Symmetric Diophantine approximation over function fields

Weidong Zhuang

Symmetric Diophantine approximation over function fields

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 3 december 2015
klokke 10:00 uur

door

Weidong ZHUANG
geboren te Jiangsu, China
in 1983

Samenstelling van de promotiecommissie:

Promotor: Prof. dr. P. Stevenhagen

Copromotor: Dr. J.-H. Evertse

Overige leden:

Prof. dr. F. Beukers (Universiteit Utrecht)
Prof. dr. Y. Bugeaud (Université Strasbourg)
Prof. dr. K. Győry (University of Debrecen)
Dr. R. de Jong
Prof. dr. B. de Smit (secretaris)
Prof. dr. A.W. van der Vaart (voorzitter)
Mw. prof. dr. J. T.-Y. Wang (Academia Sinica Taiwan)

This work was funded by the NWO vrije competitie EW 2011 project "Symmetric Diophantine Approximation" and was carried out at Universiteit Leiden.

To Yana and Zhida

Contents

Introduction

Roth's theorem gives an optimal solution to the problem how well a given algebraic number can be approximated by other algebraic numbers. A natural question is to ask how well two varying algebraic numbers can approximate each other. There is only one non-trivial result, proved by Evertse, but this is far from optimal. Its proof is based on a weak version of the abc-conjecture, which is a consequence of a generalization of Roth's Theorem, hence it is non-effective.

Let k be an algebraically closed field of characteristic 0. Over algebraic function fields of transcendence degree 1 over k there is a proved analogue of the abc-conjecture, i.e., the Mason-Stothers Theorem. This suggests that it should be possible to develop much stronger symmetric Diophantine approximation results over function fields. My research focuses mainly on this interesting problem.

To tackle this problem, one considers two cases: either the two algebraic functions that approximate each other are conjugate over the field of rational functions $k(t)$ or not.

The first case is strongly connected to the following problem: over the integers, two binary forms (i.e., homogeneous polynomials) $F, G \in \mathbb{Z}[X, Y]$ are called equivalent if $G(X, Y) = F(aX + bY, cX + dY)$ for some matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}(2, \mathbb{Z})$. Two equivalent binary forms have the same discriminant. A binary form F is called *reduced* if its height $H(F)$ (maximum of the absolute values of its coefficients) is minimal among the heights of the binary forms in its equivalence class.

Conjecture. *The height $H(F)$ of a reduced binary form F of degree $n \geqslant 4$ and non-zero discriminant D has an upper bound of the form $c_1(n)|D|^{c_2(n)}$, where $c_1(n)$, $c_2(n)$ are numbers depending only on n.*

An analogous estimate for $n = 2$ and $n = 3$ follows from work of Lagrange, Gauss and Hermite. However, the general case is still open. There is only the following much weaker effective result from [11]:

Theorem (Evertse, Győry)**.** *Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be a reduced binary form of degree $n \geqslant 2$ and discriminant $D(F) \neq 0$. Then*

$$H(F) \leqslant \exp((c_1 n)^{c_2 n^4} |D|^{8n^3}),$$

where c_1, c_2 are effectively computable, absolute constants.

More generally, we may consider the ring of integers of an algebraic number field and even the ring of S-integers instead of $\mathbb{Z}$. A weak version of Evertse [9] implies the following:

Theorem (Evertse)**.** *Let $F \in \mathbb{Z}[X,Y]$ be a reduced binary form of degree $n > 1$ with splitting field L over $\mathbb{Q}$ and non-zero discriminant. Then*

$$H(F) \leqslant C^{\mathit{ineff}}(n, L)|D(F)|^{\frac{21}{n-1}}.$$

The constant here depends on n, L and is ineffective in the sense that it is not effectively computable from the method of proof. We call this result a 'semi-effective' upper bound since it is effective in terms of $D(F)$, but ineffective in terms of n and L.

We proved an analogue of the above conjecture over $k[t]$. Our main tools are an analogue of the geometry of numbers over function fields (see Thunder [24]) and Mason's theorem which is an analogue of the abc-conjecture over function fields.

We start with some notation.

Fix $K = k(t)$ where k is an algebraically closed field of characteristic 0 and t is transcendental over k. For $x \in k[t]$, define $|x|_\infty = e^{\deg(x)}$. For $f \in k[t]\backslash\{0\}$, define $\nu_p(f)(p \in k)$ by $f = (t-p)^{\nu_p(f)} g$ where $g \in k[t]$ and $g(p) \neq 0$. We extend this to $k(t)$ by setting $\nu_p(0) := \infty$ and $\nu_p(\frac{f}{g}) = \nu_p(f) - \nu_p(g)$ for $f, g \in k[t], g \neq 0$. Define $|x|_\nu = e^{-\nu(x)}$ for $x \in K$. For a polynomial F with coefficients $a_0, \ldots, a_n$ in $k[t]$, define $H(F) := \max(|a_0|_\infty, \ldots, |a_n|_\infty)$. If a binary form F has a factorization $F(X,Y) = \prod_{i=1}^{m} (\alpha_i X + \beta_i Y)$ over $\overline{K}$, define its discriminant by $D(F) = \prod_{i<j} (\alpha_i \beta_j - \alpha_j \beta_i)^2$. For two binary forms

$$F(X,Y) = \prod_{i=1}^{m} (\alpha_i X + \beta_i Y),\ G(X,Y) = \prod_{j=1}^{n} (\gamma_j X + \delta_j Y),$$

we define their resultant by

$$R(F,G) = \prod_{i=1}^{m} \prod_{j=1}^{n} (\alpha_i \delta_j - \beta_i \gamma_j).$$

Let L be a finite extension of $K = k(t)$. We say an absolute value on $|\cdot|_\omega$ on L is an extension of $|\cdot|_\nu$ on K if $|x|_\omega = |x|_\nu^{[L_\omega : K_\nu]}$ for every $x \in K$. Here L_ω, K_ν are the completions of L, K at ω, ν respectively. Define

$$H^*(x_1, \ldots, x_n) = \Big(\prod_{\omega \in M_L} \max(1, |x_1|_\omega, \ldots, |x_n|_\omega) \Big)^{1/[L:K]} \text{ for } (x_1, \ldots, x_n) \in L^n,$$

and

$$H(F) = \max(|a_0|_\infty, \ldots, |a_m|_\infty) \text{ for } F = \sum_{i=0}^{m} a_i X^{m-i} Y^i \in k[t][X,Y].$$

For a ring R, we say that two binary forms $F, G \in R[X,Y]$ are $\mathrm{GL}\,(2,R)$-equivalent if there exists $u \in R^\times$ and $U = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in GL(2,R)$ such that $G = uF_U$, where $F_U(X,Y) = F(aX + bY, cX + dY)$. Later we will apply this definition to a polynomial ring $k[t]$ or a function field L.

We recall Mason's ABC-theorem for function fields.

Theorem (Mason)**.** *Let L be a finite extension of $K = k(t)$, g_L the genus of L and T a finite set of valuations of L. Let $\gamma_1, \gamma_2, \gamma_3$ be non-zero elements of L satisfying $\gamma_1 + \gamma_2 + \gamma_3 = 0$ and $\nu(\gamma_1) = \nu(\gamma_2) = \nu(\gamma_3)$ for every valuation $\nu \notin T$. Then either $\frac{\gamma_1}{\gamma_2} \in k$, which means $H^*(\frac{\gamma_1}{\gamma_2}) = 1$, or $H^*(\frac{\gamma_1}{\gamma_2}) \leqslant e^{(\#T+2g_L-2)/[L:K]}$.*

As a consequence we derived a non-trivial result, Theorem 5, on how well two algebraic functions that are conjugate over $k(t)$ can approximate each other. We will come back to this with more details in the next few pages.

To study how well two algebraic functions non-conjugate over $k(t)$ can approximate each other involves a study of two binary forms, and requires one to find a non-trivial lower bound for the resultant of two binary forms in terms of their heights. To obtain such a bound, we developed a generalization of Mason's theorem to more variables, based on work of Brownawell and Masser [6], J.T.-Y. Wang [25] and Zannier [26].

This dissertation is organized as follows.

Chapter 1 introduces some very standard notation and collects some results related to discriminants, resultants, valuations, heights and twisted heights.

In Chapter 2, we introduce Mason's ABC-theorem for function fields and give a generalization, which is a solid basis to build our effective results on.

In Chapter 3 we develop some geometry of numbers over the rational function field $k(t)$. The main result concerns the successive minima of a so-called S-convex symmetric body.

With the help of the results in Chapter 3, we develop in Chapter 4 a reduction theory for binary forms over the rational function field.

In Chapter 5, we first derive some consequences of the Riemann-Hurwitz formula, and by combining these with the results from Chapter 1 to 4 we prove the following effective result, which is analogous to the conjecture mentioned above. The only earlier work in this direction is due to Gaál [13]. His results are formulated differently, but they imply a similar result, with a larger upper bound in terms of $|D(F)|_\infty$ for binary forms F with $F(1,0)=1$.

Theorem 1. *Let $F \in k[t][X,Y]$ be a binary form of degree $n \geqslant 4$ with non-zero discriminant. Then F is* $\mathrm{GL}\,(2,k[t])$*-equivalent to a binary form F^* such that*

$$H(F^*) \leqslant e^{(n^2+5n-6)} |D(F)|_\infty^{20+\frac{1}{n}}.$$

In Chapter 5, we in fact deduce a general version of Theorem 1, which deals with binary forms over localizations of $k[t]$ away from a finite set of elements of k.

In Chapter 6, we focus on the finiteness of the number of equivalence classes of binary forms of given discriminant and show the following

Theorem 2. *Given $n \in \mathbb{Z}, n \geqslant 4$, non-zero $\delta \in k[t]$ and a finite extension L of K, there are only finitely many* $\mathrm{GL}\,(2,K)$*-equivalence classes of binary forms satisfying*

$$\begin{cases} F \in k[t][X,Y], D(F) \in \delta k^\times, \\ F \text{ has splitting field } L \text{ over } K, \\ \deg F = n, \\ F \text{ is not } \mathrm{GL}\,(2,L)\text{-equivalent to a binary form in } k[X,Y]. \end{cases}$$

Remark. *Theorem 2 becomes false if the last condition is replaced by F not being* $\mathrm{GL}\,(2,K)$*-equivalent to a binary form in $k[X,Y]$. A counterexample is given in Chapter 6.*

In Chapter 7, we effectively estimate the resultant of two binary forms from below in terms of their discriminants and heights. This is based on ideas of Evertse and Győry for number fields. They deduced the following:

Theorem (Evertse, Győry [12]). *Let $F \in \mathbb{Z}[X,Y]$ be a binary form of degree $m \geqslant 3$ and $G \in \mathbb{Z}[X,Y]$ a binary form of degree $n \geqslant 3$ such that FG has splitting field L over $\mathbb{Q}$ and FG is square-free. Then*

$$|R(F,G)| \geqslant C^{ineff}(m,n,L)\big(|D(F)|^{n/(m-1)}|D(G)|^{m/(n-1)}\big)^{1/18}.$$

Theorem (Evertse [10]). *Let $m,n \geqslant 3$ and let (F,G) be a pair of binary forms with coefficients in $\mathbb{Z}$ such that $\deg F = m, \deg G = n$, FG is square-free and FG has splitting field L over $\mathbb{Q}$. Then there is an $U \in \mathrm{GL}(2,\mathbb{Z})$ such that*

$$|R(F,G)| \geqslant C^{\mathrm{ineff}}(m,n,L)\big(H(F_U)^n H(G_U)^m\big)^{1/718}.$$

The ineffectivity mainly comes from Schmidt's subspace theorem from Diophantine approximation. We apply a generalization of Mason's theorem (see Chapter 2) to obtain effective results as follows.

Theorem 3. *Assume $F,G \in k[t][X,Y]$ are two binary forms such that $\deg F = m \geqslant 3, \deg G = n \geqslant 3$, FG is square-free and splits in $k(t)$. Then*

$$|R(F,G)|_\infty \geqslant |D(F)|_\infty^{\frac{n}{17(m-1)}}|D(G)|_\infty^{\frac{m}{17(n-1)}}.$$

As a consequence of Theorem 1 and Theorem 3, we also show that

Theorem 4. *Let $m,n > 2$ and let F,G be binary forms in $k[t][X,Y]$ such that FG is square-free and splits in $k(t)$. Then there exists $U \in \mathrm{GL}(2,k[t])$ such that*

$$|R(F,G)|_\infty \geqslant c_1(m,n)^{-1}H(G_U)^{\frac{m}{717}}H(F_U)^{\frac{n}{717}},$$

where

$$c_1(m,n) = \exp\Big(-\frac{mn(4m+4n+11)}{717}\Big).$$

We actually prove a more general result where FG splits over a given arbitrary finite extension L of $k(t)$.

As an application, in Chapter 8 we prove a root separation result and a symmetric improvement of a Liouville-type inequality.

A result of Mahler states that for a polynomial $f(X) = a(X-\gamma_1)\dots(X-\gamma_n)$ with complex coefficients we have

$$\min_{1\leqslant i<j\leqslant n} |\gamma_i - \gamma_j| \geqslant (n+1)^{-n-1}\frac{|D(f)|^{1/2}}{H(f)^{n-1}}.$$

In case that f has integer coefficients and non-zero discriminant this implies that

$$\min_{1\leqslant i<j\leqslant n} |\gamma_i - \gamma_j| \geqslant (n+1)^{-n-1}H(f)^{1-n}. \qquad (*)$$

This inequality is proved by an elementary argument, similar to Liouville's inequality from Diophantine approximation on the approximation of algebraic numbers by rationals. Therefore, we call $(*)$ a Liouville-type inequality.

The root separation problem is to prove a similar inequality with instead of $1-n$ a larger exponent on $H(f)$. But this is still open. The only known case is, rather surprisingly, that when $n=3$ the exponent $1-n$ is best possible. The latest result [7] of Y. Bugeaud and A. Dujella shows that for $n \geqslant 4$ the exponent cannot be bigger than $-\frac{2n-1}{3}$.

We obtain an improvement of the exponent over the rational function field as follows.

Theorem 5. *Let $K = k(t)$ and $f \in K[X]$ be a polynomial of degree $n \geqslant 4$ with splitting field L. Write $f = a\prod_{i=1}^{n}(X-\gamma_i)$ with $a \in K^*$ and $\gamma_i \in L$. Fix an extension of $|\cdot|_\infty$ to L and denote this also by $|\cdot|_\infty$. Define*

$$\Delta_\infty(f) := \min_{1\leqslant i<j\leqslant n} \frac{|\gamma_i-\gamma_j|_\infty}{\max(1,|\gamma_i|_\infty)\max(1,|\gamma_j|_\infty)}.$$

Then

$$\Delta_\infty(f) \geqslant c_3(n)^{-1}H(f)^{-n+1+\frac{n}{40n+2}},$$

where

$$c_3(n) = \exp(\frac{(n-1)(n+6)}{20+1/n}).$$

We return to number fields. If we consider two algebraic numbers α, β not conjugate to each other, the problem becomes more general. A typical result is the following generalization of (∗): for T a finite set of valuations of $K(\alpha, \beta)$, we have

$$\left(\prod_{\omega \in T} |\alpha - \beta|_\omega\right)^{1/[L:K]} \geqslant \frac{1}{2} H^*(\alpha)^{-1} H^*(\beta)^{-1},$$

where $|\cdot|_\omega := |\cdot|_p^{[L_\omega:\mathbb{Q}_p]}$ if ω lies above $p \in \{\infty\} \cup \{\text{primes}\}$. The exponents of $H^*(\alpha)$ and $H^*(\beta)$ can be improved. A generalization of Roth's theorem by S. Lang implies that there is a constant $C > 0$ depending on α and $K(\beta)$ such that

$$\left(\prod_{\omega \in T} |\alpha - \beta|_\omega\right)^{1/[L:K]} \geqslant C H^*(\beta)^{-(2/r)-\delta},$$

where $r = [K(\alpha, \beta) : K(\beta)] \geqslant 3$.

On the other hand, if we allow both α and β to vary, the problem gets more difficult. Evertse obtained the following improvement of Liouville-type inequality.

Theorem (Evertse)**.** *Let K be an algebraic number field and α, β distinct numbers algebraic over K. Let $L = K(\alpha, \beta)$. Suppose that*

$$[L : K] = [K(\alpha) : K][K(\beta) : K],\ [K(\alpha) : K] \geqslant 3,\ [K(\beta) : K] \geqslant 3.$$

Let T be a finite set of valuations of L above $\nu \in M_K$ such that

$$\varpi := \frac{1}{[L : K]} \sum_{\omega \in T} [L_\omega : K_\nu] < \frac{1}{3}.$$

Then

$$\prod_{\omega \in T} \frac{|\alpha - \beta|_\omega}{\max(1, |\alpha|_\omega) \max(1, |\beta|_\omega)} \geqslant C^{\mathit{ineff}}(L, T)\big(H^*(\alpha) H^*(\beta)\big)^{-1+\delta},$$

where $\delta = \frac{1-3\varpi}{718(1+3\varpi)}$.

Following the same idea, we give an analogous improvement of Liouville-type inequality over the rational function field, which is effective.

Let $K = k(t)$ and ξ, η be distinct and algebraic over K. Let $L = K(\xi, \eta)$ and T a finite set of valuations on L. Define

$$\Delta_T(\xi,\eta) := \left(\prod_{\omega \in T} \frac{|\xi - \eta|_\omega}{\max(1, |\xi|_\omega) \max(1, |\eta|_\omega)} \right)^{1/[L:K]}.$$

Then we have the following Liouville-type inequality

$$\Delta_T(\xi,\eta) \geqslant H^*(\xi)^{-1} H^*(\eta)^{-1}.$$

and the following effective improvement

Theorem 6. *Suppose* ξ, η *are algebraic over* $K = k(t)$ *with* $[K(\xi) : K] \geqslant 3$ *and* $[K(\eta) : K] \geqslant 3$. *Let* $L = K(\xi, \eta)$ *and assume*

$$[L : K] = [K(\xi) : K][K(\eta) : K].$$

Suppose that

$$\varpi := \frac{1}{[L:K]} \sum_{\substack{\omega|\infty \\ \omega \in T}} [L_\omega : K_\nu] < \frac{1}{3}.$$

Let g_1, g_2 *be the genera of* $K(\xi)$ *and* $K(\eta)$ *respectively. Then*

$$\Delta_T(\xi,\eta) \geqslant c_4(m,n,g_1,g_2,\varpi)^{-1} \big(H^*(\xi) H^*(\eta) \big)^{-1+\vartheta},$$

where $\vartheta = \frac{1-3\varpi}{717(1+3\varpi)}$ *and*

$$c_4(m,n,g_1,g_2,\varpi) = \exp\left(\frac{426m+426n-1677+844g_1+844g_2}{717} + (m+n)(m+n-5)(1-\vartheta) \right).$$

Last but not least, we remark that in this dissertation we prove more general versions of Theorem 3, 4, 5, 6 with multiple valuations, whilst Theorem 3 holds in a general function field of transcendent degree 1.

Chapter 1

Preliminaries

In this chapter we collect some results related to discriminants, resultants, valuations, heights and twisted heights.

Unless otherwise stated, throughout this dissertation, k will be an algebraically closed field of characteristic 0 and $K = k(t)$ the rational function field in the variable t. By a function field, we always mean a finite extension of K.

1.1 Discriminants and resultants

Let L be an arbitrary field. Let

$$F(X, Y) = a_0X^n + a_1X^{n-1}Y + \cdots + a_nY^n \in L[X, Y]$$

be a binary form of degree $n \geqslant 2$.

We have a factorization $F(X, Y) = \prod_{i=1}^{n}(\alpha_iX + \beta_iY)$ over an algebraic closure $\overline{L}$ of L. As usual, we define the discriminant of F to be

$$D(F) := \prod_{i<j}(\alpha_i\beta_j - \alpha_j\beta_i)^2.$$

This is a homogeneous polynomial of degree $2n-2$ in $\mathbb{Z}[a_0, \ldots, a_n]$. In particular, for a linear form, we define its discriminant to be 1.

It is easy to show that for $U = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}\,(2, \overline{L})$ and $\lambda \in \overline{L}^*$, we have

$$\begin{aligned} D(\lambda F) &= \lambda^{2n-2} D(F), \\ D(F_U) &= (\det U)^{n(n-1)} D(F), \end{aligned}$$

where $F_U(X,Y) = F(aX+bY, cX+dY)$.

Let $F(X,Y) = a_0X^m + a_1X^{m-1}Y + \cdots + a_mY^m$ and $G(X,Y) = b_0X^n + b_1X^{n-1}Y + \cdots + b_nY^n$ be two binary forms with coefficients in L. The resultant $R(F,G)$ of F, G is defined by the determinant

$$R(F,G) := \begin{vmatrix} a_0 & a_1 & \cdots & \cdots & a_m & & & & \\ & a_0 & a_1 & \cdots & \cdots & a_m & & & \\ & & \ddots & & & & \ddots & & \\ & & & a_0 & a_1 & \cdots & \cdots & a_m \\ b_0 & b_1 & \cdots & b_n & & & & & \\ & b_0 & b_1 & \cdots & b_s & & & & \\ & & \ddots & & & \ddots & & & \\ & & & \ddots & & & \ddots & & \\ & & & & b_0 & b_1 & \cdots & b_n \end{vmatrix}, \qquad (1.1.1)$$

where the first n rows consist of coefficients of F and the last m rows of coefficients of G.

Over the algebraic closure $\overline{L}$ of L, suppose that we have factorizations

$$F(X,Y) = \prod_{i=1}^{m}(\alpha_i X + \beta_i Y), G(X,Y) = \prod_{j=1}^{n}(\gamma_j X + \delta_j Y).$$

Then

$$R(F,G) = \prod_{i=1}^{m}\prod_{j=1}^{n}(\alpha_i\delta_j - \beta_i\gamma_j). \qquad (1.1.2)$$

Hence $R(F,G) = 0$ holds exactly when F, G have a common factor.

The resultant has the following properties:

$$R(\lambda F, \mu G) = \lambda^n \mu^m R(F, G),$$

$$R(F_1 F_2, G) = R(F_1, G) R(F_2, G),$$

$$R(G, F) = (-1)^{mn} R(F, G),$$

$$R(F, G + HF) = R(F, G),$$

where $\lambda, \mu \in L$, F, G, F_1, F_2 are binary forms and H is a binary form of degree $n - m$ if $n \geqslant m$.

For an invertible matrix $U = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, define

$$F_U(X, Y) := F(aX + bY, cX + dY).$$

Then $R(F_U, G_U) = (\det U)^{mn} R(F, G)$.

1.2 Valuations on function fields

Recall $K = k(t)$. Denote by M_K the collection of normalized discrete valuations on K that are trivial on k. This set is described as follows. For $f \in k[t] \backslash \{0\}$, define $\nu_p(f)(p \in k \cup \{\infty\})$ by $f = (t-p)^{\nu_p(f)} g$ where $g \in k[t]$ and $g(p) \neq 0$ if $p \in k$; further, define $\nu_\infty(f) = -\deg f$. We extend this to $k(t)$ by setting $\nu_p(0) := \infty$ and $\nu_p(\frac{f}{g}) = \nu_p(f) - \nu_p(g)$ for $f, g \in k[t], g \neq 0$. Then $M_K = \{\nu_p : p \in k \cup \{\infty\}\}$. In this thesis we often work with absolute values. We define the absolute value $|\cdot|_\nu$ by $e^{-\nu(\cdot)}$ for $\nu \in M_K$. These absolute values satisfy the product formula

$$\prod_{\nu \in M_K} |x|_\nu = 1$$

for every $x \in K^*$. All valuations of K are non-archimedean, so for a binary form $F \in K[X, Y]$ we have

$$|D(F)|_\nu \leqslant \max_{0 \leqslant j \leqslant n} (|a_j|_\nu^{2n-2}) \tag{1.2.1}$$

for every $\nu \in M_K$. Let S be a finite set of valuations of K, containing the 'infinite valuation' ν_∞. Define the ring of S-integers and group of S-units by

$$\begin{aligned} \mathcal{O}_S &= \{x \in K : |x|_\nu \leqslant 1 \text{ for } \nu \notin S\}, \\ \mathcal{O}_S^\times &= \{x \in K : |x|_\nu = 1 \text{ for } \nu \notin S\}. \end{aligned}$$

We define the S-norm of $x \in K$ by

$$|x|_S = \prod_{\nu \in S} |x|_\nu .$$

It is clear that $|x|_S \geqslant 1$ for $x \in \mathcal{O}_S \backslash \{0\}$ and $|x|_S = 1$ for $x \in \mathcal{O}_S^\times$.

Remark 1.2.1. *Let K be a purely transcendental extension of k of transcendence degree 1. Choose t such that $K = k(t)$. The 'infinite valuation' ν_∞ is the one with $\nu_\infty(t) < 0$. The choice of the infinite valuation depends on the choice of a transcendental element t generating K. In what follows, we make a distinction between the infinite valuation ν_∞ and the other valuations on K. But we should mention that in our arguments we could as well have chosen any other valuation to play the role of the infinite valuation.*

Recall that k is an algebraically closed field of characteristic 0, and $K = k(t)$. Let L be a finite extension of K. We say a valuation ω is normalized if $\omega(L^*) = \mathbb{Z}$. Denote by M_L the normalized valuations on L that are trivial on k. For valuations $\nu \in M_K$, $\omega \in M_L$, we say that ω lies above ν, and denote it by $\omega|\nu$, if the restriction of ω to K is a positive multiple of ν. Then for every $\nu \in M_K$, we have finitely many valuations $\omega \in M_L$ above ν. For every $\omega \in M_L$, we define the corresponding absolute value $|x|_\omega := e^{-\omega(x)}$. Then we have $\omega(x) = e(\omega|\nu)\nu(x)$ for $\omega|\nu, x \in K$, where $e(\omega|\nu)$ is called the ramification index. Let L_ω denote the completion of L at ω. In our case, k is algebraically closed with $\operatorname{char} k = 0$ and the residue field of ν is k, hence the residue degree is 1, implying that $e(\omega|\nu) = [L_\omega : K_\nu]$. Thus our chosen absolute value is a prolongation of $|\cdot|_\nu^{[L_\omega:K_\nu]}$, rather than $|\cdot|_\nu$, to L, hence by Proposition 1.2.7 of [4], we have the relation $|x|_\omega = |N_{L_\omega/K_\nu}(x)|_\nu$ for every $x \in L$. By assumption, K has characteristic 0, so the extension L/K is separable. Hence

$$N_{L/K}(x) = \prod_{\omega|\nu} N_{L_\omega/K_\nu}(x) \text{ for } x \in L,$$

so we have

$$\prod_{\omega|\nu} |x|_\omega = |N_{L/K}(x)|_\nu \text{ for } x \in L, \nu \in M_K$$

and

$$\prod_{\omega \in M_L} |x|_\omega = 1 \text{ for } x \in L^*.$$

Similarly, we define the T-norm of $x \in L$ by

$$|x|_T = \prod_{\omega \in L} |x|_\omega.$$

We recall some facts about Dedekind domains. For a non-zero fractional ideal $\mathfrak{a}$ of a Dedekind domain A and a prime ideal $\wp$ of A, we denote by $v_\wp(\mathfrak{a})$ the exponent of $\wp$ in the prime ideal factorization of $\mathfrak{a}$.

Lemma 1.2.2. *There is a bijection between the non-zero prime ideals of A and the discrete valuations of F that are non-negative on A, given by $\mathfrak{p} \mapsto \nu_\mathfrak{p}$ such that $\nu_\mathfrak{p}(a)$ is the exponent of $\mathfrak{p}$ in the unique prime ideal factorization of the ideal generated by a.*

Proof. See [1]. □

Lemma 1.2.3. *Let A be a Dedekind domain with fraction field K_1. Let L be a finite separable extension of K_1, and B the integral closure of A in L.*

Assume that L/K_1 is tamely ramified. Denote by $D_{B/A}$ the discriminant ideal and $\mathfrak{D}_{B/A}$ the different ideal of B over A. Let $\mathfrak{p}$ be a prime ideal of A, let $\wp_1, \ldots, \wp_r$ be the prime ideals of B above $\mathfrak{p}$, and ν the valuation corresponding to $\mathfrak{p}$, and ω_i corresponding to $\wp_i$ for $i = 1, \ldots, r$. Then

$$N_{L/K_1}(\mathfrak{D}_{B/A}) = D_{B/A}.$$

Further

$$\nu(D_{B/A}) = \sum_{i=1}^{r} \Big(e(\omega_i|\nu) - 1 \Big).$$

Proof. For the first part, see Proposition 6, §3, Chapter III of [22].

Since the extension L/K_1 is tamely ramified with residue degree $f(\omega_i|\nu) = 1$, we get by Proposition 13, §6, Chapter III of [22],

$$\omega_i(\mathfrak{D}_{B/A}) = e(\omega_i|\nu) - 1 \text{ for } i = 1, \ldots, r,$$

hence

$$\nu(D_{B/A}) = \nu\Big(N_{L/K_1}(\mathfrak{D}_{B/A})\Big) = \sum_{i=1}^{r} \Big(e(\omega_i|\nu) - 1 \Big),$$

which gives the claim. □

Later we will apply this lemma frequently to the case $K_1 = k(t)$, $A = k[t]$ and $K_1 = K_\nu$, the completion of K at ν and $A = R_\nu := \{x \in K_\nu : \nu(x) \geqslant 0\}$ for $\nu \in M_K$.

1.3 Polynomials and heights

Recall $K = k(t)$. For $\nu \in M_K$, denote by K_ν the completion of K at the valuation ν. Then ν has a unique extension to K_ν. Define

$$R_\nu = \{x \in K_\nu : \nu(x) \geqslant 0\}$$

to be the local ring of K_ν. Then its group of units is

$$R_\nu^\times = \{x \in K_\nu : \nu(x) = 0\}.$$

For $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in K_\nu^n$, define

$$\begin{aligned} \nu(\mathbf{x}) &= \min_{1 \leqslant i \leqslant n} \nu(x_i), \\ \|\mathbf{x}\|_\nu &= e^{-\nu(\mathbf{x})} = \max_{1 \leqslant i \leqslant n} |x_i|_\nu, \end{aligned}$$

and for $\mathbf{x} \in K^n$, define the homogeneous height and S-height

$$H_K(\mathbf{x}) = \prod_{\nu \in M_K} \|\mathbf{x}\|_\nu,$$

$$H_S(\mathbf{x}) = \prod_{\nu \in S} \|\mathbf{x}\|_\nu.$$

Clearly, the product is well-defined and $H_K(\mathbf{x}) \geqslant 1$ for every $\mathbf{x} \neq \mathbf{0}$ because of the product formula. Also, $H_K(\lambda \mathbf{x}) = H_K(\mathbf{x})$.

For a polynomial $P \in K[X_1, \ldots, X_n]$ or $P \in K_\nu[X_1, \ldots, X_n]$ we define $|P|_\nu$ to be the maximum of the $|\cdot|_\nu$-values of its coefficients.

Lemma 1.3.1 (Gauss' lemma)**.** *Let K be a field, $|\cdot|_\nu$ a non-archimedean absolute value on K, and $P = \prod_{i=1}^{t} P_i$ with $P_i \in K[X_1, \ldots, X_n]$ for $i = 1, \ldots, t$. Then*

$$|P|_\nu = \prod_{i=1}^{t} |P_i|_\nu.$$

Proof. See [14]. □

As a direct consequence, we have

Corollary 1.3.2. *Let* $F = \prod_{i=1}^{n}(\alpha_i X + \beta_i Y)$ *with* $\alpha_i, \beta_i \in K$ *for* $i = 1, \ldots, n$. *Then* $|F|_\nu = \prod_{i=1}^{n} \max(|\alpha_i|_\nu, |\beta_i|_\nu)$ *for every* $\nu \in M_K$.

For L a finite extension of K and a polynomial $P \in L[X_1, \ldots, X_m]$, we define

$$N_{L/K}(P) = \prod_{i=1}^{[L:K]} \sigma_i(P),$$

where $\sigma_1, \ldots, \sigma_{[L:K]}$ are the K-embeddings of L into $\overline{K}$, and $\sigma_i(P)$ is obtained by the action of σ_i on the coefficients of P.

1.4 Galois theory of valuations

In this section, we give a brief sketch of some aspects of Galois theory of valuations that will be needed later.

Lemma 1.4.1. *Let* K *be a field with a non-trivial absolute value* $|\cdot|_\nu$, *and* L *a finite Galois extension of* K *with Galois group* $G = \mathrm{Gal}(L/K)$. *Then for every two absolute values* $|\cdot|_\omega, |\cdot|_{\omega'}$ *on* L *prolonging* $|\cdot|_\nu$, *there is* $\sigma \in G$ *such that* $|x|_\omega = |\sigma(x)|_{\omega'}$ *for* $x \in L$.

Proof. See Corollary 1.3.5 of [4]. □

For $\nu \in M_K$ and L a Galois extension of K, denote by $\mathcal{A}(\nu)$ the set of normalized valuations of L above ν. Fix $\omega_1 \in \mathcal{A}(\nu)$. The completion L_{ω_1} of L at ω_1 is a Galois extension of K_ν. We may view L as a subfield of L_{ω_1}. As mentioned before, the absolute values on L defined above satisfy the relation $|x|_{\omega_1} = |N_{L_{\omega_1}/K_\nu}(x)|_\nu$ for $x \in L_{\omega_1}$. Without loss of generality, we may assume $K \subset K_\nu \subset L_{\omega_1} \subset \overline{K_\nu}$ and $K \subset L \subset L_{\omega_1} \subset \overline{K_\nu}$. Let $\mathcal{E}(\omega_1|\nu)$ be the set $\{\sigma \in G : \omega_1 \circ \sigma = \omega_1\}$ equipped with composition. This is by definition the decomposition group of ω_1 over ν. By, for instance, §9, Chapter II of [18], we have an isomorphism

$$\begin{aligned} \mathrm{Gal}(L_{\omega_1}/K_\nu) &\xrightarrow{\sim} \mathcal{E}(\omega_1|\nu), \\ \sigma &\longmapsto \sigma|_L. \end{aligned}$$

Thus we may view $\mathrm{Gal}(L_{\omega_1}/K_\nu)$ as a subgroup of G. Further, let

$$\mathcal{E}(\omega|\nu) = \{\sigma \in G : \omega = \omega_1 \circ \sigma\} \text{ for } \omega \in \mathcal{A}(\nu). \tag{1.4.1}$$

Since G acts transitively on $\mathcal{A}(\nu)$ (see §9, Chapter II, [18]), the sets $\mathcal{E}(\omega|\nu)$ form a partition of G, and in fact they are the right cosets of $\mathrm{Gal}(L_{\omega_1}/K_\nu)$ in G, so have the same cardinality:

$$[L_\omega : K_\nu] = [L_{\omega'} : K_\nu] \text{ for } \omega, \omega' \text{ above } \nu. \tag{1.4.2}$$

It is now reasonable to put $g_\nu := \#\mathcal{E}(\omega|\nu) = [L_{\omega_1} : K_\nu]$. If we still denote by $|\cdot|_\nu$ the prolongation of $|\cdot|_\nu$ from K to $\overline{K_\nu}$, and hence on L_{ω_1}, then $|x|_\nu = |N_{L_{\omega_1}/K_\nu}(x)|_\nu^{1/[L_{\omega_1}:K_\nu]}$ for $x \in L_{\omega_1}$. It follows that for $x \in L, \omega \in \mathcal{A}(\nu), \sigma \in \mathcal{E}(\omega|\nu)$, we have

$$|x|_\omega = |\sigma(x)|_{\omega_1} = |\sigma(x)|_\nu^{g_\nu}. \tag{1.4.3}$$

Notice that $\sigma \in \mathrm{Gal}(L/K)$, hence we may extend $\sigma \in \mathcal{E}(\omega|\nu)$ to a K_ν-isomorphism from L_ω to L_{ω_1}, by sending $\alpha = \lim_{n\to\infty} \alpha_n$ to $\sigma(\alpha) = \lim_{n\to\infty} \sigma(\alpha_n)$ where $\alpha \in L_\omega$ and $\alpha_n \in L$. Moreover, for every $x \in L_\omega$, we also have $|x|_\omega = |\sigma(x)|_{\omega_1} = |\sigma(x)|_\nu^{g_\nu}$.

1.5 Twisted heights

Let S be a finite set of valuations of K. We define the ring of S-adeles

$$\mathbb{A}_S := \prod_{\nu \in S} K_\nu = \{(x_\nu) | x_\nu \in K_\nu \text{ for every } \nu \in S\}$$

with componentwise addition and multiplication.

Further, let

$$\mathrm{GL}_n(\mathbb{A}_S) = \{(A_\nu) | A_\nu \in \mathrm{GL}_n(K_\nu) \text{ for every } \nu \in S\},$$

where $\mathrm{GL}_n(R_\nu)$ is the subgroup of $\mathrm{GL}_n(K_\nu)$ of $n \times n$ matrices whose entries are in R_ν and whose determinant is in $R_\nu^\times$.

For $A = (A_\nu) \in \mathrm{GL}_n(\mathbb{A}_S)$, define

$$|\det(A)|_S := \prod_{\nu \in S} |\det(A_\nu)|_\nu.$$

Also, we define the ν-norm of A_ν as follows: if $A_\nu = (a_{ij})_{1 \leqslant i,j \leqslant n}$, then $\|A_\nu\|_\nu = \max_{i,j} |a_{ij}|_\nu$. Given a ring R we denote by R^n the module of n-dimensional column vectors with entries in R.

Lemma 1.5.1. *Let $\nu \in M_K$. For $A_\nu \in GL_n(R_\nu)$ and $\mathbf{x} \in K_\nu^n$, we have $\nu(A_\nu \mathbf{x}) = \nu(\mathbf{x})$.*

Proof. Let $A_\nu = (a_{ij}), \mathbf{x} = (x_1, \dots, x_n) \in K^n$.

As $\min_{i,j} \nu(a_{ij}) \geqslant 0$, we have

$$\begin{aligned} \nu(A_\nu \mathbf{x}) &\geqslant \min_{1 \leqslant i \leqslant n} \nu(a_{i1}x_1 + \cdots + a_{in}x_n) \\ &\geqslant \min_{1 \leqslant i,j \leqslant n} \nu(a_{ij}x_j) \\ &\geqslant \min_{1 \leqslant j \leqslant n} \nu(x_i) + \min_{i,j} \nu(a_{ij}) \\ &\geqslant \nu(\mathbf{x}). \end{aligned}$$

Since $A_\nu^{-1} \in \mathrm{GL}_n(R_\nu)$, we have similarly for $A_\nu \in \mathrm{GL}_n(R_\nu), \mathbf{x} \in K^n$ that $\nu(\mathbf{x}) = \nu(A_\nu^{-1} A_\nu \mathbf{x}) \geqslant \nu(A_\nu \mathbf{x})$. This completes the proof. □

For $A \in \mathrm{GL}_n(\mathbb{A}_S), \mathbf{x} \in K^n$ define the divisor

$$\mathrm{div}_A(\mathbf{x}) := \sum_{\nu \in S} \nu(A_\nu \mathbf{x})\nu + \sum_{\nu \notin S} \nu(\mathbf{x})\nu$$

and its degree

$$\deg(\mathrm{div}_A(\mathbf{x})) = \sum_{\nu \in S} \nu(A_\nu \mathbf{x}) + \sum_{\nu \notin S} \nu(\mathbf{x}).$$

Also define the corresponding twisted additive height

$$h_A(\mathbf{x}) := -\deg(\mathrm{div}_A(\mathbf{x})) = -\sum_{\nu \in S} \nu(A_\nu \mathbf{x}) - \sum_{\nu \notin S} \nu(\mathbf{x}).$$

The sum is well-defined by the fact that for every $\mathbf{x} \in K^*$, we have $\nu(\mathbf{x}) = 0$ for almost all $\nu \in M_K$. Define the twisted multiplicative height for $\mathbf{x} \in K^n$ by:

$$H_A(\mathbf{x}) := \exp(h_A(\mathbf{x})) = \prod_{\nu \in S} \|A_\nu \mathbf{x}\|_\nu \prod_{\nu \notin S} \|\mathbf{x}\|_\nu.$$

It is projective in the sense that, by the product formula, $H_A(\lambda\mathbf{x}) = H_A(\mathbf{x})$ for $\mathbf{x} \in K^n, \lambda \in K^\times$.

Lastly, we define for $A \in \mathrm{GL}_n(\mathbb{A}_S)$

$$\mathrm{div}(A) := \mathrm{div}_A(K^n) := \sum_{\nu \in S} \nu(\det(A_\nu))\nu,$$

and

$$\begin{aligned} h_A(K^n) &:= -\deg(\mathrm{div}(A)), \\ H_A(K^n) &:= \exp(h_A(K^n)) = \prod_{\nu \in S} |\det A_\nu|_\nu = |\det(A)|. \end{aligned}$$

Lemma 1.5.2. *Let $A \in GL_n(\mathbb{A}_S)$. Then there exist positive constants c_1, c_2 depending on A such that $c_2 H_K(\mathbf{x}) \leqslant H_A(\mathbf{x}) \leqslant c_1 H_K(\mathbf{x})$ for all $\mathbf{x} \in K^n$. In particular, for $\mathbf{x} \neq \mathbf{0}$, we have $H_A(\mathbf{x}) \geqslant c_2$.*

Proof. Let $c_1 = \prod_{\nu\in S} \|A_\nu\|_\nu$ and $c_2 = \prod_{\nu\in S} \|A_\nu^{-1}\|_\nu^{-1}$.

Clearly, we have $\|A_\nu \mathbf{x}\|_\nu \leqslant \|A_\nu\|_\nu \|\mathbf{x}\|_\nu$ because for all $\nu \in S$, the valuation is non-archimedean. Similarly we have $\|\mathbf{x}\|_\nu = \|A_\nu^{-1}A_\nu\mathbf{x}\|_\nu \leqslant \|A_\nu^{-1}\|_\nu \|A_\nu\mathbf{x}\|_\nu$, hence $\|A_\nu^{-1}\|_\nu^{-1}\|\mathbf{x}\|_\nu \leqslant \|A_\nu\mathbf{x}\|_\nu \leqslant \|A_\nu\|_\nu\|\mathbf{x}\|_\nu$ for $\nu \in S$. By taking the product over all $\nu \in M_K$ we get $c_2 H_K(\mathbf{x}) \leqslant H_A(\mathbf{x}) \leqslant c_1 H_K(\mathbf{x})$. □

Consider a finite extension L of K. Let S be a finite subset of M_K and let $T \subset M_L$ be the set of valuations of L lying above those of S. For $x \in L$ put $|x|_T := \prod_{\omega\in T} |x|_\omega$. Define the ring of T-integers and T-units

$$\mathcal{O}_T := \{x \in L : |x|_\omega \leqslant 1 \text{ for } \omega \notin T\},$$

$$\mathcal{O}_T^\times := \{x \in L : |x|_\omega = 1 \text{ for } \omega \notin T\}.$$

Then $\mathcal{O}_T$ is the integral closure of $\mathcal{O}_S$ in L. We have

$$|x|_T = |N_{L/K}(x)|_S \text{ for } x \in L, \tag{1.5.1}$$

and in particular,

$$|x|_T = |x|_S^{[L:K]} \text{ for } x \in K. \tag{1.5.2}$$

For $\omega \in M_L$, denote by L_ω the completion of L at ω. Then there is a unique extension of ω to L_ω. For $\mathbf{x} = (x_1, \ldots, x_n)^{\mathrm{T}} \in L_\omega^n$, we define

$$\begin{aligned} \omega(\mathbf{x}) &= \min_{1\leqslant i\leqslant n} \omega(x_i), \\ \|\mathbf{x}\|_\omega &= \max_{1\leqslant i\leqslant n} |x_i|_\omega = \max_{1\leqslant i\leqslant n} e^{-\omega(x_i)}. \end{aligned}$$

Similarly as before, we define $\mathrm{div}_A(\mathbf{x}), \mathrm{div}(A)$ for $\mathbf{x} \in L^n, A \in \mathrm{GL}_n(\mathbb{A}_T)$ by replacing K, S with L, T respectively. That is,

$$\mathrm{div}_A(\mathbf{x}) := \sum_{\omega\in M_L} \omega(A_\omega \mathbf{x})\omega,$$

$$\mathrm{div}(A) := \sum_{\omega \in M_L} \omega(\det(A_\omega))\omega.$$

Define

$$\begin{aligned} h_A(\mathbf{x}) &:= -\deg(\mathrm{div}_A(\mathbf{x}))/[L:K], \\ h_A(L^n) &:= -\deg(\mathrm{div}(A))/[L:K], \end{aligned}$$

and

$$\begin{aligned} H_A(\mathbf{x}) &:= \exp(h_A(\mathbf{x})) = \Big(\prod_{\omega \in M_L} \|A_\omega \mathbf{x}\|_\omega\Big)^{\frac{1}{[L:K]}}, \\ H_A(L^n) &:= \exp(h_A(K^n)) = \Big(\prod_{\omega \in M_L} |\det A_\omega|_\omega\Big)^{\frac{1}{[L:K]}} = |\det(A)|_L^{\frac{1}{[L:K]}}. \end{aligned}$$

The height H_A on L^n is compatible with the one on K^n: $H_A(L^n) = H_A(K^n)$.

We recall Thunder's analogue of Minkowski's convex body theorem for function fields.

Lemma 1.5.3. *Let L be a finite extension of K of degree m, and H_A be the twisted height on L^n corresponding to $A \in GL_n(\mathbb{A}_S)$. Then there is a basis $\mathbf{a}_1, \ldots, \mathbf{a}_n$ of L^n satisfying*

$$\prod_{i=1}^n H_A(\mathbf{a}_i) \leqslant H_A(L^n) e^{n(g_L+m-1)/m}.$$

where g_L is the genus of L.

Proof. See Theorem 1 of [24]. □

Lemma 1.5.4. *For every basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ of L^n, we have*

$$\prod_{i=1}^n H_A(\mathbf{x}_i) \geqslant H_A(L^n).$$

In particular, there is a basis $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ *of* K^n *such that*

$$\prod_{i=1}^{n} H_A(\mathbf{a}_i) = H_A(K^n).$$

Proof. See Lemma 5 of [24] for the inequality. The equality is a combination with Lemma 1.5.3. □

Chapter 2

Height estimates for solutions of S-unit equations

Let $|\cdot|_\infty$ denote the ordinary absolute value on $\mathbb{Q}$ and for a prime p, denote by $|\cdot|_p$ the p-adic absolute value, normalized such that $|p|_p = p^{-1}$. Let K be a number field and M_K its collection of places (equivalence classes of absolute values). For every $\nu \in M_K$, choose $|\cdot|_\nu$ from ν such that if ν lies above $p \in \{\infty\} \cup \{\text{primes}\}$. Then $|x|_\nu = |x|_p^{[K_\nu:\mathbb{Q}_p]}$ for $x \in \mathbb{Q}$.

We recall the Subspace Theorem, due to Schmidt and Schlickewei.

For $\mathfrak{X} = [x_0 : \cdots : x_n] \in \mathbb{P}^n(K)$, define $|\mathfrak{X}|_\nu := \max(|x_1|_\nu, \ldots, |x_n|_\nu)$ for $\nu \in M_K$ and $H_K(\mathfrak{X}) = \prod_{\nu \in M_K} |\mathfrak{X}|_\nu$.

Subspace Theorem. *Let $n \geqslant 1$, and let S be a finite set of places of K. For $\nu \in S$, let $L_{0\nu}, \ldots, L_{n\nu}$ be linearly independent linear forms with coefficients in K. Further, let $C > 0, \delta > 0$. Then the set of solutions of the inequality*

$$\prod_{\nu \in S} \frac{|L_{0\nu}(\mathfrak{X}) \ldots L_{n\nu}(\mathfrak{X})|_\nu}{|\mathfrak{X}|_\nu^{n+1}} \leqslant C H_K(\mathfrak{X})^{-n-1-\delta}$$

in $\mathfrak{X} \in \mathbb{P}^n(K)$ is contained in a finite union of proper linear subspaces of

$\mathbb{P}^n(K)$.

This was proved by Schmidt in [20], [21] in the case that S contains only archimedean places, and by Schlickewei [19] in full generality.

As a consequence, in [9] Evertse derived the following result.

