Author: Zhuang, Weidong
Title: Symmetric diophantine approximation over function fields
Issue Date: 2015-12-03

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 $\{\omega_1, \ldots, \omega_{d_i}\}$ where $d_i = [K_i : K] = \deg F_i$. The relative discriminant $D_i = D_{K_i/K}(\omega_1, \ldots, \omega_{d_i})$ of an $\mathcal{O}_S$-basis $\omega_1, \ldots, \omega_{d_i}$ 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 $i = 1, \ldots, d$.*

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 $i = 1, \ldots, d$. Fix $i \in \{1, \ldots, n\}$. 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 $\{1, \theta_1, \ldots, \theta_{d_i-1}\}$:

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

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

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

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 $K_1, \ldots, K_d$ be as before. For $i = 1, \ldots, d$, 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

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

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 $|P|_\nu = \max(|a_1|_\nu, \ldots, |a_n|_\nu)$, where $a_1, \ldots, a_n$ are the non-zero coefficients of P. For a finite set S of valuations containing $\{\nu_\infty\}$, $P \in K[X_1, \ldots, X_m]$, 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 $P \in \mathcal{O}_S[X_1, \ldots, X_m]\setminus\{0\}$, then $|P|_S \geqslant 1$. Clearly, $|aP|_S = |a|_S|P|_S$ for $a \in K^*, P \in K[X_1, \ldots, X_m]$. Define the inhomogeneous height of $P \in K[X_1, \ldots, X_m]$ by

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

For $P \in \mathcal{O}_S[X_1, \ldots, X_m]$, 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 $P \in L[X_1, \ldots, X_m]$, 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 $P = \frac{1}{a}(b_0X^n + b_1X^{n-1}Y + \cdots + b_nY^n) \in \mathcal{O}_S[X,Y]$, where $a, b_i \in k[t] (1 \leqslant i \leqslant n)$, $\gcd(b_0, \ldots, b_n, a) = 1$ and $|\frac{b_i}{a}|_\nu \leqslant 1$ for every $\nu \notin S$. Since $\gcd(b_0, \ldots, b_n, a) = 1$ we have in fact $|a|_\nu = 1$ for $\nu \notin S$, i.e., $a \in \mathcal{O}_S^*$. Assume that $\gcd(b_0, \ldots, b_n) = b \prod_{i=1}^{l} (t - p_i)^{h_i}$ 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*

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

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

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

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 $|a'|_S \leqslant (\prod_{\nu\in M_K\setminus S} |F|_\nu)^{-1}$ and

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

□

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 $f_i(X,1)$, $c_i \in K, l_i \in L_i[X,Y]^{\text{lin}}$. By Lemma 4.2.2, we have $f_i = c_i' N_{L_i/K}(l_i')$ with $c_i' \in K, l_i' \in \mathcal{O}_T[X,Y]^{\text{lin}}$. 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 $\sigma \in \text{Gal}(L/K)$, the restriction $\sigma|_{L_i}$ is a K-isomorphism of L_i, hence σ acts as a permutation. This completes the proof. $\square$

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 $G = \text{Gal}(L/K)$. By Lemma 4.2.3 we have a factorization $F = a\prod_{i=1}^{n} l_i$ with $l_i \in L[X,Y]^{\text{lin}}$ 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 $\det(l_i, l_j)\det(\mathbf{y}_1, \mathbf{y}_2) = \det\left(\begin{smallmatrix} l_i(\mathbf{y}_1) & l_i(\mathbf{y}_2) \\ l_j(\mathbf{y}_1) & l_j(\mathbf{y}_2) \end{smallmatrix}\right)$, 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 $B_{i\nu} = A^{1/g_\nu}_{\sigma^{-1}(i),\omega}$ 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*

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

(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*

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

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 $H^*(F^*) = \prod_{\nu \in S} |F_U|_\nu$, 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

$$W := a l_1(\mathbf{a}_2) \prod_{i=2}^{n} l_i(\mathbf{a}_1) \neq 0.$$

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

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

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. $\square$