Publ. Math. Debrecen
61/3-4 (2002), 267–314

Geometry of a cubic Thue equation

By RYOTARO OKAZAKI (Kyoto)

Dedicated to Professor Kálmán Győry on his 60th birthday

Abstract. Let $f(X,Y) = aX^3+bX^2Y+cXY^2+dY^3 \in \mathbf{Z}[X,Y]$ be a binary cubic form and denote by $D = D(f)$ its discriminant. Under the hypothesis $D \geq 5.65 \cdot 10^{65}$, we shall prove that the number of representations of 1 by f, i.e., the number of integral solutions to the equation $f(x,y) = 1$, is at most 7. This improves upon the previous upper bound 10 of BENNETT [1].

Under various other conditions, we verify the Pethő–Lippok Conjecture, which asserts that if $D > 361$ then the number of representations of 1 by f is at most 5.

1. Introduction

Let $f(X,Y) = aX^3 + bX^2Y + cXY^2 + dY^3 \in \mathbf{Z}[X,Y]$ be a binary cubic form and denote by $D = D(f)$ its discriminant. A pair $(x,y) \in \mathbf{Z}^2$ is called *a representation of* 1 *by* f if $f(x,y) = 1$. Let $\mathcal{R} = \mathcal{R}(f)$ be the set of representations of 1 by f. We are interested in the number $\#\mathcal{R}$ of representations.

DELONE [6] in 1922 and NAGELL [18] in 1928 proved the best possible upper bound for $\#\mathcal{R}$ in the case that f has negative discriminant: $\#\mathcal{R} \leq 3$ if $D < -44$; $\#\mathcal{R} = 4$ if $D = -44$ or if $D = -31$; and $\#\mathcal{R} = 5$ if $D = -23$.

If f is reducible over $\mathbf{Q}$, Bezout Theorem implies $\#\mathcal{R} \leq 4$.

Mathematics Subject Classification: Primary 11D25; Secondary 11D45, 11D59.
Key words and phrases: cubic Thue equation, representation of unity, cubic form.

Therefore, we assume $D > 0$ and f is irreducible over $\mathbf{Q}$. We mention that in this case, the *Hessian* $H(f;X,Y) = -\frac{1}{4}\left(\frac{\partial^2 f(X,Y)}{\partial X^2}\frac{\partial^2 f(X,Y)}{\partial Y^2} - \left(\frac{\partial^2 f(X,Y)}{\partial X \partial Y}\right)^2\right)$ of f is positive definite quadratic form. We say f is reduced if $H(f;X,Y)$ is a reduced quadratic form (see §2 or §3).

We shall prove the following theorem:

Theorem 1.1. *If* $D \geq 5.65 \cdot 10^{65}$, *we have*

$$\#\mathcal{R} \leq 7.$$

If $f(X,Y)$ *is reduced,* $f(1,0) \neq \pm 1$ *and* $D \geq 2.15 \cdot 10^{69}$, *we have*

$$\#\mathcal{R} \leq 6.$$

The research on the number $\#\mathcal{R}$ of representations was started by THUE work [25] in 1909, who proved the finiteness of $\#\mathcal{R}$ by using Padé approximations. In 1929, SIEGEL [22] (see also [23]) used hypergeometric functions for refining Thue's Theorem and proved $\#\mathcal{R} \leq 18$ for sufficiently large D. Their method, now called the hypergeometric method, is ineffective in the sense it gives no upper bound for the size $\max\{|x|,|y|\}$. However, it can give a good universal upper bound for $\#\mathcal{R}$. Indeed, Gel'man proved $\#\mathcal{R} \leq 10$ for sufficiently large D, in his student paper of 1949 (see [7, Chap 5]). More recently, EVERTSE [9] in 1983 proved $\#\mathcal{R} \leq 12$ unconditionally and BENNETT [1, Theorem 1.4] in 2001 improved the unconditional bound to $\#\mathcal{R} \leq 10$. A general effective method was invented by Baker, who proved an effective lower bound for linear forms in logarithms of algebraic numbers. This method is usually used for computing the set $\mathcal{R}$ of representations when concrete data of the cubic field associated with f are given.

Our proof of Theorem 1.1 uses Baker's method but in such a way that we do not have to investigate the data of any specific number field. This leads to an upper bound for $\#\mathcal{R}$ which is smaller than the best known result obtained by the hypergeometric method, i.e., Bennett's result cited above.

Under various additional conditions on the cubic form f, we prove the following conjecture of PETHŐ [19] in 1987 and LIPPOK [13] in 1993:

*Conjecture 1 (*PETHŐ [19] and LIPPOK [13]*).* If $D > 361$, we have

$$\#\mathcal{R} \leq 5.$$

The first condition is on the cubic order associated with f. Without loss of generality, we assume $\#\mathcal{R} > 0$. Then, according to Theorem 4.1 below, there are algebraic integers α and β belonging to some suitable cubic field such that f factors as a *norm form*

$$f(X,Y) = \prod_{i=1}^{3} (\alpha_i X + \beta_i Y)\,, \tag{1}$$

where α_i $(i = 1,2,3)$ are the conjugates of α and β_i $(i = 1,2,3)$ those of β and such that the discriminant of the order $\mathbf{Z}[\alpha,\beta]$ is equal to the discriminant D of f. Further, it is shown in Theorem 4.1 that $\mathbf{Z}[\alpha,\beta]$ is up to isomorphism indpendent of the choice of α and β. We now put $\mathfrak{O}(f) = \mathbf{Z}[\alpha,\beta]$ and call this the order associated with f.

Let $\mathfrak{O}$ be a totally real cubic order and denote the conjugates of $\theta \in \mathfrak{O}$ by subscripts: θ_i $(i = 1,2,3)$. The element $\mathrm{sgn}(\varepsilon) = {}^t(\varepsilon_i/|\varepsilon|_i)$ of $\{\pm 1\}^3 \simeq \mathbf{F}_2^3$ is called the signature of ε when $\varepsilon \in \mathfrak{O}^\times$. The signature rank of $\mathfrak{O}^\times$ is the rank of the image $\mathrm{sgn}(\mathfrak{O}^\times)$. (Hence, the cardinality of the image is 2^s, where s denotes the signature rank.)

Theorem 1.2. *Assume that the signature rank of* $\mathfrak{O}(f)^\times$ *equals* 1. *If* $D \geq 5.65 \cdot 10^{65}$, *we have*

$$\#\mathcal{R} \leq 5.$$

An important invariant of f is the group of its automorphisms: $\mathrm{Aut}(f) = \{M \in GL_2(\mathbf{Z}) | f \circ M = f\}$, where for a matrix $M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$ we define the cubic form $f \circ M$ by $(f \circ M)(X,Y) = f(m_{11}X + m_{12}Y, m_{21}X + m_{22}Y)$. We shall prove a better upper bound for $\#\mathcal{R}$ when $\mathrm{Aut}(f)$ is non-trivial:

Theorem 1.3. *Assume* $\mathrm{Aut}(f) \neq 1$. *If* $D \geq 2.56 \cdot 10^{18}$, *we have*

$$\#\mathcal{R} \in \{0,3\}.$$

If the signature rank of $\mathfrak{O}(f)^\times$ *is less than* 3, *then we have*

$$\#\mathcal{R} = 0.$$

This is a generalization of THOMAS [24] of 1990, who determined all representations of 1 by the simplest cubic forms $f_N(X,Y) = X^3 + (N -$

$1)X^2Y-(N+2)XY^2+Y^3$ with integer parameter $N \geq 1.365 \cdot 10^7$. (See also [17] for the complete determination of all solutions in the full range, i.e., $N \neq 0, 1, 3$.) Our generalization of Thomas' result is non-trivial, since there are infinitely many cubic forms f other than the forms f_N for which $\#\mathcal{R} \geq 3$ and $\mathrm{Aut}(f) \neq 1$. (See the remark after Corollary 3.2 of [1].) Our result also improves upon the previous estimate $\#\mathcal{R} \leq 9$ of BENNETT [1, Corollary 3.2] obtained by the hypergeometric method.

Since units of trace zero in a given cubic order correspond to representations of 1 by a certain binary cubic form, our result has an application to such units. Indeed, an attempt for the following theorem was the prototype for the current research.

Theorem 1.4. *Let $\mathfrak{O}$ be a totally real cubic order and $D(\mathfrak{O})$ its discriminant. Denote by $\mathcal{T}$ the set of units of $\mathfrak{O}$ whose traces are 0 and norms are 1. Assume $D(\mathfrak{O}) \geq 7.2 \cdot 10^7$. Then, we have*

$$\#\mathcal{T} \leq 3.$$

If the signature rank of $\mathfrak{O}^\times$ is 2, we have

$$\#\mathcal{T} \leq 2.$$

If the field of fractions of $\mathfrak{O}$ is cyclic, we have $\#\mathcal{T} = 3$ or 0 according as $D(\mathfrak{O}) = 81$ or not.

We remark that if the signature rank of $\mathfrak{O}^\times$ is 1, then $\mathcal{T}$ is empty. Indeed, in this case, each element of $\mathfrak{O}^\times$ can be expressed as $\pm\varepsilon$ where ε is totally positive, hence each element of $\mathfrak{O}^\times$ has trace $\neq 0$.

The set $\mathcal{T}$ is non-empty when the order $\mathfrak{O}$ is generated by a root θ of $X^3 - kX - 1$ with $k \geq 3$. Hence, there are infinity of such orders $\mathfrak{O}$ for which $\mathcal{T}$ is non-empty. When $k = 3$, the order $\mathfrak{O}$ becomes the ring of integers of the cyclic cubic field of discriminant 81. Hence, the three conjugates of θ lie in $\mathcal{T}$.

The structure of our proof is as follows: The key idea is to use the geometry of a particular continuous plane curve which enables us to prove exponential gap principles of a geometric nature. The curve is found (§5) after interpretation of the Hessian of f as an exterior product (§3) and after giving a geometric interpretation of the representations of 1 by f as points lying in some displaced lattice in the logarithmic space (§4). Then, our gap principles assert that these points are lying very far apart. We use Baker theory (§7) to prove effective upper bounds for the sizes of these

points. This makes a difference from previous studies (including Bennett's) in which uniform upper bounds for $\#\mathcal{R}$ were obtained by means of the hypergeometric method. To obtain nice numeric constants, we use results of LAURENT–MIGNOTTE–NESTERENKO [12] and MATVEEV [15] from Baker theory. We apply geometry of numbers (§6) in the logarithmic space (or rather a modification of this) to rewrite lower bounds for linear forms in logarithms in terms of our geometric set-up. Our theorems (§8) follow by combining the exponential gap principles mentioned above with the upper bounds obtained from Baker theory.

The author thanks the referee whose comments greatly improved the paper.

2. Preliminaries

We recall here what is known about lattices, quadratic forms and cubic forms.

We begin with lattices and quadratic forms associated with them. A finitely generated module $\mathfrak{L}$ of a metric $\mathbf{R}$-linear space $\mathfrak{V}$ is called a lattice if it is discrete. A minimal set of generators $\boldsymbol{a}^{(1)}, \boldsymbol{a}^{(2)}, \dots, \boldsymbol{a}^{(m)}$ of the lattice $\mathfrak{L}$ is called a basis of $\mathfrak{L}$. The number m is called the rank of $\mathfrak{L}$. A basis always consists of $\mathbf{R}$-linearly independent vectors. Therefore, the rank m never exceeds the dimension of $\mathfrak{V}$. The volume of the fundamental parallelepiped of $\mathfrak{L}$ in the linear hull $\mathbf{R}\mathfrak{L}$ of $\mathfrak{L}$ is written $\mathrm{disc}(\mathfrak{L})$ and called the covolume of $\mathfrak{L}$ in $\mathbf{R}\mathfrak{L}$. The quadratic form $Q : (X^{(1)}, X^{(2)}, \dots, X^{(m)}) \longmapsto \|X^{(1)}\boldsymbol{a}^{(1)} + X^{(2)}\boldsymbol{a}^{(2)} + \cdots + X^{(m)}\boldsymbol{a}^{(m)}\|^2$ is said to be associated with the lattice $\mathfrak{L}$ or more precisely with the basis $\boldsymbol{a}^{(1)}, \boldsymbol{a}^{(2)}, \dots, \boldsymbol{a}^{(m)}$. The Voronoï domain of $\mathbf{0}$ in $\mathfrak{V}$ with respect to $\mathfrak{L}$ is the domain

$$\{\boldsymbol{u} \in \mathfrak{V} \mid \|\boldsymbol{u} - \boldsymbol{v}\| \leq \|\boldsymbol{u}\| \text{ for all } \boldsymbol{v} \in \mathfrak{L}\}.$$

Every point in $\mathfrak{V}$ can be translated by some lattice vector of $\mathfrak{L}$ into this Voronoï domain. (The translated point is not unique in general. However, we do not care about uniqueness in this paper.)

Vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ of a lattice $\mathfrak{L}$ of rank 2 are called successive minima of $\mathfrak{L}$ if the following two conditions are satisfied:

(i) $\|\boldsymbol{a}\| \leq \|\boldsymbol{v}\|$ for every vector $\boldsymbol{v}$ in $\mathfrak{L} \setminus \{\mathbf{0}\}$;

(ii) $\|\boldsymbol{b}\| \leq \|\boldsymbol{v}\|$ for every vector $\boldsymbol{v}$ in $\mathfrak{L} \setminus \mathbf{Z}\boldsymbol{a}$.

Successive minima of a lattice$\mathfrak{L}$ of rank 2 always form a basis of $\mathfrak{L}$, which we call a reduced basis of $\mathfrak{L}$. A reduced basis $\boldsymbol{a}$ and $\boldsymbol{b}$ of a lattice $\mathfrak{L}$ of rank 2 can be effectively computed from a given basis $\boldsymbol{a}'$ and $\boldsymbol{b}'$ of $\mathfrak{L}$. If $\|\boldsymbol{a}'\| \leq \|\boldsymbol{b}\|'$, we can easily deduce $\|\boldsymbol{a}\| \leq \|\boldsymbol{a}'\|$ and $\|\boldsymbol{b}\| \leq \|\boldsymbol{b}'\|$. This process is called Lagrange reduction.

It is easy to verify that the radius of the Voronoï domain of $\mathbf{0}$ in a plane with respect to $\mathfrak{L}$ of rank 2 is less than or equal to $\sqrt{2}\,\|\boldsymbol{b}\|$, where $\boldsymbol{b}$ is a second minimum of $\mathfrak{L}$.

A basis $\boldsymbol{a}$ and $\boldsymbol{b}$ of $\mathfrak{L}$ is reduced if and only if $2|\boldsymbol{a} \bullet \boldsymbol{b}| \leq \|\boldsymbol{a}\|^2 \leq \|\boldsymbol{b}\|^2$, where $\boldsymbol{a} \bullet \boldsymbol{b}$ denotes the inner product of $\boldsymbol{a}$ and $\boldsymbol{b}$. This in particular means that vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ form an angle between $\pi/3$ and $2\pi/3$ (inclusively). Hence,

$$\frac{\sqrt{3}}{2}\,\|\boldsymbol{a}\|^2 \leq \frac{\sqrt{3}}{2}\,\|\boldsymbol{a}\| \cdot \|\boldsymbol{b}\| \leq \operatorname{disc}(\mathfrak{L}) \leq \|\boldsymbol{a}\| \cdot \|\boldsymbol{b}\| \leq \|\boldsymbol{b}\|^2. \tag{2}$$

If $Q(X, Y) = aX^2 + bXY + cY^2 = \|X\boldsymbol{a} + Y\boldsymbol{b}\|^2$ is a quadratic form associated with the basis $\boldsymbol{a}$ and $\boldsymbol{b}$, the mentioned characterization of a reduced basis is written as $|b| \leq a \leq c$. We also say Q is reduced if this condition is satisfied. (Note: every positive definite binary quadratic form can be associated with some lattice of rank 2.) The discriminant $D(Q) = b^2 - 4ac$ of Q equals $-4\operatorname{disc}(\mathfrak{L})^2$. (The reader should be careful since the meaning of the term "discriminant" of Q in different references varies within $D(Q)$, $\operatorname{disc}(\mathfrak{L})^2$ and $\operatorname{disc}(\mathfrak{L})$. In this paper, a discriminant is a discriminant of a binary form. For avoiding confusion, we speak of covolumes of lattices.)

The reader is referred to [4, pp. 26–33] for these facts or [21] for more historical treatment.

Vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ of a lattice $\mathfrak{L}$ of rank 3 are called successive minima of $\mathfrak{L}$ if the following three conditions are satisfied:

(i) $\|\boldsymbol{a}\| \leq \|\boldsymbol{v}\|$ for every vector $\boldsymbol{v}$ in $\mathfrak{L} \setminus \{\mathbf{0}\}$;

(ii) $\|\boldsymbol{b}\| \leq \|\boldsymbol{v}\|$ for every vector $\boldsymbol{v}$ in $\mathfrak{L} \setminus \mathbf{Z}\boldsymbol{a}$;

(iii) $\|\boldsymbol{c}\| \leq \|\boldsymbol{v}\|$ for every vector $\boldsymbol{v}$ in $\mathfrak{L} \setminus (\mathbf{Z}\boldsymbol{a} + \mathbf{Z}\boldsymbol{b})$.

Successive minima of $\mathfrak{L}$ of rank 3 always form a basis of $\mathfrak{L}$, which we call a Seeber-reduced basis of $\mathfrak{L}$ (see [8] for a beautiful geometric proof). A reduced basis $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ of a lattice $\mathfrak{L}$ of rank 3 can be effectively computed from a given basis $\boldsymbol{a}'$, $\boldsymbol{b}'$ and $\boldsymbol{c}'$ of $\mathfrak{L}$ (see [26] for another beautiful geometric proof). If $\|\boldsymbol{a}'\| \leq \|\boldsymbol{b}'\| \leq \|\boldsymbol{c}'\|$, we can easily deduce $\|\boldsymbol{a}\| \leq \|\boldsymbol{a}'\|$, $\|\boldsymbol{b}\| \leq \|\boldsymbol{b}'\|$ and $\|\boldsymbol{c}\| \leq \|\boldsymbol{c}'\|$. We call this process Seeber–Vallée reduction.