Let S be a finite set of places of K containing all archimedean places. Define the ring of S-integers $\mathcal{O}_S = \{x \in K : |x|_\nu \leqslant 1 \text{ for } \nu \notin S\}$. Define

$$|x|_S := \prod_{\nu \in S} |x|_\nu \text{ for } x \in \mathcal{O}_S,$$

$$H_S(x_1, \ldots, x_n) := \prod_{\nu \in S} \max(|x_1|_\nu, \ldots, |x_n|_\nu) \text{ for } x_1, \ldots, x_n \in \mathcal{O}_S.$$

Theorem (Evertse)**.** *Let K be an algebraic number field and S a finite set of valuations of K containing those archimedean ones. Assume $x_1, \ldots, x_n \in \mathcal{O}_S$ such that $\sum_{i=1}^{n} x_i = 0$ but no non-empty proper subsum vanishes. Then for every $\varepsilon > 0$ we have*

$$H_S(x_1, \ldots, x_n) < C(n, \varepsilon, S) |\prod_{i=1}^{n} x_i|_S^{1+\varepsilon}.$$

Here $C(n, \varepsilon, S)$ is an ineffective constant. In this chapter, we are going to prove a much stronger analogue of this result over function fields.

2.1 Height estimates

Let $K = k(t)$, L a finite extension of K. For $x_1, \ldots, x_n \in L$, define

$$H_L(x_1, \ldots, x_n) = \prod_{\omega \in M_L} \max(|x_1|_\omega, \ldots, |x_n|_\omega),$$

$$H_L^*(x_1, \ldots, x_n) = \prod_{\omega \in M_L} \max(1, |x_1|_\omega, \ldots, |x_n|_\omega).$$

$$H(x_1,\ldots,x_n)=\prod_{\omega\in M_L}\max(|x_1|_\omega,\ldots,|x_n|_\omega)^{1/[L:K]},$$

$$H^*(x_1,\ldots,x_n)=\prod_{\omega\in M_L}\max(1,|x_1|_\omega,\ldots,|x_n|_\omega)^{1/[L:K]}.$$

For a finite set $T\subset M_L$, define

$$H_T(x_1,\ldots,x_n)=\prod_{\omega\in T}\max(|x_1|_\omega,\ldots,|x_n|_\omega).$$

Lemma 2.1.1 (Mason)**.** *Let L be a finite extension of $K=k(t)$, and T a finite set of valuations of L. Let $\gamma_1,\gamma_2,\gamma_3$ be non-zero elements of L satisfying $\gamma_1+\gamma_2+\gamma_3=0$ and $\nu(\gamma_1)=\nu(\gamma_2)=\nu(\gamma_3)$ for every valuation $\nu\notin T$. Then either $\frac{\gamma_1}{\gamma_2}\in k$, which means $H^*(\frac{\gamma_1}{\gamma_2})=1$, or $H^*(\frac{\gamma_1}{\gamma_2})\leqslant e^{(\#T+2g_L-2)/[L:K]}$.*

Proof. See Chapter I, Lemma 2 of [17]. □

Corollary 2.1.2. *With the above notation, we have in both the cases $\frac{\gamma_1}{\gamma_2}\in k,\frac{\gamma_1}{\gamma_2}\notin k$ that*

$$H^*(\frac{\gamma_1}{\gamma_2})\leqslant e^{(\#T+2g_L-1)/[L:K]}.$$

Proof. This follows directly from the facts that $g_L\geqslant 0$ and $\#T\geqslant 1$. □

Recall

$$\mathcal{O}_T:=\{x\in L:|x|_\omega\leqslant 1\text{ for }\omega\notin T\},$$

$$\mathcal{O}_T^\times:=\{x\in L:|x|_\omega=1\text{ for }\omega\notin T\}.$$

Note that by the product formula, we have

$$\begin{aligned}H^*(\frac{\gamma_1}{\gamma_2}) &= \Big(\prod_{\omega\in M_L}\max(1,\left|\frac{\gamma_1}{\gamma_2}\right|_\omega)\Big)^{1/[L:K]}\\ &= \Big(\prod_{\omega\in M_L}\max(|\gamma_1|_\omega,|\gamma_2|_\omega)\Big)^{1/[L:K]}=H(\gamma_1,\gamma_2),\end{aligned}$$

and if $\gamma_1, \gamma_2 \in \mathcal{O}_T$, then $H(\gamma_1,\gamma_2)^{[L:K]} \leqslant H_T(\gamma_1,\gamma_2)$.

Brownawell and Masser obtained the following generalization:

Theorem 2.1.3. *Let L be a finite extension of $K = k(t)$, and T a finite set of valuations of L. Put $g' = \max(0, 2g-2)$. Let $u_1,\dots,u_n$ be T-units in L satisfying $u_1+\dots+u_n = 0$ but $\sum_{i\in I} u_i \neq 0$ for every non-empty proper subset I of $\{1,\dots,n\}$. Then*

$$H(u_1,\dots,u_n) \leqslant e^{\frac{1}{2}(n-1)(n-2)(\#T+g')/[L:K]}.$$

Proof. See [6]. □

We deduce the following result, which will be improved in the next section.

Corollary 2.1.4. *Let L be a finite extension of $K = k(t)$, and T a finite set of valuations of L. Put $g' = \max(0, 2g-2)$. Let $u_1,\dots,u_n$ be elements of $\mathcal{O}_T$ satisfying $u_1+\dots+u_n = 0$ but $\sum_{i\in I} u_i \neq 0$ for every non-empty proper subset I of $\{1,\dots,n\}$. Then*

$$H_T(u_1,\dots,u_n) \leqslant e^{\frac{1}{2}(n-1)(n-2)(\#T+g')}\Big|\prod_{i=1}^{n} u_i\Big|_T^{\frac{(n-1)(n-2)}{2}}.$$

Proof. Let U be the collection of $\omega \in M_L\backslash T$ such that $\omega(u_i)$, $i = 1,\dots,n$, are not all equal. Then clearly $\#U < \infty$.

Now consider the complement of $T \cup U$. For every $\omega \notin T \cup U$, we have $\omega(u_1) = \dots = \omega(u_n)$. Since $u_i \in \mathcal{O}_T$, there are two cases: either $\omega(u_i) = 0$, which is the case for almost all valuations, or $\omega(u_i) > 0$. Let $V = \{\omega \notin T \cup U : \omega(u_1) = \dots = \omega(u_n) > 0\}$.

If $V = \emptyset$, then by Theorem 2.1.3, we have $H(u_1,\dots,u_n) \leqslant e^{\frac{(n-1)(n-2)(\#T+\#U+g')}{2[L:K]}}$.

If $V \neq \emptyset$, then $\frac{u_1}{u_n}+\dots+\frac{u_{n-1}}{u_n}+1 = 0$ and each nontrivial partial sum is non-zero by assumption. As $\frac{u_i}{u_n}, i = 1,\dots,n-1$, and 1 are all elements of

$\mathcal{O}^*_{T\cup U}$, and the height function H is projective, we obtain by Theorem 2.1.3

$$H(u_1,\dots,u_n)=H(\frac{u_1}{u_n},\dots,\frac{u_{n-1}}{u_n},1)\leqslant e^{\frac{(n-1)(n-2)(\#T+\#U+g')}{2[L:K]}}. \qquad (2.1.1)$$

On the other hand, since $u_i\in\mathcal{O}_T$ for $i=1,\dots,n$, we have $\max\limits_{1\leqslant i\leqslant n}|u_i|_\omega\leqslant 1$, hence $\min\limits_{1\leqslant i\leqslant n}|u_i|_\omega\leqslant e^{-1}$ for $\omega\in U$, and therefore

$$e^{\#U}\leqslant\prod_{\omega\notin T}\frac{1}{\min\limits_{1\leqslant i\leqslant n}|u_i|_\omega}. \qquad (2.1.2)$$

Combining (2.1.1) with (2.1.2) we derive that

$$\begin{aligned}H_T(u_1,\dots,u_n) &\leqslant e^{\frac{1}{2}(n-1)(n-2)(\#T+g')}\Big(\prod_{\omega\notin T}\frac{1}{\min\limits_{1\leqslant i\leqslant n}|u_i|_\omega}\Big)^{\frac{(n-1)(n-2)}{2}}\prod_{\omega\notin T}\frac{1}{\max\limits_{1\leqslant i\leqslant n}|u_i|_\omega}\\ &\leqslant e^{\frac{1}{2}(n-1)(n-2)(\#T+g')}\prod_{\omega\notin T}\prod_{i=1}^{n}|u_i|_\omega^{-\frac{(n-1)(n-2)}{2}}\\ &= e^{\frac{1}{2}(n-1)(n-2)(\#T+g')}|\prod_{i=1}^{n}u_i|_T^{\frac{(n-1)(n-2)}{2}},\end{aligned}$$

as claimed. □

2.2 S-unit equations and heights

Actually, from an effective version of the subspace theorem over function fields, we can deduce better results.

The following theorem is originally stated in terms of additive heights and over function fields K_1 associated to arbitrary nonsingular varieties. We restate it in our notation in the special case for curves, i.e., for function fields of transcendence degree 1. For $n\in\mathbb{Z}_{\geqslant 1}$, put

$$C(n)=e^{\binom{n}{2}(2g_{K_1}-2+\#S_1)}, C'(n)=e^{\binom{n}{2}\max(0,2g_{K_1}-2+\#S_1)}.$$

Theorem 2.2.1. *Let K_1 be a finite extension of $K = k(t)$ and $L_1, \dots, L_q$ hyperplanes in $\mathbb{P}^N(K_1)$ defined by linear forms with coefficients in k. Let $S_1 \subset M_{K_1}$ be a finite set of valuations. If the coordinates of $\mathfrak{X} = [x_0 : \cdots : x_N] \in \mathbb{P}^N(K_1)$ are linearly independent over k, then*

$$\prod_{\nu \in S_1} \min_J \prod_{j \in J} \frac{|L_j(\mathfrak{X})|_\nu}{|\mathfrak{X}|_\nu} \geqslant C(N+1)^{-1} H(\mathfrak{X})^{-(N+1)[K_1:k(t)]},$$

where the minimum is taken over all subsets J of $\{1, \dots, q\}$ such that the linear forms $L_j\,(j \in J)$ are linearly independent.

Proof. See Theorem 1 of [25]. □

Corollary 2.2.2. *Assume $x_1, \dots, x_n \in \mathcal{O}_{S_1}$ are k-linearly independent. Then*

$$\left(\prod_{i=1}^n |x_i|_{S_1}\right) |x_1 + \cdots + x_n|_{S_1} \geqslant C(n)^{-1} H_{S_1}(x_1, \dots, x_n).$$

Proof. We apply Theorem 2.2.1 with $N = n-1, \mathfrak{X} = [x_1 : \cdots : x_n], L_i = x_i\,(i = 1, \dots, n), L_{n+1} = x_1 + \cdots + x_n$.

For each $\nu \in S_1$, choose $t(\nu) \in \{1, \dots, n\}$ such that $|x_{t(\nu)}|_\nu = \max_{1 \leqslant i \leqslant n} (|x_i|_\nu)$, and take $J(\nu) = \{1, \dots, n+1\} \backslash \{t(\nu)\}$. Then

$$\prod_{\nu \in S_1} \prod_{j \in J(\nu)} \frac{|L_j(\mathfrak{X})|_\nu}{|\mathfrak{X}|_\nu} \geqslant C(n)^{-1} H(\mathfrak{X})^{-n[K_1:k(t)]},$$

hence as $x_i \in \mathcal{O}_{S_1}$,

$$\frac{\left(\prod_{i=1}^n |x_i|_{S_1}|\right) |x_1 + \cdots + x_n|_{S_1}}{H_{S_1}(\mathfrak{X})^{n+1}} \geqslant C(n)^{-1} H_{S_1}(\mathfrak{X})^{-n}.$$

This completes the proof. □

Actually, the condition that $x_1, \dots, x_n$ be k-linearly independent can be relaxed to the condition that $x_1 + \cdots + x_n$ have no vanishing subsum.

Corollary 2.2.3. *Let $x_1, \dots, x_n \in \mathcal{O}_{S_1}$ such that $\sum_{i \in I} x_i \neq 0$ for any non-empty subset $I \subset \{1, \dots, n\}$. Let S_1 be a finite subset of M_{K_1}, T a subset of S_1. Then*

$$\prod_{i=1}^{n} |x_i|_{S_1} |x_1 + \cdots + x_n|_T \geqslant C'(n)^{-1} H_T(x_1, \dots, x_n).$$

Proof. We proceed by induction on n. For $n = 1$ the assertion is trivial since $x_1 \in \mathcal{O}_{S_1}$. Let $N \geqslant 2$ and assume the assertion is true for $n < N$. We now consider the case $n = N$. Since each $\nu \in S_1$ is non-archimedean, i.e., $|x_1 + \cdots + x_n|_\nu \leqslant \max_{1 \leqslant i \leqslant n} |x_i|_\nu$, it suffices to deal with the special case $T = S_1$.

First suppose that $x_1, \dots, x_N$ are k-linearly independent. Then the assertion is true by Corollary 2.2.2. Next assume that $\text{rank}_k\{x_1, \dots, x_N\} < N$. Then, possibly after rearranging the indices, we may assume that $x_1 + \cdots + x_N = a_1 x_1 + \cdots + a_u x_u$, where $1 \leqslant u < N$, $a_1, \dots, a_u \in k^*$ and u is minimal with this property. Then $x_1, \dots, x_u$ are k-linearly independent and no subsum of the right-hand side is 0. Partition S_1 into two subsets

$$S^{(1)} = \{\nu \in S_1 : \max_{1 \leqslant i \leqslant N} (|x_i|_\nu) = \max_{1 \leqslant i \leqslant u} (|x_i|_\nu)\},$$

$$S^{(2)} = \{\nu \in S_1 : \max_{1 \leqslant i \leqslant N} (|x_i|_\nu) > \max_{1 \leqslant i \leqslant u} (|x_i|_\nu)\}.$$

Then we have $x_{u+1} + \cdots + x_N = (a_1 - 1)x_1 + \cdots + (a_u - 1)x_u$ and hence $|x_{u+1} + \cdots + x_N|_\nu \leqslant \max_{1 \leqslant i \leqslant u} (|x_i|_\nu)$ for $\nu \in S^{(2)}$. Combining this with the induction hypothesis, we derive that

$$\begin{aligned}
&\Big(\prod_{i=1}^{N}|x_i|_{S_1}\Big)|x_1+\cdots+x_N|_{S_1}\\
=\;&\Big(\prod_{i=1}^{u}|a_ix_i|_{S_1}|a_1x_1+\cdots+a_ux_u|_{S_1}\Big)\prod_{i=u+1}^{N}|x_i|_{S_1}\\
\geqslant\;& C(u)^{-1}H_{S_1}(a_1x_1,\ldots,a_ux_u)\prod_{i=u+1}^{N}|x_i|_{S_1}\\
=\;& C(u)^{-1}H_{S^{(1)}}(x_1,\ldots,x_u)H_{S^{(2)}}(x_1,\ldots,x_u)\prod_{i=u+1}^{N}|x_i|_{S_1}\\
\geqslant\;& C(u)^{-1}H_{S^{(1)}}(x_1,\ldots,x_N)|x_{u+1}+\cdots+x_N|_{S^{(2)}}\prod_{i=u+1}^{N}|x_i|_{S_1}\\
\geqslant\;& C(u)^{-1}H_{S^{(1)}}(x_1,\ldots,x_N)C'(N-u)^{-1}H_{S^{(2)}}(x_{u+1},\ldots,u_N)\\
\geqslant\;& C'(N)^{-1}H_{S_1}(x_1,\ldots,x_N),
\end{aligned}$$

which completes the induction step. □

With the help of two lemmas stated below and a similar idea as in the proof of Theorem 1, [26], we obtain a generalization of Theorem 1 of [26]. The following two lemmas are from [25], which deals with a more general case. We restate and prove the lemmas in our specific case. Recall that for every $z \in K_1\backslash k$ we have a derivation d/dz. For each valuation $\nu \in M_{K_1}$, we choose a local parameter $\xi = \xi_\nu$ with $\nu(\xi) = 1$. Then we have another corresponding derivation $d/d\xi$.

Let $f_1, \ldots, f_n \in K_1$ be k-linearly independent. Define the Wronskian related to z as $W = W_z(f_1, \ldots, f_n) := \det\big((d/dz)^{j-1}f_i\big)_{1\leqslant i,j\leqslant n}$. Then it is well-known that $W \neq 0$.

Lemma 2.2.4. *(i) For $h \in K_1$, we have*

$$W_z(hf_1, \ldots, hf_n) = h^n W_z(f_1, \ldots, f_n).$$

(ii) For any $\xi \in K_1 \backslash k$ we have

$$W_z(f_1, \dots, f_n) = \det\Big((\frac{d\xi}{dz})^{j-1} \cdot (d/d\xi)^{j-1} f_i \Big)_{1 \leqslant i,j \leqslant n}.$$

Proof. For (i), assume $h \neq 0$, otherwise it is trivial. By the Leibniz rule for derivatives, we have for each $1 \leqslant i \leqslant n$ that

$$(d/dz)^m (hf_i) - h(d/dz)^m f_i = \sum_{l=0}^{m-1} \binom{m}{l} (d/dz)^{m-l} h \cdot (d/dz)^l f_i,$$

is a K_1-linear combination of $h(d/dz)^l f_i, 0 \leqslant l \leqslant m-1$. The determinant remains unchanged if we recursively replace the j-th column by $(h(d/dz)^{j-1} f_1, \dots, h(d/dz)^{j-1} f_n)^{\mathrm{T}}$ for $j = 2, \dots, n$. Then the assertion follows immediately.

For (ii), we will prove by induction that $(d/dz)^m f_i - (\frac{d\xi}{dz})^m \cdot (d/d\xi)^m f_i$ is a K_1-linear combination, independent of i, of $(d/d\xi) f_i, \dots, (d/d\xi)^{m-1} f_i$. Then the assertion is clear for the same reason as in (i). By the chain rule, we know

$$(d/dz) f_i = \frac{d\xi}{dz} \cdot (d/d\xi) f_i,$$

$$(d/dz)^2 f_i = (d/dz)(\frac{d\xi}{dz}) \cdot (d/d\xi) f_i + (\frac{d\xi}{dz})^2 \cdot (d/d\xi)^2 f_i.$$

Let $m \geqslant 3$ and assume our assertion is true for $m-1$, i.e.,

$$(d/dz)^{m-1} f_i - (\frac{d\xi}{dz})^{m-1} (d/d\xi)^{m-1} f_i = \sum_{j=1}^{m-2} g_j (d/d\xi)^j f_i,$$

with each $g_j \in K_1$. Put $g_0 = 0$. Then by the chain rule, we have

$$\begin{aligned} (d/dz)^m f_i - (\frac{d\xi}{dz})^m (d/d\xi)^m f_i &= \sum_{j=1}^{m-2} \Big((d/dz) g_j + g_{j-1} \frac{d\xi}{dz} \Big) (d/d\xi)^j f_i \\ &+ \Big((d/dz)(\frac{d\xi}{dz})^{m-1} + g_{m-2} \frac{d\xi}{dz} \Big) (d/d\xi)^{m-1} f_i. \end{aligned}$$

This completes the induction and hence the proof. □

Lemma 2.2.5. *For every $\nu \in M_{K_1}$, we have*

$$\nu(W) + \binom{n}{2}\nu(\frac{dz}{d\xi_\nu}) \geqslant n \min_{1\leqslant i\leqslant n} \nu(f_i),$$

where ξ_ν is a local parameter of ν.

Proof. Let $\nu \in M_{K_1}$ and a local parameter ξ_ν of ν. For k-linearly dependent $f_1, \dots, f_n$ we have $W = 0$ and the assertion is clear. Assume that $f_1, \dots, f_n$ are k-linearly independent. Let $m = -\min_{1\leqslant i\leqslant n} \nu(f_i), l = -\nu(\frac{d\xi_\nu}{dz})$, and put $g_i = f_i\xi_\nu^m$. By Lemma 2.2.4 we have

$$\begin{aligned} W &= W_z(f_1, \dots, f_n) \\ &= \xi_\nu^{-nm} \cdot W_z(g_1, \dots, g_n) \\ &= \xi_\nu^{-nm} \cdot \det\left((\frac{d\xi_\nu}{dz})^{j-1} \cdot (d/d\xi_\nu)^{j-1} g_i \right)_{1\leqslant i,j\leqslant n} \\ &= \xi_\nu^{-nm} \cdot \xi_\nu^{-l\binom{n}{2}} \cdot \det\left(\xi_\nu^{l(j-1)} (\frac{d\xi_\nu}{dz})^{j-1} \cdot (d/d\xi_\nu)^{j-1} g_i \right)_{1\leqslant i,j\leqslant n}. \end{aligned}$$

Since $\nu(\xi_\nu^{l(j-1)}(\frac{d\xi_\nu}{dz})^{j-1}) = 0$ and $\nu(g_i) = \nu(f_i)+m \geqslant 0$, we have $\nu((d/d\xi_\nu)^{j-1}g_i) \geqslant 0$, hence $\nu(W) \geqslant -nm - l\binom{n}{2}$, as claimed. □

Lemma 2.2.6. *Let $f_1, \dots, f_n$ be k-linearly independent elements of K_1. Then for every $\nu \in M_{K_1}$, we have*

$$\nu(W) + \binom{n}{2}\nu(\frac{dz}{d\xi_\nu}) \geqslant -\binom{n}{2} + \sum_{i=1}^{n} \nu(f_i).$$

Proof. See [26] or [6]. □

Lemma 2.2.7. *Let $f_1, \dots, f_n$ be k-linearly independent elements of K_1 and $b = \sum_{i=1}^{n} f_i$. Then*

$$C(n)|b|_{S_1} \prod_{i=1}^{n} |f_i|_{S_1} \Big(\prod_{\nu\notin S_1} \max_i(|f_i|_\nu) \Big)^n \geqslant H_{S_1}(f_1, \dots, f_n).$$

Proof. Let $\nu \in S_1$, choose $j(\nu) \in \{1, \ldots, n\}$ such that $\nu(f_{j(\nu)}) = \min_i(\nu(f_i))$. Then $W_z(f_1, \ldots, f_n)$ does not change if we replace $f_{j(\nu)}$ by b. Applying Lemma 2.2.6, we get

$$\nu(W) + \binom{n}{2}\nu(\frac{dz}{d\xi_\nu}) + \binom{n}{2} \geqslant \Big(\sum_{i=1}^{n} \nu(f_i)\Big) + \nu(b) - \min_{1\leqslant i\leqslant n}(\nu(f_i)).$$

Now let $\nu \notin S_1$. Then by Lemma 2.2.5, we get

$$\nu(W) + \binom{n}{2}\nu(\frac{dz}{d\xi_\nu}) \geqslant n \min_{1\leqslant i\leqslant n} \nu(f_i).$$

Taking the sum over all $\nu \in M_{K_1}$, and noticing that $\sum\limits_{\nu\in M_{K_1}} \nu(W) = 0$, $\sum\limits_{\nu\in M_{K_1}} \nu(\frac{dz}{d\xi_\nu}) = 2g_{K_1}$, we deduce that

$$\binom{n}{2}(2g_{K_1}-2)+\binom{n}{2}\#S_1 \geqslant \sum_{\nu\in S_1}\sum_{i=1}^{n}\nu(f_i)+\sum_{\nu\in S_1}\nu(b)-\sum_{\nu\in S_1}\min_i \nu(f_i)+n\sum_{\nu\notin S_1}\min_i \nu(f_i).$$

Hence

$$C(n) \geqslant \Big(|b|_{S_1}\prod_{i=1}^{n}|f_i|_{S_1}\Big)^{-1}\prod_{\nu\in S_1}\max_i(|f_i|_\nu)\Big(\prod_{\nu\notin S_1}\max_i(|f_i|_\nu)\Big)^{-n},$$

or equivalently,

$$C(n)|b|_{S_1}\prod_{i=1}^{n}|f_i|_{S_1}\Big(\prod_{\nu\notin S_1}\max_i(|f_i|_\nu)\Big)^{n} \geqslant H_{S_1}(f_1, \ldots, f_n).$$

□

Lemma 2.2.8. *Let S_1, T be as in Corollary 2.2.3. Let $x_1, \ldots, x_n \in K_1$ be such that for each non-empty subset $I \subset \{1, \ldots, n\}$, we have $\sum\limits_{i\in I} x_i \neq 0$. Then*

$$\prod_{i=1}^{n}|x_i|_{S_1}|x_1 + \cdots + x_n|_T\Big(\prod_{\nu\notin S_1}\max_i(|x_i|_\nu)\Big)^{n} \geqslant C'(n)^{-1}H_T(x_1, \ldots, x_n),$$

with $T \subset S_1$.

This is a slight generalization of Corollary 2.2.3.

Proof. Observe that $C'(h)C'(l) \leqslant C'(h+l)$ for $h, l \geqslant 0$.

We proceed by induction on n. For $n = 1$, the assertion follows trivially from the product formula. Let $N \geqslant 2$ and assume the assertion is true for $n < N$. We prove the assertion for $n = N$. We may again assume that $T = S_1$ without loss of generality since each $\nu \in S_1$ is non-archimedean, i.e., $|x_1 + \cdots + x_n|_\nu \leqslant \max_{1\leqslant i\leqslant n} |x_i|_\nu$.

First suppose that $x_1, \ldots, x_n$ are k-linearly independent. Then the assertion is true by Lemma 2.2.7.

Now, possibly after rearranging the indices, suppose that $x_1 + \cdots + x_n = a_1x_1 + \cdots + a_ux_u$ with $1 \leqslant u < n$, $a_1, \ldots, a_u \in k^*$ where u with this property has been chosen minimally and no proper subsum of the right-hand side vanishes. Then $x_1, \ldots, x_u$ are k-linearly independent. Partition S_1 into two subsets

$$S^{(1)} = \{\nu \in S_1 : \max_{1\leqslant i\leqslant n} (|x_i|_\nu) = \max_{1\leqslant i\leqslant u} (|x_i|_\nu)\},$$

$$S^{(2)} = \{\nu \in S_1 : \max_{1\leqslant i\leqslant n} (|x_i|_\nu) > \max_{1\leqslant i\leqslant u} (|x_i|_\nu)\}.$$

Then we have $x_{u+1} + \cdots + x_n = (a_1 - 1)x_1 + \cdots + (a_u - 1)x_u$ and hence $|x_{u+1} + \cdots + x_n|_\nu \leqslant \max_{1\leqslant i\leqslant u} (|x_i|_\nu)$ for $\nu \in S^{(2)}$. Combining this with the induction hypothesis, we derive that

$$
\begin{aligned}
&\prod_{i=1}^{n}|x_i|_{S_1}|x_1+\cdots+x_n|_{S_1}\\
= \ &\Big(\prod_{i=1}^{u}|a_ix_i|_{S_1}|a_1x_1+\cdots+a_ux_u|_{S_1}\Big)\prod_{i=u+1}^{n}|x_i|_{S_1}\\
\geqslant \ &C'(u)^{-1}H_{S_1}(a_1x_1,\ldots,a_ux_u)\Big(\prod_{\nu\notin S_1}\max_{1\leqslant i\leqslant u}(|a_ix_i|_\nu)\Big)^{-u}\prod_{i=u+1}^{n}|x_i|_{S_1}\\
= \ &C'(u)^{-1}H_{S^{(1)}}(x_1,\ldots,x_u)H_{S^{(2)}}(x_1,\ldots,x_u)\prod_{i=u+1}^{n}|x_i|_{S_1}\Big(\prod_{\nu\notin S_1}\max_{1\leqslant i\leqslant u}(|x_i|_\nu)\Big)^{-u}\\
\geqslant \ &C'(u)^{-1}H_{S^{(1)}}(x_1,\ldots,x_n)\Big(|x_{u+1}+\cdots+x_n|_{S^{(2)}}\prod_{i=u+1}^{n}|x_i|_{S_1}\Big)\Big(\prod_{\nu\notin S_1}\max_{1\leqslant i\leqslant u}(|x_i|_\nu)\Big)^{-u}\\
\geqslant \ &C'(u)^{-1}H_{S^{(1)}}(x_1,\ldots,x_n)C'(n-u)^{-1}H_{S^{(2)}}(x_{u+1},\ldots,u_n)\\
&\quad\times\Big(\prod_{\nu\notin S_1}\max_{i>u}(|x_i|_\nu)\Big)^{u-n}\Big(\prod_{\nu\notin S_1}\max_{1\leqslant i\leqslant u}(|x_i|_\nu)\Big)^{-u}\\
\geqslant \ &C'(n)^{-1}H_{S_1}(x_1,\ldots,x_n)\Big(\prod_{\nu\notin S_1}\max_{1\leqslant i\leqslant n}|x_i|_\nu\Big)^{-n},
\end{aligned}
$$

which completes the induction step. □

Theorem 2.2.9. *Let S_1, T be as above. Let $x_1,\ldots,x_n\in K_1$ be such that $\sum_{i\in I}x_i\neq 0$ for any non-empty subset $I\subset\{1,\ldots,n\}$. Then*

$$\prod_{i=1}^{n}|x_i|_{S_1}\cdot|x_1+\cdots+x_n|_T\cdot\Big(\prod_{\nu\notin S_1}\max_i(|x_i|_\nu)\Big)^n\geqslant C'(u)^{-1}H_T(x_1,\ldots,x_n),$$

where $u=rank_k\{x_1,\ldots,x_n\}$.

This result improves Lemma 2.2.8 and is inspired by an idea of Zannier [26].

Proof. Recall that $C'(s) = e^{\frac{s(s-1)}{2}\max(2g_{K_1}-2+\#S_1,0)}$ for $s \in \mathbb{N}$. Then $C'(s)C'(t) \leqslant C'(s+t)$.

First notice that the special case $u = n$ is just Lemma 2.2.7.

For the general case we proceed by induction on n, the case $n = 1$ being trivial. Let $N \geqslant 2$ and assume the assertion is true for all $n < N$, now consider the case $n = N$. Like in the proof of Lemma 2.2.8, we only have to consider the special case $T = S_1$. Let $x_1, \dots, x_u$ be k-linearly independent with u maximal and assume, renumbering indices if necessary, $x_1 + \dots + x_N = \sum_{i=1}^{v} a_i x_i$ with $a_1, \dots, a_v \in k^*$ and $1 \leqslant v \leqslant u$.

First assume $v = u$. Then each x_i is a k-linear combination of $x_1, \dots, x_u$, hence

$$\max_{1\leqslant i\leqslant N}(|x_i|_\nu) = \max_{1\leqslant i\leqslant u}(|x_i|_\nu). \tag{2.2.1}$$

Then by applying Lemma 2.2.8 to $a_1x_1, \dots, a_ux_u$, $|\cdot|_\nu$ and using that $|\cdot|_\nu$ is trivial on k^* for $\nu \in M_{K_1}$, we get

$$\prod_{i=1}^{u} |x_i|_{S_1} |x_1 + \dots + x_N|_{S_1} \Big(\prod_{\nu\notin S_1} \max_{1\leqslant i\leqslant u}(|x_i|_\nu)\Big)^u \geqslant c(u)^{-1} H_{S_1}(x_1, \dots, x_u). \tag{2.2.2}$$

Clearly, $\max_{1\leqslant i\leqslant N}(|x_i|_\nu) \geqslant |x_i|_\nu$ for $i > u, \nu \notin S_1$, so

$$\Big(\prod_{\nu\notin S_1} \max_{1\leqslant i\leqslant N}(|x_i|_\nu)\Big)^{N-u} \geqslant \prod_{i>u}\prod_{\nu\notin S_1} |x_i|_\nu$$

and hence by the product formula,

$$\prod_{i>u} |x_i|_{S_1} \Big(\prod_{\nu\notin S_1} \max_{1\leqslant i\leqslant n}(|x_i|_\nu)\Big)^{N-u} \geqslant 1. \tag{2.2.3}$$

Combining (2.2.2) with (2.2.3) we derive the assertion when $v = u$.

So we assume $1 \leqslant v < u < N$. Applying again Lemma 2.2.8 to

$a_1x_1, \ldots, a_vx_v$ we get

$$\prod_{i=1}^{v} |x_i|_{S_1} |x_1 + \cdots + x_N|_{S_1} \Big(\prod_{\nu \notin S_1} \max_{1 \leqslant i \leqslant v} (|x_i|_\nu) \Big)^v \geqslant c(v)^{-1} H_{S_1}(x_1, \ldots, x_v). \tag{2.2.4}$$

We claim that there exists $h < N$ such that there are two finite sequences $\{u_l\}, \{j_l\}$ of integers of the same length h satisfying the following:

(i) $u_0 = v, u_h = u, u_l > u_{l-1}$ for $1 \leqslant l \leqslant h$,

(ii) $j_0 = 0$, $u < j_l \leqslant N$ for $l \geqslant 1$, $j_s \neq j_t$ for $s \neq t$,

(iii) there is a renumbering of the indices $v+1, \ldots, u$ such that for $l > 0$ x_{j_l} is a k-linear combination of $x_1, \ldots, x_{u_l}$ and for all $l \geqslant 0$

$$\prod_{i=0}^{l} |x_{j_i}|_{S_1} \prod_{i=1}^{u_l} |x_i|_{S_1} |x_1 + \cdots + x_N|_{S_1} \Big(\prod_{\nu \notin S_1} \max_{i \in A_l} (|x_i|_\nu) \Big)^{u_l + l} \geqslant c(u_l)^{-1} H_{S_1}(x_1, \ldots, x_{u_l}), \tag{2.2.5}$$

where we put $|x_0|_{S_1} = 1$ and $A_l = \{1, \ldots, u_l\} \cup \{j_1, \ldots, j_l\}$.

Then this construction will finish the proof in the end.

We prove this claim by induction on l. The first step when $l = 0$ is just (2.2.4). Let $r \geqslant 0$ and assume that $u_0, \ldots, u_r, j_0, \ldots, j_r$ have been constructed such that $u_r < u$ and (2.2.5) holds for $l = 0, \ldots, r$. We show the existence of u_{r+1}, j_{r+1} such that (2.2.5) holds for $l = r+1$. For any index $0 < j \leqslant N$ we have

$$x_j = \sum_{i=1}^{u} \lambda_{i,j} x_i = \sum_{i=1}^{u_r} \lambda_{i,j} x_i + \sum_{i=u_r+1}^{u} \lambda_{i,j} x_i := T_{j,r} + U_{j,r},$$

with $\lambda_{i,j} \in k$ uniquely determined.

We claim that there is j such that both $T_{j,r}$ and $U_{j,r}$ are non-zero. Assume the contrary, then for each $j \in \{1, \ldots, N\}$, either $U_{j,r} = 0$ i.e.,

$x_j = T_j$, or $T_{j,r} = 0$, that is, $x_j = U_{j,r}$. Since $x_1 + \cdots + x_N = \sum_{i=1}^{v} a_i x_i$ and $v < u_r + 1$, we derive that $\sum_{j=1}^{N} U_{j,r} = 0$, or equivalently, $\sum_{U_{j,r} \neq 0} x_j = 0$. But $u_r < u$, so we have $U_{u,r} \neq 0$, and hence $\{j : U_{j,r} \neq 0\} \neq \emptyset$. This gives a vanishing subsum, which contradicts the assumption.

Let j be the smallest index with $U_{j,r} \neq 0, T_{j,r} \neq 0$ and put $j_{r+1} = j$. Then clearly $j > u$ because $x_1, \ldots, x_u$ are k-linearly independent with u maximal. Renumbering the indices $u_r + 1, \ldots, u$, we can write

$$U_{j_{r+1},r} = \sum_{i=u_r+1}^{u_{r+1}} \lambda_{i,j_{r+1}} x_i, \tag{2.2.6}$$

where $\lambda_{i,j_{r+1}} \neq 0$ for $u_r + 1 \leqslant i \leqslant u_{r+1}$. This defines u_{r+1} satisfying $u_r < u_{r+1} \leqslant u$ and gives $x_{j_{r+1}}$ a linear combination of $x_1, \ldots, x_{u_{r+1}}$. Since for $l < r + 1$, $u_l \leqslant u_r < u_{r+1}$ and x_{j_l} is a linear combination of $x_1, \ldots, x_{u_l}$, we infer that $j_{r+1} \neq j_l$.

Put $B_l = A_{l+1} \backslash A_l = \{j_{l+1}\} \cup \{u_l + 1, \ldots, u_{l+1}\}$. The assumption in (iii) for u_r, j_r gives

$$\left(\prod_{i=0}^{r} |x_{j_i}|_{S_1}\right) \left(\prod_{i=1}^{u_r} |x_i|_{S_1}\right) |x_1 + \cdots + x_N|_{S_1} \left(\prod_{\nu \notin S_1} \max_{i \in A_r}(|x_i|_\nu)\right)^{u_r + r} \geqslant c(u_r)^{-1} H_{S_1}(x_i : i \in A_r). \tag{2.2.7}$$

Notice that $T_{j_{r+1}} = \sum_{i=1}^{u_r} \lambda_{i,j_{r+1}} x_i = x_{j_{r+1}} - \sum_{i=u_r+1}^{u_{r+1}} \lambda_{i,j_{r+1}} x_i$ as a sum of $x_{j_{r+1}}$ and $-\lambda_{i,j_{r+1}} x_i$, $u_r + 1 \leqslant i \leqslant u_{r+1}$, the assumption of Lemma 2.2.8 is satisfied and the components are indeed k-linearly independent, since $T_{j_{r+1}} \neq 0$, $x_1, \ldots, x_u$ are k-linearly independent with u maximal and $\lambda_{i,j_{r+1}} \neq 0$ for $u_r + 1 \leqslant i \leqslant u_{r+1}$. By Lemma 2.2.8, we obtain

$$\prod_{i \in B_r} |x_i|_{S_1} |T_{j_{r+1}}|_{S_1} \left(\prod_{\nu \notin S_1} \max_{i \in B_r}(|x_i|_\nu)\right)^{u_{r+1} - u_r + 1} \geqslant c(u_{r+1} - u_r + 1)^{-1} H_{S_1}(x_i : i \in B_r). \tag{2.2.8}$$

Combining this with (2.2.7), we get

$$\prod_{i=0}^{r+1}|x_{j_i}|_{S_1}\prod_{i=1}^{u_{r+1}}|x_i|_{S_1}|T_{j_{r+1}}|_{S_1}|x_1+\cdots+x_N|_{S_1}\times \tag{2.2.9}$$
$$\times\Big(\prod_{\nu\notin S_1}\max_{i\in A_r}(|x_i|_\nu)\Big)^{u_r+r}\Big(\prod_{i\in B_r}\max_{i\in B_r}(|x_i|_\nu)\Big)^{u_{r+1}-u_r+1}$$
$$\geqslant C'(u_r)^{-1}C'(u_{h+1}-u_r+1)^{-1}H_{S_1}(x_i:i\in A_r)H_{S_1}(x_i:i\in B_r).$$

Noticing that for any $\nu\in M_{K_1}$,

$$|T_{j_{r+1}}|_\nu\leqslant\min(\max(|x_i|_\nu:i\in A_r),\max(|x_i|_\nu:i\in B_r)),$$

$$\max(|x_i|_\nu:i=1,\dots,u_{r+1})\leqslant\max(\max(|x_i|_\nu:i\in A_r),\max(|x_i|_\nu:i\in B_r)),$$

we deduce that for $\nu\in M_{K_1}$

$$|T_{j_{r+1}}|_\nu\max(|x_i|_\nu:i=1,\dots,u_{r+1})\leqslant\max(|x_i|_\nu:i\in A_r)\max(|x_i|_\nu:i\in B_r)).$$

Taking the product over $\nu\in S_1$, and inserting (2.2.9), we obtain

$$\prod_{i=0}^{r+1}|x_{j_i}|_{S_1}\prod_{i=1}^{u_{r+1}}|x_i|_{S_1}\cdot|x_1+\cdots+x_N|_{S_1}\times$$
$$\times\Big(\prod_{\nu\notin S_1}\max_{i\in A_r}(|x_i|_\nu)\Big)^{u_r+r}\Big(\prod_{i\in B_r}\max_{i\in B_r}(|x_i|_\nu)\Big)^{u_{r+1}-u_r+1}$$
$$\geqslant C'(u_r)^{-1}C'(u_{r+1}-u_r+1)^{-1}H_{S_1}(x_1,\dots,x_{u_{r+1}}).$$

Observing that $C'(x)C'(y+1)\leqslant C'(x+y)$ for $x,y\geqslant 1$ and $A_l\cup B_l=A_{l+1}$, we get

$$\prod_{i=0}^{r+1}|x_{j_i}|_{S_1}\prod_{i=1}^{u_{r+1}}|x_i|_{S_1}|x_1+\cdots+x_N|_{S_1}\Big(\prod_{\nu\notin S_1}\max_{i\in A_{r+1}}(|x_i|_\nu)\Big)^{u_{r+1}+r+1}$$
$$\geqslant C'(u_{r+1})^{-1}H_{S_1}(x_1,\dots,x_{u_{r+1}}).$$

This verifies (iii) for $r+1$ in place of r (in case $u_r<u$), and completes the proof of the claim.

Now for u_h, j_h, since $x_1, \ldots, x_u$ are k-linearly independent with u maximal, we have

$$\prod_{i=0}^{h} |x_{j_i}|_{S_1} \prod_{i=1}^{N} |x_i|_{S_1} |x_1 + \cdots + x_N|_{S_1} \Big(\prod_{\nu \notin S_1} \max_{1 \leqslant i \leqslant u} (|x_i|_\nu) \Big)^{u+h} \geqslant C'(u)^{-1} H_{S_1}(x_1, \ldots, x_N). \tag{2.2.10}$$

Clearly we have

$$\prod_{\nu \notin S_1} \max_{1 \leqslant i \leqslant N} (|x_i|_\nu) \geqslant \prod_{\nu \notin S_1} |x_i|_\nu \text{ for } i \notin C := \{1, \ldots, u, j_1, \ldots, j_h\},$$

so

$$\Big(\prod_{\nu \notin S_1} \max_{1 \leqslant i \leqslant N} (|x_i|_\nu) \Big)^{N-u-h} \geqslant \prod_{\nu \notin S_1} \prod_{i \in C} |x_i|_\nu = \prod_{i \in C} |x_i|_{S_1}^{-1}. \tag{2.2.11}$$

Combining (2.2.10) and (2.2.11) completes the proof. □

Example 2.2.10. *Let x_1, x_2, x_3, x_4, x_5 be k-linearly independent, $x_6 = -x_4 - x_5 + 2x_3$. Then $n = 6, u = 5, v = 3, j_1 = 6, u_1 = u$.*

Let x_1, x_2, x_3, x_4 be k-linearly independent, $x_5 = 2x_3 - x_4, x_6 = -3x_3 + x_2$. Then $n = 6, u = 4, v = 2, j_1 = 6, u_1 = 3, j_2 = 5, u_2 = u$.

Let $x_1, \ldots, x_5$ be k-linearly independent, $x_6 = -x_3 + x_2 - x_4, x_7 = -x_5 + x_1$. Then $n = 7, u = 5, v = 2$, and we get $j_1 = 6, u_1 = 4, j_2 = 7, u_2 = u$, or we reorder x_3, x_4, x_5 by x_5', x_4', x_3', then we get $j_1 = 7, u_1 = 3, j_2 = 6, u_2 = u$.

Corollary 2.2.11. *Let $n \geqslant 3$. Assume $x_1, \ldots, x_n \in K_1$ and $\sum_{i=1}^{n} x_i = 0$ but no non-empty proper subsum vanishes. Then*

$$H_{S_1}(x_1, \ldots, x_n) \leqslant e^{\binom{n-1}{2} \max(2g_{K_1} - 2 + \#S_1, 0)} \left(\prod_{i=1}^{n} |x_i|_{S_1} \right) \left(\prod_{\nu \notin S_1} \max_i (|x_i|_\nu) \right)^{n-1}.$$

If $x_1, \ldots, x_n$ are k-linearly independent, then we can replace $\max(2g_{K_1} - 2 + \#S_1, 0)$ *by* $2g_{K_1} - 2 + \#S_1$.

Proof. Simply apply Theorem 2.2.9 for $x_1, \ldots, x_{n-1}$. □

Corollary 2.2.12. *Let $n \geqslant 3$. Assume $x_1, \ldots, x_n \in \mathcal{O}_{S_1}$ and $\sum_{i=1}^{n} x_i = 0$ but no non-empty proper subsum vanishes. Then*

$$H_{S_1}(x_1, \ldots, x_n) \leqslant e^{\binom{n-1}{2} \max(2g_{K_1}-2+\#S_1, 0)} \prod_{i=1}^{n} |x_i|_{S_1}.$$

Proof. This is a direct consequence since for $x \in \mathcal{O}_{S_1}, |x|_{S_1} \leqslant 1$. □

For $n = 4$, the constant $\frac{(n-1)(n-2)}{2}$ is best possible, as is shown by the following example from [5].

Example 2.2.13. *Let $K = k(t)$, and $x_1 = (t^r+1)^3, x_2 = -t^{3r}, x_3 = -3t^r(t^r+1), x_4 = -1$ where r is a positive integer. Take S to be the set of valuations corresponding to ∞ and the prime factors of $t(t^r+1)$. Then $\#S = r+2$, x_i $(i = 1, 2, 3, 4)$ are S-units, $H_S(x_1, x_2, x_3, x_4) = e^{3r} = e^{3(2g_K-2+\#S)}$. This implies that for $n = 4$ the constant $\frac{(n-1)(n-2)}{2}$ is best possible.*

Remark 2.2.14. *Corollary 2.2.12 is much stronger than its analogue over number fields, i.e., Lemma 2. Lemma 2 first involves an exponent $1+\varepsilon$ on $\prod_{i=1}^{n} |x_i|_{S_1}$ and second an ineffective constant $C(n, S_1, \varepsilon)$, which is caused by the ineffectivity of the Subspace Theorem. We also notice the improvement in comparison with the result of Corollary 2.1.4, with a much sharper exponent 1 instead of $\frac{(n-1)(n-2)}{2}$.*

Theorem 2.2.9 and its Corollary 2.2.11 imply the following results:

Lemma 2.2.15 (Mason, Stothers)**.** *Let L be a finite extension of $K = k(t)$, and T a finite set of valuations of L. Let $\gamma_1, \gamma_2, \gamma_3$ be non-zero elements of L satisfying $\gamma_1 + \gamma_2 + \gamma_3 = 0$ and $\nu(\gamma_1) = \nu(\gamma_2) = \nu(\gamma_3)$ for every valuation $\nu \notin T$. Then either $\frac{\gamma_1}{\gamma_2} \in k$, which means $H^*(\frac{\gamma_1}{\gamma_2}) = 1$, or*

$$H^*(\frac{\gamma_1}{\gamma_2}) \leqslant e^{(\#T+2g_L-2)/[L:K]}.$$

In particular, let $a(t), b(t), c(t)$ be coprime polynomials over k such that $a(t) + b(t) = c(t)$ and not all of them are constants. Then

$$\max(\deg a(t), \deg b(t), \deg c(t)) \leqslant \deg(\mathrm{rad}(abc)) - 1,$$

where rad(f) *denotes the product of the distinct prime factors of f.*

Proof. Assume that γ_1, γ_2 are k-linearly independent. Apply Corollary 2.2.11 with $n = 3$, $K_1 = L$, $S_1 = T$ and $x_i = \gamma_i$ $(i = 1, 2, 3)$ and apply the product formula. For the particular case that $a(t), b(t), c(t)$ are polynomials from $k[t]$ without a common factor, let S_1 be the set of valuations consisting of ν_∞ and those corresponding to the zeros of abc. Then $\#S_1 = \deg(\mathrm{rad}(abc)) - 1$ and thus our assertion follows directly from Corollary 2.2.11. □

Theorem 2.2.16 (Brownawell, Masser)**.** *Assume $u_1, \ldots, u_n$ are S_1-units satisfying $u_1 + \cdots + u_n = 0$ but no non-empty proper subsum vanishes. Then*

$$H_{K_1}(u_1, \ldots, u_n) \leqslant \exp\left(\binom{n-1}{2} \max(\#S + 2g_{K_1} - 2, 0)\right).$$

This is mentioned after Theorem B of [6].

Proof. Apply Corollary 2.2.11 by taking $T = S_1$ and noticing that for an S_1-unit x, we have $|x|_{S_1} = 1$ and $|x|_\nu = 1$ for every $\nu \notin S_1$. □

Theorem 2.2.17 (Zannier)**.** *Let $a_1, \ldots, a_n \in K_1$ be S_1-units such that $\sum_{i\in\Gamma} a_i \neq 0$ for every nonempty $\Gamma \subset \{1, \ldots, n\}$. Put $b = a_1 + \cdots + a_n$. Then*

$$\sum_{\nu\in S_1} (\nu(b) - \min \nu(a_i)) \leqslant \binom{\mu}{2} \max(\#S_1 + 2g_{K_1} - 2, 0)$$

where $\mu = \mathrm{rank}\{a_1, \ldots, a_n\}$.

