, s) \) 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

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

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

\[
h(A(\xi, \mu, s)) = (s - 1)h(\xi_0, \ldots, \xi_n),
\]
since
\[
\bigvee_{0 \leq i \leq n} \bigwedge_{1 \leq i < s} (\xi_i^s) = (s-1) \bigwedge_{0 \leq i \leq n} (\xi_i). \tag{2.20}
\]

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

\[
h(A) \leq (s - 1) h(\xi_0, \ldots, \xi_n) + (\mu_{\max}(1 : 2 : \cdots : s - 1)),
\]

where \( \mu_{\max} := \max\{\mu_0, \ldots, \mu_n\} \).

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 \( 0 \leq j \leq d \):

\[
\text{coeff} \left( f^{(j)}(\xi T) \right) \geq \text{coeff} \left( f(\xi T) \right) + j \cdot \left( \text{div} [1 : 2 : \cdots : d] - (\xi) \right).
\]

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

**Proof.** By induction on \( j \), it only suffices to consider the case \( j = 1 \). We write \( f(T) = a_0 + a_1 T + \cdots + a_d T^d \) with \( a_0, \ldots, a_d \in K \) and compute:

\[
\text{coeff} \left( f'(T) \right) = \bigwedge_{1 \leq i \leq d} (i a_i T^{i-1}) = \bigwedge_{1 \leq i \leq d} (a_i T^{i-1}) + \bigwedge_{i=1}^d (i) \\
= \bigwedge_{1 \leq i \leq d} (a_i T^{i-1}) + \bigwedge_{i=1}^d (i) - (\xi) = \text{coeff} \left( f(\xi T) \right) + \text{div} [1 : 2 : \cdots : d] - (\xi)
\]

as required. \( \square \)

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 \( 1 \leq \mu \leq s \) we have

\[
\text{div} \left( A(\xi, \mu, s) \right) = \bigwedge_{j=0}^{\mu-1} \text{coeff} \left( f^{(j)}(\xi T) \right).
\]

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

\[
\text{div} \left( A(\xi, \mu, s) \right) \geq \bigwedge_{j=0}^{\mu-1} \left( \text{coeff} \left( f(\xi T) \right) + j \cdot \left( \text{div} [1 : 2 : \cdots : s - 1] - (\xi) \right) \right) \\
= \text{coeff} \left( f(\xi T) \right) + \bigwedge_{j=0}^{\mu-1} j \cdot \text{div} [1 : 2 : \cdots : s - 1] \\
\geq \text{coeff} \left( f(\xi T) \right) + (\mu - 1) \cdot \text{div} [1 : 2 : \cdots : s - 1]. \tag{2.21}
\]
We now prove Proposition 2.32. For this, let \( \xi = (\xi_0, \ldots, \xi_n) \in K^{n+1} \) and \( \mu = (\mu_0, \ldots, \mu_n) \in \mathbb{N}^{n+1} \), 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):
\[
\text{div} \left( A(\xi, \mu, s) \right) = \bigwedge_{i=0}^{n} \text{div} \left( A(\xi_i, \mu_i, s) \right)
\geq \bigwedge_{i=0}^{n} \text{coeff} \left( f(\xi_i T) \right) + (\mu_i - 1) \cdot \text{div} [1 : 2 : \cdots : s - 1]
\geq \bigwedge_{i=0}^{n} \bigwedge_{j=0}^{s-1} (\xi_i^j) + (\mu_{\text{max}} - 1) \cdot \text{div} [1 : 2 : \cdots : s - 1].
\]

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 \( X = (X_1, \ldots, X_N) \) is a tuple of indeterminates, where \( N \geq 1 \). In this section we construct a height function on the algebraic closure \( \mathbb{Q}(X)_{\text{alg}} \) 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 \( \text{Gal}(\mathbb{Q}(X)_{\text{alg}}/\mathbb{Q}(X)) \)-invariant degree function on \( \mathbb{Q}(X)_{\text{alg}} \), and establish a few of its properties needed later on.

3.1. The arithmetic field \((\mathbb{Q}(X), M_{\text{arch}})\). Let \( M_{\text{arch}} \) be the set of all archimedean absolute values on \( \mathbb{Q}(X) \) which extend the usual archimedean absolute value \( v_\infty \) on \( \mathbb{Q} \). There is a bijective correspondence between \( M_{\text{arch}} \) and the subset
\[ U := \left\{ (z_1, \ldots, z_N) \in \mathbb{C}^N : \text{trdeg}(\mathbb{Q}(z_1, \ldots, z_N)/\mathbb{Q}) = N \right\} \]
of \( \mathbb{C}^N \) as follows: each \( z \in U \) defines an archimedean absolute value \( v_z \in M_{\text{arch}} \) on \( \mathbb{Q}(X) \) by
\[ |f/g|_z = |f/g|_{v_z} := |f(z)/g(z)| \quad \text{for } f, g \in \mathbb{Z}[X], g \neq 0, \]
and conversely, by Ostrowski’s Theorem, each \( v \in M_{\text{arch}} \) is of the form \( v = v_z \) for a uniquely determined \( z \in U \). Let \( C(U, \mathbb{R}) \) be the lattice-ordered group of continuous functions \( U \to \mathbb{R} \) (pointwise operations and ordering). The map \( z \mapsto \varphi(z) := v_z : U \to M_{\text{arch}} \) induces a homomorphism of lattice-ordered groups
\[ \varphi^* : D(\mathbb{Q}(X), M_{\text{arch}}) \to C(U, \mathbb{R}), \quad \varphi^*(D)(z) := \varphi(z)(D) = v_z(D) \text{ for all } z \in U. \]
The homomorphism \( \varphi^* \) maps the principal divisor \( (f/g) \), for non-zero \( f, g \in \mathbb{Z}[X] \), to the function
\[ z \mapsto -\log |f(z)/g(z)| : U \to \mathbb{R}. \]
The arithmetic field \((\mathbb{Q}(X), M_{\text{arch}})\) is not classical.
3.2. Integration of divisors on \((\mathbb{Q}(X), M_{\text{arch}})\). 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 \(\mu_1\) 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 \(\mu_1\) on \(\mathbb{T}^N\). Then \(\mu_N\) is a non-atomic probability measure on \(\mathbb{T}^N\). Let \(\mu = \varphi(\mu_N | U \cap \mathbb{T}^N)\) be the image measure of the restriction of \(\mu_N\) to \(U \cap \mathbb{T}^N\) under the map \(\varphi\), that is,
\[
\mu(S) = \mu_N(\varphi^{-1}(S) \cap \mathbb{T}^N)
\]
for all \(S \subseteq M_{\text{arch}}\). (3.1)

Each \(D \in \mathcal{D}(\mathbb{Q}(X), M_{\text{arch}})\) is \(\mu\)-measurable and \(\varphi^*(D) | \mathbb{T}^N\) is defined \(\mu_N\)-almost everywhere, since the set
\[
\mathbb{T}^N \setminus U = \bigcup_{f \in \mathbb{Z}[X]\setminus\{0\}} \{z \in \mathbb{T}^N : f(z) = 0\}
\]
has measure zero. Therefore,
\[
\int_{M_{\text{arch}}} D \, d\mu = \int_{U \cap \mathbb{T}^N} \varphi^*(D) | \mathbb{T}^N \, d\mu_N = \int_{\mathbb{T}^N} \varphi^*(D) | \mathbb{T}^N \, d\mu_N.
\]

For non-zero \(f(X) \in \mathbb{Z}[X]\), we have
\[
m(f) = \int_{\mathbb{T}^N} \log |f| \, d\mu_N = -\int_{M_{\text{arch}}} (f) \, d\mu,
\]
where \(m(f)\) is the Mahler measure of \(f\). We also have
\[
\int_{M_{\text{arch}}} (f)_\infty \, d\mu = \int_{\mathbb{T}^N} (-\varphi^*((f)) \vee 0) \, d\mu_N = \int_{\mathbb{T}^N} \log^+ |f| \, d\mu_N = m^+(f),
\]
using the notation \(m^+\) introduced in Section 1.2. In fact:

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

**Proof.** Since \(L^1(\mu)\) is closed under \(\wedge\) and \(\vee\), it is enough to verify that \(\text{div}(f) \in L^1(\mu)\) 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 \(\int_{M_{\text{arch}}} |\text{div}(f)| \, d\mu = -m(f) + 2m^+(f) \in \mathbb{R}\) by (3.2) and (3.3), and hence \(\text{div}(f) \in L^1(\mu)\) as claimed. \(\square\)

**Example 3.2.** For \(N = 1\) and \(f(X) = X\), we get
\[
\int_{M_{\text{arch}}} (X) \, d\mu = \int_{S^1} \log^+ |z| \, d\mu_1(z) = 0.
\]
Therefore, for any monomial \(X^\nu = X_1^{\nu_1} \cdots X_N^{\nu_N}\) with \(\nu = (\nu_1, \ldots, \nu_N) \in \mathbb{N}^N\),
\[
\int_{M_{\text{arch}}} (X^\nu) \, d\mu = 0.
\]
(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 $M_\infty$ 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 $\text{div}_\infty(a)$ for the principal divisor of $a \in K^\times$ in $D(K, M_\infty)$. Assume that $\deg_\infty : D(K, M_\infty) \to \mathbb{R}$ is a map satisfying, for each $D, E \in D(K, M_\infty)$:

$$\deg_\infty(D + E) = \deg_\infty(D) + \deg_\infty(E),$$

$$D \geq 0 \Rightarrow \deg_\infty(D) \geq 0,$$

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

Note that then $\deg_\infty\left(\text{div}_\infty(u)\right) = 0$ 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

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

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

$$\pi_\infty : D(K, M) \to D(K, M_\infty), \quad \pi_\infty(D) = D \upharpoonright M_\infty$$

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

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

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

$$h(x/y) = \deg(\left(\begin{array}{c} -\text{div}_\infty(x) \\ \vee \\ -\text{div}_\infty(y) \end{array}\right)), \quad (3.10)$$

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

$$h(x/y) = \deg(\left(\begin{array}{c} x/y \\ \vee \\ -\text{div}_\infty(x/y) \end{array}\right))$$

$$= \sum_{v \in M_R} v(y) \lambda_v + \deg_\infty(\left(\begin{array}{c} -\text{div}_\infty(x/y) \\ \vee \\ 0 \end{array}\right))$$

$$= -\deg_\infty(\text{div}_\infty(y)) + \deg_\infty(\left(\begin{array}{c} -\text{div}_\infty(x/y) \\ \vee \\ 0 \end{array}\right))$$

$$= \deg_\infty(\left(\begin{array}{c} -\text{div}_\infty(x) \\ \vee \\ -\text{div}_\infty(y) \end{array}\right)).$$

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

$$h(f) \geq \sum_{v \in M_R} \lambda_v \log^+ |f|_v. \quad (3.11)$$

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

$$h([a_0 : \cdots : a_n]) =$$

$$-\sum_{v \in M_R} \min_i v(a_i) \lambda_v + \deg_\infty(\left(\begin{array}{c} -\text{div}_\infty(a_0) \\ \vee \\ \cdots \\ \vee \\ -\text{div}_\infty(a_n) \end{array}\right)),$$

hence

$$h([a_0 : \cdots : a_n]) \leq \deg_\infty(\left(\begin{array}{c} -\text{div}_\infty(a_0) \\ \vee \\ \cdots \\ \vee \\ -\text{div}_\infty(a_n) \end{array}\right)).$$
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

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

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 $v_X : K^\times \to \mathbb{Z}$ is the **total degree valuation** on $K$, given by

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

We let

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

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

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

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 $\text{Gal}(K_{\text{alg}}|K)$-invariant height function on $K_{\text{alg}}$ is an “$F[X]$-stable degree function on $K_{\text{alg}}$” (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 $v_{X_i} : K^\times \to \mathbb{Z}$ is the $X_i$-**degree valuation** on $K$, given by

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

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

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

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

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

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 $(\mathbb{Q}(X), M, \deg_{(X)})$ 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

$$M_\infty := M_{\text{arch}} \cup \{v_X, \ldots, v_{X_K}\},$$

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

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

$$\int_{M_{\text{arch}}} (f) \, d\mu \leq 0 \text{ for all } 0 \neq f \in R.$$  \hspace{1cm} \text{(3.12)}

The measure $\mu = \varphi(\mu_N \upharpoonright U \cap \mathbb{T}^N)$ defined in (3.1), which is concentrated on the subset $\varphi(U \cap \mathbb{T}^N)$ of $M_{\text{arch}}$, 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 $\varphi(\mu_{sp,N})$, where $\mu_{sp,N}$ is the $N$-fold product of the so-called spherical measure $\mu_{sp}$ on $\mathbb{C}$, given by the density function (in polar coordinates)

$$d\mu_{sp} = \frac{r \, dr \, d\theta}{\pi (1 + r^2)^{N/2}}.$$  \hspace{1cm} \text{(3.14)}

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

For any positive measure $\mu$ on $M_{\text{arch}}$ 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 $h_\mu$ is given by

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

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

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

$$h(f/g) = h_{(X)}(f/g) + \int_{\mathbb{T}^N} \max\{\log |f|, \log |g|\} \, d\mu_N. \hspace{1cm} \text{(3.14)}$$

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

$$h(f) = \deg_{(X)} f + \int_{\mathbb{T}^N} \log^+ |f(z)| \, d\mu_N(z) = \deg_{(X)} f + m^+(f). \hspace{1cm} \text{(3.15)}$$

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

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

\[
(1 - \log 2) \deg(X) f + \log ||f||_1 \leq h(f) \leq \deg(X) f + \log ||f||_1. 
\]

These estimates imply that the global field \((\mathbb{Q}(X), M, \deg)\) has the finiteness property.