The reader is referred to [21, Chap. 10] for the characterization of a Seeber-reduced basis in terms of its associated quadratic form.

We need more sophisticated notion of successive minima and reduction, that is due to Minkowski, for lattices of higher rank. However, we do not go into its detail since we will only use lattices of rank at most 3.

Now, we move to cubic forms. Let $f(X,Y) = aX^3 + bX^2Y + cXY^2 + dY^3$ be cubic form be given. Let $f(X,Y) = \prod_{i=1}^{3}(a_iX + b_iY)$ be some factorization.Then,

$$D(f) = \prod_{i=1}^{3} \begin{vmatrix} a_i & b_i \\ a_{i+1} & b_{i+1} \end{vmatrix}^2 = b^2c^2 - 27a^2d^2 + 18abcd - 4ac^3 - 4b^3d$$

is called the discriminant of f and

$$H(f;X,Y) = -\frac{1}{4} \begin{vmatrix} \dfrac{\partial^2 f(X,Y)}{\partial X^2} & \dfrac{\partial^2 f(X,Y)}{\partial X \partial Y} \\ \dfrac{\partial^2 f(X,Y)}{\partial Y \partial X} & \dfrac{\partial^2 f(X,Y)}{\partial Y^2} \end{vmatrix}$$

$$= (b^2 - 3ac)X^2 + (bc - 9ad)XY + (c^2 - 3bd)Y^2$$

is called the Hessian of f. It is positive definite if $D(f) > 0$. Straightforward calculation implies

$$D(H(f)) = -3D(f).$$

We say f is reduced if $D(F) > 0$ and the Hessian $H(f)$ is reduced.

An element M of $GL_2(\mathbf{Z})$ acts on a binary form f by

$$(f \circ M)(X,Y) = f(Z,W), \quad \text{where} \quad \begin{pmatrix} Z \\ W \end{pmatrix} = M \begin{pmatrix} X \\ Y \end{pmatrix}.$$

Under this action, the Hessian $H(f)$ of a binary cubic form f is a covariant of f, i.e., $H(f \circ M) = H(f) \circ M$. Indeed, this fact is a motivation of the definition of reducedness of a binary cubic form.

For the elegant theory of binary forms in the greater picture, the reader is referred to [11, pp. 1–150]. For direct treatment of binary cubic forms, the reader is referred to [4, pp. 51–55], [5, pp. 400–418] or [7, pp. 166–176].

3. The Hessian as an exterior product

Let $f(X,Y) \in \mathbf{Z}[X,Y]$ be a given irreducible homogeneous cubic form. We denote its discriminant by $D = D(f)$. We assume $D > 0$. We shall associate a canonical lattice $\mathfrak{L}^\natural$ in $\mathbf{R}^3$ with f. The canonical lattice will be constructed from a factorization of f in such a way that its independence from the factorization is guaranteed. (In §4, we choose an optimal factorization for our purpose. The choice, however, is made after a formulation that is suggested by the canonical lattice.)

In the sequel, we adopt the following conventions: an element of $\mathbf{R}^3$ is considered as a column vector; subscripts denote indices of components of vectors in $\mathbf{R}^3$ and are read modulo 3; (z_i) denotes a column vector ${}^t(z_1, z_2, z_3)$ and a bold letter $\boldsymbol{z}$ denotes (z_i). We also write $\mathbf{0} = {}^t(0,0,0)$ and $\mathbf{1} = {}^t(1,1,1)$. We equip $\mathbf{R}^3$ with coordinatewise multiplication: if $\boldsymbol{a} = (a_i)$ and $\boldsymbol{b} = (b_i)$ then $\boldsymbol{ab} = (a_i b_i)$. The inner product of $\boldsymbol{a} = (a_i)$ and $\boldsymbol{b} = (b_i)$ is given by $\boldsymbol{a} \bullet \boldsymbol{b} = a_1b_1 + a_2b_2 + a_3b_3$. We also equip the space $\mathbf{R}^3$ with functions *norm* $N : (x_i) \longmapsto x_1x_2x_3$ and *trace* $\mathrm{tr} : (x_i) \longmapsto x_1 + x_2 + x_3$. When we consider a totally real cubic field $\mathfrak{K}$, we denote its embeddings into $\mathbf{R}$ by $\gamma \longmapsto \gamma_i$ $(i = 1,2,3)$ and form a vector in the way compatible to our convention: $\gamma \in \mathfrak{K} \longmapsto \boldsymbol{\gamma} = (\gamma_i) \in \mathbf{R}^3$. An embedding of the normal closure $\tilde{\mathfrak{K}}$ in $\mathbf{R}^3$ will be specified in §4.

Definition 3.1. Let

$$f(X,Y) = \prod_{i=1}^{3} (a_iX + b_iY) \tag{3}$$

be a factorization of $f(X,Y)$ in $\mathbf{R}[X,Y]$. Let $\boldsymbol{a} \times \boldsymbol{b}$ be the exterior product

$$\boldsymbol{a} \times \boldsymbol{b} = \left(\begin{vmatrix} a_{i+1} & b_{i+1} \\ a_{i+2} & b_{i+2} \end{vmatrix} \right)$$

of $\boldsymbol{a} = (a_i)$ and $\boldsymbol{b} = (b_i)$. Then, we have

$$N(\boldsymbol{a} \times \boldsymbol{b})^2 = D(f). \tag{4}$$

The lattice $\mathfrak{L}^\natural = (\boldsymbol{a} \times \boldsymbol{b})(\mathbf{Z}\boldsymbol{a} + \mathbf{Z}\boldsymbol{b})$ of rank 2 is called the canonical lattice associated with f. Here, the product of $\boldsymbol{a} \times \boldsymbol{b}$ and a lattice vector is the componentwise product.

PROOF that $\mathfrak{L}^\natural$ is well-defined. We prove uniqueness and then existence of $\mathfrak{L}^\natural$. The identity (4) will be verified in the latter step.

Assume one factorization of the form (3) is given. Then, an arbitrary factorization of $f(X,Y)$ is of the form

$$f(X,Y) = \prod_{i=1}^{3} (\zeta_i a_i X + \zeta_i b_i Y),$$

where $\boldsymbol{\zeta} = (\zeta_i) \in \mathbf{R}^3$ has norm 1.

We have $(\boldsymbol{\zeta a} \times \boldsymbol{\zeta b})\boldsymbol{\zeta a} = (\boldsymbol{a} \times \boldsymbol{b})\boldsymbol{a}$ since

$$(\boldsymbol{\zeta a} \times \boldsymbol{\zeta b})\boldsymbol{\zeta a} = \left(\begin{vmatrix} \zeta_{i+1} a_{i+1} & \zeta_{i+1} b_{i+1} \\ \zeta_{i+2} a_{i+2} & \zeta_{i+2} b_{i+2} \end{vmatrix} \zeta_i a_i \right) = \left(\zeta_i \zeta_{i+1} \zeta_{i+2} \begin{vmatrix} a_{i+1} & b_{i+1} \\ a_{i+2} & b_{i+2} \end{vmatrix} a_i \right).$$

Similarly, we have $(\boldsymbol{\zeta a} \times \boldsymbol{\zeta b})\boldsymbol{\zeta b} = (\boldsymbol{a} \times \boldsymbol{b})\boldsymbol{b}$.

The invariance of the basis of $\mathfrak{L}^\natural$ implies uniqueness of $\mathfrak{L}^\natural$.

On the other hand, there exists a factorization of the form (3). We can simply set $a_i = a = f(1,0)^{1/3}$ and $b_i = -a\omega_i$ $(i = 1,2,3)$ where ω_1, ω_2 and ω_3 are the three roots of $f(X,1)$.

Linear independence of $(\boldsymbol{a} \times \boldsymbol{b})\boldsymbol{a}$ and $(\boldsymbol{a} \times \boldsymbol{b})\boldsymbol{b}$ follows from

$$N(\boldsymbol{a} \times \boldsymbol{b})^2 = \prod_{i=1}^{3} \begin{vmatrix} a & -a\omega_{i+1} \\ a & -a\omega_{i+2} \end{vmatrix}^2 = a^{12} \prod_{i=1}^{3} (\omega_{i+1} - \omega_{i+2})^2. \tag{5}$$

Indeed, the right hand side is non-zero since f is irreducible. Non-vanishing of the norm $N(\boldsymbol{a} \times \boldsymbol{b})$ implies $\boldsymbol{a} \times \boldsymbol{b} \in (\mathbf{R}^3)^\times = \{(z_i) \mid z_1 z_2 z_3 \neq 0\}$. In particular, we get $\boldsymbol{a} \times \boldsymbol{b} \neq \mathbf{0}$. Hence, $\boldsymbol{a}$ and $\boldsymbol{b}$ are linearly independent over $\mathbf{R}$. Further, multiplication by $\boldsymbol{a} \times \boldsymbol{b} \in (\mathbf{R}^3)^\times$ preserves linear independence of vectors. Hence, $\mathfrak{L}^\natural = \mathbf{Z}(\boldsymbol{a} \times \boldsymbol{b})\boldsymbol{a} + \mathbf{Z}(\boldsymbol{a} \times \boldsymbol{b})\boldsymbol{b}$ is a lattice of rank 2.

Finally, the identity (5) implies (4). □

Remark. The exterior product $\boldsymbol{a} \times \boldsymbol{b}$ (up to orientation) is an invariant of the lattice $\mathbf{Z}\boldsymbol{a} + \mathbf{Z}\boldsymbol{b}$. Indeed, its direction is determined by the plane $\mathbf{R}\boldsymbol{a} + \mathbf{R}\boldsymbol{b}$ and its magnitude is determined by the covolume $\operatorname{disc}(\mathbf{Z}\boldsymbol{a} + \mathbf{Z}\boldsymbol{b})$ of the lattice $\mathbf{Z}\boldsymbol{a} + \mathbf{Z}\boldsymbol{b}$ in $\mathbf{R}\boldsymbol{a} + \mathbf{R}\boldsymbol{b}$.

Remark. In general, if we choose $\boldsymbol{a} = (a_i)$ and $\boldsymbol{b} = (b_i)$ as in the proof, then for $i = 1,2,3$ the field $\mathbf{Q}(a_i, b_i)$ will have degree 9. We shall later recall a way of choosing $\boldsymbol{a}$ and $\boldsymbol{b}$ such that for $i = 1,2,3$, the field $\mathbf{Q}(a_i, b_i)$

becomes cubic. Such a choice will be essential for our study. However, such a choice will depend on a representation of 1 though f. The advantage of the choice in the proof is that it does not require such a representation.

Several properties of the lattice $\mathfrak{L}^\natural$ are interesting:

Proposition 3.2. *The lattice $\mathfrak{L}^\natural$ is contained in the plane Π which is orthogonal to* $\mathbf{1}$ *and its covolume* $\mathrm{disc}(\mathfrak{L}^\natural)$ *in Π is $\sqrt{3D}$.*

If the basis $(\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{a}$ and $(\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{b}$ of $\mathfrak{L}^\natural$ is reduced by $GL_2(\mathbf{Z})$, we have

$$\|(\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{a}\|^2 \geq 3\sqrt[3]{2D} \quad \textit{and} \quad \|(\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{b}\|^2 \geq \sqrt{3D}.$$

Further, the quadratic form $\|X(\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{a}+Y(\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{b}\|^2$ is proportional to the Hessian $H(f;X,Y)$ of $f(X,Y)$:

$$\|X(\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{a}+Y(\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{b}\|^2 = 2H(f;X,Y).$$

PROOF. By definition of exterior product, we have

$$\mathbf{1}\bullet(\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{a} = (\boldsymbol{a}\times\boldsymbol{b})\bullet\boldsymbol{a} = 0,$$

where $\bullet$ denotes the inner product. Similarly, we have $\mathbf{1}\bullet(\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{b} = 0$. Therefore, the lattice $\mathfrak{L}^\natural$ is on the plane Π. Its covolume $\mathrm{disc}(\mathfrak{L}^\natural)$ in Π is the magnitude of the exterior product

$$(\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{a}\times(\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{b} = \left(\begin{vmatrix} a_{i+1}d_{i+1} & b_{i+1}d_{i+1} \\ a_{i+2}d_{i+2} & b_{i+2}d_{i+2}\end{vmatrix}\right) = N((d_i))\,\mathbf{1},$$

of its basis, where we set

$$(d_i) = \left(\begin{vmatrix} a_{i+1} & b_{i+1} \\ a_{i+2} & b_{i+2}\end{vmatrix}\right).$$

By the identity (4), the magnitude of the vector in the right hand side is $\sqrt{3D}$. Thus, the value of $\mathrm{disc}(\mathfrak{L}^\natural)$ is established.

By the identity (4) and the inequality of the arithmetic and the geometric means, we get the lower bound

$$\|(\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{a}\|^2 \geq 3\sqrt[3]{N((\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{a})^2} = 3\sqrt[3]{Df(1,0)^2} \geq 3\sqrt[3]{D},$$

which is of the same order of the corresponding lower bound of the theorem. The improved constant is obtained by Lagrange's method of unknown multipliers.

The estimate of $\|(\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{b}\|^2$ is established in the usual way of geometry of numbers:

$$\|(\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{b}\|^2 \geq \|(\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{a}\| \cdot \|(\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{b}\| \geq \operatorname{disc}(\mathfrak{L}^\natural) = \sqrt{3D}.$$

The expression for the Hessian is verified by straight-forward calculation. (Or one can appeal to the uniqueness (see [11, p. 45]) of a quadratic covariant of a cubic form.) □

The study of representations of 1 by f can now be formulated in terms of lattice points:

Proposition 3.3. *Define the curve $\mathcal{H}$ on Π by*

$$(\mathcal{H}): \qquad N(\boldsymbol{z}) = \sqrt{D}, \quad \operatorname{tr}\boldsymbol{z} = 0.$$

Representations of 1 by f and lattice points of $\mathfrak{L}^\natural$ on the curve $\mathcal{H}$ are in bijective correspondence and hence we have

$$\#\mathcal{R} = \#\left(\mathfrak{L}^\natural \cap \mathcal{H}\right).$$

Proof. We follow the notation of Definition 3.1. The map $(m,n) \in \mathbf{Z}^2 \longmapsto m(\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{a} + n(\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{b} \in \mathfrak{L}^\natural$ is obviously a bijection. On the other hand, the identity (4) implies

$$N(m(\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{a} + n(\boldsymbol{a}\times\boldsymbol{b})\boldsymbol{b}) = \pm\sqrt{D}N(m\boldsymbol{a}+n\boldsymbol{b}) = \pm\sqrt{D}f(m,n).$$

Since the lattice is closed under inversion of vectors, the assertion is now established. □

The curve $\mathcal{H}$ consists of 3 branches (connected components):

$$(\mathcal{H}^j): \qquad (z_i) \in \mathcal{H}, \quad z_j > 0.$$

Points on $\mathcal{H}^j$ satisfy $z_{j+1}, z_{j+2} < 0$ and $|z_{j+1}|, |z_{j+2}| < |z_{j+1}|+|z_{j+2}| = |z_j|$.

Each branch $\mathcal{H}^j$ has two asymptotic lines $z_{j+1} = 0$ and $z_{j+2} = 0$ on the plane Π. We cut the branch $\mathcal{H}^j$ at the "middle" and divide it in two parts. More precisely, we define the curve $\mathcal{H}_k$ by

$$(\mathcal{H}_k): \qquad (z_i) \in \mathcal{H}, \quad |z_k| \leq |z_{k+1}|,\ |z_{k+2}|$$

and set

$$\mathcal{H}_k^j = \mathcal{H}^j \cap \mathcal{H}_k$$

for $j \neq k$. (Note: $\mathcal{H}^j$ and $\mathcal{H}_j$ do not intersect.) The six pieces of the curve do not intersect except at the three points $c^{\,t}(2,-1,-1)$, $c^{\,t}(-1,2,-1)$ and $c^{\,t}(-1,-1,2)$, where c designates $\sqrt[6]{D}/\sqrt{2}$.

Let $\boldsymbol{z} = (z_i)$ be a vector of Π. We introduce local coordinates of $\boldsymbol{z}$ for each piece $\mathcal{H}_k$:

$$p = p(\boldsymbol{z}) = \frac{z_{k+1} - z_{k+2}}{\sqrt{2}}, \quad q = q(\boldsymbol{z}) = -\frac{\sqrt{6}\, z_k}{2} \tag{6}$$

and

$$\boldsymbol{e}^{\parallel}(k)=(e_i^{\parallel}(k)), \quad e_k^{\parallel}(k)=0, \quad e_{k+1}^{\parallel}(k)=\frac{1}{\sqrt{2}}, \quad e_{k+2}^{\parallel}(k)= -\frac{1}{\sqrt{2}}; \tag{7}$$

$$\boldsymbol{e}^{\perp}(k)=(e^{\perp}(k)), \quad e_k^{\perp}(k)= -\frac{2}{\sqrt{6}}, \quad e_{k+1}^{\perp}(k)=\frac{1}{\sqrt{6}}, \quad e_{k+2}^{\perp}(k) = \frac{1}{\sqrt{6}}.$$

Then, the vectors $\boldsymbol{e}^{\parallel}(k)$ and $\boldsymbol{e}^{\perp}(k)$ form an orthonormal basis of Π and we have

$$\boldsymbol{z} = p(\boldsymbol{z})\boldsymbol{e}^{\parallel}(k) + q(\boldsymbol{z})\boldsymbol{e}^{\perp}(k). \tag{8}$$

Lemma 3.4. *The piece $\mathcal{H}_k$ of the curve $\mathcal{H}$ s defined by*

$$\left(3p^2 - q^2\right) q = 3\sqrt{6D}, \quad |p| \geq \sqrt{3}\, q > 0$$

in terms of the local coordinates $p = p(\boldsymbol{z})$ and $q = q(\boldsymbol{z})$ on the plane Π. In particular, we have

$$0 < q \leq 9\sqrt{6D}/8p^2.$$

Assume $\boldsymbol{z}, \boldsymbol{z}' \in \mathfrak{L}^{\natural} \cap \mathcal{H}_k$. Assume also $|p(\boldsymbol{z}')| \geq |p(\boldsymbol{z})|$. Then, we have

$$|p(\boldsymbol{z}')| > (2\sqrt{2}/9)p(\boldsymbol{z})^2.$$

Remark. The first assertion indicates the close connection between $\mathfrak{L}^{\natural}$ and the Lagrange resolvent, which has been used together with the hypergeometric method. Indeed, we have

$$\left((p + q\sqrt{-1}\,)/\sqrt{2}\right)^3 - ((p - q\sqrt{-1})/\sqrt{2}\,)^3 = 3\sqrt{-3D}.$$

(For a general treatment of resolvent, see Lecture XX and XXI of Hilbert's lecture notes [11] of 1897. For its application in Thue equations, see §7 of [22] and §70, Chap. V of [7].)

