COMPOSITIO MATHEMATICA

J. H. EVERTSE
K. GYORY

Effective finiteness results for binary forms with given discriminant

Compositio Mathematica, tome 79, n° 2 (1991), p. 169-204
<http://www.numdam.org/item?id=CM_1991__79_2_169_0>

Compositio Mathematica **79**: 169–204, 1991.

Effective finiteness results for binary forms with given discriminant

J. H. EVERTSE[1] and K. GYŐRY[2]

[1]*University of Leiden, Department of Mathematics and Computer Science, P.O. Box 9512, 2300 RA Leiden, The Netherlands and* [2]*Kossuth Lajos University, Mathematical Institute, 4010 Debrecen, Hungary*

Received 5 January 1990; accepted 15 October 1990

1. Introduction

In 1972, Birch and Merriman [2] proved that there are only finitely many equivalence classes of binary forms with integral coefficients of given degree and given discriminant. Here equivalence is defined by transformations in $\mathrm{GL}(2, \mathbf{Z})$. They extended their result to binary forms whose coefficients belong to the ring of S-integers of an algebraic number field. Birch and Merriman proved their results in an ineffective way. In this paper we give an effective proof of the result of Birch and Merriman on binary forms with S-integral coefficients. Further, we give applications of this result to binary forms, algebraic numbers of given discriminant and discriminant form equations. Our results are formulated in a quantitative form.

Each binary form $F(X, Y) = \Sigma_{i=0}^{r} a_i X^{r-i} Y^i$ factors as $\Pi_{j=1}^{r} (\alpha_j X - \beta_j Y)$ in some finite extension of $\mathbf{Q}(a_0, \dots, a_r)$. The *discriminant* of F is defined by

$$D(F) = \prod_{1 \leqslant i < j \leqslant r} (\alpha_i \beta_j - \alpha_j \beta_i)^2.$$

The discriminant has the following properties: $D(F) \in \mathbf{Z}[a_0, \dots, a_r]$; $D(\lambda F) = \lambda^{2r-2} D(F)$ for each constant λ; if $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is any 2×2-matrix, then the transformed binary form $F_A(X, Y) = F(aX + bY, cX + dY)$ has discriminant

$$D(F_A) = (\det A)^{r(r-1)} D(F).$$

Let R be an integral domain with unit group R^*. The group of 2×2-matrices with entries in R and determinant 1 is denoted by $\mathrm{SL}(2, R)$. Two binary forms F, G in $R[X, Y]$ are called *R-equivalent* if there is a matrix U in $\mathrm{SL}(2, R)$ and

[1]The research of the first author has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences (K.N.A.W.).
[2]The research of the second author has been supported in part by Grant 273 from the Hungarian National Foundation for Scientific Research.

$\varepsilon \in R^*$, such that

$$G = \varepsilon F_U.$$

From the properties of the discriminant mentioned above, it follows easily, that when F, G are R-equivalent binary forms in $R[X, Y]$, then there is an ε in R^* such that $D(G) = \varepsilon D(F)$.

In 1773, Lagrange [17] proved that there are only finitely many **Z**-equivalence classes of binary quadratic forms in $\mathbf{Z}[X, Y]$ of given discriminant. In 1851, Hermite [15] proved the same for binary cubic forms in $\mathbf{Z}[X, Y]$. The proofs of Lagrange and Hermite were *effective*, in the sense that they provided an algorithm to determine a full system of representatives for the **Z**-equivalence classes of binary forms involved.

In 1972, Birch and Merriman [2] proved that for arbitrary $r \geqslant 4$, there are only finitely many **Z**-equivalence classes of binary forms in $\mathbf{Z}[X, Y]$ of degree r and given discriminant. Their proof was *ineffective*. The main tool in the proof of Birch and Merriman was the finiteness of the number of solutions of the so-called unit equation $\alpha x + \beta y = 1$ in units x, y of the ring of integers of some given algebraic number field (implicitly proved by Siegel in 1926 [24]). Using Baker's method on linear forms in logarithms one can solve unit equations effectively (cf. [1], [11]) but this is not sufficient to make the proof of Birch and Merriman effective.

Independently of Birch and Merriman, Győry obtained some effective results on equivalence classes of *polynomials* of given discriminant. Two polynomials $f(X)$, $g(X) \in \mathbf{Z}[X]$ are called **Z**-equivalent if $g(X) = f(X + a)$ for some $a \in \mathbf{Z}$. Note that two **Z**-equivalent polynomials have the same leading coefficient. In 1973, Győry [7] proved that every monic polynomial $f(X) \in \mathbf{Z}[X]$ of degree $r \geqslant 2$ and discriminant $D \neq 0$ has degree $r \leqslant C_1(D)$ and is **Z**-equivalent to a polynomial $g(X) = X^r + g_1 X^{r-1} + \cdots + g_r$ such that $\max(|g_1|, \ldots, |g_r|) \leqslant C_2(D)$, where $C_1(D)$ and $C_2(D)$ are effectively computable numbers depending only on D. Later, Győry [8] showed that $C_1(D) = 3 + 2\log|D|/\log 3$ and $C_2(D) = \exp\exp\{4(\log|3D|)^{13}\}$ can be taken. This implies that one can effectively determine a full set of representatives for the **Z**-equivalence classes of monic polynomials in $\mathbf{Z}[X]$ of degree $\geqslant 2$ and given discriminant.

In this paper, we give an effective proof for the result of Birch and Merriman:

THEOREM 1. *Let $F(X, Y) \in \mathbf{Z}[X, Y]$ be a binary form of degree $r \geqslant 2$ and discriminant $D \neq 0$. Then F is* **Z**-*equivalent to a form $G(X, Y) = g_0 X^r + g_1 X^{r-1} Y + \cdots + g_r Y^r$ for which*

$$\max(|g_0|, \ldots, |g_r|) \leqslant \exp\{(c_1 r)^{c_2 r^4} |D|^{8r^3}\},$$

where c_1, c_2 are effectively computable, absolute constants.

By a result of Győry ([8], Theorem 1), every binary form F in $\mathbf{Z}[X, Y]$ with non-zero discriminant D has degree at most $3+2\log|D|/\log 3$. By substituting this for r into Theorem 1 we obtain:

COROLLARY 1. *Every binary form F in $\mathbf{Z}[X, Y]$ of degree $\geqslant 2$ with discriminant $D \neq 0$ is $\mathbf{Z}$-equivalent to a form $G(X, Y) = g_0X^r + \cdots + g_rY^r$ for which*

$$\max(|g_0|, \ldots, |g_r|) \leqslant \exp\exp\{c_3(\log 3|D|)^4 \log\log(3|D|)\},$$

where c_3 is an effectively computable, absolute constant.

Corollary 1 implies that there are only finitely many $\mathbf{Z}$-equivalence classes of binary forms of degree $\geqslant 2$ with discriminant $D \neq 0$, and that a full set of representatives of these classes can be effectively determined.

We mention that our results do not imply those of Győry on polynomials of given discriminant. In our proof of Theorem 1 we used an effective result on the unit equation, but apart from that, our approach is different from that of Birch and Merriman.

Theorem 1 can be applied to algebraic numbers. To every algebraic number α we can associate a binary form $F_\alpha(X, Y) \in \mathbf{Z}[X, Y]$ such that: $F_\alpha(\alpha, 1) = 0$; F_α is irreducible; $F_\alpha(1,0) > 0$; and the coefficients of F_α have gcd 1. Let $H(\alpha)$ be the maximum of the absolute values of the coefficients of F_α, and define the *discriminant* $D(\alpha)$ of α to be the discriminant $D(F_\alpha)$ of F_α. Two algebraic numbers α, β are called *equivalent* if there are $a, b, c, d \in \mathbf{Z}$ with $ad - bc = 1$ such that

$$\beta = \frac{a\alpha + b}{c\alpha + d}.$$

It is easy to check that α is equivalent to β if and only if F_α is $\mathbf{Z}$-equivalent to F_β; in that case, $D(\alpha) = D(\beta)$. Now Theorem 1 implies at once:

COROLLARY 2. *Every algebraic number α of degree $r \geqslant 2$ and discriminant D is equivalent to an algebraic number β with*

$$H(\beta) \leqslant \exp\{(c_1 r)^{c_2 r^4} |D|^{8r^3}\}.$$

In [8], Győry proved a similar result for algebraic integers α, but with a stronger notion of equivalence: two algebraic integers α, β are called strongly equivalent if $\beta = \alpha + b$ for some $b \in \mathbf{Z}$.

Let K be an algebraic number field, and S a finite set of places on K. The ring of S-integers $\mathcal{O}_S$ is the set of those α in K which are integral at every finite place outside S; $\mathcal{O}_S^*$ is the unit group of $\mathcal{O}_S$. In their same paper [2] of 1972, Birch and Merriman proved that for every $r \geqslant 3$, there are only finitely many $\mathcal{O}_S$-equivalence classes of binary forms of degree r with discriminant in $\mathcal{O}_S^*$. In 1978, Győry [9] (see also [13]) extended his result on polynomials of given discriminant to $\mathcal{O}_S$ in the following way: let $r \geqslant 3$ and $\delta \in \mathcal{O}_S \backslash \{0\}$; then for every monic polynomial $f(X) \in \mathcal{O}_S[X]$ of degree r with discriminant in

$\delta\mathcal{O}_S^* = \{\delta\varepsilon: \varepsilon \in \mathcal{O}_S^*\}$ there are $\eta \in \mathcal{O}_S^*$, $\alpha \in \mathcal{O}_S$ such that the polynomial $g(X) = \eta^{-r}.f(\eta X + \alpha)$ has height (cf. §2) $\leqslant C$, where C is an effectively computable number depending only on K, S, r and δ. Győry also gave an explicit expression for C. In this paper we shall prove that every binary form in $\mathcal{O}_S[X, Y]$ of degree $r \geqslant 2$ with discriminant in $\delta\mathcal{O}_S^*$ is $\mathcal{O}_S$-equivalent to a binary form whose height is bounded above by an effectively computable number C' depending only on K, S, r and δ (cf. Theorem 3, §2). We also give an explicit expression for C'.

We mention that in 1984, Győry [13] succeeded in generalizing some of his effective results on polynomials in $\mathcal{O}_S[X]$ of given discriminant to the case that the coefficients of the polynomials involved belong to some ring R which is finitely generated over $\mathbf{Z}$ but may contain transcendental numbers. We do not know, if our results can be generalized to that extent.

2. Results

Before we state our results, we have to introduce some terminology. Let K be an algebraic number field and $\mathcal{O}_K$ its ring of integers. Put $d = [K:\mathbf{Q}]$. Let M_K be the collection of places (equivalence classes of multiplicative valuations) on K. Each infinite place v on K contains a valuation $|\sigma(\cdot)|$, where σ is a $\mathbf{Q}$-isomorphism: $K \hookrightarrow \mathbf{C}$ and $|\cdot|$ is the ordinary absolute value on $\mathbf{C}$. We call v a *real* place if $\sigma(K) \subset \mathbf{R}$ and a *complex* place if $\sigma(K) \not\subset \mathbf{R}$. $|\sigma_1(\cdot)|$ and $|\sigma_2(\cdot)|$ belong to the same place if and only if $\sigma_2(\alpha) = \overline{\sigma_1(\alpha)}$ for all $\alpha \in K$. If the infinite place v contains $|\sigma(\cdot)|$, then put

$$|\cdot|_v = |\sigma(\cdot)|^{1/d} \quad \text{if } v \text{ is real};$$

$$|\cdot|_v = |\sigma(\cdot)|^{2/d} \quad \text{if } v \text{ is complex}. \tag{2.1}$$

There is a one-to-one correspondence between the finite places on K and the prime ideals of $\mathcal{O}_K$ and we shall identify them. For every prime ideal $\mathbf{p}$ we define the discrete valuation $|\cdot|_{\mathbf{p}}$ by

$$|\alpha|_{\mathbf{p}} = N_{K/\mathbf{Q}}(\mathbf{p})^{-\mathrm{ord}_{\mathbf{p}}(\alpha)/d} \quad \text{for } \alpha \in K^*, \ |0|_{\mathbf{p}} = 0, \tag{2.2}$$

where $N_{K/\mathbf{Q}}(\mathbf{p})$ is the norm of $\mathbf{p}$ and $\mathrm{ord}_{\mathbf{p}}(\alpha)$ the exponent of $\mathbf{p}$ in the prime ideal decomposition of the ideal generated by α. For every $\alpha \in K$ there are only finitely many places v with $|\alpha|_v \neq 1$.

Let $\mathbf{A}$ be the field of algebraic numbers. The *height* $h(\alpha)$ of $\alpha \in \mathbf{A}$ is defined as follows: take $K = \mathbf{Q}(\alpha)$; then

$$h(\alpha) = \prod_{v \in M_K} \max(1, |\alpha|_v).$$

The height $h(F)$ of a polynomial F in $\mathbf{A}[X_1, \ldots, X_n]$ is defined as the maximum of the heights of the coefficients of F.

Let S be a finite set of places on K (which by convention contains all infinite places on K). Then the ring of S-integers $\mathcal{O}_S$ and the group of S-units $\mathcal{O}_S^*$ are defined by

$$\mathcal{O}_S = \{\xi \in K : |\xi|_v \leqslant 1 \quad \text{for } v \in M_K \backslash S\},$$
$$\mathcal{O}_S^* = \{\xi \in K : |\xi|_v = 1 \quad \text{for } v \in M_K \backslash S\},$$

respectively. By an $\mathcal{O}_S$-ideal we mean a finitely generated $\mathcal{O}_S$-submodule of K and by an integral $\mathcal{O}_S$-ideal, an $\mathcal{O}_S$-ideal that is contained in $\mathcal{O}_S$. The $\mathcal{O}_S$-ideal generated by $\alpha_1, \ldots, \alpha_n$ is denoted by $(\alpha_1, \ldots, \alpha_n)_S$. If $F \in K[X_1, \ldots, X_n]$ then $(F)_S$ denotes the $\mathcal{O}_S$-ideal generated by the coefficients of F. We consequently replace the subscript S by K when S is just the set of infinite places on K.

For $\alpha \in K$ we put

$$|\alpha|_S = \prod_{v \in S} |\alpha|_v. \tag{2.3}$$

There is a unique $\mathcal{O}_K$-ideal $\mathbf{a}^*$, composed of $(\mathcal{O}_K)$-prime ideals outside S, such that $(\alpha)_S = \mathbf{a}^* \mathcal{O}_S$. From (2.1), (2.2) it follows that $|\alpha|_S = N_{K/\mathbf{Q}}(\mathbf{a}^*)^{1/d}$. More generally if $\mathbf{a}$ is an $\mathcal{O}_S$-ideal, and $\mathbf{a}^*$ is the $\mathcal{O}_K$-ideal composed of prime ideals outside S such that $\mathbf{a} = \mathbf{a}^* \mathcal{O}_S$, we put

$$|\mathbf{a}|_S = N_{K/\mathbf{Q}}(\mathbf{a}^*)^{1/d}. \tag{2.4}$$

First we state a result about *weakly $\mathcal{O}_S$-equivalent* binary forms. Two binary forms $F, G \in K[X, Y]$ are called weakly $\mathcal{O}_S$-equivalent if there is a matrix U in $\mathrm{SL}(2, \mathcal{O}_S)$ and $\lambda \in K^*$ such that

$$G = \lambda . F_U.$$

Suppose that F is a *square-free* binary form in $K[X, Y]$, that is a binary form without multiple factors. Define the so-called *S-discriminant* of F by the $\mathcal{O}_S$-ideal

$$\mathbf{d}_S(F) = \frac{(D(F))_S}{(F)_S^{2r-2}}.$$

Note that $(F)_S^{-1}$ consists of the numbers $\alpha \in K$ such that $\alpha F \in \mathcal{O}_S[X, Y]$. Further, $D(\alpha F) = \alpha^{2r-2} D(F)$. Hence $\mathbf{d}_S(F)$ is the $\mathcal{O}_S$-ideal generated by the discriminants $D(H)$ of those forms $H = \alpha F (\alpha \in K^*)$ whose coefficients belong to $\mathcal{O}_S$. Therefore, the $\mathcal{O}_S$-ideal $\mathbf{d}_S(F)$ is integral. Now assume that F, G are two weakly $\mathcal{O}_S$-equivalent binary forms; then $G = \lambda . F_U$ for some $\lambda \in K^*$, $U \in \mathrm{SL}(2, \mathcal{O}_S)$. Let $H = \alpha F (\alpha \in K^*)$ be a binary form with coefficients in $\mathcal{O}_S$. Then H_U has its coefficients in $\mathcal{O}_S$ and $H_U = (\alpha/\lambda) G$. Hence $D(H_U) \in \mathbf{d}_S(G)$. But $D(H) = D(H_U)$. Hence $D(H) \in \mathbf{d}_S(G)$. This implies that $\mathbf{d}(F) \subseteq \mathbf{d}_S(G)$. Similarly, $\mathbf{d}_S(G) \subseteq \mathbf{d}_S(F)$. We conclude that if F, G are two weakly $\mathcal{O}_S$-equivalent binary forms in $K[X, Y]$, then

$$\mathbf{d}_S(F) = \mathbf{d}_S(G).$$

In the results mentioned below, the following notation is used: $d = [K:\mathbf{Q}]$, D_K is the discriminant of K, s is the cardinality of S, t is the number of finite places in S, P is the largest of the prime numbers lying below the finite places of S, $P = 1$ if $t = 0$, and r is an integer $\geqslant 2$. Further, $c_4, c_5, \ldots, c_{11}$ are effectively computable absolute constants.