By Theorem 2.11, \( \deg \) extends uniquely to a \( \text{Gal}(K_{\text{alg}}|K) \)-invariant degree function on \((K_{\text{alg}}, M_{\text{alg}})\), which we also denote by \( \deg \). Its corresponding height function, also denoted by \( h \), extends the absolute logarithmic height on \( \mathbb{Q}_{\text{alg}} \). Moreover, if \( N > 1 \), then the height function on \( \mathbb{Q}(X_1, \ldots, X_N)_{\text{alg}} \) so defined extends the height function on \( \mathbb{Q}(X_1, \ldots, X_{N-1})_{\text{alg}} \) 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 \( \alpha \in \mathbb{Q}(X)_{\text{alg}} \) with \( h(\alpha) \leq C \) and \([\mathbb{Q}(X, \alpha) : \mathbb{Q}] \leq d\).

### 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

\[
h(\partial f/\partial X_j) = \deg_{X_j}(\partial f/\partial X_j) + m(\partial f/\partial X_j + Y) 
\leq (\deg_{X_j}(f) - 1) + m(f(X) + X_j Y) + \log \deg_{X_j}(f)
\]

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

\[
m(f(X) + X_j Y) = 1 + m^+(f(X)),
\]

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

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

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

**Example 3.8.** Let \( s \in \mathbb{N}, s > 0 \). Then

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

by the previous lemma and (3.4).

**Example 3.9.** Let \( e \in \mathbb{N}, N > 0 \). Then

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

To see this note that

\[
\log^+ |z_i + z_N^{N-1}| \leq \log 2 \quad \text{for } (z_1, \ldots, z_N) \in \mathbb{T}^N, 1 \leq i \leq N - 1,
\]

and

\[
\deg_{X_i}(X_i + X_N^{N-1}) = e^{N-i} + 1,
\]

and apply Lemma 3.7.
Let \( f(X, Y) \in \mathbb{Z}[X, Y] \), \( Y = (Y_1, \ldots, Y_M) \), be non-zero, and write \( f(X, Y) = \sum\nu a_\nu (X) Y^\nu \) with \( a_\nu \in \mathbb{Z}[X] \). (Here \( \nu \) 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 \( h_{\text{coeff}}(f) \) 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

\[
h_{\text{coeff}}(f) \leq h(f) \leq h_{\text{coeff}}(f) + d + \log \left( \frac{M + d}{d} \right).
\]

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

\[
\max_\nu \log |a_\nu(z)| = \log \|f(z, Y)\|_\infty \leq m^+(f(z, Y)) + \deg(Y) f.
\]

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

\[
\int_{\mathbb{T}^N} \max_\nu \log |a_\nu(z)| \, d\mu_N(z) \leq \int_{\mathbb{T}^{N+M}} \log^+|f(z, w)| \, d\mu_{N+M}(z, w) + \deg(Y) f.
\]

Since

\[
\max_\nu \deg(X) a_\nu \leq \deg(X) f,
\]

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

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

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

\[
h_{\text{coeff}}(\partial f/\partial X_j) \leq h_{\text{coeff}}(f) + d + \log(d + 1) + \log \deg(X)(f).
\]

(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 \( r = \text{res}_T(f, g) \in \mathbb{Z}[X] \) be the resultant of \( f \) and \( g \) with respect to the indeterminate \( T \). Then

\[
h(r) \leq n h(f) + m h(g) + mn \log 4.
\]

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 \( m = m_v \) 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 \( \overline{\mathcal{O}} = \mathcal{O}/m \), with residue homomorphism \( a \mapsto \overline{a} : \mathcal{O} \to \overline{\mathcal{O}} \). The subset

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

of the ring \( K[[X]] = K[[X_1, \ldots, X_N]] \) 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, \( \nu = (\nu_1, \ldots, \nu_N) \) ranges over all multiindices in \( \mathbb{N}^N \), and \( |\nu| = \sum_\nu \nu_j \).
Theorem 4.1. The Gauß norm on \( K[X] \) extends to an ultrametric absolute value on the domain \( K(X) \) (called the Gauß norm on \( K(X) \)) by setting
\[
|f| := \max_{\nu} |a_{\nu}| \quad \text{for } f = \sum_{\nu} a_{\nu}X^\nu \in K(X), \quad f \neq 0.
\]
(See [10], p. 44, Corollary 2.) The set of all \( f \in K(X) \) with \( |f| \leq 1 \) forms a subring \( \mathcal{O}(X) \) of \( K(X) \) (the ring of restricted power series with coefficients in \( \mathcal{O} \)). We identify \( \mathcal{O}(X)/m\mathcal{O}(X) \) with \( \overline{\mathcal{O}[X]} \) in the natural way, and we denote the image of \( f \in \mathcal{O}(X) \) under the canonical surjection \( \mathcal{O}(X) \to \mathcal{O}(X)/m\mathcal{O}(X) = \overline{\mathcal{O}[X]} \) by \( \overline{f} \).

Suppose from now on that \( N \geq 1 \), and let \( X' := (X_1, \ldots, X_{N-1}) \). We have \( K(X') \subseteq K(X) \) and \( \mathcal{O}(X') \subseteq \mathcal{O}(X) \) in a natural way. Every element \( f \in \mathcal{O}(X) \) can be written uniquely in the form
\[
f = \sum_{i=0}^{\infty} f_i X_N^i \quad \text{with } f_i(X') \in \mathcal{O}(X') \quad \text{for all } i \in \mathbb{N},
\]
where the infinite sum converges with respect to the Gauß norm on \( K(X) \). An element \( f \) of \( \mathcal{O}(X) \), written as in (4.1), is called regular in \( X_N \) of degree \( s \) if its reduction \( \overline{f} \) is a unit-monic polynomial of degree \( s \) in \( X_N \), that is,
\begin{itemize}
  \item[(1)] \( \overline{f} \) is a unit in \( \overline{\mathcal{O}[X']} \),
  \item[(2)] \( \overline{f} \) is 0 for all \( i > s \).
\end{itemize}
If \( f \in \mathcal{O}(X'[X_N]) \) is monic of \( X_N \)-degree \( s \) (so that in particular \( f \) is regular in \( X_N \) of degree \( s \), as an element of \( \mathcal{O}(X) \)), 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 \( \mathcal{O}(X) \).) Let \( g \in \mathcal{O}(X) \) be regular in \( X_N \) of degree \( s \). Then for each \( f \in \mathcal{O}(X) \) there are uniquely determined elements \( q \in \mathcal{O}(X) \) and \( r \in \mathcal{O}(X'[X_N]) \) with \( \deg_{X_N} r < s \) such that \( f = qg + r \).

Applying Weierstraß Division with \( f = X_N^N \), we obtain:

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

Regularity can achieved by a change of variables:

**Lemma 4.3.** (Noether Normalization.) Let \( e > 1 \) and \( f \in \mathcal{O}(X) \), and suppose that \( \overline{f} \in \overline{\mathcal{O}[X]} \) is non-zero of degree \( < e \). Let \( T_e : \mathcal{O}(X) \to \mathcal{O}(X) \) be the \( \mathcal{O} \)-automorphism defined by
\[
X_i \mapsto X_i + X_N^{e-1} \quad \text{for } 1 \leq i < N \\
X_N \mapsto X_N.
\]
Then \( T_e(f) \) is regular in \( X_N \) of degree \( < e^N \).

A non-zero element \( f \in K(X) \) is called regular in \( X_N \) of degree \( s \in \mathbb{N} \) if there exists \( b \in \mathcal{O} \) such that \( bf \in \mathcal{O}(X) \) 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 \( K(X) \).) Let \( g \in K(X) \) be regular in \( X_N \) of degree \( s \). Then every \( f \in K(X) \) can be uniquely written as \( f = qg + r \) with \( q \in K(X) \) and \( r \in K(X'[X_N]) \), \( \deg_{X_N} r < s \). In particular, there are a unique
Weierstraß polynomial $w \in \mathcal{O}(X')[[X_N]]$ of degree $s$ and a unique unit $u \in K(X)$ such that $g = u \cdot w$.

Corollary 4.4 and Lemma 4.3 imply that $K\langle X \rangle$ is noetherian. (If $\mathcal{O}$ is a discrete valuation ring, then $\mathcal{O}(X)$ 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 $\alpha_1, \ldots, \alpha_n$ of an algebraic closure $F(X)_{\text{alg}}$ of $F(X)$, we write

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

If $X$ is understood, then we also write $\Delta$ for $\Delta_X$, and we abbreviate $\Delta_X$ to $\Delta'$.

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

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

Let $\alpha \in F(X)_{\text{alg}}$. 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 $X_1, \ldots, X_N$. Among the non-zero polynomials $P(T) = \sum_{i=0}^d a_i T^i$ with coefficients $a_0, \ldots, a_d \in F[X], a_d \neq 0$, 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** $\alpha$.

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

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

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

**Proof.** Let $P(X, T) \in \mathbb{Q}[X, T]$ be a minimal polynomial for $\alpha$, 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 $Q(X', \alpha(x_N)) = 0$, so $d = \Delta'(\alpha(x_N))$. 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 $\mathbb{Q}(X', \alpha_0, \ldots, \alpha_s) \supseteq \mathbb{Q}(X')$, so $\Delta'(\alpha_0, \ldots, \alpha_s) = \Delta'(\alpha(x_N))$. 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 $K\langle X \rangle_{\text{alg}}$ 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 $K\langle X \rangle_{\text{alg}}$ denotes $F_{\text{alg}} \cap K$, and not an algebraic closure $K_{\text{alg}}$ of $K$.)

We put

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

Note that $\mathcal{O}\langle X \rangle_{\text{alg}} = \text{Frac}(\mathcal{O}\langle X \rangle_{\text{alg}}) \cap \mathcal{O}\langle X \rangle$. The $\mathcal{O}$-automorphism $T_\varepsilon$ 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 $T_\varepsilon$, maps $F[X]$ into itself, so by restriction we get an $F$-automorphism of $K\langle X \rangle_{\text{alg}}$ (which in turn restricts to a $D$-automorphism of $\mathcal{O}\langle X \rangle_{\text{alg}}$).