This is Theorem 1 of [26], except that there the result was stated $\#S_1 + 2g_{K_1} - 2$ instead of $\max(\#S_1 + 2g_{K_1} - 2, 0)$. However, the proof in [26] gives only the inequality with the maximum with 0.

Proof. This follows directly by taking $T = S_1, x_i = a_i\,(i = 1, \ldots, n)$ in Theorem 2.2.9 and using the fact that $a_1, \ldots, a_n$ are S_1-units. $\square$

Theorem 2.2.18 (Davenport). *If $f(t), g(t)$ are nonzero polynomials over k such that $g(t)^2 \neq f(t)^3$, then*

$$\deg(g(t)^2 - f(t)^3) \geqslant \frac{1}{2}\deg f(t) + 1.$$

Proof. This is an analogue of Hall's conjecture over the function fields. It is first proved by Davenport in [8].

In Corollary 2.2.11, let $T = S_1$ be the set S consisting of ν_∞ and the valuations corresponding to the zeros of fg, and $x_1 = f(t)^3, x_2 = -g(t)^2, x_3 = g(t)^2 - f(t)^3$. Then

$$H_S(f(t)^3, g(t)^2) \leqslant e^{\#S-2}|g(t)^2 - f(t)^3|_S.$$

In particular, when $f(t), g(t)$ are coprime, we deduce that

$$\begin{aligned}\frac{1}{2}(3\deg f(t) + 2\deg g(t)) &\leqslant \max(\deg f(t)^3, \deg g(t)^2)\\ &\leqslant \deg(\mathrm{rad}(fg)) - 1 + \deg(g(t)^2 - f(t)^3)\\ &\leqslant \deg f(t) + \deg g(t) + \deg(g(t)^2 - f(t)^3) - 1.\end{aligned}$$

Hence

$$1 + \frac{1}{2}\deg f(t) \leqslant \deg(g(t)^2 - f(t)^3).$$

The case when $f(t), g(t)$ are not coprime is a direct consequence of the above. $\square$

Chapter 3

Geometry of numbers over function fields

Minkowski's results on successive minima of convex bodies have analogues over function fields. These are discussed in this chapter. Our main reference is Thunder [24].

3.1 Successive minima

Recall $K = k(t)$ is the rational function field over an algebraically closed field k of characteristic 0 and for $\nu \in M_K$, $R_\nu = \{x \in K_\nu : |x|_\nu \leqslant 1\}$. A subset $\mathcal{C}_\nu$ of K_ν^n is called a ν-adic convex symmetric body if it has the following properties:

- $\mathcal{C}_\nu$ is closed and bounded in the topology of K_ν^n induced by $|\cdot|_\nu$ and has $\mathbf{0}$ as an interior point;
- for every $\mathbf{x} \in \mathcal{C}_\nu, \alpha \in K_\nu$ with $|\alpha|_\nu \leqslant 1$, we have $\alpha\mathbf{x} \in \mathcal{C}_\nu$;
- for every $\mathbf{x}, \mathbf{y} \in \mathcal{C}_\nu$, we have $\mathbf{x} + \mathbf{y} \in \mathcal{C}_\nu$.

Remark 3.1.1. *These properties imply that $\mathcal{C}_\nu$ is an R_ν-module.*

Let S be a finite subset of M_K containing ν_∞ and $\mathbb{A}_S = \prod_{\nu\in S} K_\nu$ the ring of S-adeles. Consider K as a subring of $\mathbb{A}_S$ by identifying $x \in K$ with the adele $(x_\nu)_{\nu\in S}$ with $x_\nu = x$ for all $\nu \in S$. A subset $\mathcal{C}$ of $\mathbb{A}_S^n$ is called convex symmetric if $\mathcal{C} = \prod_{\nu\in S} \mathcal{C}_\nu$ with $\mathcal{C}_\nu$ ν-adic convex symmetric for $\nu \in S$. We need two lemmas.

Lemma 3.1.2. *Let $\nu \in M_K$ and $\mathcal{C}_\nu \subset K_\nu^n$. Then $\mathcal{C}_\nu$ is a ν-adic symmetric convex body if and only if $\mathcal{C}_\nu = \{\mathbf{x} \in K_\nu^n : \|A_\nu \mathbf{x}\|_\nu \leqslant 1\}$ for some $A_\nu \in GL_n(K_\nu)$.*

Proof. First, notice that from the definition, $\mathcal{C}_\nu$ is an R_ν-module. It is also bounded, hence there is a constant $C > 0$ such that $\|\mathbf{x}\|_\nu \leqslant C$ for every $\mathbf{x} \in \mathcal{C}_\nu$.

It is easy to see that $\mathcal{C}_\nu$ contains a basis of K_ν^n, since $\mathbf{0}$ is an interior point. Hence there exists a free R_ν-module $\mathcal{M}_1$ of rank n such that $\mathcal{M}_1 \subset \mathcal{C}_\nu$. Take $\alpha \in K_\nu$ with $|\alpha|_\nu \geqslant C$. Then $\mathcal{C}_\nu \subset \{\mathbf{x} \in K_\nu^n : \|\mathbf{x}\|_\nu \leqslant |\alpha|_\nu\}$. Hence $\mathcal{C}_\nu$ is contained in the free R_ν-module $\mathcal{M}_2 = \alpha R_\nu^n$ of rank n.

As is well-known, R_ν is a principal ideal domain, hence by Chapter III, Theorem 7.1 of [16], we know that $\mathcal{C}_\nu$ is also a free R_ν-module of rank n. Take an R_ν-basis of $\mathcal{C}_\nu$, let B_ν be the matrix whose columns consists of this basis, and let $A_\nu = B_\nu^{-1}$. Then $\mathcal{C}_\nu = \{B_\nu \mathbf{y} : \mathbf{y} \in R_\nu^n\} = \{\mathbf{x} \in K_\nu^n : \|A_\nu \mathbf{x}\|_\nu \leqslant 1\}$. □

Remark 3.1.3. *For $\mathcal{C}_\nu = R_\nu$ we will choose A_ν to be I_n, the $n \times n$ identity matrix. This does not change $\mathcal{C}_\nu$.*

Example 3.1.4. *Take ν be the valuation corresponding to 0. Then $K_\nu = k((t)), R_\nu = \mathbb{C}[[t]]$. Let $\mathcal{C}_\nu = \{\mathbf{x} \in k((t))^n : \|\mathbf{x}\|_\nu < 1\}$. Then $\mathcal{C}_\nu = \{(x_1, \ldots, x_n) \in \mathcal{C}((t))^n : x_i \in tk[[t]], i = 1, \ldots, n\}$. We may take $A_\nu = diag(\frac{1}{t}, \ldots, \frac{1}{t})$ and this gives $\mathcal{C}_\nu = \{\mathbf{x} \in k((t))^n : \|A_\nu \mathbf{x}\|_\nu \leqslant 1\}$.*

Lemma 3.1.5. *For $\mathbf{x} \in K^n \backslash \{\mathbf{0}\}$ there exists $f \in K$ such that $\|A_\infty(f\mathbf{x})\|_\infty = H_A(\mathbf{x})$, $\|A_\nu(f\mathbf{x})\|_\nu = 1$ for $\nu \in S, \nu \neq \nu_\infty$ and $\|f\mathbf{x}\|_\nu = 1$ for $\nu \notin S$.*

Proof. For consistency put $A_\nu = I_n$ the $n \times n$ identity matrix for $\nu \notin S$. Let $\{\nu_1, \dots, \nu_m\} \subset M_K \backslash \{\nu_\infty\}$ be the finite set of valuations such that $\|A_\nu \mathbf{x}\|_\nu \neq 1$, with corresponding uniformizers $t - p_1, \dots, t - p_m \in K$. Let $n_i = -\nu_i(A_{\nu_i}\mathbf{x})$ for $1 \leqslant i \leqslant m$ and $f = \prod_{i=1}^{m} (t - p_i)^{n_i}$. Then $\|A_\infty(f\mathbf{x})\|_\infty = H_A(\mathbf{x})$, $\|A_{\nu_i}(f\mathbf{x})\|_{\nu_i} = 1$ for $i = 1, \dots, m$ and $\|A_{\nu_i}(f\mathbf{x})\|_{\nu_i} = 1$ for $\nu \notin \{\nu_1, \dots, \nu_m, \nu_\infty\}$, as claimed. □

Let $\mathcal{C} \subset \mathbb{A}_S^n$ be an S-convex symmetric body and $\lambda \in e^{\mathbb{Z}}$. By Lemma 3.1.2, there exists $A \in \mathrm{GL}_n(\mathbb{A}_S)$, such that $\mathcal{C}_\nu = \{\mathbf{x} \in K_\nu^n : \|A_\nu \mathbf{x}\|_\nu \leqslant 1\}$ for each $\nu \in S$. We view $\mathcal{O}_S^n$ as a subset of $\prod_{\nu \in S} K_\nu^n$ via the diagonal embedding.

For $\lambda \in e^{\mathbb{Z}}$, define

$$\lambda\mathcal{C}_\infty := \{a\mathbf{x} : \mathbf{x} \in \mathcal{C}_\infty, a \in K_\infty, |a|_\infty \leqslant \lambda\}, \tag{3.1.1}$$

$$\lambda\mathcal{C} := (\lambda\mathcal{C}_\infty) \times \prod_{\nu \in S, \nu \neq \nu_\infty} \mathcal{C}_\nu. \tag{3.1.2}$$

Then

$$\lambda\mathcal{C}_\infty = \{\mathbf{x} \in K_\infty^n : \|A_\infty \mathbf{x}\|_\infty \leqslant \lambda\},$$

and

$$\lambda\mathcal{C} \cap \mathcal{O}_S^n := \{\mathbf{x} \in \mathcal{O}_S^n : H_A(x) \leqslant \lambda\}.$$

Remark that by Lemma 3.1.5, for every $\mathbf{x} \in \mathcal{O}_S^n$ with $H_A(\mathbf{x}) \leqslant \lambda$, there exists $f \in K$ such that $\|f\mathbf{x}\|_\nu = 1$ for $\nu \neq \nu_\infty$ and $\|A_\infty(f\mathbf{x})\|_\infty = H_A(\mathbf{x}) \leqslant \lambda$. In particular, $f\mathbf{x} \in k[t]^n$.

Definition 3.1.6. *The i-th successive minimum λ_i of $\mathcal{C}$ is the minimum of all $\lambda \in e^{\mathbb{Z}}$ such that $\lambda\mathcal{C} \cap \mathcal{O}_S^n$ contains at least i K-linearly independent points.*

Clearly, given $\lambda \in e^{\mathbb{Z}}$ and $\mathbf{x} \in \lambda\mathcal{C}$, we have $H_A(\mathbf{x}) \leqslant \lambda$.

Theorem 3.1.7. *The successive minima exist and* $0 < \lambda_1 \leqslant \cdots \leqslant \lambda_n < \infty$. *Moreover, there exists a basis* $\mathbf{x}_1, \ldots, \mathbf{x}_n$ *of* K^n *such that* $\mathbf{x}_i \in \lambda_i \mathcal{C} \cap \mathcal{O}_S^n$, *and* $\|A_\infty(\mathbf{x}_i)\|_\infty = H_A(\mathbf{x}_i) = \lambda_i$ *for all* $i = 1, \ldots, n$.

Proof. For every $\mathbf{x} \in K^n \backslash \{0\}$, we have $H_A(\mathbf{x}) \in e^{\mathbb{Z}}$ and also $H_A(\mathbf{x}) \geqslant c_2 > 0$ unless $\mathbf{x} = \mathbf{0}$ by Lemma 1.5.2. Hence there is $\mathbf{x}_1 \in K^n \backslash \{0\}$ such that $H_A(\mathbf{x}_1)$ is minimal. Further, by Lemma 3.1.5 we may choose $\mathbf{x}_1$ such that $\|A_\infty(\mathbf{x}_1)\|_\infty = H_A(\mathbf{x}_1)$. Then automatically, $\lambda_1 = H_A(\mathbf{x}_1)$ is the first successive minimum. Successively, for $j = 1, \ldots, n-1$, we take $\mathbf{x}_{j+1}$ to be a point $\mathbf{x} \in K^n$ such that $\mathbf{x}$ is K-linearly independent of $\mathbf{x}_1, \ldots, \mathbf{x}_j$ and $H_A(\mathbf{x}_{j+1})$ is minimal with this property, and we may also assume that $\|A_\nu(\mathbf{x}_{j+1})\|_\nu = 1$ for $\nu \neq \nu_\infty$ by Lemma 3.1.5. With this choice, $H_A(\mathbf{x}_1) \leqslant \cdots \leqslant H_A(\mathbf{x}_n)$ and $\mathbf{x}_i \in H_A(\mathbf{x}_i)\mathcal{C}$ for $i = 1, \ldots, n$, hence λ_i exists and $\lambda_i \leqslant H_A(\mathbf{x}_i)$ for $i = 1, \ldots, n$. We claim that $\lambda_i = H_A(\mathbf{x}_i)$ and $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are as required. Assume the contrary, let i be the smallest index such that $H_A(\mathbf{x}_i) > \lambda_i$. There are K-linearly independent points $\mathbf{y}_1, \ldots, \mathbf{y}_i$ in $K^n \cap \lambda_i \mathcal{C}$. Clearly, we have $H_A(\mathbf{x}_i) > \lambda_i \geqslant H_A(\mathbf{y}_j)$ for $j = 1, \ldots, i$. So by our choice of $\mathbf{x}_i$, we know that each $\mathbf{y}_j$ is K-linearly dependent of $\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}$, which contradicts the fact that $\mathbf{y}_1, \ldots, \mathbf{y}_i$ are K-linearly independent. This completes the proof of the claim. □

Theorem 3.1.8. *Let* $\mathcal{C}$ *be an* S*-convex body. Then there is an* $\mathcal{O}_S$*-module basis* $\mathbf{a}_1, \ldots, \mathbf{a}_n$ *of* $\mathcal{O}_S^n$ *such that* $\|A_\infty(\mathbf{a}_i)\|_\infty = H_A(\mathbf{a}_i) = \lambda_i$ *for* $i = 1, \ldots, n$. *Also, we have* $\prod_{i=1}^{n} \lambda_i = |\det A|$.

Proof. By Lemma 1.5.4, we can choose a K-basis of column vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ of K^n such that $\prod_{i=1}^{n} H_A(\mathbf{a}_i) = H_A(K^n)$.

Let $\mu_i = H_A(\mathbf{a}_i)$, and assume that $\mu_1 \leqslant \cdots \leqslant \mu_n$ without loss of generality. By Lemma 3.1.5 we may also assume that $\|A_\infty(\mathbf{a}_i)\|_\infty = H_A(\mathbf{a}_i) = \mu_i$, whence $\mathbf{a}_i \in \mathcal{O}_S^n$ for $i = 1, \ldots, n$. Then $H_A(K^n) = \prod_{i=1}^{n} \mu_i$ and $\mathbf{a}_1, \ldots, \mathbf{a}_i \in$

$\mu_i\mathcal{C}$. By the definition of successive minima, we have $\lambda_i \leqslant \mu_i$ for all i, hence $\prod_{i=1}^{n} \lambda_i \leqslant H_A(K^n)$.

On the other hand, by Theorem 3.1.7, we may take a K-basis $\mathbf{x}_1, \ldots, \mathbf{x}_n$ with $\mathbf{x}_i \in \lambda_i\mathcal{C}, H_A(\mathbf{x}_i) = \lambda_i$ for all i. By lemma 1.5.4, we have $\prod_{i=1}^{n} \lambda_i \geqslant H_A(K^n)$. Therefore, we get $\mu_i = \lambda_i$ and $\prod_{i=1}^{n} \lambda_i = H_A(K^n) = |\det A|$.

For $\nu \in S$ let $l_{i\nu}(\mathbf{x}) = \sum_j a_{ij\nu}x_j$ and let $A_\nu = (a_{ij\nu})_{i,j}$ be the $n \times n$ matrix with the coefficients of $l_{i\nu}$ on the i-th row. Let $\Delta_\nu = \det(l_{i\nu}(\mathbf{a}_j))$. Note that $\|A_\nu \mathbf{a}_j\|_\nu = \max_{1\leqslant i\leqslant n} |l_{i\nu}(\mathbf{a}_j)|_\nu$. By the rules of matrix multiplication, we have $A_\nu(\mathbf{a}_1, \ldots, \mathbf{a}_n) = (l_{i\nu}(\mathbf{a}_j))_{i,j} = \Delta_\nu$, where $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ is the matrix with columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$. Then by taking determinants, we get

$$\begin{aligned}
\prod_{\nu\in S} |\det A_\nu|_\nu \prod_{\nu\in S} |\det(\mathbf{a}_1, \ldots, \mathbf{a}_n)|_\nu &= \prod_{\nu\in S} |\Delta_\nu|_\nu \leqslant \prod_{\nu\in S}\prod_{j=1}^{n} \max_{1\leqslant i\leqslant n} |l_{i\nu}(\mathbf{a}_j)|_\nu \\
&= \prod_{j=1}^{n}(\prod_{\nu\in S} \|A_\nu \mathbf{a}_j\|_\nu) = \prod_{j=1}^{n} H_A(\mathbf{a}_j) \\
&= |\det A| = \prod_{\nu\in S} |\det A_\nu|_\nu.
\end{aligned}$$

Thus we deduce that $\prod_{\nu\in S} |\det(\mathbf{a}_1, \ldots, \mathbf{a}_n)|_\nu \leqslant 1$. Since $\mathbf{a}_i \in \mathcal{O}_S^n$, we have $|\det(\mathbf{a}_1, \ldots, \mathbf{a}_n)|_\nu \leqslant 1$ for $\nu \notin S$. By the product formula we have $\prod_{\nu\in S} |\det(\mathbf{a}_1, \ldots, \mathbf{a}_n)|_\nu = 1$, hence $|\det(\mathbf{a}_1, \ldots, \mathbf{a}_n)|_\nu = 1$ for $\nu \notin S$. This implies that $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is an $\mathcal{O}_S$-module basis of $\mathcal{O}_S^n$. □

3.2 A generalization

For an arbitrary field L, we denote by $L[X_1, \ldots, X_n]^{\text{lin}}$ the L-vector space of linear forms in n variables with coefficients in L. Recall that $K = k(t)$ and S a finite set of valuations of K containing ν_∞. For each $\nu \in S$, let

$m_{1\nu}, \ldots, m_{n\nu}$ be linearly independent linear forms from $K_\nu[X_1, \ldots, X_n]^{\mathrm{lin}}$ and define

$$\mathcal{C}_\nu = \{\mathbf{x} \in K_\nu^n : \max_{1 \leqslant i \leqslant n} |m_{i\nu}(\mathbf{x})|_\nu \leqslant 1\}.$$

By Lemma 3.1.2, this is indeed a symmetric ν-adic convex body. Then for $\mathcal{C} = \prod_{\nu \in S} \mathcal{C}_\nu$, Theorem 3.1.8 gives the equality

$$\prod_{i=1}^{n} \lambda_i = \prod_{\nu \in S} |\det(m_{1\nu}, \ldots, m_{n\nu})|_\nu,$$

with $\lambda_i, i = 1, \ldots, n$ the successive minima of $\prod_{\nu \in S} \mathcal{C}_\nu$. We may generalize this result as follows.

Let S be a finite set of valuations of K containing the infinite valuation ∞. For every $\nu \in S$, $|\cdot|_\nu$ has a unique extension to the algebraic closure $\overline{K_\nu}$. Let $l_{1\nu}, \ldots, l_{m\nu}, m \geqslant n$ be a set of linear forms in $\overline{K_\nu}[X_1, \ldots, X_n]^{\mathrm{lin}}$, with rank n. Let $\mathcal{C}_\nu = \{\mathbf{x} \in K_\nu^n : |l_{i\nu}(\mathbf{x})|_\nu \leqslant 1 \, (1 \leqslant i \leqslant m)\}$ and $\mathcal{C} = \prod_{\nu \in S} \mathcal{C}_\nu$. Since $\mathrm{rank}(l_{1\nu}, \ldots, l_{m\nu}) = n$, $\mathcal{C}$ is indeed a convex symmetric body. We say that $\{l_{1\nu}, \ldots, l_{m\nu}\}$ is $\mathrm{Gal}(\overline{K_\nu}/K_\nu)$-*symmetric*, if for every $\sigma \in \mathrm{Gal}(\overline{K_\nu}/K_\nu)$, the linear forms $\sigma(l_{1\nu}), \ldots, \sigma(l_{m\nu})$ are a permutation of $l_{1\nu}, \ldots, l_{m\nu}$. With this setting, we have the following result, which is a function field analogue of a result from the geometry of numbers over number fields by Evertse [9].

Theorem 3.2.1. *Let $\mathcal{L}_\nu = \{l_{1\nu}, \ldots, l_{m\nu}\} \subset \overline{K_\nu}[X_1, \ldots, X_n]$ be a $\mathrm{Gal}(\overline{K_\nu}/K_\nu)$-symmetric set of linear forms of rank n for each $\nu \in S$. Let $\lambda_1, \ldots, \lambda_n$ be the successive minima of $\mathcal{C}$. Then*

$$\prod_{i=1}^{n} \lambda_i \geqslant \prod_{\nu \in S} \max_{1 \leqslant i_1 < \cdots < i_n \leqslant m} |\det(l_{i_1\nu}, \ldots, l_{i_n\nu})|_\nu,$$

$$\prod_{i=1}^{n} \lambda_i \leqslant e^{(m-1)n\#S/2} \prod_{\nu \in S} \max_{1 \leqslant i_1 < \cdots < i_n \leqslant m} |\det(l_{i_1\nu}, \ldots, l_{i_n\nu})|_\nu.$$

Proof. Let ν in S and define $G_\nu := \mathrm{Gal}(\overline{K_\nu}/K_\nu)$. Since $\mathcal{L}_\nu = \{l_{1\nu}, \ldots, l_{m\nu}\}$ is $\mathrm{Gal}(\overline{K_\nu}/K_\nu)$-symmetric, we have an action of G_ν on $\mathcal{L}_\nu$. Consider the G_ν-orbits and without loss of generality, assume that $l_{1\nu}, \ldots, l_{r\nu}$ are representatives for the orbits. Let $K_{i\nu}$ be the field over K_ν generated by the coefficients of $l_{i\nu}$, and $\sigma_{i\nu}^{(1)}, \ldots, \sigma_{i\nu}^{(m_{i\nu})}$ the K_ν-isomorphic embeddings of $K_{i\nu}$ into $\overline{K_\nu}$, where $m_{i\nu} = [K_{i\nu} : K_\nu]$. Then it is clear that

$$\mathcal{L}_\nu = \bigcup_{i=1}^{r} \{\sigma_{i\nu}^{(1)}(l_{i\nu}), \ldots, \sigma_{i\nu}^{(m_{i\nu})}(l_{i\nu})\}. \tag{3.2.1}$$

This implies

$$\mathcal{C}_\nu = \{\mathbf{x} \in K_\nu^n : |l_{i\nu}(\mathbf{x})|_\nu \leqslant 1\,(1 \leqslant i \leqslant r)\}.$$

Let $\mathcal{O}_{i\nu}$ be the integral closure of R_ν in $K_{i\nu}$. Then it is a free R_ν-module of rank $[K_{i\nu} : K_\nu]$ (see [22], Chap. II, Prop. 3). Let $\omega_{i\nu}^{(1)}, \ldots, \omega_{i\nu}^{(m_{i\nu})}$ be an R_ν-basis of $\mathcal{O}_{i\nu}$. Then it is also a K_ν-basis of $K_{i\nu}$. Hence we may write

$$l_{i\nu} = \sum_{j=1}^{m_{i\nu}} \omega_{i\nu}^{(j)} M_{i\nu}^{(j)}, \tag{3.2.2}$$

where $M_{i\nu}^{(j)} \in K_\nu[X_1, \ldots, X_n]^{\mathrm{lin}}$. By the choice of our R_ν-basis, it is easy to see that for $y = \sum_{j=1}^{m_{i\nu}} \omega_{i\nu}^{(j)} x_j$, with $x_j \in K_\nu$, we have $|y|_\nu \leqslant 1$ if and only if $|x_j|_\nu \leqslant 1$ for $j = 1, \ldots, m_{i\nu}$. Hence $|l_{i\nu}(\mathbf{x})|_\nu \leqslant 1$ if and only if $|M_{i\nu}^{(j)}(\mathbf{x})|_\nu \leqslant 1$ for $j = 1, \ldots, m_{i\nu}$, and therefore

$$\mathcal{C}_\nu = \{\mathbf{x} \in K_\nu^n : |M_{i\nu}^{(j)}(\mathbf{x})|_\nu \leqslant 1\,(1 \leqslant i \leqslant r, 1 \leqslant j \leqslant m_{i\nu})\}.$$

By (3.2.1), we have $\sum_{i=1}^{r} m_{i\nu} = m$. Let $\{M_{1\nu}, \ldots, M_{m\nu}\}$ be the linear forms $M_{i\nu}^{(j)}(1 \leqslant i \leqslant r, 1 \leqslant j \leqslant m_{i\nu})$ in some order. Then

$$\mathcal{C}_\nu = \{\mathbf{x} \in K_\nu^n : |M_{i\nu}(\mathbf{x})|_\nu \leqslant 1\,(1 \leqslant i \leqslant m)\}.$$

Without loss of generality, we may assume that for each $\nu \in S$,

$$|\det(M_{1\nu}, \ldots, M_{n\nu})|_\nu = \max_{1 \leqslant i_1 < \cdots < i_n \leqslant m} |\det(M_{i_1\nu}, \ldots, M_{i_n\nu})|_\nu.$$

By Lemma 3.1.2, we must have $|\det(M_{1\nu}, \ldots, M_{n\nu})|_\nu > 0$, so we may write $M_{j\nu} = \sum_{h=1}^{n} \xi_{jh} M_{h\nu}$, with $\xi_{jh} \in K_\nu$ for all $j = 1, \ldots, m$. By Cramer's rule, we have $\xi_{jh} = \frac{\det(M_{1\nu},\ldots,M_{j\nu},\ldots,M_{n\nu})}{\det(M_{1\nu},\ldots,M_{h\nu},\ldots,M_{n\nu})}$ and hence $|\xi_{jh}|_\nu \leqslant 1$. By the ultrametric inequality, we have $|M_{j\nu}(\mathbf{x})|_\nu \leqslant \max_{1 \leqslant i \leqslant n} |M_{i\nu}(\mathbf{x})|_\nu$ for every $\mathbf{x} \in K_\nu^n$. Therefore, we have

$$\mathcal{C}_\nu = \{\mathbf{x} \in K_\nu^n : |M_{i\nu}(\mathbf{x})|_\nu \leqslant 1 \, (1 \leqslant i \leqslant n)\}.$$

By Theorem 3.1.8, we have for the successive minima of $\mathcal{C}$

$$\begin{aligned} \prod_{i=1}^{n} \lambda_i &= \prod_{\nu \in S} |\det(M_{1\nu}, \ldots, M_{n\nu})|_\nu \\ &= \prod_{\nu \in S} \max_{1 \leqslant i_1 < \cdots < i_n \leqslant m} |\det(M_{i_1\nu}, \ldots, M_{i_n\nu})|_\nu. \end{aligned} \tag{3.2.3}$$

By applying $\sigma_{i\nu}^{(1)}, \ldots, \sigma_{i\nu}^{(m_{i\nu})} : K_{i\nu} \hookrightarrow \overline{K_\nu}$ to (3.2.2), we infer that

$$\sigma_{i\nu}^{(h)}(l_{i\nu}) = \sum_{j=1}^{m_{i\nu}} \sigma_{i\nu}^{(h)}(\omega_{i\nu}^{(j)}) M_{i\nu}^{(j)}, 1 \leqslant h \leqslant m_{i\nu}, 1 \leqslant i \leqslant r.$$

In view of (3.2.1) we may write this in a matrix form

$$\begin{pmatrix} l_{1\nu} \\ \vdots \\ l_{m\nu} \end{pmatrix} = \Omega_\nu \begin{pmatrix} M_{1\nu}^{(1)} \\ \vdots \\ M_{r\nu}^{(m_{r\nu})} \end{pmatrix},$$

or simply $\mathcal{L}_\nu = \Omega_\nu \mathcal{M}_\nu$, where $\Omega_\nu = \mathrm{diag}(B_{1\nu}, \ldots, B_{r\nu})$ is a block matrix with $B_{i\nu} = (\sigma_{i\nu}^{(h)}(\omega_{i\nu}^{(j)}))_{h,j}$. Since $\omega_{i\nu}^{(1)}, \ldots, \omega_{i\nu}^{(m_{i\nu})}$ is an R_ν-basis of $\mathcal{O}_{i\nu}$, and integral over R_ν, we know that their conjugates $\sigma_{i\nu}^{(h)}(\omega_{i\nu}^{(j)})$ are also integral over R_ν, and moreover that every matrix $B_{i\nu}$ is invertible. Further, every entry of Ω_ν^{-1} is of the form $\frac{\mu}{\det \Omega_\nu}$ with $|\mu|_\nu \leqslant 1$.

Now we have $|\det \Omega_\nu|_\nu = \prod_{i=1}^{r} |\det B_{i\nu}|_\nu$ and as is well known, $(\det B_{i\nu})^2$ generates the ideal $D_{K_{i\nu}/K_\nu}$, where $D_{K_{i\nu}/K_\nu}$ is the local discriminant of $K_{i\nu}/K_\nu$. Recall that K_ν is complete, hence there is exactly one valuation $V_{i\nu}$ on $K_{i\nu}$ above ν, with ramification index $e_{i\nu} = m_{i\nu}$. By Lemma 1.2.3,

$$2 \cdot \nu(\det B_{i\nu}) = e_{i\nu} - 1.$$

We deduce that

$$\prod_{i=1}^{r} |\det B_{i\nu}|_\nu = e^{-\sum_{i=1}^{r}(e_{i\nu}-1)/2} = e^{\frac{-m+r}{2}} \geqslant e^{\frac{-m+1}{2}}.$$

Hence $\Omega_\nu^{-1} = (\omega_\nu^{ij})_{i,j}$ with $|\omega_\nu^{ij}|_\nu \leqslant e^{\frac{m-1}{2}}$. From $\mathcal{M}_\nu = \Omega_\nu^{-1}\mathcal{L}_\nu$ we know that each $M_{i\nu}$ is a linear combination of the linear forms $l_{i\nu}$ with coefficients whose $|\cdot|_\nu$-value is at most $e^{(m-1)/2}$. Combining this with (3.2.3), and applying the Cauchy-Binet formula from linear algebra which is valid over any field, we conclude that

$$\prod_{i=1}^{n} \lambda_i \leqslant e^{(m-1)n\#S/2} \prod_{\nu \in S} \max_{1 \leqslant i_1 < \cdots < i_n \leqslant m} |\det(l_{i_1\nu}, \ldots, l_{i_n\nu})|_\nu.$$

On the other hand, each entry of $\mathcal{M}_\nu$ has $|\cdot|_\nu$-value no more than 1, hence similarly as above, we have

$$\max_{1 \leqslant i_1 < \cdots < i_n \leqslant m} |\det(l_{i_1\nu}, \ldots, l_{i_n\nu})|_\nu \leqslant \max_{1 \leqslant i_1 < \cdots < i_n \leqslant m} |\det(M_{i_1\nu}, \ldots, M_{i_n\nu})|_\nu,$$

which combined with (3.2.3) gives

$$\prod_{i=1}^{n} \lambda_i \geqslant \prod_{\nu \in S} \max_{1 \leqslant i_1 < \cdots < i_n \leqslant m} |\det(l_{i_1\nu}, \ldots, l_{i_n\nu})|_\nu.$$

□

Remark 3.2.2. *The proof of Theorem 3.2.1 remains valid if for* $\nu \in S$, *we take sets of linear forms* $\{l_{1\nu}, \dots, l_{m(\nu),\nu}\}$ *of different cardinalities* $m(\nu) \geqslant n$, *and different numbers* $r(\nu)$ *of* G_ν*-orbits. In that case, our estimate for* $\prod_{i=1}^{n} \lambda_i$ *becomes*

$$\prod_{i=1}^{n} \lambda_i \leqslant \prod_{\nu \in S} e^{n(m(\nu)-r(\nu))/2} \max_{i_1,\dots,i_n} |\det(l_{i_1\nu}, \dots, l_{i_n\nu})|_\nu.$$

Now let L be a finite extension of K of degree m and of genus g_L.

Lemma 3.2.3. *Let* $c = e^{2g_L}$. *Then for every tuple* $(\alpha_\omega : \omega \in M_L)$ *such that*

$$\alpha_\omega \in e^{\mathbb{Z}} \text{ for } \omega \in M_L, \alpha_\omega = 1 \text{ for almost all } \omega, \prod_{\omega \in M_L} \alpha_\omega \geqslant c,$$

there is an $x \in L^*$ *such that* $|x|_\omega \leqslant \alpha_\omega$ *for all* $\omega \in M_L$.

Proof. Let $\alpha_\omega = e^{r_\omega}$ for $\omega \in M_L$ with $r_\omega \in \mathbb{Z}$ and $r_\omega = 0$ for almost all ω. Consider the divisor $D = \sum_{\omega \in M_L} r_\omega \omega$. By the Riemann-Roch theorem, if $\deg D = \sum_{\omega \in M_L} r_\omega \geqslant 2g_L$, then the dimension $\dim_k\{x \in L : \omega(x) \geqslant -r_\omega \text{ for } \omega \in M_L \text{ or } x = 0\}$ is positive, hence there exists $x \in L^*$ such that $\omega(x) \geqslant -r_\omega$, i.e., $|x|_\omega \leqslant \alpha_\omega$ for all $\omega \in M_L$. □

Chapter 4

Reduction theory for binary forms over $k(t)$

In this chapter we work out a reduction theory for binary forms over $k(t)$. This is a function field analogue of the reduction theory over number fields developed in [9]. We follow the arguments from [9].

Recall that $K = k(t)$ and S a finite set of valuations of K containing the infinite valuation ν_∞. For a binary form $F(X,Y) = a_0X^n + a_1X^{n-1}Y + \cdots + a_nY^n \in \mathcal{O}_S[X,Y]$, let

$$H_S(F) = \prod_{\nu \in S} \max(|a_0|_\nu, \ldots, |a_n|_\nu).$$

We say that two binary forms $F, G \in \mathcal{O}_S[X,Y]$ are $\mathrm{GL}\,(2, \mathcal{O}_S)$-equivalent if for some $u \in \mathcal{O}_S^\times$ and $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathcal{O}_S)$, we have

$$G(X,Y) = uF(aX + bY, cX + dY).$$

This equivalence relation preserves the S-value of the discriminant: $|D(F)|_S = |D(G)|_S$.

Definition 4.0. *A binary form $F \in \mathcal{O}_S[X,Y]$ is called S-reduced if $H_S(F) \leqslant H_S(G)$ for each binary form G that is $\mathrm{GL}\,(2, \mathcal{O}_S)$-equivalent to F.*

This is well-defined since $H_S(F)$ always lies in $e^{\mathbb{Z}}$ and for $F \in \mathcal{O}_S[X,Y]$ we have $H_S(F) \geqslant 1$.

Remark that by (1.2.1), we have $|D(F)|_S \leqslant H_S(F)^{2n-2}$.

4.1 Discriminant and genus

Let $F \in \mathcal{O}_S[X,Y]$ be a binary form with $D(F) \neq 0$ and $\deg F = n$. The ring $\mathcal{O}_S$ is a localization of $k[t]$, hence it is a principal ideal domain. So we may factor F as $F = F_1 \cdots F_d$ where $F_i \in \mathcal{O}_S[X,Y]$ is an irreducible binary form over K. If $F_i(1,0) \neq 0$ we may assume that $F_i = F_i(1,0)\mathrm{N}_{K_i/K}(X - \alpha_i Y)$ with $K_i = K(\alpha_i)$, where α_i is a root of $F_i(X,1)$. Let $\mathcal{O}_i$ be the integral closure of $\mathcal{O}_S$ in K_i. Since $\mathcal{O}_S$ is a principal ideal domain, $\mathcal{O}_i$ is a free $\mathcal{O}_S$-module of rank $[K_i : K]$. Assume it has an $\mathcal{O}_S$-basis $\{\omega_1, \dots, \omega_{d_i}\}$ where $d_i = [K_i : K] = \deg F_i$. The relative discriminant $D_i = D_{K_i/K}(\omega_1, \dots, \omega_{d_i})$ of an $\mathcal{O}_S$-basis $\omega_1, \dots, \omega_{d_i}$ is determined up to a multiplication by an element of $\mathcal{O}_S^{\times}$, hence the discriminant ideal $D_{\mathcal{O}_i/\mathcal{O}_S}$ of $\mathcal{O}_i$ over $\mathcal{O}_S$ generated by D_i is uniquely determined.

Lemma 4.1.1. *With the notation as above, we have $D_i | D(F_i)$ for $i = 1, \dots, d$.*

Proof. The proof is similar to that of Lemma 3 of [2]. We have included it for convenience of the reader.

We may assume without loss of generality that $F(1,0) \neq 0$ for if not, we may replace F by $F(X, mX+Y)$ for some integer m with $F(1,m) \neq 0$, which does not affect F_i and $D(F_i)$ for $i = 1, \dots, d$. Fix $i \in \{1, \dots, n\}$. If F_i has degree 1 then $(D_i) = (1), D(F_i) = 1$. Assume that F_i has degree $d_i \geqslant 2$. By assumption $F(1,0) \neq 0$, hence

$$F_i = b_0 X^{d_i} + b_1 X^{d_i - 1} Y + \cdots + b_{d_i} Y^{d_i} = b_0 N_{K_i/K}(X - \alpha_i Y),$$

where $b_j \in \mathcal{O}_S$ and $b_0 = F_i(1,0) \neq 0$.

Let

$$\begin{aligned}
\theta_1 &= b_0\alpha_i + b_1, \\
\theta_2 &= b_0\alpha_i^2 + b_1\alpha_i + b_2, \\
&\vdots \\
\theta_{d_i-1} &= b_0\alpha_i^{d_i-1} + b_1\alpha_i^{d_i-2} + \cdots + b_{d_i-1}.
\end{aligned}$$

We claim that they are integral over $\mathcal{O}_S$. This is equivalent to the assertion that $\theta_j - b_j$ is integral over $\mathcal{O}_S$ for $j = 1, \ldots, d_i - 1$; we prove this by induction on j. For $j = 1$, since $\sum_{h=0}^{d_i} b_h\alpha_i^{d_i-h} = 0$, we have $\sum_{h=0}^{d_i} b_h b_0^{h-1}(b_0\alpha_i)^{d_i-h} = 0$, hence $\theta_1 - b_1$ is integral over $\mathcal{O}_S$. Now let $j \geqslant 2$ and suppose the claim is true for $j-1$. Then using $\theta_j = \alpha_i\theta_{j-1} + b_j$ and $\theta_{j-1}\alpha_i^{d_i-j+1} = \sum_{h=d_i-j+1}^{d_i} b_{d_i-h}\alpha_i^h$, we deduce from $\sum_{h=0}^{d_i} b_h\alpha_i^{d_i-h} = 0$ that

$$\begin{aligned}
& (\theta_j - b_j)^{d_i-j+1} + \sum_{h=0}^{d_i-j} b_{d_i-h}\theta_{j-1}^{d_i-j-h}(\theta_j - b_j)^h \\
= \;& \theta_{j-1}^{d_i-j+1}\alpha_i^{d_i-j+1} + \sum_{h=0}^{d_i-j} \theta_{j-1}^{d_i-j} b_{d_i-h}\alpha_i^h \\
= \;& \theta_{j-1}^{d_i-j+1}\alpha_i^{d_i-j+1} - \theta_{j-1}^{d_i-j} \sum_{h=d_i-j+1}^{d_i} b_{d_i-h}\alpha_i^h \\
= \;& 0.
\end{aligned}$$

Therefore $\theta_j - b_j$ is integral over $\mathcal{O}_S[\theta_{j-1}]$, and hence it is integral over $\mathcal{O}_S$ by the induction hypothesis. This completes the induction hypothesis.

Consider the relative discriminant of $\{1, \theta_1, \dots, \theta_{d_i-1}\}$:

$$\begin{aligned} D_{K_i/K}(1, \theta_1, \dots, \theta_{d_i-1}) &= \det \begin{pmatrix} 1 & 0 & \cdots & 0 \\ b_1 & b_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ b_{d_i-1} & \cdots & b_1 & b_0 \end{pmatrix}^2 D_{K_i/K}(1, \alpha_i, \dots, \alpha_i^{d_i-1}) \\ &= b_0^{2d_i-2} \prod_{1\leqslant h<l\leqslant d_i} (\alpha_i^{(h)} - \alpha_i^{(l)})^2 \\ &= D(F_i), \end{aligned} \tag{4.1.1}$$

where $\alpha_i^{(h)}$ denotes the h-th conjugate of α_i in K_i, and the last equality comes from the definition. Also, we have $\theta_j = \sum_h a_{jh}\omega_h$ with $a_{jh} \in \mathcal{O}_S$. Then we have

$$D_{K_i/K}(1, \theta_1, \dots, \theta_{d_i-1}) = \det(a_{jh})^2 D_{K_i/K}(\omega_1, \dots, \omega_{d_i}). \tag{4.1.2}$$

Now (4.1.1) and (4.1.2) complete the proof. □

Because taking the discriminant commutes with localization (see [15]), the ideal $D_{\mathcal{O}_i/\mathcal{O}_S}$ of $\mathcal{O}_S$ is also generated by the relative discriminant ideal $D_{\mathcal{O}_{K_i}}$ of the integral closure $\mathcal{O}_{K_i}$ of $k[t]$ in K_i, so $D_{\mathcal{O}_i/\mathcal{O}_S} = D_{\mathcal{O}_{K_i}/k[t]}\mathcal{O}_S$. See also Chapter III, §2, [18].

Lemma 4.1.2. *Let $K_1, \dots, K_d$ be as before. For $i = 1, \dots, d$, let g_{K_i} be the genus of K_i. If $\#S > 1$, then*

$$\prod_{i=1}^{d} e^{2g_{K_i}} \leqslant e^{(\#S-2)(n-d)}|D(F)|_S.$$

Proof. By Lemma 1.2.3, we have an element p of k such that if $\nu = \nu_p$ is its corresponding valuation,

$$\nu(D_{\mathcal{O}_{K_i}}) = \sum_{\omega|\nu} \nu(\mathfrak{D}_{\mathcal{O}_{K_i}/k[t]}) = \sum_{\omega|\nu}(e(\omega|\nu) - 1).$$

Further, by the Riemann-Hurwitz formula,

$$\begin{aligned} 2g_{K_i}-2 &= [K_i:K](2g_K-2)+\sum_{\nu}\sum_{\omega|\nu}(e(\omega|\nu)-1) \\ &= -2d_i+\sum_{\nu\in S}\sum_{\omega|\nu}(e(\omega|\nu)-1)+\sum_{\nu\notin S}\nu(D_{\mathcal{O}_{K_i}}) \\ &\leqslant -2d_i+\sum_{\nu\in S}\sum_{\omega|\nu}(e(\omega|\nu)-1)+\sum_{\nu\notin S}\nu(D(F_i)), \end{aligned}$$

where the last inequality comes from Lemma 4.1.1.

Since $f(\omega|\nu)=1$ for each $\omega|\nu$, we have $\sum_{\omega|\nu} e(\omega|\nu)=d_i$. By the definition of the resultant, we have

$$D(F)=\prod_{1\leqslant i<j\leqslant r} R(F_i,F_j)^2\prod_{i=1}^{d} D(F_i), \tag{4.1.3}$$

where $R(F_i,F_j)\in\mathcal{O}_S$. Hence $\prod_{i=1}^{d} D(F_i)|D(F)$.

Using $\sum_{i=1}^{d} d_i=n$, we get

$$\begin{aligned} \sum_{i=1}^{d}(2g_{K_i}-2) &\leqslant \sum_{i=1}^{d}(-2d_i+\sum_{\nu\in S}(d_i-1))+\sum_{\nu\notin S}\nu(D(F)) \\ &= (n-d)\#S-2n-\sum_{\nu\in S}\nu(D(F)). \end{aligned}$$

Thus, we conclude that $\prod_{i=1}^{d} e^{2g_{K_i}}\leqslant e^{(\#S-2)(n-d)}|D(F)|_S$. □

4.2 Preparations on polynomials

Let $K=k(t)$. We still denote by $|\cdot|_\nu$ the unique extension of $|\cdot|_\nu$ to $\overline{K_\nu}$. Recall that for $P\in\overline{K_\nu}[X_1,\ldots,X_m]$ we have defined $|P|_\nu=$

$\max(|a_1|_\nu, \dots, |a_n|_\nu)$, where $a_1, \dots, a_n$ are the non-zero coefficients of P. For a finite set S of valuations containing $\{\nu_\infty\}$, $P \in K[X_1, \dots, X_m]$, define

$$|P|_S = (\prod_{\nu \in M_K \setminus S} |P|_\nu)^{-1} \text{ for } P \neq 0,$$

and $|0|_S = 0$ by convention. This is well-defined since $|P|_\nu = 1$ for almost all $\nu \in M_K$. For $P = a$ a constant, we have by the product formula $|P|_S = \prod_{\nu \in S} |a|_\nu$. If $P \in \mathcal{O}_S[X_1, \dots, X_m] \setminus \{0\}$, then $|P|_S \geqslant 1$. Clearly, $|aP|_S = |a|_S |P|_S$ for $a \in K^*, P \in K[X_1, \dots, X_m]$. Define the inhomogeneous height of $P \in K[X_1, \dots, X_m]$ by

$$H^*(P) = \prod_{\nu \in M_K} \max(1, |P|_\nu).$$

For $P \in \mathcal{O}_S[X_1, \dots, X_m]$, we have $|P|_\nu \leqslant 1$ for every $\nu \notin S$, hence

$$H^*(P) = \prod_{\nu \in S} \max(1, |P|_\nu).$$

Similarly, for a finite extension L of K, and $P \in L[X_1, \dots, X_m]$, we define

$$H^*(P) = \left(\prod_{\omega \in M_L} \max(1, |P|_\omega) \right)^{1/[L:K]}.$$

Lemma 4.2.1. *Let $P \in \mathcal{O}_S[X, Y]$ be a binary form. Then there exists $u \in \mathcal{O}_S^*$ such that $H^*(uP) = \prod_{\nu \in S} |P|_\nu$.*

Proof. We may write $P = \frac{1}{a}(b_0 X^n + b_1 X^{n-1} Y + \dots + b_n Y^n) \in \mathcal{O}_S[X, Y]$, where $a, b_i \in k[t] (1 \leqslant i \leqslant n)$, $\gcd(b_0, \dots, b_n, a) = 1$ and $|\frac{b_i}{a}|_\nu \leqslant 1$ for every $\nu \notin S$. Since $\gcd(b_0, \dots, b_n, a) = 1$ we have in fact $|a|_\nu = 1$ for $\nu \notin S$, i.e., $a \in \mathcal{O}_S^*$. Assume that $\gcd(b_0, \dots, b_n) = b \prod_{i=1}^{l} (t - p_i)^{h_i}$ with $h_i > 0, p_i \in S, 1 \leqslant i \leqslant l$ and $b \in k[t]$ a polynomial with zeros outside S. Let

$$b_i' = \frac{b_i}{\prod_{i=1}^{l} (t - p_i)^{h_i}}, \quad u = \frac{a}{\prod_{i=1}^{l} (t - p_i)^{h_i}}.$$

Then

$$b_i' \in \mathcal{O}_S \cap k[t](0 \leqslant i \leqslant n),\ u \in \mathcal{O}_S^*$$

and

$$P = \frac{1}{u}(b_0'X^n + b_1'X^{n-1}Y + \cdots + b_n'Y^n).$$

We deduce that

$$H^*(uP) = \prod_{\nu \in S} \max(1, |uP|_\nu) = \max(1, |uP|_\infty) = \max_{0\leqslant i\leqslant n}(e^{\deg b_i'}).$$

On the other hand, we have that $\gcd(b_0', \ldots, b_n') = b$ is coprime with $t - p$ for each $p \in S$ with $p \neq \infty$, hence $\max_{1\leqslant i\leqslant n}(|b_i'|_\nu) = 1$ for $\nu \in S\backslash\{\infty\}$. Recalling that $u \in \mathcal{O}_S^*$, we see that

$$\begin{aligned} \prod_{\nu\in S}|P|_\nu &= \prod_{\nu\in S}\max_{0\leqslant i\leqslant n}(|\frac{b_i'}{u}|_\nu) = \frac{\prod_{\nu\in S}\max_{0\leqslant i\leqslant n}(|b_i'|_\nu)}{\prod_{\nu\in S}|u|_\nu} \\ &= \prod_{\nu\in S}\max_{0\leqslant i\leqslant n}(|b_i'|_\nu) = H^*(uP). \end{aligned}$$

□

Clearly, this result only depends on the coefficients and hence can be extended for polynomials in more variables.