THEOREM 2. *Every square-free binary form $F \in K[X, Y]$ of degree r with $\mathbf{d}_S(F) = \mathbf{d}$ is weakly $\mathcal{O}_S$-equivalent to a form G in $\mathcal{O}_S[X, Y]$ for which*

$$h(G) \leqslant \exp\{(c_4 rs)^{c_5 dr^4(r^2t+1)^2} P^{2dr^4(r^2t+1)^2} |D_K|^{2r^4(r^4t+4)} . |\mathbf{d}|_S^{2dr^3(r^4t+4)}\}.$$

From this theorem we shall derive a similar result for $\mathcal{O}_S$-equivalent forms.

THEOREM 3. *Let $\delta \in \mathcal{O}_S \backslash \{0\}$, and let $F \in \mathcal{O}_S[X, Y]$ be a binary form of degree r with $D(F) \in \delta\mathcal{O}_S^*$. Then F is $\mathcal{O}_S$-equivalent to a form G in $\mathcal{O}_S[X, Y]$ for which*

$$h(G) \leqslant \exp\{(c_6 rs)^{c_7 dr^4(r^2t+1)^2} P^{2dr^4(r^2t+1)^2} |D_K|^{2r^4(r^4t+4)} |\delta|_S^{2dr^3(r^4t+4)}\}.$$

Theorem 1 follows at once from Theorem 3, by substituting $d = 1$, $s = 1$, $t = 0$, $P = 1$ and $D_K = 1$. Theorems 2 and 3 can be generalized to binary forms with multiple factors, provided that the definition of the discriminant is appropriately modified. We shall not work this out. We shall derive Theorem 3 from Theorem 2, but it is not difficult to show that the statements of Theorems 2 and 3 are in fact equivalent. The form G in Theorem 3 can be expressed as $\varepsilon . G_U$, where $\varepsilon \in \mathcal{O}_S^*$ and $U \in \mathrm{SL}(2, \mathcal{O}_S)$. A slight variation on Theorem 3, which is easy to derive from Theorem 3, states that for every binary form $F \in \mathcal{O}_S[X, Y]$ of degree r and with $D(F) = \delta$, there is a matrix $U \in \mathrm{SL}(2, \mathcal{O}_S)$ such that $h(F_U) \leqslant C$, where C is the same expression as the upper bound for $h(G)$ in Theorem 3, but with $|\delta|_S$ replaced by $h(\delta)$ and c_6, c_7 by other constants. This implies an (ineffective) result of Birch and Merriman ([2], Thm. 2), that up to transformations by matrices in $\mathrm{SL}(2, \mathcal{O}_K)$, there are only finitely many binary forms in $\mathcal{O}_K[X, Y]$ of given degree r and given discriminant δ.

Theorem 3 can be used to compute a representative from each $\mathcal{O}_S$-equivalence class, provided that the elements of K are representable in such a way that one can do computations in K. For this we assume that an irreducible polynomial $f(X) \in \mathbf{Z}[X]$ is given such that $K = \mathbf{Q}(\alpha)$ for some zero α of f. Then every β in K can be expressed uniquely as $(\Sigma_{i=0}^{d-1} a_i \cdot \alpha^i)/c$, with $a_0, \ldots, a_{d-1}, c \in \mathbf{Z}$, $c > 0$ and $\gcd(a_0, \ldots, a_{d-1}, c) = 1$. The tuple $(a_0, \ldots, a_{d-1}, c)$ is called a *representation* for β. By saying that certain numbers of K are given (or computable), we mean that representations for these numbers are given (or computable). Thus, if $\beta_1, \beta_2 \in K$ are given, then $\beta_1 + \beta_2$, $\beta_1 - \beta_2$, $\beta_1 \cdot \beta_2$ and β_1/β_2 (if $\beta_2 \neq 0$) are computable. For every $C \geqslant 1$ it is possible to compute a finite subset of K such that each α in K with $h(\alpha) \leqslant C$ belongs to that subset. We assume that for every prime ideal in S, a set of generators is given. Then for any given β in K it can be decided whether $\beta \in \mathcal{O}_S$.

COROLLARY 3. *Let $r \geqslant 2$ be an integer, and $\delta \in \mathcal{O}_S \backslash \{0\}$. Then there are only finitely many $\mathcal{O}_S$-equivalence classes of binary forms F in $\mathcal{O}_S[X, Y]$ of degree r with $D(F) \in \delta \mathcal{O}_S^*$, and there exists an algorithm that computes a set consisting of exactly one form from each class.*

Corollary 3 does not follow at once from Theorem 3 since some of the forms G with small height mentioned in Theorem 3 might be $\mathcal{O}_S$-equivalent. In §7 we prove Corollary 3 by showing that there exists an algorithm that can decide whether two given binary forms are $\mathcal{O}_S$-equivalent.

Every binary form $F(X, Y) \in K[X, Y]$ can be factored as

$$\lambda F_1(X, Y) \cdots F_m(X, Y),$$

where $\lambda \in K^*$ and $F_1, \dots, F_m$ are irreducible forms in $K[X, Y]$. For $j = 1, \dots, m$, let $M_j = K(\alpha_j)$ where α_j is one of the zeros of $F_j(X, 1)$, $M_j = K$ if $F_j = Y$. $(M_1, \dots, M_m)$ is called a system of fields associated to F and it is determined by F up to conjugation over K. If we restrict ourselves to binary forms associated to a given system of fields, then the bounds in Theorems 2 and 3 can be replaced by bounds depending only polynomially on $|\mathbf{d}|_S$, $|\delta|_S$, respectively. In the statements below, D_M denotes the discriminant of the extension $M/\mathbf{Q}$.

THEOREM 2′. *Let $F(X, Y) \in K[X, Y]$ be a binary form of degree $r \geqslant 2$ with $\mathbf{d}_S(F) = \mathbf{d}$, and suppose that F is associated to the system of fields $(M_1, \dots, M_m)$. Put $D = |D_{M_1} \cdots D_{M_m}|$. Then F is weakly $\mathcal{O}_S$-equivalent to a form G in $\mathcal{O}_S[X, Y]$ for which*

$$h(G) \leqslant \exp\{(c_8 rs)^{c_9 d r^4 (r^4 t + 1)} . P^{r^4 d} . D^{2r^3 (r^4 t + 3)} . (D^{2r^3} + \log |\mathbf{d}|_S)\}.$$

THEOREM 3′. *Let $F(X, Y) \in \mathcal{O}_S[X, Y]$ be a binary form of degree $r \geqslant 2$ with $D(F) \in \delta \mathcal{O}_S^*$, where $\delta \in \mathcal{O}_S^*$, and let $M_1, \dots, M_m, D$ have the same meaning as in Theorem 2. Then F is $\mathcal{O}_S$-equivalent to a form G for which*

$$h(G) \leqslant \exp\{(c_{10} rs)^{c_{11} d r^4 (r^4 t + 1)} P^{r^4 d} D^{2r^3 (r^4 t + 3)} (D^{2r^3} + \log |\delta|_S)\}.$$

In the proof of Theorem 2′, we reduce the problem of finding all weak $\mathcal{O}_S$-equivalence classes of binary forms $F \in K[X, Y]$ of degree r, associated to $(M_1, \dots, M_m)$, with $\mathbf{d}_S(F) = \mathbf{d}$, to solving a number of equations of the form $\alpha x + \beta y = 1$ in $x, y \in \mathcal{O}_T^*$, where T is the set of places on some finite extension of K lying above S. Using Baker's theory on linear forms in logarithms and its p-adic analogue, it is possible to compute an upper bound for the heights of the solutions of such equations (this was already implicitly proved in [3], but Győry ([10], Lemma 6) was the first to work this out in detail). From these effective upper bounds we derive Theorem 2′ (cf. §5). Theorem 3′ will be derived from Theorem 2′. We shall derive Theorems 2 and 3 from Theorems 2′ and 3′ by estimating D from above in terms of $r, s, t, P, |D_K|$ and $|\mathbf{d}|_S$ (cf. §6).

In our proof of Theorem 2′, we do not need the effective result on the T-unit equations mentioned above for $r=2,3$, hence in that case, a much better upper bound for $h(G)$ can be obtained which depends only polynomially on D; thus one gets upper bounds in Theorems 2 and 3 depending only polynomially on $|\mathbf{d}|_S$, $|\delta|_S$, respectively. We do not work this out. To prove Theorem 2′ for $r \geqslant 4$, it seems necessary to apply some effective result for certain T-unit equations. In fact, one gets back an effective upper bound for the heights of the solutions of $\xi+\eta=1$ in ξ, $\eta \in \mathcal{O}_S^*$ by applying Theorem 2′ to the binary forms $XY(X+Y)(\xi X-\eta Y)$ with discriminant $\{\xi\eta(\xi+\eta)\}^2 \in \mathcal{O}_S^*$. We work this out in more detail in §10. The upper bound in Theorem 2′ depends, among others, on r. It is possible to estimate r from above in terms of $\mathbf{d}=\mathbf{d}_S(F)$ and the splitting field of F (the composite of $M_1,\ldots,M_m$ and their conjugates) over K. If $\mathbf{a}$ is an integral $\mathcal{O}_S$-ideal, then it can be expressed uniquely as

$$\mathbf{a} = \mathbf{q}_1^{k_1} \cdots \mathbf{q}_u^{k_u}\mathcal{O}_S,$$

where $\mathbf{q}_1,\ldots,\mathbf{q}_u$ are distinct prime ideals outside S and $k_1,\ldots,k_u \in \mathbf{Z}_{>0}$. Put $\Omega_S(\mathbf{a}) = k_1 + \cdots + k_u$, $\omega_S(\mathbf{a}) = u$. Then we have

THEOREM 4. *Let $F(X,Y) \in K[X,Y]$ be a square-free binary form of degree r and splitting field L over K, and put $g = [L:K]$, $\mathbf{d} = \mathbf{d}_S(F)$. Then*

$$r \leqslant 3(7^{g(d+2s)} + 1) + \sqrt{3g\Omega_S(\mathbf{d})}\,\cdot$$

The proof of Theorem 4 is based on a result of the first author, [4] on the number of solutions of the equation $\alpha x + \beta y = 1$ in $x, y \in \mathcal{O}_S^*$. If $\mathbf{q}_1,\ldots,\mathbf{q}_u$ are the prime ideals outside S that divide $\mathbf{d}$, and $S' = S \cup \{\mathbf{q},\ldots,\mathbf{q}_u\}$, then $\Omega_{S'}(\mathbf{d}) = 0$. Hence, by Theorem 4 with S' instead of S, we get

$$r \leqslant 3(7^{g(d+2s+2\omega_S(\mathbf{d}))} + 1).$$

Theorem 4 is useless for irreducible F, but it is of some interest for instance when F factors into linear forms over K in which case $g = 1$. The upper bound in Theorem 4 can not be replaced by one depending only on $|\mathbf{d}|_S$. Let, for example, $K = \mathbf{Q}$, $S = \{\infty, p\}$, where ∞ is the infinite place on $\mathbf{Q}$ and p is an odd prime, $L_t = \mathbf{Q}(e^{2\pi i/p^t})$, $f_t(X)$ is the p^t-cyclotomic polynomial, and

$$F_t(X,Y) = Y^{\deg f_t}\cdot f_t(X/Y).$$

Then

$$\deg(F_t) = (p-1)p^{t-1}, \qquad D(F_t) = D_{L_t} = \pm p^{t-1}(pt-t-1)$$

(cf. [18], Chap. 4, §1, Thms. 1,3). Hence $\deg(F_t) \to \infty$ as $t \to \infty$, and $|\mathbf{d}(F_t)|_S = |D(F_t)|_S = 1$ for $t = 1,2,3\ldots$. Theorem 4 will be proved in §8.

3. Applications

In this section we give some applications of the results mentioned in §2. The results mentioned in §3 will be proved in §9. K, d, D_K, S, s, t, P will have the same meaning as in the statements of Theorems 2 and 3. Further, whenever F is a binary form associated to the system of fields $(M_1, \ldots, M_m)$, we put

$$D = |D_{M_1} \cdots D_{M_m}|.$$

$c_{12}(\ldots), c_{13}(\ldots), \ldots$ will denote positive effectively computable numbers depending only on the parameters between the parentheses; if there are no parameters, these numbers are absolute constants.

The first application concerns the following problem. Suppose that F is a binary form in $K[X, Y]$ and that

$$\mathbf{d}_S(F) = \mathbf{q}_1^{k_1} \cdots \mathbf{q}_u^{k_u} \cdot \mathcal{O}_S,$$

where $\mathbf{q}_1, \ldots, \mathbf{q}_u$ are prime ideals outside S and $k_1, \ldots, K_u \in \mathbf{Z}_{>0}$; put

$$C_S(F) = N_{K/\mathbf{Q}}(\mathbf{q}_1 \cdots \mathbf{q}_u);$$

$C_S(F)$ is called the *S-conductor* of F. Can we give an upper bound for $|\mathbf{d}_S(F)|_S$ in terms of $K, S, C_S(F)$? In general, such a bound need not exist, but such a bound does exist when F has *minimal* S-discriminant. This means that with $T = S \cup \{\mathbf{q}_1, \ldots, \mathbf{q}_u\}$, we have

$$|\mathbf{d}_S(F)|_S \leqslant |\mathbf{d}_S(G)|_S$$

for every binary form G that is weakly $\mathcal{O}_T$-equivalent to F. In the proof of the corollary below we used that

$$\omega_S(\mathbf{d}_S(F)) = u \leqslant c_{12}(K)\{\log(3C_S(F))/\log\log(3C_S(F))\}.$$

COROLLARY 4. *Let $F(X, Y) \in K[X, Y]$ be a binary form of degree $r \geqslant 2$ with minimal S-discriminant, and put $u = \omega_S(\mathbf{d}_S(F))$. Then*

(i) $\log\log|\mathbf{d}_S(F)|_S \leqslant c_{13}(K, S, r)(u+1)^2 \log(3C_S(F))$
$\qquad \leqslant c_{14}(K, S, r)\{\log(3C_S(F))\}^3/\{\log\log(3C_S(F))\}^2$

(ii) $\log\log|\mathbf{d}_S(F)|_S \leqslant c_{15}(K, S, r, D) \log(3C_S(F))$.

Corollary 4 provides some information about the arithmetical properties of minimal S-discriminants.

The second application deals with the problem to find a value of a binary form with "small" height.

COROLLARY 5. *Let $F \in \mathcal{O}_S[X, Y]$ be a square-free binary form of degree $r \geqslant 3$, and put $\mu_S(F) = \min\{h(F(x, y))\colon x, y \in \mathcal{O}_S, F(x, y) \neq 0\}$. Then*

(i) $\mu_S(F) \leqslant \exp\{(c_{16}rs)^{C_{17}dr^4(r^2t+1)^2} P^{2dr^4(r^2t+1)^2} |D_K|^{2r^4(r^4t+4)} |D(F)|_S^{2dr^3(r^4t+4)}\}$

and

(ii) $\mu_S(F) \leqslant \exp\{(c_{18}rs)^{c_{19}dr^4(r^4t+1)} P^{dr^4} D^{2r^3(r^4t+3)} (D^{2r^3} + \log|D(F)|_S)\}$.

Much better upper bounds are known for the quantity

$$\lambda(F) = \min\{|F(x, y)|\colon (X, Y) \in \mathbf{Z}^2 \backslash \{\mathbf{0}\}\}$$

which is defined for binary forms $F \in \mathbf{R}[X, Y]$ of degree $r = 2$ or 3 and with discriminant $D \neq 0$: $\lambda(F) \leqslant (-D/3)^{1/2}$ for $r = 2$, $D < 0$ [6]; $\lambda(F) \leqslant (D/5)^{1/2}$ for $r = 2$, $D > 0$ [16], [22]; $\lambda(F) \leqslant (-D/23)^{1/4}$ for $r = 3$, $D < 0$ [23]; $\lambda(F) \leqslant (D/49)^{1/4}$ for $r = 3$, $D > 0$ [23]. These bounds are best possible.

A consequence of Theorem 2 concerns equivalence of algebraic numbers. Two numbers $\theta_1, \theta_2 \in \mathbf{A}^*$ are called $\mathcal{O}_S$-equivalent if there are $\alpha, \beta, \gamma, \delta \in \mathcal{O}_S$, with $\alpha\delta - \beta\gamma = 1$, such that

$$\theta_2 = \frac{\alpha\theta_1 + \beta}{\gamma\theta_1 + \delta}.$$

To every $\theta \in \mathbf{A}^*$ of degree $r \geqslant 1$ over K we can associate the binary form

$$F_\theta(X, Y) = \prod_{i=1}^{r} (X - \theta^{(i)} Y), \tag{3.1}$$

where $\theta^{(1)} = \theta, \theta^{(2)}, \ldots, \theta^{(r)}$ are the conjugates of θ over K. It is easy to check that θ_1, θ_2 are $\mathcal{O}_S$-equivalent if and only if F_{θ_1} and F_{θ_2} are weakly $\mathcal{O}_S$-equivalent. We define the S-discriminant of $\theta \in \mathbf{A}^*$ by

$$\mathbf{d}_S(\theta) = (1)_S \quad \text{if } r = 1; \qquad \mathbf{d}_S(\theta) = \mathbf{d}_S(F_\theta) \quad \text{if } r \geqslant 2.$$

Thus, $\mathcal{O}_S$-equivalent numbers have the same S-discriminant.

COROLLARY 6. *Let $\theta \in \mathbf{A}^*$ have degree $r \geqslant 1$ over K and put $\mathbf{d}_S(\theta) = \mathbf{d}$, $|D_{K(\theta)}| = D_0$. Then θ is $\mathcal{O}_S$-equivalent to a number $\theta^* \neq 0$ for which*

(i) $h(\theta^*) \leqslant \exp\{(c_{19} rs)^{c_{20} dr^4(r^2t+1)} P^{dr^4(r^2t+1)^2} |D_K|^{2r^4(r^4t+4)} |\mathbf{d}|_S^{2dr^3(r^4t+4)}\}$

and

(ii) $h(\theta^*) \leqslant \exp\{(c_{21} rs)^{c_{22} dr^4(r^4t+1)} P^{dr^4} D_0^{2r^3(r^4t+3)} (D_0^{2r^3} + \log |\mathbf{d}|_S)\}$.