In view of the applications in the later sections, we now make the following simplifying assumption: for every $a \in \mathcal{O}$ there exists $b \in \mathcal{O} \cap F_{\text{alg}}$ with $[a] = [b]$. (So for every $f \in K\langle X \rangle_{\text{alg}}$ which is regular in $X_N$ of degree $s \in \mathbb{N}$ there exists $b \in \mathcal{O} \cap F_{\text{alg}}$ such that $bf \in \mathcal{O}\langle X \rangle_{\text{alg}}$ 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 $K\langle X \rangle_{\text{alg}}$.) In Corollary 4.4, if in addition $f$ and $g$ are in $K\langle X \rangle_{\text{alg}}$, then $q \in K\langle X \rangle_{\text{alg}}$ and $r \in K\langle X \rangle_{\text{alg}}[X_N]$. If moreover we write

$$r = r_0 + r_1 X_N + \cdots + r_{s-1} X_N^{s-1} \quad \text{with } r_0, \ldots, r_{s-1} \in K\langle X \rangle_{\text{alg}},$$

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

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

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 $w \in \mathcal{O}\langle X \rangle_{\text{alg}}[X_N]$ a Weierstraß polynomial of degree $s$. We denote by $\Omega$ an algebraic closure of $\text{Frac}(K\langle X \rangle')$, and we fix a zero $\xi \in \Omega$ of $w$. (We will sometimes just say that $\xi$ 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 $\Omega$ is indeed just obtained by substitution of $\xi$ for $X_N$ into the polynomial $f$. For notational simplicity, from now on we just write $f(\xi)$ for $f(X', \xi)$. The map $f \mapsto f(\xi) : K\langle X \rangle \to \Omega$ clearly is $K\langle X' \rangle$-linear. In fact, we have:

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

**Proof.** We have to show that

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

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 $f_i = q_i g + r_f$ for $i = 1, 2$. Let also $q \in K\langle X \rangle$ be the quotient obtained by dividing $f_1 f_2$ by $g$, so $f_1 f_2 = qg + r_{f_1 f_2}$. We also have $f_1 f_2 = (q_1 g + r_{f_1})(q_2 g + r_{f_2})$, and
subtracting these two representations of \( f_1f_2 \) one sees that \( r_{f_1f_2} - r_{f_1}r_{f_2} \in K(X)[X_N] \) is a multiple of \( g \) in \( K(X) \), 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 \( f \in K(X)_{\text{alg}} \) be non-zero, and let \( P(T) = \sum_{i=0}^d a_i T^i \in F[X,T] \), with \( a_0, \ldots, a_d \in F[X], a_d \neq 0 \), be a minimal polynomial for \( f \). Note that \( a_0 \neq 0 \). Since \( P(f) = 0 \) and \( f \mapsto f(\xi) \) is a \( K(X') \)-algebra homomorphism, we have

\[
0 = (P(f))(\xi) = \sum_{i=0}^d a_i (\xi^i) = \sum_{i=0}^d a_i (\xi^i) (f(\xi))^i.
\]  

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

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

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

\[
\Delta(\xi_0, \ldots, \xi_{s-1}) \leq \left( \deg_{X_N} P(0) \right)^s.
\]  

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

\[
[F(\xi', \xi_0, \ldots, \xi_{s-1}, f(\xi_0), \ldots, f(\xi_{s-1})): F(\xi', \xi_0, \ldots, \xi_{s-1})] \leq \Delta(f)^s.
\]

Moreover:

**Lemma 4.8.** Let \( f_1, \ldots, f_m \in K(X)_{\text{alg}}, m \geq 1 \). There exists \( \alpha \in K(X)_{\text{alg}} \) such that

\[
\Delta(\alpha) \leq \Delta(f_1, \ldots, f_m)
\]

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

**Proof.** Put \( \alpha = f_1 + c_2f_2 + \cdots + c_m f_m \in K(X)_{\text{alg}} \) with as yet undetermined \( c_2, \ldots, c_m \in F \). Then \( \Delta(\alpha) \leq \Delta(f_1, \ldots, f_m) \). Moreover, \( \alpha(\xi) \) is a primitive element of the field extension \( F(X', f_1(\xi), \ldots, f_m(\xi)) \supseteq F(X') \) for all but finitely many \( (c_2, \ldots, c_m) \in F^{m-1} \).

Since \( g \) only has finitely (namely, at most \( s \)) many zeros, we can choose \( (c_2, \ldots, c_m) \in F^{m-1} \) such that \( \alpha(\xi) \) has the required property for every zero \( \xi \) of \( g \). \( \square \)

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

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

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

**Lemma 4.9.** Let \( f \in K(X)_{\text{alg}} \) and \( g_1, \ldots, g_N \in \Theta(Y)_{\text{alg}} \). Then \( f(g_1, \ldots, g_N) \in K(Y)_{\text{alg}} \) with \( \Delta_Y(f(g_1, \ldots, g_N)) \leq \Delta(f) \Delta_Y(g_1) \cdots \Delta_Y(g_N) \).
Proof. We prove the following slightly more general statement, by induction on $N$: if $f \in K\langle X, Y \rangle_{\text{alg}}$ and $g_1, \ldots, g_N \in O(Y)_{\text{alg}}$, then $f(g_1, \ldots, g_N, Y) \in K(Y)_{\text{alg}}$ with $\Delta_Y(f(g_1, \ldots, g_N, Y)) \leq \Delta_X(f) \Delta_Y(g_1) \cdots \Delta_Y(g_N)$. For $N = 0$, this is trivial. Let $N > 0$ and assume the corresponding statement is true for $N - 1$ instead of $N$. Let $f \in K\langle X, Y \rangle_{\text{alg}}$ and $g_1, \ldots, g_N \in O(Y)_{\text{alg}}$. We set $p(X_N) := X_N - g_N \in O(X', Y)[X_N]$, 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 $f(X', g_N, Y) \in \Delta_X(f) \Delta_Y(g_N)$. Now apply the inductive hypothesis to $f(X', g_N, Y)$.

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

Proof. Let $f \in K\langle X \rangle_{\text{alg}}$ and $j \in \{1, \ldots, N\}$. Let $P(T) = \sum_{i=0}^d a_i T^i \in F[X, T]$, $a_0, \ldots, a_d \in F[X]$, 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 $\text{char } F = 0$), and since $\partial a_i / \partial X_j \in F[X]$ for $i = 0, \ldots, d$, this shows that $\partial f / \partial X_j \in F(X, f)$.

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

$$(\partial f / \partial X_N^i)(\xi) = F(X', \xi, f(\xi))$$

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 $P(X, T) = \sum_{i=0}^d a_i(X) T^i \in F[X, T]$, $a_0, \ldots, a_d \in F[X]$, $a_d(X', \xi) \neq 0$, be a polynomial of minimal degree in $T$ such that $P(X', \xi, f(\xi)) = 0$. 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 $\xi$ 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 $q := \sum_{i=1}^d i a_i(X', \xi) f(\xi)^{i-1}$ is non-zero by choice of $P$, and hence

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

as required.

Corollary 4.12. If $f = \sum_{i=0}^\infty f_i X_N^i \in K\langle X \rangle_{\text{alg}}$ with $f_i \in K\langle X' \rangle$ for all $i \in \mathbb{N}$, then $f_i \in K\langle X' \rangle_{\text{alg}}$ with $\Delta'(f_i) \leq \Delta(f)$ for all $i \in \mathbb{N}$.
Proof. This follows from \( f_i = \frac{1}{n} \frac{\partial f_i}{\partial X_N}(X', 0) \) 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
\[
g(\xi) = \left( \frac{\partial g}{\partial X_N} \right)(\xi) = \cdots = \left( \frac{\partial^{n-1} g}{\partial X_N^{n-1}} \right)(\xi) = 0, \quad \left( \frac{\partial^n g}{\partial X_N^n} \right)(\xi) \neq 0.
\]

4.2. Proof of the Weierstraß Division Theorem for \( K(X)_{\text{alg}} \). Let \( f \) and \( g \) be as in Theorem 4.6. By Weierstraß Division in \( K(X) \),
\[
f = qg + r \quad \text{with} \quad q \in K(X), \ r \in K(X)[X_N], \ \deg_{X_N} r < s. \quad (4.5)
\]
Write \( r = r(X_N) = \sum_{j=0}^{s-1} r_j X_N^j \) with \( r_0, \ldots, r_{s-1} \in K(X') \). For the first statement in Theorem 4.6, it suffices to show that \( q = (f - r)/g \in \text{Frac}(K(X)_{\text{alg}}) \cap K(X) = K(X)_{\text{alg}} \).

Let now \( Q(T) \in F[X, T] \) be a minimal polynomial for \( g \). We need to prove that
\[
r_0, \ldots, r_{s-1} \in F(X')_{\text{alg}}
\]
and
\[
\Delta'(r_0, \ldots, r_{s-1}) \leq (\Delta(f) \deg_{X_N} Q(0))^s.
\]
To see this, let \( \xi_0, \ldots, \xi_{s-1} \) be the zeros of \( g \) in \( F(X')_{\text{alg}} \). Using Lemma 4.7 and (4.5), we get
\[
f(\xi_i) = r(\xi_i) = \sum_{j=0}^{s-1} r_j \xi_i^j \quad \text{for} \quad i = 0, \ldots, s - 1.
\]

We have \( f(\xi_i) \in F(X')_{\text{alg}} \) by Lemma 4.7, since \( f \in K(X)_{\text{alg}} \) and \( \xi_i \in F(X')_{\text{alg}} \). Hence by (4.3), (4.4) it suffices to show that
\[
r_0, \ldots, r_{s-1} \in F(f(\xi_0), \ldots, f(\xi_{s-1}), \xi_0, \ldots, \xi_{s-1}). \quad (4.6)
\]

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:
\[
\begin{bmatrix}
1 & \xi_0 & \cdots & \xi_0^{s-1} \\
1 & \xi_1 & \cdots & \xi_1^{s-1} \\
\vdots & \vdots & \ddots & \vdots \\
1 & \xi_{s-1} & \cdots & \xi_{s-1}^{s-1}
\end{bmatrix}
\begin{bmatrix}
r_0 \\
r_1 \\
\vdots \\
r_{s-1}
\end{bmatrix}
=
\begin{bmatrix}
f(\xi_0) \\
f(\xi_1) \\
\vdots \\
f(\xi_{s-1})
\end{bmatrix} \quad (4.7)
\]
(The Vandermonde matrix in (4.7) is non-singular.) So (4.6) follows by Cramer’s Rule.

Consider now the case that the zero \( \xi_0 \) of \( g \) has multiplicity 2, and \( \xi_1, \ldots, \xi_{s-2} \) each has multiplicity 1. Then we may use the identity
\[
\frac{\partial f}{\partial X_N} = q \cdot \frac{\partial g}{\partial X_N} + \frac{\partial q}{\partial X_N} \cdot g + (r_1 + 2r_2 X_N + \cdots + (s - 1)r_{s-1} X_N^{s-2})
\]
to get the following system of linear equations:
\[
\begin{bmatrix}
1 & \xi_0 & \xi_0^2 & \cdots & \xi_0^{s-1} \\
0 & 1 & 2\xi_0 & \cdots & (s - 1)\xi_0^{s-2} \\
1 & \xi_1 & \xi_1^2 & \cdots & \xi_1^{s-1} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & \xi_{s-2} & \xi_{s-2}^2 & \cdots & \xi_{s-2}^{s-1}
\end{bmatrix}
\begin{bmatrix}
r_0 \\
r_1 \\
r_2 \\
\vdots \\
r_{s-1}
\end{bmatrix}
=
\begin{bmatrix}
f(\xi_0) \\
(\partial f/\partial X_N)(\xi_0) \\
f(\xi_1) \\
\vdots \\
f(\xi_{s-2})
\end{bmatrix} \quad (4.8)
\]
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 $\mu$, suppose that each $\xi_i$ has multiplicity $\mu_i$, and let $\xi = (\xi_0, \ldots, \xi_n)$. Then $[r_0, \ldots, r_{s-1}]^T$ 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.) \hfill \Box

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 $f_i, g_i \in K(X)_{\text{alg}}$ and $r_{i0}, \ldots, r_{i,s-1} \in K(X')_{\text{alg}}$ such that

\[ f_i = q_i + r_i \quad \text{with} \quad r_i = r_{i0} + r_{i1}X_N + \cdots + r_{i,s-1}X_N^{s-1} \in K(X)_{\text{alg}}. \]

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

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

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

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

Now choose $\alpha \in K(X)_{\text{alg}}$ with the properties stated in Lemma 4.8. Then

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

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

As in the case of Weierstraß Division over $K(X)$, from Theorem 4.6 we get:

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

\[ w = w_0 + w_1X_N + \cdots + w_{s-1}X_N^{s-1} + X_N^s \]