For $F(X,Y) = a_0X^n + a_1X^{n-1}Y + \cdots + a_nY^n \in \mathcal{O}_S[X,Y]$, let L be its splitting field over K, and $G = \mathrm{Gal}(L/K)$ the corresponding Galois group. In this case, $N_{L/K}(P) = \prod_{\sigma\in G}\sigma(P)$.

Lemma 4.2.2. *Let $F = aN_{L/K}(l)$. Then there are $a' \in K^*$ and $\lambda \in L^*$ such that $F = a'N_{L/K}(l')$ where $l' = \lambda l \in \mathcal{O}_T[X,Y]$, and*

$$e^{-2g_L}\prod_{\nu\in M_K\backslash S}|F|_\nu^{-1} \leqslant |a'|_S \leqslant \prod_{\nu\in M_K\backslash S}|F|_\nu^{-1}.$$

Proof. Notice that by section 1.4 the sets $\mathcal{E}(\omega|\nu)\,(\omega|\nu)$ are a partition of $G = \mathrm{Gal}(L/K)$, so

$$|N_{L/K}(l)|_\nu = \prod_{\sigma\in G} |\sigma(l)|_\nu = \prod_{\omega|\nu}\prod_{\sigma\in\mathcal{E}(\omega|\nu)} |\sigma(l)|_\nu = \prod_{\omega|\nu} |l|_\omega.$$

Let $\omega_0 \in T$. Then by Lemma 3.2.3, there exists $\lambda \in L^*$ such that

$$\begin{cases} |\lambda|_{\omega_0} \leqslant e^{2g_L} \prod_{\nu\notin S} |N_{L/K}(l)|_\nu, & \\ |\lambda|_\omega \leqslant 1 & \text{for } \omega \in T\backslash\{\omega_0\}, \\ |\lambda|_\omega \leqslant |l|_\omega^{-1} & \text{for } \omega \in M_L\backslash T. \end{cases}$$

For this λ and $a' = aN_{L/K}(\lambda)^{-1}$, we see that $F = a'N_{L/K}(\lambda l)$ and the coefficients of λl are in $\mathcal{O}_T$. Hence, we have $N_{L/K}(l') \in \mathcal{O}_S[X,Y]$. So we have

$$|F|_\nu = |a'|_\nu |N_{L/K}(l')|_\nu \leqslant |a'|_\nu \text{ for } \nu \notin S.$$

From the product formula, we deduce that $|a'|_S \leqslant (\prod_{\nu\in M_K\backslash S} |F|_\nu)^{-1}$ and

$$\begin{aligned} |a'|_S &= |a|_S |N_{L/K}(\lambda)|_S^{-1} = |a|_S \prod_{\omega\in T} |\lambda|_\omega^{-1} \\ &\geqslant e^{-2g_L} |a|_S \prod_{\nu\in M_L\backslash S} |N_{L/K}(l)|_\nu^{-1} \\ &= e^{-2g_L} \prod_{\nu\in M_K\backslash S} |F|_\nu^{-1}. \end{aligned}$$

□

Lemma 4.2.3. *Let* $F(X,Y) = a_0X^n + a_1X^{n-1}Y + \cdots + a_nY^n \in \mathcal{O}_S[X,Y]$ *be a binary form with* $D(F) \neq 0$. *Then we have a factorization* $F = a\prod_{i=1}^n l_i$, *where* $a \in K^*$ *and the* l_i *are linear forms in* $\mathcal{O}_T[X,Y]$ *such that for every* $\sigma \in G$, $\sigma(l_1), \ldots, \sigma(l_n)$ *is a permutation of* $l_1, \ldots, l_n$.

Proof. Since $K[X,Y]$ is a UFD, we may assume $F = f_1 \cdots f_g$ with f_i irreducible over K, $1 \leqslant i \leqslant g$.

For a fixed i with $1 \leqslant i \leqslant g$, if $f_i \neq Y$, we may write $f_i = c_i N_{L_i/K}(l_i)$, with L_i a subfield of L/K generated by a root of $f_i(X,1)$, $c_i \in K, l_i \in L_i[X,Y]^{\mathrm{lin}}$. By Lemma 4.2.2, we have $f_i = c_i' N_{L_i/K}(l_i')$ with $c_i' \in K, l_i' \in \mathcal{O}_T[X,Y]^{\mathrm{lin}}$. So we have $F = a \prod_{i=1}^{g} N_{L_i/K}(l_i')$ with $a \in K, l_i' \in \mathcal{O}_T[X,Y]$. This gives a factorization into linear forms of $\mathcal{O}_T[X,Y]$, up to a scalar in K.

For every $\sigma \in \mathrm{Gal}(L/K)$, the restriction $\sigma|_{L_i}$ is a K-isomorphism of L_i, hence σ acts as a permutation. This completes the proof. □

Remark 4.2.4. *In accordance with Lemma 4.2.3, later we will view $\sigma \in G$ as a permutation of $(1,\ldots,n)$ such that $\sigma(l_i) = l_{\sigma(i)}$ for $i = 1,\ldots,n$.*

4.3 Reduced binary forms and successive minima

Let $F(X,Y) \in \mathcal{O}_S[X,Y]$ be a binary form of degree $n > 1$ with $D(F) \neq 0$, and let L be the splitting field of $F(X,Y)$ over K and $G = \mathrm{Gal}(L/K)$. By Lemma 4.2.3 we have a factorization $F = a \prod_{i=1}^{n} l_i$ with $l_i \in L[X,Y]^{\mathrm{lin}}$ and for each $\sigma \in G$ a permutation $\sigma(l_1),\ldots,\sigma(l_n)$ of $l_1,\ldots,l_n$.

For $\omega \in M_L$ and $\sigma \in G$, there is $\omega \circ \sigma \in M_L$ such that $|x|_{\omega\circ\sigma} = |\sigma(x)|_\omega$ for $x \in L$, and $\omega \circ \sigma \in T$ if and only if $\omega \in T$.

Definition 4.3.1. *We call $\mathbb{A} = (A_{i\omega} : \omega \in T, i = 1,\ldots,n)$ an admissible tuple if $A_{i\omega} > 0$ and $A_{\sigma(i),\omega} = A_{i,\omega\circ\sigma}$ for $\omega \in T, \sigma \in G, i = 1,\ldots,n$.*

For $\nu \in S$, denote by $\mathcal{A}(\nu)$ the set of valuations of L lying above ν, and put

$$\mathcal{C}_\nu = \{\mathbf{x} \in K_\nu^2 : |l_i(\mathbf{x})|_\omega \leqslant A_{i\omega} \text{ for } i = 1,\ldots,n,\ \omega|\nu\}. \tag{4.3.1}$$

It is easy to check that this is a ν-adic symmetric convex body since $D(F) \neq 0$. Consider $\mathcal{C} = \prod_{\nu \in S} \mathcal{C}_\nu$ and let λ_1, λ_2 be the successive minima of $\mathcal{C}$. Here $\mathcal{C}_\nu$ and $\mathcal{C}$ depend on A, but for convenience we omit the subscript A here. To estimate $\lambda_1\lambda_2$, we try to rewrite $\mathcal{C}_\nu$ so that Theorem 3.2.1 can be applied to it.

Lemma 4.3.2. *Let $\mathbb{A}$ be an admissible tuple and let λ_1, λ_2 be the successive minima of $\mathcal{C}$. Assume $n \geqslant 2$. Then*

$$\lambda_1\lambda_2 \geqslant \left(\prod_{\omega \in T} \max_{1 \leqslant i < j \leqslant n} \frac{|\det(l_i, l_j)|_\omega}{A_{i\omega} A_{j\omega}} \right)^{1/[L:K]}, \tag{4.3.2}$$

$$\lambda_1\lambda_2 \leqslant e^{(n+1)\#S} \left(\prod_{\omega \in T} \max_{1 \leqslant i < j \leqslant n} \frac{|\det(l_i, l_j)|_\omega}{A_{i\omega} A_{j\omega}} \right)^{1/[L:K]}. \tag{4.3.3}$$

Proof. First, let $s(\omega) = [L_\omega : K_\infty]$ if $\omega | \infty$ and $s(\omega) = 0$ otherwise. As $\mathcal{C}_\nu = \{\mathbf{x} \in K_\nu^2 : |l_i(\mathbf{x})|_\omega \leqslant A_{i\omega}$ for $i = 1, \ldots, n$, $\omega | \nu\}$, we have

$$\lambda \mathcal{C}_\nu = \{|l_i(\mathbf{x})|_\omega \leqslant \lambda^{s(\omega)} A_{i\omega} \text{ for } i = 1, \ldots, n,\ \omega | \nu\}\}.$$

By Theorem 3.1.8, we can choose an $\mathcal{O}_S$-basis $\{\mathbf{y}_1, \mathbf{y}_2\}$ of $\mathcal{O}_S^2$ such that $\mathbf{y}_i \in \lambda_i \mathcal{C}$, $i = 1, 2$. Since $\det(l_i, l_j) \det(\mathbf{y}_1, \mathbf{y}_2) = \det \begin{pmatrix} l_i(\mathbf{y}_1) & l_i(\mathbf{y}_2) \\ l_j(\mathbf{y}_1) & l_j(\mathbf{y}_2) \end{pmatrix}$, we deduce that

$$\begin{aligned} \frac{|\det(l_i, l_j)|_\omega}{A_{i\omega} A_{j\omega}} &= \frac{1}{|\det(\mathbf{y}_1, \mathbf{y}_2)|_\omega A_{i\omega} A_{j\omega}} \left| \det \begin{pmatrix} l_i(\mathbf{y}_1) & l_i(\mathbf{y}_2) \\ l_j(\mathbf{y}_1) & l_j(\mathbf{y}_2) \end{pmatrix} \right|_\omega \\ &\leqslant \frac{(\lambda_1\lambda_2)^{s(\omega)}}{|\det(\mathbf{y}_1, \mathbf{y}_2)|_\omega}. \end{aligned}$$

Hence

$$\prod_{\omega \in T} \max_{1 \leqslant i < j \leqslant n} \frac{|\det(l_i, l_j)|_\omega}{A_{i\omega} A_{j\omega}} \leqslant \frac{(\lambda_1\lambda_2)^{\sum_{\omega|\infty} [L_\omega : K_\infty]}}{|\det(\mathbf{y}_1, \mathbf{y}_2)|_T} = (\lambda_1\lambda_2)^{[L:K]}.$$

This gives (4.3.2).

For the second inequality, put $B_{i\nu} = A_{\sigma^{-1}(i),\omega}^{1/g_\nu}$ with corresponding $\omega \in \mathcal{A}(\nu)$ and $\sigma \in \mathcal{E}(\omega|\nu)$. We show that this is independent of the choice of ω, σ. Let ω', σ' be another pair with $\omega' \in \mathcal{A}(\nu)$ and $\sigma' \in \mathcal{E}(\omega'|\nu)$. Then $\omega \circ \tau = \omega'$ for $\tau = \sigma^{-1}\sigma'$, and by the admissibility of $\mathbb{A}$,

$$A_{\sigma'^{-1}(i),\omega'} = A_{\tau^{-1}\sigma^{-1}(i),\omega'} = A_{\sigma^{-1}(i),\omega'\circ\tau^{-1}} = A_{\sigma^{-1}(i),\omega},$$

hence the $B_{i\nu}$ are well-defined. Moreover, since $\mathcal{E}(\omega|\nu)$ is a right-coset of $\mathrm{Gal}(L_{\omega_1}/K_\nu)$, if $j = \tau(i)$ for $\tau \in \mathrm{Gal}(L_{\omega_1}/K_\nu)$, then $B_{i\nu} = B_{j\nu}$.

With this notation, by (1.4.3) we have that for $\mathbf{x} \in K_\nu^2$ the condition

$$|l_i(\mathbf{x})|_\omega \leqslant A_{i\omega} \text{ for } 1 \leqslant i \leqslant n, \omega \in \mathcal{A}(\nu)$$

is equivalent to the condition

$$|\sigma(l_i)(\mathbf{x})|_\nu \leqslant B_{\sigma(i),\nu} \text{ for } 1 \leqslant i \leqslant n, \omega \in \mathcal{A}(\nu), \sigma \in \mathcal{E}(\omega|\nu),$$

that is,

$$|l_{\sigma(i)}(\mathbf{x})|_\nu \leqslant B_{\sigma(i),\nu} \text{ for } 1 \leqslant i \leqslant n, \sigma \in \mathrm{Gal}(L/K),$$

which is equivalent to the condition

$$|l_i(\mathbf{x})|_\nu \leqslant B_{i\nu} \text{ for } 1 \leqslant i \leqslant n.$$

Altogether, we get

$$\mathcal{C}_\nu = \{\mathbf{x} \in K_\nu^2 : |l_i(\mathbf{x})|_\nu \leqslant B_{i\nu} \text{ for } 1 \leqslant i \leqslant n\}.$$

Since $|\cdot|_\nu$ is normalized, the value set of K_ν^* is $e^{\mathbb{Z}}$, hence for $\nu \in S$, we can choose $a_{i\nu} \in K_\nu^*, 1 \leqslant i \leqslant n$ satisfying

$$\begin{cases} B_{i\nu}/e < |a_{i\nu}|_\nu \leqslant B_{i\nu} \, (1 \leqslant i \leqslant n) \\ a_{i\nu} = a_{j\nu} \text{ if } i = \tau(j) \text{ for } \tau \in \mathrm{Gal}(L_{\omega_1}/K_\nu). \end{cases}$$

Put $m_{i\nu} = a_{i\nu}^{-1} l_i$ for $\nu \in S, 1 \leqslant i \leqslant n$. By the choice of l_i and $a_{i\nu}$, the system $\{m_{1\nu}, \ldots, m_{n\nu}\}$ is $\mathrm{Gal}(\overline{K_\nu}/K_\nu)$-symmetric. Further, let

$$\mathcal{C}'_\nu = \{\mathbf{x} \in K_\nu^2 : |m_{i\nu}(\mathbf{x})|_\nu \leqslant 1 \text{ for } 1 \leqslant i \leqslant n\}.$$

Then $\mathcal{C}'_\nu \subset \mathcal{C}_\nu$. Hence, the successive minima λ'_1, λ'_2 of $\prod_{\nu\in S} \mathcal{C}'_\nu$ satisfy $\lambda_i \leqslant \lambda'_i$ for $i = 1, 2$. By Theorem 3.2.1, we have

$$\begin{aligned}
\lambda_1\lambda_2 \leqslant \lambda'_1\lambda'_2 &\leqslant e^{(n-1)\#S} \prod_{\nu\in S} \max_{1\leqslant i<j\leqslant n} |\det(m_{i,\nu}, m_{j,\nu})|_\nu \\
&= e^{(n-1)\#S} \prod_{\nu\in S} \max_{1\leqslant i<j\leqslant n} \frac{|\det(l_i, l_j)|_\nu}{|a_{i\nu}a_{j\nu}|_\nu} \\
&\leqslant e^{(n+1)\#S} \prod_{\nu\in S} \max_{1\leqslant i<j\leqslant n} \frac{|\det(l_i, l_j)|_\nu}{B_{i\nu}B_{j\nu}},
\end{aligned} \tag{4.3.4}$$

Finally, by (1.4.3) we have

$$|\det(l_i, l_j)|_\omega = |\sigma(\det(l_i, l_j))|_\nu^{g_\nu} = |\det(l_{\sigma(i)}, l_{\sigma(j)})|_\nu^{g_\nu}$$

for $\omega|\nu$ and $\sigma \in \mathcal{E}(\omega|\nu)$, where $g_\nu = \#\mathcal{E}(\omega|\nu)$. This leads to

$$\begin{aligned}
\prod_{\omega|\nu} \max_{1\leqslant i<j\leqslant n} \frac{|\det(l_i, l_j)|_\omega}{A_{i\omega}A_{j\omega}} &= \prod_{\omega|\nu} \prod_{\sigma\in\mathcal{E}(\omega|\nu)} \max_{1\leqslant i<j\leqslant n} \frac{|\det(l_{\sigma(i)}, l_{\sigma(j)})|_\nu}{B_{\sigma(i),\nu}B_{\sigma(j),\nu}} \\
&= \prod_{\sigma\in\mathrm{Gal}(L/K)} \max_{1\leqslant i<j\leqslant n} \frac{|\det(l_{\sigma(i)}, l_{\sigma(j)})|_\nu}{B_{\sigma(i),\nu}B_{\sigma(j),\nu}} \\
&= \left(\max_{1\leqslant i<j\leqslant n} \frac{|\det(l_i, l_j)|_\nu}{B_{i\nu}B_{j\nu}}\right)^{[L:K]},
\end{aligned}$$

hence we deduce that

$$\prod_{\nu\in S} \max_{1\leqslant i<j\leqslant n} \frac{|\det(l_i, l_j)|_\nu}{B_{i\nu}B_{j\nu}} = \left(\prod_{\omega\in T} \max_{1\leqslant i<j\leqslant n} \frac{|\det(l_i, l_j)|_\omega}{A_{i\omega}A_{j\omega}}\right)^{1/[L:K]}.$$

Together with (4.3.4), this implies (4.3.3), and we complete the proof of our lemma.

□

Using Lemma 4.3.2, we can prove the following

Theorem 4.3.3. *Let $F \in \mathcal{O}_S[X,Y]$ be a binary form of degree n with non-zero discriminant and with splitting field L over K, and choose a factorization $F = a\prod_{i=1}^{n} l_i$ with $a \in K^*, l_i \in L[X,Y]^{lin}$ such that for every $\sigma \in G$, $(\sigma(l_1),\ldots,\sigma(l_n))$ is a permutation of $(l_1,\ldots,l_n)$. Put*

$$M = \prod_{\omega\in T}\prod_{i=1}^{n} A_{i\omega},$$

$$R = \prod_{\omega\in T}\max_{1\leqslant i<j\leqslant n}\frac{|\det(l_i,l_j)|_\omega}{A_{i\omega}A_{j\omega}}.$$

(i) If $n \geqslant 2$ and F has no factor in $K[X,Y]^{lin}$, then F is $\mathrm{GL}(2,\mathcal{O}_S)$-equivalent to a binary form F^* such that

$$H^*(F^*) \leqslant e^{n(n+1)\#S}|a|_S^2 R^{n/[L:K]}M^{2/[L:K]}.$$

(ii) If $n \geqslant 3$ and F does have a factor in $K[X,Y]^{lin}$, then F is $\mathrm{GL}(2,\mathcal{O}_S)$-equivalent to a binary form F^* such that

$$H^*(F^*) \leqslant \left(e^{n(n+1)\#S}|a|_S^2 R^{n/[L:K]}M^{2/[L:K]}\right)^{(n-1)/(n-2)}.$$

Proof. By Theorem 3.1.8, we have a basis $\mathbf{a}_1 = (a_{11},a_{21}), \mathbf{a}_2 = (a_{12},a_{22})$ of $\mathcal{O}_S^2$ such that $\mathbf{a}_i \in \lambda_i\prod_{\nu\in S}\mathcal{C}_\nu$ for $i = 1,2$. Hence we have

$$\begin{aligned} |l_i(\mathbf{a}_1)|_\omega &\leqslant \lambda_1^{s(\omega)}A_{i\omega},\\ |l_i(\mathbf{a}_2)|_\omega &\leqslant \lambda_2^{s(\omega)}A_{i\omega}, \end{aligned} \tag{4.3.5}$$

for $1 \leqslant i \leqslant n, \omega \in T$, with $s(\omega) = [L_\omega : K_\infty]$ if $\omega|\nu_\infty$ and zero otherwise. Take $U = \left(\begin{smallmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{smallmatrix}\right)$. Then $U \in \mathrm{GL}(2,\mathcal{O}_S)$, and $F_U = a\prod_{i=1}^{n} m_i$ with $m_i = l_i(\mathbf{a}_1)X + l_i(\mathbf{a}_2)Y$ for $i = 1,\ldots,n$. We deduce that for $\omega \in T$,

$$\begin{aligned} |F_U|_\omega &\leqslant |a|_\omega\prod_{i=1}^{n}\max(|l_i(\mathbf{a}_1)|_\omega,|l_i(\mathbf{a}_2)|_\omega)\\ &\leqslant |a|_\omega\lambda_2^{ns(\omega)}\prod_{i=1}^{n}A_{i\omega}. \end{aligned}$$

Also, we have

$$\prod_{\omega|\nu}|a|_\omega = |a|_\nu^{[L:K]}, \ \prod_{\omega|\nu}|F_U|_\omega = |F_U|_\nu^{[L:K]}$$

and

$$\sum_{\omega\in T} s(\omega) = [L:K],$$

therefore, we get

$$\prod_{\nu\in S}|F_U|_\nu = (\prod_{\omega\in T}|F_U|_\omega)^{1/[L:K]} \leqslant |a|_S\lambda_2^n M^{1/[L:K]}.$$

By Lemma 4.2.1, there exists $u \in \mathcal{O}_S^*$ such that $F^* = uF_U$ satisfies $H^*(F^*) = \prod\limits_{\nu\in S}|F_U|_\nu$, hence

$$H^*(F^*) \leqslant |a|_S\lambda_2^n M^{1/[L:K]}. \tag{4.3.6}$$

What remains is to estimate λ_2. First assume that F has no linear factor in $K[X,Y]$, so $F(\mathbf{a}_1) \in \mathcal{O}_S\backslash\{0\}$. Now by (4.3.5) we have

$$1 \leqslant \prod_{\omega\in T}|F(\mathbf{a}_1)|_\omega = \prod_{\omega\in T}|a|_\omega \cdot \prod_{\omega\in T}\prod_{i=1}^{n}|l_i(\mathbf{a}_1)|_\omega \leqslant |a|_S^{[L:K]}\lambda_1^{n[L:K]}M.$$

Together with Lemma 4.3.2, we deduce that

$$\lambda_2^n \leqslant e^{n(n+1)\#S}|a|_S R^{n/[L:K]}M^{1/[L:K]},$$

and therefore by (4.3.6),

$$H^*(F^*) \leqslant e^{n(n+1)\#S}|a|_S^2 R^{n/[L:K]}M^{2/[L:K]}.$$

Next assume that F does have a linear factor in $K[X,Y]$. If $F(\mathbf{a}_1) \neq 0$, we still have the above result. Assume $F(\mathbf{a}_1) = 0$ and $n \geqslant 3$. Without loss of generality, let $l_1(\mathbf{a}_1) = 0$. Since $D(F) \neq 0$, we have

$$W := al_1(\mathbf{a}_2)\prod_{i=2}^{n} l_i(\mathbf{a}_1) \neq 0.$$

As $\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{O}_S$, we have by Gauss' Lemma

$$|W|_\omega \leqslant |a|_\omega \prod_{i=1}^{n} |l_i|_\omega = |F|_\omega \leqslant 1 \text{ for } \omega \notin T.$$

Hence, using (4.3.5) we deduce that

$$\begin{aligned} 1 &\leqslant \prod_{\omega \in T} |W|_\omega \\ &\leqslant (\lambda_1^{n-1}\lambda_2)^{[L:K]} \prod_{\omega \in T} |a|_\omega M \\ &= |a|_S^{[L:K]} (\lambda_1^{n-1}\lambda_2)^{[L:K]} M. \end{aligned}$$

Then together with Lemma 4.3.2, we obtain

$$\begin{aligned} \lambda_2^{n-2} &\leqslant \lambda_2^{n-2} \cdot (\lambda_1^{n-1}\lambda_2)|a|_S M^{\frac{1}{[L:K]}} \\ &\leqslant |a|_S e^{(n^2-1)\#S} M^{1/[L:K]} R^{(n-1)/[L:K]}, \end{aligned}$$

and finally, by (4.3.6)

$$H^*(F^*) \leqslant \left(e^{n(n+1)\#S} |a|_S^2 R^{n/[L:K]} M^{2/[L:K]}\right)^{(n-1)/(n-2)}.$$

□

Remark 4.3.4. *The binary form F^* depends on the admissible tuple $\mathbb{A}$. We say that F^* is associated with $\mathbb{A}$. By taking the special case $A_{i\omega} = 1$ for $1 \leqslant i \leqslant n, \omega \in T$, we obtain:*

Corollary 4.3.5. *Let $F \in \mathcal{O}_S[X,Y]$ be a binary form of degree n with non-zero discriminant. Then with the same factorization of F as in Theorem 4.3.3,*

(i) if $n \geqslant 2$ and F has no factor in $K[X,Y]^{lin}$, then F is $\mathrm{GL}(2, \mathcal{O}_S)$-equivalent to a binary form F^* such that

$$H^*(F^*) \leqslant e^{n(n+1)\#S} |a|_S^2 \Big(\prod_{\omega \in T} \max_{1 \leqslant i < j \leqslant n} |\det(l_i, l_j)|_\omega \Big)^{n/[L:K]}.$$

(ii) if $n \geqslant 3$ and F does have a factor in $K[X,Y]^{lin}$, then F is $\mathrm{GL}(2,\mathcal{O}_S)$-equivalent to a binary form F^* such that

$$H^*(F^*) \leqslant \left(e^{n(n+1)\#S}|a|_S^2\Big(\prod_{\omega\in T}\max_{1\leqslant i<j\leqslant n}|\det(l_i,l_j)|_\omega\Big)^{n/[L:K]}\right)^{(n-1)/(n-2)}.$$

Corollary 4.3.6. *Let $F \in \mathcal{O}_S[X,Y]$ be a binary quadratic form of non-zero discriminant $D(F)$. Then F is $\mathrm{GL}(2,\mathcal{O}_S)$-equivalent to a binary form F^* such that $H^*(F^*) \leqslant e^{6\#S}|D(F)|_S$.*

Proof. If F is irreducible over K, then we may factor as $F = al_1l_2$ with $a \in K^*, l_1, l_2 \in L[X,Y]^{\mathrm{lin}}$ conjugate over K and in this case, $n = 2, [L:K] = 2$ and $D(F) = a^2\det(l_1,l_2)^2$. Take $A_{1\omega} = A_{2\omega} = 1$ for every $\omega \in T$. By Theorem 4.3.3, there exists a binary form F^* equivalent to F such that

$$H^*(F^*) \leqslant e^{6\#S}|D(F)|_S.$$

However, if F is reducible over K, then $L = K, T = S$. We follow the idea in the proof of Theorem 4.3.3. We may factor F as $F = l_1l_2$ with $l_1, l_2 \in K[X,Y]^{\mathrm{lin}}$. Take $A_{1\infty} = |l_1|_S, A_{2\infty} = |l_2|_S$, $A_{i\nu} = 1$ for $\nu \in S\backslash\infty, i = 1,2$. Further, take $\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{O}_S^2$ as in proof of Theorem 4.3.3. Then one of $l_1(\mathbf{a}_1), l_2(\mathbf{a}_1)$ is non-zero, say, $l_1(\mathbf{a}_1) \neq 0$, and we have

$$\begin{aligned} 1 &= \prod_{\nu\in M_K}|l_1(\mathbf{a}_1)|_\nu \\ &\leqslant \prod_{\nu\in S}|l_1(\mathbf{a}_1)|_\nu\prod_{\nu\notin S}|l_1|_\nu \\ &\leqslant \lambda_1|l_1|_S\prod_{\nu\notin S}|l_1|_\nu \\ &= \lambda_1. \end{aligned}$$

Applying Lemma 4.3.2, we get

$$\lambda_2 \leqslant \lambda_1\lambda_2 \leqslant e^{3\#S}|\det(l_1,l_2)|_S/|l_1l_2|_S.$$

Hence there exists F^* equivalent to F such that

$$H^*(F^*) \leqslant e^{6\#S} |\deg(l_1, l_2)|_S^2 = e^{6\#S} |D(F)|_S.$$

□

Corollary 4.3.7. *Let $F \in \mathcal{O}_S[X,Y]$ be a binary cubic form of non-zero discriminant $D(F)$. Then F is* $\mathrm{GL}\,(2, \mathcal{O}_S)$*-equivalent to a binary form F^* such that*

(i) if F is irreducible over K, then $H^*(F^*) \leqslant e^{12\#S}|D(F)|_S^{\frac{1}{2}}$;

(ii) if F is reducible over K, then $H^*(F^*) \leqslant e^{12\#S}|D(F)|_S$.

Proof. Factor as $F = al_1l_2l_3$. Take $A_{i\omega} = |\det(l_j, l_h)|_\omega^{-1}$ for $i = 1, 2, 3, \omega \in T$ with $\{i, j, h\} = \{1, 2, 3\}$. This gives an admissible tuple. Indeed, for $\sigma \in \mathrm{Gal}(L/K), \omega \in T$ and $i = 1, 2, 3$, we have

$$\begin{aligned} A_{\sigma(i),\omega} &= |\det(l_{\sigma(j)}, l_{\sigma(h)})|_\omega^{-1} \\ &= |\sigma(\det(l_j, l_h))|_\omega^{-1} \\ &= |\det(l_j, l_h)|_{\omega\circ\sigma}^{-1} \\ &= A_{i,\omega\circ\sigma}. \end{aligned}$$

By $\prod\limits_{\omega|\nu} |a|_\omega = |a|_\nu^{[L:K]}$, we have

$$\prod_{\omega\in T}\prod_{i=1}^{3} A_{i\omega} = (\prod_{\nu\in S} |\det(l_1, l_2)\det(l_2, l_3)\det(l_3, l_1)|_\nu^{-[L:K]}),$$

and further,

$$\max_{1\leqslant i<j\leqslant 3} \frac{|\det(l_i, l_j)|_\omega}{A_{i\omega}A_{j\omega}} = |\det(l_1, l_2)\det(l_2, l_3)\det(l_3, l_1)|_\omega,$$

$$a^4(\det(l_1, l_2)\det(l_2, l_3)\det(l_3, l_1))^2 = D(F).$$

Now an application of Theorem 4.3.3 gives the desired result. □

Chapter 5

Height estimates in terms of the discriminant

We are going to prove a generalization of Theorem 1 in the introduction.

Main Theorem. *Let $K = k(t)$ and S a finite set of valuations of K containing ν_∞. Let $F \in \mathcal{O}_S[X,Y]$ be a binary form of degree $n \geqslant 4$ with non-zero discriminant. Then F is* $\mathrm{GL}\,(2, \mathcal{O}_S)$*-equivalent to a binary form F^* such that*

$$H^*(F^*) \leqslant e^{(n-1)\left(\#S(n+11)-5\right)} |D(F)|_S^{20+\frac{1}{n}}.$$

We mainly follow the arguments from [9].

5.1 Consequences of the Riemann-Hurwitz formula

First we deduce some consequences of the Riemann-Hurwitz formula. In this section, let K_1 be a finite extension of $k(t)$, unless otherwise stated. Here k is an algebraically closed field of characteristic 0, this implies that all residue degrees are 1. Let L be a finite extension of K_1. Then by the Riemann-Hurwitz formula, we have

$$2g_L - 2 = [L : K_1](2g_{K_1} - 2) + \sum_{\nu \in M_{K_1}} \sum_{V|\nu} (e(V|\nu) - 1).$$

We denote by S_1 a finite set of valuations of K_1, and by T the set of valuations of L above S_1. Clearly, we have $\#T \leqslant [L : K_1]\#S_1$. For $\nu \in M_{K_1}$, we put $R_{L/K_1,\nu} = \sum_{V|\nu}(e(V|\nu) - 1)$, where the sum is taken over all valuations V of L lying above ν. Then $0 \leqslant R_{L/K_1,\nu} \leqslant [L : K_1] - 1$.

Lemma 5.1.1. $2g_L - 2 + \#T = [L : K_1](2g_{K_1} - 2 + \#S_1) + \sum_{\nu \notin S_1} R_{L/K_1,\nu}$.

Proof. By the Riemann-Hurwitz formula, we have

$$\begin{aligned} 2g_L - 2 + \#T &= [L : K_1](2g_{K_1} - 2) + \#T + \sum_{\nu \in S_1} R_{L/K_1,\nu} + \sum_{\nu \notin S_1} R_{L/K_1,\nu} \\ &= [L : K_1](2g_{K_1} - 2) + \#T + \sum_{\nu \in S_1} \Big(\sum_{V|\nu} e(V|\nu) \Big) - \#T + \sum_{\nu \notin S_1} R_{L/K_1,\nu} \\ &= [L : K_1](2g_{K_1} - 2) + \sum_{\nu \in S_1} [L : K_1] + \sum_{\nu \notin S_1} R_{L/K_1,\nu} \\ &= [L : K_1](2g_{K_1} - 2 + \#S_1) + \sum_{\nu \notin S_1} R_{L/K_1,\nu}, \end{aligned}$$

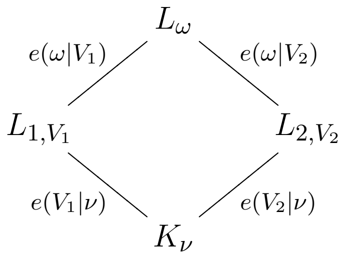
as claimed. □

Consider the compositum L of finite extensions $L_1, \ldots, L_r$ of K_1.

Lemma 5.1.2. *Let* $\nu \in M_{K_1}$. *Then* $\frac{R_{L/K_1,\nu}}{[L:K_1]} \leqslant \sum_{i=1}^{r} \frac{R_{L_i/K_1,\nu}}{[L_i:K_1]}$.

Proof. It suffices to prove this in the case $r = 2$. Then the general statement follows easily by induction.

So assume $r = 2$. Let $\omega \in M_L$ with $\omega|\nu$ and let $V_1 \in M_{L_1}, V_2 \in M_{L_2}$ be such that $\omega|V_1, \omega|V_2$. We have the diagram



Since all residue degrees are equal to 1, every ramification index is equal to the extension degree. By general theory of field extensions, we know that $e(\omega|V_2) \leqslant e(V_1|\nu), e(\omega|V_1) \leqslant e(V_2|\nu)$ and $e(\omega|\nu) = e(\omega|V_1)e(V_1|\nu)$. On the other hand, since every ramification index is a positive integer, we deduce that

$$(e(\omega|V_2) - 1)(e(\omega|V_1) - 1) \geqslant 0,$$

hence

$$e(\omega|V_2)e(V_2|\nu) - e(\omega|V_1) - e(\omega|V_2) \geqslant -1,$$

and therefore

$$e(\omega|V_1)(e(V_1|\nu) - 1) + e(\omega|V_2)(e(V_2|\nu) - 1) \geqslant e(\omega|\nu) - 1.$$

By taking the sum we deduce that

$$\begin{aligned}\sum_{\substack{\omega \in M_L \\ \omega|\nu}} \Big(e(\omega|\nu) - 1\Big) \;\leqslant\; & \sum_{\substack{V_1 \in M_{L_1} \\ V_1|\nu}} \Big(\sum_{\substack{\omega \in M_L \\ \omega|V_1}} e(\omega|V_1)(e(V_1|\nu) - 1)\Big) \\ + & \sum_{\substack{V_2 \in M_{L_2} \\ V_2|\nu}} \Big(\sum_{\substack{\omega \in M_L \\ \omega|V_2}} e(\omega|V_2)(e(V_2|\nu) - 1)\Big).\end{aligned}$$

Noticing that $\sum_{\omega|V_i} e(\omega|V_i) = [L : L_i]$ for $i = 1, 2$, this leads to

$$R_{L/K_1,\nu} \leqslant [L : L_1]R_{L_1/K_1,\nu} + [L : L_2]R_{L_2/K_1,\nu}.$$

which implies the desired result. □

We deduce some other genus estimates that will be needed later.

Recall $K = k(t)$. Assume that $F(1,0) \neq 0$. Then we may assume that $F = a\prod_{j=1}^{n}(X - \alpha_i Y), a \in K^*$ and that for every $\sigma \in \mathrm{Gal}(\overline{K}/K)$, there is a permutation of $(1,\ldots,n)$, also denoted by σ, such that for $j = 1,\ldots,n$ we have $\sigma(\alpha_j) = \alpha_{\sigma(j)}$. Suppose we have a factorization $F = \prod_{i=1}^{d} F_i$ where $F_i \in \mathcal{O}_S[X,Y]$ is a primitive irreducible binary form of degree n_i. Let $\alpha_{i,j}, j = 1,\ldots,n_i$ be the zeros of $F_i(X,1)$ among $\alpha_1,\ldots,\alpha_n$. Then all terms $R_{K(\alpha_{i,j})/K,\nu}, 1 \leqslant j \leqslant n_i$ are equal. We put $L_{ij} = K(\alpha_{i,j})$ for $i = 1,\ldots,d, j = 1,\ldots,n_i$, and $I = \{1,\ldots,n\}$. Then

$$\sum_{i=1}^{d}\sum_{j=1}^{n_i}\frac{R_{L_{ij}/K,\nu}}{[L_{ij}:K]} = \sum_{i=1}^{d} R_{L_{i1}/K,\nu}. \tag{5.1.1}$$

For such a field L_{ij} and a valuation ω of $\mathcal{O}_{L_{ij}}$, lying above the valuation ν of $k[t]$, we have by Lemma 1.2.3

$$\nu(D_{\mathcal{O}_{L_{ij}}/k[t]}) = \sum_{\omega|\nu}(e(\omega|\nu) - 1).$$

Further, by Lemma 4.1.1, we have $R_{L_{ij}/K,\nu} = \nu(D_{\mathcal{O}_{L_{ij}}/k[t]}) \leqslant \nu(D(F_i))$ for $\nu \notin S$, hence

$$\sum_{\nu\notin S} R_{L_{ij}/K,\nu} \leqslant \sum_{\nu\notin S}\nu(D(F_i)) = -\sum_{\nu\in S}\nu(D(F_i)). \tag{5.1.2}$$

For any set of indices $J = \{i_1,\ldots,i_m\} \subset \{1,\ldots,n\}$ we put $L_J = K(\alpha_{i_1},\ldots,\alpha_{i_m})$, and let T_J be the set of valuations of L_J above S. For each $i \in \{1,\ldots,n\}$ we choose $\bar{i} \in \{1,\ldots,d\}$ such that α_i is a root of $F_{\bar{i}}$. Recall that $F = F_1\cdot\cdots\cdot F_d$, where F_i is an irreducible factor of F in $\mathcal{O}_S[X,Y]$ with $F_i(\alpha_{\bar{i}},1) = 0$. Then by Lemma 5.1.1 with $g_K = 0$ and Lemma 5.1.2, we have

$$
\begin{aligned}
\frac{2g_{L_J} - 2 + \#T_J}{[L_J : K]} &= (-2 + \#S) + \sum_{\nu \notin S} \frac{R_{L_J/K,\nu}}{[L_J : K]} \\
&\leqslant -2 + \#S + \sum_{\nu \notin S} \sum_{i \in J} \frac{R_{K(\alpha_i)/K,\nu}}{[K(\alpha_i) : K]} \\
&= -2 + \#S + \sum_{i \in J} \frac{1}{[K(\alpha_i) : K]} \sum_{\nu \notin S} R_{K(\alpha_i)/K,\nu} \\
&\leqslant -2 + \#S - \sum_{i \in J} \sum_{\nu \in S} \frac{\nu(D(F_{\bar{i}}))}{\deg F_{\bar{i}}}. \qquad (5.1.3)
\end{aligned}
$$

Applying this to $J = I$ and combining this with (5.1.1) and (5.1.2), we obtain for the splitting field L of F over K,

$$
\begin{aligned}
\frac{2g_L - 2 + \#T}{[L : K]} &\leqslant (-2 + \#S) - \sum_{\nu \in S} \sum_{i=1}^{n} \frac{\nu(D(F_{\bar{i}}))}{\deg F_{\bar{i}}} \\
&= (-2 + \#S) - \sum_{\nu \in S} \nu\Big(\prod_{i=1}^{d} D(F_i)\Big) \\
&\leqslant (-2 + \#S) - \sum_{\nu \in S} \nu(D(F)),
\end{aligned}
$$

where the last step follows from (4.1.3).

On the other hand, by Lemma 5.1.2, we have

$$
\begin{aligned}
\frac{2g_L - 2}{[L : K]} &= -2 + \sum_{\nu \in M_K} \frac{R_{L/K,\nu}}{[L : K]} \\
&\leqslant -2 + \sum_{\nu \in M_K} \sum_{i=1}^{d} \sum_{j=1}^{n_i} \frac{R_{L_{ij}/K,\nu}}{[L_{ij} : K]} \\
&= -2 + \sum_{\nu \in M_K} \sum_{i=1}^{d} R_{L_{i1}/K,\nu}.
\end{aligned}
$$

Let g_i be the genus of $L_{i1}, i = 1, \ldots, d$. Then

$$
2g_i - 2 = -2[L_{i1} : K] + \sum_{\nu \in M_K} R_{L_{i1}/K,\nu},\ i = 1, \ldots, d,
$$

so

$$\frac{2g_L - 2}{[L:K]} \leqslant -2 + \sum_{i=1}^{d} (2g_i - 2 + 2n_i), \tag{5.1.4}$$

or,

$$\frac{2g_L - 2}{[L:K]} \leqslant 2 \deg F - 2 + \sum_{i=1}^{d} (2g_i - 2). \tag{5.1.5}$$

5.2 A few lemmas

In the proof of Theorem 5 we follow the idea of [9]. We are going to construct a special admissible tuple $\mathbb{A}$ as in definition 4.3.1 with some good properties.

Let $F \in \mathcal{O}_S[X,Y]$ be a binary form of degree $n \geqslant 4$. Assume we have a factorization into linear forms $F = al_1 \dots l_n$ with $l_i = X - \alpha_i Y$. Denote by Δ_{ij} the determinant $\det(l_i, l_j)$. Then

$$\Delta_{ij}\Delta_{hl} + \Delta_{jh}\Delta_{il} + \Delta_{hi}\Delta_{jl} = 0. \tag{5.2.1}$$

We will use this identity and apply Lemma 2.1.1 (Mason's Theorem) to it. Let L' be the splitting field L of F or the field $L_{ijhl} = K(\alpha_i, \alpha_j, \alpha_h, \alpha_l)$, and T' the set of valuations of L' lying above those in S. The case when $L' = L_{ijhl}$ is prepared for Theorem 1, whilst the case $L' = L$ is a variation on Theorem 1, which will be needed as well.

We introduce some auxiliary quantities that will be used in the proof of Theorem 1. For $i = 1, \dots, n$, let

$$\xi_{i\omega} = \max(|\alpha_i|_\omega, 1) \text{ for } \omega \notin T,$$

$$\xi_{i\omega} = (\prod_{\omega' \notin T} \xi_{i\omega'})^{-1/\#T} \text{ for } \omega \in T.$$

Then $\prod_{\omega\in M_L} \xi_{i\omega} = 1$ for $i = 1, \ldots, n$. We also have for $\omega \notin T$,

$$\begin{aligned} \xi_{\sigma(i),\omega} &= \max(|\alpha_{\sigma(i)}|_\omega, 1) \\ &= \max(|\alpha_i|_{\omega_\sigma}, 1) \\ &= \xi_{i,\omega_\sigma}. \end{aligned}$$

Hence $\xi_{\sigma(i)\omega} = \xi_{i\omega_\sigma}$ for each $\omega \in T$ as well.

Next we put

$$\theta_{ij\omega} = \frac{|\det(l_i, l_j)|_\omega}{\xi_{i\omega}\xi_{j\omega}}, i \neq j.$$

We have $\theta_{ij\omega} \leqslant 1$ for $\omega \notin T$, $\theta_{\sigma(i),\sigma(j),\omega} = \theta_{ij\omega\circ\sigma}$ for $i, j \in \{1, \ldots, n\}$ with $i \neq j$ and $\sigma \in \mathrm{Gal}(L/K)$, and

$$\prod_{\omega\in M_L} \theta_{ij\omega} = 1.$$

Further, let $U' \subset M_{L'} \backslash T'$ be the set of valuations ω such that $|\Delta_{ij}\Delta_{hl}|_\omega$, $|\Delta_{jh}\Delta_{il}|_\omega$, $|\Delta_{hi}\Delta_{jl}|_\omega$ are not all equal. Then clearly $\#U' < \infty$.

Note that for $\gamma_1, \gamma_2 \in L^*$, we have by the product formula

$$\begin{aligned} H^*(\frac{\gamma_1}{\gamma_2}) &= \Big(\prod_{\omega\in M_L} \max(1, |\frac{\gamma_1}{\gamma_2}|_\omega)\Big)^{1/[L:K]} \\ &= \Big(\prod_{\omega\in M_L} \max(|\gamma_1|_\omega, |\gamma_2|_\omega)\Big)^{1/[L:K]} = H(\gamma_1, \gamma_2). \end{aligned}$$

Hence by (5.2.1) and Lemma 2.1.1, we have

$$\begin{aligned} &H(\Delta_{ij}\Delta_{hl}, \Delta_{jh}\Delta_{il}, \Delta_{hi}\Delta_{jl}) = H(\Delta_{ij}\Delta_{hl}, \Delta_{jh}\Delta_{il}) \\ &\leqslant e^{\max(2g_{L'}-2+\#T'+\#U',0)/[L':K]} \\ &\leqslant e^{\max(2g_{L'}-2+\#T',0)/[L':K]} e^{\#U'/[L':K]}. \end{aligned} \tag{5.2.2}$$

Let $U \subset M_L$ be the set of valuations outside T such that $|\Delta_{ij}\Delta_{hl}|_\omega$, $|\Delta_{jh}\Delta_{il}|_\omega$, $|\Delta_{hi}\Delta_{jl}|_\omega$ are not all equal. Then $\#U \leqslant [L : L']\#U'$.

Put $\theta'_{ij\omega'} = \frac{|\Delta_{ij}|_{\omega'}}{\xi_{i\omega'}\xi_{j\omega'}}$ with $\xi_{i\omega'} = \max(|\alpha_i|_{\omega'}, 1)$ for $\omega' \notin T'$. Then $\theta'_{ij\omega'} \leqslant 1$ and

$$U' \subset \{\omega' \notin T' : \min(\theta'_{ij\omega'}\theta'_{hl\omega'}, \theta'_{jh\omega'}\theta'_{il\omega'}, \theta'_{hi\omega'}\theta'_{jl\omega'}) < 1\}.$$

For $\omega' \in U'$, the minimum is in fact at most e^{-1}, so we have

$$\begin{aligned} e^{\#U'} &\leqslant \prod_{\omega' \in U'} \frac{1}{\min(\theta'_{ij\omega'}\theta'_{hl\omega'}, \theta'_{jh\omega'}\theta'_{il\omega'}, \theta'_{hi\omega'}\theta'_{jl\omega'})} \\ &\leqslant \prod_{\omega' \notin T'} \frac{1}{\min(\theta'_{ij\omega'}\theta'_{hl\omega'}, \theta'_{jh\omega'}\theta'_{il\omega'}, \theta'_{hi\omega'}\theta'_{jl\omega'})}, \end{aligned}$$

hence

$$\begin{aligned} e^{\#U'/[L':K]} &\leqslant \prod_{\omega' \notin T'} \left(\frac{\prod_{t \in \{i,j,h,l\}} \max(|\alpha_t|_{\omega'}, 1)}{\min(|\Delta_{ij}\Delta_{hl}|_{\omega'}, |\Delta_{jh}\Delta_{il}|_{\omega'}, |\Delta_{hi}\Delta_{jl}|_{\omega'})} \right)^{1/[L':K]} \\ &= \prod_{\omega \notin T} \left(\frac{\prod_{t \in \{i,j,h,l\}} \max(|\alpha_t|_{\omega}, 1)}{\min(|\Delta_{ij}\Delta_{hl}|_{\omega}, |\Delta_{jh}\Delta_{il}|_{\omega}, |\Delta_{hi}\Delta_{jl}|_{\omega})} \right)^{1/[L:K]} \\ &= \prod_{\omega \notin T} \left(\frac{1}{\min(\theta_{ij\omega}\theta_{hl\omega}, \theta_{jh\omega}\theta_{il\omega}, \theta_{hi\omega}\theta_{jl\omega})} \right)^{1/[L:K]}, \end{aligned} \qquad (5.2.3)$$

where the first equality comes from the fact $\prod_{\omega|\omega'} |x|_\omega = |x|_{\omega'}^{[L:L']}, x \in L'$.