Remark. The gap principle of Lemma 3.4 corresponds to Lemma 5.2 of [1], i.e., $|(p(\boldsymbol{z}')+q(\boldsymbol{z}')\sqrt{-1})/\sqrt{2}| \geq (1-\epsilon)|(p(\boldsymbol{z})+q(\boldsymbol{z})\sqrt{-1})/\sqrt{2}|^2$ for $|(p(\boldsymbol{z})+q(\boldsymbol{z})\sqrt{-1})/\sqrt{2}| \gg_\epsilon 0$ with explicit specification of the constant of "$\gg_\epsilon$" for a suitable ϵ. Our proof of the last assertion is already known. However, we included it since the inequality (9) occurring in it will be used in the proof of Lemma 5.4 below.

Remark. We consider Lemma 3.4 as a linear gap principle in the logarithmic space. Indeed, the lower bound reads $\log|p(\boldsymbol{z}')| \geq 2\log|p(\boldsymbol{z})| - \epsilon$ where the logarithms are close to the coordinates to be introduced in §5. In contrast, our geometric gap principles Theorems 5.5 and 5.6 are exponential.

PROOF. Write $\boldsymbol{z} = {}^t(z_1, z_2, z_3)$. Let (i,j,k) be a permutation of $(1,2,3)$ such that $z_i \leq z_k \leq 0 < z_j$. Then, we have $|z_j| = |z_i| + |z_k| \geq 2|z_k|$. Hence, $p = (|z_{k+1}| + |z_{k+2}|)/\sqrt{2} \geq 3|z_k|/\sqrt{2} = \sqrt{3}q > 0$.

The equation follows immediately after substitution of (8) in the definition of $\mathcal{H}$. Then, the estimate of q follows.

The vectors $\boldsymbol{z}$ and $\boldsymbol{z}'$ are linearly independent when they are on the curve $\mathcal{H}$. (Note: $N(c\boldsymbol{z}) = c^3 N(\boldsymbol{z})$ for arbitrary $c \in \mathbf{R}$.) Hence, we have

$$\sqrt{3D} = \operatorname{disc}(\mathfrak{L}^\natural) \leq \|\boldsymbol{z} \times \boldsymbol{z}'\| = \left|\det\begin{pmatrix} p & p' \\ q & q' \end{pmatrix}\right|, \tag{9}$$

where $p = p(\boldsymbol{z})$, $q = q(\boldsymbol{z})$; $p' = p(\boldsymbol{z}')$, $q' = q(\boldsymbol{z}')$. We now use the estimate of q just established to show

$$\left|\det\begin{pmatrix} p & p' \\ q & q' \end{pmatrix}\right| \leq \frac{9\sqrt{6D}\,|p'|}{8p^2} + \frac{9\sqrt{6D}\,|p|}{8(p')^2} \leq \frac{9\sqrt{6D}\,|p'|}{4p^2}.$$

The lower bound for $p(\boldsymbol{z}')$ is now obvious. □

4. Displaced lattice in the logarithmic space

Let $f(X,Y) \in \mathbf{Z}[X,Y]$ be a given irreducible homogeneous cubic form of positive discriminant such that $\#\mathcal{R} > 0$. We associate with f a displaced lattice in the plane $\Pi_{\log}$ which is orthogonal to $\mathbf{1}$ in the space of logarithms. We map the representations into this displaced lattice. We will also study the geometric properties of the displaced lattice.

Our interpretation of the Hessian, i.e., $\mathfrak{L}^\natural$ of rank 2 in $\mathbf{R}^3$, nables us to exploit the multiplicative structure of $\mathbf{R}^3$ whileits implementation as a lattice in the complex plane $\mathbf{C}$ (as in [22]) enables us to exploit the multiplicative structure of $\mathbf{C}$. Our method thus enables us to exploit the multiplicative action of the unit group associated with f on the space $\mathbf{R}^3$.

The multiplicative action comes from a Minkowski embedding ($\gamma \in \mathfrak{K} \longmapsto \boldsymbol{\gamma} = (\gamma_i) \in \mathbf{R}^3$) where $\mathfrak{K}$ is a totally real cubic field. Obviously, the Minkowski embedding is an injective ring homomorphism of $\mathfrak{K}$ into the ring $\mathbf{R}^3$ equipped with componentwise operation. Hence, the unit group of (a given order of) $\mathfrak{K}$ acts on $\mathbf{R}^3$. The norm and the trace on $\mathfrak{K}$ are compatible with those on $\mathbf{R}^3$: $N(\gamma) = N(\boldsymbol{\gamma})$ and $\mathrm{tr}(\gamma) = \mathrm{tr}(\boldsymbol{\gamma})$.

The associated order associated with f and its effect are determined by the following:

Theorem 4.1. *Assume $\#\mathcal{R} > 0$. Then, the following three assertions hold:*

(i) *There is a pair of algebraic integers α and β such that*

$$\left.\begin{array}{ll} & \mathbf{Q}(\alpha,\beta) \text{ is a totally real cubic field,} \\ (10) & \text{the order } \mathbf{Z}[\alpha,\beta] \text{ has discriminant } D(\mathbf{Z}[\alpha,\beta]) \text{ equal to } D, \\ & f(X,Y) = \prod_{i=1}^{3} (\alpha_i X + \beta_i Y)\,, \end{array}\right\}$$

where subscripts i designate the three real embeddings of $\mathbf{Q}(\alpha,\beta)$.

(ii) *If the norm of $\zeta \in \mathbf{Z}[\alpha,\beta]$ to $\mathbf{Q}$ equals 1, then the pair $(\mu,\nu) = (\zeta\alpha, \zeta\beta)$ satisfies* (10) *and $\mathbf{Z}[\mu,\nu] = \mathbf{Z}[\alpha,\beta]$.*

(iii) *Moreover, the order $\mathbf{Z}[\alpha,\beta]$ is uniquely determined up to isomorphism, i.e., independent of the choice of α and β with* (10). *We denote the order $\mathbf{Z}[\alpha,\beta]$ by $\mathfrak{O}(f)$ and call it the associated order with f. Further, we write $\mathfrak{K} = \mathfrak{K}(f)$ for the field $\mathbf{Q}(\alpha,\beta)$.*

(iv) *Every lattice point of $\mathfrak{L}^\natural$ lying on the curve $\mathcal{H}$ can be expressed as $(\boldsymbol{\alpha} \times \boldsymbol{\beta})\boldsymbol{\varepsilon}$ for some $\varepsilon \in \mathfrak{O}(f)^\times$.*

PROOF. It is convenient to start from (i), although it is well-known. Let $(x,y) \in \mathcal{R}$. Then, there is a matrix

$$M = \begin{pmatrix} x & * \\ y & * \end{pmatrix} \in SL_2(\mathbf{Z}).$$

We have

$$(f \circ M)(1,0) = f(x,y) = 1.$$

Therefore, $(f \circ M)(Z, W)$ factors as

$$(f \circ M)(Z, W) = \prod_{i=1}^{3} (1, -\theta_i) \begin{pmatrix} Z \\ W \end{pmatrix},$$

where $(\theta_1, \theta_2, \theta_3)$ is the conjugate triple of a totally real cubic integer θ. Set

$$(\alpha, \beta) = (1, \theta) M^{-1}. \tag{11}$$

Then, $\mathbf{Q}(\alpha, \beta)$ is clearly a totally real cubic field and we get

$$f(X, Y) = \prod_{i=1}^{3} (1, -\theta_i) M^{-1} \begin{pmatrix} X \\ Y \end{pmatrix} = \prod_{i=1}^{3} (\alpha_i, \beta_i) \begin{pmatrix} X \\ Y \end{pmatrix}.$$

The equality (11) also implies $\mathbf{Z}\alpha + \mathbf{Z}\beta = \mathbf{Z}1 + \mathbf{Z}\theta$. Hence, we get $\mathbf{Z}[\alpha, \beta] = \mathbf{Z}[\mathbf{Z}\alpha + \mathbf{Z}\beta] = \mathbf{Z}[\mathbf{Z}1 + \mathbf{Z}\theta] = \mathbf{Z}[\theta]$. Now, invariance of the discriminant under the action of $GL_2(Z)$ implies $D = D(f) = D(f \circ M) = D(\mathbf{Z}[\theta]) = D(\mathbf{Z}[\alpha, \beta])$. This proves (i).

We now prove (ii). Let $\zeta \in \mathbf{Z}[\alpha, \beta] = \mathbf{Z}[\theta]$ satisfy $N(\zeta) = 1$. Then, we obviously have

$$f(X, Y) = \prod_{i=1}^{3} (\zeta_i \alpha_i X + \zeta_i \beta_i Y).$$

On the other hand, we have

$$\mathbf{Z}[\zeta\alpha, \zeta\beta] = \mathbf{Z}[\mathbf{Z}\zeta\alpha + \mathbf{Z}\zeta\beta] = \mathbf{Z}[\mathbf{Z}\zeta + \mathbf{Z}\zeta\theta] = \mathbf{Z}[\zeta, \zeta\theta] = \mathbf{Z}[\zeta][\zeta\theta] = \mathbf{Z}[\zeta][\theta],$$

where the last identity follows from $\zeta \in \mathbf{Z}[\zeta]^{\times}$. Thus, we get

$$\mathbf{Z}[\zeta\alpha, \zeta\beta] = \mathbf{Z}[\theta][\zeta] = \mathbf{Z}[\theta].$$

We now prove (iii). Let (μ, ν) be any other pair of algebraic integers satisfying (10). Write $(\mu, \nu)M = (\lambda, -\lambda\vartheta)$. Then, we have $\prod_{i=1}^{3}(Z - \theta_i W) = \prod_{i=1}^{3}(\lambda_i Z - \lambda_i \vartheta_i W)$. Hence, we can identify ϑ with θ by isomorphism. The fact $N(\lambda) = 1$ is obvious. We follow calculation of the previous paragraph to get $\mathbf{Z}[\mu, \nu] = \mathbf{Z}[\theta][\lambda] = \mathbf{Z}[\alpha, \beta][\lambda]$. Now, equality of discriminants implies $\mathbf{Z}[\mu, \nu] = \mathbf{Z}[\alpha, \beta]$.

We now porve (iv). Let $(\boldsymbol{\alpha} \times \boldsymbol{\beta})(x\boldsymbol{\alpha} + y\boldsymbol{\beta})$ be a point of $\mathfrak{L}^{\natural}$ lying on the curve $\mathcal{H}$. By definition of $\mathcal{H}$, we have $N((\boldsymbol{\alpha} \times \boldsymbol{\beta})(x\boldsymbol{\alpha} + y\boldsymbol{\beta})) = \sqrt{D}$.

On the other hand, we have $N(\boldsymbol{\alpha} \times \boldsymbol{\beta}) = \pm\sqrt{D}$ by (4). Therefore, we get $N(x\boldsymbol{\alpha} + y\boldsymbol{\beta}) = \pm 1$. Put $\varepsilon = \pm(x\alpha + y\beta)$ with the suitable sign. Then, we now have $\varepsilon \in \mathfrak{O}(f)^{\times}$. □

Remark. We shall use the second assertion and normalize $\boldsymbol{a} \times \boldsymbol{b}$ by a unit of $\mathbf{Z}[\alpha, \beta]$ later in this section. Then, $\mathbf{Z}\boldsymbol{a} + \mathbf{Z}\boldsymbol{b}$ will also be normalized.

Remark. Multiplication by a possible unit of $\mathbf{Q}(\alpha, \beta)$ outside $\mathbf{Z}[\alpha, \beta]$ can change the order generated by α and β. Indeed, we have

$$\mathbf{Z}[\lambda, \lambda\theta] = \mathbf{Z}[\lambda][\lambda\theta] = \mathbf{Z}[\lambda][\theta] = \mathbf{Z}[\theta][\lambda] \neq \mathbf{Z}[\theta]$$

if λ is a unit of $\mathbf{Q}(\alpha, \beta)$ outside $\mathbf{Z}[\alpha, \beta]$. However, the change is detected by smaller discriminant.

Remark. The proof of uniqueness of $\mathbf{Z}[\alpha, \beta]$ also implies that $\mathbf{Z}[\boldsymbol{\alpha}, \boldsymbol{\beta}]$ (up to permutation of components) is determined by f.

Remark. In general, if $\varepsilon \in \mathfrak{O}(f)^{\times}$, then the coordinatewise product $(\boldsymbol{\alpha} \times \boldsymbol{\beta})\boldsymbol{\varepsilon}$ lies outside the image of the Minkowski embedding $\gamma \longmapsto (\gamma_1, \gamma_2, \gamma_3)$ of $\mathfrak{K}(f)$ into $\mathbf{R}^3$ unless $\mathfrak{K}(f)/\mathbf{Q}$ is cyclic. This is most easily seen when $(\alpha, \beta) = (1, -\theta)$ and $\varepsilon = 1$. In this situation, $\boldsymbol{\alpha} \times \boldsymbol{\beta} = (\theta_{i+2} - \theta_{i+1})$. However, $\theta_{i+2} - \theta_{i+1}$ generates the normal closure of $\mathfrak{K}(f)$ over $\mathfrak{K}(f)$.

Hereafter, we apply the machinery of §3 by choosing $\boldsymbol{a} = \boldsymbol{\alpha}$ and $\boldsymbol{b} = \boldsymbol{\beta}$. We identify $\mathfrak{K}(f)$ with its image of the Minkowski embedding in $\mathbf{R}^3$. We also identify the normal closure $\tilde{\mathfrak{K}}$ of $\mathfrak{K}(f)$ with $\mathfrak{K}(\boldsymbol{\alpha} \times \boldsymbol{\beta})$ in $\mathbf{R}^3$. (This embedding is suitable for our problem although it is not a Minkowski embedding in general.) Due to our identifications, we have inclusions $\mathbf{Q1} \subset \mathfrak{K}(f) \subset \tilde{\mathfrak{K}} \subset \mathbf{R}^3$, where $\mathbf{Q1} = \{(a, a, a) \mid a \in \mathbf{Q}\}$ is the diagonal embedding of $\mathbf{Q}$, and where the first three sets are fields of which the addition and multiplication are coordinatewise.

The rotation $\sigma : (w_i) \in \mathbf{R}^3 \longmapsto (w_{i+1}) \in \mathbf{R}^3$ of order 3 around the diagonal line $\mathbf{R1}$ induces an automorphism of order 3 of $\tilde{\mathfrak{K}}$. This is verified as follows: Let $\boldsymbol{\theta}$ be a generator of $\mathfrak{K}$ over the diagonal embedding $\mathbf{Q1}$ of $\mathbf{Q}$. Then, the normal closure $\tilde{\mathfrak{K}}$ is isomorphic to $\mathfrak{J} = \mathbf{Q}(\theta_1, \theta_2, \theta_3)$. Let τ be the conjugation of $\mathfrak{J}$ induced by $\tau(\theta_i) = \theta_{i+1}$ $(i = 1, 2, 3)$ and embed $\mathfrak{J}$ in $\mathbf{R}^3$ by $\iota : \gamma \longmapsto (\gamma^{\tau^{i-1}})$. We have $\iota(\gamma^{\tau}) = \iota(\gamma)^{\sigma}$. It now suffices to show $\iota(\mathfrak{J}) = \tilde{\mathfrak{K}}$. Indeed, we have $\boldsymbol{\theta} = \iota(\theta_1) \in \iota(\mathfrak{J})$ and hence $\boldsymbol{\alpha} = \iota(\alpha_1), \boldsymbol{\beta} = \iota(\beta_1) \in \iota(\mathfrak{J})$. Thus, $\boldsymbol{\alpha} \times \boldsymbol{\beta} = \boldsymbol{\alpha}^{\sigma}\boldsymbol{\beta}^{\sigma^2} - \boldsymbol{\alpha}^{\sigma^2}\boldsymbol{\beta}^{\sigma} \in \iota(\mathfrak{J})$. Therefore, $\tilde{\mathfrak{K}} \subset \iota(\mathfrak{J})$. We now get $\tilde{\mathfrak{K}} = \iota(\mathfrak{J})$ by comparing degrees. (One might expect interchange of two

coordinates induces an automorphism of $\tilde{\mathfrak{K}}$. This is however not the case since $(\theta_i) - {}^t(\theta_1, \theta_3, \theta_2) = {}^t(0, \theta_2 - \theta_3, \theta_3 - \theta_2)$ is a zero divisor that cannot lie in any field.)

Let

$$\log : (z_i) \in (\mathbf{R}^\times)^3 \longmapsto (\log |z_i|) \in \mathbf{R}^3$$

be the log-map of Dirichlet. Then, by Dirichlet Unit Theorem, the set

$$\mathfrak{E}(f) = \log(\mathfrak{O}(f)^\times)$$

is a lattice of rank 2 in the plane $\Pi_{\log} = \{(w_i) \in \mathbf{R}^3 \mid w_1 + w_2 + w_3 = 0\}$.