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

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

As a consequence, the ring $K(X)_{\text{alg}}$ 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 $(\mathbb{Q}_p)_{\text{alg}}$ of the field $\mathbb{Q}_p$ of $p$-adic numbers. We denote by $\nu_p$ the unique extension of the $p$-adic valuation on $\mathbb{Q}$ to a valuation on $\mathbb{C}_p$, and we put $|\cdot|_p = |\cdot|_{\log p}$. We take $D = \mathbb{Z}$; then the ring $\mathbb{C}_p(X)_{\text{alg}}$ consists of all power series in $\mathbb{C}_p(X)$ which are algebraic over $\mathbb{Q}(X)$. Let $f \in \mathbb{C}_p(X)_{\text{alg}}$ 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

\[ P(T) = a_0 + a_1T + \cdots + a_dT^d \]


with $a_0, \ldots, a_d \in \mathbb{Z}[X]$ without common factor in $\mathbb{Z}[X]$. By Corollary 2.22
\[ h_{\text{coeff}}(P) = h([a_0 : \cdots : a_d]) \leq d(h(f) + \log 2). \] (4.9)
On the other hand, for all $j = 0, \ldots, d$ with $a_j \neq 0$,
\[ h([a_0 : \cdots : a_d]) \geq \deg_{(X')} a_j + m(a_j) \geq (1 - \log 2) \deg_{(X')} a_j + \log \|a_j\|_1 \geq (1 - \log 2)(\deg_{(X')} a_j + \log \|a_j\|_1), \] (4.10)
by Lemmas 3.7 and 1.4, (5). Combining (4.9) and (4.10) we get
\[ \deg_{(X')} a_j \leq d(h(f) + \log 2), \quad \log \|a_j\|_1 \leq d(h(f) + \log 2). \] (4.11)
Using (3.17) we also obtain from (4.10):
\[ h(a_j) \leq \frac{d}{1 - \log 2}(h(f) + \log 2) \quad \text{for all } j = 0, \ldots, d. \]
Moreover, by Proposition 3.10
\[ h(P) \leq d(h(f) + \log 2 + 1) + \log(d + 1) \leq 4d(h(f) + \log 2). \] (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 $f_1, \ldots, f_m, g \in \mathbb{C}_p(X)_{\text{alg}}, m \geq 1$, and suppose $g$ is regular in $X_N$ of degree $s \in \mathbb{N}$. For each $i = 1, \ldots, m$ let $q_i \in \mathbb{C}_p(X)_{\text{alg}}$ and $r_i \in \mathbb{C}_p(X')_{\text{alg}}[X_N]$ such that $f_i = q_i g + r_i$ and $\deg_{X_N} r_i < s$. Write
\[ r_i = r_{i0} + r_{i1} X_N + \cdots + r_{is-1} X_N^{s-1} \] with $r_{ij} \in \mathbb{C}_p(X')_{\text{alg}}$ for all $i, j$.

Then
\[ \Delta'(r_{ij} : 1 \leq i \leq m, \ 0 \leq j < s) \leq (\Delta(f_1, \ldots, f_m) \Delta(g)(h(g) + \log 2))^s \]
and
\[ h_{\text{max}}(r_{ij} : 1 \leq i \leq m, \ 0 \leq j < s) \leq O(1)^s (\Delta(f_1, \ldots, f_m))^s \Delta(g)(h_{\text{max}}(f_1, \ldots, f_m) + \log 2)(h(g) + \log 2). \]

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

**Lemma 4.16.** For each $j \in \mathbb{N}$, we have
\[ h(\partial^j f / \partial X_N^j) + \log 2 \leq (6d)^j(h(f) + \log 2). \]

**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)
\[ h_{\text{coeff}}(P') \leq h_{\text{coeff}}(P) + \log d \leq d(h(f) + \log 2) + \log d. \]
Hence by (2.17)
\[ h(P'(f)) \leq (d - 1)h(f) + h_{\text{coeff}}(P') + \log d \leq (2d - 1)h(f) + \log(2^d d^2). \]
By (3.19) and (4.11) we have
\[ h_{\text{coeff}}(P*) \leq h_{\text{coeff}}(P) + d + \log(d + 1) + \log d(h(f) + \log 2) \]
and hence by (2.17), (4.9) and using that \( \log(x + \log 2) \leq x \) and \( x \leq \log 2^x \) for all \( x \in \mathbb{R}^\geq 0 \) we get
\[
(2d + 1)h(f) + \log(2^d(1+d)^2d) + \log(h(f) + \log 2).
\]
So we have
\[
h(\partial f / \partial X_N) \leq h(P^*(f)) + h(P'(f)) \leq 4dh(f) + \log(2^{2d+1}d^3(d + 1)^2).
\]
A computation shows that \( \log(2^{2d+1}d^3(d + 1)^2) \leq (6d - 1) \log 2 \), hence
\[
h(\partial f / \partial X_N) + \log 2 \leq 6d(h(f) + \log 2).
\]
The lemma now follows by induction on \( j \geq 1 \).

For the proof of Theorem 4.15, let \( g \in \mathbb{C}_p(X)_{\text{alg}} \) be regular in \( X_N \) of degree \( s \). Let
\[
Q(T) = b_0 + b_1T + \cdots + b_eT^e
\]
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
\[
\text{deg}(X) b_j \leq \Delta(g)(h(g) + \log 2), \quad \log ||b_j||_1 \leq \Delta(g)(h(g) + \log 2) \quad (4.14)
\]
and
\[
h(b_j) \leq \frac{\Delta(g)}{1 - \log 2}(h(g) + \log 2) \quad (4.15)
\]
for all \( j = 0, \ldots, e \). (By (4.11), applied to \( f \) in place of \( f \).)

For each \( i = 1, \ldots, m \) let \( q_i \in \mathbb{C}_p(X)_{\text{alg}} \) and let \( r_i \in \mathbb{C}_p(X')_{\text{alg}}[X_N] \) 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
\[
\Delta'(r_{ij} : 1 \leq i \leq m, 0 \leq j < s) \leq (\Delta(f_1, \ldots, f_m)\Delta(g)(h(g) + \log 2))^s.
\]
This is the first estimate in Theorem 4.15.

Let \( \xi_0, \ldots, \xi_{s-1} \) be the zeros of \( g \) in \( \mathbb{Q}(X')_{\text{alg}} \). Recall that for \( h \in \mathbb{C}_p(X) \) we write \( h(\xi_i) \) for \( r_h(\xi_i) \), where \( r_h \in \mathbb{C}_p(X')[X_N] \) is the remainder obtained through Weierstraß Division of \( h \) by \( g \). For each \( \xi_i \) 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)
\[
\sum_{i=0}^{s-1} h(\xi_i) \leq h_{\text{coeff}}(b_0) + \text{deg}_{X_N} b_0 \log 2 \leq 4e(h(g) + \log 2).
\]
(Here, \( h_{\text{coeff}}(b_0) \) 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 $\mathbb{C}_p(X)_{\text{alg}}$ and
\[ w = w_0 + \cdots + w_{s-1}X_N^{s-1} + X_N^s \quad \text{with} \quad w_0, w_1, \ldots, w_{s-1} \in \mathbb{C}_p(X)_{\text{alg}} \]
be a Weierstraß polynomial of degree $s$ such that $g = u \cdot w$. Then
\[ \Delta'(w_0, \ldots, w_{s-1}) \leq \left( \Delta(g)(h(g) + \log 2) \right)^s, \]
\[ h(w_0, \ldots, w_{s-1}) \leq s \log 2 + 4\Delta(g)(h(g) + \log 2). \]

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
\[ h(f(\xi)) \leq C_0 \cdot \Delta(f)\Delta(g)(h(f) + \log 2)(h(g) + \log 2). \]

Hence for all $j \in \mathbb{N}$
\[ h((\partial^j f/\partial X_N^j)(\xi)) \leq C_0 \cdot 6^j\Delta(f)^{j+1}\Delta(g)(h(f) + \log 2)(h(g) + \log 2). \]

Proof. We have $b_0(X', \xi) = P(X', \xi, f(\xi)) = 0$. So if $r(T) = \text{res}_{X_N}(b_0, P) \in \mathbb{Z}[X', T]$ 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 $r(f(\xi)) = 0$. 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):
\[ h_{\text{coeff}}(r) \leq 9 \deg h(f) + \log 2)(h(g) + \log 2). \]

So by Corollary 2.22, for some constant $C_0 > 0$:
\[ h(f(\xi)) \leq h_{\text{coeff}}(r) + \log 2 \leq C_0 \cdot \deg h(f) + \log 2)(h(g) + \log 2) \]
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 $\mu_i := \text{the multiplicity of } \xi_i$. (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 $\mathbb{Q}(X')_{\text{alg}}$). If $s = 1$ then $h(A) = 0$; otherwise, from Proposition 2.32, Lemma 2.26, (4), and $h([1 : 2 : \cdots : s - 1]) = \log(s - 1)$ we get that
\[ h(A) \leq (s - 1)(h(\xi_0, \ldots, \xi_n) + \log(s - 1)) \]
\[ \leq (s - 1)(4e(h(g) + \log 2) + \log(s - 1)). \]

Hence from Corollary 2.31 we obtain
\[ h(A^{-1}) \leq s \left( \frac{1}{2} \log s + h(A) \right) \]
\[ \leq \frac{1}{2}s \log s + 4e(h(g) + \log 2) \]
\[ \leq s(s - 1/2) \log s + 4e(h(g) + \log 2). \]
hence there exists a universal constant $C_1 > 0$ with
$$h(A^{-1}) + \log s \leq C_1 s^3 e(h(g) + \log 2).$$

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
$$\left[ f_i(\xi_0), \frac{\partial f_i}{\partial X^N}(\xi_0), \ldots, \frac{\partial^{\mu_1 - 1} f_i}{\partial X^{\mu_1 - 1}}(\xi_0), \ldots, f_i(\xi_n), \frac{\partial f_i}{\partial X^N}(\xi_n), \ldots, \frac{\partial^{\mu_n - 1} f_i}{\partial X^{\mu_n - 1}}(\xi_n) \right].$$

From Lemma 4.19 we get
$$h_{\text{max}}(B) = \max_{1 \leq i \leq m, 0 \leq j \leq n} h((\partial^k f_i / \partial X^N)(\xi_j))$$
$$\leq C_0 s \cdot (6d)^s e(h_{\text{max}}(f_1, \ldots, f_m) + \log 2)(h(g) + \log 2).$$

Since $R = A^{-1}B$, by Lemma 2.30 therefore
$$h_{\text{max}}(R) \leq s(h_{\text{max}}(A^{-1}) + h_{\text{max}}(B)) + \log s$$
$$\leq (C_2 d)^s e(h_{\text{max}}(f_1, \ldots, f_m) + \log 2)(h(g) + \log 2)$$
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 $y_0, \ldots, y_{s-1} \in \mathbb{C}_p(X')_{\text{alg}}, s > 0$. Then
$$h(y) \leq h(y_0, \ldots, y_{s-1}) + (s - 1 + \log s)$$
by (2.17), Example 3.8 and Lemma 2.26, (3). Therefore the bound on the quantity $h(r_{ij}, \ldots, h_{r_{s-1}})$ in Theorem 4.15 also entails an upper bound on $h(r_{ij})$.

(2) Theorem 4.15 in conjunction with Lemma 2.26 immediately gives a bound on the affine heights in the Weierstraß Division Theorem:
$$h(r_{ij}) : 1 \leq i \leq m, 0 \leq j < s \leq mO(1)^s (\Delta(f_1, \ldots, f_m))^s \Delta(g)(h(f_1, \ldots, f_m) + \log 2)(h(g) + \log 2).$$

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 $f \in \mathbb{C}_p(X')_{\text{alg}}$, 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 $\mathbb{C}_p(X)$ given by
$$X_i \mapsto X_i + X_N^{N-i} \quad \text{for } 1 \leq i < N, \quad X_N \mapsto X_N.$$
(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 $T_e(a_d) \neq 0$. In particular, $\Delta(T_e(f)) = d$.