Lemma 5.2.1. *We have*

(i) $\theta_{il\omega}\theta_{jh\omega} \leqslant \max(\theta_{ij\omega}\theta_{hl\omega}, \theta_{ih\omega}\theta_{jl\omega})$ *for all* $\omega \in T$*;*

(ii) $\prod_{\omega \in T} \max(\theta_{ij\omega}\theta_{hl\omega}, \theta_{ih\omega}\theta_{jl\omega}, \theta_{il\omega}\theta_{jh\omega})$
$\leqslant e^{\max(2g_{L'}-2+\#T', 0)[L:L']} \cdot \Big(\prod_{\omega \in T} \theta_{ij\omega}\theta_{hl\omega}\theta_{ih\omega}\theta_{jl\omega}\theta_{il\omega}\theta_{jh\omega} \Big)$*;*

(iii) $(\prod_{\omega \in T} \prod_{1 \leqslant i < j \leqslant n} \theta_{ij\omega})^{1/[L:K]} = |D(F)|_S^{1/2}$ *if* F *is primitive.*

Proof. (i) It is easy to see that (i) is a direct consequence of (5.2.1).

(ii) We have

$$
\begin{aligned}
&\left(\prod_{\omega\in T}\max(\theta_{ij\omega}\theta_{hl\omega},\theta_{ih\omega}\theta_{jl\omega},\theta_{il\omega}\theta_{jh\omega})\right)^{1/[L:K]}\\
=\;&\left(\frac{\prod_{\omega\in M_L}\max(\theta_{ij\omega}\theta_{hl\omega},\theta_{ih\omega}\theta_{jl\omega},\theta_{il\omega}\theta_{jh\omega})}{\prod_{\omega\notin T}\max(\theta_{ij\omega}\theta_{hl\omega},\theta_{ih\omega}\theta_{jl\omega},\theta_{il\omega}\theta_{jh\omega})}\right)^{1/[L:K]}\\
=\;&H(\Delta_{ij}\Delta_{hl},\Delta_{jh}\Delta_{il},\Delta_{hi}\Delta_{jl})\Big(\prod_{\omega\notin T}\frac{1}{\max(\theta_{ij\omega}\theta_{hl\omega},\theta_{ih\omega}\theta_{jl\omega},\theta_{il\omega}\theta_{jh\omega})}\Big)^{1/[L:K]}\\
\overset{(1)}{\leqslant}\;&e^{\max(2g_{L'}-2+\#T',0)/[L':K]}\times\\
\times\;&\prod_{\omega\notin T}\Big(\frac{1}{\min(\theta_{ij\omega}\theta_{hl\omega},\theta_{jh\omega}\theta_{il\omega},\theta_{ih\omega}\theta_{jl\omega})}\frac{1}{\max(\theta_{ij\omega}\theta_{hl\omega},\theta_{ih\omega}\theta_{jl\omega},\theta_{il\omega}\theta_{jh\omega})}\Big)^{1/[L:K]}\\
\overset{(2)}{\leqslant}\;&e^{\max(2g_{L'}-2+\#T',0)/[L':K]}\left(\prod_{\omega\in T}\theta_{ij\omega}\theta_{hl\omega}\theta_{ih\omega}\theta_{jl\omega}\theta_{il\omega}\theta_{jh\omega}\right)^{1/[L:K]},
\end{aligned}
$$

where (1) follows from (5.2.2) and (5.2.3), and (2) is deduced from the product formula and the simple fact that if $a,b,c\leqslant 1$, then $abc\leqslant\max(a,b,c)\min(a,b,c)$. This gives (ii).

(iii) If F is primitive, we have $|a|_\omega\prod_{i=1}^{n}\xi_{i\omega}=1$ for $\omega\notin T$. So

$$\prod_{\omega\notin T}(|a|_\omega\prod_{i=1}^{n}\xi_{i\omega})=1,$$

which implies that

$$\prod_{\omega\in T}(|a|_\omega\prod_{i=1}^{n}\xi_{i\omega})=1. \tag{5.2.4}$$

Notice that

$$D(F)=a^{2n-2}\prod_{1\leqslant i<j\leqslant n}\det(l_i,l_j)^2,$$

$$|D(F)|_T = |D(F)|_S^{[L:K]}$$

and

$$\prod_{\omega \in T} \prod_{1 \leqslant i < j \leqslant n} \xi_{i\omega} \xi_{j\omega} = \prod_{\omega \in T} \prod_{i=1}^{n} (\xi_{i\omega})^{n-1}.$$

Hence

$$\begin{aligned} \prod_{\omega \in T} \prod_{1 \leqslant i < j \leqslant n} \theta_{ij\omega} &= \prod_{\omega \in T} \prod_{1 \leqslant i < j \leqslant n} \frac{|\det(l_i, l_j)|_\omega}{\xi_{i\omega} \xi_{j\omega}} \\ &= \frac{\prod_{\omega \in T} |D(F)/a^{2n-2}|_\omega^{\frac{1}{2}}}{\prod_{\omega \in T} \prod_{i=1}^{n} \xi_{i\omega}^{n-1}} \\ &= |D(F)|_T^{\frac{1}{2}}. \end{aligned}$$

This completes the proof. □

Let

$$M = \prod_{\omega} \prod_{i=1}^{n} A_{i\omega},$$

$$R = \prod_{\omega \in T} \max_{1 \leqslant i < j \leqslant n} \frac{|\det(l_i, l_j)|_\omega}{A_{i\omega} A_{j\omega}},$$

$$A'_{i\omega} = \frac{A_{i\omega}}{\xi_{i\omega}},$$

$$M' = \prod_{\omega \in T} \prod_{i=1}^{n} A'_{i\omega}.$$

Let $F = a l_1 \cdots l_n, n \geqslant 3$, where $l_i = X - \alpha_i Y$. By Theorem 4.3.3, F is equivalent to a binary form F^* such that

$$H^*(F^*) \leqslant (e^{n(n+1)\#S} |a|_S^2 R^{n/[L:K]} M^{2/[L:K]})^{(n-1)/(n-2)}. \qquad (5.2.5)$$

We now state our important proposition.

Proposition 5.2.2. *Suppose F is primitive. Then there is an admissible tuple $\mathbb{A}$ such that*

(i) $\prod_{\omega\in T}\max_{1\leqslant i<j\leqslant n}\frac{|\det(l_i,l_j)|_\omega}{A_{i\omega}A_{j\omega}}\leqslant 1$;

(ii) $(\prod_{i=1}^{n}\prod_{\omega\in T}A_{i\omega})^{\frac{2(n-1)}{n-2}}\leqslant |a|_T^{-\frac{2(n-1)}{n-2}}e^{[L:K](\#S-1)5(n-1)}|D(F)|_T^{20+\frac{1}{n}}$.

Before prove this proposition, we sketch the rough idea behind it.

Without loss of generality, let us assume for the moment that we have to deal with only one absolute value, simply denoted by $|\cdot|$. We are aiming at minimizing $M^2R^n=(A_1\cdots A_n)^2\big(\max_{1\leqslant i<j\leqslant n}\frac{|\det(l_i,l_j)|}{A_iA_j}\big)^n$. By replacing A_i by λA_i for $i=1,\ldots,n$, we may assume $\max_{1\leqslant i<j\leqslant n}\frac{|\det(l_i,l_j)|}{A_iA_j}=1$. So we aim at minimizing $A_1\cdots A_n$ subject to $\max_{1\leqslant i<j\leqslant n}\frac{|\det(l_i,l_j)|}{A_iA_j}=1$.

By taking logarithms this translates into a linear programming problem. Let $x_i=\log A_i,\delta_{ij}=\log|\Delta_{ij}|$. We want to minimize $x_1+\cdots+x_n$ subject to $\max_{1\leqslant i<j\leqslant n}(\delta_{ij}-x_i-x_j)\leqslant 0$, which is to say, $x_i+x_j\geqslant\delta_{ij}$ for all i,j with $1\leqslant i<j\leqslant n$.

We also have some conditions

$$\delta_{pq}+\delta_{ij}\leqslant\max(\delta_{pi}+\delta_{qj},\delta_{pj}+\delta_{qi})$$

for all distinct i,j,p,q by (5.2.1), and by Mason's result,

$$\max(\delta_{ij}+\delta_{pq},\delta_{iq}+\delta_{jp})\leqslant\delta_{ij}+\delta_{pq}+\delta_{iq}+\delta_{jp}+\delta_{ip}+\delta_{jq}+C$$

for all distinct i,j,p,q.

We want to estimate $x_1+\cdots+x_n$ in terms of $\sum_{1\leqslant i<j\leqslant n}\delta_{ij}$ because $D(F)=\prod_{1\leqslant i<j\leqslant n}|\Delta_{ij}|^2=\exp\big(2\sum_{1\leqslant i<j\leqslant n}\delta_{ij}\big)$.

Our idea is as follows. Fix p,q and let $x_p^{(pq)}=\frac{1}{2}\delta_{pq}+z^{(pq)},x_q^{(pq)}=\frac{1}{2}\delta_{pq}-z^{(pq)}$, where $z^{(pq)}$ will be determined later. Then $x_p^{(pq)}+x_q^{(pq)}=\delta_{pq}$.

We need $x_i^{(pq)} + x_p^{(pq)} \geqslant \delta_{ip}, x_i^{(pq)} + x_q^{(pq)} \geqslant \delta_{iq}$

So we take $x_i^{(pq)} = \max(\delta_{ip} - x_p^{(pq)}, \delta_{iq} - x_q^{(pq)}) = \max(\delta_{ip} - \frac{1}{2}\delta_{pq} - z^{(pq)}, \delta_{iq} - \frac{1}{2}\delta_{pq} + z^{(pq)})$ for $i = 1, \ldots, n, i \neq p, q$.

Thus $x_i^{(pq)} + x_j^{(pq)} \geqslant \delta_{ip} - \frac{1}{2}\delta_{pq} - z^{(pq)} + \delta_{jq} - \frac{1}{2}\delta_{pq} + z^{(pq)} = \delta_{ip} + \delta_{jq} - \delta_{pq} \geqslant \delta_{ij}$.

Now appropriate choices of $z^{(pq)}$ $(1 \leqslant p < q \leqslant n)$ and $x_i = \frac{2}{n(n-1)} \sum_{1 \leqslant p < q \leqslant n} x_i^{(pq)}$ $(i = 1, \ldots, n)$ will give a nearby solution.

Lemma 5.2.3. *If F is primitive, then $M'^{1/[L:K]} = |a|_S M^{1/[L:K]}$.*

Proof. We deduce that

$$\left(\frac{M'}{M}\right)^{1/[L:K]} = \left(\prod_{i=1}^{n} \prod_{\omega} \xi_{i\omega}\right)^{-1/[L:K]} = \left(\prod_{i=1}^{n} \prod_{\omega \notin T} \xi_{i\omega}\right)^{1/[L:K]}.$$

By Gauss' Lemma, we have

$$1 = |F|_\omega = |a|_\omega \prod_{i=1}^{n} \max(|\alpha_i|_\omega, |\beta|_\omega) = |a|_\omega \prod_{i=1}^{n} \xi_{i\omega}$$

for $\omega \notin T$, hence

$$\left(\frac{M'}{M}\right)^{1/[L:K]} = \left(\prod_{\omega \notin T} \frac{1}{|a|_\omega}\right)^{1/[L:K]} = \left(\prod_{\omega \in T} |a|_\omega\right)^{1/[L:K]} = |a|_S.$$

□

Hence we have $H^*(F^*) \leqslant (e^{n(n+1)\#S}(M'^2 R^n)^{1/[L:K]})^{(n-1)/(n-2)}$.

We can rewrite R as $R = \prod_{\omega \in T} \max_{1 \leqslant i < j \leqslant n} \frac{\theta_{ij\omega}}{A'_{i\omega} A'_{j\omega}}$. Now it is clear that Proposition 5.2.2 is equivalent to the following:

Proposition 5.2.4. *Suppose F is primitive. Then there is an admissible tuple $\mathbb{A}' = (A'_{i\omega} : \omega \in T, i = 1, \ldots, n)$ such that*

(i) $\prod_{\omega \in T} \max_{1 \leqslant i < j \leqslant n} \frac{\theta_{ij\omega}}{A'_{i\omega} A'_{j\omega}} \leqslant 1$;

(ii) $$\left(\prod_{i=1}^{n}\prod_{\omega\in T} A'_{i\omega}\right)^{\frac{2(n-1)}{n-2}} \leqslant e^{[L:K](\#S-1)5(n-1)}|D(F)|_T^{20+\frac{1}{n}}.$$

We prove (i) first. For the proof of (ii), we need some further preparations.

Proof of (i). Fix two distinct elements $p, q \in \{1, \ldots, n\}$ and $\omega \in T$.

Define

$$\Phi_{pq\omega}(x) = \prod_{h\neq p,q} \max(\theta_{ph\omega}e^{-x}, \theta_{qh\omega}e^{x}).$$

This function is continuous over the reals $\mathbb{R}$ and goes to infinity as x tends to $\pm\infty$. Hence, $\Phi_{pq\omega}$ assumes a minimum; let $x_{pq\omega}$ be the smallest real number at which $\Phi_{pq\omega}(x)$ assumes this minimum.

Put

$$A_{p\omega}^{'(pq)} = \theta_{pq\omega}^{\frac{1}{2}} e^{x_{pq\omega}},$$
$$A_{q\omega}^{'(pq)} = \theta_{pq\omega}^{\frac{1}{2}} e^{-x_{pq\omega}},$$
$$A_{h\omega}^{'(pq)} = \theta_{pq\omega}^{-\frac{1}{2}} \max(\theta_{ph\omega}e^{-x_{pq\omega}}, \theta_{qh\omega}e^{x_{pq\omega}}) \text{ for } h \neq p, q.$$

Next, put $A'_{i\omega} = \{\prod_{p\neq q} A_{i\omega}^{'(pq)}\}^{\frac{1}{n(n-1)}}$ for $i = 1, \ldots, n$, where the product is taken over all distinct pairs (p, q) with $1 \leqslant p, q \leqslant n$ and $p \neq q$. Finally, put $\mathbb{A}' = (A'_{i\omega} : i = 1, \ldots, n, \omega \in T)$.

We claim that this $\mathbb{A}'$ is admissible.

For every $\sigma \in \mathrm{Gal}(L/K), \omega \in T$ and $p, q \in \{1, \ldots, n\}, p \neq q$, we have

$$0 < \theta_{\sigma(i),\sigma(j),\omega} = \frac{|\Delta_{\sigma(i),\sigma(j)}|_\omega}{\xi_{\sigma(i)\omega}\xi_{\sigma(j)\omega}} = \frac{|\sigma(\Delta_{ij})|_\omega}{\xi_{\sigma(i)\omega}\xi_{\sigma(j)\omega}} = \frac{\Delta_{ij\omega_\sigma}}{\xi_{i\omega_\sigma}\xi_{j\omega_\sigma}} = \theta_{ij\omega_\sigma}.$$

From this, we deduce that

$$\begin{aligned}\Phi_{\sigma(p),\sigma(q),\omega}(x) &= \prod_{h\neq p,q} \max(\theta_{\sigma(p),\sigma(h),\omega}e^{-x}, \theta_{\sigma(q),\sigma(h),\omega}e^{x}) \\ &= \prod_{h\neq p,q} \max(\theta_{ph\omega_\sigma}e^{-x}, \theta_{qh\omega_\sigma}e^{x}) \\ &= \Phi_{pq\omega_\sigma}(x).\end{aligned}$$

Therefore, $x_{\sigma(p),\sigma(q),\omega} = x_{pq\omega_\sigma}$.

So we get $A^{'(\sigma(p)\sigma(q))}_{\sigma(h)\omega} = A^{'(pq)}_{h\omega_\sigma}$ for $h = 1,\dots,n, p,q \in \{1,\dots,n\}, p \neq q$ and hence $A'_{\sigma(i)\omega} = A'_{i\omega_\sigma}$. This shows that $\mathbb{A}'$ is admissible. Notice that

$$\theta_{pq\omega} = A^{'(pq)}_{p\omega} A^{'(pq)}_{q\omega},$$
$$\theta_{ph\omega} \leqslant A^{'(pq)}_{p\omega} A^{'(pq)}_{h\omega}, \theta_{qh\omega} \leqslant A^{'(pq)}_{q\omega} A^{'(pq)}_{h\omega} \text{ for } h \neq p,q.$$

Further, by $\Delta_{pq}\Delta_{ij} = \Delta_{pi}\Delta_{qj} - \Delta_{pj}\Delta_{qi}$, we have

$$\theta_{pq\omega}\theta_{ij\omega} \leqslant A^{'(pq)}_{p\omega} A^{'(pq)}_{q\omega} A^{'(pq)}_{i\omega} A^{'(pq)}_{j\omega},$$

hence

$$\theta_{ij\omega} \leqslant A^{'(pq)}_{i\omega} A^{'(pq)}_{j\omega}.$$

By taking the geometric means over all pairs p,q we get $\theta_{ij\omega} \leqslant A'_{i\omega}A'_{j\omega}$ for $\omega \in T$. This proves (i). □

We proceed to prove (ii). This will be much more involved, and requires some extra results.

For $p,q \in \{1,\dots,n\}, p \neq q, \omega \in T$, set

$$\phi_{pq\omega} = \Phi_{pq\omega}(x_{pq\omega}),\ \phi_{pq} = \prod_{\omega\in T} \phi_{pq\omega}.$$

We have

$$\begin{aligned} \prod_{h=1}^{n} A^{'(pq)}_{h\omega} &= \theta_{pq\omega} \prod_{h\neq p,q} \theta_{pq\omega}^{-\frac{1}{2}} \max(\theta_{ph\omega}e^{-x_{pq\omega}}, \theta_{qh\omega}e^{x_{pq\omega}}) \\ &= \theta_{pq\omega}^{-\frac{n}{2}+2}\phi_{pq\omega} \\ &= \theta_{pq\omega}^{-\frac{n}{2}+2}\phi_{pq\omega}, \end{aligned}$$

and

$$\prod_{p\neq q}\prod_{\omega\in T} \theta_{pq\omega} = |D(F)|_T,$$

since F is primitive by Lemma 5.2.1 (iii).

Hence

$$\begin{aligned} M' &= \prod_{i=1}^{n}\prod_{\omega\in T} A'_{i\omega} \\ &= \left(\prod_{p\neq q}\prod_{\omega\in T}\prod_{h=1}^{n} A'^{(pq)}_{h\omega}\right)^{\frac{1}{n(n-1)}} \\ &= \left(\prod_{p\neq q}\prod_{\omega\in T} \theta_{pq\omega}^{-\frac{n}{2}+2}\phi_{pq\omega}\right)^{\frac{1}{n(n-1)}} \\ &= \left(|D(F)|_T^{-\frac{n}{2}+2}\prod_{p\neq q}\phi_{pq}\right)^{\frac{1}{n(n-1)}}. \end{aligned} \tag{5.2.6}$$

We estimate $\prod_{p\neq q}\phi_{pq}$. To this end, we need the following notation and a lemma.

For a fixed pair $\{p,q\}$ with $p,q\in\{1,\dots,n\}, p\neq q$, put $W_{pq}=\{1,\dots,n\}\backslash\{p,q\}$. For $J\subset W_{pq},\omega\in T$, define the quantities $M_\omega(J)$ as follows.

If $\#J=0, J=\emptyset$, put $M_\omega(J)=1$;

If $\#J=1, J=\{j\}$, put $M_\omega(J)=\sqrt{\theta_{pj\omega}\theta_{qj\omega}}$;

If $\#J\geqslant 2$, put

$$M_\omega(J)=\max\left\{\prod_{h\in I}\theta_{ph\omega}\prod_{h\in J\backslash I}\theta_{qh\omega} : I\subset J, \#I=\frac{1}{2}\#J\right\} \text{ if } \#J \text{ is even,}$$

$$M_\omega(J)=\sqrt{M_{1\omega}(J)M_{2\omega}(J)} \text{ if } \#J \text{ is odd,}$$

where

$$M_{1\omega(J)}=\max\left\{\prod_{h\in I}\theta_{ph\omega}\prod_{h\in J\backslash I}\theta_{qh\omega} : I\subset J, \#I=\tfrac{1}{2}(\#J+1)\right\},$$

$$M_{2\omega(J)}=\max\left\{\prod_{h\in I}\theta_{ph\omega}\prod_{h\in J\backslash I}\theta_{qh\omega} : I\subset J, \#I=\tfrac{1}{2}(\#J-1)\right\}.$$

Finally, put $M(J) = \prod_{\omega\in T} M_\omega(J)$.

Lemma 5.2.5. $\phi_{pq} = M(W_{pq})$.

Proof. This is taken from [9], which deals with number fields. But over function fields, the argument is the same. We repeat it again here.

It suffices to prove $\phi_{pq\omega} = M_\omega(W_{pq})$ for every $\omega \in T$.

Take $f(x) = \log \Phi_{pq\omega}(x) = \sum_{h\in W_{pq}} \max(f_{ph} - x, f_{qh} + x)$ where $f_{ph} = \log \theta_{ph\omega}, f_{qh} = \log \theta_{qh\omega}$.

We can express $f(x)$ as

$$f(x) = \max\{C_0-(n-2)x, C_1-(n-2)x, \cdots, C_{n-3}+(n-4)x, C_{n-2}+(n-2)x\},$$

where

$$\begin{aligned} C_s &= \max\Big\{\sum_{h\in I} f_{ph} + \sum_{h\in W_{pq}\setminus I} f_{qh} : I \subset W_{pq}, \#I = n-2-s\Big\} \\ &= \log\max\Big\{\prod_{h\in I} \theta_{ph\omega} \prod_{h\in W_{pq}\setminus I} \theta_{qh\omega} : I \subset W_{pq}, \#I = n-2-s\Big\}, \end{aligned}$$

for $s = 0, \ldots, n-2$.

Let

$$I_s = \{x \in \mathbb{R} : f(x) = C_s - (n-2-2s)x\}\,(s = 0, \ldots, n-2).$$

We first show that I_s is nonempty.

Clearly, $I_0 = \{x \in \mathbb{R} : f(x) = C_0 - (n-2)x\} \neq \emptyset$, and similarly $I_{n-2} \neq \emptyset$.

We show that $I_s \neq \emptyset$ for $s \in \{1, \ldots, n-3\}$. Choose $I \subset W_{pq}$, with $\#I = n-2-s$ such that

$$C_s = \sum_{h\in I} f_{ph} + \sum_{h\in W_{pq}\setminus I} f_{qh}.$$

Take $i \in I, j \in W_{pq} \backslash I$ and consider the same sum but with $I' = \{j\} \cup I \backslash \{i\}$ instead of I. This sum is at most C_s, and so $f_{pi} + f_{qj} \geqslant f_{pj} + f_{qi}$ and hence $f_{pi} - f_{qi} \geqslant f_{pj} - f_{qj}$. So there exists $x \in \mathbb{R}$ such that

$$\max_{j \in I^c} \frac{1}{2}(f_{pj} - f_{qj}) \leqslant x \leqslant \min_{i \in I} \frac{1}{2}(f_{pi} - f_{qi}).$$

For this specific x, we have $f_{pi} - x \geqslant f_{qi} + x$ and $f_{pj} - x \leqslant f_{qj} + x$ for any $i \in I, j \in I^c$, and hence

$$f(x) = \sum_{i \in I} f_{pi} + \sum_{j \in I^c} f_{qj} - (n - 2 - 2s)x = C_s - (n - s - 2s)x.$$

So indeed, $I_s \neq \emptyset$.

Now we may use Lemma 12 of [9] to conclude that

$$\log \phi_{pq\omega} = \min\{f(x) : x \in \mathbb{R}\} = C_{\frac{1}{2}(n-2)} = \log M_\omega(W_{pq})$$

when n is even; and similarly

$$\log \phi_{pq\omega} = \frac{1}{2}(C_{\frac{1}{2}(n-1)} + C_{\frac{1}{2}(n-3)}) = \log M_\omega(W_{pq})$$

when n is odd. This completes the proof.

□

Next, we estimate $M(J)$ from above by induction on $\#J$, and eventually deduce an upper bound for $M(W_{pq}) = \phi_{pq}$.

Put $\Theta_p(J) = \Theta_q(J) = 1, D(J) = 1$ if $J = \emptyset$; $D(J) = 1$ if $\#J = 1$; and

$$\begin{cases} \Theta_p(J) = \prod_{\omega \in T} \prod_{h \in J} \theta_{ph\omega}, \\ \Theta_q(J) = \prod_{\omega \in T} \prod_{h \in J} \theta_{qh\omega}, \\ D(J) = \prod_{\omega \in T} \prod_{h \neq l \in J} \theta_{hl\omega} \end{cases} \tag{5.2.7}$$

if $\#J \geqslant 2$.

For $s \geqslant 2$, let

$$d(s) = \begin{cases} \frac{1}{s} & s \text{ even} \\ \frac{s}{s^2-1} & s \text{ odd,} \end{cases}$$

and

$$a(0) = 0, a(1) = 0, a(s) = a(s-2) + 1 + 4(s-2)d(s), s \geqslant 2;$$

$$b(0) = 0, b(1) = \frac{1}{2}, b(s) = \frac{s-2}{s}b(s-2) + \frac{2}{s} + \frac{4(s-2)}{s}d(s);$$

$$c(0) = 0, c(1) = 0, c(s) = \frac{(s-2)(s-3)}{s(s-1)}c(s-2) + \frac{1+4(s-2)d(s)}{s(s-1)}.$$

It is not difficult to show, by a straightforward computation

$$a(s) \leqslant 1 + \frac{5}{2}(s-2), b(s) \leqslant 3, c(s) < \frac{5}{2s-2}.$$

Take $c(s)$ as an example. We have

$$s(s-1)c(s) = (s-2)(s-3)c(s-2) + 1 + 4d(s)(s-2).$$

When s is even, we have

$$\begin{aligned} s(s-1)c(s) &= \sum_{h=1}^{s/2} 1 + 4(2h-2)d(2h) \\ &= \sum_{h=1}^{s/2} (5 - \frac{4}{h}) \\ &= \frac{5s}{2} - 4\sum_{h=1}^{s/2} \frac{1}{h} \\ &< \frac{5s}{2}, \end{aligned}$$

hence $c(s) < \frac{5}{2s-2}$.

When s is odd, we derive $c(s) < \frac{5}{2s-2}$ by a similar computation.

Lemma 5.2.6. *Put* $C_{pqij} = e^{\max(2g_{L_{pqij}}-2+\#T_{pqij},0)[L:L_{pqij}]}$. *We have*

$$M(J) \leqslant C(J)(\prod_{\omega\in T}\theta_{pq\omega})^{a(s)}(\Theta_p(J)\Theta_q(J))^{b(s)}D(J)^{c(s)},$$

where $J\subset W_{pq}, s=\#J$ *and*

$$C(J) = \begin{cases} 1 & \text{for } s=0,1; \\ \Big(\prod_{i\neq j\in J} C_{pqij}\Big)^{c(s)} & \text{for } s\geqslant 2. \end{cases} \tag{5.2.8}$$

where the notation $\prod_{i\neq j\in J}$ *means that the product is taken over all ordered pairs* (i,j) *with* $i,j\in J$ *and* $i\neq j$.

Proof. If $s=0$, then $M(J)=1$, and if $s=1$, $J=\{j\}$, then $M(J)=\prod_{\omega\in T}\sqrt{\theta_{pj\omega}\theta_{qj\omega}}$. So in these cases, Lemma 5.2.6 is trivial.

Let $s\geqslant 2$ and assume the assertion is true for sets J of cardinality strictly smaller than s. Let $J\subset W_{pq}, \#J=s$. Fix $i,j\in J$ with $i\neq j$, let $J_{ij}=J\backslash\{i,j\}$ and fix any $\omega\in T$.

We distinguish the cases s even and s odd.

First suppose s is even.

Let $I\subset J, \#I=\frac{1}{2}s\geqslant 1$, $g(I)=\prod_{h\in I}\theta_{ph\omega}\cdot\prod_{h\in J\backslash I}\theta_{qh\omega}$.

If $i\in I, j\in J\backslash I$, then since

$$M_\omega(J_{ij}) \geqslant \prod_{h\in K\backslash\{i\}}\theta_{ph\omega}\prod_{h\in J\backslash\{I\cup\{j\}\}}\theta_{qh\omega},$$

we have

$$g(I)\leqslant \theta_{pi\omega}\theta_{qj\omega}M_\omega(J_{ij}).$$

Hence

$$g(I)\leqslant \max(\theta_{pi\omega}\theta_{qj\omega},\theta_{pj\omega}\theta_{qi\omega})M_\omega(J_{ij}). \tag{5.2.9}$$

This is also true if $j\in I, i\in J\backslash I$.

If $i, j \in I$, then pick $l \in J\backslash I$ such that $\frac{\theta_{pj\omega}\theta_{ql\omega}}{\theta_{pl\omega}\theta_{qj\omega}}$ is minimal for all $l \in J\backslash I$. Since $J\backslash I \subset J_{ij}$, we have

$$\begin{aligned}
\frac{\theta_{pj\omega}\theta_{ql\omega}}{\theta_{pl\omega}\theta_{qj\omega}} &\leqslant \left(\prod_{h\in J\backslash I} \max(1, \frac{\theta_{pj\omega}\theta_{qh\omega}}{\theta_{ph\omega}\theta_{qj\omega}})\right)^{2/s} \\
&\leqslant \left(\prod_{h\in J_{ij}} \max(1, \frac{\theta_{pj\omega}\theta_{qh\omega}}{\theta_{ph\omega}\theta_{qj\omega}})\right)^{2/s}.
\end{aligned}$$

Take $I' = I \cup \{l\}\backslash\{j\}$. Then $\#I' = \frac{s}{2}, i \in I', j \in J\backslash I'$. From (5.2.9), we get

$$\begin{aligned}
g(I) &= \frac{\theta_{pj\omega}\theta_{ql\omega}}{\theta_{pl\omega}\theta_{qj\omega}} g(I') \\
&\leqslant \max(\theta_{pi\omega}\theta_{qj\omega}, \theta_{pj\omega}\theta_{qi\omega}) M_\omega(J_{ij}) \left(\prod_{h\in J_{ij}} \max(1, \frac{\theta_{pj\omega}\theta_{qh\omega}}{\theta_{ph\omega}\theta_{qj\omega}})\right)^{2/s}.
\end{aligned}$$

Similarly, we have

$$g(I) \leqslant \max(\theta_{pi\omega}\theta_{qj\omega}, \theta_{pj\omega}\theta_{qi\omega}) M_\omega(J_{ij}) \left(\prod_{h\in J_{ij}} \max(1, \frac{\theta_{pi\omega}\theta_{qh\omega}}{\theta_{ph\omega}\theta_{qi\omega}})\right)^{2/s}.$$

So we get

$$g(I) \leqslant \left(\prod_{h\in J_{ij}} \max(1, \frac{\theta_{pi\omega}\theta_{qh\omega}}{\theta_{ph\omega}\theta_{qi\omega}}) \max(1, \frac{\theta_{pj\omega}\theta_{qh\omega}}{\theta_{ph\omega}\theta_{qj\omega}})\right)^{1/s} \max(\theta_{pi\omega}\theta_{qj\omega}, \theta_{pj\omega}\theta_{qi\omega}) M_\omega(J_{ij}).$$

If $i, j \in J\backslash I$, by interchanging $I, J\backslash I$ and also p, q, we have

$$g(I) \leqslant \left(\prod_{h\in J_{ij}} \max(1, \frac{\theta_{ph\omega}\theta_{qi\omega}}{\theta_{pi\omega}\theta_{qh\omega}}) \max(1, \frac{\theta_{ph\omega}\theta_{qj\omega}}{\theta_{pj\omega}\theta_{qh\omega}})\right)^{1/s} \max(\theta_{pi\omega}\theta_{qj\omega}, \theta_{pj\omega}\theta_{qi\omega}) M_\omega(J_{ij}).$$

This gives altogether

$$M_\omega(J) = \max g(I) \leqslant H_\omega^{\frac{1}{s}} \max(\theta_{pi\omega}\theta_{qj\omega}, \theta_{pj\omega}\theta_{qi\omega}) M_\omega(J_{ij}),$$

where

$$H_\omega = \prod_{h\in J_{ij}} \max\left(1, \frac{\theta_{pi\omega}\theta_{qh\omega}}{\theta_{ph\omega}\theta_{qi\omega}}\right) \max\left(1, \frac{\theta_{pj\omega}\theta_{qh\omega}}{\theta_{ph\omega}\theta_{qj\omega}}\right) \max\left(1, \frac{\theta_{ph\omega}\theta_{qi\omega}}{\theta_{pi\omega}\theta_{qh\omega}}\right) \max\left(1, \frac{\theta_{ph\omega}\theta_{qj\omega}}{\theta_{pj\omega}\theta_{qh\omega}}\right).$$

If $s \geqslant 3$ is odd, the argument is similar.

Finally, we have

$$M_\omega(J) \leqslant H_\omega^{d(s)} \max(\theta_{pi\omega}\theta_{qj\omega}, \theta_{pj\omega}\theta_{qi\omega}) M_\omega(J_{ij}).$$

Let $H = \prod_{\omega\in T} H_\omega$. Note that this quantity depends on J_{ij}. We have

$$M(J) \leqslant H^{d(s)} \prod_{\omega\in T} \max(\theta_{pi\omega}\theta_{qj\omega}, \theta_{pj\omega}\theta_{qi\omega}, \theta_{pq\omega}\theta_{ij\omega}) M(J_{ij}). \qquad (5.2.10)$$

By Lemma 5.2.1, we have

$$\prod_{\omega\in T} \max(\theta_{pq\omega}\theta_{ij\omega}, \theta_{pi\omega}\theta_{qj\omega}, \theta_{pj\omega}\theta_{qi\omega}) \leqslant C_{pqij} \prod_{\omega\in T} \theta_{pq\omega}\theta_{ij\omega}\theta_{pi\omega}\theta_{qj\omega}\theta_{pj\omega}\theta_{qi\omega},$$

where $L_{pqij} = K(\alpha_i, \alpha_j, \alpha_p, \alpha_q)$ and T_{pqij} is the set of valuations in L_{pqij} above those in T.

Using Lemma 5.2.1 again, we obtain

$$\begin{aligned}
H &= \prod_{h\in J_{ij}} \prod_{\omega\in T} \left(\frac{\max(\theta_{pi\omega}\theta_{qh\omega}, \theta_{ph\omega}\theta_{qi\omega})^2 \max(\theta_{pj\omega}\theta_{qh\omega}, \theta_{ph\omega}\theta_{qj\omega})^2}{\theta_{ph\omega}^2\theta_{qh\omega}^2\theta_{qi\omega}\theta_{pi\omega}\theta_{qj\omega}\theta_{pj\omega}} \right) \\
&\leqslant \prod_{h\in J_{ij}} C_{pqih}^2 C_{pqjh}^2 \times \\
&\quad \times \prod_{h\in J_{ij}} \prod_{\omega\in T} \left(\frac{(\theta_{pi\omega}\theta_{qh\omega}\theta_{ph\omega}\theta_{qi\omega}\theta_{pq\omega}\theta_{ih\omega})^2 (\theta_{pj\omega}\theta_{qh\omega}\theta_{ph\omega}\theta_{qj\omega}\theta_{pq\omega}\theta_{jh\omega})^2}{\theta_{ph\omega}^2\theta_{qh\omega}^2\theta_{qi\omega}\theta_{pi\omega}\theta_{qj\omega}\theta_{pj\omega}} \right) \\
&= \prod_{h\in J_{ij}} \left(C_{pqih}^2 C_{pqjh}^2 \right) \cdot \prod_{h\in J_{ij}} \prod_{\omega\in T} \left(\theta_{pq\omega}^4 \theta_{qh\omega}^2 \theta_{ph\omega}^2 \theta_{ih\omega}^2 \theta_{jh\omega}^2 \theta_{pi\omega}\theta_{qi\omega}\theta_{pj\omega}\theta_{qj\omega} \right).
\end{aligned}$$

By substituting these estimates into (5.2.10), we get

$$
\begin{aligned}
M(J) \leqslant\ & \Big(\prod_{h\in J_{ij}} C_{pqih}^2 C_{pqjh}^2\Big)^{d(s)} \times \\
\times\ & \left(\prod_{h\in J_{ij}}\prod_{\omega\in T} \theta_{pq\omega}^4 \theta_{qh\omega}^2 \theta_{ph\omega}^2 \theta_{ih\omega}^2 \theta_{jh\omega}^2 \theta_{pi\omega}\theta_{qi\omega}\theta_{pj\omega}\theta_{qj\omega}\right)^{d(s)} \times \\
\times\ & C_{pqij}\cdot \left(\prod_{\omega\in T} \theta_{pq\omega}\theta_{ij\omega}\theta_{pi\omega}\theta_{qj\omega}\theta_{pj\omega}\theta_{qi\omega}\right) M(J_{ij}).
\end{aligned}
$$

This inequality is valid for each pair $i,j \in J$ with $i \neq j$. By taking the geometric means over these pairs we get

$$
\begin{aligned}
M(J) \leqslant\ & \left(\prod_{i\neq j\in J}\prod_{h\in J_{ij}} C_{pqih}^2 C_{pqjh}^2\right)^{\frac{d(s)}{s(s-1)}} \times \\
\times\ & \left(\prod_{i\neq j\in J}\prod_{h\in J_{ij}}\prod_{\omega\in T} \theta_{pq\omega}^4 \theta_{qh\omega}^2 \theta_{ph\omega}^2 \theta_{ih\omega}^2 \theta_{jh\omega}^2 \theta_{pi\omega}\theta_{qi\omega}\theta_{pj\omega}\theta_{qj\omega}\right)^{\frac{d(s)}{s(s-1)}} \times \\
\times\ & \left(\prod_{i\neq j\in J} C_{pqij}\Big(\prod_{\omega\in T} \theta_{pq\omega}\theta_{ij\omega}\theta_{pi\omega}\theta_{qj\omega}\theta_{pj\omega}\theta_{qi\omega}\Big) M(J_{ij})\right)^{\frac{1}{s(s-1)}}.
\end{aligned}
$$

By inserting the upper bound for $M(J_{ij})$ following from the induction hypothesis, we get

$$\begin{aligned} M(J) &\leqslant \left(\prod_{i \neq j \in J} \prod_{h \in J_{ij}} C_{pqih}^2 C_{pqjh}^2 \right)^{\frac{d(s)}{s(s-1)}} \times \\ &\times \left(\prod_{i \neq j \in J} \prod_{h \in J_{ij}} \prod_{\omega \in T} \theta_{pq\omega}^4 \theta_{qh\omega}^2 \theta_{ph\omega}^2 \theta_{ih\omega}^2 \theta_{jh\omega}^2 \theta_{pi\omega} \theta_{qi\omega} \theta_{pj\omega} \theta_{qj\omega} \right)^{\frac{d(s)}{s(s-1)}} \times \\ &\times \left(\prod_{i \neq j \in J} C_{pqij} \cdot \Big(\prod_{\omega \in T} \theta_{pq\omega} \theta_{ij\omega} \theta_{pi\omega} \theta_{qj\omega} \theta_{pj\omega} \theta_{qi\omega} \Big) \right)^{\frac{1}{s(s-1)}} \times \\ &\times \left(\prod_{i \neq j \in J} C(J_{ij}) (\prod_{\omega \in T} \theta_{pq\omega})^{a(s-2)} (\Theta_p(J_{ij}) \Theta_q(J_{ij}))^{b(s-2)} D(J_{ij})^{c(s-2)} \right)^{\frac{1}{s(s-1)}}, \end{aligned}$$

where $C(J_{ij})$ is defined by (5.2.8) with J_{ij} replacing J.

Put

$$C'(J) = \Big(\prod_{i \neq j \in J} C(J_{ij}) \Big)^{\frac{1}{s(s-1)}} \Big(\prod_{i \neq j \in J} C_{pqij} \Big)^{\frac{1+4d(s)(s-2)}{s(s-1)}}.$$

Then by the previous inequality, we get

$$\begin{aligned} M(J) &\leqslant C'(J) \Big(\prod_{\omega \in T} \theta_{pq\omega} \Big)^{a(s)} \Big(\prod_{i \neq j} \prod_{h \in J_{ij}} \prod_{\omega \in T} \theta_{qh\omega} \theta_{ph\omega} \theta_{ih\omega} \theta_{jh\omega} \Big)^{\frac{2d(s)}{s(s-1)}} \times \\ &\times \Big(\prod_{i \neq j \in J} \prod_{\omega \in T} \prod_{h \in J_{ij}} \theta_{ph\omega} \theta_{qh\omega} \Big)^{\frac{b(s-2)}{s(s-1)}} \Big(\prod_{i \neq j} D(J_{ij}) \Big)^{\frac{c(s-2)}{s(s-1)}} \times \\ &\times \Big(\prod_{i \neq j \in J} \prod_{\omega \in T} \theta_{ij\omega} \Big)^{\frac{1}{s(s-1)}} \Big(\prod_{i \neq j} \prod_{\omega \in T} \theta_{pi\omega} \theta_{qi\omega} \theta_{pj\omega} \theta_{qj\omega} \Big)^{\frac{(s-2)d(s)+1}{s(s-1)}}, \end{aligned}$$

hence

$$\begin{aligned} M(J) &= C'(J)\Big(\prod_{\omega\in T}\theta_{pq\omega}\Big)^{a(s)}(\Theta_p(J)\Theta_q(J))^{2d(s)}\times \\ &\times\Big(\prod_{i\neq j}\prod_{h\in J_{ij}}\prod_{\omega\in T}\theta_{ih\omega}\theta_{jh\omega}\Big)^{\frac{2d(s)}{s(s-1)}}\Big(\prod_{i\neq j}\prod_{\omega\in T}\prod_{h\in J_{ij}}\theta_{ph\omega}\theta_{qh\omega}\Big)^{\frac{b(s-2)}{s(s-1)}}\times \\ &\times\Big(\prod_{i\neq j}D(J_{ij})\Big)^{\frac{c(s-2)}{s(s-1)}}\Big(\prod_{i\neq j\in J}\prod_{\omega\in T}\theta_{ij\omega}\Big)^{\frac{1}{s(s-1)}}\times \\ &\times\Big(\prod_{i\neq j}\prod_{\omega\in T}\theta_{pi\omega}\theta_{qi\omega}\theta_{pj\omega}\theta_{qj\omega}\Big)^{\frac{(s-4)d(s)+1}{s(s-1)}}. \end{aligned}$$

Now by (5.2.8) with J_{ij} replacing J, we have

$$C'(J)=\Big(\prod_{i\neq j\in J}\prod_{h\neq l\in J_{ij}}C_{pqhl}\Big)^{\frac{c(s-2)}{s(s-1)}}\Big(\prod_{i\neq j\in J}C_{pqij}\Big)^{\frac{1+4d(s)(s-2)}{s(s-1)}},$$

hence

$$\begin{aligned} C'(J) &= \Big(\prod_{i\neq j\in J}C_{pqij}\Big)^{((s-2)(s-3)c(s-2)+1+4(s-2)d(s))/(s(s-1))} \\ &= \Big(\prod_{i\neq j\in J}C_{pqij}\Big)^{c(s)} \\ &= C(J). \end{aligned} \tag{5.2.11}$$

Now combining the just established upper bound for $M(J)$ with (5.2.7) and the obvious identities

$$\Big(\prod_{i\neq j}\theta_{pi\omega}\theta_{qi\omega}\theta_{pj\omega}\theta_{qj\omega}\Big)^{\frac{1}{s(s-1)}}=\Big(\prod_{h\in J}\prod_{\omega\in T}\theta_{ph\omega}\theta_{qh\omega}\Big)^{\frac{2}{s}},$$

$$\Big(\prod_{i\neq j}\prod_{\omega\in T}\prod_{h\in J_{ij}}\theta_{ph\omega}\theta_{qh\omega}\Big)^{\frac{1}{s(s-1)}}=\Big(\prod_{\omega\in T}\prod_{h\in J}\theta_{ph\omega}\theta_{qh\omega}\Big)^{\frac{s-2}{s}},$$

$$(\prod_{i\neq j}\prod_{h\in J_{ij}}\prod_{\omega\in T}\theta_{ih\omega}\theta_{jh\omega})^{\frac{1}{s(s-1)}} = (D(J))^{\frac{2(s-2)}{s(s-1)}},$$

$$(\prod_{i\neq j}\prod_{\omega\in T}\theta_{ij\omega})^{\frac{1}{s(s-1)}} = (D(J))^{\frac{1}{s(s-1)}},$$

$$(\prod_{i\neq j}D(J_{ij}))^{\frac{1}{s(s-1)}} = (D(J))^{\frac{(s-2)(s-3)}{s(s-1)}}.$$

we deduce that

$$M(J) \leqslant C(J)(\prod_{\omega\in T}\theta_{pq\omega})^{a(s)}(\Theta_p(J)\Theta_q(J))^{b(s)}D(J)^{c(s)},$$

which completes the induction step and the proof of Lemma 5.2.6. □

Proof of (ii) of Proposition 5.2.4. Now Lemma 5.2.6 with $J = W_{pq}$ and Lemma 5.2.5 give that

$$\begin{aligned}\phi_{pq} = M(W_{pq}) \quad &\leqslant \quad \Big(\prod_{i\neq j\in W_{pq}} C_{pqij}\Big)^{c(n-2)} \\ &\times \quad \Big(\prod_{\omega\in T}\theta_{pq\omega}\Big)^{a(n-2)}\Big(\Theta_p(W_{pq})\Theta_q(W_{pq})\Big)^{b(n-2)}D(W_{pq})^{c(n-2)}.\end{aligned}$$

Notice that since F is primitive, it follows from (5.2.4) that $|a|_T = \prod_{i=1}^{n}\prod_{\omega\in T}\frac{1}{\xi_{i\omega}}$, so by an easy computation, we have

$$\prod_{p\neq q}\prod_{\omega\in T}\theta_{pq\omega} = |D(F)|_T,$$

$$\prod_{p\neq q}\Theta_p(W_{pq})\Theta_q(W_{pq}) = |D(F)|_T^{2n-4},$$

$$\prod_{p\neq q} D(W_{pq}) = |D(F)|_T^{(n-2)(n-3)}.$$

Thus, we deduce that

$$\prod_{\substack{p,q\in\{1,\dots,n\}\\ p\neq q}} \phi_{pq} \ \leqslant\ U\cdot\Big(\prod_{p\neq q}\prod_{i\neq j\in W_{pq}} C_{pqij}\Big)^{c(n-2)}, \tag{5.2.12}$$

where $U = |D(F)|_T^{a(n-2)+(2n-4)b(n-2)+(n-2)(n-3)c(n-2)}$.

We need to estimate $\prod_{p\neq q}\prod_{i\neq j\in W_{pq}} C_{pqij}$ from above.

Denote the field $K(\alpha_h)$ by L_h. By (5.1.3), we have

$$\frac{2g_{pqij}-1+\#T_{pqij}}{[L_{pqij}:K]} \ \leqslant\ -1+\#S-\sum_{h\in\{p,q,i,j\}}\sum_{\nu\in S}\frac{\nu(D(F_{\overline{h}}))}{\deg F_{\overline{h}}}.$$

so we get

$$(2g_{pqij}-1+\#T_{pqij})[L:L_{pqij}] \leqslant [L:K]\Big(-1+\#S-\sum_{h\in\{p,q,i,j\}}\sum_{\nu\in S}\frac{\nu(D(F_{\overline{h}}))}{\deg F_{\overline{h}}}\Big).$$

Hence

$$\begin{aligned}
&\sum_{p\neq q}\sum_{i\neq j\in W_{pq}} \max(2g_{pqij}-2+\#T_{pqij},0)[L:L_{pqij}]\\
\leqslant\ &\sum_{p\neq q}\sum_{i\neq j\in W_{pq}} (2g_{pqij}-1+\#T_{pqij})[L:L_{pqij}]\\
\leqslant\ &[L:K]\sum_{p\neq q}\sum_{\substack{i,j\neq p,q\\ i\neq j}}\Big(-1+\#S-\sum_{h\in\{p,q,i,j\}}\sum_{\nu\in S}\frac{\nu(D(F_{\overline{h}}))}{\deg F_{\overline{h}}}\Big)\\
=\ &[L:K]\Big(n(n-1)(n-2)(n-3)(\#S-1)\\
&\qquad -4(n-1)(n-2)(n-3)\sum_{i=1}^{d}\sum_{\nu\in S}\nu(D(F_i))\Big)\\
\leqslant\ &[L:K]\Big(n(n-1)(n-2)(n-3)(\#S-1)\\
&\qquad -4(n-1)(n-2)(n-3)\sum_{\nu\in S}\nu(D(F))\Big).
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
\prod_{p\neq q}\phi_{pq} &\leqslant \left(e^{[L:K](n-1)(n-2)(n-3)(n(\#S-1)-4\sum\limits_{\nu\in S}\nu(D(F)))}\right)^{c(n-2)} \times \\
&\quad \times |D(F)|_T^{a(n-2)+(2n-4)b(n-2)+(n-2)(n-3)c(n-2)} \\
&= e^{[L:K]n(n-1)(n-2)(n-3)(\#S-1)c(n-2)} \times \\
&\quad \times |D(F)|_T^{4(n-1)(n-2)(n-3)c(n-2)+a(n-2)+(2n-4)b(n-2)+(n-2)(n-3)c(n-2)}.
\end{aligned}$$

As $a(n-2)\leqslant \frac{5n}{2}-9, b(n-2)\leqslant 3, c(n-2)<\frac{5}{2n-6}$ we conclude from (5.2.6) that

$$\begin{aligned}
M'^{\frac{2(n-1)}{n-2}} &\leqslant \left(|D(F)|_T^{-\frac{n}{2}+2}\prod_{p\neq q}\phi_{pq}\right)^{\frac{2}{n(n-2)}} \\
&\leqslant e^{[L:K](\#S-1)5(n-1)}|D(F)|_T^{20+\frac{1}{n}}.
\end{aligned}$$

This gives Proposition 5.2.4 (ii). □

5.3 Completion of the Proof of the Main Theorem

Main Theorem. *Let $F\in\mathcal{O}_S[X,Y]$ be a binary form of degree $n\geqslant 4$ with non-zero discriminant. Then F is* $\mathrm{GL}(2,\mathcal{O}_S)$*-equivalent to a binary form F^* such that*

$$H^*(F^*)\leqslant e^{(n-1)\left(\#S(n+11)-5\right)}|D(F)|_S^{20+\frac{1}{n}}.$$

Proof. When F is primitive, this follows directly from Proposition 5.2.4 and Theorem 4.3.3.