Define the modified log-map by

$$\phi : \boldsymbol{z} \in (\mathbf{R}^\times)^3 \longmapsto \log\left(D^{-1/6}\,\boldsymbol{z}\right) \in \mathbf{R}^3.$$

When $(\boldsymbol{\alpha} \times \boldsymbol{\beta})\boldsymbol{\varepsilon} \in \mathfrak{L}^\natural \cap \mathcal{H}$, then by the identity (4), the image

$$\phi((\boldsymbol{\alpha} \times \boldsymbol{\beta})\boldsymbol{\varepsilon}) = \log \boldsymbol{\varepsilon} + \phi(\boldsymbol{\alpha} \times \boldsymbol{\beta})$$

is contained in the displaced lattice

$$\mathfrak{E}(f) + \phi(\boldsymbol{\alpha} \times \boldsymbol{\beta}) \subset \Pi_{\log}.$$

We normalize the displacement $\phi(\boldsymbol{\alpha} \times \boldsymbol{\beta})$. Let $\boldsymbol{\zeta} \in \mathfrak{O}(f)^\times$ be a unit of norm 1 such that $\log \boldsymbol{\zeta}$ is a closest point of $\mathfrak{E}(f)$ to $\phi(\boldsymbol{\alpha} \times \boldsymbol{\beta})$. We can replace $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ by $\boldsymbol{\zeta}\boldsymbol{\alpha}$ and $\boldsymbol{\zeta}\boldsymbol{\beta}$. The replacement preserves the lattice $\mathfrak{L}^\natural$ (by Definition 3.1) and the order $\mathfrak{O}(f)$ (by Theorem 4.1). The exterior product $\boldsymbol{\alpha} \times \boldsymbol{\beta}$ is replaced with $\boldsymbol{\zeta}^{-1}(\boldsymbol{\alpha} \times \boldsymbol{\beta})$. (This is already verified in the proof of Definition 3.1.) Obviously, $\log(\boldsymbol{\zeta}^{-1}(\boldsymbol{\alpha} \times \boldsymbol{\beta}))$ is in the Voronoï domain of $\mathbf{0}$ on $\Pi_{\log}$ with respect to $\mathfrak{E}(f)$.

Therefore, we assume that $\phi(\boldsymbol{\alpha} \times \boldsymbol{\beta})$ is in the Voronoï domain of $\mathbf{0}$ on $\Pi_{\log}$ with respect to $\mathfrak{E}(f)$. We set

$$\boldsymbol{\delta} = \boldsymbol{\delta}(f) = \boldsymbol{\alpha} \times \boldsymbol{\beta}.$$

Remark. Exterior product has been used in Diophantine equations (see e.g. [10]). However, we use it in a way different from previous investigations. Indeed, we shall use the exterior product as a connection between the representations of 1 by f and a certain continuous curve contained in $\Pi_{\log}$.

Define the rotation σ of order 3 by $\sigma : (w_i) \in \mathbf{R}^3 \longmapsto (w_{i+1}) \in \mathbf{R}^3$. We let the group ring $\mathbf{Z}[\sigma]$ on the logarithmic space $\mathbf{R}^3$ from left, i.e., $(a + b\sigma + c\sigma^2)\boldsymbol{w} = a\boldsymbol{w} + b\sigma\boldsymbol{w} + c\sigma^2\boldsymbol{w}$ for $a, b, c \in \mathbf{Z}$.

Automorphisms of f are reflected in our displaced lattice as follows:

Theorem 4.2. *Assume* $\mathrm{Aut}(f) \neq 1$ *and* $\#\mathcal{R} > 0$. *Then, the cubic field* $\mathfrak{K}(f)$ *is cyclic. Let* $\boldsymbol{z} \in \mathfrak{L}^\natural \cap \mathcal{H}$. *Then,* $\mathbf{Z}\phi(\boldsymbol{z}) + \mathfrak{E}(f)$ *is a lattice of rank* 2 *in* $\Pi_{\log}$ *and the index* $[\mathbf{Z}\phi(\boldsymbol{z}) + \mathfrak{E}(f) : \mathfrak{E}(f)]$ *is either* 1 *or* 3. *The vectors* $(1 - \sigma)\phi(\boldsymbol{z})$ *and* $(1 - \sigma^2)\phi(\boldsymbol{z})$ *belong to the lattice* $\mathfrak{E}(f)$. *The statement holds with* $\phi(\boldsymbol{\delta})$ *in place of* $\phi(\boldsymbol{z})$ *since* $\phi(\boldsymbol{z}) - \phi(\boldsymbol{\delta}) \in \mathfrak{E}(f)$.

Moreover, the lattice $\mathfrak{L}^\natural$ *is invariant under the rotation* σ.

Proof. Let M be a non-trivial automorphism of f and put $(\mu, \nu) = (\alpha, \beta)M$. Then, $\prod_{i=1}^3(\mu_i X + \nu_i Y) = \prod_{i=1}^3(\alpha_i X + \beta_i Y)$. Therefore, we get $\mu_i X + \nu_i Y = \zeta_i\alpha_{\tau(i)}X + \zeta_i\beta_{\tau(i)}Y$ $(i = 1, 2, 3)$ for some permutation τ and some $(\zeta_i) \in \mathbf{R}^3$ such that $N((\zeta_i)) = 1$. Thus, we have

$$(\alpha_i, \beta_i)M = (\mu_i, \nu_i) = (\zeta_i\alpha_{\tau(i)}, \zeta_i\beta_{\tau(i)}) \quad \text{for } i = 1, 2, 3.$$

If τ is trivial, the matrix M has eigenvectors (α_i, β_i) of different directions. (Note: $-\beta_i/\alpha_i$ $(i = 1, 2, 3)$ are distinct roots of $f(X, 1)$.) Since the matrix M of degree 2, this implies that M a unit matrix or its opposite. The former is against our assumption on M. The latter is impossible since $(X, Y) \longmapsto (-X, -Y)$ does not preserve f. This contradiction implies that τ is non-trivial.

Assume $\tau(j) = k \neq j$. Then, we have $\nu_j/\mu_j = \beta_k/\alpha_k$. Hence, we get the inclusion $\mathbf{Q}(\alpha_k, \beta_k) = \mathbf{Q}(\beta_k/\alpha_k) = \mathbf{Q}(\nu_j/\mu_j) \subset \mathbf{Q}(\mu_j, \nu_j) = \mathbf{Q}(\alpha_j, \beta_j)$ of cubic fields. Hence, we get the identity $\mathbf{Q}(\alpha_j, \beta_j) = \mathbf{Q}(\alpha_k, \beta_k)$. Thus, a non-trivial automorphism $\bar{\sigma}$ of $\mathbf{Q}(\alpha_j, \beta_j)$ is induced by $\alpha_j, \beta_j \longmapsto \alpha_k, \beta_k$. Since the degree 3 of $\mathbf{Q}(\alpha_j, \beta_j)$ is a prime number, the fixed field of $\bar{\sigma}$ is $\mathbf{Q}$. Therefore, the order of $\bar{\sigma}$ equals $[\mathbf{Q}(\alpha_j, \beta_j) : \mathbf{Q}] = 3$. In particular, the field $\mathfrak{K}(f) \simeq \mathbf{Q}(\alpha_j, \beta_j)$ is a cyclic cubic field. Applying $\bar{\sigma}$ to the identity $(\alpha_j, \beta_j)M = (\zeta_j\alpha_k, \zeta_j\beta_k)$, we get $(\alpha_k, \beta_k)M = (\zeta_j^{\bar{\sigma}}\alpha_k^{\bar{\sigma}}, \zeta_j^{\bar{\sigma}}\beta_k^{\bar{\sigma}})$. Comparing this with $(\alpha_k, \beta_k)M = (\zeta_k\alpha_{\tau(k)}, \zeta_k\beta_{\tau(k)})$, we get $\beta_{\tau(k)}/\alpha_{\tau(k)} = (\beta_k/\alpha_k)^{\bar{\sigma}} = (\beta_j/\alpha_j)^{\bar{\sigma}^2}$. Since the $\bar{\sigma}$ is an automorphism of order 3 of $\mathbf{Q}(\alpha_j, \beta_j) = \mathbf{Q}(\beta_j/\alpha_j)$, the right hand side differs from β_j/α_j and β_k/α_k. Hence, the permutation τ is a cyclic permutation of order 3. We can assume $\tau(i) =$

$i+1 \pmod 3$ by inverting M if necessary. Now, we have

$$(\boldsymbol{\alpha}, \boldsymbol{\beta})\, M = (\boldsymbol{\mu}, \boldsymbol{\nu}) = (\zeta \boldsymbol{\alpha}^{\sigma}, \zeta \boldsymbol{\beta}^{\sigma})\,.$$

Hence, M^3 preserves $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ and hence is a unit matrix. In particular, the matrix M belongs to $SL_2(\mathbf{Z})$ and hence preserves the exterior product $\boldsymbol{\alpha} \times \boldsymbol{\beta}$. Thus, we get

$$\boldsymbol{\delta \mu} = (\boldsymbol{\mu} \times \boldsymbol{\nu})\, \boldsymbol{\mu} = (\zeta \boldsymbol{\alpha}^{\sigma} \times \zeta \boldsymbol{\beta}^{\sigma})\, \zeta \boldsymbol{\alpha}^{\sigma} = (\boldsymbol{\alpha}^{\sigma} \times \boldsymbol{\beta}^{\sigma})\, \boldsymbol{\alpha}^{\sigma} = (\boldsymbol{\delta \alpha})^{\sigma}$$

by the proof that Definition 3.1 makes sense. The corresponding identity for $\boldsymbol{\nu}$ is verified similarly. Therefore, the map ${}^t(m,n) \in \mathbf{Z}^2 \longmapsto M\,{}^t(m,n) \in \mathbf{Z}^2$ induces the rotation σ of order 3 on Π. This argument implies that the rotation σ preserves $\mathfrak{L}^{\natural}$.

We now consider $\boldsymbol{z}^{\sigma} \in \mathfrak{L}^{\natural}$. Obviously, $N(\boldsymbol{z}^{\sigma}) = N(\boldsymbol{z})$ and hence $\boldsymbol{z}^{\sigma} \in \mathcal{H}$. Thus, Theorem 4.1 implies $\boldsymbol{z} = \boldsymbol{\delta \varepsilon}$ and $\boldsymbol{z}^{\sigma} = \boldsymbol{\delta \varepsilon'}$ for some units $\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon'} \in \mathfrak{O}(f)^{\times}$. Thus, $(1-\sigma)\phi(\boldsymbol{z}) = \log \boldsymbol{\varepsilon} - \log \boldsymbol{\varepsilon'} \in \mathfrak{E}(f)$. Similarly, we see $(1-\sigma^2)\phi(\boldsymbol{z}) \in \mathfrak{E}(f)$. It is now easy to see $3\phi(\boldsymbol{z}) \in \mathfrak{E}(f)$. Hence, $\mathbf{Z}\phi(\boldsymbol{z}) + \mathfrak{E}(f)$ is a lattice of rank 2 in $\Pi_{\log}$ and the index $[\mathbf{Z}\phi(\boldsymbol{z}) + \mathfrak{E}(f) : \mathfrak{E}(f)]$ is 1 or 3. □

We shall continue to use the notation σ both for the rotation $\sigma : (w_i) \in \mathbf{R}^3 \longmapsto (w_{i+1}) \in \mathbf{R}^3$ of order 3 and for the induced automorphism of order 3 of $\tilde{\mathfrak{K}}$.

5. Geometric gap principles

We shall prove gap principles of points on $\phi(\mathfrak{L}^{\natural} \cap \mathcal{H})$. The most important tool is the continuous curve given by

$$(\mathcal{C}): \qquad \pm \exp(u_j) \pm \exp(u_{j+1}) \pm \exp(u_{j+2}) = 0, \quad u_1 + u_2 + u_3 = 0,$$

on the plane $\Pi_{\log}$ (see Figure 1: "The Continuous Curve $\mathcal{C}$"), or more precisely $\mathcal{C} = \mathcal{C}^1 \cup \mathcal{C}^2 \cup \mathcal{C}^3$, where the branch $\mathcal{C}^j$ is defined by

$$(\mathcal{C}^j): \qquad \exp(u_j) - \exp(u_{j+1}) - \exp(u_{j+2}) = 0, \quad u_1 + u_2 + u_3 = 0.$$

The curve $\mathcal{C}$ and its branches are related to $\phi(\mathfrak{L}^{\natural} \cap \mathcal{H})$ as follows:

Proposition 5.1. *The curve $\mathcal{C}$ is the image of $\mathcal{H}$ under ϕ. More precisely, ϕ induces a bijection from the branch $\mathcal{H}^j$ to the branch $\mathcal{C}^j$. Thus, we have*

$$\#\left(\mathfrak{L}^\natural \cap \mathcal{H}\right) = \#\left(\left(\phi\left(\boldsymbol{\delta}\right) + \mathfrak{E}(f)\right) \cap \mathcal{C}\right).$$

PROOF. The first assertion is obvious. The second assertion follows from the fact that the signature of $\boldsymbol{z} \in \mathcal{H}$ can be recovered from the condition $z_i > 0 \Longleftrightarrow u_i = \max\{u_1, u_2, u_3\}$. The last assertion now follows from Theorem 4.1 as we have discussed in §4. □

We set $\mathcal{C}_k = \phi(\mathcal{H}_k)$, for $k = 1, 2, 3$, i.e.,

$$(\mathcal{C}_k): \qquad (u_i) \in \mathcal{C}, \quad u_k \leq u_{k+1}, u_{k+2}$$

and $\mathcal{C}_k^j = \phi(\mathcal{H}_k^j)$, i.e.,

$$\mathcal{C}_k^j = \mathcal{C}^j \cap \mathcal{C}_k.$$

We introduce local coordinates for each piece $\mathcal{C}_k$

$$s = s(\boldsymbol{u}) = \frac{u_{k+1} - u_{k+2}}{\sqrt{2}}, \quad t = t(\boldsymbol{u}) = -\frac{\sqrt{6}\, u_k}{2} \tag{12}$$

so that

$$\boldsymbol{u} = s(\boldsymbol{u})\boldsymbol{e}^{\|}(k) + t(\boldsymbol{u})\boldsymbol{e}^{\perp}(k), \tag{13}$$

where the notation is defined by (7). (The notation is unfortunate and the asymptotic line of $\mathcal{C}_k$ is $\mathbf{R}\boldsymbol{e}^{\perp}(k)$. However, the asymptotic line of $\mathcal{H}$ is $\mathbf{R}\boldsymbol{e}^{\|}(k)$. We cannot make $\mathbf{R}\boldsymbol{e}^{\|}(k)$ the notation for all important asymptotic lines.)

We now study the geometry of $\mathcal{C}_3^1$. Since $\mathcal{C}$ is invariant under permutation of the three coordinates, our results for $\mathcal{C}_3^1$ can be translated to each $\mathcal{C}_k^j$. The curve $\mathcal{C}_3^1$ is described in terms of the coordinates u_i's by

$$(\mathcal{C}_3^1): \ \exp(u_1) - \exp(u_2) - \exp(u_3) = 0, \ u_1 + u_2 + u_3 = 0, \ u_1 \geq u_2 \geq u_3.$$

Since $u_1 = s/\sqrt{2} + t/\sqrt{6}, u_2 = -s/\sqrt{2} + t/\sqrt{6}$ and $u_3 = -2t/\sqrt{6}$, the curve $\mathcal{C}_3^1$ is described by the local equation

$$2\sinh\left(s/\sqrt{2}\right) = \exp\left(-\sqrt{6}\, t/2\right) \tag{14}$$

and the inequality

$$0 \le s \le \sqrt{3}\, t. \tag{15}$$

The local equation has the obvious explicit analytic solution

$$s = \sqrt{2} \sinh^{-1} \left(\exp\left(-\sqrt{6}\, t/2 \right) /2 \right). \tag{16}$$

Since this function is decreasing, we can replace the inequality (15) with

$$t \ge (\log 2)/\sqrt{6}, \tag{17}$$

where the right hand side is the t-coordinate of the boundary point of the curve $\mathcal{C}_3^1$. The other side of the curve is open and asymptotic to the line $s = 0$. The local equation (14) implies a precise estimate of s:

$$0 < s < \sqrt{2} \sinh\left(s/\sqrt{2} \right) = \exp\left(-\sqrt{6}\, t/2 \right) /\sqrt{2}. \tag{18}$$

The explicit analytic solution (16) justifies implicit differentiation of the local equation (14):

$$\frac{ds}{dt} \cosh\left(s/\sqrt{2} \right) = -\frac{\sqrt{3}}{2} \exp\left(-\sqrt{6}\, t/2 \right).$$

Substituting the local equation (14) into the right hand side, we get

$$\frac{ds}{dt} = -\sqrt{3} \tanh\left(s/\sqrt{2} \right) < 0.$$

Further, we get

$$\frac{d^2 s}{dt^2} = -\frac{\sqrt{3}}{\sqrt{2} \cosh^2\left(s/\sqrt{2} \right)} \frac{ds}{dt} > 0.$$

The function

$$-\frac{1}{s}\frac{ds}{dt} = \frac{\sqrt{3} \tanh\left(s/\sqrt{2} \right)}{s}$$

of s is decreasing in $s > 0$ since the derivative

$$\sqrt{3} \left(\sqrt{2} s - \sinh\left(\sqrt{2} s \right) \right) / 2 s^2 \cosh^2\left(s/\sqrt{2} \right)$$

of the right hand side is negative. Therefore, we have

$$\frac{2}{\sqrt{6}\log 2} \leq -\frac{1}{s}\frac{ds}{dt} \leq \frac{\sqrt{6}}{2} \tag{19}$$

in the range

$$0 < s \leq (\log 2)/\sqrt{2}. \tag{20}$$

Note: this range corresponds to the range (17), i.e., all range of t.

We summarize the properties of the curve $\mathcal{C}_k^j$ for later reference.