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

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

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

Lemma 3.7 and (4.11) yield

\[
h_{\text{coeff}}(\phi_0, \ldots, \phi_d) = \max_i \log ||a_i||_\infty \leq d(h(f) + \log 2).
\]

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

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

hence by Corollary 2.22 we get

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

Remark 4.22. Similarly, one shows that \( h(T_e^{-1}(f)) = O(e^{N-1}(h(f) + \log 2)) \), 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.** \( |v_p(f)| \leq \Delta(f)(h(f) + \log 2)/\log p. \)

Let \( j \in \{0, \ldots, d\}. \) 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 \( a_j: \)

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

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

\[
Q(T) = b_0 + b_1T + \cdots + b_nT^n \in K[T] \quad (b_0, \ldots, b_n \in K, b_n \neq 0).
\]

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

**Proof.** We have

\[
|v(\alpha)| \leq \max \left\{ \left| \frac{v(b_j) - v(b_i)}{i - j} \right| : i \neq j, b_i, b_j \neq 0 \right\} \leq \max\{|v(b_i)| : b_i \neq 0\},
\]
as required. \( \square \)

**Remark 4.25.** If \( f \) lies in \( (\mathbb{C}_p)_{\text{alg}} \) 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 \( f \in (\mathbb{C}_p)_{\text{alg}}, \) then by (2.15)

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

so \( |v_p(f)| \leq \frac{\Delta(f)h(f)}{\log p} \) 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. \( \square \)
Lemma 4.26. Suppose $|f|_p = 1$. Then

$$\deg_X(\overline{f}) \leq 3d(h(f) + \log 2), \quad \deg_X f \leq d(h(f) + \log 2).$$

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

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

Moreover, for all $j = 0, \ldots, d$ with $a_j \neq 0$,

$$h_X(a_j) = \deg_X(a_j) \leq \deg_X(a_j) \leq d(h(f) + \log 2),$$

by (4.11). Therefore, by Corollary 2.22:

$$\deg_X(\overline{f}) = h_X(\overline{f}) \leq \frac{1}{d} h_X(\overline{a_0} : \cdots : \overline{a_d}) + \log 2 \leq \frac{1}{d} \sum_j \deg_X(\overline{a_j}) + \log 2 \leq (d + 1)(h(f) + \log 2) + \log 2.$$

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 $Q(T) = P(T) \in K[T]$, and $\alpha = \overline{f}$. \hfill \square

Now assume that $f \in \mathbb{Z}_p(X)_{\text{alg}}$. Then there are uniquely determined polynomials $f_0, f_1, \ldots$ in $\mathbb{Z}[X]$ such that $||f_j||_\infty < p$ for all $j$ and

$$f = f_0 + pf_1 + p^2 f_2 + \cdots \in \mathbb{Z}_p(X).$$

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

$$\deg_X(f_0) = \deg_X(\overline{f}) \leq 3\Delta(\overline{f})(h(f) + \log 2),$$

so by (3.18):

$$h(f_0) \leq 2 \deg_X(f_0) + \log ||f_0||_\infty \leq 6d(h(f) + \log 2) + \log p.$$

Now put $f_1 := (f - f_0)/p \in \mathbb{Z}_p(X)_{\text{alg}}$. Then

$$h(f_1) \leq h(f) + h(f_0) + \log 2 + \log p \leq h(f) + 6d(h(f) + \log 2) + \log 2 + 2 \log p.$$

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

$$h(f_1) + \log 2 \leq C \log p \Delta(f)(h(f) + \log 2)$$

and hence, inductively,

$$\deg_X(f_j) \leq (C \log p)^j \Delta(f)^{j+1}(h(f) + \log 2) \quad \text{for all } j \in \mathbb{N}. \quad (4.16)$$

Now let $j \geq 1$ and put $f_j := f_0 + pf_1 + \cdots + p^{j-1} f_{j-1} \in \mathbb{Z}[X]$. Then $f \equiv f_j \mod p^j$ and $||f_j|| < p^j$, and by (3.18), (4.16):

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

These bounds may be used to show that power series like $\sum_{i=0}^{\infty} p^i X^{2^i} \in \mathbb{Z}_p(X)$ (where $X$ is a single indeterminate) are not algebraic over $\mathbb{Q}(X)$. AN EFFECTIVE WEIERSTRASS DIVISION THEOREM 45
5. Hermann’s Method for Restricted Power Series

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

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

and the vector $b = [b_1, \ldots, b_m]^\text{tr}$ have entries from the ring $\mathbb{Z}_p(X)$ of restricted power series with coefficients in $\mathbb{Z}_p$. A necessary condition for solvability of the system $Ay = b$ in $\mathbb{Z}_p(X)$ is certainly that this system is solvable in $\mathbb{Q}_p(X)$, and that for all $e \geq 1$, the system $\overline{A}y = \overline{b}$, obtained from $Ay = b$ by applying the canonical homomorphism $a \mapsto \overline{a} = a \mod p^e$ to the coefficients of $A$ and $b$, is solvable in $\mathbb{Z}_p(X)$ mod $p^e = (\mathbb{Z}/p^e\mathbb{Z})[X]$.

(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 $\mathbb{Q}_p(X)$ and in $\mathbb{Z}_p(X)$ mod $p^e$ for large enough $e$ implies solvability of $Ay = b$ in $\mathbb{Z}_p(X)$. 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 $d, m \geq 1$, $h \geq 0$ there exists a positive integer $E = E(N, d, h, m)$ with

$$2^{O(1)N(m^2d(h+1))^N} (N \text{ many 2's})$$

such that the following holds: for every $A \in (\mathbb{Z}_p(X)_{\text{alg}})^{m \times n}$ with $\Delta(A) \leq d$ and $h(A) \leq h$ and every $b \in \mathbb{Z}_p(X)^m$, the system

$$Ay = b$$

is solvable in $\mathbb{Z}_p(X)$ if and only if it is solvable in $\mathbb{Q}_p(X)$ and solvable mod $p^E$.

Provided that $b \in (\mathbb{Z}_p(X)_{\text{alg}})^n$, we will also show: if $Ay = b$ is solvable in $\mathbb{Z}_p(X)$, then there exists a solution $y \in (\mathbb{Z}_p(X)_{\text{alg}})^n$ 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 $f_0, f_1, \ldots, f_n \in \mathbb{Z}_p(X)_{\text{alg}}$ with $f_0 \in (f_1, \ldots, f_n)\mathbb{Q}_p(X)$ and $\Delta(f_1, \ldots, f_n) \leq d$, $h(f_1, \ldots, f_n) \leq h$, we have

$$f_0 \in (f_1, \ldots, f_n)\mathbb{Z}_p(X) \iff \mathcal{F}_0 \in (\mathcal{F}_1, \ldots, \mathcal{F}_n)(\mathbb{Z}/p^E\mathbb{Z})[X],$$

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 $b = [b_1, \ldots, b_m]^\text{tr} \in D^m$. We are mostly interested in the case where $D = \mathbb{Z}_p(X)$ or $D = \mathbb{Z}_p(X)_{\text{alg}}$; 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

$$[\begin{array}{cccc} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{array}] [\begin{array}{c} y_1 \\ \vdots \\ y_n \end{array}] = [\begin{array}{c} b_1 \\ \vdots \\ b_m \end{array}]$$

(1)

(or: $Ay = b$) has a solution $y = [y_1, \ldots, y_n]^\text{tr} \in D^n$, and if it does, effectively find such a solution. Of course, we may assume $A \neq 0$, so the rank $r = \text{rank}_F(A)$ of $A$ (considered as a matrix over $F$) is $\geq 1$. Let $\Delta$ be an $r \times r$-submatrix of $A$ with $\delta = \det \Delta \neq 0$. 

After rearranging the order of the equations and permuting the unknowns \( y_1, \ldots, y_n \) in (I) we may assume that \( \Delta = (a_{ij})_{1 \leq i, j \leq r} \). A necessary condition for (I) to have a solution \( y \in D^n \) is clearly that

\[
\text{rank}_F(A) = \text{rank}_F(A, b). \tag{NC}
\]

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 \( \Delta^{-1} \) of \( \Delta \), (I) turns into the system

\[
\begin{bmatrix}
  \delta & c_{1,r+1} & \cdots & c_{1,n} \\
  \delta & c_{2,r+1} & \cdots & c_{2,n} \\
  \vdots & \vdots & \ddots & \vdots \\
  \delta & c_{r,r+1} & \cdots & c_{rn}
\end{bmatrix}
\begin{bmatrix}
  y_1 \\
  y_2 \\
  \vdots \\
  y_n
\end{bmatrix}
= \begin{bmatrix}
  d_1 \\
  d_2 \\
  \vdots \\
  d_r
\end{bmatrix} \tag{S}
\]

(with \( c_{ij}, d_i \in D \) for \( 1 \leq i \leq r < j \leq n \)) which has the same solutions in every domain extending \( D \) as (I). Clearly, a sufficient condition for (S) to have a solution \( y = [y_1, \ldots, y_n]^T \in D^n \) is that \( d_1, \ldots, d_r \) are each divisible by \( \delta \). This will be the case if \( \delta \) is a unit; then a solution to (S) (and hence to (I)) is given by

\[
y_j = \begin{cases} 
d_j/\delta & \text{for } 1 \leq j \leq r, \\
0 & \text{for } r < j \leq n. 
\end{cases}
\]

Suppose \( \delta \) is not a unit, so \( \overline{D} = D/\delta D \neq 0 \). Then, reducing the coefficients in (S) modulo \( \delta \), the system (S) turns into

\[
\begin{bmatrix}
  \overline{c_{1,r+1}} & \cdots & \overline{c_{1,n}} \\
  \vdots & \ddots & \vdots \\
  \overline{c_{r,r+1}} & \cdots & \overline{c_{rn}}
\end{bmatrix}
\begin{bmatrix}
  y_{r+1} \\
  \vdots \\
  y_n
\end{bmatrix}
= \begin{bmatrix}
  \overline{d_1} \\
  \vdots \\
  \overline{d_r}
\end{bmatrix} \tag{\overline{S}}
\]

over \( \overline{D} \). (Here \( \overline{a} \) denotes the image of \( a \in D \) in \( \overline{D} \).) For any \( y_{r+1}, \ldots, y_n \in \overline{D} \) with the property that \( [\overline{y_{r+1}}, \ldots, \overline{y_n}]^T \) is a solution of the reduced system (\( \overline{S} \)) there are uniquely determined \( y_1, \ldots, y_r \in D \) such that

\[
[y_1, \ldots, y_r, y_{r+1}, \ldots, y_n]^T \in D^n
\]

is a solution of (S), and hence of (I). In particular, (I) is solvable in \( D \) if and only if (\( \overline{S} \)) 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 \( \delta \) with \( \delta \bmod p \neq 0 \). Then, after applying the \( \mathbb{Z}_p \)-automorphism \( T_d \) of \( D \) given by Lemma 4.3, where \( d > \deg X \bmod p \), to all coefficients of (S), we may even assume that \( \delta \) is regular in \( X_N \) of some degree \( s \). By Weierstraß Division we now have

\[
\overline{D} = D/\delta D \cong D' \oplus D'X_N^s \oplus \cdots \oplus D'X_N^{s-1}
\]
as $D'$-algebras. This allows us to replace the system $(\mathcal{S})$ with an equivalent system over $D'$ as follows. Each $\tilde{c}_{ij}$ can be written uniquely as

$$c_{ij} = c_{ij,0} + c_{ij,1}X_N + c_{ij,2}X_N^2 + \cdots + c_{ij,s-1}X_N^{s-1}$$

with $c_{ij,0}, \ldots, c_{ij,s-1} \in D'$, $1 \leq i \leq r < j \leq n$. We also write each $\tilde{d}_i$ and each power $X_N$ ($t \geq s$) in this way,

$$\tilde{d}_i = d_{i0} + d_{i1}X_N + d_{i2}X_N^2 + \cdots + d_{is}X_N^{s-1}$$

with $d_{ik} \in D'$, $1 \leq i \leq r$, $0 \leq k < s,

$$X_N^{-t} = \xi_{00} + \xi_{01}X_N + \cdots + \xi_{0,s-1}X_N^{s-1}$$

(5.1)

with $\xi_{ik} \in D'$, and also each unknown $y_j$ as

$$y_j = y_{j0} + y_{j1}X_N + \cdots + y_{js}X_N^{s-1}$$

with new unknowns $y_{jk}$ ($r < j \leq n$, $0 \leq k < s$) ranging over $D'$. The coefficient matrix of the system $(\mathcal{S})$ above may then be written as

$$C(0) + C(1)X_N + \cdots + C(s-1)X_N^{s-1},$$

where

$$C(k) = \left(c_{ij,k}\right)_{1 \leq i \leq r, 1 \leq j \leq n} \in (D')^{r \times (n-r)}$$

for $k = 0, \ldots, s - 1$. Similarly we may write

$$\begin{bmatrix} d_1 \\ \vdots \\ d_r \end{bmatrix} = d(0) + d(1)X_N + \cdots + d(s-1)X_N^{s-1}, \quad d(k) = \begin{bmatrix} d_{1k} \\ \vdots \\ d_{rk} \end{bmatrix},$$

and also

$$\begin{bmatrix} y_{r+1} \\ \vdots \\ y_n \end{bmatrix} = y(0) + y(1)X_N + \cdots + y(s-1)X_N^{s-1}, \quad y(k) = \begin{bmatrix} y_{r+1,k} \\ \vdots \\ y_{nk} \end{bmatrix}.$$
for \(k = 0, \ldots, s - 1\), where we put \(C(t) := 0\) for \(t < 0\). Combining these systems into a single one, we obtain a system

\[
A' y' = b',
\]

where

\[
A' \in (D')^{m' \times n'}, \quad b' \in (D')^{m'}, \quad m' = rs, \quad n' = s(n - r),
\]

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 \(D\), whose solutions in \(D\) are in one-to-one correspondence with the solutions in \(D'\) of \(A'y' = b'\). Therefore, we have

\[
Ay \in \mathbb{Q}_p(X), \quad \text{by Weierstraß Division for } \mathbb{Q}_p(X).
\]

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 \(\bigcap_{n \in \mathbb{N}} p^nD = (0)\). 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 \(v(0) := -\infty \in \mathbb{N}\). We also put \(F = \text{Frac}(D)\) and \(F(p) = \text{Frac}(D/pD)\).