In the general case, write $F = a\tilde{F}$ such that $a \in \mathcal{O}_S$ and $\tilde{F}$ is primitive. Then there exists $\tilde{F}_1$ that is $\mathrm{GL}_n(2, \mathcal{O}_S)$-equivalent to $\tilde{F}$ such that

$$H^*(\tilde{F}_1) \leqslant e^{(n-1)\left(\#S(n+11)-5\right)} |D(\tilde{F})|_S^{20+\frac{1}{n}}.$$

Let $F_1 = a\tilde{F}_1$. Since $\tilde{F}_1$ is a binary form over $\mathcal{O}_S$, $H_S(\tilde{F}_1) \leqslant H^*(\tilde{F}_1)$. Noticing that $D(F) = a^{2n-2}D(\tilde{F})$, we deduce that

$$\begin{aligned} H_S(F_1) &= |a|_S H_S(\tilde{F}_1) \\ &\leqslant |a|_S e^{(n-1)\left(\#S(n+11)-5\right)} |D(\tilde{F})|_S^{20+\frac{1}{n}} \\ &= e^{(n-1)\left(\#S(n+11)-5\right)} |D(F)|_S^{20+\frac{1}{n}} |a|_S^{1-(2n-2)(20+1/n)} \\ &\leqslant e^{(n-1)\left(\#S(n+11)-5\right)} |D(F)|_S^{20+\frac{1}{n}}. \end{aligned}$$

By Lemma 4.2.1, there exists $u \in \mathcal{O}_S^*$ such that $H^*(uF_1) = H_S(F_1)$. Put $F^* = uF_1 = au\tilde{F}_1$. Then F^* is $\mathrm{GL}_n(2, \mathcal{O}_S)$-equivalent to F and

$$H^*(F^*) \leqslant e^{(n-1)\left(\#S(n+11)-5\right)} |D(F)|_S^{20+\frac{1}{n}},$$

as claimed. □

We need a variation of the Main Theorem. To get this, in the proof of Lemma 5.2.6 we repeat all computations but with all fields L_{pqij} replaced by L. Then we get Lemma 5.2.6 with C_{pqij} replaced by $e^{\max(2g_L-2+\#T,0)}$. This gives, instead of (5.2.12),

$$\begin{aligned} \prod_{p \neq q \in \{1,\ldots,n\}} \phi_{pq} &< e^{\max(2g_L-2+\#T,0)\binom{n}{4}c(n-2)} \times \\ &\quad \times |D(F)|_T^{a(n-2)+(2n-4)b(n-2)+(n-2)(n-3)c(n-2)} \\ &\leqslant e^{\max(2g_L-2+\#T,0)\binom{n}{4}c(n-2)} |D(F)|_T^{11n-26}. \end{aligned} \tag{5.3.1}$$

Similarly as before, this leads together with (5.2.6), to the following:

Proposition 5.3.1. *Suppose F is primitive with splitting field L. Then there is an admissible tuple $\mathbb{A}' = (A'_{i\omega} : \omega \in T, i = 1, \ldots, n)$ such that* *(i)* $\prod_{\omega\in T}\max_{1\leqslant i<j\leqslant n}\frac{\theta_{ij\omega}}{A'_{i\omega}A'_{j\omega}}\leqslant 1;$

(ii) $\left(\prod_{i=1}^{n}\prod_{\omega\in T}A'_{i\omega}\right)^{\frac{2(n-1)}{n-2}}\leqslant e^{\frac{5}{24}(n-1)\max(2g_L-2+\#T,0)}|D(F)|_T^{\frac{21}{n}}.$

Now with the same idea as in the proof of the Main Theorem, we have

Theorem 5.3.2. *Let $F\in\mathcal{O}_S[X,Y]$ be a binary form of degree $n\geqslant 4$ with non-zero discriminant. Then F is $\mathrm{GL}\,(2,\mathcal{O}_S)$-equivalent to a binary form F^* such that*

$$H^*(F^*)\leqslant\exp\left((n^2+6n-7)\#S+\frac{(5n-5)(2g_L-1)}{24[L:K]}\right)|D(F)|_S^{\frac{21}{n}}.$$

Proof. First, if F is primitive, this follows from Proposition 5.3.1 and Theorem 4.3.3 by a direct computation.

If F is not primitive, we assume that $F=a\tilde{F}$ with $a\in\mathcal{O}_S$ and $\tilde{F}$ primitive. Then there exists $\tilde{F}_1$ that is $\mathrm{GL}_n(2,\mathcal{O}_S)$-equivalent to $\tilde{F}$ such that

$$H^*(\tilde{F}_1)\leqslant\exp\left((n^2+6n-7)\#S+\frac{(5n-5)(2g_L-1)}{24[L:K]}\right)|D(\tilde{F})|_S^{\frac{21}{n}}.$$

Let $F_1=a\tilde{F}_1$. Then

$$\begin{aligned}H_S(F_1)&=|a|_S H_S(\tilde{F}_1)\\&\leqslant |a|_S\exp\left((n^2+6n-7)\#S+\frac{(5n-5)(2g_L-1)}{24[L:K]}\right)|D(\tilde{F})|_S^{\frac{21}{n}}\\&=\exp\left((n^2+6n-7)\#S+\frac{(5n-5)(2g_L-1)}{24[L:K]}\right)|D(F)|_S^{\frac{21}{n}}|a|_S^{1-42(n-1)/n}\\&\leqslant\exp\left((n^2+6n-7)\#S+\frac{(5n-5)(2g_L-1)}{24[L:K]}\right)|D(F)|_S^{\frac{21}{n}}.\end{aligned}$$

By Lemma 4.2.1, there is $u\in\mathcal{O}_S^*$ such that $H^*(uF_1)=H_S(F_1)$. Take $F^*=uF_1=au\tilde{F}_1$, then it is $\mathrm{GL}_n(2,\mathcal{O}_S)$-equivalent to $F=a\tilde{F}$ and

$$H^*(F^*)\leqslant\exp\left((n^2+6n-7)\#S+\frac{(5n-5)(2g_L-1)}{24[L:K]}\right)|D(F)|_S^{\frac{21}{n}}.$$

□

Remark 5.3.3. *This result is weaker than the Main Theorem in the sense that the constant depends on the splitting field L of F as well; however, it is apparently stronger because the exponent of $|D(F)|_S$ is much smaller and tends to zero when n goes to infinity.*

Chapter 6

Finiteness for the number of equivalence classes

It is known that if $\mathcal{O}_S$ is the ring of S-integers in an algebraic number field K, then there are only finitely many $\mathrm{GL}\,(2, \mathcal{O}_S)$-equivalence classes of binary forms with coefficients in $\mathcal{O}_S$ of given degree and given discriminant. As it turns out, the analogous statement over function field is false. However we shall show that if $K = k(t)$ with k an algebraically closed fields of characteristic 0 and $\mathcal{O}_S$ is the ring of S-integers in K, then under certain conditions the binary forms with coefficients in $\mathcal{O}_S$ and of given degree and discriminant lie in finitely many $\mathrm{GL}\,(2, K)$-equivalence classes.

6.1 $\mathrm{GL}\,(2, K)$-equivalence classes

Let as usual k be a field with $k = \overline{k}, \mathrm{char}\, k = 0$ and $K = k(t)$, S a finite set of valuations containing ∞. Let $\alpha_1, \ldots, \alpha_s \in k$ be distinct and $p_i = t - \alpha_i, i = 1, \ldots, s$. Let $F \in \mathcal{O}_S[X, Y]$ and $\delta \in \mathcal{O}_S \backslash \{0\}$. Let L be a finite extension of K. For two binary forms $F_1, F_2 \in \mathcal{O}_S[X, Y]$ we say they are GL $(2, K)$-equivalent if there exists $U \in \mathrm{GL}\,(2, K)$ and $\lambda \in K^*$ such that $F_1 = \lambda (F_2)_U$, and they are GL $(2, L)$-equivalent if the same holds when we replace K by L.

Fix $n \geqslant 4$, and consider the following two conditions:

$$\begin{cases} F \in \mathcal{O}_S[X, Y], D(F) \in \delta \mathcal{O}_S^{\times}, \\ F \text{ has splitting field } L \text{ over } K, \\ \deg F = n, \end{cases} \tag{6.1.1}$$

$$F \text{ is not } \mathrm{GL}\,(2, L)\text{-equivalent to a binary form in } k[X, Y]. \tag{6.1.2}$$

Theorem 6.1.1. *There are only finitely many* GL $(2, K)$*-equivalence classes of binary forms satisfying (6.1.1) and (6.1.2).*

Proof. We reduce the GL $(2, K)$-equivalence classes to GL $(2, L)$-equivalence classes. We prove first that every GL $(2, L)$-equivalence class of binary forms F with (6.1.1) is a union of finitely many GL $(2, K)$-equivalence classes. Then it suffices to prove that there are only finitely many GL $(2, L)$-equivalence classes of binary forms F with (6.1.1) and (6.1.2).

Fix a binary form F satisfying (6.1.1). It has a factorization

$$\begin{cases} F = a \displaystyle\prod_{i=1}^{n} (\alpha_i X + \beta_i Y), a \in K^* \\ (\sigma(\alpha_i), \sigma(\beta_i)) = (\alpha_{\sigma(i)}, \beta_{\sigma(i)}) \text{ for } i = 1, \dots, n, \sigma \in \mathrm{Gal}(L/K), \end{cases} \tag{6.1.3}$$

where $(\sigma(1), \dots, \sigma(n))$ is a permutation of $(1, \dots, n)$ depending on F. For each $\sigma \in \mathrm{Gal}(L/K)$, there are only finitely many possibilities for the permutation of $(1, \dots, n)$ associated with σ. So we may subdivide those GL $(2, L)$-equivalence classes into subclasses under consideration such that two binary binary forms belong to the same subclass if and only if they satisfy (6.1.3) with the same permutation $(\sigma(1), \dots, \sigma(n))$ for each $\sigma \in \mathrm{Gal}(L/K)$.

Now consider all binary forms in the same subclass. These are all GL $(2, L)$-equivalent to one another and satisfy (6.1.3) with the same permutation $(\sigma(1), \ldots, \sigma(n))$. Fix one of such, $F_0 = a_0 \prod_{i=1}^{n} (\alpha_{0i}X - \beta_{0i}Y)$. Let $F = a \prod_{i=1}^{n} (\alpha_i X + \beta_i Y)$ be any other binary form in the same subclass. Then by definition there exists $V = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{PGL}\,(2, L)$ such that

$$V[\alpha_{0i} : \beta_{0i}] = [\alpha_{\tau(i)} : \beta_{\tau(i)}]\, (i = 1, \ldots, n), \tag{6.1.4}$$

with τ a permutation of $(1, \ldots, n)$. We divide our subclass into finitely many smaller subclasses, such that two binary forms in the same smaller subclass satisfy (6.1.4) with the same permutation τ.

Let F_1, F_2 be two binary forms in the same smaller subclass, i.e., they are GL $(2, L)$-equivalent and satisfy (6.1.3) with the same permutations $(\sigma(1), \ldots, \sigma(n))\, (\sigma \in \mathrm{Gal}(L/K))$ and (6.1.4) with the same τ. Assume

$$F_1 = a_1 \prod_{i=1}^{n} (\alpha_{1i}X + \beta_{1i}Y),$$
$$F_2 = a_2 \prod_{i=1}^{n} (\alpha_{2i}X + \beta_{2i}Y).$$

Then there exists $U \in \mathrm{PGL}\,(2, L)$ such that

$$U[\alpha_{1i} : \beta_{1i}] = [\alpha_{2i} : \beta_{2i}]\, (i = 1, \ldots, n), \tag{6.1.5}$$

because (6.1.4) holds true for the same τ. Without loss of generality, we assume that the U is represented by a matrix one of whose elements equals 1.

Applying each $\sigma \in \mathrm{Gal}(L/K)$ to (6.1.5), we obtain

$$\sigma(U)[\sigma(\alpha_{1i}) : \sigma(\beta_{1i})] = [\sigma(\alpha_{2i}) : \sigma(\beta_{2i})] \;\; (i = 1, \ldots, n, \sigma \in \mathrm{Gal}(L/K)).$$

By (6.1.3) and our subdivision we derive that

$$\sigma(U)[\alpha_{1\sigma(i)} : \beta_{1\sigma(i)}] = [\alpha_{2\sigma(i)} : \beta_{2\sigma(i)}]\, (\sigma \in \mathrm{Gal}(L/K), i = 1, \ldots, n).$$

Hence

$$\sigma(U)[\alpha_{1i}:\beta_{1i}] = [\alpha_{2i}:\beta_{2i}]\,(\sigma \in \mathrm{Gal}(L/K), i = 1,\dots,n). \tag{6.1.6}$$

Now from (8.2.3) and (8.3.3) it follows that the images of $[\alpha_{1i} : \beta_{1i}]\,(i = 1,\dots,n)$ under the projective transformation U and $\sigma(U)$ are equal. Since $n \geqslant 3$ and one of the entries of U is 1, this implies $\sigma(U) = U$ for any $\sigma \in \mathrm{Gal}(L/K)$. Hence $U \in \mathrm{PGL}\,(2, K)$. This means that F_1, F_2 are actually $\mathrm{GL}\,(2, K)$-equivalent, which proves the claim.

What remains is to prove that the binary forms with (6.1.1), (6.1.2) and (6.1.3) lie in only finitely many $\mathrm{GL}\,(2, L)$-equivalence classes.

Write $F = a\prod_{i=1}^{n}(\alpha_i X + \beta_i Y)$ with $a \in K^*$. Suppose $D(F) \in \delta\mathcal{O}_S^\times$. Let $R' = \mathcal{O}_S[\delta^{-1}]$. Then $D(F) \in R'^\times$. Let R'_L be the integral closure of R' in L. For $\theta_1,\dots,\theta_r \in L$ we denote by $(\theta_1,\dots,\theta_r)$ the fractional ideal with respect to R'_L generated by $\theta_1,\dots,\theta_r$. Further, for a given polynomial P we denote by (P) the ideal of R'_L generated by the coefficients of P. Then by Gauss' Lemma we have

$$(F) = (a)\prod_{i=1}^{n}(\alpha_i, \beta_i). \tag{6.1.7}$$

Let $\Delta_{ij} = \alpha_i\beta_j - \alpha_j\beta_i$. Then

$$(\Delta_{ij}) \subseteq (\alpha_i, \beta_i)(\alpha_j, \beta_j) \text{ for } i, j = 1,\dots,n, i \neq j. \tag{6.1.8}$$

Now we have

$$\begin{aligned}
(1) \;&\supseteq\; \prod_{1\leqslant i<j\leqslant n}\Big(\frac{(\Delta_{ij})}{(\alpha_i,\beta_i)(\alpha_j,\beta_j)}\Big)^2 \\
&=\; \frac{(a^{-2n+2}D(F))}{(a^{-2n+2})(F)^{2n-2}} \\
&=\; \frac{(D(F))}{(F)^{2n-2}} \\
&\supseteq\; (1),
\end{aligned}$$

where the last equality is implied by the fact that $D(F) \in R'^{\times}$ and $F \in R'[X,Y]$. So we derive that (6.1.8) is actually an equality for every pair (i,j). Define the cross ratio

$$\rho_{ijhl}(F) := \frac{\Delta_{ij}\Delta_{hl}}{\Delta_{ih}\Delta_{jl}}.$$

Then $\rho_{ijhl}(F) \in R_L'^{\times}$ for all distinct $i,j,h,l \in \{1,\ldots,n\}$.

Lemma 6.1.2. *Let L be a finite extension of $k(t)$ and $\mathcal{O}_L$ the integral closure of $k[t]$ in L. The unit equation $x+y=1$ has only finitely many solutions x,y with $x,y \in \mathcal{O}_L \backslash k$ and all of them can be determined effectively in principle.*

Proof. See Theorem 1 and Theorem 2 of [17]. □

Lemma 6.1.3. *Suppose that $\frac{\Delta_{ij}\Delta_{hl}}{\Delta_{ih}\Delta_{jl}}$ lies in k^* for all tuples (i,j,h,l) in $\{1,\ldots,n\}$ with i,j,h,l distinct. Then F is* GL $(2,L)$*-equivalent to a binary form in $k[X,Y]$.*

Proof. Let $F = a\prod_{i=1}^{n}(\alpha_i X + \beta_i Y)$. Then there exists $U \in \mathrm{PGL}\,(2,L)$ such that

$$\begin{cases} U[\alpha_1 : \beta_1] = [1:0], \\ U[\alpha_2 : \beta_2] = [0:1], \\ U[\alpha_3 : \beta_3] = [1:1]. \end{cases} \tag{6.1.9}$$

So F is GL $(2,L)$-equivalent to a binary form of the shape

$$F' = a'XY(X+Y)\prod_{i=4}^{n}(\alpha_i' X + \beta_i' Y),$$

with $a' \in L^*$. Since the cross ratios remain invariant under a projective transformation, we have $\rho_{123i}(F') = 1 - \frac{\beta_i'}{\alpha_i'} \in k$ for $i \geqslant 4$. Hence $\frac{\beta_i'}{\alpha_i'} \in k$ and therefore $F' = bP$ with $b \in L^*, P \in k[X,Y]$. This proves the assertion. □

Now consider $F = a\prod_{i=1}^{n}(\alpha_i X + \beta_i Y), a \in K^*, n \geqslant 4$ with $D(F) \in R'^{\times}$. By (6.1.2) and Lemma 6.1.3, we may assume without loss of generality that $\rho_{1234} \notin k$. Since

$$\Delta_{12}\Delta_{34} + \Delta_{14}\Delta_{23} = \Delta_{13}\Delta_{24},$$

we have

$$\rho_{1234}(F) + \rho_{1432}(F) = 1.$$

But $\rho_{1234}(F), \rho_{1432}(F) \in R_L'^{\times}$, hence by Lemma 6.1.2, we know that there are only finitely many possibilities for $\rho_{1234}(F)$. For each choice $\lambda \in L\backslash k$ of $\rho_{1234}(F)$, consider all binary forms F with $\rho_{1234}(F) = \lambda$. There exists $U \in \mathrm{PGL}\,(2, L)$ such that (6.1.9) holds. So F is $\mathrm{GL}\,(2, L)$-equivalent to $XY(X+Y)\prod_{i=4}^{n}(\alpha'_i X + \beta'_i Y)$ with $\alpha'_i \neq 0$ for $i = 1, \ldots, n$. Since we have $\rho_{1234}(F) = 1 - \frac{\beta'_4}{\alpha'_4}$, we deduce that F is $\mathrm{GL}\,(2, L)$-equivalent to $XY(X+Y)(X+(\lambda+1)Y)$ if $n = 4$ or $XY(X+Y)(X+(\lambda+1)Y)\prod_{i=5}^{n}(X - \gamma_i Y)$ if $n \geqslant 5$. When $n > 4$, observe that for $i > 4$ we have $\rho_{123i}(F) = 1 + \gamma_i$ and $\rho_{124i}(F) = -1 - \frac{\lambda+1}{\gamma_i}$. These quantities cannot lie in k simultaneously since $\lambda \notin k$. Hence by applying Lemma 6.1.2 again, we infer that there are only finitely possibilities for $\gamma_i, i > 4$. It follows that there are only finitely many $\mathrm{GL}\,(2, L)$-equivalence classes of binary forms with (6.1.1), (6.1.2) and (6.1.3). This completes the proof.

□

Remark 6.1.4. The condition (6.1.2) cannot be relaxed to the condition that F not be $\mathrm{GL}\,(2, K)$-equivalent to a binary form in $k[X, Y]$. Here is a counter-example: fix $b \in K\backslash K^2$, consider all binary forms $F = X^4 + abX^2Y^2 + b^2Y^4, a \in k, a^2 \neq 4$. First, notice that the splitting field of such an F over K is $L = K(\sqrt{b})$, so F is $\mathrm{GL}\,(2, L)$-equivalent to $G = X^4 + aX^2Y^2 + Y^4 \in k[X, Y]$. However, F is not $\mathrm{GL}\,(2, K)$-equivalent to a binary form in $k[X, Y]$, since otherwise F would split into linear factors in

K, contradicting the fact that $b \notin K^2$. Clearly, $F_a = X^4 + abX^2Y^2 + b^2Y^4$ and $F_{a'} = X^4 + a'bX^2Y^2 + b^2Y^4$ satisfy (6.1.1). Suppose F_a and $F_{a'}$ are GL $(2, K)$-equivalent. Then $G_a = X^4 + aX^2Y^2 + Y^4$ and $G_{a'} = X^4 + a'X^2Y^2 + Y^4$ are GL $(2, L)$-equivalent, hence being GL $(2, k)$-equivalent. Let $c = \sqrt{a^2 - 4} \in k$. Then

$$G_a = (X - \lambda_1 Y)(X - \lambda_2 Y)(X - \lambda_3 Y)(X - \lambda_4 Y),$$

where $\lambda_1 = \sqrt{\frac{-a+c}{2}}, \lambda_2 = -\sqrt{\frac{-a+c}{2}}, \lambda_3 = \sqrt{\frac{-a-c}{2}}, \lambda_4 = -\sqrt{\frac{-a-c}{2}}$. The cross-ratio of $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ is

$$\lambda = \frac{(\lambda_1 - \lambda_2)(\lambda_4 - \lambda_3)}{(\lambda_1 - \lambda_3)(\lambda_4 - \lambda_2)} = \frac{4}{a+2}.$$

The cross ratios of the permutations of $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ are

$$\begin{aligned} \lambda &= \frac{4}{a+2}, & \frac{1}{\lambda} &= \frac{a+2}{4}, \\ 1 - \lambda &= \frac{a-2}{a+2}, & \frac{1}{1-\lambda} &= \frac{a+2}{a-2}, \\ \frac{\lambda}{\lambda - 1} &= \frac{4}{2-a}, & \frac{\lambda - 1}{\lambda} &= \frac{2-a}{4}. \end{aligned}$$

These are all one-to-one functions of a. Therefore, if $G_{a'}$ is GL $(2, L)$-equivalent to G_a for some $a' \in k$, the corresponding cross-ratios remain the same, so there are at most six choices of a' such that $G_{a'}$ and G_a are GL $(2, k)$-equivalent. This implies that when a runs through k, there are infinitely many GL $(2, K)$-equivalence classes of binary forms of the form $F = X^4 + abX^2Y^2 + b^2Y^4$.

6.2 $\mathrm{GL}(2, \mathcal{O}_S)$-equivalence classes

Let $K = k(t)$ and S a finite set of valuations of K. We now show that a $\mathrm{GL}(2, K)$-equivalence class of binary forms with (6.1.1) is in general not a union of finitely many $\mathrm{GL}(2, \mathcal{O}_S)$-equivalence classes.

Lemma 6.2.1. *Let $F \in K[X, Y]$ with degree $\deg F \geqslant 3$ and $\mathrm{Aut}(F) := \{W \in \mathrm{PGL}(2, K) :$ there exists $\lambda \in K^*$ such that $F_W = \lambda F\}$. Then $\mathrm{Aut}(F)$ is finite.*

Proof. Let $W = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{PGL}(2, K), \lambda \in K^*$ such that $F_W = \lambda F$ and $F = \prod_{i=1}^{n} (\alpha_i X + \beta_i Y)$ with $\alpha_i, \beta_i \in \overline{K}$. Then $F_W = \prod_{i=1}^{n} \big((a\alpha_i + c\beta_i)X + (b\alpha_i + d\beta_i)Y\big)$. So there is a permutation σ of $(1, \ldots, n)$ such that $[\sigma(\alpha_i) : \sigma(\beta_i)] = [\alpha_i : \beta_i]W$ for $i = 1, \ldots, n$. That is, W maps $n \geqslant 3$ distinct points in $\mathbb{P}^1(\overline{K})$ to n other distinct points. Hence W depends only on σ. Therefore $\#\mathrm{Aut}(F) \leqslant n!$. □

Let $U_1, U_2 \in \mathrm{GL}(2, K)$ with entries in $\mathcal{O}_S$. If F_{U_1} and F_{U_2} are $\mathrm{GL}(2, \mathcal{O}_S)$-equivalent, then by definition $F_{U_1} = \varepsilon F_{U_2 V}$ for some $V \in \mathrm{GL}(2, \mathcal{O}_S), \varepsilon \in k^*$. Then $F_{(U_2 V)^{-1} U_1} = \varepsilon F$ and so $(U_2 V)^{-1} U_1 \in \mathrm{Aut}(F)$, hence $U_1 = U_2 V W$ for some $W \in \mathrm{Aut}(F)$, in this case we say U_1 is *related* to U_2 *associated to* W and write $U_1 \equiv U_2\,(W)$.

Lemma 6.2.2. *Let $F \in K[X, Y]$ and $U \in \mathrm{GL}(2, K)$ with entries in $\mathcal{O}_S$ and $\det U = \delta$. Assume U_1, U_2 are related to U associated to the same $W \in \mathrm{Aut}(F)$ with $\det U_1, \det U_2 \in \delta \mathcal{O}_S^{\times}$. Then we have $U_1 U_2^{-1} \in \mathrm{GL}(2, \mathcal{O}_S)$.*

Proof. By assumption we have

$$U_1^{-1} U = V_1(\lambda_1 W),\; U_2^{-1} U = V_2(\lambda_2 W),$$

for $V_1, V_2 \in \mathrm{GL}(2,\mathcal{O}_S)$ and $\lambda_1, \lambda_2 \in K^*$. Then $U_1^{-1}U_2 = \frac{\lambda_1}{\lambda_2}V_1V_2^{-1}$. But $\frac{\det U_1}{\det U_2} \in \mathcal{O}_S^\times$, hence $\frac{\lambda_1}{\lambda_2} \in \mathcal{O}_S^\times$. Therefore $U_1^{-1}U_2 \in GL(2,\mathcal{O}_S)$. This completes the proof. $\square$

Theorem 6.2.3. *Let $F \in \mathcal{O}_S[X,Y]$ be a binary form of degree $n \geqslant 3$ and non-zero discriminant. Then there exists $D \in \mathcal{O}_S\backslash\{0\}$ with the following property: the binary forms $F' \in \mathcal{O}_S[X,Y]$ with*

$$\begin{cases} D(F') \in D\mathcal{O}_S^\times, \\ F' \text{ is } \mathrm{GL}(2,K)\text{-equivalence to } F \end{cases} \tag{6.2.1}$$

lie in infinitely many GL$(2,\mathcal{O}_S)$*-equivalence classes.*

Proof. Suppose $S = \{\infty, p_1, \ldots, p_h\}$ and take $T = t$ if $S = \{\infty\}$ or $T = \prod_{i=1}^{h}(t-p_i)$ otherwise. Consider all binary forms F_U where $U \in \mathrm{GL}(2,K)$ has entries in $\mathcal{O}_S$ and $\det U = T^2 - 1$. Let $D = (T^2-1)^{n(n-1)}D(F)$. Suppose there are only finitely many GL$(2,\mathcal{O}_S)$-equivalence classes of binary forms in $\mathcal{O}_S[X,Y]$ with the property (6.2.1). Then for every binary form F_V there exists U and $W \in \mathrm{Aut}(F)$ such that $V \equiv U\,(W)$.

Choose $U_1 = \left(\begin{smallmatrix} aT & 1 \\ 1 & bT \end{smallmatrix}\right)$, $U_2 = \left(\begin{smallmatrix} a'T & 1 \\ 1 & b'T \end{smallmatrix}\right)$ with $a, b, a', b' \in k$ satisfying $ab = a'b' = 1, a \neq a'$. Then U_1, U_2 have entries in $\mathcal{O}_S$ and $F_{U_1}, F_{U_2} \in \mathcal{O}_S[X,Y]$. But we have

$$U_1U_2^{-1} = \frac{1}{T^2-1}\begin{pmatrix} ab'T^2 - 1 & a'T - aT \\ b'T - bT & a'bT^2 - 1 \end{pmatrix}.$$

This is not in GL$(2,\mathcal{O}_S)$ because for each $i = 1, \ldots, h$, $t - p_i$ is coprime with $T^2 - 1 = (T-1)(T+1)$.

Since k is algebraically closed, k is an infinite field, hence there are infinitely many matrices of the form U_1 and U_2. So there must be two matrices V, V' of form U_1, U_2 and $U \in \mathrm{GL}(2,K), W \in \mathrm{Aut}(F)$ such that $V \equiv U\,(W), V' \equiv U\,(W)$. This is a contradiction of the above and Lemma 6.2.2. $\square$

Chapter 7

Lower bounds for resultants

Evertse and Győry deduced some semi-effective lower bounds of resultants over number fields in [10], [12]. Apart from two theorems mentioned in the introduction, they have also established the following:

Theorem (Evertse, Győry). *Let $F, G \in \mathbb{Z}[X, Y]$ be two binary forms of degree $m > 1, n > 2$ such that FG has splitting field L over $\mathbb{Q}$ and is square-free, and $F(1,0) = G(1,0) = 1$. Then*

$$|R(F,G)| \geqslant C(m,n,L) \max(|D(F)|_{S_1}^{\frac{n}{7(m-1)}}, |D(G)|_{S_1}^{\frac{m}{7(n-1)}}),$$

where $C(m,n,L)$ depends on m, n and L.

The constant $C(m,n,L)$ cannot be effectively computed from their method of proof. In this chapter, we deduce effective analogous results over function fields, with the help of outcome derived before.

7.1 Monic binary forms

Recall that $K = k(t)$ and S is a finite set of valuations of K containing ν_∞. Let L be a finite extension of $K = k(t)$ of genus g_L. Let T be a finite set of places of L above those in S. Denote by $\mathcal{O}_T$ the integral closure of $\mathcal{O}_S$.

A binary form of degree n is called X-monic if the leading coefficient of X^n is 1. We call two X-monic quadratic forms related if the coefficients of the term XY are the same, and unrelated if otherwise.

Lemma 7.1.1. *Let F, G be two binary quadratic forms over the ring $\mathcal{O}_T$ satisfying $F(1,0) = G(1,0) = 1$, and suppose that FG is square-free and splits into linear forms over L. Then we have*

(i) $|D(F)|_T \leqslant e^{2(\#T+\max(0,2g_L-2))}|R(F,G)|_T|D(G)|_T$, *if F, G are related;*

(ii) $|D(F)|_T \leqslant e^{6(\#T+\max(0,2g_L-2))}|R(F,G)|_T^2$, *if F, G are unrelated.*

Proof. Put $g' = \max(0, 2g_L - 2)$. Since $F(1,0) = G(1,0) = 1$ and FG splits into linear factors over L, we may assume that

$$F(X,Y) = (X - \alpha_1 Y)(X - \alpha_2 Y),$$

$$G(X,Y) = (X - \beta_1 Y)(X - \beta_2 Y).$$

where $\alpha_1, \alpha_2, \beta_1, \beta_2$ are distinct elements of L.

We actually have $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathcal{O}_T$, since $\mathcal{O}_T$ is integrally closed.

Now, we have

$$D(F) = (\alpha_1 - \alpha_2)^2, D(G) = (\beta_1 - \beta_2)^2,$$

$$R(F,G) = (\alpha_1 - \beta_1)(\alpha_1 - \beta_2)(\alpha_2 - \beta_1)(\alpha_2 - \beta_2).$$

If F, G are related, i.e., $\alpha_1 + \alpha_2 = \beta_1 + \beta_2$, then $\alpha_1 - \alpha_2 = (\beta_1 - \alpha_2) + (\beta_2 - \alpha_2)$. Considering the identity $(\beta_1 - \alpha_2) - (\beta_2 - \alpha_2) - (\beta_1 - \beta_2) = 0$, and applying Corollary 2.2.11, we have

$$H_T(\beta_1-\alpha_2, -(\beta_2-\alpha_2), -(\beta_1-\beta_2)) \leqslant e^{\#T+g'}|(\beta_1-\alpha_2)(\beta_2-\alpha_2)(\beta_1-\beta_2))|_T,$$

hence

$$\begin{aligned} |D(F)|_T^{\frac{1}{2}} &= |\alpha_1 - \alpha_2|_T \\ &= |\beta_1 - \alpha_2 + \beta_2 - \alpha_2|_T \\ &\leqslant e^{\#T+g'}|(\beta_1 - \alpha_2)(\beta_2 - \alpha_2)(\beta_1 - \beta_2))|_T \\ &= e^{\#T+g'}|(\beta_1 - \alpha_2)(\beta_2 - \alpha_2)|_T |D(G)|_T^{\frac{1}{2}}. \end{aligned}$$

Similarly, we have

$$|D(F)|_T^{\frac{1}{2}} \leqslant e^{\#T+g'}|(\beta_1 - \alpha_1)(\beta_2 - \alpha_1)|_T |D(G)|_T^{\frac{1}{2}}.$$

Hence

$$|D(F)|_T \leqslant e^{2(\#T+g')}|R(F,G)|_T |D(G)|_T.$$

If F, G are unrelated, i.e., $\alpha_1 + \alpha_2 \neq \beta_1 + \beta_2$, then we consider the identity

$$(\alpha_1 - \beta_1) - (\alpha_1 - \beta_2) - (\alpha_2 - \beta_1) + (\alpha_2 - \beta_2) = 0,$$

which satisfies the condition of Corollary 2.2.11. We derive that

$$H_T(\alpha_1 - \beta_1, \alpha_1 - \beta_2, \alpha_2 - \beta_1, \alpha_2 - \beta_2) \leqslant e^{3(\#T+g')}|R(F,G)|_T.$$

Hence

$$\begin{aligned} |D(F)|_T^{\frac{1}{2}} &= |\alpha_1 - \alpha_2|_T \\ &= |(\alpha_1 - \beta_1) - (\alpha_2 - \beta_1)|_T \\ &\leqslant H_T(\alpha_1 - \beta_1, \alpha_1 - \beta_2, \alpha_2 - \beta_1, \alpha_2 - \beta_2) \\ &\leqslant e^{3(\#T+g')}|R(F,G)|_T, \end{aligned}$$

and

$$|D(F)|_T \leqslant e^{6(\#T+g')}|R(F,G)|_T^2.$$

□

Theorem 7.1.2. *Let K_1 be a finite extension of $K = k(t)$, S_1 a finite set of valuations of K_1. Let $F, G \in \mathcal{O}_{S_1}[X,Y]$ be binary forms satisfying the following conditions:*

$$\begin{cases} \deg F = m \geqslant 2, \deg G = n \geqslant 3, F(1,0) = G(1,0) = 1, \\ FG \text{ has splitting field } L \text{ over } K_1 \text{ and } FG \text{ is square-free.} \end{cases}$$

Then we have

$$|R(F,G)|_{S_1} \geqslant e^{-\frac{mn}{2}(\#S_1 + \frac{\max(0,2g_L-2)}{[L:K_1]})} \max(|D(F)|_{S_1}^{\frac{n}{6(m-1)}}, |D(G)|_{S_1}^{\frac{m}{6(n-1)}}).$$

Proof. Let $T \subset M_L$ be the set of valuations over those in S_1. Over L, we have

$$F(X,Y) = \prod_{i=1}^{m}(X - \alpha_i Y),\ G(X,Y) = \prod_{j=1}^{n}(X - \beta_j Y).$$

Since $F, G \in \mathcal{O}_{S_1}[X,Y]$, we have $\alpha_i, \beta_j \in \mathcal{O}_T$, $i = 1, \ldots, m, j = 1, \ldots, n$.

Let

$$F_{pq}(X,Y) = (X - \alpha_p Y)(X - \alpha_q Y)\,(1 \leqslant p < q \leqslant m),$$

$$G_{ij}(X,Y) = (X - \beta_i Y)(X - \beta_j Y)\,(1 \leqslant i < j \leqslant n).$$

Now fix a pair $p < q$. Let I_{pq} be the collection of pairs $i < j$ such that G_{ij} is related to F_{pq}: $\alpha_p + \alpha_q = \beta_i + \beta_j$. Then each two pairs in I_{pq} are disjoint since FG is square-free, and hence $\#I_{pq} \leqslant [\frac{n}{2}] \leqslant \frac{n(n-1)}{6}$.

Put $g' = \max(0, 2g_L - 2)$. By Lemma 7.1.1 we get

$$|D(F_{pq})|_T \leqslant e^{6(\#T+g')}|R(F_{pq}, G_{ij})|_T^2 \text{ for } (i,j) \notin I_{pq}.$$

Hence we have

$$
\begin{aligned}
|D(F_{pq})|_T &\leqslant e^{6(\#T+g')}\Big(\prod_{(i,j)\notin I_{pq}} |R(F_{pq},G_{ij})|_T\Big)^{\frac{2}{\frac{n(n-1)}{2}-\#I_{pq}}} && (7.1.1)\\
&\leqslant e^{6(\#T+g')}\Big(\prod_{(i,j)\notin I_{pq}} |R(F_{pq},G_{ij})|_T\Big)^{\frac{6}{n(n-1)}} && (7.1.2)\\
&\leqslant e^{6(\#T+g')}\Big(\prod_{1\leqslant i<j\leqslant n} |R(F_{pq},G_{ij})|_T\Big)^{\frac{6}{n(n-1)}} && (7.1.3)\\
&= e^{6(\#T+g')}|R(F_{pq},G)|_T^{\frac{6}{n}}. && (7.1.4)
\end{aligned}
$$

So

$$
\begin{aligned}
|D(F)|_T &= \prod_{1\leqslant p<q\leqslant m} |D(F_{pq})|_T\\
&\leqslant e^{3m(m-1)(\#T+g')}\Big(\prod_{1\leqslant p<q\leqslant m} |R(F_{pq},G)|_T\Big)^{\frac{6}{n}}\\
&\leqslant e^{3m(m-1)(\#T+g')}|R(F,G)|_T^{\frac{6(m-1)}{n}}.
\end{aligned}
$$

Similarly,

$$|D(G)|_T \leqslant e^{3n(n-1)(\#T+g')}|R(F,G)|_T^{\frac{6(n-1)}{m}}.$$

Then it follows from $|x|_T = |x|_{S_1}^{[L:K_1]}, x \in K_1$ and $\#T \leqslant [L:K_1]\#S_1$. □

Remark 7.1.3. *If $m = n = 2$, the results above are not valid. Simply take $K = k(t), S = \nu_\infty$. Let $u, v \in k[t]$ be a solution of $x^2 - (t^2-1)y^2 = 1$ and put $F(X,Y) = X^2 - u^2Y^2, G(X,Y) = X^2 - (t^2-1)v^2Y^2$. Then it is easy to check that $D(F) = 4u^2, D(G) = 4v^2(t^2-1), R(F,G) = 1$ and FG is square free with splitting field $K(\sqrt{t^2-1})$. However, since $u = t, v = 1$ is a solution of $x^2 - (t^2-1)y^2 = 1$, we can find infinitely many solutions $u, v \in k[t]$ satisfying $u + \sqrt{t^2-1}v = (t+\sqrt{t^2-1})^j$ with $|u|_\infty$ goes to infinity.*

7.2 Results for binary cubic forms

Recall that K_1 is a finite extension of $K = k(t)$ with genus g_{K_1}, S_1 a finite set of valuations on K_1. Consider two binary forms $F, G \in K_1[X, Y]$ such that

$$F(X,Y) = \prod_{i=1}^{3}(\alpha_i X - \beta_i Y),$$

$$G(X,Y) = \prod_{i=1}^{3}(\gamma_i X - \delta_i Y),$$

where $\alpha_i, \beta_i, \gamma_j, \delta_j \in K_1, i, j = 1, 2, 3$, and FG is square-free. In this section we prove

Proposition 7.2.1. *With the same setting as above, we have*

$$|R(F,G)|_{S_1} \geqslant e^{-\frac{90}{17}(2g_{K_1}-1+\#S_1)}|D(F)D(G)|_{S_1}^{\frac{3}{34}}|FG|_{S_1}^{\frac{45}{17}}.$$

Before proving this result we start with some preliminaries and a lemma.

Put $\Delta_{ij} = \alpha_i\delta_j - \beta_i\gamma_j, F_{ij} = \alpha_i\beta_j - \alpha_j\beta_i, G_{ij} = \gamma_i\delta_j - \gamma_j\delta_i$ for $i, j = 1, 2, 3$. Then by direct calculation

$$\det\begin{pmatrix} \Delta_{11} & \Delta_{12} & \Delta_{13} \\ \Delta_{21} & \Delta_{22} & \Delta_{23} \\ \Delta_{31} & \Delta_{32} & \Delta_{33} \end{pmatrix} = 0.$$

Put

$$u_1 = \Delta_{11}\Delta_{22}\Delta_{33},\ u_2 = -\Delta_{11}\Delta_{23}\Delta_{32},$$
$$u_3 = \Delta_{12}\Delta_{23}\Delta_{31},\ u_4 = -\Delta_{12}\Delta_{21}\Delta_{33},$$
$$u_5 = \Delta_{13}\Delta_{21}\Delta_{32},\ u_6 = -\Delta_{13}\Delta_{22}\Delta_{31}.$$

Then

$$u_1 + u_2 + u_3 + u_4 + u_5 + u_6 = 0, \tag{7.2.1}$$

$$u_1 u_3 u_5 = -u_2 u_4 u_6 = R(F, G) \neq 0. \qquad (7.2.2)$$

Hence $u_i \neq 0$ for $i = 1, \ldots, 6$. Also

$$R(F, G) = \prod_{i=1}^{3} \prod_{j=1}^{3} \Delta_{ij}, D(F) = (F_{12} F_{23} F_{13})^2, D(G) = (G_{12} G_{23} G_{13})^2.$$

Similarly as in [12], we have

$$(D(F)D(G))^{\frac{3}{2}} = \pm R(F, G)^{-1} \prod_{\substack{1 \leqslant p < q \leqslant 6 \\ p \not\equiv q (\mathrm{mod}\, 2)}} (u_p + u_q). \qquad (7.2.3)$$

Hence

$$u_p + u_q \neq 0 \text{ for } 1 \leqslant p < q \leqslant 6, p \not\equiv q (\mathrm{mod}\, 2). \qquad (7.2.4)$$

Put $c_n = e^{\binom{n}{2}(\max(2g_{K_1} - 2 + \# S_1, 0)}$. Analogously to Lemma 5, [12], we have

Lemma 7.2.2. *For $(u_1, \ldots, u_6)$ satisfying (7.2.1), (7.2.2), (7.2.4), we have*

$$\prod_{\substack{1 \leqslant p < q \leqslant 6 \\ p \not\equiv q (\mathrm{mod}\, 2)}} H_{S_1}(u_p, u_q) \leqslant c_5^9 |R(F, G)|_{S_1}^{18} |FG|_{S_1}^{-45}.$$

Proof. We adapt the idea in the proof of Lemma 5 of [12].

By symmetry, we have to consider only the following four cases:

(i) $\sum_{i=1}^{6} u_i$ has no vanishing proper subsum;

(ii) $u_1 + u_3 = 0, u_2 + u_4 + u_5 + u_6 = 0$ with no vanishing proper subsum;

(iii) $u_1 + u_2 + u_3 = u_4 + u_5 + u_6 = 0$ with no vanishing proper subsum;

(iv) $u_1 + u_3 + u_5 = u_2 + u_4 + u_6 = 0$ with no vanishing proper subsum.

First, since

$$|\Delta_{ij}|_\nu = |\alpha_i\delta_j - \beta_i\gamma_j|_\nu \leqslant \max(|\alpha_i|_\nu, |\beta_i|_\nu)\max(|\gamma_j|_\nu, |\delta_j|_\nu) \text{ for } \nu \in M_{K_1}, i, j = 1, 2, 3,$$

we have

$$|u_1|_\nu \leqslant \prod_{i=1}^{3} \max(|\alpha_i|_\nu, |\beta_i|_\nu) \prod_{j=1}^{3} \max(|\gamma_j|_\nu, |\delta_j|_\nu) = |FG|_\nu,$$

and similarly for $i = 1, \ldots, 6$

$$|u_i|_\nu \leqslant |FG|_\nu \text{ for } \nu \in M_{K_1}. \tag{7.2.5}$$

For case (i), by applying Corollary 2.2.11, we get for $p < q$ with $p \not\equiv q (\operatorname{mod} 2)$ that

$$\begin{aligned} H_{S_1}(u_p, u_q) &\leqslant H_{S_1}(u_1, \ldots, u_6) \\ &\leqslant c_5 \prod_{i=1}^{6} |u_i|_{S_1} \left(\prod_{\nu \notin S_1} \max_{1 \leqslant i \leqslant 6} (|u_i|_\nu) \right)^5 \\ &\leqslant c_5 |R(F,G)|_{S_1}^2 |FG|_{S_1}^{-5}. \end{aligned}$$

hence

$$\prod_{\substack{1 \leqslant p < q \leqslant 6 \\ p \not\equiv q (\operatorname{mod} 2)}} H_{S_1}(u_p, u_q) \leqslant c_5^9 |R(F,G)|_{S_1}^{18} |FG|_{S_1}^{-45}.$$

For case (ii), we apply Corollary 2.2.11 to $u_2 + u_4 + u_5 + u_6 = 0$ and derive that for $(p, q) = (2, 5), (4, 5), (5, 6)$,

$$\begin{aligned} H_{S_1}(u_p, u_q) &\leqslant H_{S_1}(u_2, u_4, u_5, u_6) \\ &\leqslant c_3 |u_2 u_4 u_5 u_6|_{S_1} \left(\prod_{\nu \notin S_1} \max(|u_2|_\nu, |u_4|_\nu, |u_5|_\nu, |u_6|_\nu) \right)^3 \\ &\leqslant c_3 \prod_{i=1}^{6} |u_i|_{S_1} \left(\prod_{\nu \notin S_1} \max_{1 \leqslant i \leqslant 6} (|u_i|_\nu) \right)^5 \\ &\leqslant c_3 |R(F,G)|_{S_1}^2 |FG|_{S_1}^{-5}, \end{aligned} \tag{7.2.6}$$

where in the penultimate inequality we have used the consequence of the product formula that $|u_j|_{S_1}\left(\prod_{\nu\notin S_1}\max_{1\leqslant i\leqslant 6}(|u_i|_\nu)\right)\geqslant 1$ for $j=1,3$, while in the last inequality we used (7.2.2) and (7.2.5).

For $(p,q)=(1,2),(1,4),(1,6),(2,3),(3,4),(3,6)$, we combine $u_1+u_3=0$ with (7.2.2) and get, for example in the case $(p,q)=(1,2)$, using $(u_1^2,u_2^2)=\frac{u_2}{u_5}(u_4u_6,u_2u_5)$,

$$\begin{aligned} H_{S_1}(u_1,u_2)^2 &= H_{S_1}(u_1^2,u_2^2)\\ &\leqslant \left|\frac{u_2}{u_5}\right|_{S_1} H_{S_1}(u_4,u_2)H_{S_1}(u_6,u_5)\\ &\leqslant \left|\frac{u_2}{u_5}\right|_{S_1} H_{S_1}(u_2,u_4,u_5,u_6)^2. \end{aligned}$$

By Corollary 2.2.11, this is at most

$$\begin{aligned} &\left|\frac{u_2}{u_5}\right|_{S_1} c_3^2|u_2u_4u_5u_6|_{S_1}^2\left(\prod_{\nu\notin S_1}\max(|u_2|_\nu,|u_4|_\nu,|u_5|_\nu,|u_6|_\nu)\right)^6\\ = \;& c_3^2|R(F,G)|_{S_1}^2|u_2u_5|_{S_1}\left(\prod_{\nu\notin S_1}\max(|u_2|_\nu,|u_4|_\nu,|u_5|_\nu,|u_6|_\nu)\right)^6\\ \leqslant\;& c_3^2|R(F,G)|_{S_1}^2|u_1u_2u_3u_4u_5u_6|_{S_1}\left(\prod_{\nu\notin S_1}\max_{1\leqslant i\leqslant 6}(|u_i|_\nu)\right)^{10}\\ \leqslant\;& c_3^2|R(F,G)|_{S_1}^4|FG|_{S_1}^{-10}, \end{aligned}$$

where in the penultimate inequality we have used that $|u_j|_{S_1}\left(\prod_{\nu\notin S_1}\max_{1\leqslant i\leqslant 6}(|u_i|_\nu)\right)\geqslant 1$ for $j=1,3,4,6$, and in the last inequality again (7.2.2) and (7.2.5).