Corollaries 3 and 6 imply that there are only finitely many $\mathcal{O}_S$-equivalence classes of algebraic numbers of degree r and discriminant $\mathbf{d}$, and that a set of distinct representatives for these classes can be determined effectively. We mention that Corollary 4 has an analogue for algebraic numbers.

Two algebraic numbers θ_1, θ_2 are called *strongly* $\mathcal{O}_S$-equivalent if there are $\alpha \in \mathcal{O}_S$, $\varepsilon \in \mathcal{O}_S^*$ such that $\theta_2 = \varepsilon\theta_1 + \alpha$. Győry [9] (see also [13]) proved that every algebraic number θ of degree $r \geqslant 3$ over K that is integral over $\mathcal{O}_S$ is strongly $\mathcal{O}_S$-equivalent to a number θ^* for which $h(\theta^*) \leqslant C$, where C is an effectively computable number of a similar form as the first bound in Corollary 6.

Let M/K be a finite extension and let $\{\omega_1, \ldots, \omega_r\}$ be a K-basis of M. Then every $\alpha \in M$ can be expressed uniquely as $x_1\omega_1 + \cdots + x_r\omega_r$ with $x_1, \ldots, x_r \in K$. Put $A = \max_{1 \leqslant i \leqslant r} h(\omega_i)$. We consider the *discriminant form inequality*

$$0 < |\mathbf{d}_S(x_1\omega_1 + \cdots + x_r\omega_r)|_S \leqslant C, \quad \text{in } x_1, \ldots, x_r \in K. \tag{3.2}$$

Two solutions $\boldsymbol{x} = (x_1, \ldots, x_r)$, $\boldsymbol{y} = (y_1, \ldots, y_r)$ cf. (3.2) are called $\mathcal{O}_S$-equivalent if there are $\alpha, \beta, \gamma, \delta \in \mathcal{O}_S$ such that $\alpha\delta - \beta\gamma = 1$ and

$$\sum_{i=1}^{r} y_i\omega_i = \frac{\alpha(\sum_{i=1}^{r} x_i\omega_i) + \beta}{\gamma(\sum_{i=1}^{r} x_i\omega_i) + \delta}.$$

From Corollary 6 we shall derive

COROLLARY 7. *Every solution $\mathbf{x} = (x_1, \ldots, x_r)$ of* (3.2) *is $\mathcal{O}_S$-equivalent to a solution $\mathbf{x}^* = (x_1^*, \ldots, x_r^*)$ for which*

$$\max_i h(x_i^*) \leqslant \exp\{(c_{23}rs)^{c_{24}dr^4(r^4t+1)}p^{dr^4}.|D_M|^{2r^3(r^4t+3)}(|D_M|^{2r^3} + \log(AC))\}.$$

We mention that previously, Győry and Papp [14] considered inequalities similar to (3.2) but with unknowns $x_1, \ldots, x_r$ in $\mathcal{O}_S$. One of their results implies the following: assume that $\omega_1 = 1$, and that $\omega_2, \ldots, \omega_r$ are integral over $\mathcal{O}_S$; then for every solution $(x_1, \ldots, x_r) \in \mathcal{O}_S^r$ of (3.2) with $x_1 = 0$, there is an $\varepsilon \in \mathcal{O}_S^*$ such that $\max_{i \geqslant 2} h(\varepsilon x_i) \leqslant C_2$, where C_2 is an effectively computable number of a similar form as the upper bound in Corollary 7.

4. Auxiliary results

In this section we state and prove some auxiliary results that will be used throughout this paper. Let K be an algebraic number field and $|\cdot|_v$ ($v \in M_K$) the valuations defined by (2.1), (2.2). It is easy to check that these valuations satisfy the *product formula*

$$\prod_{v \in M_K} |\alpha|_v = 1 \quad \text{for } \alpha \in K^*.$$

Further, they satisfy the *extension formula*: if L/K is a finite extension, then

$$\prod_{w|v} |\alpha|_w = |N_{L/K}(\alpha)|_v^{1/[L:K]} \quad \text{for } \alpha \in L,\ v \in M_K,$$

where the product is taken over all places w on L lying above v. Using the extension formula one can show that if $\alpha \in \mathbf{A}$ and K is *any* finite extension of $\mathbf{Q}(\alpha)$, then

$$h(\alpha) = \prod_{v \in M_K} \max(1, |\alpha|_v).$$

We recall that the height $h(F)$ of a polynomial $F \in \mathbf{A}[X, \ldots, X_n]$ is the maximum of the heights of the coefficients of F. Similarly, we define the heights $h(\boldsymbol{a})$, $h(A)$ of a vector $\boldsymbol{a}$ and a matrix A with algebraic entries, respectively, as the maxima of the heights of the entries of $\boldsymbol{a}, A$, respectively. The maximum of the absolute values of the conjugates of an algebraic number α is denoted by $\overline{|\alpha|}$.

LEMMA 1. (i) *If* α, β, $\alpha_1, \dots, \alpha_n$ *are algebraic numbers with* $\alpha \neq 0$, *then*

$$h(\alpha^k) = h(\alpha)^{|k|} \quad \text{for } k \in \mathbf{Z}; \qquad h(\alpha\beta) \leqslant h(\alpha)h(\beta);$$

$$h(\beta/\alpha) \leqslant h(\alpha)h(\beta); \qquad h(\alpha_1 + \cdots + \alpha_n) \leqslant nh(\alpha_1)\cdots h(\alpha_n).$$

(ii) *If* $F(X_1, \dots, X_n) \in \mathbf{A}[X_1, \dots, X_n]$ *has exactly* R *non-zero coefficients and degree* d_j *in* X_j *for* $j = 1, \dots, n$, *then*

$$h(F(\alpha_1, \dots, \alpha_n)) \leqslant Rh(F)^R h(\alpha_1)^{d_1} \cdots h(\alpha_n)^{d_n} \quad \text{for } \alpha_1, \dots, \alpha_n \in \mathbf{A}.$$

(iii) *If* α, β *are conjugate algebraic numbers over* $\mathbf{Q}$, *then* $h(\alpha) = h(\beta)$.

(iv) *If* α *is a non-zero algebraic integer, then* $h(\alpha) \leqslant \overline{|\alpha|}$.

(v) *If* K *is an algebraic number field,* S *is a finite set of places on* K *and* $\alpha \in K$, *then* $|\alpha|_S \leqslant h(\alpha)$.

(vi) *If* $F(X) \in \mathbf{A}[X]$ *and* $F(\theta) = 0$, *then* $h(\theta) \leqslant \{4h(F)\}^{\deg F + 1}$.

Proof. Choose a number field K containing all numbers, coefficients of polynomials etc., appearing in the statement of the lemma.

(i) In [19], p. 51, it is proved that $h(\alpha^{-1}) = h(\alpha)$, $h(\alpha^n) = h(\alpha)^n$ for $n \in \mathbf{Z}$, $n \geqslant 0$, and $h(\alpha\beta) \leqslant h(\alpha)h(\beta)$. These results imply that $h(\alpha^n) = h(\alpha)^{|n|}$ for $n \in \mathbf{Z}$, and $h(\beta/\alpha) \leqslant h(\alpha).h(\beta)$. The last inequality of (i) is a special case of (ii).

(ii) Put $s(v) = 1/[K:\mathbf{Q}]$ if v is a real place, $s(v) = 2/[K:\mathbf{Q}]$ if v is complex and $s(v) = 0$ if v is finite. Then

$$\sum_{v \in M_K} s(v) = 1$$

and

$$|\beta_1 + \cdots + \beta_m|_v \leqslant m^{s(v)} \max(|\beta|_v, \dots, |\beta_m|_v) \quad \text{for } \beta_1, \dots, \beta_m \in K,\ v \in M_K.$$

Suppose that $F(X_1, \dots, X_n) = \Sigma_{i=1}^R a_i X_1^{k(i,1)} \cdots X_n^{k(i,n)}$, where $c_i \in K^*$ and $0 \leqslant k(i,j) \leqslant d_j$ for $i = 1, \dots, R$, $j = 1, \dots, n$. It is easy to check that

$$|a_i \alpha_1^{k(i,1)} \cdots \alpha_n^{k(i,n)}|_v$$

$$\leqslant \left(\prod_{i=1}^R \max(1, |a_i|v)\right)\left(\prod_{j=1}^n \{\max(1, |\alpha_j|_v)\}^{d_j}\right) \quad \text{for } v \in M_K.$$

Hence

$$\max(1, |F(\alpha_1, \dots, \alpha_n)|_v)$$

$$\leqslant R^{s(v)} \cdot \left(\prod_{i=1}^R \max(1, |a_i|_v)\right)\left(\prod_{j=1}^n \{\max(1, |\alpha_j|_v\}^d\right) \quad \text{for } v \in M_K.$$

Now (ii) follows, by taking the product over all v.

(iii) Suppose that α has degree d over $\mathbf{Q}$ and let

$$f(X) = a_d X^d + \cdots + a_0 \in \mathbf{Z}[X]$$

be its minimal polynomial, with $a_d > 0$, and $\gcd(a_0, \ldots, a_d) = 1$. If $\alpha_1, \ldots, \alpha_d$ are the zeros of f, then

$$h(\alpha) = \left\{ a_d \cdot \prod_{i=1}^{d} \max(1, |\alpha_i|) \right\}^{1/d}$$

(cf. [19], Remark pp. 53–54). This implies that $h(\alpha) = h(\beta)$.

(iv), (v) Straightforward consequences of definitions of valuations and height.

(vi) Let $F(X) = a_d X^d + \cdots + a_0$ and put $\tilde{h}(F) = \Pi_{v \in M_K} \max(|a_0|_v, \ldots, |a_d|_v)$. By [19], Lemma 2.2, p. 57, we have

$$h(\theta) \leqslant 4^{d+1} \tilde{h}(F).$$

Now (vi) follows from the obvious inequality $\tilde{h}(F) \leqslant h(F)^{d+1}$. □

In what follows, K will be an algebraic number field of degree d and discriminant D_K and S is a finite set of places on K. Letting $|\alpha|_S$ (for $\alpha \in K$) and $|\mathbf{a}|_S$ (for $\mathcal{O}_S$-ideals $\mathbf{a}$) be the quantities defined by (2.3) and (2.4), respectively, we have $|\alpha|_S = |(\alpha)_S|_S$. Hence $|\varepsilon|_S = 1$ for every S-unit ε. If F is a polynomial in $K[X_1, \ldots, X_n]$, then $(F)_S$ denotes the $\mathcal{O}_S$-ideal generated by the coefficients of F and we put $|F|_S := |(F)_S|_S$. Similarly, if $\boldsymbol{a} \in K^n$ is a vector, then $(\boldsymbol{a})_S$ is the $\mathcal{O}_S$-ideal generated by the coordinates of $\boldsymbol{a}$ and we put $|\underline{a}|_S := |(\underline{a})_S|_S$. We shall frequently use that for every two $\mathcal{O}_S$-ideals $\mathbf{a}, \mathbf{b}$,

$$|\mathbf{a} \cdot \mathbf{b}|_S = |\mathbf{a}|_S \cdot |\mathbf{b}|_S, \qquad |\mathbf{a}|_S \leqslant |\mathbf{b}|_S \quad \text{if } \mathbf{b} \subseteq \mathbf{a}. \tag{4.1}$$

Further, if L/K is a finite extension and T is the set of places on L lying above those in S, then

$$|\mathbf{a}\mathcal{O}_T|_T = |\mathbf{a}|_S \quad \text{for every } \mathcal{O}_S\text{-ideal } \mathbf{a}. \tag{4.2}$$

LEMMA 2 (Gauss' Lemma). *Let F, $G \in K[X_1, \ldots, X_n]$. Then*

(i) $(FG)_S = (F)_S . (G)_S$;
(ii) $|F \cdot G|_S = |F|_S \cdot |G|_S$.

Proof. Letting $c_1, \ldots, c_r$ be the non-zero coefficients of F, we put

$$|F|_v = \max(|c_1|_v, \ldots, |c_r|_v)$$

for every finite place v on K. We define $|G|_v$ similarly. Then $|F \cdot G|_v = |F|_v \cdot |G|_v$ for every finite place v (cf. [19], p. 55, Prop. 2.1). By applying this for $v \in M_K \backslash S$ we get (i). We obtain (ii) by applying (4.1) to (i). □

LEMMA 3. *Let $F(X, Y) \in K[X, Y]$ be a square-free binary form of degree r.*

Then

$$|\mathbf{d}_S(F)|_S \leqslant r^{2r-1}h(F)^{2r^2-2}.$$

Proof. It suffices to prove Lemma 3 for the case that S is the set of infinite places on K. For assume that

$$|\mathbf{d}_K(F)|_K \leqslant r^{2r-1}h(F)^{2r^2-2} \tag{4.3}$$

and let $\mathbf{d}^*$ be the $\mathcal{O}_K$-ideal, composed of prime ideals outside S, such that $\mathbf{d}_S(F) = \mathbf{d}^*\mathcal{O}_S$. Then $\mathbf{d}^*$ is an integral $\mathcal{O}_K$-ideal dividing $\mathbf{d}_K(F)$, hence $|\mathbf{d}_S(F)|_S = N_{K/\mathbf{Q}}(\mathbf{d}^*)^{1/d} \leqslant N_{K/\mathbf{Q}}(\mathbf{d}_K(F))^{1/d} = |\mathbf{d}_K(F)|_K$.

We now prove (4.3). By (4.1) and the definition of $\mathbf{d}_K(F)$ we have

$$|\mathbf{d}_K(F)|_K = |F|_K^{-(2r-2)} \cdot |D(F)|_K. \tag{4.4}$$

We shall estimate both terms in the right-hand side from above. Let

$$F(X, Y) = a_r X^r + \cdots + a_0 Y^r$$

and put

$$|F|_v = \max(|a_0|_v, \ldots, |a_r|_v) \quad \text{for } v \in M_K.$$

By the definition of $|F|_K$ and (2.2) we have

$$|F|_K = \{N_{K/\mathbf{Q}}((a_0, \ldots, a_r)_K)\}^{1/d} = \prod_{\mathbf{p}} (N_{K/\mathbf{Q}}(\mathbf{p})^{1/d})^{\min(\mathrm{ord}_{\mathbf{p}}(a_0), \ldots, \mathrm{ord}_{\mathbf{p}}(a_r))}$$
$$= \left(\prod_{v \nmid \infty} |F|_v\right)^{-1},$$

where the product is taken over all prime ideals $\mathbf{p}$ of $\mathcal{O}_K$. Hence

$$|F|_K^{-(2r-2)} \leqslant \prod_{v \nmid \infty} |F|_v^{2r-2}. \tag{4.5}$$

We now estimate $|D(F)|_K$ from above. By (2.3) we have

$$|D(F)|_K = \prod_{v | \infty} |D(F)|_v.$$

Fix an infinite place v on K and put $s(v) = 1/d$ if v is real and $s(v) = 2/d$ if v is complex. Then we may assume that $|\cdot|_v^{1/s(v)}$ is the ordinary absolute value $|\cdot|$ on $\mathbf{R}$ or $\mathbf{C}$. By using arguments of Lewis and Mahler (cf. [20], p. 335, formula (1)) we obtain

$$|D(F)| \leqslant r^{2r-1}\{\max(|a_0|, \ldots, |a_r|)\}^{2r-2},$$

hence

$$|D(F)|_v \leqslant (r^{2r-1})^{s(v)}|F|_v^{2r-2}.$$

By combining this with (4.4), (4.5), we get

$$|\mathbf{d}_K(F)|_K = \left(\prod_{v|\infty} |D(F)|_v\right).|F|_K^{-(2r-2)} \leqslant r^{2r-1}.\left(\prod_{v\in M_K} |F|_v\right)^{2r-2}.$$

It is easy to check that $|F|_v \leqslant \Pi_{i=1}^r \max(1, |a_i|_v)$ for $v \in M_K$, whence

$$\prod_{v\in M_K} |F|_v \leqslant h(a_0)\cdots h(a_r) \leqslant h(F)^{r+1}.$$

This proves Lemma 3. □

LEMMA 4. *Let* $\mathbf{a}$ *be an* $\mathcal{O}_S$*-ideal. Then there is an* $\alpha \in \mathbf{a}$ *with* $\alpha \neq 0$ *and* $|\alpha|_S \leqslant |D_K|^{1/2d}|\mathbf{a}|_S$.

Proof. If S is the set of infinite places on K, then Lemma 4 follows from the fact that every $\mathcal{O}_K$-ideal $\mathbf{a}$ contains an $\alpha \neq 0$ for which $|N_{K/\mathbf{Q}}(\alpha)| \leqslant |D_K|^{1/2}N_{K/\mathbf{Q}}(\mathbf{a})$ (cf. [18], p. 119, for a better estimate). Suppose that S contains also finite places, and let $\mathbf{a}^*$ be the $\mathcal{O}_K$-ideal, composed of prime ideals outside S, such that $\mathbf{a} = \mathbf{a}^*.\mathcal{O}_S$. Choose $\alpha \in \mathbf{a}^*$ with $\alpha \neq 0$ such that $|\alpha|_K \leqslant |D_K|^{1/2d}\cdot|\mathbf{a}^*|_K$. Then $\operatorname{ord}\mathbf{p}(\alpha) \geqslant 0$ for $\mathbf{p} \in S$, $\mathbf{p} \nmid \infty$, hence $|\alpha|_S \leqslant |\alpha|_K$. Further $|\mathbf{a}^*|_K = |\mathbf{a}|_S$. Now Lemma 4 follows by combining these inequalities. □

LEMMA 5. *Let* $\mathbf{a}$ *be an integral* $\mathcal{O}_K$*-ideal. Then* K *has a* $\mathbf{Q}$*-basis* $\{\omega_1, \ldots, \omega_d\}$ *such that* $\omega_i \in \mathbf{a}$ *and* $\overline{|\omega_i|} \leqslant |D_K|^{1/2}\cdot|\mathbf{a}|_K$ *for* $i = 1, \ldots, d$.