Theorem 5.2. *We consider the curve $\mathcal{C}_k$ with $k \in \{1,2,3\}$. Let $j \in \{1,2,3\}$ be different from k. Define the local coordinates s and t for $\mathcal{C}_k$ by (12). In terms of s and t, the part $\mathcal{C}_k^j = \mathcal{C}^j \cap \mathcal{C}_k$ of $\mathcal{C}$ is given by $s = \pm g(t)$, $t \geq (\log 2)/\sqrt{6}$, where the sign is positive or negative according as $j \equiv k-1 \pmod 3$ or $j \equiv k+1 \pmod 3$, and g is a function independent of k with the following properties:*

(i) *g is differentiable and convex;*

(ii) *$g((\log 2)/\sqrt{6} = (\log 2)/\sqrt{2})$ and g monotonously decreases to 0 as t tends to $+\infty$;*

(iii) *the convergence of $g(t)$ to 0 is quantitatively expressed by*

$$0 < |s| = g(t) < \exp\left(-\sqrt{6}\,t/2\right)/\sqrt{2}; \tag{21}$$

(iv) *for the logarithmic derivative of f, we have*

$$\frac{2}{\sqrt{6}\log 2} \leq -\frac{\dot{g}(t)}{g(t)} \leq \frac{\sqrt{6}}{2}; \tag{22}$$

(v) *the ratio $g(t)/t$ assumes its maximum $\sqrt{3}$ at $t = (\log 2)/\sqrt{6}$ and decreases as t increases;*

(vi) *put $r(t) = (t^2 + g(t)^2)^{1/2}$. Then, $t \longmapsto r(t)/t$ is a decreasing function. In particular, we have $t < r(t) < 1.01t$ if $r(t) \geq 1.2$ or $t \geq 1.2$. Further, we have $t < r(t) < 1 + 10^{-10}$ if $t \geq 8$.*

(We use the notation $\dot{g}$ for derivative $dg(t)/dt$, since we are using the notation x', y' for different representations of 1.)

We combine geometry of numbers on $\mathfrak{L}^\natural$ with geometric knowledge about $\mathcal{C}$ to infer two gap principles. The following gap principle is useful when $f(X,Y)$ is reduced but not monic.

Lemma 5.3. *Assume $\boldsymbol{\delta\alpha}$ and $\boldsymbol{\delta\beta}$ form a reduced basis of $\mathfrak{L}^\natural$. Let $\boldsymbol{z} \in \mathfrak{L}^\natural \cap \mathcal{H}$ satisfy $\|\boldsymbol{z}\| \geq \|\boldsymbol{\delta\beta}\|$. Then, we have*

$$\|\phi(\boldsymbol{z})\| > t(\phi(\boldsymbol{z})) \geq \frac{1}{2\sqrt{6}} \log \frac{D}{27}.$$

If further $t \geq 3$, we have

$$\|\phi(\boldsymbol{z})\| > t(\phi(\boldsymbol{z})) \geq \frac{1}{2\sqrt{6}} \log(0.4D).$$

PROOF. By Proposition 3.2, the assumption implies $\sqrt{3D} \leq \|\boldsymbol{\delta\beta}\|^2 \leq \|\boldsymbol{z}\|^2$. Let $s = s(\phi(\boldsymbol{z}))$ and $t = t(\phi(\boldsymbol{z}))$. Then, we have

$$\begin{aligned} D^{-1/3}\|\boldsymbol{z}\|^2 &= 2e^{2t/\sqrt{6}} \cosh\left(\sqrt{2}s\right) + e^{-4t/\sqrt{6}} \\ &= 4e^{2t/\sqrt{6}} \sinh^2\left(s/\sqrt{2}\right) + 2e^{2t/\sqrt{6}} + e^{-4t/\sqrt{6}} \\ &= 2e^{2t/\sqrt{6}} + 2e^{-4t/\sqrt{6}} \leq 3e^{2t/\sqrt{6}}, \end{aligned}$$

where the local equation (14) is used to prove the last equality and (17) is used to prove the last inequality. The assertions are now immediate. □

The next gap principle, which we call a linear gap principle, is a translation of Lemma 3.4. It guarantees a space of constant size between points of $\phi(\mathfrak{L}^\natural \cap \mathcal{H})$ that is necessary for proper use of Theorem 5.5. (For a weaker but easier alternative, see the end of this section.)

Lemma 5.4. *Assume distinct points $\boldsymbol{z}$ and $\boldsymbol{z}'$ of $\mathfrak{L}^\natural$ lie on the same piece $\mathcal{H}_k$ of the curve $\mathcal{H}$. Assume $t(\phi(\boldsymbol{z}')) \geq t(\phi(\boldsymbol{z}))$. Then, we have*

$$t(\phi(\boldsymbol{z}')) \geq 2t(\phi(\boldsymbol{z})) + \frac{1}{\sqrt{6}} \log D - \frac{3}{\sqrt{6}} \log \frac{9}{2}.$$

The last term can be replaced with -10^{-3}if $t \geq 4.8$.

PROOF. Put $p = p(\boldsymbol{z})$, $p' = p(\boldsymbol{z}')$, $\boldsymbol{u} = \phi(\boldsymbol{z}))$, $\boldsymbol{u}' = \phi(\boldsymbol{z}')$, $s = s(\boldsymbol{u})$, $s' = s(\boldsymbol{u}')$, $t = t(\boldsymbol{u})$ and $t' = t(\boldsymbol{u}')$. Then, we have

$$|p| = \sqrt{2}D^{1/6}\, e^{t/\sqrt{6}} \cosh\left(s/\sqrt{2}\right), \quad |q| = (\sqrt{6}/2)D^{1/6}e^{-2t/\sqrt{6}}.$$

Substituting these in (9), we get

$$D^{1/6} \le e^{(t'-2t)/\sqrt{6}} \cosh(s'/\sqrt{2}\,) + e^{(t-2t')/\sqrt{6}} \cosh(s/\sqrt{2}\,).$$

By expressing cosh in terms of sinh and then substituting (14), we get

$$D^{1/6} \le e^{(t'-2t)/\sqrt{6}} \left(1 + e^{-3(t'-t)/\sqrt{6}}\right) \sqrt{1 + e^{-\sqrt{6}\,t}/4}. \tag{23}$$

Substituting (17), we get the first assertion. It implies $t'-t \ge t-0.26$ since $D \ge 49$. Substituting this in (23), we can improve the last term to -10^{-2}. Hence, we get $t' - t \ge t + 1.578$. The assertion follows after repeating the same argument. □

We now state two exponential gap principles of geometric nature, which constitute the key step of this paper.

Theorem 5.5. *Let $\mathfrak{M}$ be a lattice of rank 2 in $\Pi_{\log}$. Assume distinct points $\boldsymbol{u}$, $\boldsymbol{u}'$ and $\boldsymbol{u}''$ of $\phi(\boldsymbol{\delta}) + \mathfrak{M}$ lie on the same piece $\mathcal{C}_k$ of the curve $\mathcal{C}$. Set $t = t(\boldsymbol{u})$, $t' = t(\boldsymbol{u}')$ and $t'' = t(\boldsymbol{u}'')$. Assume $t'' - \sqrt{6}\log 2 \ge t' \ge t$. Then, we have*

$$t'' \ge \frac{\sqrt{2}\,\mathrm{disc}(\mathfrak{M}) \exp\left(\sqrt{6}\,t/2\right)}{1 + \exp\left(-2(t'-t)/\sqrt{6}\log 2\right)}.$$

Theorem 5.6. *Let $\mathfrak{M}$ be a lattice of rank 2 in $\Pi_{\log}$. Assume distinct points $\boldsymbol{u}$, and $\boldsymbol{u}'$ of $\mathfrak{M}$ lie on the same piece $\mathcal{C}_k$ of the curve $\mathcal{C}$. Set $t = t(\boldsymbol{u})$ and $t' = t(\boldsymbol{u}')$. Assume $t' \ge t$. Then, we have*

$$t' \ge \frac{\sqrt{2}\,\mathrm{disc}(\mathfrak{M}) \exp\left(\sqrt{6}\,t/2\right)}{1 + \exp\left(-2(t'-t)/\sqrt{6}\log 2\right)}.$$

These gap principles can be rewritten in terms of $r = \|\boldsymbol{u}\|$, $r' = \|\boldsymbol{u}'\|$ and $r'' = \|\boldsymbol{u}''\|$ since $r'' \ge t''$, $r' \ge t'$ and $t \ge r/2$ (or $t \ge r/1.01$ if $r \ge 1.2$).

PROOF of Theorem 5.5. By symmetry, we assume $k = 3$ and $\boldsymbol{u} \in \mathcal{C}_3^1$.

We shall later show that the three points are not collinear. In particular, the three points $\boldsymbol{u}$, $\boldsymbol{u}'$ and $\boldsymbol{u}''$ form a triangle of positive area, which we denote by Δ. Since $\boldsymbol{u}' - \boldsymbol{u}$ and $\boldsymbol{u}'' - \boldsymbol{u}$ are vectors in $\mathfrak{M}$, this implies the lattice constraint

$$\Delta \ge \mathrm{disc}(\mathfrak{M})/2 \tag{24}$$

on the area Δ.

We express everything in terms of the (t, s)-coordinates of $\mathcal{C}_3$ defined by (12). Then, with respect to these coordinates, $\mathcal{C}_3^1$ may be viewed as the graph of g and $\mathcal{C}_3^2$ as the graph of $-g$, where g is the function of Theorem 5.2. Thus, if we put $s = s(\boldsymbol{u})$, $s' = s(\boldsymbol{u}')$ and $s'' = s(\boldsymbol{u}'')$, we have $s = \pm g(t)$, $s' = \pm g(t')$ and $s'' = \pm g(t'')$, where the signs depend on whether the points lie on $\mathcal{C}_3^1$ or $\mathcal{C}_3^2$. Recall that g is a decreasing function assuming only positive values. Therefore, the triangle formed by $\boldsymbol{u}$, $\boldsymbol{u}'$ and $\boldsymbol{u}''$ is contained in a rectangle of sides $t'' - t$ and $|s| + |s'|$ (see Figure 2: "Area Estimate"). Hence, we have:

$$\Delta \leq (t'' - t)(|s| + |s'|)/2.$$

Here, we have $|s'| < |s| e^{-2(t'-t)/\sqrt{6}\log 2}$ by integrating (19). Therefore, the inequality (21) now implies the area estimate

$$\Delta \leq t'' \exp\left(-\sqrt{6}\, t/2\right)\left(1 + e^{-2(t'-t)/\sqrt{6}\log 2}\right)/2\sqrt{2}. \tag{25}$$

The lower bound for t'' is an immediate consequence of the lattice constraint (24) and the area estimate (25).

We now show that the three points $\boldsymbol{u}$, $\boldsymbol{u}'$ and $\boldsymbol{u}''$ are not collinear. Suppose the contrary, i.e., they lie on the same line ℓ and investigate its inclination against the t-axis.

First of all, the line ℓ cannot be perpendicular to the t-axis since a perpendicular line intersect $\mathcal{C}_k^j$ at only one point for each $j = 1, 2$. Hence its inclination with respect the t-axis makes sense. Thus, inclination of ℓ is equals to both $G = (s' - s)/(t' - t)$ and $G' = (s'' - s')/(t'' - t)$.

We now show that $G = G'$ is impossible. Without loss of generality, assume $s > 0$. There are four possibilities of the combination of signs of s' and s''.

When $s' < 0$ and $s'' < 0$, the fact g monotonously decreases implies $G' > 0 > G$.

When $s' < 0$ and $s'' > 0$, we obviously have $G' > 0 > G$.

When $s' > 0$ and $s'' < 0$, we have $0 < -G' < 4s'/\sqrt{6}\log 2 < -\dot{g}(t')$ by the inequality (22) of Theorem 5.2. On the other hand, the convexity of g (Theorem 5.2 again) implies $\dot{g}(t') > G$. Therefore, we get $G' > \dot{g}(t') > G$. (See Figure 3: "Proper Triangle across Axis".)

When $s' > 0$ and $s'' > 0$, the convexity of g implies $G' > \dot{g}(t') > G$. (See Figure 4: "Proper Triangle in One Side".)

We now arrive at the desired contradiction. □

PROOF of Theorem 5.6. The proof is similar to the proof of Theorem 5.5. We easily verify that $\mathbf{0}$, $\boldsymbol{u}$ and $\boldsymbol{u}'$ are not collinear. Indeed, the function g is a decreasing function of positive value (Theorem 5.2). Hence, two distinct points on the union of the graphs of g and $-g$ cannot be on the same line passing through the origin.

The rest is the same as the proof of Theorem 5.5. □

Before closing this section, we give two lemmata which are also shown by using the geometry of continuous curves. The following lemma is a refinement of Pohst's lower bound [20]:

Lemma 5.7. *Let $\boldsymbol{\zeta}$ be a totally real cubic unit of discriminant $D(\boldsymbol{\zeta})$. Then, we have*

$$\|\log \boldsymbol{\zeta}\| \geq \frac{1}{2\sqrt{2}} \log \frac{D(\boldsymbol{\zeta})}{P(D(\boldsymbol{\zeta}))},$$

where $P(D(\boldsymbol{\zeta})) = 1.01$ if $D(\boldsymbol{\zeta}) \geq 10^{12}$ or $P(D(\boldsymbol{\zeta})) = 4$ otherwise.

Further, we have for a totally real cubic field $\mathfrak{K}$,

$$\operatorname{disc}(\mathfrak{E}(\mathfrak{K})) \geq \frac{\sqrt{3}}{2} \log^2 \frac{D(\mathfrak{K})}{P(D(\mathfrak{K}))} \geq 0.6.$$

PROOF. Without loss of generality, we assume $|\zeta_1| \geq 1 \geq |\zeta_2| \geq |\zeta_3|$. Put $l = \|\log \boldsymbol{\zeta}\|$. Then, the point $(U_1, U_2, U_3) = \log \boldsymbol{\zeta}$ is on the arc $\mathcal{A}$ of the circle $U_1^2 + U_2^2 + U_3^2 = l^2$ (on the plane $U_1 + U_2 + U_3 = 0$) cut by the sector $U_3 \leq U_2 \leq 0$.

Then, we have $\sqrt{D(\boldsymbol{\zeta})} = |\zeta_1^2\zeta_2(1 - \zeta_3/\zeta_1)(1 - \zeta_2/\zeta_1)(1 - \zeta_1/\zeta_2)| \leq |\zeta_1^2\zeta_2|(1-|\zeta_3/\zeta_1|)(1+|\zeta_2/\zeta_1|)(1+|\zeta_3/\zeta_2|)$ since at most two quotients among ζ_2/ζ_1, ζ_3/ζ_2 and ζ_3/ζ_1 are negative. Hence, the inequality

$$W(U_1, U_2, U_3) \geq \sqrt{D(\boldsymbol{\zeta})}$$

holds at some point on the arc $\mathcal{A}$, where

$$W = W(U_1, U_2, U_3) = e^{2U_1+U_2}(1 - e^{U_3-U_1})(1 + e^{U_2-U_1})(1 + e^{U_3-U_2}).$$

We make a change of variables $T = -\sqrt{3}\,U_2/2$ and $S = (U_1 - U_3)/2$, i.e., $U_1 = S + T/\sqrt{3}$, $U_2 = -2T/\sqrt{3}$ and $U_3 = -S + T/\sqrt{3}$. The equation

of the arc $\mathcal{A}$ becomes $S^2 + T^2 = l^2/2$ with $0 \le \sqrt{3}T \le S$. We have

$$W = W(S,T) = 4\left(\cosh(S) + \cosh(\sqrt{3}T)\right)\sinh(S).$$

Substituting $0 \le \sqrt{3}T \le S \le l/\sqrt{2}$, we get $W \le 4\cosh(2S) < 2e^{\sqrt{2}l}$. The inequality with $P(D(\boldsymbol{\zeta})) = 4$ now follows immediately.

We now assume $D(\boldsymbol{\zeta}) \ge 10^{12}$. Then, we have $l \ge 9.27$ and hence $S \ge 4\log 4$ by the inequality just established. Put $T(S) = \sqrt{l^2/2 - S^2}$. Then,

$$\frac{d}{dS}\left(\cosh(S) + \cosh\left(\sqrt{3}T(S)\right)\right) = S\left(\frac{\sinh(S)}{S} - 3\frac{\sinh(\sqrt{3}T(S))}{\sqrt{3}T(S)}\right).$$

When $\sqrt{3}T = S - \log 4$, we have $\sinh(S) > 4\sinh(\sqrt{3}T)$, $S \le (4/3)\sqrt{3}T$ and hence $\sinh(S)/S > 3\sinh(\sqrt{3}T)/\sqrt{3}T$. The last inequality is also valid for T in the range $0 \le \sqrt{3}T \le S - \log 4$ since $\sinh(\sqrt{3}T)/\sqrt{3}T$ is an increasing function of T. Hence, we get $\cosh(S) + \cosh\left(\sqrt{3}T(S)\right) \le \cosh\left(l/\sqrt{2}\right) + 1$ and hence $W \le W(l/\sqrt{2}, 0)$.

When $S - \sqrt{3}T(S) < \log 4$, we have $\sqrt{3}T \le S \le (\sqrt{6}/4)l + (1/4)\log 4$ and hence $W(S, T(S)) \le 4\cosh((\sqrt{6}l + \log 4)/2)$. Thus, $W(S, T(S)) \le 0.7W(l/\sqrt{2}, 0)$.

We now have

$$W \le W(l/\sqrt{2}, 0) = \left(1 + e^{-l/\sqrt{2}}\right)^2\left(1 - e^{-2l/\sqrt{2}}\right)e^{\sqrt{2}l} \le 1.004e^{\sqrt{2}l}.$$

Therefore, $P(D(\boldsymbol{\zeta}))$ can be improved to 1.01. This proves the first assertion.

The second assertion follows from the first assertion on noting (2). □

An immediate application of this lower bound is the following initial gap principle:

Lemma 5.8. *Let $\boldsymbol{z} \in \mathfrak{L}^{\natural} \cap \mathcal{H}$. If $\boldsymbol{z} \ne \boldsymbol{\delta}$, we have*

$$\|\phi(\boldsymbol{z})\| \ge \frac{1}{4\sqrt{2}}\log\frac{D}{P(D)}.$$

If $D \ge 10^{13}$ and there is some point $\boldsymbol{z}^{\dagger} \in \mathfrak{L}^{\natural} \cap \mathcal{H}$ ($\boldsymbol{z}^{\dagger} \ne \boldsymbol{z}$) satisfying

$\|\phi(\boldsymbol{z})\| \geq \|\phi(\boldsymbol{z}^\dagger)\|$, *we have*

$$\|\phi(\boldsymbol{z})\| \geq \frac{1}{2\sqrt{6}} \log \frac{D}{1.02}.$$

If $\mathrm{Aut}(f) \neq 1$, *we have*

$$\|\phi(\boldsymbol{z})\| \geq \frac{1}{2\sqrt{6}} \log \frac{D}{P(D)}.$$

Note: the conditions of the three inequalities are independent.