This also gives

$$H_{S_1}(u_1,u_2)\leqslant c_3|R(F,G)|_{S_1}^2|FG|_{S_1}^{-5}.$$

In the same way, this inequality holds true for $(p,q)=(1,4),(1,6),(2,3),(3,4),(3,6)$, and therefore

$$\prod_{\substack{1\leqslant p<q\leqslant 6\\ p\not\equiv q (\mathrm{mod}\, 2)}} H_{S_1}(u_p,u_q)\leqslant c_3^9|R(F,G)|_{S_1}^{18}|FG|_{S_1}^{-45}.$$

For case (iii), first we apply Corollary 2.2.11 to $u_1+u_2+u_3=0$ and $u_4+u_5+u_6=0$ and obtain

$$\begin{aligned}
& H_{S_1}(u_1,u_2)H_{S_1}(u_2,u_3)H_{S_1}(u_4,u_5)H_{S_1}(u_5,u_6)\\
\leqslant\ & H_{S_1}(u_1,u_2,u_3)^2H_{S_1}(u_4,u_5,u_6)^2\\
\leqslant\ & c_2^4\prod_{i=1}^{6}|u_i|_{S_1}^2\left(\prod_{\nu\notin S_1}\max_{1\leqslant i\leqslant 3}(|u_i|_\nu)\right)^4\left(\prod_{\nu\notin S_1}\max_{4\leqslant i\leqslant 6}(|u_i|_\nu)\right)^4\\
\leqslant\ & c_2^4|R(F,G)|_{S_1}^4|FG|_{S_1}^{-8}. \qquad (7.2.7)
\end{aligned}$$

We estimate $H_{S_1}(u_p,u_q)$ for $(p,q)=(1,4),(1,6),(3,4),(3,6)$. When $(p,q)=(1,4)$, we have by (7.2.2), for instance in the case $(p,q)=(1,4)$, that $(u_1,u_4)=\frac{u_1u_4}{R(F,G)}(-u_2u_6,u_3u_5)$. Hence by corollary 2.2.11 we have

$$\begin{aligned}
H_{S_1}(u_1,u_4) &\leqslant |u_1u_4|_{S_1}|R(F,G)|_{S_1}^{-1}H_{S_1}(u_2,u_3)H_{S_1}(u_6,u_5)\\
&\leqslant |u_1u_4|_{S_1}|R(F,G)|_{S_1}^{-1}H_{S_1}(u_1,u_2,u_3)H_{S_1}(u_4,u_5,u_6)\\
&\leqslant |u_1u_4|_{S_1}|R(F,G)|_{S_1}^{-1}c_2^2|u_1u_2u_3|_{S_1}\left(\prod_{\nu\notin S_1}\max_{1\leqslant i\leqslant 3}(|u_i|_\nu)\right)^2\\
&\qquad\times|u_4u_5u_6|_{S_1}\left(\prod_{\nu\notin S_1}\max_{4\leqslant i\leqslant 6}(|u_i|_\nu)\right)^2\\
&\leqslant c_2^2|u_1u_4|_{S_1}|R(F,G)|_{S_1}\left(\prod_{\nu\notin S_1}\max_{1\leqslant i\leqslant 6}(|u_i|_\nu)\right)^4.
\end{aligned}$$

For $(p,q)=(1,6),(3,4),(3,6)$ we obtain similar estimates. Therefore

$$\begin{aligned}
& H_{S_1}(u_1,u_4)H_{S_1}(u_1,u_6)H_{S_1}(u_3,u_4)H_{S_1}(u_3,u_6) \\
\leqslant\ & c_2^8|u_1u_4u_3u_6|_{S_1}^2|R(F,G)|_{S_1}^4\left(\prod_{\nu\notin S_1}\max_{1\leqslant i\leqslant 6}(|u_i|_\nu)\right)^{16} \\
\leqslant\ & c_2^8\left(\prod_{i=1}^{6}|u_i|_{S_1}^2\right)|R(F,G)|_{S_1}^4\left(\prod_{\nu\notin S_1}\max_{1\leqslant i\leqslant 6}(|u_i|_\nu)\right)^{20} \\
=\ & c_2^8|R(F,G)|_{S_1}^8|FG|_{S_1}^{-20}, \qquad (7.2.8)
\end{aligned}$$

where in the penultimate inequality we have used that $|u_j|_{S_1}\left(\prod_{\nu\notin S_1}\max_{1\leqslant i\leqslant 6}(|u_i|_\nu)\right)\geqslant 1$ for $j=2,5$.

We still have to estimate $H_{S_1}(u_2,u_5)$. Since $(u_2,u_5)=R(F,G)^{-1}(-u_2^2u_4u_6,u_1u_3u_5^2)$, we obtain in a similar way, using corollary 2.2.11, that

$$\begin{aligned}
H_{S_1}(u_2,u_5) \leqslant\ & |R(F,G)|_{S_1}^{-1}H_{S_1}(u_2,u_1)H_{S_1}(u_2,u_3)H_{S_1}(u_4,u_5)H_{S_1}(u_6,u_5) \\
\leqslant\ & |R(F,G)|_{S_1}^{-1}H_{S_1}(u_1,u_2,u_3)^2H_{S_1}(u_4,u_5,u_6)^2 \\
\leqslant\ & c_2^4|R(F,G)|_{S_1}^{-1}|u_1u_2u_3|_{S_1}^2\left(\prod_{\nu\notin S_1}\max_{1\leqslant i\leqslant 3}(|u_i|_\nu)\right)^4\times \\
& \times|u_4u_5u_6|_{S_1}^2\left(\prod_{\nu\notin S_1}\max_{4\leqslant i\leqslant 6}(|u_i|_\nu)\right)^4 \\
\leqslant\ & c_2^4|R(F,G)|_{S_1}^3|FG|_{S_1}^{-8}. \qquad (7.2.9)
\end{aligned}$$

This leads to

$$\prod_{\substack{1\leqslant p<q\leqslant 6\\ p\not\equiv q (\mathrm{mod}\, 2)}} H_{S_1}(u_p,u_q)\leqslant c_2^{16}|R(F,G)|_{S_1}^{15}|FG|_{S_1}^{-36}.$$

Finally, for case (iv), using the same idea we deduce that

$(u_1^3, u_2^3) = \frac{u_1 u_2}{R(F,G)}(-u_1^2 u_4 u_6, u_2^2 u_3 u_5)$ and so

$$\begin{aligned}
& H_{S_1}(u_1,u_2)^3 = H_{S_1}(u_1^3,u_2^3) \\
\leqslant\ & |u_1u_2R(F,G)^{-1}|_{S_1} H_{S_1}(u_1,u_3)H_{S_1}(u_1,u_5)H_{S_1}(u_4,u_2)H_{S_1}(u_6,u_2) \\
\leqslant\ & |u_1u_2R(F,G)^{-1}|_{S_1} H_{S_1}(u_1,u_3,u_5)^2 H_{S_1}(u_2,u_4,u_6)^2 \\
\leqslant\ & c_2^4 |R(F,G)|_{S_1}^{-1} |u_1u_2|_{S_1} |u_1u_3u_5|_{S_1}^2 \left(\prod_{\nu \notin S_1} \max(|u_1|_\nu, |u_3|_\nu, |u_5|_\nu) \right)^4 \times \\
& \qquad \times |u_2u_4u_6|_{S_1}^2 \left(\prod_{\nu \notin S_1} \max(|u_2|_\nu, |u_4|_\nu, |u_6|_\nu) \right)^4 \\
\leqslant\ & c_2^4 |u_1u_2R(F,G)^3|_{S_1} |FG|_{S_1}^{-8}. \qquad (7.2.10)
\end{aligned}$$

Similar inequalities hold true for the other pairs (p,q) under consideration. Combining with (7.2.2), we have

$$\prod_{\substack{1 \leqslant p < q \leqslant 6 \\ p \not\equiv q (\mathrm{mod}\, 2)}} H_{S_1}(u_p,u_q) \leqslant c_2^{12} |R(F,G)|_{S_1}^{11} |FG|_{S_1}^{-24}.$$

This finishes our proof. □

Proof of Proposition 7.2.1. This is a combination of (7.2.3) and Lemma 7.2.2, applying the ultra-metric inequality for non-archimedean valuations. □

Remark 7.2.3. *In this section we assumed only $F, G \in K_1[X,Y]$. If we require $F, G \in \mathcal{O}_{S_1}[X,Y]$, then $|FG|_{S_1} \geqslant 1$ and so Proposition 7.2.1 gives*

$$|R(F,G)|_{S_1} \geqslant e^{-\frac{90}{17}(2g_{K_1}-1+\#S_1)} |D(F)D(G)|_{S_1}^{\frac{3}{34}}.$$

7.3 Binary forms of arbitrary degree

Again, recall that K_1 is a finite extension of $K = k(t)$ with genus g_{K_1}, and S_1 a finite set of valuations on K_1.

Theorem 7.3.1. *Assume $F, G \in K_1[X,Y]$ are two binary forms such that $\deg F = m \geqslant 3, \deg G = n \geqslant 3$, FG is square-free and has splitting field L over K_1. Then*

$$|R(F,G)|_{S_1} \geqslant e^{-\frac{10mn}{17}(\frac{2g_L-1}{[L:K_1]}+\#S_1)}|D(F)|_{S_1}^{\frac{n}{17(m-1)}}|D(G)|_{S_1}^{\frac{m}{17(n-1)}}\left(|F|_{S_1}^{\frac{15n}{17}}|G|_{S_1}^{\frac{15m}{17}}\right).$$

In particular, if F, G are irreducible, let L' be the field generated by one root of $F(X,1)$ and $G(Y,1)$, Then

$$|R(F,G)|_{S_1} \geqslant e^{-\frac{10mn}{17}(\frac{6g_{L'}+4m^2n^2}{[L':K_1]}+\#S_1)}\times$$
$$\times|D(F)|_{S_1}^{\frac{n}{17(m-1)}}|D(G)|_{S_1}^{\frac{m}{17(n-1)}}\left(|F|_{S_1}^{\frac{15n}{17}}|G|_{S_1}^{\frac{15m}{17}}\right).$$

Lemma 7.3.2 (Castelnuovo's Inequality)**.** *Let F be a function field of transcendence degree 1 over k. Let F_1, F_2 be two finite extensions of $k(t)$ and F their compositum. Suppose that*

(i) $F = F_1F_2$ is the compositum of F_1 and F_2,

(ii) $[F : F_i] = n_i$ and F_i has genus g_i (i=1,2).

Then the genus g of F is bounded by

$$g \leqslant n_1g_1 + n_2g_2 + (n_1-1)(n_2-1).$$

Proof. See Theorem 3.11.3 of [23]. □

Proof of Theorem 7.3.1. Let T be the set of valuations in L above those in S_1. Assume $F(X,Y) = \prod_{i=1}^{m}(\alpha_iX-\beta_iY), G(X,Y) = \prod_{j=1}^{n}(\gamma_jX-\delta_jY)$. We make a reduction to the case of cubic binary forms. Let

$$F_{pqr}(X,Y) = (\alpha_pX-\beta_pY)(\alpha_qX-\beta_qY)(\alpha_rX-\beta_rY) \text{ for } 1 \leqslant p < q < r \leqslant m,$$

$$G_{ijh}(X,Y) = (\gamma_iX-\delta_iY)(\gamma_jX-\delta_jY)(\gamma_hX-\delta_hY) \text{ for } 1 \leqslant i < j < h \leqslant n.$$

By Proposition 7.2.1, we have

$$|R(F_{pqr}, G_{ijh})|_T \geqslant e^{-\frac{90}{17}(2g_L-1+\#T)}|D(F_{pqr})D(G_{ijh})|_T^{\frac{3}{34}}|F_{pqr}G_{ijh}|_T^{\frac{45}{17}}.$$

Observ that

$$\prod_{1\leqslant p<q<r\leqslant m} F_{pqr} = F^{\binom{m-1}{2}},$$

$$\prod_{1\leqslant i<j<h\leqslant n} G_{ijh} = G^{\binom{n-1}{2}},$$

$$\prod_{1\leqslant p<q<r\leqslant m}\prod_{1\leqslant i<j<h\leqslant n} R(F_{pqr}, G_{ijh}) = R(F,G)^{\frac{(m-1)(n-1)(m-2)(n-2)}{4}},$$

$$\prod_{1\leqslant p<q<r\leqslant m} D(F_{pqr}) = D(F)^{m-2},$$

$$\prod_{1\leqslant i<j<h\leqslant n} D(G_{ijh}) = D(G)^{n-2}.$$

Hence, by taking the products, we deduce that

$$|R(F,G)|_T \geqslant e^{-\frac{10mn}{17}(2g_L-1+\#T)}|D(F)|_T^{\frac{n}{17(m-1)}}|D(G)|_T^{\frac{m}{17(n-1)}}\left(|F|_T^{\frac{15n}{17}}|G|_T^{\frac{15m}{17}}\right).$$

As $\#T \leqslant [L:K_1]\#S_1$ and $|x|_T = |x|_{S_1}^{[L:K_1]}$ for $x \in K_1$, we conclude that

$$|R(F,G)|_{S_1} \geqslant e^{-\frac{10mn}{17}(\frac{2g_L-1}{[L:K_1]}+\#S_1)}|D(F)|_{S_1}^{\frac{n}{17(m-1)}}|D(G)|_{S_1}^{\frac{m}{17(n-1)}}\left(|F|_{S_1}^{\frac{15n}{17}}|G|_{S_1}^{\frac{15m}{17}}\right).$$

If F, G are irreducible, write $F = a\prod_{i=1}^{m}(X-\gamma_i Y), G = b\prod_{j=1}^{n}(X-\delta_j Y)$, then all fields $K_1(\gamma_i, \delta_j)$ are isomorphic. Without loss of generality, assume $L' = K_1(\gamma_1, \delta_1)$.

Let

$$F_{pqr}(X,Y) = (X-\gamma_p Y)(X-\gamma_q Y)(X-\gamma_r Y) \text{ for } 1 \leqslant p < q < r \leqslant m,$$

$$G_{ijh}(X,Y) = (X-\delta_i Y)(X-\delta_j Y)(X-\delta_h Y) \text{ for } 1 \leqslant i < j < h \leqslant n.$$

Let $M = K_1(\gamma_p, \gamma_q, \gamma_r, \delta_i, \delta_j, \delta_h)$ and T be the set of valuations of M above those in S_1. Here we omit the subscript because all such field are isomorphic and all T have the same cardinality. By Proposition 7.2.1,

$$|R(F_{pqr}, G_{ijh})|_T \geqslant e^{-\frac{90}{17}(2g_M - 1 + \#T)} |D(F_{pqr})D(G_{ijh})|_T^{\frac{3}{34}} |F_{pqr}G_{ijh}|_T^{\frac{45}{17}}.$$

Applying Lemma 7.3.2 to $L_1 = K_1(\gamma_p, \delta_p, \gamma_q, \delta_q)$ and its subfields $K_1(\gamma_p, \delta_p), K_1(\gamma_q, \delta_q)$ we obtain

$$g_{L_1} \leqslant 2dg_{L'} + (d-1)^2,$$

where

$$\begin{aligned} d &= [L_1 : K_1(\gamma_q, \delta_q)] \\ &= [L_1 : K_1(\gamma_p, \delta_p)] \\ &\leqslant [L_1 : K_1(\gamma_p, \delta_p, \gamma_q)][K_1(\gamma_p, \delta_p, \gamma_q) : K_1(\gamma_p, \delta_p)] \\ &\leqslant [K_1(\delta_p, \delta_q) : K_1(\delta_p)][K_1(\gamma_p, \gamma_q) : K_1(\gamma_p)] \\ &\leqslant (m-1)(n-1) \\ &< mn. \end{aligned}$$

Observing that $[M : L_1] \leqslant d$ and $[M : K_1(\gamma_r, \delta_r)] \leqslant d^2$, and applying Lemma 7.3.2 to M and its subfields $L_1, K_1(\gamma_r, \delta_r)$ we obtain

$$g_M \leqslant [M : K_1(\gamma_r, \delta_r)]g_{L'} + [M : L_1]g_{L_1} + (d-1)(d^2-1).$$

Hence

$$\begin{aligned} g_M &\leqslant 3[M : K_1(\gamma_p, \delta_p)]g_{L'} + (2d+1)(d-1)^2 \\ &< \frac{3[M : K_1]}{[L' : K_1]} g_{L'} + 2m^2n^2 \frac{[M : K_1]}{[L' : K_1]} \end{aligned}$$

and

$$|R(F_{pqr}, G_{ijh})|_{S_1} \geqslant e^{-\frac{90}{17}(\frac{6g_{L'} + 4m^2n^2}{[L' : K_1]} + \#S_1)} |D(F_{pqr})D(G_{ijh})|_{S_1}^{\frac{3}{34}} |F_{pqr}G_{ijh}|_{S_1}^{\frac{45}{17}}.$$

By taking the products over all triple (p, q, r) and (i, j, h) we deduce that

$$|R(F,G)|_{S_1} \geqslant e^{-\frac{10mn}{17}(\frac{6g_{L'}+4m^2n^2}{[L':K_1]}+\#S_1)}|D(F)|_{S_1}^{\frac{n}{17(m-1)}}|D(G)|_{S_1}^{\frac{m}{17(n-1)}}\left(|F|_{S_1}^{\frac{15n}{17}}|G|_{S_1}^{\frac{15m}{17}}\right).$$

This completes the proof. □

Corollary 7.3.3. *Let $F, G \in \mathcal{O}_{S_1}[X, Y]$ be two binary forms such that $\deg F = m \geqslant 3, \deg G = n \geqslant 3$, FG is square-free and has splitting field L over K_1. Then*

$$|R(F,G)|_{S_1} \geqslant e^{-\frac{10mn}{17}(\frac{2g_L-1}{[L:K_1]}+\#S_1)}|D(F)|_{S_1}^{\frac{n}{17(m-1)}}|D(G)|_{S_1}^{\frac{m}{17(n-1)}}.$$

In particular, if F, G are irreducible, let L' be the field generated by one root of $F(X, 1)$ and $G(Y, 1)$ instead, then

$$|R(F,G)|_{S_1} \geqslant e^{-\frac{10mn}{17}(\frac{6g_{L'}+4m^2n^2}{[L':K_1]}+\#S_1)}|D(F)|_{S_1}^{\frac{n}{17(m-1)}}|D(G)|_{S_1}^{\frac{m}{17(n-1)}},$$

Proof. Since $F, G \in \mathcal{O}_{S_1}[X, Y]$, $|F|_{S_1} \geqslant 1, |G|_{S_1} \geqslant 1$. Then apply Theorem 7.3.1. In particular, if F, G are irreducible, then $F(1, 0), G(1, 0) \in \mathcal{O}_{S_1}$ and the rest is clear. □

Remark 7.3.4. *Theorem 7.3.1 and Corollary 7.3.3 do not hold if $m = 2$ or $n = 2$. For instance, if $m = 2, n > 2$, take $F = X^2 - (t^2 - 1)Y^2, G = \prod_{i=1}^{n}(a_iX - b_iY)$ where $a_i, b_i\ (i = 1, \ldots, n)$ satisfy $a_i^2 - (t^2 - 1)b_i^2 = 1$. Say, $u_j, v_j \in k[t]$ are the unique solution of $u_j + v_j\sqrt{t^2-1} = (t + \sqrt{t^2-1})^j\ (j \in \mathbb{N})$ and $a_i = u_{l_i}, b_i = v_{l_i}\ (i = 1, \ldots, n)$ with $l_1 < \cdots < l_n$. Then $R(F, G) = 1, D(F) = 4(t^2 - 1)$ and*

$$\begin{aligned} D(G) &= \prod_{1\leqslant i<j\leqslant n}(a_ib_j - a_jb_i)^2 \\ &= \frac{1}{4(t^2-1)}\prod_{1\leqslant i<j\leqslant n}\left((t+\sqrt{t^2-1})^{l_i-l_j} - (t-\sqrt{t^2-1})^{l_i-l_j}\right)^2 \\ &= \frac{1}{(t^2-1)}\prod_{1\leqslant i<j\leqslant n}u_{l_i-l_j}^2. \end{aligned}$$

It is easy to check $\deg(u_j) = j$, *hence* $|D(G)|_\infty \to \infty$ *while* $\max\limits_{1 \leqslant i < j \leqslant n} (l_i - l_j) \to \infty$. *This gives a counter-example.*

7.4 A result on Thue-Mahler equations

The idea of the following sections comes from [10]. We work out an analogue for function fields. Let L be a finite extension of $K = k(t)$, and $T \subset M_L$ a finite set of valuations. As in K, for a binary form F with coefficients $a_0, \dots, a_n$, put

$$H_T(F) = \prod_{\omega \in T} \max(|a_0|_\omega, \dots, |a_n|_\omega),$$

$$H_L(F) = \prod_{\omega \in L} \max(|a_0|_\omega, \dots, |a_n|_\omega),$$

$$H(F) = H_L(F)^{1/[L:K]}.$$

Lemma 7.4.1. *Let* $F(X,Y) \in L[X,Y]$ *be a binary form of degree* $m \geqslant 3$ *with* $D(F) \neq 0$. *Let* $A \geqslant 1$ *and suppose* F *splits in* L. *Then every solution* $(x,y) \in L^2$ *of the Thue-Mahler equation*

$$|F(x,y)|_T = A \tag{7.4.1}$$

satisfies

$$H_T(x,y) \leqslant e^{2g_L - 1 + \#T} \prod_{\nu \notin T} \max(|x|_\nu, |y|_\nu)^2 \left(A \frac{H_L^2(F)}{H_T(F)} \right)^{3/m}.$$

In particular, if $F(X,Y) \in \mathcal{O}_T[X,Y]$ *and* $(x,y) \in \mathcal{O}_T^2$, *then*

$$H_T(x,y) \leqslant e^{2g_L - 1 + \#T} \left(A \cdot H_T(F) \right)^{3/m}.$$

Proof. Suppose we have a factorization $F(X,Y) = \prod_{i=1}^{m}(\alpha_i X + \beta_i Y)$ in L.

Put $\Delta_{ij} = \alpha_i\beta_j - \alpha_j\beta_i$ for $i,j = 1,\dots,m$. Then

$$|\Delta_{ij}|_\omega \leqslant \max(|\alpha_i|_\omega, |\beta_i|_\omega)\max(|\alpha_j|_\omega, |\beta_j|_\omega) \text{ for } \omega \in M_L.$$

Fix an arbitrary triple $\{r,s,t\} \subset \{1,\dots,m\}$ and $(x,y) \in L^2$, and put

$$A_r = \Delta_{st}(\alpha_r X + \beta_r Y),\ a_r = A_r(x,y)$$

and similarly for A_s, A_t, a_s, a_t. Then observe that

$$\begin{aligned} A_r + A_s + A_t &= 0, \\ a_r + a_s + a_t &= 0. \end{aligned} \tag{7.4.2}$$

Applying Corollary 2.2.11 to (7.4.2) , we obtain

$$H_T(a_r,a_s,a_t) \leqslant e^{\max(2g_L-2+\#T,0)}|a_ra_sa_t|_T\left(\prod_{\omega\notin T}\max(|a_r|_\omega,|a_s|_\omega,|a_t|_\omega)\right)^2, \tag{7.4.3}$$

where

$$|a_ra_sa_t|_T = |\Delta_{rs}\Delta_{st}\Delta_{tr}|_T \prod_{i\in\{r,s,t\}} |\alpha_i x + \beta_i y|_T,$$

$$\prod_{\omega\notin T}\max(|a_r|_\omega,|a_s|_\omega,|a_t|_\omega) \leqslant \prod_{\omega\notin T}\Big(\max(|x|_\omega,|y|_\omega)\prod_{i\in\{r,s,t\}}\max(|\alpha_i|_\omega,|\beta_i|_\omega)\Big).$$

Also,

$$\begin{aligned} \Delta_{rs}\Delta_{st}\Delta_{tr}X &= \Delta_{tr}\beta_s A_r - \Delta_{st}\beta_r A_s, \\ \Delta_{rs}\Delta_{st}\Delta_{tr}Y &= -\Delta_{tr}\alpha_s A_r + \Delta_{st}\alpha_r A_s. \end{aligned} \tag{7.4.4}$$

Then for each solution (x,y) of (7.4.1) and each $\omega \in T$,

$$|\Delta_{rs}\Delta_{st}\Delta_{tr}|_\omega \max(|x|_\omega,|y|_\omega) \leqslant \left(\prod_{i\in\{r,s,t\}}\max(|\alpha_i|_\omega,|\beta_i|_\omega)\right)\max(|a_r|_\omega,|a_s|_\omega).$$

Hence

$$|\Delta_{rs}\Delta_{st}\Delta_{tr}|_T H_T(x,y) \leqslant \Big(\prod_{i\in\{r,s,t\}} H_T(\alpha_i,\beta_i)\Big) H_T(a_r,a_s). \qquad (7.4.5)$$

Noticing that $H_T(a_r,a_s) = H_T(a_r,a_s,a_t)$, combining (7.4.5) with (7.4.3), we deduce that

$$\begin{aligned} H_T(x,y) \ \leqslant\ & e^{2g_L-1+\#T}\prod_{\omega\notin T}\max(|x|_\omega,|y|_\omega)^2 \\ & \times \prod_{i\in\{r,s,t\}}\Big(|\alpha_i x+\beta_i y|_T \prod_{\omega\in T}\max(|\alpha_i|_\omega,|\beta_i|_\omega)\prod_{\omega\notin T}\max(|\alpha_i|_\omega,|\beta_i|_\omega)^2\Big). \end{aligned}$$

However, by Gauss' lemma

$$\prod_{i=1}^{m}\max(|\alpha_i|_\omega,|\beta_i|_\omega) = |F|_\omega.$$

Then by taking the products over all triples $\{r,s,t\}\subset\{1,\dots,m\}$ and (7.4.1), we deduce that

$$H_T(x,y) \leqslant e^{2g_L-1+\#T}\prod_{\omega\notin T}\max(|x|_\omega,|y|_\omega)^2\left(A\frac{H_L^2(F)}{H_T(F)}\right)^{3/m}, \qquad (7.4.6)$$

if $F(X,Y)\in\mathcal{O}_T[X,Y]$ and $(x,y)\in\mathcal{O}_T^2$, then we get $H_T(F)\geqslant 1, \max(|x|_\omega,|y|_\omega)\leqslant 1$ for $\omega\notin T$, hence

$$H_T(x,y)\leqslant e^{2g_L-1+\#T}\left(A\cdot H_T(F)\right)^{3/m}.$$

□

7.5 Lower bounds for resultants in terms of heights

In this section we estimate the resultants from below in terms of heights. Again let $K = k(t)$, and let S be a finite set of valuations of M_K. Further, let $F, G \in K[X,Y]$ be two binary forms of degree m, n respectively. Recall that for $U = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with $\det U \neq 0$, define $F_U(X,Y) = F(aX+bY, cX+dY)$ and the same for G_U. Then

$$R(F_U, G_U) = (\det U)^{mn} R(F,G).$$

By the definition of resultant (1.1.1) and the non-archimedean property of the absolute values $|\cdot|_\nu$ on K, we have

$$|R(F,G)|_\nu \leqslant |F|_\nu^n |G|_\nu^m \text{ for } \nu \in M_K,$$

and hence

$$|R(F,G)|_S \leqslant |F|_S^n |G|_S^m.$$

Theorem 7.5.1. *Let $m, n > 2$ and let F, G be binary forms in $\mathcal{O}_S[X,Y]$ such that FG is square-free and with splitting field L over K. Then there exists $U \in GL_2(\mathcal{O}_S)$ such that*

$$|R(F,G)|_S \geqslant c(m,n,S,L)^{-1} H_S(G_U)^{\frac{m}{717}} H_S(F_U)^{\frac{n}{717}},$$

where

$$c(m,n,S,L) = \exp\left(\frac{422mn(2g_L-1)}{717[L:K]} + mn(4m+4n+433)\frac{\#S}{717}\right).$$

Lemma 7.5.2. *Let $F \in \mathcal{O}_S[X,Y]$ be a binary form of degree m with non-zero discriminant. Then there exists $U \in GL_2(\mathcal{O}_S)$such that*

$$|D(F)|_S \geqslant e^{-\frac{m-1}{21}\left((m^2+6m-7)\#S + \frac{(5m-5)(2g_L-1)}{24[L:K]}\right)} H_S(F_U)^{\frac{m-1}{21}}.$$

Proof. This follows from Theorem 5.3.2, observing that $H_S(F') \leqslant H^*(F')$ for any $F' \in \mathcal{O}_S[X,Y]$. □

Proof of Theorem 7.5.1. Without loss of generality, assume that $|D(F)|_S^{\frac{n}{m-1}} \leqslant |D(G)|_S^{\frac{m}{n-1}}$.

By Lemma 7.5.2, there exists $U \in \mathrm{GL}_2(\mathcal{O}_S)$ such that

$$|D(F)|_S \geqslant e^{-\frac{m-1}{21}\left((m^2+6m-7)\#S+\frac{(5m-5)(2g_L-1)}{24[L:K]}\right)} H_S(F_U)^{\frac{m-1}{21}}.$$

Combining this with Corollary 7.3.3, we deduce that

$$\begin{aligned} |R(F,G)|_S &\geqslant e^{-\frac{10mn}{17}(\frac{2g_L-1}{[L:K]}+\#S)}|D(F)|_S^{\frac{n}{17(m-1)}}|D(G)|_S^{\frac{m}{17(n-1)}} \\ &\geqslant e^{-\frac{10mn}{17}(\frac{2g_L-1}{[L:K]}+\#S)}|D(F)|_S^{\frac{2n}{17(m-1)}} \\ &\geqslant C(L,S)H_S(F_U)^{\frac{2n}{357}}, \end{aligned} \tag{7.5.1}$$

where $C(L,S) = e^{-\frac{10mn}{17}(\frac{2g_L-1}{[L:K]}+\#S)-\frac{2n}{357}\left((m^2+6m-7)\#S+\frac{(5m-5)(2g_L-1)}{24[L:K]}\right)}$.

On the other hand, let $T \subset M_L$ be the set of valuations above those in S. Assume F_U, G_U factor in L as

$$F_U(X,Y) = \prod_{i=1}^{m}(\alpha_i X + \beta_i Y),$$

$$G_U(X,Y) = \prod_{j=1}^{n}(\gamma_j X + \delta_j Y).$$

Then

$$|R(F,G)|_S^{[L:K]} = |R(F_U,G_U)|_S^{[L:K]} = |R(F_U,G_U)|_T = \prod_{j=1}^{n}|F_U(\delta_j,-\gamma_j)|_T.$$

By Lemma 7.4.1, we obtain for $j = 1,\dots,n$ that

$$H_T(\delta_j,-\gamma_j) \leqslant e^{2g_L-1+\#T}\prod_{\omega\notin T}\max(|\delta_j|_\omega,|\gamma_j|_\omega)^2\left(|F_U(\delta_j,-\gamma_j)|_T\frac{H_L(F_U)^2}{H_T(F_U)}\right)^{3/m}.$$

Combining this with Gauss' lemma, we deduce that

$$\begin{aligned}
H_T(G_U) &= \prod_{j=1}^{n} H_T(\gamma_j, \delta_j) \\
&\leqslant e^{n(2g_L-1+\#T)} \prod_{\omega\notin T}\prod_{j=1}^{n} \max(|\delta_j|_\omega, |\gamma_j|_\omega)^2 \Big(\prod_{j=1}^{n} |F_U(\delta_j, -\gamma_j)|_T \frac{H_L(F_U)^{2n}}{H_T(F_U)^n}\Big)^{3/m} \\
&= e^{n(2g_L-1+\#T)} \prod_{\omega\notin T} |G_U|_\omega^2 \Big(|R(F_U, G_U)|_T \frac{H_L(F_U)^{2n}}{H_T(F_U)^n}\Big)^{3/m}.
\end{aligned}$$

Therefore

$$H_S(G_U) \leqslant e^{n(\frac{2g_L-1}{[L:K]}+\#S)} \prod_{\nu\notin S} |G_U|_\nu^2 \Big(|R(F,G)|_S \frac{H_K(F_U)^{2n}}{H_S(F_U)^n}\Big)^{3/m}.$$

Noticing that $F_U, G_U \in \mathcal{O}_S[X,Y]$, we obtain

$$H_S(G_U) \leqslant e^{n(\frac{2g_L-1}{[L:K]}+\#S)} \Big(|R(F,G)|_S H_S(F_U)^n\Big)^{3/m}. \tag{7.5.2}$$

Combining (7.5.1) with (7.5.2) we conclude that

$$\begin{aligned}
H_S(G_U)^m H_S(F_U)^n &\leqslant e^{mn(\frac{2g_L-1}{[L:K]}+\#S)} |R(F,G)|_S^3 H_S(F_U)^{4n} \\
&\leqslant c|R(F,G)|_S^{717},
\end{aligned}$$

where

$$\begin{aligned}
c &= \exp\Big(\Big(421mn + \frac{5n(m-1)}{6}\Big)\frac{2g_L-1}{[L:K]} + \Big(421mn + 4n(m^2+6m-7)\Big)\#S\Big) \\
&< \exp\Big(\frac{422mn(2g_L-1)}{[L:K]} + mn(4m+4n+433)\#S\Big).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&H_S(G_U)^m H_S(F_U)^n \\
&\quad < \exp\Big(\frac{422mn(2g_L-1)}{[L:K]} + mn(4m+4n+433)\#S\Big)|R(F,G)|_S^{717}.
\end{aligned}$$

□

Corollary 7.5.3. *If F, G are irreducible over K, let L_1 be an extension of K generated by a root of $F(X,1)$ and L_2 an extension of K by a root of $G(X,1)$, and suppose that L_i has genus g_i for $i = 1, 2$. Then*

$$|R(F,G)|_S \geqslant c(m,n,S,L_1,L_2)^{-1} H_S(G_U)^{\frac{m}{717}} H_S(F_U)^{\frac{n}{717}},$$

where

$$c(m,n,S,L_1,L_2) = \exp\left(\frac{422mn(m+n-5+2g_1+2g_2)}{717} + mn(4m+4n+433)\frac{\#S}{717}\right).$$

Proof. When F, G are irreducible, the claim is a combination of Theorem 7.5.1 and (5.1.5). □

Chapter 8

Distances between algebraic functions

Let $K = k(t)$. In section 8.1 we give a lower bound for the distance tetween two roots of a polynomial $f \in k[t][X]$, and in section 8.3 we derive such a lower bound between roots of different polynomials. We follow [9], [10] where similar results have been derived over number fields.

8.1 Root separation of polynomials

Let $K = k(t)$ and let $f \in K[X]$ be a polynomial of degree $n \geqslant 4$ with splitting field L and non-zero discriminant. Assume that $f = a \prod_{i=1}^{n} (X - \gamma_i)$ with $a \in K^*$ and $\gamma_i \in L$ for $i = 1, \dots, n$. Let S be a finite set of valuations on K and let T be the set of valuations on L above those in S. For each $\nu \in S$ fix a prolongation of $|\cdot|_\nu$ to L, also denoted by $|\cdot|_\nu$. Define

$$\Delta_S(f) := \prod_{\nu \in S} \min_{1 \leqslant i < j \leqslant n} \frac{|\gamma_i - \gamma_j|_\nu}{\max(1, |\gamma_i|_\nu) \max(1, |\gamma_j|_\nu)}.$$

Since L/K is a Galois extension, this quantity $\Delta_S(f)$ is independent of the choices of the extensions of $|\cdot|_\nu$ to L. To be specific, by (1.4.3) we have for $\omega \in \mathcal{A}(\nu)$ and $\sigma \in \mathcal{E}(\omega|\nu)$ that

$$\begin{aligned}
&\min_{1\leqslant i<j\leqslant n} \frac{|\gamma_i-\gamma_j|_\omega}{\max(1,|\gamma_i|_\omega)\max(1,|\gamma_j|_\omega)}\\
&= \left(\min_{1\leqslant i<j\leqslant n} \frac{|\sigma(\gamma_i-\gamma_j)|_\nu}{\max(1,|\sigma(\gamma_i)|_\nu)\max(1,|\sigma(\gamma_j)|_\nu)}\right)^{g_\nu}\\
&= \left(\min_{1\leqslant i<j\leqslant n} \frac{|\gamma_{\sigma(i)}-\gamma_{\sigma(j)}|_\nu}{\max(1,|\gamma_{\sigma(i)}|_\nu)\max(1,|\gamma_{\sigma(j)}|_\nu)}\right)^{g_\nu}\\
&= \left(\min_{1\leqslant i<j\leqslant n} \frac{|\gamma_i-\gamma_j|_\nu}{\max(1,|\gamma_i|_\nu)\max(1,|\gamma_j|_\nu)}\right)^{g_\nu},
\end{aligned}$$

since $\sigma \in \mathrm{Gal}(L/K)$ acts on $1,\ldots,n$ as a permutation and $g_\nu = [L_\omega : K_\nu]$ is independent of ω. Hence

$$\Delta_S(f) = \prod_{\omega\in T}\left(\min_{1\leqslant i<j\leqslant n} \frac{|\gamma_i-\gamma_j|_\omega}{\max(1,|\gamma_i|_\omega)\max(1,|\gamma_j|_\omega)}\right)^{1/[L:K]}. \qquad (8.1.1)$$

Put $H(f) = \prod_{\nu\in M_K} |f|_\nu$. Then clearly $H(f) \geqslant 1$.

Theorem 8.1.1. *Let* $c_4(n) = \exp\left(\frac{(n-1)\left((n+11)\#S-5\right)}{20+1/n}\right)$. *We have*

$$\Delta_S(f) \geqslant c_4(n)^{-1} H(f)^{-n+1+\frac{n}{40n+2}}.$$

Proof. Homogenize $f = a_0X^n + a_1X^{n-1} + \cdots + a_n$ and choose

$$F(X,Y) = b(a_0X^n + a_1X^{n-1}Y + \cdots + a_nY^n)$$

with $b \in K^*$ such that

$$|b|_\infty = |f|_\infty^{-1} H(f),\ |b|_\nu = |f|_\nu^{-1} \text{ for } \nu \neq \nu_\infty.$$

The existence of b is guaranteed because $\prod_{\nu\in M_K} |f|_\nu^{-1} H(f) = 1$. So we get $F \in \mathcal{O}_S[X,Y]$, $|F|_\infty = H(f)$ and hence

$$H^*(F) = \max(1, |F|_\infty) = H(f).$$

Factor F in L as $F = \prod_{i=1}^{n} (\alpha_i X + \beta_i Y)$. Then $\gamma_i = -\frac{\beta_i}{\alpha_i}$. Put

$$\delta_\omega = \min_{1 \leqslant i < j \leqslant n} \frac{|\alpha_i \beta_j - \alpha_j \beta_i|_\omega}{|\alpha_i, \beta_i|_\omega |\alpha_j, \beta_j|_\omega} \ (\omega \in T).$$

Then

$$\Delta_S(f) = \prod_{\omega \in T} \delta_\omega^{1/[L:K]}.$$

Let $F^*(X,Y) = F(aX + bY, cX + dY)$ with $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}\,(2, \mathcal{O}_S)$ be such that F^* is reduced. Then $F^*(X,Y) = \prod_{i=1}^{n} (\alpha_i^* X + \beta_i^* Y)$ where $(\alpha_i^*, \beta_i^*) = (\alpha_i, \beta_i) \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right), i = 1, \ldots, n$.

Now for $\omega \in T$ put $f_{i\omega} := |\alpha_i, \beta_i|_\omega, f_{i\omega}^* := |\alpha_i^*, \beta_i^*|_\omega$ and $\zeta_{ij\omega} := |\alpha_i \beta_j - \alpha_j \beta_i|_\omega$. Then $\prod_{i=1}^{n} f_{i\omega} = |F|_\omega, \prod_{i=1}^{n} f_{i\omega}^* = |F^*|_\omega$ and $\prod_{1 \leqslant i < j \leqslant n} \zeta_{ij\omega} = |D(F)|_\omega^{1/2}$.

By the ultrametric inequality we have $\zeta_{ij\omega} \leqslant f_{i\omega} f_{j\omega}$, and

$$\zeta_{ij\omega} = |ad - bc|_\omega^{-1} |\alpha_i^* \beta_j^* - \alpha_j^* \beta_i^*|_\omega \leqslant |ad - bc|_\omega^{-1} f_{i\omega}^* f_{j\omega}^*,$$

So

$$\zeta_{ij\omega} \leqslant \min(f_{i\omega} f_{j\omega}, |ad - bc|_\omega^{-1} f_{i\omega}^* f_{j\omega}^*) \ \text{ for } 1 \leqslant i < j \leqslant n, \omega \in T. \quad (8.1.2)$$

We are going to bound δ_ω from below for each $\omega \in T$. Let $\omega \in T$, and assume, without loss of generality, that $\delta_\omega = \frac{\zeta_{12\omega}}{f_{1\omega} f_{2\omega}}$. Then

$$\delta_\omega \geqslant \frac{\zeta_{12\omega}}{f_{1\omega} f_{2\omega}} \prod_{\substack{1 \leqslant i < j \leqslant n \\ (i,j) \neq (1,2)}} \frac{\zeta_{ij\omega}}{\min(f_{i\omega} f_{j\omega}, |ad - bc|_\omega^{-1} f_{i\omega}^* f_{j\omega}^*)} = \frac{|D(F)|_\omega^{1/2}}{\Lambda_\omega},$$

with $\Lambda_\omega = f_{1\omega} f_{2\omega} \prod_{\substack{1 \leqslant i < j \leqslant n \\ (i,j) \neq (1,2)}} \min(f_{i\omega} f_{j\omega}, |ad - bc|_\omega^{-1} f_{i\omega}^* f_{j\omega}^*)$.

We claim that

$$\Lambda_\omega \leqslant |F|_\omega |F^*|_\omega^{n-2} |ad - bc|_\omega^{-n(n-2)/2}. \quad (8.1.3)$$

Then

$$\delta_\omega \geqslant \frac{|D(F)|_\omega^{1/2}|ad-bc|_\omega^{n(n-2)/2}}{|F|_\omega |F^*|_\omega^{n-2}}.$$

By the Main Theorem, we have

$$|D(F)|_S^{1/2} \geqslant H^*(F^*)^{n/(40n+2)} e^{\frac{(1-n)\big((n+11)\#S-5\big)}{20+1/n}}. \tag{8.1.4}$$

Using $ad-bc \in \mathcal{O}_S^*$, $H_S(F) \leqslant H^*(F), H_S(F^*) = H^*(F^*) \leqslant H^*(F) = H(f)$, we deduce that

$$\begin{aligned}
\Delta_S(f) &\geqslant \left(\prod_{\omega\in T} \frac{|D(F)|_\omega^{1/2}|ad-bc|_\omega^{n(n-2)/2}}{|F|_\omega |F^*|_\omega^{n-2}}\right)^{1/[L:K]} \\
&= \frac{|D(F)|_S^{1/2}}{H_S(F)H_S(F^*)^{n-2}} \\
&\geqslant \exp\left(-\frac{(n-1)\big((n+11)\#S-5\big)}{20+1/n}\right)\frac{1}{H(f)}H^*(F^*)^{\frac{n}{40n+2}-n+2} \\
&\geqslant \exp\left(-\frac{(n-1)\big((n+11)\#S-5\big)}{20+1/n}\right)H(f)^{-n+1+\frac{n}{40n+2}}. \qquad (8.1.5)
\end{aligned}$$

Finally, to prove (8.1.3), we have to distinguish two cases. First let $n \geqslant 4$ be even. Take $I = \{(1,2),\ldots,(n-1,n)\}$. Then

$$\begin{aligned}
\Lambda_\omega &\leqslant \prod_{i=1}^n f_{i\omega} \prod_{\substack{1\leqslant i<j\leqslant n \\ (i,j)\notin I}} |ad-bc|_\omega^{-1} f_{i\omega}^* f_{j\omega}^* \\
&= \prod_{i=1}^n f_{i\omega}\Big(\prod_{i=1}^n f_{i\omega}^*\Big)^{n-2}|ad-bc|_\omega^{-n(n-2)/2} \\
&= |F|_\omega |F^*|_\omega^{n-2}|ad-bc|_\omega^{-n(n-2)/2}.
\end{aligned}$$

Next let $n \geqslant 5$ be odd. Take

$$I = \{(1,2),\ldots,(n-2,n-1),(n-2,n),(n-1,n)\}.$$

Then

$$\begin{aligned}\Lambda_\omega &\leqslant \prod_{i=1}^{n-3} f_{i\omega} \prod_{n-2\leqslant i<j\leqslant n} \left(f_{i\omega} f_{j\omega} f_{i\omega}^* f_{j\omega}^* |ad-bc|_\omega^{-1}\right)^{1/2} \prod_{\substack{1\leqslant i<j\leqslant n\\ (i,j)\notin I}} |ad-bc|_\omega^{-1} f_{i\omega}^* f_{j\omega}^* \\ &= \prod_{i=1}^{n} f_{i\omega} \Big(\prod_{i=1}^{n} f_{i\omega}^*\Big)^{n-2} |ad-bc|_\omega^{-n(n-2)/2} \\ &= |F|_\omega |F^*|_\omega^{n-2} |ad-bc|_\omega^{-n(n-2)/2}.\end{aligned}$$

□

As a direct consequence, we obtain the following result on simultaneous root separation for various absolute values.

Corollary 8.1.2. *We have*

$$\prod_{\nu\in S} \min_{1\leqslant i<j\leqslant n} |\gamma_i-\gamma_j|_\nu \geqslant \exp\Big(\frac{-(n-1)\big((n+11)\#S-5\big)}{20}\Big) H(f)^{-n+1+\frac{n}{40n+2}}.$$

Proof. Since the denominator of $\Delta_S(f)$ is at least 1, this is a direct consequence of Theorem 8.1.1 and the fact $|x|_S^{[L:K]} = |x|_T$. □

Corollary 8.1.3.

$$\Delta_S(f) \geqslant \exp\Big(-\frac{n-1}{100}\big(5n(n+7)\#S + \frac{2g_L-1}{[L:K]}\big)\Big) H(f)^{-n+1+\frac{n}{42}}.$$

Proof. It is similar with proof of Theorem 8.1.1, but replace (8.1.4) by using Theorem 5.3.2. □

8.2 Two lemmas

We need some preparations for the next section where we consider distances between algebraic function that are roots of different polynomials.

Let $K = k(t)$. Let $H^*(\gamma) = \prod_{\omega \in M_L} \max(1, |\gamma|_\omega)^{1/[L:K]}$ for any $\gamma \in L$ algebraic over K. This is independent of the choice of L.