Proof. This is a special case of Satz 6 of [21]. □

In the following lemmas we write $\alpha \equiv \beta \bmod \mathbf{a}$, if $\alpha - \beta$ belongs to the $\mathcal{O}_S$-ideal $\mathbf{a}$, and $\alpha \equiv \beta \bmod \gamma$ if $\alpha - \beta \in (\gamma)_S$.

LEMMA 6. *Let* $\mathbf{a}$ *be an integral* $\mathcal{O}_S$*-ideal and* $\beta \in \mathcal{O}_S$. *Then there is an* $\alpha \in \mathcal{O}_K$ *such that*

$$\alpha \equiv \beta \bmod \mathbf{a}, \qquad h(\alpha) \leqslant d|D_K|^{1/2}|\mathbf{a}|_S, \qquad \alpha \neq 0.$$

Proof. We assume that S is the set of infinite places on K. According to an argument in the proof of Lemma 2 of [5], this is no restriction. Let $\mathbf{a}$ be an integral $\mathcal{O}_K$-ideal, $\beta \in \mathcal{O}_K$ and $\{\omega_1, \ldots, \omega_d\}$ a $\mathbf{Q}$-basis of K satisfying the conditions of Lemma 5. Then $\beta = \Sigma_{i=1}^d x_i\omega_i$ for certain $x_1, \ldots, x_d \in \mathbf{Q}$. Choose $y_1, \ldots, y_d \in \mathbf{Q}$ such that $0 < y_i \leqslant 1$ and $x_i - y_i \in \mathbf{Z}$ for $i = 1, \ldots, d$ and put $\alpha = \Sigma_{i=1}^d y_i\omega_i$. Then $\alpha - \beta = \Sigma_{i=1}^d (y_i - x_i)\omega_i \in \mathbf{a}$. Further, by Lemma 5 and Lemma 1, (iv),

$$h(\alpha) \leqslant \overline{|\alpha|} \leqslant \left(\sum_{i=1}^d \overline{|\omega_i|}\right) \leqslant d|D_K|^{1/2}.|\mathbf{a}|_K. \qquad \square$$

The next lemma is a more explicit version of Lemma 3 of [5]. The set of $m \times n$-matrices with entries in some set R is denoted by $R^{m,n}$. For convenience, we shall denote the matrix $\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}$ by $(\alpha, \beta; \gamma, \delta)$.

LEMMA 7. *Let $A \in \mathcal{O}_S^{2,2}$ have determinant $\Delta \neq 0$. Then there is a matrix U in* $\mathrm{SL}(2, \mathcal{O}_S)$ *such that*

$$h(UA) \leqslant 2d^5 |D_K|^{5/2} h(\Delta)^8.$$

Proof. We only indicate where the arguments in the proof of Lemma 3 of [5] have to be modified. All matrices we shall consider belong to $\mathcal{O}_S^{2,2}$. We write $(\alpha, \beta; \gamma, \delta) \sim (\alpha_1, \beta_1; \gamma_1, \delta_1)$ if there is a U in $\mathrm{SL}(2, \mathcal{O}_S)$ such that

$$(\alpha_1, \beta_1; \gamma_1, \delta_1) = U(\alpha, \beta; \gamma, \delta).$$

Let $A = (\alpha, \beta; \gamma, \delta)$. From the arguments in step 1 of the proof of Lemma 3 of [5] it follows that for every $\theta \in (\alpha, \gamma)_S$ and for every $\alpha_1 \in \mathcal{O}_S$ with $\alpha_1 \equiv \alpha \bmod \theta^2$, $\alpha_1 \neq 0$ there are $\beta^{(1)}, \gamma^{(1)}, \delta^{(1)} \in \mathcal{O}_S$ such that $(\alpha, \beta; \gamma, \delta) \sim (\alpha_1, \beta^{(1)}; \gamma^{(1)}, \delta^{(1)})$. From the arguments in step 2 of the proof of that lemma we infer that for every $\gamma_1 \in \mathcal{O}_S$ with $\gamma_1 \equiv \gamma^{(1)} \bmod \alpha_1$, there are $\beta^{(2)}, \delta^{(2)} \in \mathcal{O}_S$ such that

$$(\alpha_1, \beta^{(1)}; \gamma^{(1)}, \delta^{(1)}) \sim (\alpha_1, \beta^{(2)}; \gamma_1, \delta^{(2)}).$$

Finally, from the arguments in step 3 of the proof of Lemma 3 of [5], we conclude that for every $\beta_1 \in \mathcal{O}_S$ with $\beta_1 \equiv \beta^{(2)} \bmod \Delta\alpha_1$, there is a $\delta_1 \in \mathcal{O}_S$ such that $(\alpha_1, \beta^{(2)}; \gamma_1, \delta^{(2)}) \sim (\alpha_1, \beta_1; \gamma_1, \delta_1)$.

Note that $\Delta \in (\alpha, \gamma)_S$, and that $|\Delta|_S \leqslant h(\Delta)$ by Lemma 1, (v). By Lemma 6 we can choose $\alpha_1, \beta_1, \gamma_1$ such that

$$h(\alpha_1) \leqslant \frac{d}{2} |D_K|^{1/2} |\Delta^2|_S \leqslant \frac{d}{2} |D_K|^{1/2} h(\Delta)^2; \ \alpha_1 \neq 0;$$

$$h(\gamma_1) \leqslant d|D_K|^{1/2} |\alpha_1|_S \leqslant d|D_K|^{1/2} h(\alpha_1) \leqslant d^2 |D_K| h(\Delta)^2;$$

$$h(\beta_1) \leqslant d|D_K|^{1/2} |\Delta\alpha_1|_S \leqslant d|D_K|^{1/2} h(\Delta) h(\alpha_1) \leqslant d^2 |D_K| h(\Delta)^3.$$

Lemma 1, (i) implies that

$$h(\delta_1) \leqslant h\left(\frac{\Delta + \beta_1\gamma_1}{\alpha_1}\right) \leqslant 2h(\Delta)h(\beta_1)h(\gamma_1)h(\alpha_1) \leqslant 2d^5 |D_K|^{5/2} h(\Delta)^8.$$

This proves Lemma 7. □

We need some estimates for S-units. $c_{25}, \ldots, c_{33}$ denote effectively computable, positive absolute constants. Let h_K, R_K be the class number and regulator of K, respectively. We assume that S has cardinality s, that S contains exactly t finite places and that P is the largest prime number lying below a finite place of S with the convention that $P = 1$ if $t = 0$. We need the following estimates for h_K and R_K in terms of d and D_K.

LEMMA 8.

$$h_K R_K \leqslant c_{25}|D_K|^{1/2}(\log|D_K|)^{d-1}; \qquad R_K \geqslant c_{26};$$
$$\max(h_K, R_K) \leqslant c_{27}|D_K|^{1/2}(\log|D_K|)^{d-1}.$$

Proof. The first inequality follows from [25], Satz 1, the second from [28] Korollar, p. 375, and the third is a trivial consequence of the first two inequalities. □

LEMMA 9. *Let* $\alpha \in \mathcal{O}_S\backslash\{0\}$, $n \in \mathbf{N}$. *Then there is a* $\pi \in \mathcal{O}_S^*$ *such that*

$$\pi^n\alpha \in \mathcal{O}_K, \qquad |\pi^n\alpha|_K \leqslant P^{nth_K}\cdot|\alpha|_S.$$

Proof. Let $\mathbf{p}_1, \ldots, \mathbf{p}_t$ be the prime ideals in S. There are integers $k_1, \ldots, k_t$ and an integral $\mathcal{O}_K$-ideal $\mathbf{a}^*$ composed of prime ideals outside S, such that

$$(\alpha)_K = \mathbf{p}_1^{k_1}\cdots\mathbf{p}_t^{k_t}\mathbf{a}^*$$

where $(\alpha)_K$ is the $\mathcal{O}_K$-ideal generated by α. Define integers a_i, b_i by

$$k_i = nh_K a_i + b_i, \qquad 0 \leqslant b_i < nh_K \quad \text{for } i = 1, \ldots, d.$$

There is a $\pi \in \mathcal{O}_S^*$ such that $(\pi)_K = (\mathbf{p}_1^{-a_1}\cdots\mathbf{p}_t^{-a_t})^{k_K}$. For this π,

$$(\pi^n\alpha)_K = \mathbf{p}_1^{b_1}\cdots\mathbf{p}_t^{b_t}\cdot\mathbf{a}^* \subseteq \mathcal{O}_K.$$

Hence $\pi^n\alpha \in \mathcal{O}_K$. Further, by (2.4)

$$|\pi^n\alpha|_K = \{N_{K/\mathbf{Q}}(\mathbf{p}_1^{b_1}\cdots\mathbf{p}_t^{b_t})\cdot N_{K/\mathbf{Q}}(\mathbf{a}^*)\}^{1/d}$$
$$\leqslant N_{K/\mathbf{Q}}(\mathbf{p}_1\cdots\mathbf{p}_t)^{nh_K/d}\cdot|\alpha|_S \leqslant P^{n.th_K}\cdot|\alpha|_S. \qquad \square$$

LEMMA 10. *Put* $C(d) = (6d^3)^d$. *Then for every* $\alpha \in \mathcal{O}_S\backslash\{0\}$ *and* $n \in \mathbf{N}$, *there is an* $\varepsilon \in \mathcal{O}_S^*$ *such that*

$$h(\varepsilon^n\alpha) \leqslant |\alpha|_S\cdot\{e^{nC(d)R_K}P^{nth_K}\}$$
$$\leqslant |\alpha|_S\cdot\exp\{n.(c_{28}s)^{c_{29}d}|D_K|\cdot\log(2P)\}.$$

Proof. We first prove this for the case that S consists only of the infinite places of K. So let $\alpha \in \mathcal{O}_K\backslash\{0\}$ and put $M = |\alpha|_K = |N_{K/\mathbf{Q}}(\alpha)|^{1/d}$. Denote the conjugates of $\beta \in K$ by $\beta^{(1)} = \beta, \ldots, \beta^{(d)}$. By Lemma 3 of [11] there is an $\varepsilon \in \mathcal{O}_K^*$ such that

$$\left|\log|M^{-1}(\alpha\varepsilon^n)^{(j)}|\right| \leqslant nC(d)R_K \quad \text{for } j = 1, \ldots, d.$$

Putting $s(v) = 1/d$ if v is real, $s(v) = 2/d$ if v is complex, this implies, by (2.1),

$$|\log|\alpha\varepsilon^n|_v| \leqslant s(v)\{\log M + nC(d)R_K\}.$$

Hence

$$h(\alpha\varepsilon^n) \leqslant M\exp\{nC(d)R_K\} = |\alpha|_K\cdot\exp\{nC(d)R_K\}.$$

Now assume that $\alpha \in \mathcal{O}_S \backslash \{0\}$. By Lemma 9 there is an S-unit π such that $\pi^n \cdot \alpha \in \mathcal{O}_K$ and $|\pi^n \alpha|_K \leqslant P^{nth_K} \cdot |\alpha|_S$. By what we just proved, there is an $\eta \in \mathcal{O}_K^*$ such that

$$h(\eta^n \pi^n \alpha) \leqslant |\pi^n \alpha|_K \cdot \exp\{nC(d)R_K\}$$
$$\leqslant |\alpha|_S \cdot \exp\{nC(d)R_K\} P^{nth_K}.$$

This implies the first estimate of Lemma 10 with $\varepsilon = \eta\pi$. The second estimate follows by using the estimates for h_K and R_K in Lemma 8. □

The main tool in our proofs is an effective result of Győry ([10], Lemma 6) on the homogeneous S-unit equation in three variables. We give a slight reformulation of Győry's result.

LEMMA 11. *Let $A \geqslant 1$, and let x_0, x_1, x_2 be non-zero S-integers such that*

$$x_0 + x_1 + x_2 = 0, \qquad |x_i|_S \leqslant A \quad \text{for } i = 0, 1, 2. \tag{4.6}$$

Then, for every $\varepsilon > 0$,

$$\max_{i,j} h(x_i/x_j) \leqslant \exp\left\{\left(\frac{c_{20}s}{\varepsilon}\right)^{c_{31}d(t+1)} \cdot P^{d+\varepsilon} |D_K|^{(l+3)/2+\varepsilon} (|D_K|^{1/2} + \log A)\right\}.$$

Proof. Győry's result applies to equations of the form

$$u_0 v_0 + u_1 v_1 + u_2 v_2 = 0$$

with

$$u_i \in \mathcal{O}_K \cap \mathcal{O}_S^*, \; v_i \in \mathcal{O}_K, \; |N_{K/\mathbf{Q}}(v_i)| \leqslant N \quad \text{for } i = 0, 1, 2$$

and fixed N, and we shall transform (4.6) into such an equation. By Lemma 9, there are $\pi_0, \pi_1, \pi_2 \in \mathcal{O}_S^*$ such that $v_i := \pi_i x_i \in \mathcal{O}_K \backslash \{0\}$ and $|N_{K/\mathbf{Q}}(v_i)| \leqslant P^{tdh_K} A^d$ for $i = 0, 1, 2$. Choose $\pi \in \mathcal{O}_S^*$ such that $u_i := \pi/\pi_i \in \mathcal{O}_K \cap \mathcal{O}_S^*$ for $i = 0, 1, 2$. Then

$$\pi x_i = u_i v_i, \qquad u_i \in \mathcal{O}_K \cap \mathcal{O}_S^*, \qquad v_i \in \mathcal{O}_K,$$
$$|N_{K/\mathbf{Q}}(v_i)| \leqslant N_i := P^{tdh_K} A^d \quad \text{for } i = 0, 1, 2 \tag{4.7}$$

and

$$u_0 v_0 + u_1 v_1 + u_2 v_2 = 0. \tag{4.8}$$

By Lemma 6 of [10], there are $\sigma \in \mathcal{O}_K$ and $\rho_0, \rho_1, \rho_2 \in \mathcal{O}_K \cap \mathcal{O}_S^*$ such that

$$u_i v_i = \sigma \rho_i, \qquad \overline{|\rho_i|} \leqslant \exp\{C(N)\} \quad \text{for } i = 0, 1, 2, \tag{4.9}$$

where

$$C(N) = (c_{32}s)^{c_{33}d(t+1)} \times (3P)^d \log(2P)$$
$$\times [1 + t(R_K + h_K \log P) \log(1 + th_K R_K)]$$

$$\times R_K[(t+1)R_K+th_K \log P][R_K+th_K \log P]^t$$
$$\times [\log(R_K+1)+t \log(1+h_K R_K \log P)]^2$$
$$\times [R_K+th_K \log P+\log N].$$

By (4.7), (4.8), (4.9) and Lemma 1 this gives, for $i,j\in\{0,1,2\}$,

$$h(x_i/x_j) = h\left(\frac{u_i v_i}{u_j v_j}\right) = h\left(\frac{\sigma\rho_i}{\sigma\rho_j}\right) = h(\rho_i/\rho_j)$$
$$\leqslant h(\rho_i)h(\rho_j) \leqslant \lceil\overline{\rho_i}\rceil \cdot \lceil\overline{\rho_j}\rceil \leqslant \exp\{2C(N)\}.$$

We can estimate $C(N)$ from above by using the inequalities from Lemma 8. Further, we can remove all logarithmic factors by using the inequality

$$(\log x)^B \leqslant (B/2\varepsilon)^B x^\varepsilon \quad \text{for } B>0,\ x>0,\ \varepsilon>0.$$

Thus we obtain the upper bound for $h(x_i/x_j)$ mentioned in the statement of Lemma 11. □

We mention that results similar to Lemmas 9, 10 and 11 were proved in [12].

5. Proofs of Theorems 2′ and 3′

We use the same notation as before: K is an algebraic number field of degree d with discriminant D_K, S is a finite set of places on K of cardinality s, t is the number of finite places in S, and P is the largest of the prime numbers lying below the prime ideals in S, $P=1$ if $t=0$. It will be convenient to express a binary form $F(X, Y)$ as $F(\boldsymbol{X})$ or $\lambda\prod_{i=1}^r(\boldsymbol{a}_i, \boldsymbol{X})$, where λ is a non-zero constant, $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r$ are two-dimensional column vectors, $\boldsymbol{X}$ denotes the column vector $(X, Y)^T$, and $(.,.)$ is the scalar product of two column vectors. If $\boldsymbol{a}=(\alpha, \beta)^T$ has its coordinates in some extension L of K, and σ is a K-isomorphism of L, then we put $\sigma(\boldsymbol{a}) = (\sigma(\alpha), \sigma(\beta))^T$.