PROOF. By triangle inequality and the choice of $\boldsymbol{\delta}$, we have

$$\|\log(\boldsymbol{\varepsilon})\| \leq \|\log(\boldsymbol{\varepsilon}) - \phi(\boldsymbol{z})\| + \|\phi(\boldsymbol{z})\| = \|\phi(\boldsymbol{\delta})\| + \|\phi(\boldsymbol{z})\| \leq 2\|\phi(\boldsymbol{z})\|,$$

where $\boldsymbol{\varepsilon} = \boldsymbol{z}\boldsymbol{\delta}^{-1} \in \mathfrak{O}(f)^\times$. Substituting Lemma 5.7 and recalling $D(\boldsymbol{\varepsilon}) \geq D(\mathfrak{O}(f)) = D$

(Theorem 4.1), we establish the first lower bound.

Let $\mathfrak{D}$ denote the closed disc of radius $r = \|\phi(\boldsymbol{z})\|$ centered at the origin. Two points of the part of $\mathcal{C}$ intersecting with $\mathfrak{D}$ which have largest between each other lie on the intersection of $\mathcal{C}$ with the boundary of $\mathfrak{D}$. Indeed, we have seen in Theorem 5.2 that every branch $\mathcal{C}^j$ is convex and each piece $\mathcal{C}^j_k$ is described by an explicit function $t \longmapsto s$. Thus, the convex hull of $\mathcal{C} \cap \mathfrak{D}$ is a hexagon. Hence, the two points of f $\mathcal{C} \cap \mathfrak{D}$ having largest distance between each other are among the vertices of this hexagon.

By using the rotational symmetry of order 3, and the estimate (21), we see that the distance between those points is less than $\sqrt{3}r + \sqrt{2}\exp\left(-\sqrt{6}\,r/2.02\right)$. Therefore, $\|\phi(\boldsymbol{z}) - \phi(\boldsymbol{z}')\|$ is smaller than this quantity. Hence, Lemma 5.7 implies

$$\sqrt{3}r \geq (1/2\sqrt{2}\,)\log(D/P(D)) - \sqrt{2}e^{-\sqrt{6}\,r/2.02}.$$

(Note: $\phi(\boldsymbol{z}) - \phi(\boldsymbol{z}') \in \mathfrak{E}(f)$.) By estimating r by the first assertion, we see that the second term is less than 0.0024 when $D \geq 10^{13}$. The second assertion follows immediately.

Now, assume $\mathrm{Aut}(f) \neq 1$. Then, Theorem 4.2 implies $(1-\sigma)\phi(\boldsymbol{z}) \in \mathfrak{E}(f)$. Since the map $1-\sigma$ expands the lengths of the vectors in $\Pi_{\log}$ by a factor of $\sqrt{3}$, Lemma 5.7 implies the third assertion provided $\phi(\boldsymbol{z}) \neq 0$. This condition is indeed guaranteed since $\boldsymbol{z} \in \Pi$. □

Remark. Lemma 5.7 also implies an alternative for Lemma 5.4: $t(\phi(\boldsymbol{z}'))-t(\phi(\boldsymbol{z})) \geq \sqrt{(1/8)\log^2(D/4) - 2\log^2 2}$. Indeed, this follows from Lemma 5.7, Pythagoras' theorem and the inequalities $|s'|, |s| \leq \sqrt{2}\log 2$ (Theorem 5.2).

6. Arithmetic invariants

We can write a coordinate of $(1-\sigma)\phi(\boldsymbol{\delta\varepsilon})$ as a linear form in three logarithms since $(\boldsymbol{\delta\varepsilon})^{1-\sigma}$ belongs to the division group of the multiplicative group generated by $\boldsymbol{\delta}^{1-\sigma}$ and $(\mathfrak{O}(f)^\times)^{1-\sigma}$. In this section, we investigate the invariants of the division group, hich will be substituted in Baker theory for estimating the geometric size $\|\phi(\boldsymbol{\delta\varepsilon})\|$ in §7.

The reason for studying the division group is the dependence of Matveev's lower bound [15] on the "Kummer condition". Indeed, Matveev's result gives a lower bound for $|l\log\tilde{\delta}_i + m\log\tilde{\xi}_i + n\log\tilde{\eta}_i|$, where the group generated by $\tilde{\boldsymbol{\delta}}$, $\tilde{\boldsymbol{\xi}}$ and $\tilde{\boldsymbol{\eta}}$ essentially coincides with its division group in a suitable field.

The goal of this section is to estimate the heights of suitable generators of the division group in terms of geometric sizes, i.e., L^2-norms. This establishes a connection between geometric gap principles of §5 and the lower bounds for linear forms in logarithms stated in §7 which are written in terms of heights, i.e., normalized L^1-norms.

By height, we mean the absolute Weil height

$$h(\tilde{\boldsymbol{\gamma}}) = \frac{1}{2[\tilde{\mathfrak{K}} : \mathbf{Q}]} \sum_{v \in M(\tilde{\mathfrak{K}})} |\log \|\tilde{\boldsymbol{\gamma}}\|_v|$$

of an arbitrary algebraic number $\tilde{\boldsymbol{\gamma}} \in \tilde{\mathfrak{K}}^\times$, where v runs through the set $M(\tilde{\mathfrak{K}})$ of all places of $\tilde{\mathfrak{K}}$ and $\|\cdot\|_v$ denotes the normalized v-adic valuation with respect to the product formula. The normalization of each v-adic valuations is chosen so that the restriction of $\|\cdot\|$ to $\mathbf{Q}$ equals $\|\cdot\|_p^{[(\tilde{\mathfrak{K}})_v:\mathbf{Q}_p]}$, where $\|\cdot\|_p$ denotes the standard p-adic valuation on $\mathbf{Q}$.

We prepare some notation for the unit group of $\mathfrak{K}$ since they are more closely related with the division group than the group $\mathfrak{O}(f)^\times$ is. We denote the maximal order of $\mathfrak{K}$ by $\mathfrak{O}(\mathfrak{K})$ and the lattice $\log(\mathfrak{O}(\mathfrak{K})^\times)$ on $\Pi_{\log}$ by $\mathfrak{E}(\mathfrak{K})$. Of course, the unit group of $\mathfrak{K}$ is $\mathfrak{O}(\mathfrak{K})^\times$.

We consider the multiplicative group $\Gamma(\boldsymbol{\omega})$ generated by $\mathfrak{O}(\mathfrak{K})^\times$ and a given element $\boldsymbol{\omega}$ of $\tilde{\mathfrak{K}}$. We denoted by $\tilde{\Gamma}(\boldsymbol{\omega})$ the division group of $\Gamma(\boldsymbol{\omega})^{1-\sigma}$ in $\tilde{\mathfrak{K}}$. Then, for $\boldsymbol{\varepsilon} \in \mathfrak{O}(\mathfrak{K})^\times$, $(1-\sigma)\phi(\boldsymbol{\delta\varepsilon})$ belongs to the module $\log \Gamma(\boldsymbol{\delta})^{1-\sigma}$ and hence to $\log \tilde{\Gamma}(\boldsymbol{\delta})$. The division group $\tilde{\Gamma}(\boldsymbol{\delta})$ is the group that will be used. The reason (apart from the use of $\tilde{\Gamma}(\mathbf{1})$) for allowing other value of $\boldsymbol{\omega}$ than $\boldsymbol{\delta}$ is the following identity:

$$\tilde{\Gamma}(\boldsymbol{\delta}) = \tilde{\Gamma}(\boldsymbol{\delta}/\sqrt{D}). \tag{26}$$

This identity is important since the right hand side is a division group of a subgroup of $(\mathfrak{K}^\times)^{1-\sigma}$. (Note: $\boldsymbol{\delta}/\sqrt{D} \in \mathfrak{K}^\times$.)

Its consequence is the following:

Lemma 6.1. *If $\tilde{\boldsymbol{\gamma}} \in \tilde{\Gamma}(\boldsymbol{\delta})$ is a unit of the maximal order of $\tilde{\mathfrak{K}}$, then $\tilde{\boldsymbol{\gamma}}^3$ or $-\tilde{\boldsymbol{\gamma}}^3$ belongs to $(\mathfrak{O}(\mathfrak{K})^\times)^{1-\sigma}$.*

PROOF. When $\tilde{\mathfrak{K}} = \mathfrak{K}$, we have $\tilde{\boldsymbol{\gamma}}^{(1-\sigma^2)(1-\sigma)} = \tilde{\boldsymbol{\gamma}}^{3-(1+\sigma+\sigma^2)} = \pm\tilde{\boldsymbol{\gamma}}^3$. This implies the assertion since $\tilde{\boldsymbol{\gamma}}^{1-\sigma^2} \in \mathfrak{O}(\mathfrak{K})^\times$.

We now assume $\tilde{\mathfrak{K}} \neq \mathfrak{K}$. By (26), there is a pair of a positive integer m and an element $\boldsymbol{\gamma}$ of $\mathfrak{K}^\times$ such that $\tilde{\boldsymbol{\gamma}}^m = \boldsymbol{\gamma}^{1-\sigma}$. Let τ be a non-trivial conjugation of $\tilde{\mathfrak{K}}/\mathfrak{K}$. Then, we have $\tilde{\boldsymbol{\gamma}}^{m(1+\tau)} = \boldsymbol{\gamma}^{1+\tau-\sigma-\sigma\tau} = \boldsymbol{\gamma}^{1+\tau-\sigma-\tau\sigma^2}$ since the Galois group of $\tilde{\mathfrak{K}}$ is a dihedral group. Noting that τ preserves the element $\boldsymbol{\gamma}$ of $\mathfrak{K}$, we get $\tilde{\boldsymbol{\gamma}}^{m(1+\tau)} = \boldsymbol{\gamma}^{3-(1+\sigma+\sigma^2)} = \boldsymbol{\gamma}^3 N_{\mathfrak{K}/\mathbf{Q}}\boldsymbol{\gamma}^{-1}$. Thus, $\tilde{\boldsymbol{\gamma}}^{m(1+\tau)(1-\sigma)} = \boldsymbol{\gamma}^{3(1-\sigma)} = \tilde{\boldsymbol{\gamma}}^{3m}$. Since $\mathfrak{K}$ is totally real, this implies $\tilde{\boldsymbol{\gamma}}^{(1+\tau)(1-\sigma)} = \pm\boldsymbol{\gamma}^3$. On the other hand, $\tilde{\boldsymbol{\gamma}}^{1+\tau} \in \mathfrak{O}(\mathfrak{K})^\times$ since $\tilde{\boldsymbol{\gamma}} \in \mathfrak{O}(\tilde{\mathfrak{K}})^\times$. The lemma is now established. □

The division group $\tilde{\Gamma}(\boldsymbol{\delta})$ always contains $\tilde{\Gamma}(\mathbf{1})$ and it coincides with $\tilde{\Gamma}(\mathbf{1})$ in some important cases. Let $\tilde{\mathfrak{G}} = \log \tilde{\Gamma}(\mathbf{1})$. Then, Lemma 6.1 implies

$$3\tilde{\mathfrak{G}} \subset (1-\sigma)\mathfrak{E}(\mathfrak{K}) \subset \tilde{\mathfrak{G}}. \tag{27}$$

In particular, we get the isomorphism

$$\tilde{\Gamma}(\mathbf{1}) \simeq \mathbf{Z}^2 \oplus (\mathbf{Z}/2\mathbf{Z}) \tag{28}$$

of groups. The following lemma discriminates when $\tilde{\Gamma}(\boldsymbol{\delta}) = \tilde{\Gamma}(\mathbf{1})$:

Lemma 6.2. *The following five conditions are equivalent:*

(i) $\boldsymbol{\delta}^{1-\sigma}$ *is a unit;*

(ii) $\tilde{\Gamma}(\boldsymbol{\delta}) = \tilde{\Gamma}(\mathbf{1})$;

(iii) $\tilde{\Gamma}(\boldsymbol{\delta}) \simeq \mathbf{Z}^2 \oplus (\mathbf{Z}/2\mathbf{Z})$;

(iv) $\log \tilde{\Gamma}(\boldsymbol{\delta})$ *is a lattice of rank* 2 *in* $\Pi_{\log}$;

(v) $\mathbf{Z}\phi(\boldsymbol{\delta}) + \mathfrak{E}(\mathfrak{K})$ *is a lattice of rank* 2 *in* $\Pi_{\log}$.

The index $[\mathbf{Z}\phi(\boldsymbol{\delta}) + \mathfrak{E}(\mathfrak{K}) : \mathfrak{E}(\mathfrak{K})]$ *is either* 1 *or* 3 *when these conditions hold.*

PROOF. By Lemma 6.1, the condition (i) implies (ii). By isomorphism (28), condition (ii) implies (iii). By isomorphism (28) again, condition (iii) implies finiteness of the index $[\tilde{\Gamma}(\boldsymbol{\delta}) : \tilde{\Gamma}(\mathbf{1})]$. Thus, it implies (iv) by (27). Since $1-\sigma$ is a similarity on $\Pi_{\log}$, $\mathbf{Z}\phi(\boldsymbol{\delta})+\mathfrak{E}(\mathfrak{K})$ is a lattice if the submodule $(1-\sigma)(\mathbf{Z}\phi(\boldsymbol{\delta}) + \mathfrak{E}(\mathfrak{K}))$ of $\log \tilde{\Gamma}(\boldsymbol{\delta})$ is a lattice. Hence, (iv) implies (v).

We now assume (v) and prove (i) and the last assertion on the index. Let $m = [\mathbf{Z}\phi(\boldsymbol{\delta}) + \mathfrak{E}(\mathfrak{K}) : \mathfrak{E}(\mathfrak{K})] < \infty$. Then, there exist $\boldsymbol{\zeta} \in \mathfrak{O}(\mathfrak{K})^{\times}$ such that $m\phi(\boldsymbol{\delta}) = \log \boldsymbol{\zeta}$. Thus, $m \log \boldsymbol{\delta}^{1-\sigma} = (1-\sigma)m\phi(\boldsymbol{\delta}) = \log \boldsymbol{\zeta}^{1-\sigma}$. Therefore, $(\boldsymbol{\delta}^{1-\sigma})^{2m} = (\boldsymbol{\zeta}^2)^{1-\sigma}$ is a unit and so is $\boldsymbol{\delta}^{1-\sigma}$. By Lemma 6.1, this implies $\boldsymbol{\delta}^{3(1-\sigma)} \in (\mathfrak{O}(\mathfrak{K})^{\times})^{1-\sigma}$, or equivalently $3(1-\sigma)\phi(\boldsymbol{\delta}) \in (1-\sigma)\mathfrak{E}(\mathfrak{K})$. Since $1-\sigma$ is a similarity on $\Pi_{\log}$, we now get $3\phi(\boldsymbol{\delta}) \in \mathfrak{E}(\mathfrak{K})$. □

Lemma 6.3. *Assume* $\boldsymbol{\delta}^{1-\sigma}$ *is a unit. Let* $\boldsymbol{\xi}$ *and* $\boldsymbol{\eta}$ *be independent units of* $\mathfrak{K}$ *such that* $\|\log \boldsymbol{\xi}\| \leq \|\log \boldsymbol{\eta}\|$. *Then, there is a pair* $\tilde{\boldsymbol{\xi}}$ *and* $\tilde{\boldsymbol{\eta}}$ *of elements of* $\tilde{\Gamma}(\boldsymbol{\delta}) = \tilde{\Gamma}(\mathbf{1})$ *such that* $\log \tilde{\boldsymbol{\xi}}$ *and* $\log \tilde{\boldsymbol{\eta}}$ *form a reduced basis for* $\tilde{\mathfrak{G}}$. *We have*

$$\frac{1}{3}|\log \tilde{\boldsymbol{\xi}}|_{\infty} = h(\tilde{\boldsymbol{\xi}}) \leq \frac{\sqrt{6}}{9}\|\log \tilde{\boldsymbol{\xi}}\| \leq \frac{\sqrt{2}}{3}\|\log \boldsymbol{\xi}\|;$$

$$\frac{1}{3}|\log \tilde{\boldsymbol{\eta}}|_{\infty} = h(\tilde{\boldsymbol{\eta}}) \leq \frac{\sqrt{6}}{9}\|\log \tilde{\boldsymbol{\eta}}\| \leq \frac{\sqrt{2}}{3}\|\log \boldsymbol{\eta}\|.$$

Moreover, we have

$$\|m \log \tilde{\boldsymbol{\xi}}\| + \|n \log \tilde{\boldsymbol{\eta}}\| \leq 2\|m \log \tilde{\boldsymbol{\xi}} + n \log \tilde{\boldsymbol{\eta}}\|$$

for an arbitrary pair of integers m *and* n.

PROOF. By Lemma 6.2, $\log \tilde{\Gamma}(\boldsymbol{\delta})$ is a lattice of rank 2 in $\Pi_{\log}$. Choose elements $\tilde{\boldsymbol{\xi}}$ and $\tilde{\boldsymbol{\eta}}$ of $\tilde{\Gamma}(\boldsymbol{\delta})$ so that $\log \tilde{\boldsymbol{\xi}}$ and $\log \tilde{\boldsymbol{\eta}}$ form a reduced basis of

$\log \tilde{\Gamma}(\boldsymbol{\delta})$. By Lemma 6.2, we have $\tilde{\Gamma}(\boldsymbol{\delta}) = \tilde{\Gamma}(\mathbf{1})$ and hence $\log \tilde{\Gamma}(\boldsymbol{\delta}) = \tilde{\mathfrak{G}}$. Therefore, $\log \tilde{\boldsymbol{\xi}}$ and $\log \tilde{\boldsymbol{\eta}}$ form a reduced basis of $\tilde{\mathfrak{G}}$.