Let \(A = (a_{ij})\) be an \(m \times n\)-matrix over \(D\), of rank \(r = \text{rank}_F(A)\), and let \(b = [b_1, \ldots, b_m]^\text{tr} \in D^m\). 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 \(c = [c_1, \ldots, c_m]^\text{tr} \in D^m\) with the following properties:

1. \(r = \text{rank}_F(B) = \text{rank}_{F(p)}(B \mod p)\), 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 \(\Delta\) 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, \(\Delta = (a_{ij})_{1 \leq i, j \leq r}\).

As above, consider now the system

\[
\begin{bmatrix}
\delta \\
\delta \\
\cdot \\
\delta \\
\end{bmatrix}
\begin{bmatrix}
c_{1,r+1} & \cdots & c_{1,n} \\
c_{2,r+1} & \cdots & c_{2,n} \\
\vdots & \ddots & \vdots \\
c_{r,r+1} & \cdots & c_{n,n} \\
\end{bmatrix}
\begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_n \\
\end{bmatrix}
= \begin{bmatrix}
d_1 \\
d_2 \\
\vdots \\
d_r \\
\end{bmatrix}
\]

which is obtained by multiplying both sides of \(Ay = b\) from the left with the adjoint of \(\Delta\). 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, \(v(c_{ij}) \geq v(\delta)\) for all \(i, j\), by choice of \(\Delta\). Therefore,
if \( v(d_i) < v(\delta) \) for some \( i \in \{1, \ldots, m\} \), 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 \( i = 1, \ldots, m \). Dividing all coefficients \( \delta, c_{ij} \) 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 \( Ay = b \mod p^e \), where \( t \geq e \), then clearly \( y \) is also a solution to \( By = c \mod p^{t-e} \). Hence:

**Lemma 5.3.** If \( Ay = b \) is solvable mod \( p^e \), 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 \( p^{t-e(A)} \). \( \square \)

### 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 \( D_{\text{alg}} = \mathbb{Z}_p \langle X \rangle_{\text{alg}} \), of rank \( r \) over \( F = \text{Frac}(D) \), and \( b = (b_1, \ldots, b_m)^t \) will denote a vector with \( b_1, \ldots, b_m \in D \). Note that the entries of the matrix \( B \) produced by \( p \)-desingularization lie in \( D_{\text{alg}} \), and if \( b \in (D_{\text{alg}})^m \), then \( c \in (D_{\text{alg}})^m \) as well. We now want to estimate the height of \( B \). For every \( r \times r \)-submatrix \( S \) of \( A \) we have

\[
h(\det S) + \log 2 \leq r(\log r + h_{\text{max}}(S)) + \log 2 \leq r^2(h_{\text{max}}(A) + \log 2). \tag{5.4}
\]

(By Lemma 2.28 and the crude estimate \( r \log r + \log 2 \leq r^2 \log 2 \).) Hence

\[
e(A) \leq r^2 \Delta(A)(h_{\text{max}}(A) + \log 2)/\log p \tag{5.5}
\]

by Lemma 4.23. Therefore, if \( Ay = b \) is solvable mod \( p^s \), with \( t \geq s \), where \( s \) is the integral part

\[
s = \lfloor \frac{m^2 \Delta(A)(h_{\text{max}}(A) + \log 2)}{\log p} \rfloor,
\]

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

\[
m^2(h_{\text{max}}(A) + \log 2) - \log 2, \tag{5.4}
\]

by (5.4). So the matrix \( B \) produced by \( p \)-desingularization satisfies:

\[
h_{\text{max}}(B) + \log 2 \leq e(A) \log p + m^2(h_{\text{max}}(A) + \log 2) \leq 2m^2 \Delta(A)(h_{\text{max}}(A) + \log 2).
\]

Similarly, if all entries \( b_1, \ldots, b_m \) of the vector \( b \) are algebraic over \( \mathbb{Q}(X) \), one gets

\[
h_{\text{max}}(B, c) + \log 2 \leq 2m^2 \Delta(A, b)(h_{\text{max}}(A, b) + \log 2).
\]

We summarize this discussion:

**Proposition 5.4.** If the system \( Ay = b \) is solvable mod \( p^s \), where

\[
t \geq s = \lfloor \frac{m^2 \Delta(A)(h_{\text{max}}(A) + \log 2)}{\log p} \rfloor,
\]

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

\[
h_{\text{max}}(B) \leq 2m^2 \Delta(A)(h_{\text{max}}(A) + \log 2) - \log 2
\]

and if the entries of \( b \) are from \( D_{\text{alg}} \), then so are the entries of \( c \), with

\[
h_{\text{max}}(B, c) \leq 2m^2 \Delta(A, b)(h_{\text{max}}(A, b) + \log 2) - \log 2.
\]

\( \square \)
5.3. Height and degree bounds in Hermann’s method. We now want to determine how $m', h_{\text{max}}(A')$ and $\Delta'(A')$ depend on $m, h_{\text{max}}(A)$ and $\Delta(A)$. In order to do this efficiently, we introduce the following ad-hoc notation:

**Notation.** We call the real number

$$c(A) := \Delta(A)(h_{\text{max}}(A) + 1)$$

the complexity of $A$.

Note that always $c(A) \geq 1$, and $c(S) \leq c(A)$ for every submatrix $S$ of $A$ (in particular $\Delta(a)(h(a) + 1) \leq c(A)$ for every entry $a$ of $A$). For given real number $\beta$ and integers $m, n > 0$, there exist only finitely many $m \times n$-matrices $A$ over $\mathbb{Z}_p(X)_{\text{alg}}$ with $c(A) \leq \beta$. We also abbreviate $m^2 c(A)$ by $c^*(A)$.

Rather than studying how $\Delta(A)$ and $h_{\text{max}}(A)$ 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 $s = \lceil c^*(A)/\log p \rceil$, and the matrix $B$ obtained in this way satisfies $h_{\text{max}}(B) \leq 2c^*(A) - \log 2$. If $b \in (D_{\text{alg}})^m$, then similarly $h_{\text{max}}(B, c) \leq 2c^*(A, b) - \log 2$.

5.3.2. Complexity of $T_d(B)$. We now want to bound $c\left(T_d(B)\right)$. Recall that the system $By = c$ obtained from $Ay = b$ by $p$-desingularization has the shape (5.3), where $\delta \equiv 0 \mod p$. By Lemma 4.26 we have

$$\deg_X \delta \mod p \leq \Delta(\delta)(h(\delta) + \log 2) \leq c(B) \leq 2\Delta(A)c^*(A).$$

(5.6)

Let $d = \lceil 2\Delta(A)c^*(A) \rceil + 1$ (so that $d > \deg_X \delta \mod p$ by (5.6)). Then $T_d(\delta)$ is $X_N$-distinguished of some degree $s$ with $s < d^N \leq \left(4\Delta(A)c^*(A)\right)^N$, and by Proposition 4.21,

$$h(T_d(f)) = O(1)^{N-1}\left(\Delta(A)c^*(A)\right)^{N-1}(h(f) + \log 2)$$

for all $f \in D_{\text{alg}}$.

We conclude that

$$c\left(T_d(B)\right) = O(1)^{N-1}\left(\Delta(A)c^*(A)\right)^N.$$  

(5.7)

Similarly, if $b \in (D_{\text{alg}})^m$, then applying $T_d$ to all coefficients of $By = c$ yields a system of linear equations over $D_{\text{alg}}$ whose extended coefficient matrix has complexity

$O(1)^{N-1}\left(\Delta(A, b)c^*(A, b)\right)^N$.

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 $A' \in (D_{\text{alg}}')^{m' \times n'}$ and $b' \in (D_{\text{alg}}')^{m'}$ as above. Let $f_1, \ldots, f_K \in D_{\text{alg}} = \mathbb{Z}_p(X)_{\text{alg}}$, and write each $f_k = \text{canonical image of } f_k$ in $D_{\text{alg}}/T_d(\delta)D_{\text{alg}}$ as

$$\overline{f}_k = f_{k0} + f_{k1}X_N + \cdots + f_{k,s-1}X_N^{s-1} \quad \text{with } f_{kl} \in D_{\text{alg}}'.$$

Then by Theorem 4.15 and the estimate (5.7):

$$\Delta'(f_{kl} : 1 \leq k \leq K, 0 \leq l < s) = O(1)^{(N-1)s}\Delta(f_1, \ldots, f_K)^s (\Delta(A)c^*(A))^N$$

and

$$h_{\text{max}}(f_{kl} : 1 \leq k \leq K, 0 \leq l < s) =$$

$$O(1)^{N+s-1}\left(\Delta(f_1, \ldots, f_K)^s(h_{\text{max}}(f_1, \ldots, f_K) + \log 2)\right) (\Delta(A)c^*(A))^N.$$
In addition, let $\Delta^t(C(k), \xi_{tk} : 0 \leq k < s \leq t \leq 2(s - 1)) = O(1)^{Ns} \Delta(A)^{(N+1)s} c^*(A)^{Ns}$ and
\[
h_{\max}(C(k) : 0 \leq k < s) = O(1)^{Ns} \Delta(A)^{2N+s} c^*(A)^{2N}, \tag{5.8}
\]
and since $h_{\max}(X_N^k, \ldots, X_N^{2(s-1)}) = 2(s - 1)$, we also have
\[
h_{\max}(\xi_{tk} : 0 \leq k < s \leq t \leq 2(s - 1)) = O(1)^{Ns} (\Delta(A) c^*(A))^N. \tag{5.9}
\]
Moreover, for $0 \leq k, l < s$ we get
\[
h_{\max} \left( C(k-l) + \sum_{t=s}^{2(s-1)} C(t-l) \xi_{tk} \right) \leq h_{\max}(C(k-l)) + \sum_{t=s}^{2(s-1)} h_{\max}(C(t-l)) + h(\xi_{tk}) + \log s.
\]
The right-hand side in this inequality may be bounded from above by
\[
s \cdot h_{\max}(C(0), \ldots, C(s-1)) + (s-1) \cdot h_{\max}(\xi_{sk}, \ldots, \xi_{2(s-1),k}) + \log s,
\]
so with (5.8), (5.9):
\[
h_{\max}(A') = O(1)^{Ns} \Delta(A)^{2N+s} c^*(A)^{2N}.
\]
This yields $c(A') = O(1)^{Ns} c^*(A)^{O(1)Ns}$, and therefore, since $m' \leq ms$ and $s = O(1)^N c^*(A)^N$:
\[
c^*(A') = O(1)^{Ns} c^*(A)^{O(1)Ns}.
\]
We can generously estimate
\[
c^*(A') \leq 2^{O(1)^N c^*(A)^{N+1}}.
\]
Now suppose that $b \in (D_{\text{alg}})^n$. Then similarly as above we obtain
\[
c(A', b') \leq 2^{O(1)^N c^*(A,b)^{N+1}}.
\]
In addition, let $y' = (y_{jk}) \in (D'_{\text{alg}})^n$ be a solution to the system $A'y' = b'$, and $y \in (D_{\text{alg}})^n$ the corresponding solution to the original system $Ay = b$. We have, for $j = r + 1, \ldots, n$,
\[
T_d(y_j) = y_{j0} + y_{j1} X_N + \cdots + y_{j,s-1} X_N^{s-1},
\]
and thus, by Remarks 4.20, (1)
\[
h(T_d(y_j)) \leq (s - 1) + s \cdot h_{\max}(y') + \log s.
\]
So we get
\[
h(y_j) = O(d^{N-1} (h(T_d(y_j)) + \log 2)) = O(1)^N c^*(A)^{3N-1} (h_{\max}(y') + 1).
\]
For $i = 1, \ldots, r$, 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 \( c^* (A, b) \), as is \( h(\delta) \). (See (5.4).) Hence

\[
h(y_i) \leq h(\delta) + \sum_{j=r+1}^{n} (h(c_{ij}) + h(y_j)) + h(d_i) + \log(n - r + 1)
\]

\[
\leq (n - r + 2) c^* (A, b) + (n - r) h_{\max}(y_{r+1}, \ldots, y_n) + \log(n - r + 1)
\]

\[
\leq O(1)^n n c^* (A, b)^{3N-1} (h_{\max}(y') + 1),
\]

and also \( \Delta(y_i) \leq \Delta(A, b) \Delta(y') \). We put everything together:

**Proposition 5.5.** Let \( A \in D_{\text{alg}}^{m \times n}, b \in D^m \), and assume that \( N > 0 \). Suppose the system \( Ay = b \) is solvable mod \( p^t \), where \( t \geq s = \lceil c^*(A) / \log p \rceil \). Let \( A'y' = b' \) with \( A' \in (D'_{\text{alg}})^{m' \times n'} \), \( b' \in (D')^{m'} \) be the system obtained from \( Ay = b \) by the procedure sketched above. Then \( A'y' = b' \) is solvable mod \( p^{t-s} \), and

\[
c^*(A') = 2 O(1)^N c^*(A)^{N+1}, \quad n' = O(1)^n n c^*(A)^N.
\]

*If in addition \( b \in (D_{\text{alg}})^m \), then*

\[
c^*(A', b') = 2 O(1)^N c^*(A, b)^{N+1}, \quad y' \in (D_{\text{alg}})^n \text{ is a solution to } A'y' = b' \text{ and } y \in (D_{\text{alg}})^n \text{ the corresponding solution to } Ay = b, \]

\[
\Delta(y) \leq \Delta(A, b) \Delta(y').
\]

\[\square\]

### 5.4. Towers of exponentials

Before we go on, let us introduce the following useful notation: for a real number \( a \), we define recursively

\[2 \uparrow 0 a := a, \quad 2 \uparrow^{n+1} a := 2^{2^{n} a}.
\]

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 \( (n, a) \mapsto 2 \uparrow^n a : \mathbb{N} \times \mathbb{N} \to \mathbb{N} \) is primitive recursive. (However, it is not elementary recursive in the sense of Kalmár, see [19].) Note that for \( a \leq b \in \mathbb{R} \) and \( n \leq m \), we have \( 2 \uparrow^n a \leq 2 \uparrow^m b \). We note a few other basic estimates:

**Lemma 5.6.** Let \( a, b \in \mathbb{R}, a, b \geq 1 \). Then

1. \((2 \uparrow n a) + b \leq 2 \uparrow^n (a + b)\).
2. \((2 \uparrow n a) \cdot b \leq 2 \uparrow^n (a + b), \text{ for } n \geq 1\).
3. \((2 \uparrow n a)^{b} \leq 2 \uparrow^n (a + b), \text{ for } n \geq 2\).

*Proof.* We prove (1) by induction on \( n \). The case \( n = 0 \) being trivial, suppose (1) holds for a certain \( n \). We have \( 1 + b \leq 2^b \), so that

\[(2 \uparrow^{n+1} a) + b = 2 \uparrow^n a + b \leq 2 \uparrow^n a (1 + b) \leq 2 \uparrow^n a 2^b = 2 \uparrow^{n+1} a + b,
\]

and by applying the inductive hypothesis, the claim follows. Property (2) is now a consequence of (1), since

\[(2 \uparrow^{n} a) \cdot b \leq 2 \uparrow^{n-1} a \cdot 2^b \leq 2 \uparrow^{n-1} a (a + b) = 2 \uparrow^n (a + b),
\]

for \( n \geq 1 \). Similarly, (3) follows from (2), since

\[(2 \uparrow^{n} a)^{b} = 2 \uparrow^{n} (a) a \cdot b \leq 2 \uparrow^{n-1} a (a + b) = 2 \uparrow^{n} (a + b)
\]

for all \( n \geq 2 \). \[\square\]
5.5. Proof of Theorem 5.1. Suppose the system $Ay = b$ has a solution in $\mathbb{Q}_p(X)$ (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

$$A^{(N)} y^{(N)} = b^{(N)}$$

(S$_N$)

$$\vdots$$

$$A^{(0)} y^{(0)} = b^{(0)},$$

(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 $\mathbb{Z}_p(X_1, \ldots, X_\nu)_{\text{alg}}$,
3. $y^{(\nu)} = [y_1^{(\nu)}, \ldots, y_{n(\nu)}^{(\nu)}]^T$ is a vector of unknowns, and
4. $b^{(\nu)} = [b_1^{(\nu)}, \ldots, b_{m(\nu)}^{(\nu)}]^T$ is a vector with coordinates in $\mathbb{Z}_p(X_1, \ldots, X_\nu)$.

The initial equation (S$_N$) is just $Ay = b$, and if $\nu > 0$, the equation (S$_{\nu-1}$) is obtained from (S$_\nu$) by the procedure described earlier in this section: Weierstraß Division by a suitable minor, after a preliminary $p$-desingularization. The matrices $A^{(N-1)}, \ldots, A^{(0)}$ 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 $b^{(\nu)} = b$ are algebraic over $\mathbb{Q}(X)$, then the entries of $b^{(\nu)}$ are algebraic over $\mathbb{Q}(X_1, \ldots, X_\nu)$, for $0 \leq \nu \leq N$, and if (S$_\nu$) is solvable in $\mathbb{Q}_p(X_1, \ldots, X_\nu)$, then (S$_{\nu-1}$) is solvable in $\mathbb{Q}_p(X_1, \ldots, X_{\nu-1})$. 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 $e_\nu$ the integral part $[c^*(A^{(\nu)})] / \log p]$. 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 $\nu = 0, \ldots, N$, we have

$$c^*(A^{(\nu)}) = 2 \uparrow^{N-\nu} O(1)^N c^*(A)^{N+1},$$

and if $b \in (D_{\text{alg}})^m$, then

$$c^*(A^{(\nu)}, b^{(\nu)}) = 2 \uparrow^{N-\nu} O(1)^N c^*(A, b)^{N+1}.$$ 

Proof. Let $C \geq 1$ be the universal constant such that $c^*(A') \leq 2^{C^N} c^*(A)^{N+1}$ from Proposition 5.5, and put $c^{**}(A) := C c^*(A)$. Since $Cx \leq x^C$ for all $x \in \mathbb{R}$, $x \geq 2$, we have the simple estimate $c^{**}(A') \leq 2 c^{**}(A)^{N+1}$. We now claim that

$$c^{**}(A^{(\nu)}) \leq 2 \uparrow^{N-\nu} ((\nu + 2) + \cdots + (N - 1) + N c^{**}(A)^{N+1})$$

for all $\nu = 0, \ldots, N$. This is clear for $\nu = N$. If $\nu = N - 1$, then

$$c^{**}(A^{(N-1)}) \leq 2 c^{**}(A)^{N+1} \leq 2^N c^{**}(A)^{N+1},$$

which completes the proof.
and if \( \nu = N - 2 \), then
\[
c^{**}(A^{(N-2)}) \leq 2c^{**}(A^{(N-1)})^N \leq 2^{2N}c^{**}(A)^{N+1}.
\]

Suppose we have proved the inequalities in question for some \( \nu \) with \( 0 < \nu \leq N - 2 \). Then by inductive hypothesis and Lemma 5.6, (3), we have
\[
c^{**}(A^{(\nu-1)}) \leq 2c^{**}(A^{(\nu)})^{\nu+1} \leq 2^{2^{N-\nu}}((\nu+2)\cdots+(N-1)+N)c^{**}(A)^{N+1})^{\nu+1} \leq 2^{N-\nu+1}((\nu+1)+(\nu+2)\cdots+(N-1)+N)c^{**}(A)^{N+1}).
\]
This shows the claim, which in turn easily yields the desired bound on \( c^*(A^{(\nu)}) \). For \( c^*(A^{(\nu)}) \), in case \( b \in (D_{\text{alg}})^m \), one argues similarly.

Proceeding in this way we ultimately arrive at the last equation \((S_0)\) over \((\mathbb{Z}_p)_{\text{alg}}\), which will be solvable mod \( p^{E-(\epsilon_1+\cdots+\epsilon_{N})} \). 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 \( \mathbb{Q} \), \( b = [b_1, \ldots, b_m]^t \in \mathbb{Z}_p^m \), and suppose \( \text{rank}_{\mathbb{Q}_p}(A) = \text{rank}_{\mathbb{Q}_p}(A, b) \), that is, the system \( Ay = b \) is solvable in \( \mathbb{Q}_p \).

**Proposition 5.8.** Let \( e_0 := [c^*(A)/\log p] \). 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 \( b_1, \ldots, b_m \in (\mathbb{Z}_p)_{\text{alg}} \), then the system \( Ay = b \) has a solution \( y = [y_1, \ldots, y_n]^t \) with \( y_1, \ldots, y_n \in (\mathbb{Z}_p)_{\text{alg}} \) and
\[
\Delta(y) \leq \Delta(A, b), \quad h_{\max}(y) \leq 4c^*(A, b).
\]

**Proof.** We may assume \( m = r = \text{rank}_{\mathbb{Q}_p}(A) \). 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 \( \delta \) is a non-zero \( r \times r \)-minor of \( A \). Now \( v_p(\delta) \leq c(A) \leq e_0 \), so \( \mathbb{Z}_p/\delta \mathbb{Z}_p = \mathbb{Z}/p^\varepsilon \mathbb{Z} \) for some \( \varepsilon \leq e_0 \). 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
\[
\begin{bmatrix}
\delta \\
\vdots \\
\delta 
\end{bmatrix}
\begin{bmatrix}
c_{1,r+1} & \cdots & c_{1,n} \\
\vdots & \ddots & \vdots \\
c_{r,r+1} & \cdots & c_{r,n} 
\end{bmatrix}
\begin{bmatrix}
y_1 \\
y_2 \\
y_n 
\end{bmatrix}
= 
\begin{bmatrix}
d_1 \\
d_2 \\
d_r 
\end{bmatrix}
\]
with the same solutions in \( \mathbb{Z}_p \), where \( c_{ij}, d_i \in (\mathbb{Z}_p)_{\text{alg}} \) for \( 1 \leq i \leq r < j \leq n \) and \( \delta \in (\mathbb{Z}_p)_{\text{alg}} \) is a unit, whose heights are bounded from above by \( 2c^*(A, b) \). This system has a solution \( y = [y_1, \ldots, y_n]^t \) given by \( y_j = d_j/\delta \) for \( 1 \leq j \leq r \) and \( y_j = 0 \) for \( r < j \leq n \), and we have
\[
h(y_j) \leq h(d_j) + h(\delta) \leq 4c^*(A, b)
\]
for \( 1 \leq j \leq r \), as required.