Let $F(X, Y)\in K[X, Y]$ be a square-free binary form of degree $r\geqslant 2$, associated to the system of fields $(M_1, \ldots, M_m)$. Then F can be expressed as

$$F(\boldsymbol{X}) = \lambda F_1(\boldsymbol{X})\cdots F_m(\boldsymbol{X}), \tag{5.1}$$

where $\lambda\in K^*$ and $F_j=Y$ or $F_j\in K[X, Y]$ is an irreducible binary form such that M_j contains a zero of $F_j(X, 1)$ $(j=1, \ldots, m)$. Put $r_j=\deg F_j=[M_j\colon K]$. Further, let $\sigma_{j1}, \ldots, \sigma_{j,r_j}$ be the distinct K-isomorphisms of M_j, where σ_j, is the identity. For convenience we introduce a function f, such that

$$(f(1, 1), \ldots, f(1, r_1), f(2, 1), \ldots, f(2, r_2), \ldots, f(m, 1), \ldots, f(m, r_m))$$

is a permutation of $(1, \ldots, r)$ and $f(j, 1)=j$ for $j=1, \ldots, m$. Define the fields $M_1, \ldots, M_r$ by $M_{f(j,k)} = \sigma_{jk}(M_j)$ for $j=1, \ldots, m$, $k=1, \ldots, r_j$. By changing λ if necessary, we can find vectors $\boldsymbol{a}_1, \dots, \boldsymbol{a}_r$ such that

$$F(X) = \lambda \prod_{i=1}^{r} (\boldsymbol{a}_i, X), \qquad \boldsymbol{a}_i \in M_i^2 \quad \text{for } i = 1, \dots, r$$

$$\boldsymbol{a}_{f(j,k)} = \sigma_{jk}(\boldsymbol{a}_j) \quad \text{for } j = 1, \dots, m,\ k = 1, \dots, r_j. \tag{5.2}$$

Obviously, we may assume that

$$F_j(X) = \prod_{k=1}^{r_j} (\boldsymbol{a}_{f(j,k)}, X). \tag{5.3}$$

Any tuple of vectors $(\boldsymbol{a}_1, \dots, \boldsymbol{a}_r)$ satisfying (5.2) for some $\lambda \in K^*$ is called a *proper factorization* of F. We shall prove that F has a proper factorization with certain prescribed properties. Put

$$\Delta_{ij} = \det(\boldsymbol{a}_i, \boldsymbol{a}_j), \qquad \alpha_i = \prod_{k \neq i} \Delta_{ik}, \qquad A = \prod_{i=1}^{r} \left(\prod_{k \neq i} \Delta_{ik} \right) = \alpha_1 \cdots \alpha_r,$$

$$D = \prod_{j=1}^{m} |D_{M_j}|, \qquad D_0 = |\mathbf{d}_S(F)|_S.$$

Further, let T_i be the set of places on M_i lying above the places in S. $c_{34}, \dots, c_{47}$ denote effectively computable, absolute constants.

LEMMA 12. *F has a proper factorization $(\boldsymbol{a}_1, \dots, \boldsymbol{a}_r)$ such that*

$$\boldsymbol{a}_i \in \mathcal{O}_{T_i}^2, \qquad h(\alpha_i) \leqslant D_0 \exp\{(c_{34} rs)^{c_{35} rd} D \log(2P)\} \quad \textit{for } i = 1, \dots, r,$$
$$|A|_S \leqslant D_0 D^{(r-1)/d}.$$

Proof. Let $(\boldsymbol{b}_1, \dots, \boldsymbol{b}_r)$ be any proper factorization of F. By Lemma 4, we can choose $\delta_j \in (\boldsymbol{b}_j)_{T_j}^{-1}$, $\delta_j \neq 0$ for $j = 1, \dots, m$ such that

$$|\delta_j|_{T_j} \leqslant |D_{M_j}|^{1/2r_j d} |\mathbf{b}_j|_{T_j}^{-1}.$$

Define $\delta_{m+1}, \dots, \delta_r$ by $\delta_{f(j,k)} = \sigma_{jk}(\delta_j)$ for $j = 1, \dots, m$, $k = 2, \dots, r_j$, and put $\boldsymbol{c}_i = \delta_i \boldsymbol{b}_i$ for $i = 1, \dots, r$. Then $(\boldsymbol{c}_1, \dots, \boldsymbol{c}_r)$ is a proper factorization of F such that

$$\boldsymbol{c}_i \in \mathcal{O}_{T_i}^2, \qquad |\boldsymbol{c}_i|_{T_i} \leqslant |D_{M_i}|^{1/(2d[M_i:K])} \quad \text{for } i = 1, \dots, r. \tag{5.4}$$

Put $\tilde{F}(X) = \prod_{i=1}^{r} (\boldsymbol{c}_i, X)$. Since $|(\boldsymbol{c}_i, X)|_{T_i} = |\boldsymbol{c}_i|_{T_i}$ for $i = 1, \dots, m$ we have, by Lemma 2, (ii), (4.2) and (5.4), that

$$|\tilde{F}|_S = \prod_{i=1}^{r} |\boldsymbol{c}_i|_{T_i} \leqslant \prod_{i=1}^{r} |D_{M_i}|^{1/(2d[M_i:K])} = \prod_{j=1}^{m} \left(\prod_{k=1}^{r_j} |D_{\sigma_{jk}(M_j)}|^{1/2dr_j} \right) = D^{1/2d}.$$

Together with $(D(\tilde{F}))_S = \mathbf{d}_S(\tilde{F})(\tilde{F})_S^{2r-2} = \mathbf{d}_S(F)\mathbf{c}(\tilde{F})_S^{2r-2}$ and (4.1) this implies that

$$|D(\tilde{F})|_S \leqslant D_0 D^{(r-1)/d}. \tag{5.5}$$

Put $\gamma_i = \prod_{k \neq i} \det(\boldsymbol{c}_i, \boldsymbol{c}_k)$. In the remainder of the proof, we distinguish between the cases $r = 2$ and $r > 2$.

r = 2. Note that $D(\tilde{F}) \in \mathcal{O}_S$. Hence by Lemma 10, there is an $\varepsilon \in \mathcal{O}_S^*$ such that

$$h(\varepsilon^4 D(\tilde{F})) \leqslant |D(\tilde{F})|_S \exp\{4(c_{36}s)^{c_{37}d}|D_K| \log(2P)\}.$$

Together with (5.5) and $|D_K| \leqslant D$, this implies that

$$h(\varepsilon^4 D(\tilde{F})) \leqslant D_0 \exp\{(c_{38}s)^{c_{39}d} D \log(2P)\}.$$

Put $\boldsymbol{a}_i = \varepsilon \boldsymbol{c}_1$, $\boldsymbol{a}_2 = \varepsilon \boldsymbol{c}_2$. Then $(\boldsymbol{a}_1, \boldsymbol{a}_2)$ is a proper factorization of F with $\boldsymbol{a}_i \in \mathcal{O}_{T_i}^2$ for $i = 1, 2$. Further, $\alpha_2 = -\alpha_1$ and $A = \alpha_1\alpha_2 = -\varepsilon^4 D(\tilde{F})$. Hence

$$\max(h(\alpha_1), h(\alpha_2)) \leqslant D_0^{1/2} \exp\{(c_{40}s)^{c_{41}d} D \log(2P)\}$$

and, by (5.5),

$$|A|_S = |D(\tilde{F})|_S \leqslant D_0 D^{1/d}.$$

r > 2. Let $\boldsymbol{c}_i = (\lambda_i, \mu_i)^T$ for $i = 1, \dots, r$. Then

$$\frac{\partial \tilde{F}}{\partial X}(-\mu_i, \lambda_i) = \lambda_i \gamma_i, \qquad \frac{\partial \tilde{F}}{\partial y}(-\mu_i, \lambda_i) = \mu_i \gamma_i \quad \text{for } i = 1, \dots, r. \tag{5.6}$$

Since $\frac{\partial \tilde{F}}{\partial X}, \frac{\partial \tilde{F}}{\partial Y} \in K[X, Y]$ and the coordinates of $\boldsymbol{c}_1, \dots, \boldsymbol{c}_r$ are integral over $\mathcal{O}_S$, we have $\gamma_i \in \mathcal{O}_{T_i}$ for $i = 1, \dots, r$. By Lemma 10, there are $\varepsilon_j \in \mathcal{O}_{T_j}^*$ $(j = 1, \dots, m)$ such that

$$h(\varepsilon_j^{(r-2)(2r-2)}\gamma_j) \leqslant |\gamma_j|_{T_j} \exp\{(r-2)(2r-2)(c_{41}r_j s)^{c_{42}r_j d}|D_{M_j}| \log(2P)\},$$

where $r_j = [M_j : K]$. Since $\gamma_1, \dots, \gamma_r$ are integral over $\mathcal{O}_S$ and $D(\tilde{F}) = \gamma_1 \cdots \gamma_r$ we have $\gamma_j | D(\tilde{F})$ in O_{T_j}. Hence, by (4.1), (4.2), (5.5)

$$|\gamma_j|_{T_j} \leqslant |D(\tilde{F})|_{T_j} = |D(\tilde{F})|_S \leqslant D_0 D^{(r-1)/d} \quad \text{for } j = 1, \dots, m$$

Further, $r_j \leqslant r$, $|D_{M_j}| \leqslant D$. Hence, for $j = 1, \dots, m$,

$$h(\varepsilon_j^{(r-2)(2r-2)}\gamma_j) \leqslant D_0 \exp\{(c_{44}rs)^{c_{45}rd} D \log(2P)\} \tag{5.7}$$

By (5.6) we have $\gamma_{f(j,k)} = \sigma_{jk}(\gamma_j)$ for $j = 1, \dots, m$, $k = 2, \dots, r_j$. Put $\varepsilon_{f(j,k)} = \sigma_{jk}(\varepsilon_j)$ for $j = 1, \dots, m$, $k = 2, \dots, r_j$. Thus, $\varepsilon_i \in \mathcal{O}_{T_i}^*$ for $i = 1, \dots, r$. Now Lemma 1, (iii) implies that (5.7) is also true for $j = m + 1, \dots, r$. Put

$$\boldsymbol{a}_i = \frac{\varepsilon_i^{2r-2}}{\varepsilon_i \cdots \varepsilon_r} \boldsymbol{c}_i \quad \text{for } i = 1, \dots, r.$$

Since $\varepsilon_{f(j,k)} = \sigma_{jk}(\varepsilon_j)$ for $j = 1, \dots, m$, $k = 2, \dots, r_j$, the product $\varepsilon_1 \cdots \varepsilon_r$ belongs to $\mathcal{O}_S^*$. Hence $(\boldsymbol{a}_1, \dots, \boldsymbol{a}_r)$ is a proper factorization of F, with $\boldsymbol{a}_i \in \mathcal{O}_{T_i}^2$. Further,

$$\alpha_i = \prod_{k \neq i} \det(\boldsymbol{a}_i, \boldsymbol{a}_k) = \left(\frac{\varepsilon_i^{2r-2}}{\varepsilon_1 \cdots \varepsilon_r}\right)^{r-2} \prod_{k=1}^{r} \frac{\varepsilon_k^{2r-2}}{\varepsilon_1 \cdots \varepsilon_r} \prod_{k \neq i} \det(\boldsymbol{c}_i, \boldsymbol{c}_k)$$

$$= \varepsilon_i^{(r-2)(2r-2)}\gamma_i \quad \text{for } i = 1, \dots, r.$$

Hence $h(\alpha_i)$ is bounded above by the right-hand side of (5.7), for $i=1,\dots,r$. Further, by (5.5),

$$|A|_S = |\alpha_1 \cdots \alpha_r|_S = |(\varepsilon_1 \cdots \varepsilon_r)^{(r-2)(2r-2)}\gamma_1 \cdots \gamma_r|_S = |\gamma_1 \cdots \gamma_r|_S$$
$$= |D(\tilde{F})|_S \leqslant D_0 D^{(r-1)/d}.$$

This proves Lemma 12 for $r > 2$. □

In the sequel, $(\boldsymbol{a}_1, \dots, \boldsymbol{a}_r)$ will be a proper factorization of F with the properties stated in Lemma 12. The following lemma is the part of our proof of Theorem 1 in which we apply Győry's effective result on the S-unit equation. We recall that $\Delta_{ij} := \det(\boldsymbol{a}_i, \boldsymbol{a}_j)$ for $i, j \in \{1, \dots, r\}$.

LEMMA 13. *Assume that $r \geqslant 4$. Then for each subset $\{i,j,k,l\}$ of $\{1,\dots,r\}$ we have*

$$h\left(\frac{\Delta_{ij}\Delta_{kl}}{\Delta_{ik}\Delta_{jl}}\right) \leqslant \exp\{(c_{46} rs)^{c_{47} d r^4 (r^4 t + 1)} P^{r^4 d} D^{2r^3(r^4 t + 3)}(D^{2r^3} + \log D_0)\}. \tag{5.8}$$

Proof. We use the identity $\Delta_{ij}\boldsymbol{a}_k + \Delta_{jk}\boldsymbol{a}_i + \Delta_{ki}\boldsymbol{a}_j = \boldsymbol{0}$. By taking the determinant of the left-hand side with $\boldsymbol{a}_l$, we get

$$\Delta_{ij}\Delta_{kl} + \Delta_{jk}\Delta_{il} + \Delta_{ki}\Delta_{jl} = 0. \tag{5.9}$$

Let $\tilde{M}$ be the extension of K generated by the coordinates of $\boldsymbol{a}_i$, $\boldsymbol{a}_j$, $\boldsymbol{a}_k$, $\boldsymbol{a}_l$, and let $\tilde{T}$ be the set of places on $\tilde{M}$ lying above the places in S. Since the coordinates of $\boldsymbol{a}_1, \dots, \boldsymbol{a}_r$ are integral over $\mathcal{O}_S$, each determinant Δ_{pq} is integral over $\mathcal{O}_S$. Hence the three terms in (5.9) all divide A. Together with (4.1), (4.2), Lemma 12, this implies that

$$\max(|\Delta_{ij}\Delta_{kl}|_{\tilde{T}}, |\Delta_{jk}\Delta_{il}|_{\tilde{T}}, |\Delta_{ki}\Delta_{jl}|_{\tilde{T}}) \leqslant D_0 D^{(r-1)/d}. \tag{5.10}$$

We apply Lemma 11 to (5.9). Put $f(r) = r(r-1)(r-2)(r-3)$. Then $[\tilde{M}:\mathbf{Q}] \leqslant f(r)d$, $\tilde{T}$ has cardinality at most $f(r)s$, and T contains at most $f(r)t$ finite places. By a result of Stark ([26], Lemma 7) we have

$$D_{\tilde{M}} \,\Big|\, \prod_{p=i,j,k,l} D_{M_p}^{[\tilde{M}:M_p]}.$$

Obviously, $[\tilde{M}: M_p] \leqslant (r-1)(r-2)(r-3)$, and M_p is a conjugate over K of one of the fields $M_1, \dots, M_m$ for $p = i, j, k, l$. Hence

$$|D_{\tilde{M}}| \leqslant D^{4(r-1)(r-2)(r-3)}.$$

Now Lemma 13 follows by applying Lemma 11 to (5.9), (5.10) and replacing the parameters $d, s, t, |D_K|$ in the upper bound in Lemma 10 by the upper estimates obtained above for $[\tilde{M}: \mathbf{Q}]$, the cardinality of $\tilde{T}$, the number of finite places in $\tilde{T}$ and $|D_{\tilde{M}}|$, respectively. □

We now turn to the proof of Theorem 2'. $C_1, \ldots, C_{12}$ will denote expressions of the same form as the upper bound in Theorem 2' (and hence as the right-hand side of (5.8)), but with other effectively computable numbers instead of c_8, c_9 (or c_{46}, c_{47}). We recall that the height $h(\boldsymbol{a})$ of a vector $\boldsymbol{a}$ with algebraic coordinates is the maximum of the heights of the coordinates of $\boldsymbol{a}$.

Proof of Theorem 2'. We have to prove that F is weakly $\mathcal{O}_S$-equivalent to a binary form G in $\mathcal{O}_S[X, Y]$ for which $h(G) \leqslant C_1$. To this end, we first prove that there are vectors $\boldsymbol{b}_1, \boldsymbol{b}_2 \in \mathcal{O}_S^2$, and algebraic numbers $\lambda_{i1}, \lambda_{i2}$, such that

$$\begin{aligned} &\boldsymbol{a}_i = \lambda_{i1}\boldsymbol{b}_1 + \lambda_{i2}\boldsymbol{b}_2, \quad \max(h(\lambda_{i1}), h(\lambda_{i2})) \leqslant C_2 \quad \text{for } i = 1, \ldots, r; \\ &h(\det(\boldsymbol{b}_1, \boldsymbol{b}_2)) \leqslant C_3. \end{aligned} \tag{5.11}$$

To this end, we use for $r \geqslant 3$ the identity

$$\Delta_{ij}^{(r-1)(r-2)} = \pm \frac{(\alpha_i\alpha_j)^{r-1} \Pi^* (\Delta_{kh}\Delta_{ij}/\delta_{ik}\Delta_{jh})}{\alpha_1 \cdots \alpha_r} \tag{5.12}$$

for $i, j \in \{1, \ldots, r\}$, $i \neq j$, where the product Π^* is taken over all pairs (k, h) with $k \neq h$, $k \neq i, j$ and $h \neq i, j$. By Lemma 12, we have $h(\alpha_k) \leqslant C_4$ for $k = 1, \ldots, r$, and by Lemma 13 we have $h(\Delta_{kh}\Delta_{ij}/\Delta_{ik}\Delta_{jh}) \leqslant C_5$ for all k, h. Together with (5.12) and Lemma 1, (ii) this implies that for $r \geqslant 3$,

$$h(\Delta_{ij}) \leqslant C_6 \quad \text{for } i, j \in \{1, \ldots, r\} \text{ with } i \neq j. \tag{5.13}$$

Formula (5.13) is also true for $r = 2$ since $\Delta_{12} = \alpha_1$. First suppose that $\boldsymbol{a}_1, \boldsymbol{a}_2 \in \mathcal{O}_S^2$. We have

$$\boldsymbol{a}_i = \frac{\Delta_{2i}}{\Delta_{21}} \boldsymbol{a}_1 + \frac{\Delta_{i1}}{\Delta_{21}} \boldsymbol{a}_2 \quad \text{for } i = 1, \ldots, r. \tag{5.14}$$

By (5.13) and Lemma 1, (i), we have $h(\Delta_{2i}/\Delta_{21}) \leqslant C_7$, $h(\Delta_{i1}/\Delta_{21}) \leqslant C_7$ and $h(\det(\boldsymbol{a}_1, \boldsymbol{a}_2)) = h(\Delta_{12}) \leqslant C_6$. Now (5.11) follows at once.