Obviously, $(1-\sigma)\log \boldsymbol{\xi}$ and $(1-\sigma)\log \boldsymbol{\eta}$ belong to $\tilde{\mathfrak{G}}$. Therefore, Lagrange reduction implies $\|\log \tilde{\boldsymbol{\xi}}\| \leq \|(1-\sigma)\log \boldsymbol{\xi}\| = \sqrt{3}\,\|\log \boldsymbol{\xi}\|$ and $\|\log \tilde{\boldsymbol{\eta}}\| \leq \sqrt{3}\,\|\log \boldsymbol{\eta}\|$.

For $\tilde{\boldsymbol{\gamma}} \in \tilde{\Gamma}(\mathbf{1})$, we have $h(\tilde{\boldsymbol{\gamma}}) = |\log \tilde{\boldsymbol{\gamma}}|_1/6$, where $|\cdot|_1$ denotes the L^1-norm. This is obvious if $[\tilde{\mathfrak{K}} : \mathfrak{K}] = 1$. Otherwise, $h(\tilde{\boldsymbol{\gamma}}) = |(\log \tilde{\boldsymbol{\gamma}}, -\log \tilde{\boldsymbol{\gamma}})|_1/12 = |\log \tilde{\boldsymbol{\gamma}}|_1/6$. On the other hand, we get $|\log \tilde{\boldsymbol{\gamma}}|_1 \leq 2\sqrt{6}/3 \cdot \|\log \tilde{\boldsymbol{\gamma}}\|$ by applying Lagrange's method of unknown multipliers to the function $w_1 - w_2 - w_3$ defined on the circle $w_1^2 + w_2^2 + w_3^2 = \|\log \tilde{\boldsymbol{\gamma}}\|^2$ on $\mathit{\Pi}_{\log}$. We arrive at the second assertion after observing the elementary fact $|\log \tilde{\boldsymbol{\gamma}}|_\infty = |\log \tilde{\boldsymbol{\gamma}}|_1/2$ which follows from $\log \tilde{\boldsymbol{\gamma}} \in \mathit{\Pi}_{\log}$.

The last assertion follows from observing that the angle formed by the reduced basis $\log \tilde{\boldsymbol{\xi}}$ and $\log \tilde{\boldsymbol{\eta}}$ is between $\pi/3$ and $2\pi/3$. □

We now assume $\boldsymbol{\delta}^{1-\sigma}$ is not unit. Then, the module $\log \tilde{\Gamma}(\boldsymbol{\delta})$ is not a lattice by Lemma 6.2. Therefore, we construct a suitable space so that we can capture arithmetic information about $\tilde{\Gamma}(\boldsymbol{\delta})$ in a geometric way. We fix some convention here since we will use many norms: $\|\cdot\|$ denotes the Euclidean (L^2-)norm on $\mathbf{R}^3$; $\|\cdot\|_v$ the v-adic valuation of $\tilde{\mathfrak{K}}$ normalized such that the product formula holds; $|\cdot|$ the absolute value of a real number; and $|\cdot|_p$ the L^p-norm on a Cartesian power of $\mathbf{R}$ or on $\mathrm{Map}_{\mathrm{cpt}}(\mathcal{M}, \mathbf{R})$, $(p = 1, 2, \infty)$. Note: $\mathrm{Map}_{\mathrm{cpt}}(\mathcal{M}, \mathbf{R})$ denotes the space of functions from $\mathcal{M}$ to $\mathbf{R}$ with compact supports (or finite support if $\mathcal{M}$ is discrete). Note also: $\mathrm{Map}_{\mathrm{cpt}}(\mathcal{M}, \mathbf{R})$ is equipped with the L^p norm $|F|_p = (\sum_{v \in \mathcal{M}} |F(v)|^p)^{1/p}$ if p is finite or $|F|_\infty = \max_{v \in \mathcal{M}} |F(v)|\}$.

Let $M^\infty(\tilde{\mathfrak{K}})$ be the set of all infinite places of $\tilde{\mathfrak{K}}$, $M^0(\tilde{\mathfrak{K}})$ the set of all finite places of $\tilde{\mathfrak{K}}$ and $M(\tilde{\mathfrak{K}})$ their union $M^\infty(\tilde{\mathfrak{K}}) \cup M^0(\tilde{\mathfrak{K}})$. Define

$$\tilde{\psi}_{\mathcal{M}} : \tilde{\boldsymbol{\gamma}} \in \tilde{\Gamma}(\boldsymbol{\delta}) \longmapsto (\log \|\tilde{\boldsymbol{\gamma}}\|_v)_{v \in \mathcal{M}} \in \mathrm{Map}_{\mathrm{cpt}}(\mathcal{M}, \mathbf{R})$$

for an arbitrary subset $\mathcal{M}$ of $M(\tilde{\mathfrak{K}})$. Set $\tilde{\psi}^\infty = \tilde{\psi}_{M^\infty(\tilde{\mathfrak{K}})}$, $\tilde{\psi}^0 = \tilde{\psi}_{M^0(\tilde{\mathfrak{K}})}$ and $\tilde{\psi} = \tilde{\psi}_{M(\tilde{\mathfrak{K}})}$. Then, $\tilde{\psi}(\tilde{\Gamma}(\boldsymbol{\delta}))$ is a lattice of rank 3 in $\mathrm{Map}_{\mathrm{cpt}}(M(\tilde{\mathfrak{K}}), \mathbf{R})$, equipped with the L^2-metric mentioned above.

We want to control the L^1 and the L^2-norms of suitable generators of

this lattice. For this purpose, it is preferable to work in a linear space of small finite dimension. Therefore, we shall factor $\tilde{\psi}$ by a map $\psi : \tilde{\Gamma}(\boldsymbol{\delta}) \longrightarrow \mathbf{R}^4$ and investigate the L^1 and the L^2-norms of suitable generators of the module $\psi(\tilde{\Gamma}(\boldsymbol{\delta}))$.

The image $\tilde{\psi}^0(\tilde{\Gamma}(\boldsymbol{\delta}))$ is contained in $\mathbf{Q}\tilde{\psi}^0(\boldsymbol{\delta}^{1-\sigma})$ since $\tilde{\psi}^0((\mathfrak{O}(\mathfrak{K})^\times)^{1-\sigma}) = 0$. This image is non-trivial since $\boldsymbol{\delta}^{1-\sigma}$ is not a unit. Noting that $\tilde{\psi}^0(\tilde{\Gamma}(\boldsymbol{\delta}))$ is discrete with respect to the L^1-topology, we see that $\tilde{\psi}^0(\tilde{\Gamma}(\boldsymbol{\delta})) = \mathbf{Z}\tilde{\psi}^0(\tilde{\boldsymbol{\kappa}})$ for some $\tilde{\kappa} \in \tilde{\Gamma}(\boldsymbol{\delta})$.

The division group $\tilde{\Gamma}(\boldsymbol{\delta})$ is generated by $\tilde{\boldsymbol{\kappa}}$ and $\ker(\tilde{\psi}^0)$, where $\ker(\tilde{\psi}) = \tilde{\Gamma}(\mathbf{1})$ by Lemma 6.1.

We define the homomorphism $\psi^0 : \tilde{\Gamma}(\boldsymbol{\delta}) \longrightarrow \mathbf{R}$ by

$$\psi^0(\tilde{\boldsymbol{\kappa}}) = \frac{1}{[\tilde{\mathfrak{K}} : \mathfrak{K}]} |\tilde{\psi}^0(\tilde{\boldsymbol{\kappa}})|_1$$

and $\psi^0(\tilde{\Gamma}(\mathbf{1})) = 0$. We also define maps $\psi^\infty : \tilde{\boldsymbol{\gamma}} \in \tilde{\Gamma}(\boldsymbol{\delta}) \longmapsto \log \tilde{\boldsymbol{\gamma}} \in \mathbf{R}^3$ and $\psi : \tilde{\boldsymbol{\gamma}} \in \tilde{\Gamma}(\boldsymbol{\delta}) \longmapsto (\psi^\infty(\tilde{\boldsymbol{\gamma}}), \psi^0(\tilde{\boldsymbol{\gamma}})) \in \mathbf{R}^4$. We see

$$|\tilde{\psi}(\tilde{\boldsymbol{\gamma}})|_1 = [\tilde{\mathfrak{K}} : \mathfrak{K}] \cdot |\psi(\tilde{\boldsymbol{\gamma}})|_1 \tag{29}$$

by noting that $\tilde{\psi}^\infty(\tilde{\boldsymbol{\gamma}}) = \log \tilde{\boldsymbol{\gamma}}$ if $[\tilde{\mathfrak{K}} : \mathfrak{K}] = 1$ or $\tilde{\psi}^\infty(\tilde{\boldsymbol{\gamma}}) = (\log \tilde{\boldsymbol{\gamma}}, -\log \tilde{\boldsymbol{\gamma}})$ if $[\tilde{\mathfrak{K}} : \mathfrak{K}] = 2$. The module $\psi(\tilde{\Gamma}(\boldsymbol{\delta})) = (\tilde{\mathfrak{G}} \oplus 0) + \mathbf{Z}\psi(\tilde{\boldsymbol{\kappa}})$ is a lattice of rank 3 since $\psi(\tilde{\boldsymbol{\kappa}}) \notin \mathbf{R}^3 \oplus 0$.

Lemma 6.4. *Assume $\boldsymbol{\delta}^{1-\sigma}$ is not a unit. Let $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ be independent units of $\mathfrak{K}$ such that $\|\log \boldsymbol{\xi}\| \leq \|\log \boldsymbol{\eta}\|$. Then, there is a triple $\tilde{\boldsymbol{\delta}}$, $\tilde{\boldsymbol{\xi}}$ and $\tilde{\boldsymbol{\eta}}$ of generators of $\tilde{\Gamma}(\boldsymbol{\delta})$ (modulo $\{+1, -1\}$) such that (a permutation of) $\psi(\tilde{\boldsymbol{\delta}})$, $\psi(\tilde{\boldsymbol{\xi}})$ and $\psi(\tilde{\boldsymbol{\eta}})$ form a reduced basis for $\psi(\tilde{\Gamma}(\boldsymbol{\delta}))$ and their L^2-norms satisfy the following inequalities:*

$$\begin{cases} 3h(\tilde{\boldsymbol{\delta}}) \leq |\psi(\tilde{\boldsymbol{\delta}})|_2 \leq |\psi(\boldsymbol{\delta}^{1-\sigma})|_2 \leq \sqrt{3}\sqrt{\|\phi(\boldsymbol{\delta})\|^2 + (1/3)\log^2 D}; \\ 3h((\tilde{\boldsymbol{\xi}}) \leq |\psi(\tilde{\boldsymbol{\xi}})|_2 \leq |\psi(\boldsymbol{\xi}^{1-\sigma})|_2 \leq \sqrt{3}\,\|\log \boldsymbol{\xi}\|; \\ 3h(\tilde{\boldsymbol{\eta}}) \leq |\psi(\tilde{\boldsymbol{\eta}})|_2 \leq |\psi(\boldsymbol{\eta}^{1-\sigma})|_2 \leq \sqrt{3}\,\|\log \boldsymbol{\eta}\|. \end{cases} \tag{30}$$

Let $\boldsymbol{u} \in \phi(\boldsymbol{\delta}) + \mathbf{Z}\log \boldsymbol{\xi} + \mathbf{Z}\log \boldsymbol{\eta}$ and write $(1-\sigma)\boldsymbol{u} = l\log \tilde{\boldsymbol{\delta}} + m\log \tilde{\boldsymbol{\xi}} + n\log \tilde{\boldsymbol{\eta}}$ with integers l, m, and n. Then, we have

$$\max\left\{|l\psi(\tilde{\boldsymbol{\delta}}))|_2, |m\psi(\tilde{\boldsymbol{\xi}}))|_2, |n\psi(\tilde{\boldsymbol{\eta}}))|_2\right\} \leq \sqrt{8\|\boldsymbol{u}\|^2 + (8/3)\log^2 D}. \tag{31}$$

PROOF. Firstly, the existence of a basis ($\psi(\tilde{\boldsymbol{\delta}})$, $\psi(\tilde{\boldsymbol{\xi}})$ and $\psi(\tilde{\boldsymbol{\eta}})$) of $\psi(\tilde{\Gamma}(\boldsymbol{\delta}))$ satisfying the three inequalities in the middle column of (30) is guaranteed by Seeber–Vallée reduction.

Secondly, we show that $\tilde{\boldsymbol{\delta}}$, $\tilde{\boldsymbol{\xi}}$ and $\tilde{\boldsymbol{\eta}}$ satisfy the other inequalities of (30).

The identity (29) implies $h(\tilde{\boldsymbol{\gamma}}) = |\psi(\tilde{\boldsymbol{\gamma}})|_1/6$ for arbitrary $\tilde{\gamma} \in \tilde{\Gamma}(\boldsymbol{\delta})$. On the other hand, we have the usual estimate $|\psi(\tilde{\boldsymbol{\gamma}})|_1 \leq 2|\psi(\tilde{\boldsymbol{\gamma}})|_2$ of the L^1-norm in terms of the L^2-norm. (Note: the image of ψ is contained in $\mathbf{R}^4$.) Therefore, the three inequalities on the left hold.

By Pythagoras' theorem $|\psi(\boldsymbol{\delta}^{1-\sigma})|_2^2 = \|(1-\sigma)\phi(\boldsymbol{\delta})\|^2 + \|\psi^0(\boldsymbol{\delta}^{1-\sigma})\|_1^2$. For the first term, we have $\|(1-\sigma)\phi(\boldsymbol{\delta})\| = \sqrt{3}\,\|\phi(\boldsymbol{\delta})\|$ and for the second sum we have $|\psi^0(\boldsymbol{\delta}^{1-\sigma})|_1 = |\tilde{\psi}^0(\boldsymbol{\delta}^{1-\sigma})|_1/[\tilde{\mathfrak{K}} : \mathfrak{K}] \leq 2|\tilde{\psi}^0(\boldsymbol{\delta})|_1/[\tilde{\mathfrak{K}} : \mathfrak{K}]$. Further, the product formula implies $|\tilde{\psi}^0(\boldsymbol{\delta})|_1 = [\tilde{\mathfrak{K}} : \mathfrak{K}]\log(\delta_1\delta_2\delta_3) = ([\tilde{\mathfrak{K}} : \mathfrak{K}]/2)\log D$. Hence, we get $|\psi^0(\boldsymbol{\delta}^{1-\sigma})|_1 \leq \log D$. Therefore, the inequality in the top of the right column holds. The two other inequalities in the right column are more easily proved.

We lastly show the inequality (31). Write $\boldsymbol{u} = \phi(\boldsymbol{\delta}) + I\log\boldsymbol{\xi} + J\log\boldsymbol{\eta}$. Set $\tilde{\boldsymbol{u}} = \psi((\boldsymbol{\delta}\boldsymbol{\xi}^I\boldsymbol{\eta}^J)^{1-\sigma})$ and $\tilde{r} = |\tilde{\boldsymbol{u}}|_2$. Then, we get $\tilde{r}^2 = 3\|\boldsymbol{u}\|^2 + |\psi^0((\boldsymbol{\delta})^{1-\sigma})|^2$ by following the argument for the upper estimate of $|\psi(\boldsymbol{\delta}^{1-\sigma})|_2$ and noting $\psi^0(\boldsymbol{\xi}^I\boldsymbol{\eta}^J)$=0. Hence, we have $\tilde{r}^2 = 3\|\boldsymbol{u}\|^2 + |\psi^0(\boldsymbol{\delta}^{1-\sigma})|^2$.

It now suffices to estimate the left hand side of (31) by $(2\sqrt{2}/\sqrt{3})\tilde{r}$. Without loss of generality, we assume $|\psi(\tilde{\boldsymbol{\delta}})|_2 \leq |\psi(\tilde{\boldsymbol{\xi}})|_2 \leq |\psi(\tilde{\boldsymbol{\eta}})|_2$. Decompose $\psi(\tilde{\boldsymbol{\eta}})$ in a sum of an orthogonal vector and parallel vector to the plane spanned by $\psi(\tilde{\boldsymbol{\delta}})$ and $\psi(\tilde{\boldsymbol{\xi}})$ as $\psi(\tilde{\boldsymbol{\eta}}) = \psi(\tilde{\boldsymbol{\eta}})^{\perp} + \psi(\tilde{\boldsymbol{\eta}})^{\|}$. We obviously have $4|\psi(\tilde{\boldsymbol{\eta}})^{\|}|_2^2 \leq |\psi(\tilde{\boldsymbol{\delta}})|_2^2 + |\psi(\tilde{\boldsymbol{\xi}})|_2^2 \leq 2|\psi(\tilde{\boldsymbol{\eta}})|_2^2$. Thus, the angle formed by $\psi(\tilde{\boldsymbol{\eta}})$ and $\psi(\tilde{\boldsymbol{\eta}})^{\|}$ is between $\pi/4$ and $3\pi/4$. So is the angle formed by $\psi(\tilde{\boldsymbol{\eta}})$ and an arbitrary linear combination $\tilde{\boldsymbol{w}}$ of $\psi(\tilde{\boldsymbol{\delta}})$ and $\psi(\tilde{\boldsymbol{\xi}})$. We set $\tilde{\boldsymbol{w}} = l\psi(\tilde{\boldsymbol{\delta}}) + m\psi(\tilde{\boldsymbol{\xi}})$. Then, the estimate of the angle implies that the diameter of the circle passing through 0, $\tilde{\boldsymbol{w}}$ and $\tilde{\boldsymbol{u}}$ is at most $\sqrt{2}\,\tilde{r}$. (Note: consider circumferential angle.) Therefore, we get $\max\{|\tilde{\boldsymbol{w}}|_2, |n\tilde{\psi}(\tilde{\boldsymbol{\eta}})|_2\} \leq \sqrt{2}\,\tilde{r}$. Applying the same argument to $\tilde{\boldsymbol{w}}$ and noting that the angle of $\psi(\tilde{\boldsymbol{\delta}})$ and $\psi(\tilde{\boldsymbol{\xi}})$ is between $\pi/3$ and $2\pi/3$, we get $\max\{|l\psi(\tilde{\boldsymbol{\delta}})|_2, |m\psi(\tilde{\boldsymbol{\xi}})|_2\} \leq (2/\sqrt{3})|\boldsymbol{w}|_2$. The desired estimate now follows immediately. □

The following lemma estimates the heights of generators of $\tilde{\Gamma}(\boldsymbol{\delta})$ from below:

Lemma 6.5. *Let $\tilde{\boldsymbol{\gamma}} \in \tilde{\Gamma}(\boldsymbol{\delta})$. Then, we have*

$$|\psi(\tilde{\boldsymbol{\gamma}})|_2 = \|\log \tilde{\boldsymbol{\gamma}}\| \geq \frac{1}{2\sqrt{2}} \log \frac{49}{4} = 0.8858\cdots$$

if $\boldsymbol{\delta}^{1-\sigma}$ is a unit and $[\tilde{\mathfrak{K}} : \mathfrak{K}] = 1$; or

$$|\psi(\tilde{\boldsymbol{\gamma}})|_2 = \|\log \tilde{\boldsymbol{\gamma}}\| \geq \frac{1}{2\sqrt{6}} \log \frac{148}{4} = 0.7370\cdots$$

if $\boldsymbol{\delta}^{1-\sigma}$ is a unit and $[\tilde{\mathfrak{K}} : \mathfrak{K}] = 2$; or

$$|\psi(\tilde{\boldsymbol{\gamma}})|_2 \geq \log 2 = 0.693\cdots$$

otherwise.