We return to the general case, i.e., \( N \geq 0 \). 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 \( e_0 = [c^*(A^{(0)})/\log p] \). So it suffices to take the exponent \( E \) such that
\[
E \geq \sum_{\nu=0}^{N} \left[ c^*(A^{(\nu)})/\log p \right].
\]
By Lemma 5.7, we have
\[ c^*(A^{(\nu)}) = 2^{N-\nu}O(1)^N c^*(A)^{N+1} \quad \text{for all } \nu = 0, \ldots, N, \]
and hence we see that for some universal constant \( C > 0, \)
\[ E(N, d, h, m) := 2^N \left[ C^N (m^2 d (h + 1))^{N+1} \right] \]
has the properties required in Theorem 5.1. Suppose now that \( b \in (D_{\text{alg}})^m. \) Then
\[ c^*(A^{(\nu)}, b^{(\nu)}) = 2^{N-\nu} O(1)^N c^*(A, b)^{N+1} \quad \text{for all } \nu = 0, \ldots, N. \]

Let \( y^{(0)} \in (\mathbb{Z}_p)^n \) be a solution to (S0) with
\[ \Delta(y^{(0)}) \leq \Delta(A^{(0)}, b^{(0)}), \quad h_{\text{max}}(y^{(0)}) \leq 4 c^*(A^{(0)}, b^{(0)}) \]
as in Proposition 5.8, and \( y^{(1)}, \ldots, y^{(N)} = y \) the corresponding solutions to the systems (S1), \ldots, (S_N), respectively. Using Proposition 5.5, we obtain:
\[ \Delta(y) = 2^N O(1)^N c^*(A, b)^{N+1}. \] (5.10)

We leave it to the reader to deduce a similar bound on \( h_{\text{max}}(y) \) (involving \( N, m, n, \Delta(A) \) and \( h_{\text{max}}(A) \)). 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 \((N, \beta) \mapsto \epsilon(N, \beta)\) 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 \( \text{deg}_{(X)} A \leq d, \) \( \log \|A\|_\infty \leq h, \) where \( d \in \mathbb{N}, h \in \mathbb{R}, d, h > 0. \) There exist positive integers \( \delta \) and \( E_1, E_2 \) with
\[ E_1 = 2^N O(1)^N (m^2 (2d + h + 1))^{N+1}, \]
\[ E_2 = (2md)^{O(N \log(N+1))} (h + 1)^{2N+1}, \]

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 \( \delta^{E_1}; \)
2. \( \) if \( \text{deg} b \leq d \) and \( \|b\|_\infty \leq h, \) 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 \( \delta^{E_2} \) of degree at most \( (2md)^{O(N \log(N+1))} (h + 1). \)

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 \( b = [b_1, \ldots, b_m]^\text{tr} \) with \( b_1, \ldots, b_m \in R[X], \) and suppose the \( a_{ij} \) and the \( b_i \) all have total degree \( \leq d. \) Let \( \lambda, \mu, \nu \) 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 \( y = [y_1, \ldots, y_n]^\text{tr}, \) where
\[ y_j = \sum_{|\nu| \leq \gamma} y_{j, \nu} X^\nu \]
with new indeterminates $y_{j,\nu}$ ranging over $R$. A polynomial in $X_1, \ldots, X_N$ of degree at most $d$ has at most $M(N, d) \leq \binom{N+d}{N}$ monomials. Hence for every $R$-algebra $S$, the solutions in $S[X]$ of every equation

$$a_{i1}y_1 + \cdots + a_{in}y_n = b_i \quad (1 \leq i \leq m)$$

such that $\deg y_j \leq \gamma$ for all $j = 1, \ldots, n$ are in one-to-one correspondence with the solutions in $S$ of the system

$$\sum_{\mu + \nu = \lambda} b_{i,\lambda} = 0 \quad (|\lambda| \leq \gamma + d)$$

of $M(N, \gamma + d)$ equations in the $n \cdot M(N, \gamma)$ many variables $y_{j,\nu}$, with coefficients $a_{i,j,\mu}$, $b_{i,\lambda}$ in $R$. So the entire system $A\gamma = b$ of degree at most $\gamma$, uniformly for all $R$-algebras. Note that if $R$ is a subring of $\mathbb{Q}$, then $||A\gamma||_\infty = ||A||_\infty$ and $||b\gamma||_\infty = ||b||_\infty$. We use this discussion to show:

**Lemma 6.2.** Let $d, h \in \mathbb{N}$. There exist positive integers $\gamma_0, \gamma_1$ with

$$\gamma_0 \leq (2md)^{O(N \log(N+1))} (h+1)^{2N+1}, \quad \gamma_1 \leq (2md)^{O(N \log(N+1))} (h+1)$$

and having the following properties: Let $A = (a_{ij})$ be a non-zero $m \times n$-matrix and $b = [b_1, \ldots, b_m]^t$ be a vector with entries $a_{ij}, b_i \in \mathbb{Z}[X]$ of degree at most $d$ and

$$\log ||a_{ij}||_\infty, \log ||b_i||_\infty \leq h.$$  

Then for every prime number $p$, the system $A\gamma = b$ has a solution in $\mathbb{Z}_p[X]$ if and only if $A\gamma = b$ has a solution modulo $p^{\gamma_0}$ of degree at most $\gamma_1$.

**Proof.** By Theorem 8.6 in [8], the system $A\gamma = b$ has a solution in $\mathbb{Z}_p[X]$ if and only if it has such a solution of degree at most $\gamma_1$. As we’ve seen above, there exists a certain system $A\gamma^* = b^*$ 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 $A\gamma = b$ in $\mathbb{Z}_p[X]$ of degree $\leq \gamma_1$ are in one-to-one correspondence with the solutions of $A\gamma^* = b^*$ in $\mathbb{Z}_p$, and for any $e \geq 1$, the solutions to $A\gamma = b$ in $(\mathbb{Z}/p^e\mathbb{Z})[X]$ of degree $\leq \gamma_1$ are in one-to-one correspondence with the solutions of $A\gamma^* = b^*$ in $\mathbb{Z}/p^e\mathbb{Z}$. By Lemma 5.8, the system $A\gamma^* = b^*$ has a solution in $\mathbb{Z}_p[X]$ provided it has a solution modulo $p^e$ where

$$e = \left[ \left( m \cdot M(N, \gamma_1 + d) \right)^2 \left( h_{\max}(A\gamma^*) + 1 \right) \right].$$

Now $h_{\max}(A\gamma^*) = \log ||A, b||_\infty \leq h$ and for $d > 0$

$$m^2M(N, \gamma_1 + d)^2 \leq m^2(\gamma_1 + d + 1)^2N \leq (2md)^{O(N \log(N+1))} (h+1)^{2N},$$

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 $\delta$ such that

$$MQ[X] \cap \mathbb{Z}[X]^m = (M : \delta) := \{ y \in \mathbb{Z}[X]^m : \delta y \in M \}.$$
If $p_1, \ldots, p_K$ are the distinct prime factors of a number $\delta$ with this property, then
\[
M = M\mathbb{Q}[X] \cap M\mathbb{Z}_{(p_1)}[X] \cap \cdots \cap M\mathbb{Z}_{(p_K)}[X] = M\mathbb{Q}[X] \cap M\mathbb{Z}_{p_1}(X) \cap \cdots \cap M\mathbb{Z}_{p_K}(X).
\]

(6.1)

Proof. The existence of the integer $\delta$ 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 $\delta$ and $p_1, \ldots, p_K$ 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 $\mathbb{Z}_{p_k}(X)$ for all $k = 1, \ldots, K$. For given positive integer $E$, by the Chinese Remainder Theorem we have a surjection
\[
(\mathbb{Z}/\delta E\mathbb{Z})[X] \to (\mathbb{Z}/p_1^E\mathbb{Z})[X] \times \cdots \times (\mathbb{Z}/p_k^E\mathbb{Z})[X],
\]

for an integer $E$ such that $\delta \equiv 0 \pmod{p_i^E}$ for all $i$. Thus, for $a \bmod{\delta E} \mapsto (a \bmod{p_1^E}, \ldots, a \bmod{p_k^E})$.

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 $D$ 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 $b = (b_1, \ldots, b_m)^T \in D^m$.

We want to consider the following two problems:

$H_e(A)$: Find a finite set of generators for the submodule

$S_e(A) := \{y \in D^n : Ay = 0 \bmod{p^e}\}$

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 $b \in D^n$, with $m, n \geq 1$. Suppose $e > 1$, and let
\[
g_1 = \begin{bmatrix} g_{11} \\ \vdots \\ g_{n1} \end{bmatrix}, \ldots, g_r = \begin{bmatrix} g_{1r} \\ \vdots \\ g_{nr} \end{bmatrix}
\]

be vectors in $D^n$ generating the submodule $S_1(A)$ 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 $H_{e-1}(A^{(1)})$, that is, we can find
\[
h_1 = \begin{bmatrix} h_{11} \\ \vdots \\ h_{r1} \end{bmatrix}, \ldots, h_s = \begin{bmatrix} h_{1s} \\ \vdots \\ h_{rs} \end{bmatrix} \in D^r
\]

generating $S_{e-1}(A^{(1)})$. 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 $S_e(A)$. 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 $A \equiv b - pb^{(1)}$. Let $G = (g_{ij}) \in D^{n \times r}$ as above be a matrix with entries from $D$ whose column vectors generate $S_1(A)$, 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

$$A^{(1)}y^{(1)} = b^{(1)} \mod p^{e-1} \iff A(Gy^{(1)} + z) = b \mod p^e.$$

So if $A^{(1)}y^{(1)} = b^{(1)} \mod p^{e-1}$, then $y = Gy^{(1)} + z$ solves $Ay = b \mod p^e$. Conversely, if $y \in D^n$ is such that $Ay = b \mod p^e$, then $Ay = b \mod p$, so $y - z \in S_1(A)$ and thus $y = z + Gy^{(1)}$ for some $y^{(1)} \in D^{r}$; so $A^{(1)}y^{(1)} = b^{(1)} \mod p^{e-1}$. \hfill \Box

Using the inductive hypothesis for $I_{e-1}(A^{(1)}, b^{(1)})$, 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 $\leq (2md)^2N$. Then $A^{(1)}$ has degree $\leq d + (2md)^2N$, and if the $h_1, \ldots, h_s$ are of degree $\leq d'$, then $GH$ is of degree $\leq d' + (2md)^2N$. Let $\gamma_e(N, d, m)$ be the smallest natural number such that $S_e(A)$ for $A \in D^{m \times n}$ of degree $\leq d$ is generated by elements of degree $\leq \gamma_e(N, d, m)$. We have the recursive relation

$$\gamma_e(N, d, m) \leq \gamma_{e-1}(N, d + (2md)^2N, m) + (2md)^2N \quad \text{for } e > 1.$$

Using that $\gamma_{e-1}$ is increasing in the second variable and that $d + (2md)^2N \leq (2md)^2N + 1$ for $d > 0$, we see that for $d, e > 0$,

$$\gamma_e(N, d, m) \leq (2md)^2N + (2md)^2N(2N^2 + 2) + \cdots + (2md)^2N(2N^2 + 2)^{e-1}.$$

Since $(2md)^2N(2N^2 + 2) \leq (2md)^2N(2N^2 + 2)^{e-1}$ for $i = 0, \ldots, e - 1$, we get for $N, d, e > 0$,

$$\gamma_e(N, d, m) \leq e(2md)^2N(2N^2 + 2)^{e-1} \leq e(2md)^2N(2N^2 + 2)^{e-1} \leq e(2md)^2N(2N^2 + 2)^{e-1}.$$

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 $\leq d$, and the system of congruence equations $Ay = b \mod p^e$ 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 $\gamma_e$ 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 $\leq d$, and $e \geq 1$.

1. There exist $y^{(1)}, \ldots, y^{(r)} \in D^n$ of degree $\leq e(2md)^2N(2N^2 + 2)$ such that

   $$Ay^{(k)} = 0 \mod p^e \quad \text{for } k = 1, \ldots, r,$$

   and every $y \in D^n$ with $Ay = 0 \mod p^e$ is a $D$-linear combination of the $y^{(k)}$'s.

2. If there exists $y \in D^n$ such that $Ay = b \mod p^e$, then there exists such a $y$ of degree $\leq e(2md)^2N(2N^2 + 2)$.
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 \( \delta \) 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 \( \leq e(\delta)(2md)^{e(\delta)(N+1)} \) such that
   \[
   Ay^{(k)} \equiv 0 \mod \delta \quad \text{for } k = 1, \ldots, r,
   \]
   and every \( y \in D^n \) with \( Ay = 0 \mod \delta \) is a \( D \)-linear combination of the vectors \( y^{(1)}, \ldots, y^{(r)} \).

2. If there exists \( y \in D^n \) such that \( Ay = b \mod \delta \), then there exists such a \( y \) of degree \( \leq e(\delta)(2md)^{e(\delta)(N+1)} \).

Taking \( R = \mathbb{Z} \) in this proposition and noting that \( e(\delta) \leq \log_2 |\delta| \) yields Proposition 0.4.

**REFERENCES**


Email 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.