Now suppose that one of the vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_1$ say, does not belong to $\mathcal{O}_S^2$. Denote the conjugates of $\alpha \in M_1$ by $\alpha^{(1)} = \alpha, \ldots, \alpha^{(r_1)}$, where $r_1 = [M_1 : K]$, and consequently denote the conjugates of $\boldsymbol{a}_1$ by $\boldsymbol{a}_1^{(1)}, \ldots, \boldsymbol{a}_1^{(r)}$. By Lemma 5, M_1 has a **Q**-basis $\{\omega_1, \ldots, \omega_{r_1 d}\}$ such that $\omega_i \in \mathcal{O}_{M_1}$ and $\overline{|\omega_i|} \leqslant |D_{M_1}|^{1/2}$, for $i = 1, \ldots, r_1 d$. We may assume that $\{\omega_1, \ldots, \omega_{r_1}\}$ is a K-basis of M_1, and this basis satisfies

$$\omega_i \in \mathcal{O}_{M_1}, \qquad \overline{|\omega_i|} \leqslant D^{1/2} \quad \text{for } i = 1, \ldots, r_1. \tag{5.15}$$

There are vectors $\mathbf{d}_1, \ldots, \mathbf{d}_{r_1} \in K^2$ such that $\boldsymbol{a}_1 = \Sigma_{i=1}^{r_1} \omega_i \boldsymbol{d}_i$. Hence

$$\boldsymbol{a}_1^{(j)} = \sum_{i=1}^{r} \omega_i^{(j)} \boldsymbol{d}_i \quad \text{for } j = 1, \ldots, r_1. \tag{5.16}$$

Put $\Delta = \{\det(\omega_i^{(j)})\}^2$ and $\boldsymbol{b}_i = \Delta \boldsymbol{d}_i$ for $i = 1, \ldots, r_1$. Further, denote the matrix

$\Delta(\omega_i^{(j)})^{-1}$ by (ω^{ij}). Then $\boldsymbol{b}_k \in K^2$, each ω^{ij} is an algebraic integer and, by (5.16),

$$\boldsymbol{b}_i = \sum_{j=1}^{r_1} \omega^{ij} \boldsymbol{a}_1^{(j)} \quad \text{for } i = 1, \ldots, r_1. \tag{5.17}$$

Hence $\boldsymbol{b}_i \in \mathcal{O}_S^2$ for $i = 1, \ldots, r_1$. By (5.15) and Lemma 1, (iv), we have $h(\omega_i^{(j)}) \leqslant D^{1/2}$ for $i = 1, \ldots, r_1, j = 1, \ldots, r_1$. Each number ω^{ij} is the product of two determinants of sizes $r-1$ and r, respectively, with entries from the set of $\omega_i^{(j)}$. Together with Lemma 1, (ii) this implies that $h(\omega^{ij}) \leqslant C_8$ for all i, j. By substituting (5.14) into (5.17), using that each vector $\boldsymbol{a}_1^{(j)}$ is equal to some $\boldsymbol{a}_i$, and applying (5.13) and Lemma 1, (i), we get

$$\boldsymbol{b}_i = \kappa_{i1}\boldsymbol{a}_1 + \kappa_{i2}\boldsymbol{a}_2, \qquad \max(h(\kappa_{i1}), h(\kappa_{i2})) \leqslant c_9 \quad \text{for } i = 1, \ldots, r_1. \tag{5.18}$$

We may assume that $\{\boldsymbol{b}_1, \boldsymbol{b}_2\}$ are linearly independent. Then (5.18) and Lemma 1, (i) imply that $\boldsymbol{a}_1 = \lambda_{11}\boldsymbol{b}_1 + \lambda_{12}\boldsymbol{b}_2$, $\boldsymbol{a}_2 = \lambda_{21}\boldsymbol{b}_1 + \lambda_{22}\boldsymbol{b}_2$ for certain $\lambda_{11}, \ldots, \lambda_{22}$ with height $\leqslant C_{10}$. By substituting this into (5.14) and applying Lemma 1, (i) again, we get

$$\boldsymbol{a}_i = \lambda_{i1}\boldsymbol{b}_1 + \lambda_{i2}\boldsymbol{b}_2, \qquad \max(h(\lambda_{i1}), h(\lambda_{i2})) \leqslant C_2 \quad \text{for } i = 1, \ldots, r.$$

Further, $\Delta_{12} = \det(\boldsymbol{a}_1, \boldsymbol{a}_2) = (\lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21})\det(\boldsymbol{b}_1, \boldsymbol{b}_2)$. Together with (5.13) and Lemma 1, (i) this implies that $h(\det(\boldsymbol{b}_1, \boldsymbol{b}_2)) \leqslant C_3$. This proves (5.11).

Let B be the matrix with columns $\boldsymbol{b}_1, \boldsymbol{b}_2$. By Lemma 7 and (5.11), there is a matrix $U \in \mathrm{SL}(2, \mathcal{O}_S)$ such that $h(UB) \leqslant C_{11}$. By Lemma 12, the form

$$G(\boldsymbol{X}) = \prod_{i=1}^{r} (U\boldsymbol{a}_i, \boldsymbol{X}) = \lambda^{-1} F(U^T \boldsymbol{X})$$

belongs to $\mathcal{O}_S[X, Y]$ and is weakly $\mathcal{O}_S$-equivalent to F. Further, by (5.11),

$$U\boldsymbol{a}_i = \lambda_{i1}(U\boldsymbol{b}_1) + \lambda_{i2}(U\boldsymbol{b}_2) \quad \text{for } i = 1, \ldots, r.$$

We have $h(U\boldsymbol{b}_1) \leqslant C_{11}$, $h(U\boldsymbol{b}_2) \leqslant C_{11}$. Together with $h(\lambda_{i1})$, $h(\lambda_{i2}) \leqslant C_2$ (cf. (5.11)) and Lemma 1, (i) this implies that $h(U\boldsymbol{a}_i) \leqslant C_{12}$ for $i = 1, \ldots, r$. Each coefficient of $G(\boldsymbol{X})$ is a homogeneous polynomial of degree r in the coordinates of $U\boldsymbol{a}_1, \ldots, U\boldsymbol{a}_r$. Together with Lemma 1, (ii) this implies that $h(G) \leqslant C_1$. This proves Theorem 2′. □

Proof of Theorem 3′. Assume that $F(\boldsymbol{X})$ is a binary form of degree $r \geqslant 2$ with coefficients in $\mathcal{O}_S$, and with $D(F) \in \delta\mathcal{O}_S^*$, where $\delta \in \mathcal{O}_S \backslash \{0\}$ is fixed. We have to prove that F is $\mathcal{O}_S$-equivalent to a form G, with $h(G) \leqslant C_{13}$, where C_{13} is the upper bound of $h(G)$ in Theorem 3′. $C_{14}, \ldots, C_{17}$ will denote expressions of the same form as C_{13}, but with other effectively computable absolute constants instead of c_{10}, c_{11}. By (4.1) we have

$$1 \leqslant |\mathbf{d}(F)|_S \leqslant |D(F)|_S = |\delta|_S.$$

Together with Theorem 2′ this implies that there are a binary form

$H \in \mathcal{O}_S[X, Y]$, a matrix $U \in \mathrm{SL}(2, \mathcal{O}_S)$ and $\lambda \in K^*$ such that

$$F(X) = \lambda H_U(X), \qquad h(H) \leqslant C_{14}. \tag{5.19}$$

By (4.1) and the fact that $\mathbf{d}_S(F)$ is an integral $\mathcal{O}_S$-ideal, we have

$$1 \leqslant |F|_S = \left\{\frac{|D(F)|_S}{|\mathbf{d}_S(F)|_S}\right\}^{1/(2r-2)} \leqslant |\delta|_S^{1/(2r-2)}.$$

Together with (5.19) and the fact that $H_U \in \mathcal{O}_S[X, Y]$, this yields

$$|\lambda|_S = \frac{|F|_S}{|H_U|_S} \leqslant |F|_S \leqslant |\delta|_S^{1/(2r-2)}. \tag{5.20}$$

Let β be a non-zero coefficient of H. Then, by (5.19) and Lemma 1, (v), we have $|\beta|_S \leqslant h(\beta) \leqslant C_{14}$. Together with (5.20) this implies that $|\lambda\beta|_S \leqslant C_{15}$. Note that $\lambda\beta \in \mathcal{O}_S$. Hence by Lemma 10, there is an $\varepsilon \in \mathcal{O}_S^*$ such that $h(\varepsilon\lambda\beta) \leqslant C_{16}$. Together with $h(\beta) \leqslant C_{14}$ and Lemma 1, (i) this implies that

$$h(\varepsilon\lambda) \leqslant C_{17}. \tag{5.21}$$

Put $G(X) = \varepsilon\lambda H(X)$. Then $G_U = \varepsilon F$, hence F and G are $\mathcal{O}_S$-equivalent. Further, by (5.21), (5.19) and Lemma 1, (i), we have $h(G) \leqslant C_{13}$. This completes the proof of Theorem 3′. □

6. Proof of Theorems 2 and 3

We use the same notation as in §5, so K, d, S, s, t, P and D have the same meaning as before. We derive Theorems 2 and 3 from Theorems 2′ and 3′ by estimating D from above. We need some basic results from algebraic number theory whose proofs do not seem to be available in the literature.

Let M_0/K be any finite extension with $[M_0 : K] = r_0$, and let T_0 be a set of places in M_0 lying above the places in S. Put $\mathbf{d}_{T_0/S} = \mathbf{d}_{M_0/K}\mathcal{O}_S$, where $\mathbf{d}_{M_0/K}$ is the relative discriminant of M_0/K, $\mathbf{d}_{M_0/K} = (1)_K$ if $M_0 = K$.

LEMMA 14. $|D_{M_0}| \leqslant \{|D_K|^{1/d}(r_0 P)^t\}^{r_0 d} |\mathbf{d}_{T_0/S}|_S^d$.

Proof. Let $\mathbf{p}_1, \ldots, \mathbf{p}_t$ be the prime ideals in S. First we have

$$D_{M_0} = N_{K/\mathbb{Q}}(\mathbf{d}_{M_0/K}) D_K^{r_0}$$

and second

$$\mathbf{d}_{M_0/K} = \mathbf{p}_1^{k_1} \cdots \mathbf{p}_t^{k_t} \mathbf{a},$$

where $k_1, \ldots, k_t$ are non-negative integers and $\mathbf{a}$ is an integral $\mathcal{O}_K$-ideal composed of prime ideals outside S. By (2.4) we have $N_{K/\mathbb{Q}}(\mathbf{a}) = |\mathbf{d}_{T_0/S}|_S^d$. Hence it suffices to prove that

$$N_{K/\mathbf{Q}}(\mathfrak{p}_1^{k_1}\cdots\mathfrak{p}_t^{k_t}) \leqslant (r_0P)^{trod}.$$

Take a prime ideal $\mathfrak{p}$ from $\{\mathfrak{p}_1,\ldots,\mathfrak{p}_t\}$ and suppose that $\mathfrak{p}$ lies above the prime number p. Let $e_{\mathfrak{p}}, f_{\mathfrak{p}}$ be the ramification index and residue class degree, respectively, of $\mathfrak{p}$ over p. Further, let $\mathbf{P}_1,\ldots,\mathbf{P}_g$ be the prime ideals in M_0 lying above $\mathfrak{p}$, and denote the ramification index and residue class degree of $\mathbf{P}_i$ over $\mathfrak{p}$ by e_i and f_i, respectively. Then the exponent of $\mathfrak{p}_i$ in the prime ideal decomposition of the different $\mathbf{D}_{M_0/K}$ of M_0/K satisfies (cf. [27]. Thm. 3-7-23, p. 113)

$$\begin{aligned}\mathrm{ord}_{\mathbf{P}_i}(\mathbf{D}_{M_0/K}) &\leqslant e_i - 1 + \mathrm{ord}_{\mathbf{P}_i}(e_i) = e_i - 1 + e_{\mathfrak{p}}e_i\,\mathrm{ord}_p(e_i)\\ &\leqslant e_i(1 + e_{\mathfrak{p}}(\log e_i/\log p))\\ &= e_i(1 + e_{\mathfrak{p}}f_{\mathfrak{p}}(\log e_i/\log N_{K/\mathbf{Q}}(\mathfrak{p})))\\ &\leqslant e_i\left(1 + d\frac{\log r_0}{\log N_{K/\mathbf{Q}}(\mathfrak{p})}\right).\end{aligned}$$

Since $\mathbf{d}_{M_0/K} = N_{M_0/K}(\mathbf{D}_{M_0/K})$ this implies

$$\mathrm{ord}_{\mathfrak{p}}(\mathbf{d}_{M_0/K}) \leqslant \left(\sum_{i=1}^{g} e_if_i\right)\left(1 + d\frac{\log r_0}{\log N_{K/\mathbf{Q}}(\mathfrak{p})}\right) = r_0\left(1 + d\frac{\log r_0}{\log N_{K/\mathbf{Q}}(\mathfrak{p})}\right).$$

Hence

$$N_{K/\mathbf{Q}}(\mathfrak{p}_1^{k_1}\cdots\mathfrak{p}_t^{k_t}) \leqslant N_{K/\mathbf{Q}}(\mathfrak{p}_1\cdots\mathfrak{p}_t)^{r_0}r_0^{trod} \leqslant (r_0P)^{trod}. \qquad \square$$

Let $F(X, Y)\in K[X, Y]$ be a square-free binary form of degree r and let $M_1,\ldots, M_m$, $T_1,\ldots, T_m$ have the same meaning as in the beginning of §5.

LEMMA 15. $\mathbf{d}_S(F) \subseteq \mathbf{d}_{T_1/S}\cdots\mathbf{d}_{T_m/S}$.

Proof. We use the notation of §5. Let $(\boldsymbol{a}_1,\ldots,\boldsymbol{a}_r)$ be a proper factorization of F and define $\lambda\in K^*$ by $F(\boldsymbol{X}) = \lambda\prod_{j=1}^{r}(\boldsymbol{a}_j, \boldsymbol{X})$. Further, let C_j be the set of subscripts $\{f(j, k)\colon k = 1,\ldots, r_j\}$, so that the sets $C_1,\ldots, C_m$ are pairwise disjoint. Then the forms

$$F_j(\boldsymbol{X}) = \prod_{k\in C_j}(\boldsymbol{a}_k, \boldsymbol{X})\ (j = 1,\ldots, m)$$

are irreducible in $K[X, Y]$, and one of the vectors $\boldsymbol{a}_k (k\in C_j)$ belongs to M_j^2. Further, $F = \lambda F_1\cdots F_m$. Let L/K be the extension generated by the coordinates of $\boldsymbol{a}_1,\ldots,\boldsymbol{a}_r$ and T the set of places on L lying above those in S. Then Lemma 2, (i) implies that $(F)_T = (\lambda)_T(\boldsymbol{a}_1)_T\cdots(\boldsymbol{a}_r)_T$. Define the $\mathcal{O}_T$-ideals

$$\mathbf{d}_{ij} = \frac{(\det(\boldsymbol{a}_i, \boldsymbol{a}_j))_T}{(\boldsymbol{a}_i)_T(\boldsymbol{a}_j)_T}\ (1 \leqslant i, j \leqslant r,\ i \neq j).$$

Then

$$\mathbf{d}_S(F)\mathcal{O}_T = \frac{(D(F))_T}{(F)_T^{2r-2}} = \prod_{p \neq q} \mathbf{d}_{pq}. \tag{6.1}$$

Similarly,

$$\mathbf{d}_S(F_j)\mathcal{O}_T = \prod_{\substack{p \neq q \\ p,q \in C_j}} \mathbf{d}_{pq} \quad \text{if deg } F_j \geqslant 2 \ (j = 1, \ldots, m).$$

Each $\mathcal{O}_T$-ideal $\mathbf{d}_{ij}$ is integral. Hence

$$\mathbf{d}_S(F) \subseteq \mathbf{d}_S(F_1) \cdots \mathbf{d}_S(F_m),$$

where we put $\mathbf{d}_S(F_j) = (1)_S$ if F_j is linear. Therefore, it suffices to show that

$$\mathbf{d}_S(F_j) \subseteq \mathbf{d}_{T_j/S} \quad \text{for } j = 1, \ldots, m. \tag{6.2}$$

Fix a subscript j for which $[M_j : K] = 2$ (if $M_j = K$, then (6.2) is trivial). Write F', M', T', r' instead of F_j, M_j, T_j, r_j. In §2 we remarked that $\mathbf{d}_S(F')$ is the $\mathcal{O}_S$-ideal generated by the discriminants $D(G)$ of the binary forms $G = \lambda F$ with $\lambda \in K^*$, $\lambda F \in \mathcal{O}_S[X, Y]$. Hence it suffices to prove the following: let $G(X, Y) \in \mathcal{O}_S[X, Y]$ be an irreducible binary form and suppose that there is a number θ with $G(\theta, 1) = 0$, $M' = K(\theta)$. Then

$$D(G) \in \mathbf{d}_{T'/S}. \tag{6.3}$$

Put $G(X, Y) = a_0 X^{r'} + a_1 X^{r'-1} Y + \cdots + a_r' Y^{r'}$. Arguing similarly as in the proof of Lemma 3 in [2], we infer that

$$\{1,\ a_0\theta,\ a_0\theta^2 + a_1\theta, \ldots, a_0\theta^{r-1} + \cdots + a_{r-2}\theta\}$$

is a K-basis of M – consisting of elements from the integral closure $\mathcal{O}_T$ of $\mathcal{O}_S$ in M', and the discriminant of this basis is just $D(G)$. Hence $D(G)$ belongs to the discriminant of $\mathcal{O}_{T'}$ over $\mathcal{O}_S$, which is $\mathbf{d}_{T'/S}$. This proves (6.3). □

Proof of Theorems 2 *and* 3. Suppose that $\mathbf{d}_S(F) = \mathbf{d}$. Using that $r_j \leqslant r$ for $j = 1, \ldots, m$, we get

$$\begin{aligned} D = \prod_{j=1}^{m} |D_{M_j}| &\leqslant \prod_{j=1}^{m} \{(|D_K|^{1/d}(rP)^t)^{r_j d} |\mathbf{d}_{T_j/S}|_S^d\} \text{ by Lemma 14} \\ &= |D_K|^r (rP)^{rtd} |\mathbf{d}_{T_1/S} \cdots \mathbf{d}_{T_m/S}|_S^d \\ &\leqslant |D_K|^r (rP)^{rtd} |\mathbf{d}|_S^d \text{ by Lemma 15 and (4.1).} \end{aligned}$$

Now Theorems 2 and 3 follow by inserting this upper bound for D into the upper bounds in the statements of Theorems 2′ and 3′, respectively. □

7. Proof of Corollary 3

K and S have the same meaning as before. We shall derive Corollary 3 from the following lemma.