Let ξ, η be distinct and algebraic over K. Let $L = K(\xi, \eta)$ and T a finite set of valuations on L. Define

$$\Delta_T(\xi, \eta) := \left(\prod_{\omega \in T} \frac{|\xi - \eta|_\omega}{\max(1, |\xi|_\omega) \max(1, |\eta|_\omega)} \right)^{1/[L:K]}.$$

Then clearly

$$\begin{aligned} \Delta_T(\xi, \eta) &= \left(\prod_{\omega \notin T} \frac{\max(1, |\xi|_\omega) \max(1, |\eta|_\omega)}{|\xi - \eta|_\omega} \right)^{1/[L:K]} H^*(\xi)^{-1} H^*(\eta)^{-1} \\ &\geqslant H^*(\xi)^{-1} H^*(\eta)^{-1}. \end{aligned}$$

This is a type of Liouville-type inequality. Recall that for a matrix $A = (a_{ij})_{i,j}$, we have defined its ν-value $|A|_\nu = \max_{i,j}(|a_{ij}|_\nu)$ for $\nu \in M_K$. In this way, we also define

$$H_S(A) = \prod_{\nu \in S} |A|_\nu.$$

Lemma 8.2.1. *Let $F(X,Y) \in \mathcal{O}_S[X,Y]$ be a binary form of degree $n \geqslant 3$ with non-zero discriminant. Then for any $U \in \mathrm{GL}(2, \mathcal{O}_S)$, we have*

$$\frac{H_S(F_U)}{H_S(F)} \leqslant H_S(U) \leqslant \big(H_S(F) H_S(F_U)\big)^{3/n}.$$

Proof. Let T be the set of valuations on the splitting field L lying above the valuations in S, write $F(X,Y) = a_0 \prod_{i=1}^{n} (\alpha_i X + \beta_i Y)$ with $a_0 \in K^*, \alpha_i, \beta_i \in \mathcal{O}_T$ and $F_U(X,Y) = a_0 \prod_{i=1}^{n} (\alpha_i^* X + \beta_i^* Y)$ with

$$(\alpha_i^*, \beta_i^*) = (\alpha_i, \beta_i) U, \; i = 1, \ldots, n.$$

Let $U = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Then

$$\begin{cases} a\alpha_i + c\beta_i = \alpha_i^* \\ b\alpha_i + d\beta_i = \beta_i^* \end{cases} \text{for } i = 1, \ldots, n.$$

From the non-archimedean property, it easily follows that

$$\max(|\alpha_i^*|_\omega, |\beta_i^*|_\omega) \leqslant |U|_\omega \max(|\alpha_i|_\omega, |\beta_i|_\omega) \text{ for } \omega \in T,$$

hence by Gauss' lemma we have

$$H_T(F_U) \leqslant |U|_T H_T(F),$$

which gives

$$H_S(F_U) \leqslant |U|_S H_S(F).$$

Take any three indices i, j, l and consider the system of equations

$$A\mathbf{x} = \mathbf{0}, \tag{8.2.1}$$

where $\mathbf{x} = (x_1, \ldots, x_7)^{\mathrm{T}}$ and

$$A = \begin{pmatrix} \alpha_i & \beta_i & 0 & 0 & \alpha_i^* & 0 & 0 \\ 0 & 0 & \alpha_i & \beta_i & \beta_i^* & 0 & 0 \\ \alpha_j & \beta_j & 0 & 0 & 0 & \alpha_j^* & 0 \\ 0 & 0 & \alpha_j & \beta_j & 0 & \beta_j^* & 0 \\ \alpha_l & \beta_l & 0 & 0 & 0 & 0 & \alpha_l^* \\ 0 & 0 & \alpha_l & \beta_l & 0 & 0 & \beta_l^*. \end{pmatrix}$$

Put $X = \left(\begin{smallmatrix} x_1 & x_3 \\ x_2 & x_4 \end{smallmatrix}\right)$. Then

$$\begin{aligned} -x_5(\alpha_i^*, \beta_i^*) &= (\alpha_i, \beta_i)X, \\ -x_6(\alpha_j^*, \beta_j^*) &= (\alpha_j, \beta_j)X, \\ -x_7(\alpha_l^*, \beta_l^*) &= (\alpha_l, \beta_l)X. \end{aligned}$$

However, $D(F) \neq 0$, so X maps three pairwise non-parallel vectors to three other pairwise non-parallel vectors. Such a matrix X is unique up to a scalar if it exists. But we already know that $X = U$ with $x_5 = x_6 = x_7 = -1$ is a solution, therefore the solution space of (8.2.1) is one-dimensional and hence for any solution there exists λ such that $U = \lambda X$. Let Δ_s be the determinant of the matrix obtained by removing the s-th column of A. We claim that $(\Delta_1, -\Delta_2, \ldots, \Delta_7)$ is a solution of the system of linear equations. To see this, we make an extra seventh row by copying an row and thus obtain a square matrix with determinant 0. By Laplace's formula, expanding this determinant along the seventh row, we immediately get the result. So $U = \lambda \left(\begin{smallmatrix} \Delta_1 & \Delta_3 \\ -\Delta_2 & -\Delta_4 \end{smallmatrix}\right)$. By the ultrametric inequality and again Laplace's formula, it is easy to see that

$$|\Delta_r|_\omega \leqslant \prod_{s=i,j,h} \max(|\alpha_s^*|_\omega, |\beta_s^*|_\omega) \max(|\alpha_s|_\omega, |\beta_s|_\omega), \omega \in M_L \text{ for } r = 1, 2, 3, 4.$$

Hence

$$|U|_\omega \leqslant |\lambda|_\omega \prod_{s=i,j,h} \max(|\alpha_s^*|_\omega, |\beta_s^*|_\omega) \max(|\alpha_s|_\omega, |\beta_s|_\omega) \ (\omega \in M_L).$$

Therefore, by taking the product over $\omega \in M_L$,

$$\prod_{\omega \in M_L} |U|_\omega \leqslant \prod_{s=i,j,h} H_L(\alpha_s, \beta_s) H_L(\alpha_s^*, \beta_s^*).$$

By taking the geometric means over all triples (i, j, h) and going back from L to K, we obtain that

$$\begin{aligned} \prod_{\nu \in M_K} |U|_\nu &= \Big(\prod_{\omega \in M_L} |U|_\omega \Big)^{1/[L:K]} \\ &\leqslant \Big(H_K(F) H_K(F_U) \Big)^{\binom{n-1}{2}/\binom{n}{3}} \\ &= \Big(H_K(F) H_K(F_U) \Big)^{3/n}. \end{aligned}$$

Since $U \in \mathrm{GL}\,(2, \mathcal{O}_S)$, we have $|U|_\nu = 1$ for $\nu \notin S$. Further, $F, F_U \in \mathcal{O}_S[X, Y]$. Hence

$$H_S(U) \leqslant \Big(H_S(F) H_S(F_U) \Big)^{3/n}.$$

□

Lemma 8.2.2. *Let L be a finite extension of K of degree n and T the set of valuations on L above those in S. For $x \in L$, denote by $\sigma_i, i = 1, \ldots, n$ the K-embeddings of L into its algebraic closure, with σ_1 the identity. Then for $x \in K^*$, there exists $\alpha, \beta \in \mathcal{O}_T$ such that $\frac{\alpha}{\beta} = x$ and for $F = \prod_{i=1}^{n}(\sigma_i(\alpha)X + \sigma_i(\beta)Y)$ we have*

$$e^{-\frac{2g_L}{n}} H_S(F)^{\frac{1}{n}} \leqslant H^*(x) \leqslant H_S(F)^{\frac{1}{n}}.$$

Proof. First pick $\alpha', \beta' \in L$ such that $x = \frac{\alpha'}{\beta'}$. By Lemma 3.2.3, there is $\theta \in L^*$ such that

$$\begin{aligned} &|\theta|_\omega \leqslant \min\left(\tfrac{1}{|\alpha'|_\omega}, \tfrac{1}{|\beta'|_\omega}\right) \text{ for } \omega \notin T \\ &|\theta|_\omega \leqslant A_\omega \text{ for } \omega \in T, \end{aligned}$$

where $A_\omega \in e^{\mathbb{Z}}, \omega \in T$ satisfy $\prod_{\omega \in T} A_\omega = e^{2g_L} \prod_{\omega \notin T} \max(|\alpha'|_\omega, |\beta'|_\omega)$.

Let $\alpha = \theta\alpha', \beta = \theta\beta'$. Then $\alpha, \beta \in \mathcal{O}_T$ and so $F \in \mathcal{O}_S[X, Y]$ and $x = \frac{\alpha}{\beta}$. Also, we have

$$\begin{aligned} 1 &\geqslant \prod_{\omega \notin T} \max(|\alpha|_\omega, |\beta|_\omega) \\ &= \prod_{\omega \notin T} |\theta|_\omega \prod_{\omega \notin T} \max(|\alpha'|_\omega, |\beta'|_\omega) \\ &= \frac{1}{\prod_{\omega \in T} |\theta|_\omega} \prod_{\omega \notin T} \max(|\alpha'|_\omega, |\beta'|_\omega) \\ &\geqslant \frac{1}{\prod_{\omega \in T} A_\omega} \prod_{\omega \notin T} \max(|\alpha'|_\omega, |\beta'|_\omega) \\ &= e^{-2g_L}. \end{aligned} \tag{8.2.2}$$

Let M be a normal extension of K containing L, and U the set of valuations above those in S. By Lemma 1.4.1 we have

$$\begin{aligned}
\prod_{\nu\notin S}|F|_\nu &= \Big(\prod_{\omega\notin U}|F|_\mu\Big)^{\frac{1}{[M:K]}}\\
&= \Big(\prod_{\mu\notin U}\prod_{i=1}^{n}\max(|\sigma_i(\alpha)|_\mu,|\sigma_i(\beta)|_\mu)\Big)^{\frac{1}{[M:K]}}\\
&= \Big(\prod_{\mu\notin U}\max(|\alpha|_\mu,|\beta|_\mu)\Big)^{\frac{n}{[M:K]}}\\
&= \Big(\prod_{\omega\notin T}\max(|\alpha|_\omega,|\beta|_\omega)\Big)^{\frac{n[M:L]}{[M:K]}}\\
&= \prod_{\omega\notin T}\max(|\alpha|_\omega,|\beta|_\omega). \qquad (8.2.3)
\end{aligned}$$

Combining (8.2.2) with (8.2.3) we derive that

$$e^{-2g_L}\leqslant\frac{H(F)}{H_S(F)}\leqslant 1.$$

By the product formula we have

$$\begin{aligned}
H^*(x) &= \Big(\prod_{i=1}^{n}H^*(\sigma_i(x))\Big)^{\frac{1}{n}}\\
&= \Big(\prod_{i=1}^{n}\prod_{\omega\in M_L}\max(|\sigma_i(\alpha)|_\omega,|\sigma_i(\beta)|_\omega)\Big)^{\frac{1}{n[L:K]}}\\
&= \Big(\prod_{\omega\in M_L}|F|_\omega\Big)^{\frac{1}{n[L:K]}}\\
&= H(F)^{\frac{1}{n}}.
\end{aligned}$$

This implies that

$$e^{-\frac{2g_L}{n}}H_S(F)^{\frac{1}{n}}\leqslant H^*(x)\leqslant H_S(F)^{\frac{1}{n}}.$$

□

8.3 A symmetric improvement of the Liouville-type inequality

Theorem 8.3.1. *Suppose ξ, η are algebraic over K. Let $L = K(\xi, \eta)$ and assume*

$$[K(\xi) : K] \geqslant 3,\ [K(\eta) : K] \geqslant 3,\ [L : K] = [K(\xi) : K][K(\eta) : K].$$

Let S be a finite set of valuations on K, T_0 the set of valuations on L lying above those in S and $T \subset T_0$ such that

$$\varpi := \max_{\nu \in S} \frac{1}{[L : K]} \sum_{\substack{\omega | \nu \\ \omega \in T}} [L_\omega : K_\nu] < \frac{1}{3}.$$

Let g_1, g_2 be the genera of $K(\xi)$ and $K(\eta)$ respectively. Then

$$\Delta_T(\xi, \eta) \geqslant C_5^{-1} \big(H^*(\xi) H^*(\eta)\big)^{-1+\vartheta},$$

where $\vartheta = \frac{1-3\varpi}{717(1+3\varpi)}$ and

$$C_5 = \exp\left(\frac{422(m+n-5+2g_1+2g_2)}{717} + (4m+4n+433)\frac{\#S}{717} + (m+n)(m+n-5)(1-\vartheta)\right).$$

Proof. Assume $[K(\xi) : K] = m$, $[K(\eta) : K] = n$. Then $[L : K] = mn$. Without loss of generality, suppose $\nu_\infty \in S$. For if $\nu_\infty \notin S$, then adding ν_∞ to S does not affect ϖ. Let $\sigma_1, \ldots, \sigma_m$ and $\tau_1, \ldots, \tau_n$ be the K-isomorphic embeddings of $K(\xi)$ and $K(\eta)$ respectively into M.

By Lemma 8.2.2 there are $\alpha, \beta \in K(\xi)$ and $\gamma, \delta \in K(\eta)$ that are integral over $\mathcal{O}_S$ such that $\xi = \frac{\alpha}{\beta}, \eta = \frac{\gamma}{\delta}$, and the corresponding binary forms $F(X, Y) = \prod_{i=1}^{m} \big(\sigma_i(\alpha)X + \sigma_i(\beta)Y\big), G(X, Y) = \prod_{j=1}^{n} \big(\tau_j(\gamma)X + \tau_j(\delta)Y\big)$ satisfy

$$\begin{aligned} e^{-\frac{2g_L}{m}} H_S(F)^{\frac{1}{m}} \leqslant H^*(\xi) \leqslant H_S(F)^{\frac{1}{m}}, \\ e^{-\frac{2g_L}{n}} H_S(G)^{\frac{1}{n}} \leqslant H^*(\eta) \leqslant H_S(G)^{\frac{1}{n}}. \end{aligned} \tag{8.3.1}$$

Moreover, the assumption implies that ξ, η are not conjugate over K and hence F, G are irreducible and FG is square-free. By Theorem 7.5.1, there exists $U \in \mathrm{GL}\,(2, \mathcal{O}_S)$ such that

$$|R(F,G)|_S \geqslant C' H_S(G_U)^{\frac{m}{717}} H_S(F_U)^{\frac{n}{717}}, \tag{8.3.2}$$

where

$$C' = \exp\Big(-\frac{422mn(m+n-5+2g_1+2g_2)}{717} - mn(4m+4n+433)\frac{\#S}{717}\Big).$$

Notice that

$$F_U(X,Y) = \prod_{i=1}^{m} \big(\sigma_i(\alpha')X + \sigma_i(\beta')Y\big),$$

$$G_U(X,Y) = \prod_{j=1}^{n} \big(\tau_j(\gamma')X + \tau_j(\delta')Y\big),$$

where

$$(\alpha', \beta') = (\alpha, \beta)U,\ (\gamma', \delta') = (\gamma, \delta)U.$$

Let $V \in \mathrm{GL}\,(2, \mathcal{O}_S)$ be the inverse of U. Then

$$\begin{aligned} \alpha\delta - \beta\gamma &= (\det V)(\alpha'\delta' - \beta'\gamma'), \\ \max(|\alpha|_\omega, |\beta|_\omega) &\leqslant |V|_\omega \max(|\alpha'|_\omega, |\beta'|_\omega), \\ \max(|\gamma|_\omega, |\delta|_\omega) &\leqslant |V|_\omega \max(|\gamma'|_\omega, |\delta'|_\omega). \end{aligned}$$

For $\omega \in M_L$, put

$$\Delta_\omega(\xi, \eta) := \frac{|\xi - \eta|_\omega}{\max(1, |\xi|_\omega)\max(1, |\eta|_\omega)},$$

$$\Delta'_\omega(\xi, \eta) := \frac{|\alpha'\delta' - \beta'\gamma'|_\omega}{\max(|\alpha'|_\omega, |\beta'|_\omega)\max(|\gamma'|_\omega, |\delta'|_\omega)}.$$

Then $\Delta_\omega(\xi,\eta) \leqslant 1, \Delta'_\omega(\xi,\eta) \leqslant 1$. From what we mentioned above we have

$$\begin{aligned}\Delta_\omega(\xi,\eta) &= \frac{|\alpha\delta-\beta\gamma|_\omega}{\max(|\alpha|_\omega,|\beta|_\omega)\max(|\gamma|_\omega,|\delta|_\omega)}\\ &\geqslant \frac{|\det V|_\omega|\alpha'\delta'-\beta'\gamma'|_\omega}{|V|_\omega^2\max(|\alpha'|_\omega,|\beta'|_\omega)\max(|\gamma'|_\omega,|\delta'|_\omega)}\\ &= \frac{|\det V|_\omega}{|V|_\omega^2}\Delta'_\omega(\xi,\eta)\\ &= \frac{|\det V|_\nu^{[L_\omega:K_\nu]}}{|V|_\nu^{2[L_\omega:K_\nu]}}\Delta'_\omega(\xi,\eta).\end{aligned}$$

Since $|\det V|_\nu \leqslant |V|_\nu^2$ for any $\nu \in M_K$ and $V \in \mathrm{GL}\,(2,\mathcal{O}_S)$, we derive that

$$\begin{aligned}\prod_{\omega\in T}\Delta_\omega(\xi,\eta) &\geqslant \prod_{\nu\in S}\prod_{\substack{\omega\in T\\ \omega|\nu}}\Big(\frac{|\det V|_\nu}{|V|_\nu^2}\Big)^{[L_\omega:K_\nu]}\prod_{\omega\in T}\Delta'_\omega(\xi,\eta)\\ &\geqslant \prod_{\nu\in S}\Big(\frac{|\det V|_\nu}{|V|_\nu^2}\Big)^{[L:K]\varpi}\prod_{\omega\in T}\Delta'_\omega(\xi,\eta)\\ &= \frac{1}{H_S(V)^{2[L:K]\varpi}}\prod_{\omega\in T}\Delta'_\omega(\xi,\eta).\end{aligned}$$

By Lemma 8.2.1 we have

$$\begin{aligned}H_S(V) &\leqslant \big(H_S(F_U)H_S(F_{UV})\big)^{3/m}\\ &= \big(H_S(F)H_S(F_U)\big)^{3/m},\end{aligned}$$

and

$$H_S(V) \leqslant \big(H_S(G)H_S(G_U)\big)^{3/n},$$

and from these inequalities we deduce that

$$\prod_{\omega\in T}\Delta_\omega(\xi,\eta) \geqslant \Big(\frac{1}{H_S(F)^{1/m}H_S(G)^{1/n}H_S(F_U)^{1/m}H_S(G_U)^{1/n}}\Big)^{3[L:K]\varpi}\prod_{\omega\in T}\Delta'_\omega(\xi,\eta).$$

By taking $\varepsilon = \frac{1}{717(1+3\varpi)} < 1$ and

$$H = H_S(F)^{1/m}H_S(G)^{1/n},\ H' = H_S(F_U)^{1/m}H_S(G_U)^{1/n},$$

we conclude that

$$\begin{aligned}\prod_{\omega\in T}\Delta_\omega(\xi,\eta) &\geqslant (HH')^{-3[L:K]\varepsilon\varpi}\prod_{\omega\in T}\Big(\Delta_\omega(\xi,\eta)^{1-\varepsilon}\Delta'_\omega(\xi,\eta)^{\varepsilon}\Big)\\ &\geqslant (HH')^{-3mn\varepsilon\varpi}\prod_{\omega\in T_0}\Big(\Delta_\omega(\xi,\eta)^{1-\varepsilon}\Delta'_\omega(\xi,\eta)^{\varepsilon}\Big) \quad (8.3.3)\end{aligned}$$

However, since $[L:K]=[K(\xi):K][K(\eta):K]$ we have

$$R(F,G)=\prod_{i=1}^{m}\prod_{j=1}^{n}\big(\sigma_i(\alpha)\tau_j(\delta)-\sigma_i(\beta)\tau_j(\gamma)\big)=N_{L/K}(\alpha\delta-\beta\gamma).$$

This implies that

$$|R(F,G)|_\nu=\prod_{\omega|\nu}|\alpha\delta-\beta\gamma|_\nu \text{ for } \nu\in M_K.$$

Similarly to (8.2.3), we have $H_S(F)=H_{T_0}(\alpha,\beta)^{\frac{m}{[L:K]}}$, $H_S(G)=H_{T_0}(\gamma,\delta)^{\frac{n}{[L:K]}}$. Combining this with (8.3.2) we deduce that

$$\begin{aligned}\prod_{\omega\in T_0}\Delta_\omega(\xi,\eta)&=\frac{|R(F,G)|_S}{H_{T_0}(\alpha,\beta)H_{T_0}(\gamma,\delta)}\\ &=\frac{|R(F,G)|_S}{\Big(H_S(F)^{1/m}H_S(G)^{1/n}\Big)^{[L:K]}}\\ &\geqslant \exp\Big(-\frac{422mn(m+n-5+2g_1+2g_2)}{717}-mn(4m+4n+433)\frac{\#S}{717}\Big)\times\\ &\quad\times\frac{H_S(G_U)^{\frac{m}{717}}H_S(F_U)^{\frac{n}{717}}}{\Big(H_S(F)^{1/m}H_S(G)^{1/n}\Big)^{[L:K]}}\\ &=\exp\Big(-\frac{422mn(m+n-5+2g_1+2g_2)}{717}-mn(4m+4n+433)\frac{\#S}{717}\Big)\times\\ &\quad\times\left(\frac{\Big(H_S(G_U)^{\frac{1}{n}}H_S(F_U)^{\frac{1}{m}}\Big)^{\frac{1}{717}}}{H_S(F)^{1/m}H_S(G)^{1/n}}\right)^{mn}. \quad (8.3.4)\end{aligned}$$

Similarly, we have

$$\begin{aligned}
\prod_{\omega\in T_0}\Delta'_\omega(\xi,\eta) &= \frac{|R(F_U,G_U)|_S}{H_{T_0}(\alpha',\beta')H_{T_0}(\gamma',\delta')} \\
&= \frac{|R(F,G)|_S}{\Big(H_S(F_U)^{1/m}H_S(G_U)^{1/n}\Big)^{[L:K]}} \\
&\geqslant \exp\Big(-\frac{422mn(m+n-5+2g_1+2g_2)}{717}-mn(4m+4n+433)\frac{\#S}{717}\Big)\times \\
&\quad\times\left(\frac{\Big(H_S(G_U)^{\frac{1}{n}}H_S(F_U)^{\frac{1}{m}}\Big)^{\frac{1}{717}}}{H_S(F_U)^{1/m}H_S(G_U)^{1/n}}\right)^{mn}. \qquad (8.3.5)
\end{aligned}$$

Substituting (8.3.4) and (8.3.5) into (8.3.3), we conclude that

$$\begin{aligned}
\prod_{\omega\in T}\Delta_\omega(\xi,\eta) &\geqslant (HH')^{-3mn\varepsilon\varpi}\frac{H'^{\frac{mn}{717}}}{H^{mn(1-\varepsilon)}H'^{mn\varepsilon}}\times \\
&\quad\times\exp\Big(-\frac{422mn(m+n-5+2g_1+2g_2)}{717}-mn(4m+4n+433)\frac{\#S}{717}\Big) \\
&= \exp\Big(-\frac{422mn(m+n-5+2g_1+2g_2)}{717}-mn(4m+4n+433)\frac{\#S}{717}\Big)H^{mn(-1+\vartheta)} \\
&\geqslant \exp\Big(-\frac{422mn(m+n-5+2g_1+2g_2)}{717}-mn(4m+4n+433)\frac{\#S}{717}\Big)\times \\
&\quad\times\Big(H^*(\xi)H^*(\eta)e^{2g_L(\frac{1}{m}+\frac{1}{n})}\Big)^{mn(-1+\vartheta)}.
\end{aligned}$$

where the equality is because of the choice of ε, which makes the exponent of H' to be 0, and the last inequality is due to (8.3.1). This implies that

$$\Delta_T(\xi,\eta)\geqslant D^{-1}\Big(H^*(\xi)H^*(\eta)\Big)^{-1+\vartheta},$$

where

$$D=\exp\Big(\frac{422(m+n-5+2g_1+2g_2)}{717}+(4m+4n+433)\frac{\#S}{717}+2g_L(\frac{1}{m}+\frac{1}{n})(1-\vartheta)\Big).$$

□

Notice that $\vartheta < 1$ and by (5.1.4),

$$\frac{2g_L - 2}{mn} \leqslant m + n - 6,$$

we conclude that $D \leqslant C_5$ where

$$C_5 = \exp\left(\frac{422(m+n-5+2g_1+2g_2)}{717} + (4m+4n+433)\frac{\#S}{717} + (m+n)(m+n-5)(1-\vartheta)\right).$$

Bibliography

[1] M.F. Atiyah and I.G. MacDonald. *Introduction to commutative algebra.* Westview Press, 1994.

[2] B.J. Birch and J.R. Merriman. Finiteness theorems for binary forms with given discriminant. *Proc. London Math. Soc.*, 24(3):385–394, 1972.

[3] H. F. Blichfeldt. A new principle in the geometry of numbers, with some applications. *Trans. Amer. Math. Soc.*, 15:227–235, 1914.

[4] E. Bombieri and W. Gubler. *Heights in Diophantine geometry.* Cambridge University Press, Cambridge.

[5] J. Browkin and J. Brzeziński. Some remarks on the *abc*-conjecture. *Math. Comp.*, 62(206):931–939, 1994.

[6] W.D. Brownawell and D.W. Masser. Vanishing sums in function fields. *Math. Proc. Camb. Phil. Soc.*, 100(3):427–434, 1986.

[7] Y. Bugeaud and A. Dujella. Root separation for reducible integer polynomials. *Acta Arith.*, 162:393–403, 2014.

[8] H. Davenport. On $f^3(t) - g^2(t)$. *Norske Vid. Selsk. Forh. (Trondheim)*, 38:86–87, 1965.

[9] J.-H. Evertse. Estimates for reduced binary forms. *J. Reine Angew. Math.*, 434:159–190, 1993.

[10] J.-H. Evertse. Lower bounds for resultants. II. In K. Győry, A. Pethő, and V.T. Sós, editors, *Number Theory, Diophantine, Computational and Algebraic Aspects, proc. conf. Eger 1996*, pages 181–198. Walter de Gruyter, 1998.

[11] J.-H. Evertse and K. Győry. Effective finiteness results for binary forms with given discriminant. *Compositio Math*, 79(2):169–204, 1991.

[12] J.-H. Evertse and K. Győry. Lower bounds for resultants. I. *Compositio Math*, 88(1):1–23, 1993.

[13] I. Gaál. Integral elements with given discriminant over function fields. *Acta Math. Hung.*, 52:133–146, 1988.

[14] S. Lang. *Fundamentals of Diophantine geometry*. Springer-Verlag, New York, 1983.

[15] S. Lang. *Algebraic Number Theory*, volume 110 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1994.

[16] S. Lang. *Algebra*, volume 211 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 2002.

[17] R.C. Mason. *Diophantine equations over function fields*, volume 96 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1984.

[18] J. Neukirch. *Algebraic Number Theory*. Springer-Verlag, Berlin, 1999.

[19] H. P. Schlickewei. The $\mathfrak{p}$-adic thue-siegel-roth-schmidt theorem. *Arch. Math. (Basel)*, 29:267–270, 1977.

[20] W. M. Schmidt. Norm form equations. *Ann. of Math.*, 96(2):526–551, 1972.

[21] W. M. Schmidt. Simultaneous approximation to algebraic numbers by elements of a number field. *Monatsh. Math.*, 79:55–66, 1975.

[22] J.-P. Serre. *Corps locaux*, volume VIII of *Publications de l'Universit de Nancago*. Hermann, Paris, deuxième edition, 1968.

[23] H. Stichtenoth. *Algebraic function fields and codes*, volume 254 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, second edition, 2009.

[24] J.L. Thunder. Siegel's lemma for function fields. *Michigan Math. J.*, 42(1):147–162, 1995.

[25] J.T.-Y. Wang. An effective Schmidt's subspace theorem over function fields. *Math. Z.*, 246(4):811–844, 2004.

[26] U. Zannier. Some remarks on the S-unit equation in function fields. *Acta Arith.*, 64(1):87–98, 1993.

Abstract

Let $F \in \mathbb{Z}[X,Y]$ be a *binary form*, i.e., a homogeneous polynomial in two variables. We denote the discriminant of F by $D(F)$ and its height, i.e., the maximum of the absolute values of its coefficients, by $H(F)$. Two binary forms $F, G \in \mathbb{Z}[X,Y]$ are called $\mathrm{GL}(2,\mathbb{Z})$-equivalent if $G = \pm F_U$ for some matrix $U \in \mathrm{GL}(2,\mathbb{Z})$. Here $F_U(X,Y) = F(aX + bY, cX + dY)$ for $U = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Two $\mathrm{GL}(2,\mathbb{Z})$-equivalent binary forms have the same discriminant. A binary form $F \in \mathbb{Z}[X,Y]$ is called reduced if its height cannot be made smaller by replacing it by a $\mathrm{GL}(2,\mathbb{Z})$-equivalent form. A conjecture formulated by Evertse but probably much older asserts that if $F \in \mathbb{Z}[X,Y]$ is a reduced binary form of degree $n \geqslant 2$ and non-zero discriminant, then $H(F) \leqslant c_1(n)|D(F)|^{c_2(n)}$ where $c_1(n), c_2(n)$ depend on n only. This conjecture follows from work of Lagrange (1773) and Gauss (1801) for $n = 2$ and Hermite (1851) for $n = 3$, but for $n \geqslant 4$ it is still open. The best known result towards this conjecture is due to Evertse [9] who derived a similar inequality but with c_1 depending on n and the splitting field of F. This constant c_1 cannot be computed effectively from Evertse's method of proof. Further, Evertse and Győry [11] obtained an inequality $H(F) \leqslant \exp\left(c_1(n)|D(F)|^{c_2(n)}\right)$.

In this thesis, we consider binary forms with coefficients in the polynomial ring $k[t]$, where k is an algebraically closed field of characteristic 0. If we define an absolute value $|\cdot|$ on $k[t]$ by setting $|f| := e^{\deg f}$ for $f \in k[t]$, we can formulate an analogue of Evertse's conjecture for binary forms in

$k[t][X, Y]$. In this thesis, we give a proof of this analogue. To achieve this, we first generalized Mason's ABC-theorem using work of Brownawell and Masser [6], Zannier [26] and J.T.-Y. Wang [25], then we developed an analogue over function fields of a theorem of Evertse from the geometry of numbers and subsequently a reduction theory for binary forms over function fields. As an application, we then derived results on the root separation problem over function fields, which is another interesting problem from Diophantine approximation. An elementary inequality of Mahler (1964) states that if $f \in \mathbb{Z}[X]$ is a polynomial of degree $n \geqslant 2$ of non-zero discriminant, then for any two distinct roots $\alpha, \beta \in \mathbb{C}$ of f we have $\alpha - \beta| \geqslant c(n)H(f)^{1-n}$ where $c(n) > 0$ depends on n only. The root separation problem is to prove a similar inequality with instead of $1 - n$ a larger exponent. This is still open. In this thesis, we consider the analogous problem for polynomials in $k[t][X]$, and in this setting we managed to solve the root separation problem.

This thesis is organized as follows. In Chapter 1, we introduce standard notation and collect some results needed later. In Chapter 2, we recall Mason's ABC-theorem and deduce a generalization. Then in Chapter 3, we develop an analogue of the geometry of numbers over function fields. This is applied in Chapter 4 to develop a reduction theory for binary forms over function fields. Combining the results of Chapter $1-4$, we prove in Chapter 5 a function field analogue, in fully effective form, of Evertse's conjecture mentioned above. In Chapter 6, we consider the number of equivalence classes of binary forms of given discriminant, under certain conditions. In the last two chapters, we derive an effective inequality concerning the resultant of this binary forms and derive an effective lower bound for the distance between two algebraic functions, where we make a distinction between the cases that they are conjugate over $k(t)$ or not.

Samenvatting

Een binaire vorm van graad n is een homogeen polynoom in twee variabelen van graad n. We bekijken voorlopig binaire vormen van graad n met geheeltallige coëfficiënten. Een belangrijke invariant van een binaire vorm is zijn *discriminant*. Dit is een homogeen polynoom van graad $2n-2$ in de coëfficiënten van F. We geven met $D(F)$ de discriminant van zo'n binaire vorm F aan, en met $H(F)$ de *hoogte*, dat wil zeggen het maximum van de absolute waarden van de coëfficiënten van F. Dan is $|D(F)| \leqslant c(n)H(F)^{2n-2}$ waarbij $c(n)$ alleen van n afhangt. We zeggen dat twee binaire vormen F en G equivalent zijn, als $G = \pm F_U$ voor zekere matrix $U \in \mathrm{GL}\,(2, \mathbb{Z})$. Hier is $F_U(X,Y) = F(aX+bY, cX+dY)$ voor $U = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Twee equivalante binaire vormen hebben dezelfde discriminant. We kunnen de hoogte van een binaire vorm steeds kleiner proberen te maken door hem te vervangen door een equivalente binaire vorm. Wanneer de hoogte van een binaire vorm op die manier niet meer kleiner kan worden gemaakt noemen we hem gereduceerd. Een vermoeden geformuleerd door Evertse maar waarschijnlijk al veel ouder, zegt dat van elke gereduceerde binaire vorm $F \in \mathbb{Z}[X,Y]$ van graad $n \geqslant 2$ met discriminant $\neq 0$ de hoogte $H(F)$ kan worden afgeschat als $H(F) \leqslant c_1(n)|D(F)|^{c_2(n)}$, waarbij $c_1(n)$ en $c_2(n)$ alleen van n afhangen. Dit vermoeden is voor $n = 2$ en $n = 3$ bewezen. Voor $n = 2$ volgt het uit werk van Lagrange (1773) en Gauss (1801) en voor $n = 3$ uit werk van Hermite (1851) maar voor $n \geqslant 4$ is het nog open. Evertse bewees in 1993 een zwakkere verse van bovenstaand vermoeden met in plaats van c_1 een constante die afhangt van zowel n als het splitsingslichaam van F, en die niet effectief te berekenen is uit het gegeven bewijs. Verder bewezen Evertse en Győry in 1991 een andere zwakkere versie van bovenstaand vermoeden, met een bovengrens voor $H(F)$ van de vorm $\exp\left(c_1(n)|D(F)|^{c_2(n)}\right)$, waarbij $c_1(n), c_2(n)$ effectief uit het bewijs kunnen worden berekend.

In dit proefschrift bekijken we binaire vormen met coëfficiënten in de ring $\mathbb{C}[t]$ van polynomen met complexe coëfficiënten (of meer algemeen met coëfficiënten in een algebraïsch afgesloten lichaam van karakteristiek 0). De ring $\mathbb{C}[t]$ heeft veel eigenschappen gemeen met $\mathbb{Z}$, bijvoorbeeld eenduidige priemontbinding. Verder kunnen we op $\mathbb{C}[t]$ een absolute waarde definiëren, namelijk $|f| := e^{\text{graad}\, f}$ voor $f \in \mathbb{C}[t]$. We kunnen nu een analoge versie van Evertse's vermoeden formuleren voor binaire vormen in $\mathbb{C}[t][X,Y]$. In dit proefschrift geven we een bewijs voor die analoge versie, met expliciete waarden voor $c_1(n)$ en $c_2(n)$. Om een idee van het proefschrift te geven gaan we dieper in op twee belangrijke aspecten vsn het bewijs.

Het eerste aspect betreft de meetkunde der getallen. We geven een idee van die theorie aan de hand van twee voorbeelden. Bekijk een schoolbord met daarop twee coördinaatassen getekend, de x-as en de y-as. Teken alle roosterpunten op dit bord, dat wil zeggen met $x, y \in \mathbb{Z}$, bijvoorbeeld $(0,1), (2,3), (-5,4)$, etc. Kunnen we vier roosterpunten bedekken met een stuk papier in de vorm van een cirkel met straal 1? Het is niet moeilijk te laten zien dat dit inderdaad kan. Kan dit met een driehoekig stuk papier met oppervlakte π? Of met een stuk papier van oppervlakte π van een willekeurige vorm? Blichfeldt [3] bewees in 1914 dat je met een stuk papier van oppervlakte k, na indien nodig een verschuiving, altijd $k+1$ roosterpunten kan bedekken. Bekijk nu een vierkant stuk papier met zijdelengte gelijk aan 2, maar speld het middelpunt van de vierkant vast op de oorsprong $(0,0)$, dat wil zeggen het snijpunt van de coördinaatassen. Dus we kunnen dit stuk papier wel draaien maar niet verschuiven. Ligt er altijd een ander roosterpunt dan $(0,0)$ onder het stuk papier, ongeacht hoe we het draaien? Wat als we in plaats van een vierkant stuk papier een rechthoekig stuk paper nemen met het middelpunt vastgespeld op $(0,0)$? Of een ellipsvormig stuk papier van oppervlakte 4 met middelpunt, dat wil zeggen het snijpunt van de korte as en de lange as vastgespeld op $(0,0)$? Minkowski bewees in 1896 dat een convexvormig stuk papier van oppervlakte minstens 4, dat spiegelsymmetrisch is ten opzichte van zijn zwaartepunt en waarvan het zwaartepunt op $(0,0)$ is vastgespeld, afgezien van $(0,0)$ altijd een ander roosterpunt bedekt. Dit is de zogenaamde eerste stelling van Minkowski over convexe gebieden. Deze stelling is in zekere zin kwalitatief. Later, in 1910, bewees Minkowski zijn tweede stelling over convexe gebieden. In termen van het stuk papier, kan deze als volgt worden geformuleerd. Neem weer een convexvormig stuk papier waarvan het zwaartepunt is vastgespeld op $(0,0)$ en dat spiegelsymmetrisch is ten opzichte van zijn zwaartepunt. We kunnen dit stuk met een factor λ "vermenigvuldigen" door het in alle richtingen ten opzichte van $(0,0)$ met een factor λ uit te rekken (waarbij een uitrekking met een factor $1/2$ op hetzelfde neerkomt als een inkrimping met een factor 2). Noem λ_1 de kleinste factor waarmee we het stuk papier moeten vermenigvuldigen opdat het naast $(0,0)$ nog een ander roosterpunt bedekt. Noem λ_2 de kleinste factor waarmee we het stuk papier moeten vermenigvuldigen opdat het naast $(0,0)$ nog twee andere roosterpunten bedekt die niet samen met $(0,0)$ op dezelfde lijn liggen. Dan zegt de stelling van Minkowski voor convexe gebieden dat $\frac{2}{S} \leqslant \lambda_1\lambda_2 \leqslant \frac{4}{S}$. Minkowski bewees bovengenoemde stellingen niet alleen voor het tweedimensionale geval dat we boven hebben beschreven, maar ook voor dimensies $3, 4, \ldots$. Deze resultaten blijken erg krachtig te zijn, zelfs in het onderzoek van vandaag in de Diophantische meetkunde. In hoofdstukken 3 en 4 van dit proefschrift passen we een analoge theorie van de meetkunde der getallen over $\mathbb{C}[t]$ toe en leiden daaruit een reductietheorie voor binaire vormen over $\mathbb{C}[t]$ af.

Het tweede aspect van ons bewijs heeft betrekking op het ABC-vermoeden voor algebraïsche getallen, en een analoge versie daarvan voor algebraïsche functies, die wel bewezen is. Het ABC-vermoeden gaat over drie positieve gehele getallen a, b, c met $a + b = c$ zodat a, b en c geen factor gemeenschappelijk hebben. Noem d het product van de verschillende priemdelers van abc. Het ABC-vermoeden zegt ruwweg, dat c niet te groot kan zijn ten opzichte van d. Dus wanneer a, b deelbaar zijn door hoge machten van priemgetallen, dan kan c niet deelbaar zijn door hoge machten van priemgetallen. Het ABC-vermoeden, dat geformuleerd is door Oesterlé en later op een preciezere manier door Masser in 1986, zegt het volgende:

ABC-Vermoeden. *Voor elke $\varepsilon > 0$ zijn er maar eindig veel drietallen a, b, c van positieve gehele getallen, zodat a, b, c geen factor gemeen hebben en zodat $c > d^{1+\varepsilon}$, waarbij d het product is van de priemgetallen die abc delen.*

Dit vermoeden ziet er eenvoudig uit, maar het bleek extreem moeilijk te zijn. In 1996 beschreef de Amerikaanse wiskundige Goldfeld het als "het belangrijkste onopgeloste probleem in de Diophantische analyse." Het vermoeden is nog steeds open. De Japanse wiskundige Mochizuki beweerde in 2012 een bewijs voor het ABC-vermoeden gevonden te hebben, maar experts hebben nog niet kunnen bevestigen of zijn bewijs correct is of niet. Wanneer het ABC-vermoeden juist is, heeft dit erg veel gevolgen, bijvoorbeeld allerlei generalisaties van de laatste stelling van Fermat, verscherpingen van de Stelling van Roth over hoe goed algebraïsche getallen door rationale getallen kunnen worden benaderd, en nog veel meer.

Een analoge versie van het ABC-vermoeden voor polynomen en meer algemeen algebraïsche functies is onafhankelijk van elkaar bewezen door Stothers in 1981 en Mason in 1983. Het bewijs van deze ABC-stelling voor algebraïsche functies is niet zo moeilijk. Een eenvoudige versie van deze stelling is als volgt. Zijn $a(t), b(t), c(t)$ drie polynomen met complexe coëfficiënten zodat $a(t) + b(t) = c(t)$ en zodat $a(t), b(t), c(t)$ geen gemeenschappelijk nulpunt hebben. Zij S het aantal verschillende nulpunten van

$a(t)b(t)c(t)$. Dan hebben $a(t), b(t), c(t)$ allemaal graad hoogstens $S - 1$, tenzij $a(t), b(t), c(t)$ allemaal constant zijn. In hoofdstuk 2 van dit proefschift bewijzen we onder meer een veralgemening van de ABC-stelling voor sommen $a_1(t) + \cdots + a_n(t) = c(t)$, gebaseerd op werk van Brownawell and Masser [6], Zannier [26], en J. T-Y. Wang [25], en passen dit resultaat toe in hoofdstuk 7.

Een ander probleem dat in dit proefschrift wordt bekeken is hoever nulpunten van een polynoom van elkaar af kunnen liggen. Een elementaire ongelijkheid van Mahler (1964) zegt het volgende: zij $f \in \mathbb{Z}[X]$; dan geldt voor alle nulpunten α, β van f dat $|\alpha - \beta| \geqslant c(n)H(f)^{1-n}$, waarbij $c(n)$ een getal > 0 is dat alleen van n afhangt. Hier is $H(f)$ de hoogte van f, dat wil zeggen het maximum van de absolute waarden van de coëfficiënten van f. Het probleem is om een soortgelijke ongelijkheid te bewijzen met in plaats van $1 - n$ een grotere exponent. En wat is de grootst mogelijke exponent? Hierbij spelen de bovengenoemde afschattingen voor gereduceerde binaire vormen een belangrijke rol. Voor polynomen met coëfficiënten in $\mathbb{Z}$ is dit nog open. In dit proefschrift hebben we het analoge probleem bekeken voor polynomen met coëfficiënten in $\mathbb{C}[t]$, en bewezen dat voor polynomen $f(X) \in \mathbb{C}[t][X]$ van graad $n \geqslant 4$ in X de exponent $1 - n$ inderdaad kan worden verbeterd.

Het bovenstaande probleem ligt in het verlengde van de Stelling van Roth uit 1955 die gaat over de benadering van een vast algebraïsch getal γ door rationale getallen die we vrij laten variëren. De stelling zegt dat er voor elke $\varepsilon > 0$ een getal $c(\gamma, \varepsilon) > 0$ zodat $|\gamma - p/q| > c(\gamma, \varepsilon)q^{-2-\varepsilon}$ voor alle gehele getallen p en q met $q > 0$. Voor deze stelling kreeg Roth de Fieldsmedaille.

In het *symmetrische approximatieprobleem* kijken we naar twee algebraïsche getallen α en β die we vrij laten variëren. Neem aan dat α, β nulpunten zijn van respectievelijk de polynomen $f, g \in \mathbb{Z}[X]$. We vragen naar afschattingen $|\alpha - \beta| \geqslant cH(f)^{-\delta}H(g)^{-\eta}$ met zo klein mogelijke waarden voor δ en η in termen van de hoogtes van f en g, waarbij c alleen afhangt van het getallenlichaam dat door α en β wordt voortgebracht. In dit proefschrift bewijzen we een stelling over het analoge probleem voor algebraïsche functies in plaats van algebraïsche getallen, met effectieve constanten c, δ en η.

De opzet van dit proefschrift is als volgt. In hoofdstuk 1 introduceren we de benodigde notatie, en verzamelem we enkele hulpresultaten die later worden gebruikt. In hoofdstuk 2 noemen we de ABC-stelling voor algebraïsche functies van Mason en een generalisatie daarvan van Brownawell en Masser, en leiden een verdere generalisatie af. Vervolgens leiden we in hoofdstuk 3 een analogon voor algebraïsche functies af van een stelling van Evertse in de meetkunde der getallen die een toepassing is van de tweede stelling van Minkowski voor convexe gebieden. Dit gebruiken we in hoofdstuk 4 om een reductietheorie voor binaire vormen over $\mathbb{C}[t]$ af te leiden. In hoofdstuk 5 bewijzen we het analogon van Evertse's vermoeden voor gereduceerde binaire vormen over $\mathbb{C}[t]$ door de resultaten uit de eerdere hoofdstukken te combineren. in hoofdstuk 6 kijken we naar het aantal equivalentieklassen van binair vormen over $\mathbb{C}[t]$ van gegeven discriminant. In de laatste twee hoofdstukken bekijken we de (goed gedefinieerde) afstand tussen twee algebraïsche functies α en β en leiden hiervoor een effectieve ondergrens af, eerst in het geval dat α en β geconjugeerd zijn over $\mathbb{C}(t)$, en daarna wanneer ze niet geconjugeerd zijn over $\mathbb{C}(t)$.

Acknowledgements

This dissertation is the product of my research at Leiden University. Along this way, many people have helped and encouraged me. I would like to express my gratitude to all of them.

My deepest debt of gratitude goes to Jan-Hendrik Evertse, for leading me to this fantastic mathematical subject, for his help with both my research and my life, and for his patience and kindness to me.

I also wish to thank Peter Stevenhagen for being my promotor and all his help and support since I came to Leiden five years ago.

I sincerely thank the promotion committee for their willingness and time to read my thesis, and for all their comments, suggestions and questions. Many thanks to Frits Beukers, Yann Bugeaud, Kálmán Győry, Aad van der Vaart and Julie Tzu-Yueh Wang. I thank Leiden University and NWO for supporting my PhD research.

In these years at the Mathematical Institute, I have benefited a lot from those brilliant brains of the algebra and number theory group. It has been an honor to work with them. I would like to thank Robin de Jong for his generous help since I started my master studies, as well as being a member of the promotion committee with valuable comments. I also want to express my gratitude to Bas Edixhoven for all his kind help since my first day in Leiden and for the delicious food and wine he offered us. Also, many thanks to Hendrik Lenstra, Ronald van Luijk, Bart de Smit and Lenny Taelman for encouraging, motivating and inspiring me.

There are always people coming and leaving. Many colleagues at Leiden University have been an excellent source of support. I want to express my deep gratitude to Junjiang Liu for both the inspiring academic discussions held and the cheerful memories shared. I want to express my gratitude to Michiel Kosters and Dino Festi for all their generous help. I also want to thank Maarten Kampert for his help with many practical issues. Further, I would like to thank my fellow colleagues Athanasios Angelakis, Samuele Anni, Mai Hoang Bien, Owen Biesel, Guilin Chen, Iuliana Ciocanea, Maarten Derickx, Martin Djukanovic, Krzysztof Dorobisz, Fengnan Gao, Ziyang Gao, Alberto Gioia, Albert Gunawan, David Holmes, Abtien Javanpeykar, Ariyan Javanpeykar, Jinbi Jin, Pinar Kilicer, Chloe Martindale, Djordjo Milovic, Maxim Mornev, Rachel Newton, Carlo Pagano, Rene Pannekoek, Andrea Siviero, Mima Stanojkovski, Erik Visse, Qijun Yan, Chao Zhang, Yan Zhao and Wouter Zomervrucht for for the great time we spent together.

Special thanks to Junyi Xie for the long-lasting friendship. I also thank my friends Xiaohua Ai, Xueyu Chen, Lie Fu, Jacopo Griggio, Benben Liao, Shengfa Miao, Depeng Zhao and Bo Zhou in Leiden and Paris for those good memories.

I sincerely thank Yvonne Groenewegen, Liesbeth den Hollander, Martin Lübke and Kathelijne Smits for their great help with administration and visa issues.

Last but not least, I want to express my appreciation for my beloved wife Yana Zuo. Her love, patience and support have led me this far and she is still making me a better person everyday. My gratitude also goes to our son Zhida who has brought a new dimension to my life. I thank my parents and parents-in-law for their generous help with many things.

There have been innumerable people who have helped me. I would like to express my sincere thanks to all the people and Leiden University for being part of my wonderful life in the last few years.

Curriculum Vitae

Weidong Zhuang was born on November 23, 1983. He enrolled as a student in mathematics in University of Science and Technology of China in 2005. He obtained his bachelor's degree in 2009 and was admitted in the Erasmus-Mundus ALGANT joint master program. He studied in Université Paris-Sud 11, Orsay for the first year and went to Leiden for his second year. He wrote his master thesis "Hasse-Weil Zeta-Function in a Special Case" under the supervision of Professor M. Harris. He received his master's degree in 2011 and started his PhD project, supported by NWO, under the supervision of Dr. Jan-Hendrik Evertse in Leiden University.

Weidong Zhuang married Yana Zuo in 2012. Their son Zhida was born in 2014.

Index