LEMMA 16. *Let F, G be two $\mathcal{O}_S$-equivalent binary forms of non-zero discriminant. Then there are $\varepsilon \in \mathcal{O}_S^*$ and $U \in \mathrm{SL}(2, \mathcal{O}_S)$ such that*

$$G = \varepsilon F_U, \qquad h(U) \leqslant C_1,$$

where C_1 is an effectively computable number depending only on F and G.

Proof. In what follows, $C_2, \ldots, C_{21}$ will denote effectively computable numbers depending only on F and G. F and G have the same splitting field, L say, and one has $[L\colon \mathbf{Q}] \leqslant C_2$ and $|D_L| \leqslant C_3$. Let T be the set of places on L lying above the places in S. By assumption, there are $\varepsilon_0 \in \mathcal{O}_S^*$ and $U_0 \in \mathrm{SL}(2, \mathcal{O}_S)$ such that $G = \varepsilon_0 F_{U_0}$. We first show that F has a proper factorization (cf. §5) $(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r)$, and G has a proper factorization $(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_r)$ such that

$$\boldsymbol{a}_i \in \mathcal{O}_T^2, \qquad h(\boldsymbol{a}_i) \leqslant C_4, \qquad \boldsymbol{b}_i \in \mathcal{O}_T^2, \qquad h(\boldsymbol{b}_i) \leqslant C_4,$$

$$\lambda_i \boldsymbol{b}_i = U_0^T \boldsymbol{a}_i \quad \text{for } i = 1, \ldots, r, \tag{7.1}$$

where $\lambda_1, \ldots, \lambda_r \in L^*$. It is obvious that F has a proper factorization $(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_r)$ such that either $\boldsymbol{c}_i = (0, 1)^T$ or $\boldsymbol{c}_i = (1, \theta_i)^T$ for some $\theta_i \in L$. By Lemma 1, (vi) we have $h(\boldsymbol{c}_i) \leqslant C_5$ for $i = 1, \ldots, r$. F has a non-zero coefficient, θ say, such that $\boldsymbol{a}_i := \theta \boldsymbol{c}_i \in \mathcal{O}_T^2$ for $i = 1, \ldots, r$. Now Lemma 1, (i) implies that $h(\boldsymbol{a}_i) \leqslant C_4$ for $i = 1, \ldots, r$. $G(\boldsymbol{X})$ is a constant multiple of $\Pi_{i=1}^r (\boldsymbol{a}_i, U_0 \boldsymbol{X}) = \Pi_{i=1}^r (U_0^T \boldsymbol{a}_i, \boldsymbol{X})$. Hence G has a proper factorization $(\boldsymbol{d}_1, \ldots, \boldsymbol{d}_r)$ such that either $\boldsymbol{d}_i = (0, 1)^T$ or $\mathbf{d}_i = (1, \kappa_i)^T$ for some $\kappa_i \in L$, and $\boldsymbol{d}_i$ is a constant multiple of $U^T \boldsymbol{a}_i$. Now construct the proper factorization $(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_r)$ from $(\boldsymbol{d}_1, \ldots, \boldsymbol{d}_r)$ in a similar way as $(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r)$ from $(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_r)$. Then (7.1) is obviously satisfied.

We shall frequently use the following. Since $\det U_0 = 1$, we have $\det(\lambda_i \boldsymbol{b}_i, \lambda_j \boldsymbol{b}_j) = \det(\boldsymbol{a}_i, \boldsymbol{a}_j)$ hence $\lambda_i \lambda_j = \det(\boldsymbol{a}_i, \boldsymbol{a}_j)/\det(\boldsymbol{b}_i, \boldsymbol{b}_j)$. Together with Lemma 1, (i) and (7.1) this implies that

$$h(\lambda_i \lambda_j) \leqslant C_6 \quad \text{for } 1 \leqslant i < j \leqslant r \tag{7.2}$$

We distinguish between the cases $r = 2$ and $r > 2$.

$\mathbf{r = 2}$. We have either $L = K$ or $[L\colon K] = 2$. In the latter case, λ_1, λ_2 are each other's conjugates over K. There is a positive integer $R \leqslant C_7$ such that for every $\eta \in \mathcal{O}_T^*$:

$$\eta^R \equiv 1 \bmod \det(\boldsymbol{b}_1, \boldsymbol{b}_2). \tag{7.3}$$

Since $U_0 \in \mathrm{SL}(2, \mathcal{O}_S)$ and $\boldsymbol{a}_i, \boldsymbol{b}_i \in \mathcal{O}_T^2$ for $i = 1, 2$ we have, by (4.1), (7.1) and Lemma 1, (v),

$$|\lambda_i|_T = |U_0^T \boldsymbol{a}_i|_T / |\boldsymbol{b}_i|_T \leqslant |\boldsymbol{a}_i|_T \leqslant h(\boldsymbol{a}_i) \leqslant C_8 \quad \text{for } i = 1, 2.$$

Letting β_i be a non-zero coordinate of $\boldsymbol{b}_i$ we have, by (7.1) and Lemma 1, (v), that $\lambda_i\beta_i \in \mathcal{O}_T\backslash\{0\}$ and $|\lambda_i\beta_i|_T \leqslant C_9$ for $i = 1, 2$. Now Lemma 10 implies that there are $\varepsilon_1, \varepsilon_2 \in \mathcal{O}_T^*$ such that

$$h(\varepsilon_i^{2R}\lambda_i\beta_i) \leqslant C_{10} \quad \text{for } i = 1, 2.$$

Together with (7.1) this implies that

$$(\varepsilon_i^{2r}\lambda_i) \leqslant C_{11} \quad \text{for } i = 1, 2. \tag{7.4}$$

By Lemma 1, (iii) we may assume that if $[L: K] = 2$, then $\varepsilon_1, \varepsilon_2$ are each other's conjugates over K. Now Lemma 1, (i) implies that

$$h(\varepsilon_1\varepsilon_2)^{2R}\lambda_1\lambda_2) \leqslant C_{12}.$$

On using (7.2) and Lemma 1, (i) again, we get

$$h(\varepsilon_1\varepsilon_2) \leqslant C_{13}. \tag{7.5}$$

Put $\eta_i = \varepsilon_i^2/\varepsilon_1\varepsilon_2$ for $i = 1, 2$. Then (7.4), (7.5) and Lemma 1, (i) yield

$$h(\eta_i^R\lambda_i) \leqslant C_{14} \quad \text{for } i = 1, 2, \qquad \eta_1\eta_2 = 1,$$

$$\eta_1, \eta_2 \text{ are each other's conjugates over } K \quad \text{if } [L: K] = 2. \tag{7.6}$$

Let B_1 be the matrix with columns $\boldsymbol{b}_1, \boldsymbol{b}_2$, and B_2 the matrix with columns $\eta_1^R\boldsymbol{b}_1$, $\eta_2^R\boldsymbol{b}_2$. We claim that $B_2B_1^{-1} \in \mathrm{SL}(2, \mathcal{O}_S)$. First, $B_2B_1^{-1} \in K^{2,2}$. This is obvious if $L = K$; if $[L: K] = 2$, this follows from the fact that η_1, η_2 are each other's conjugates over K. Second, $\det(B_2B_1^{-1}) = 1$ since $\eta_1\eta_2 = 1$. Third, by (7.3) there is a matrix $C \in \mathcal{O}_T^{2,2}$ such that $B_2 = B_1 + \det(B_1)C_1$. Hence, if I is the identity matrix, then

$$B_2B_1^{-1} = I + C\{(\det B_1)B_1^{-1}\} \in \mathcal{O}_T^{2,2}.$$

This proves our claim. Putting $\mu_i = \eta_i^R\lambda_i$ $(i = 1, 2)$, $U = U_0(B_2B_1^{-1})^T$ we obtain, by (7.6),

$$\mu_i\boldsymbol{b}_i = U^T\boldsymbol{a}_i, \qquad h(\mu_i) \leqslant C_{15} \quad \text{for } i = 1, 2. \tag{7.7}$$

Together with Lemma 1 (i) and (7.1) this implies $h(U) \leqslant C_{16}$. Further, $U \in \mathrm{SL}(2, \mathcal{O}_S)$. Note that $(\boldsymbol{a}_1, X)(\boldsymbol{a}_2, X)$ is a constant multiple of F, hence $(U^T\boldsymbol{a}, X)(U^T\boldsymbol{a}_2, X) = (\boldsymbol{a}_1, UX)(\boldsymbol{a}_2, UX)$ is a constant multiple of F_U. Further, $(\boldsymbol{b}_1, X)(\boldsymbol{b}_2, X)$ is a constant multiple of G. Now (7.7) implies that $G = \lambda F_U$ for some $\lambda \in K^*$. But $(G)_S = (\varepsilon_0 F_{U_0})_S = (F_{U_0})_S = (F)_S, (F_U)_S = (F)_S$. Hence $(\lambda)_S = (1)_S$. This implies that $\lambda =: \varepsilon \in \mathcal{O}_S^*$. This proves Lemma 16 for $r = 2$.

r > 2. There are $\alpha_1, \alpha_2, \beta_1, \beta_2 \in L$ such that

$$\boldsymbol{a}_3 = \alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2, \qquad \boldsymbol{b}_3 = \beta_1\boldsymbol{b}_1 + \beta_2\boldsymbol{b}_2,$$

and by Lemma 1, (i) and (7.1) we have

$$h(\alpha_1), h(\alpha_2), h(\beta_1), h(\beta_2) \leqslant C_{17}.$$

Again by (7.1) we have

$$\lambda_3 \boldsymbol{b}_3 = U_0^T \boldsymbol{a}_3 = \alpha_1 U_0^T \boldsymbol{a}_1 + \alpha_2 U_0^T \boldsymbol{a}_2 = \alpha_1 \lambda_1 \boldsymbol{b}_1 + \alpha_2 \lambda_2 \boldsymbol{b}_2.$$

Hence $\lambda_1/\lambda_3 = \beta_1/\alpha_1$, $\lambda_2/\lambda_3 = \beta_2/\alpha_2$. Together with (7.8) and Lemma 1, (i), this implies that

$$h(\lambda_1/\lambda_2) \leqslant C_{18}. \tag{7.9}$$

By (7.2) we have $h(\lambda_1\lambda_2) \leqslant C_6$. Together with (7.9) and Lemma 1, (i), this implies that

$$h(\lambda_1),\ h(\lambda_2) \leqslant C_{19}.$$

By (7.1) we have $\lambda_i \boldsymbol{b}_i = U_0^T \boldsymbol{a}_i$, $h(\boldsymbol{a}_i)$, $h(\boldsymbol{b}_i) \leqslant C_4$ for $i = 1, 2$. Now Lemma 1, (i) gives

$$h(U_0) \leqslant C_{20}.$$

This proves Lemma 16 also for $r > 2$. □

Proof of Corollary 3. Assume that an irreducible polynomial $f(X) \in \mathbf{Z}[X]$ is given such that $K = \mathbf{Q}(\alpha)$ for some zero α of f, and that for each prime ideal in S a set of generators is given. By Theorem 3 (and the remarks made in §2 before the statement of Corollary 3), there is an effectively computable finite set of binary forms in $K[X, Y]$ of degree $r \geqslant 2$ such that every binary form F in $\mathcal{O}_S[X, Y]$ of degree r with $D(F) \in \mathcal{O}_S^*$ is $\mathcal{O}_S$-equivalent to one of these forms. For each form in the finite set it can be checked whether its coefficients belong to $\mathcal{O}_S$ and its discriminant belongs to $\delta\mathcal{O}_S^*$. Thus we get a finite set of binary forms containing at least (but possibly more than) one form from each $\mathcal{O}_S$-equivalent class. We can decide in the following way whether any two binary forms F, G in that finite set are $\mathcal{O}_S$-equivalent. Compute a finite set of matrices $U \in K^{2,2}$, containing all matrices with $h(U) \leqslant C_1$, where C_1 is the constant in Lemma 16, and check for each of these matrices U whether $U \in \mathrm{SL}(2, \mathcal{O}_S)$ and there is an $\varepsilon \in K^*$ with $S = \varepsilon F_U$; then check if $\varepsilon \in \mathcal{O}_S$. In this way, one can compute a set of binary forms, containing exactly one form from each class. □

8. Proof of Theorem 4

K, d, S, s have the same meaning as in §2. F is a square free binary form in $K[X, Y]$ with $\deg F = r$ and the splitting field L of F over K has degree g over K. It suffices to prove Theorem 4 for $L = K$. For assume that Theorem 4 holds in this case. Let T be the set of places on L lying above those in S. Then, by our assumption with L instead of K,

$$r \leqslant 3 \times (7^{[L:\mathbf{Q}]+2|T|} + 1) + \sqrt{3\Omega_T(\mathbf{d}_T(F))}, \tag{8.1}$$

where Ω_T is defined similarly as Ω_S. Now Theorem 4 follows in full generality by inserting the inequalities $[L:\mathbf{Q}]=gd$, $|T|\leqslant gs$ and $\Omega_T(\mathbf{d}_T(F))\leqslant g\Omega_S(d_S(F))$ into (8.1).

So from now on, we assume that F factors into linear forms in $K[X, Y]$ (and hence that $g = 1$), and we shall prove Theorem 4 in this case. Further, we assume that $r \geqslant 4$ which is obviously no restriction. Put

$$A = 3 \times (7^{d+2s} + 1), \qquad \Omega = \Omega_S(\mathbf{d}_S(F)).$$

We need the following lemma.

LEMMA 17. *Let* $\mathbf{a}, \mathbf{b}$ *be fixed* $\mathcal{O}_S$*-ideals. Then the number of pairs* (x, y) *with*

$$x + y = 1, \qquad (x)_S = \mathbf{a},\ (y)_S = \mathbf{b} \tag{8.2}$$

is at most $3 \times 7^{d+2s}$.

Proof. Suppose that (8.2) is solvable, and fix a solution (α, β) of (8.2). Then every solution (x, y) of (8.2) can be expressed uniquely as $x = \alpha\xi$, $y = \beta\eta$ with $\xi, \eta \in \mathcal{O}_S^*$. Now Lemma 17 follows at once from Theorem 1 of [4], which states that $\alpha\xi + \beta\eta = 1$ in $\xi, \eta \in \mathcal{O}_S^*$ has at most $3 \times 7^{d+2s}$ solutions. □

Proof of Theorem 4. By our assumption, F can be expressed as

$$F(X) = \prod_{i=1}^{r} (\boldsymbol{a}_i, X)$$

with $\boldsymbol{a}_i \in K^2$. Put $\Delta_{ij} = \det(\boldsymbol{a}_i, \boldsymbol{a}_j)$ and define the $\mathcal{O}_S$-ideals

$$\mathbf{d}_{ij} = \frac{(\Delta_{ij})_S}{(\boldsymbol{a}_i)_S(\boldsymbol{a}_j)_S} \quad \text{for } i, j \in \{1, \ldots, r\},\ i \neq j.$$

Then each ideal $\mathbf{d}_{ij}$ is $\mathcal{O}_S$-integral. Further, by (6.1) (cf. proof of Lemma 15) we have

$$\mathbf{d}_S(F) = \prod_{1\leqslant i<j\leqslant r} \mathbf{d}_{ij}^2. \tag{8.3}$$

For each triple $1 \leqslant i < j < k \leqslant r$, we define the integral $\mathcal{O}_S$-ideal

$$\mathbf{c}(i, j, k) = \prod_{l\neq i,j,k} (\mathbf{d}_{il}\mathbf{d}_{jl}\mathbf{d}_{kl}).$$

Then, by (8.3)

$$\prod_{1\leqslant i<j<k\leqslant r} \mathbf{c}(i, j, k) = \mathbf{d}_S(F)^{\binom{r-2}{2}}.$$

Put $\mathbf{c} = \mathbf{c}(1,2,3)$. We assume that

$$\Omega_S(\mathbf{c}) \leqslant \frac{\binom{r-2}{2}}{\binom{r}{3}}\Omega_S(\mathbf{d}_S(F)) \leqslant \frac{3}{r}\Omega, \tag{8.4}$$

which is no restriction, by the fact that $\Omega_S(\mathbf{ab}) = \Omega_S(\mathbf{a})\Omega_S(\mathbf{b})$ for any two integral $\mathcal{O}_S$-ideals $\mathbf{a}, \mathbf{b}$.

We partition $\{1, \ldots, r\}$ into three sets:

$$\mathscr{A}_1 = \{1, 2, 3\}, \qquad \mathscr{A}_2 = \{l: 4 \leqslant l \leqslant r: \mathbf{d}_{1l}\mathbf{d}_{2l}\mathbf{d}_{3l} = (1)_S\},$$

$$\mathscr{A}_3 = \{1, \ldots, r\} \backslash (\mathscr{A}_1 \cup \mathscr{A}_2).$$

For each l in $\mathscr{A}_3$ we have $l \geqslant 4$ and $\Omega_S(\mathbf{d}_{1l}\mathbf{d}_{2l}\mathbf{d}_{3l}) \geqslant 1$. Together with (8.4) and the additivity of Ω_S this implies that

$$|\mathscr{A}_3| \leqslant \Omega_S(\mathbf{c}) \leqslant \frac{3}{r}\Omega. \tag{8.5}$$

We now estimate $|\mathscr{A}_2|$ from above. By (5.9) (cf. proof of Lemma 13) we have

$$\Delta_{23}\Delta_{1l} + \Delta_{31}\Delta_{2l} + \Delta_{12}\Delta_{3l} = 0,$$

or

$$x_l + y_l = 1 \quad \text{for } l \in \mathscr{A}_2, \tag{8.6}$$

where

$$x_l = \frac{\Delta_{23}\Delta_{1l}}{\Delta_{21}\Delta_{3l}}, \qquad y_l = \frac{\Delta_{31}\Delta_{2l}}{\Delta_{21}\Delta_{3l}}.$$

Note that

$$(x_l)_S = \frac{\mathbf{d}_{23}\mathbf{d}_{1l}}{\mathbf{d}_{21}\mathbf{d}_{3l}}, \qquad (y_l)_S = \frac{\mathbf{d}_{31}\mathbf{d}_{2l}}{\mathbf{d}_{21}\mathbf{d}_{3l}}.$$

Hence

$$(x_l)_S = \mathbf{d}_{23}/\mathbf{d}_{21}, \qquad (y_l)_S = \mathbf{d}_{31}/\mathbf{d}_{21} \quad \text{for } l \in \mathscr{A}_2. \tag{8.7}$$

The pairs (x_l, y_l), $(l \in \mathscr{A}_2)$ are distinct. Indeed, suppose that $(x_p, y_p) = (x_q, y_q)$ for some $p, q \in \mathscr{A}_2$. Then $x_p/y_p = x_q/y_q$, whence $\Delta_{1p}/\Delta_{2p} = \Delta_{1q}/\Delta_{2q}$. But

$$\boldsymbol{a}_p = \frac{\Delta_{2p}}{\Delta_{21}}\left(\boldsymbol{a}_1 - \frac{\Delta_{1p}}{\Delta_{2p}}\boldsymbol{a}_2\right), \qquad \boldsymbol{a}_q = \frac{\Delta_{2q}}{\Delta_{21}}\left(\boldsymbol{a}_1 - \frac{\Delta_{1q}}{\Delta_{2q}}\boldsymbol{a}_2\right).$$

Hence $\boldsymbol{a}_p, \boldsymbol{a}_q$ are linearly dependent. This implies that $p = q$. Now Lemma 17 implies, in view of (8.6), (8.7), that

$$|\mathscr{A}_2| \leqslant 3 \times 7^{d+2s}.$$

By combining this with (8.5), we get

$$r = |\mathscr{A}_1| + |\mathscr{A}_2| + |\mathscr{A}_3| \leqslant A + \frac{3\Omega}{r},$$

or

$$r^2 - Ar - 3\Omega \leqslant 0.$$

This yields

$$r \leqslant \tfrac{1}{2}(A + \sqrt{A^2 + 12\Omega}) \leqslant A + \sqrt{3\Omega},$$

as required. □

9. Proof of Corollaries 4, 5, 6, and 7

K, d, S, s will have the same meaning as in §3. Further, if $F(X, Y) \in K[X, Y]$ is a binary form associated to the system of fields $(M_1, \ldots, M_m)$, then we put

$$r = \deg F, \qquad D = |D_{M_1} \cdots D_{M_m}|.$$

Of Corollaries 4, 5, and 6 we shall prove only the parts (ii), by using Theorems 2′ or 3′; the parts (i) can be derived in the same way from Theorems 2 or 3.

Proof of Corollary 4. $C_1, \ldots, C_6$ denote effectively computable numbers depending only K, S, r, D. Let $F(X, Y) \in K[X, Y]$ be a binary form of degree $r \geqslant 2$ with minimal S-discriminant and put $u = \omega_S(\mathbf{d}_S(F))$. Let T and $C_S(F)$ have the same meaning as in the statement of Corollary 4. Further, let Q be the largest of the prime numbers lying below the prime ideals in T. Note that $\mathbf{d}_T(F) = (1)_T$. By Theorem 2′, F is weakly $\mathcal{O}_T$-equivalent to a binary form $G \in K[X, Y]$ for which

$$\log\log h(G) \leqslant C_1(u + 1)\log(u + 1) + C_2 \log Q.$$

Together with the inequality $u \leqslant C_3 \log(3C_S(F))/\log\log(3C_S(F))$ and the obvious inequality $\log Q \leqslant C_4 \log(3C_S(F))$, this gives

$$\log\log h(G) \leqslant C_5 \log(3C_S(F)). \tag{9.1}$$

By Lemma 3, we have $|\mathbf{d}_S(G)|_S \leqslant r^{2r-1} h(G)^{2r^2-2}$. Further, $|\mathbf{d}_S(F)|_S \leqslant |\mathbf{d}_S(G)|_S$ since F has minimal S-discriminant. By combining these inequalities with (9.1), we get

$$\log\log |\mathbf{d}_S(F)|_S \leqslant C_6 \log(3C_S(F)),$$

as required. □

Proof of Corollary 5. Let $C_1, \ldots, C_4$ be effectively computable numbers of the same form as the upper bound for $h(G)$ in Theorem 3′, but with other absolute constants. Let $F(X, Y) \in \mathcal{O}_S[X, Y]$ be a binary form of degree $r \geqslant 2$ with discriminant $D(F) \neq 0$. By Theorem 3′, there are an S-unit ε, a matrix $U \in \mathrm{SL}(2, \mathcal{O}_S)$, and a binary form $G \in \mathcal{O}_S[X, Y]$, such that $F_U = \varepsilon G$, $h(G) \leqslant C_1$. By Lemma 10, there are $y_1, y_2 \in \mathcal{O}_S^*$ such that

$$\varepsilon = \eta_1 \eta_2^r, \qquad h(\eta_1) \leqslant C_2.$$

Put $G' = \eta_1 G$. Then $F_{\eta_2^{-1}U} = G'$ and $h(G') \leqslant C_3$. There is a rational integer a with $0 \leqslant a \leqslant r$ such that $G(1,a) \neq 0$. Let $(x_0, y_0)^T = \eta_2^{-1}U(1,a)^T$. Then, by Lemma 1, (ii),

$$h(F(x_0, y_0)) = h(F_{\eta_2^{-1}U}(1, a)) = h(G'(1, a)) \leqslant C_4. \qquad \square$$

Proof of Corollary 6. $C_1, \ldots, C_6$ will be effectively computable numbers of the same form as the upper bound for $h(\theta^*)$ in Corollary 6, (ii), but with other absolute constants. First we assume that θ has degree $r \geqslant 2$ over K. By Theorem 2′, the binary form F_θ (cf. (3.1)) is weakly $\mathcal{O}_S$-equivalent to a form G with $h(G) \leqslant C_1$. One of the zeros of the polynomial $G(X, 1)$, θ^* say, must be $\mathcal{O}_S$-equivalent to θ and by Lemma 1, (vi) we have $h(\theta^*) \leqslant C_2$.

Now suppose that $\theta \in K^*$. By Lemma 4, the $\mathcal{O}_S$-ideal $\mathbf{a} = (1, \theta)_S^{-1}$ contains a number $\alpha \neq 0$ with $|\alpha|_S \leqslant |D_K|^{1/2d}|\mathbf{a}|_S$, where $d = [K:Q]$. Put $\gamma = \alpha\theta$, $\mathbf{b} = (\alpha, \gamma)_S$. Then $\mathbf{b} \subseteq \mathcal{O}_S$ and $|\mathbf{b}|_S \leqslant |D_K|^{1/2d}$. Again by Lemma 4, $\mathbf{b}$ contains a non-zero number ω with $|\omega|_S \leqslant |D_K|^{1/d}$. By Lemma 10, there is an $\varepsilon \in \mathcal{O}_S^*$ such that $h(\varepsilon\omega) \leqslant C_3$. Put $\Delta = \varepsilon\omega$. Then $\Delta \in (\alpha, \gamma)_S$, hence there are $\beta, \delta \in \mathcal{O}_S$ such that $\alpha\delta - \beta\gamma = \Delta$. Let $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$. By Lemma 7, there is a matrix $U \in \mathrm{SL}(2, \mathcal{O}_S)$ such that $h(UA) \leqslant C_4$. At least one of the matrices UA, $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} UA$, $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} UA$ has the property that the two entries of its first column are both non-zero. Together with Lemma 1, (i), this implies that there is a matrix $U_1 \in \mathrm{SL}(2, \mathcal{O}_S)$ such that $h(U_1 A) \leqslant C_5$, and the entries in the first column of $U_1 A$, α_1, γ_1, say, are both non-zero. Put $\theta^* = \gamma_1/\alpha_1$. Then θ^* is $\mathcal{O}_S$-equivalent to $\gamma/\alpha = \theta$, and $h(\theta^*) \leqslant C_6$ by Lemma 1, (i). $\square$

Proof of Corollary 7. Let M/K be an extension of degree r, and $\{\omega_1, \ldots, \omega_r\}$ a K-basis of M. Let $C_1, \ldots, C_4$ be effectively computable numbers of the same form as the upper bound in Corollary 7, but with other absolute constants. By Corollary 6, every solution $\boldsymbol{x} = (x_1, \ldots, x_r) \in K^r$ of (3.2) is $\mathcal{O}_S$-equivalent to a solution $\boldsymbol{x}^* = (x_1^*, \ldots, x_r^*) \in K^r$ of (3.2) for which

$$h(x_1^*\omega_1 + \cdots + x_r^*\omega_r) \leqslant C_1. \tag{9.2}$$

Let $\sigma_1, \ldots, \sigma_r$ be the K-isomorphisms of M, and let $\Sigma = (\sigma_i(\omega_j))$. Then by (9.2) and Lemma 1, (ii),

$$\Sigma \boldsymbol{x}^* = \boldsymbol{a}, \qquad h(\boldsymbol{a}) \leqslant C_2.$$

Note that $\det(\Sigma) \neq 0$. Hence, by Lemma 1, $h(\Sigma^{-1}) \leqslant C_3$. Now another application of Lemma 1, yields that

$$h(\boldsymbol{x}^*) = h(\Sigma^{-1}\boldsymbol{a}) \leqslant C_4. \qquad \square$$

10. Relationship between binary forms of given discriminant and *S*-unit equations

Let K be an algebraic number field and S a finite set of places on K. C_1, C_2, C_3 are effectively computable numbers depending only on K and S. We shall show that Theorem 2 with $r \geqslant 4$ implies that every solution of the S-unit equation

$$\xi + \eta = 1 \quad \text{in } \xi, \eta \in \mathcal{O}_S^* \tag{10.1}$$

satisfies $\max(h(\xi), h(\eta)) \leqslant C_1$. We use some properties of cross ratios. The cross ratio of a binary form $F(X) = \lambda \Pi_{i=1}^4 (\boldsymbol{a}_i, X)$ of degree 4 is defined by

$$\rho(F) = \frac{\det(\boldsymbol{a}_1, \boldsymbol{a}_2) \cdot \det(\boldsymbol{a}_3, \boldsymbol{a}_4)}{\det(\boldsymbol{a}_1, \boldsymbol{a}_4) \cdot \det(\boldsymbol{a}_2, \boldsymbol{a}_3)}.$$

Note that $\rho(F)$ is independent of the choice of $\boldsymbol{a}_1, \dots, \boldsymbol{a}_4$. Further, for each constant α and each non-singular 2×2-matrix A one has $\rho(\alpha F_A) = \rho(F)$. Each $\boldsymbol{a}$ can be chosen as $(0, 1)^T$ or $(1, \theta_i)^T$. Hence if F has algebraic coefficients then by Lemma 1, (i), (vi), $h(\rho(F))$ can be effectively estimated from above in terms of $h(F)$.

To each solution (ξ, η) of (10.1) we associate the binary form $F(X, Y) = XY(X + Y)(\xi X - \eta Y)$. This form has discriminant

$$D(F) = (\xi\eta(\xi + \eta))^2 \in \mathcal{O}_S^*.$$

By Theorem 2, there are a binary form $G \in K[X, Y]$, $\lambda \in K^*$ and $U \in \mathrm{SL}(2, \mathcal{O}_S)$ such that

$$F = \lambda G_U, \qquad h(G) \leqslant C_2.$$

Hence

$$(h(\rho(F)) = h(\rho(G)) \leqslant C_3.$$

But $\rho(F) = -\xi/\eta$, hence $h(\xi/\eta) \leqslant C_3$. Together with (10.1) and Lemma 1, (i) this implies that $\max(h(\xi), h(\eta)) \leqslant C_1$. □

References

[1] A. Baker, Contributions to the theory of Diophantine equations. I. On the representation of integers by binary forms, *Philos. Trans. Roy. Soc. London Ser.* A 263 (1968), 173–191.

[2] B. J. Birch and J. R. Merriman, Finiteness theorems for binary forms with given discriminant, *Proc. London Math. Soc.* 25 (1972), 385–394.

[3] J. Coates, An effective p-adic analogue of a theorem of Thue, *Acta Arith.* 15 (1969), 275–305.

[4] J. H. Evertse, On equations in S-units and the Thue–Mahler equation, *Invent. Math.* 75 (1984), 561–584.

[5] J. H. Evertse and K. Győry, Thue–Mahler equations with a small number of solutions, *J. Reine Angew. Math.* 399 (1989), 60–80.

[6] C. F. Gauss, Disquisitiones Arithmeticae, 1801 (German translation, 2nd edn, reprinted, Chelsea Publ. New York, 1981).

[7] K. Győry, Sur les polynômes à coefficients entiers et de discriminant donné, *Acta Arith.* 23 (1973), 419–426.

[8] K. Győry, Sur les polynômes à coefficients entiers et de discriminant donné. II. *Publ. Math. Debrecen* 21 (1974), 125–144.

[9] K. Győry, On polynomials with integer coefficients and given discriminant, V, p-adic generalizations, *Acta Math. Acad. Sci. Hungar.* 32 (1978), 175–190.

[10] K. Győry, On the number of solutions of linear equations in units of an algebraic number field, *Comment. Math. Helvetici* 54 (1979), 583–600.

[11] K. Győry, On the solutions of linear diophantine equations in algebraic integers of bounded norm, Ann. Univ. Sci. Budapest. *Eötvös. Sect. Math.* 22–23 (1979–80), 225–233.

[12] K. Győry, Effective finiteness theorems for Diophantine problems and their applications, Academic Doctor's thesis, *Debrecen*, 1983 (in Hungarian).

[13] K. Győry, Effective finiteness theorems for polynomials with given discriminant and integral elements with given discriminant over finitely generated domains, *J. Reine Angew. Math.* 346 (1984), 54–100.

[14] K. Győry and Z. Z. Papp, On discriminant form and index form equations, *Studia Scient. Math. Hung.* 12 (1977), 47–60.

[15] C. Hermite, Sur l'introduction des variables continues dans la théorie des nombres, *J. Reine Angew. Math.* 41 (1851), 191–216.

[16] A. Korkine and G. Zolotareff, Sur les formes quadratiques, *Math. Ann.* 6 (1873), 366–389.

[17] J. L. Lagrange, Recherches d'arithmétique, Nouv. Mém. Acad. Berlin, 1773, 265–312, Oeuvres, III, 693–758.

[18] S. Lang, Algebraic Number Theory, Addison-Wesley Publ., Reading, Mass., 1970.

[19] S. Lang, Fundamentals of Diophantine Geometry, Springer Verlag, New York, 1983.

[20] D. J. Lewis and K. Mahler, Representation of integers by binary forms, *Acta. Arith.* 6 (1961), 333–363.

[21] K. Mahler, Über die Annäherung algebraischer Zahlen durch periodische Algorithmen, *Acta Math.* 68 (1937), 109–144.

[22] A. Markoff, Sur les forms quadratiques binaires indéfinies, *Math. Ann.* 15 (1879), 381–406.

[23] L. J. Mordell, On numbers represented by binary cubic forms, *Proc. London Math. Soc.* 48 (1945), 198–228.

[24] C. L. Siegel (Under the pseudonym X), The integer solutions of the equation $y^2 = ax^n + bx^{n-1} + \cdots + k$, *J. London Math. Soc.* 1 (1926), 66–68.

[25] C. L. Siegel, Abschätzung von Einheiten, *Nachr. Göttingen Math. Phys. Kl.* (1969), 71–86.

[26] H. M. Stark, Some effective cases of the Brauer–Siegel Theorem, *Invent. Math.* 23 (1974), 135–152.

[27] E. Weiss, Algebraic Number Theory, McGraw-Hill, New York, 1963.

[28] R. Zimmert, Ideale kleiner Norm in Idealklassen und eine Regulatorabschätzung, *Invent. Math.* 62 (1981), 367–380.