PROOF. Firstly, we consider the case that $\tilde{\boldsymbol{\gamma}}$ is a unit. This is always the case when $\boldsymbol{\delta}^{1-\sigma}$ is a unit (Lemma 6.2). If $[\tilde{\mathfrak{K}} : \mathfrak{K}] = 1$, then Lemma 5.7 and $D \geq 49$ imply the first inequality.

If $[\tilde{\mathfrak{K}} : \mathfrak{K}] = 2$, choose a unit $\boldsymbol{\gamma} \in \mathfrak{K}$ by Lemma 6.1 so that $3\|\log \tilde{\boldsymbol{\gamma}}\| = \|(1-\sigma)\log \boldsymbol{\gamma}\| = \sqrt{3}\,\|\log \boldsymbol{\gamma}\|$. Since $D \geq 148$, Lemma 5.7 implies $\|\log \boldsymbol{\gamma}\| \geq (1/2\sqrt{2})\log(148/4)$. Thus, the second inequality follows.

The third inequality follows from one of the two inequalities.

Secondly, we consider the case that $\tilde{\boldsymbol{\gamma}}$ is not a unit. Then, neither is $\boldsymbol{\delta}^{1-\sigma}$. Thus, it suffice to prove the third inequality. We have $\log \|\tilde{\boldsymbol{\gamma}}\|_v \neq 0$ for some $v \in M^0(\tilde{\mathfrak{K}})$ since $\tilde{\boldsymbol{\gamma}}$ is not a unit. Such a v cannot be unique since $\tilde{\gamma}_1\tilde{\gamma}_2\tilde{\gamma}_3 = 1$. (Note: we can use either the product formula or pull back $\log \|\tilde{\boldsymbol{\gamma}}^{\sigma(1+\sigma)}\|_v = -\log \|\tilde{\boldsymbol{\gamma}}\|_v$.) Therefore, $|\tilde{\psi}^0(\tilde{\boldsymbol{\gamma}})|_1 \geq 2\log 2$. Hence, $|\psi(\tilde{\boldsymbol{\gamma}})|_2 \geq \log 2$ follows from the definition of the map ψ. □

7. Upper bound for geometric sizes

We shall bound the geometric sizes $\|\phi(\boldsymbol{z})\|$ of a given point $\boldsymbol{z}$ of $\mathfrak{L}^{\natural} \cap \mathcal{H}$ by using Baker theory, i.e., lower bounds for linear forms in logarithms of algebraic numbers. In this paper, the logarithms of scalars are always understood as real logarithms.

Theorem 7.1. *Set $\mathfrak{K} = \mathfrak{K}(f)$. Assume $\boldsymbol{\delta}^{1-\sigma}$ is not a unit. Let $\boldsymbol{z} \in \mathfrak{L}^\natural \cap \mathcal{H}$, $t = t(\phi(\boldsymbol{z}))$. Then, we have*

$$\frac{t}{\operatorname{disc}(\mathfrak{E}(\mathfrak{K}))} < 3.418 \cdot 10^{11} \cdot L \log^2(1.01 \cdot 10^{10} \cdot L \| \log \boldsymbol{\xi} \|), \tag{32}$$

where $\| \log \boldsymbol{\xi} \|$ is the first minimum of $\mathfrak{E}(\mathfrak{K})$ and $L = \sqrt{\|\phi(\boldsymbol{\delta})\|^2 + (1/3)\log^2 D}$.

Theorem 7.2. *Let $\mathfrak{K} = \mathfrak{K}(f)$ and assume $\boldsymbol{\delta}^{1-\sigma}$ is a unit. Let $\boldsymbol{z} \in \mathfrak{L}^\natural \cap \mathcal{H}$ and $t = t(\phi(\boldsymbol{z}))$. Then, we have*

$$\frac{t}{\operatorname{disc}(\mathbf{Z}\phi(\boldsymbol{\delta}) + \mathfrak{E}(\mathfrak{K}))} \leq 3.54 \cdot 10^5.$$

The right hand side can be replaced with $5.04 \cdot 10^4$ if the field $\mathfrak{K}$ is cyclic.

Remark. If we use Theorem 2.1 of [16], we can show

$$\frac{t}{\operatorname{disc}(\mathfrak{E}(\mathfrak{K}))} < 3.11 \cdot 10^{13} \cdot L \log(2.69 \cdot 10^{15} \cdot L \| \log \boldsymbol{\xi} \|). \tag{33}$$

However, the extra log-factor of (32) will be less than 33 in the critical situation. Therefore, the upper bound (32) will be roughly 1/3 of the upper bound (33).

Remark. The explicit dependence of (32) on quantities L and $\| \log \boldsymbol{\xi} \|$ will be utilized in our proof of Theorem 1.1, where we control them by a "small" solution. Theorem 7.1 is actually specialized for this purpose. For other purposes, its dependence on the invariants of f is not nice. Here, we explain this fact by doing a little exercise on (33) which has better dependences than (32). We replace $\phi(\boldsymbol{\delta})$ with its translation by $\mathfrak{E}(\mathfrak{K})$ into the Voronoï domain of $\mathbf{0}$ with respect to $\mathfrak{E}(\mathfrak{K})$ and estimate it with $\| \log \boldsymbol{\eta} \| / \sqrt{2}$, where $\| \log \boldsymbol{\eta} \|$ denotes the second minimum of $\mathfrak{E}(\mathfrak{K})$. We then use Lemma 5.7 and an analytic estimate of the regulator (see the proof of Theorem 1.1) to show $t \leq 2 \cdot 10^{15} \operatorname{disc}(\mathfrak{E}(\mathfrak{K}))^2 + 10^{15} \operatorname{disc}(\mathfrak{E}(\mathfrak{K})) \log D \times \log\log D$. We now follow the proof for Lemma 5.3 to deduce an upper bound for $\|\boldsymbol{z}\|$ and use the property of the reduced basis of $\mathfrak{L}^\natural$ (see Proposition 3.2) to show

$$\log \max\{|x|, |y|\} \leq 2.5 \cdot 10^{15} R(\mathfrak{K})(R(\mathfrak{K}) + \log D \log\log D)$$

if we assume f is reduced. ($R(\mathfrak{K}) = \mathrm{disc}(\mathfrak{E}(\mathfrak{K}))/\sqrt{3}$ denotes the regulator of $\mathfrak{K}$.) Here, the factor $\log\log D$ is responsible for the case $R(\mathfrak{K}) = o(\log^2 D)$. The form of this estimate is not better than the previously known estimates

$$\log\max\{|x|,|y|\} \leq 7.07 \cdot 10^{44} R(\mathfrak{K})(R(\mathfrak{K}) + \log H)\log R(\mathfrak{K}),$$

of BUGEAUD–GYŐRY [3] (H is the maximum magnitude of the coefficients of f) nor

$$\log\max\{|x|,|y|\} \leq 1.01 \cdot 10^{56} R(\mathfrak{K})(h(\boldsymbol{\alpha}) + h(\boldsymbol{\beta})) + 1.01 \cdot 10^{114} R(\mathfrak{K})^2.$$

BUGEAUD [2] (Note: trivial estimates of $\log D$ are $\log D \leq \log(64H^4/3)$ and $\log D \leq 12(h(\boldsymbol{\alpha}) + h(\boldsymbol{\beta}) + \log\sqrt{2}\,)$.)

For proving Theorem 7.1, we use Matveev's lower bound, which is quoted below with some restrictions and simplifications. (The role of the subscripts differs from the other part of this paper.)

Theorem 7.3 (MATVEEV [15]). *Let $\Lambda = b_1\log\gamma_1 + b_2\log\gamma_2 + b_3\log\gamma_3$ be a linear combination with integer coefficients in logarithms of multiplicatively independent positive totally real algebraic numbers γ_1, γ_2 and γ_3. Assume the group generated by γ_1, γ_2, γ_3 and -1 coincides with its division group in $\mathbf{Q}(\gamma_1,\gamma_2,\gamma_3)^\times$.*

Choose parameters d, A_1, A_2, A_3 and B such that

$$d \geq [\mathbf{Q}(\gamma_1,\gamma_2,\gamma_3) : \mathbf{Q}];$$

$$dA_i \geq \max\{dh(\gamma_i), |\log\gamma_i|, 0.56\}, \quad (i = 1,2,3);$$

$$B \geq \max\{|b_1|A_1, |b_2|A_2, |b_3|A_3\}.$$

If $b_3 \neq 0$, we have

$$\log|\Lambda| > -C_1 d^5 A_1A_2A_3 \log(C_2 d^4 A_1A_2)\log(2eB/A_3),$$

where $C_1 = 1.9546 \cdot 10^8$ and $C_2 = 1.2032 \cdot 10^5$.

If $b_3 = 0$ and $b_2 \neq 0$, we have

$$\log|\Lambda| > -C_3 d^4 A_1A_2 \log(C_4 d^3 A_1)\log(e2B/A_2),$$

where $C_3 = 1.3608 \cdot 10^7$ and $C_4 = 1.1958 \cdot 10^4$.

PROOF of Theorem 7.1. By Lemma 5.7 and $D \geq 49$, the right hand side of (32) is greater than $10^{11} \log D$ and $\mathrm{disc}(\mathfrak{E}(\mathfrak{K})) \geq 0.6$. Hence, it suffice to discuss the case $t \geq 4 \log D$. Set $r = \|\phi(\boldsymbol{z})\|$. Then, we have $r \leq 1.01t$ by Theorem 5.2.

Without loss of generality, we assume $\boldsymbol{z} \in \mathcal{C}_3$. We consider

$$\Lambda = \sqrt{2} s(\phi(\boldsymbol{z})).$$

Its absolute value is small. Indeed, we have

$$\log |\Lambda| < -(\sqrt{6}/2)\, t \tag{34}$$

by the inequality (21) of Theorem 5.2.

By Theorem 4.1, there is a unit $\varepsilon \in \mathfrak{O}(f)^\times$ such that $\boldsymbol{z} = \boldsymbol{\delta}\boldsymbol{\varepsilon}$. Let $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ be a fundamental pair of $\mathfrak{O}(\mathfrak{K})^\times$ such that $\log \boldsymbol{\xi}$ and $\log \boldsymbol{\eta}$ form a reduced basis of $\mathfrak{E}(\mathfrak{K})$. Choose $\tilde{\boldsymbol{\delta}}$, $\tilde{\boldsymbol{\xi}}$ and $\tilde{\boldsymbol{\eta}}$ by Lemma 6.4 and write $(1-\sigma) \log \boldsymbol{z} = l \log \tilde{\boldsymbol{\delta}} + m \log \tilde{\boldsymbol{\xi}} + n \log \tilde{\boldsymbol{\eta}}$ with suitable integers l, m and n. Then, we have

$$\Lambda = l \log \left|\tilde{\delta}_1\right| + m \log \left|\tilde{\xi}_1\right| + n \log |\tilde{\eta}_1|\,. \tag{35}$$

We discuss the case $n \neq 0$ since the other case is easier.

We specify parameters for Theorem 7.3. We set $\gamma_1 = |\tilde{\delta}_1|$, $\gamma_2 = |\tilde{\xi}_1|$, $\gamma_3 = |\tilde{\eta}_1|$, $b_1 = l$, $b_2 = m$ and $b_3 = n$. Then, γ_1, γ_2 and γ_3 satisfy the assumption of Theorem 7.3 by Lemma 6.4. We set $d = 6$, so that $[\tilde{\mathfrak{K}} : \mathbf{Q}] \leq d$. We set $A_1 = |\psi(\tilde{\boldsymbol{\delta}})|_2/3$, $A_2 = |\psi(\tilde{\boldsymbol{\xi}})|_2/3$ and $A_3 = |\psi(\tilde{\boldsymbol{\eta}})|_2/3$. Lemma 6.4 guarantees $A_i \geq h(\gamma_i)$ $(i = 1, 2, 3)$. Since γ_i is totally positive, the inequality $|\log \gamma_i| \leq [\mathbf{Q}(\gamma_i) : \mathbf{Q}] h(\gamma_i) \leq d h(\gamma_i)$ is easily verified by noting the product formula. Lemma 6.5 implies $dA_i \geq 2 \log 2$. Therefore, our A_i's are suitable for Theorem 7.3. Set $B = t$. Lemma 6.4, the assumption $t \geq 4 \log D$ and the inequality $r \leq 1.01t$ imply that B is suitable for Theorem 7.3.

Theorem 7.3 and (34) imply

$$\frac{2et/A_3}{\log(2et/A_3)} < 6.75 \cdot 10^{12} \Omega \log(1.56 \cdot 10^8 \Omega),$$

where we abbreviated $\Omega = A_1 A_2$. Since we already verified $A_i \geq (\log 2)/3$, the right hand side is larger than $5.73 \cdot 10^{12}$. Thus, we get

$$t \leq 1.391 \cdot 10^{12} \cdot \Omega A_3 \log(1.56 \cdot 10^8 \Omega) \log(6.75 \cdot 10^{12} \cdot \Omega \log(1.56 \cdot 10^8 \Omega)).$$

Hence, we get

$$t \leq 1.538 \cdot 10^{12} \cdot \Omega A_3 \log^2(3.25 \cdot 10^{10}\Omega). \tag{36}$$

by using the inequality of the arithmetic and the geometric means.

On the other hand, we have $A_2 A_3 \leq (1/3)\|\log \boldsymbol{\xi}\| \cdot \|\log \boldsymbol{\eta}\|$ by Lemma 6.4. Since $\log \boldsymbol{\xi}$ and $\log \boldsymbol{\eta}$ form a reduced basis for $\mathfrak{E}(\mathfrak{K})$, this implies

$$A_2 A_3 \leq (2/3\sqrt{3})\,\mathrm{disc}(\mathfrak{E}(\mathfrak{K})).$$

Substituting this inequality in (36), we get

$$\frac{t}{\mathrm{disc}(\mathfrak{E}(\mathfrak{K}))} \leq \frac{2}{3\sqrt{3}} \cdot 1.538 \cdot 10^{12} \cdot A_1 \log^2(3.25 \cdot 10^{10}\Omega).$$

The theorem follows after substituting estimates of A_1 and A_2 (Lemma 6.4 again) in this inequality. □

For proving Theorem 7.2, we use the lower bound due to Laurent–Mignotte–Nesterenko quoted below with some restrictions and simplifications. (The role of the subscripts again differs from the other part of this paper.)

Theorem 7.4 (Laurent–Mignotte–Nesterenko [12]). *Let $\Lambda = b_1 \log \gamma_1 + b_2 \log \gamma_2$ be a linear combination with integer coefficients in logarithms of multiplicatively independent positive totally real algebraic numbers γ_1 and γ_2.*

Let $d = [\mathbf{Q}(\gamma_1, \gamma_2) : \mathbf{Q}]$ and $\lambda \geq 0.3$. Choose parameters A_1, A_2, B, b and b' such that

$$A_i \geq \max\left\{(e^{\lambda} - 1)|\log \gamma_i| + 2dh(\gamma_i), 2\lambda, 2\right\}, \quad (i = 1, 2);$$

$$B \geq (|b_1|A_1 + |b_2|A_2)/A_1 A_2;$$

$$b \geq \max\{d(\log B + \log \lambda + 1.56), d/2, 5\lambda\}$$

$$b' = b + \lambda + \lambda^2/4b.$$

Then, we have

$$\frac{\log|\Lambda|}{A_1 A_2 b'^2} \geq -\frac{16}{9\lambda^3} - \frac{8(A_1+A_2)+6\lambda}{3\lambda A_1 A_2 b'} - \frac{16\sqrt{2}}{3\sqrt{\lambda^3 A_1 A_2 b'}} - \frac{\log\left(\lambda^{-3} A_1 A_2 b'^2\right)}{A_1 A_2 b'^2}.$$

PROOF of Theorem 7.2. Without loss of generality, we assume $t/\operatorname{disc}(\mathbf{Z}\phi(\boldsymbol{\delta}) + \mathfrak{E}(\mathfrak{K})) \geq 10^4$. Then, Lemma 5.7 and 6.2 imply $t \geq 10^3$.

Without loss of generality, we assume $\boldsymbol{z} \in \mathcal{C}_3$. We consider

$$\Lambda = \sqrt{2}s(\phi(\boldsymbol{z})).$$

Then, we have

$$\log|\Lambda| < -(\sqrt{6}/2)\,t \tag{37}$$

by the inequality (21) of Theorem 5.2.

Let $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ be elements of $\mathfrak{O}(\mathfrak{K})^\times$ such that $\log\boldsymbol{\xi}$ and $\log\boldsymbol{\eta}$ form a reduced basis for $\mathfrak{E}(\mathfrak{K})$. Choose generators $\tilde{\boldsymbol{\xi}}$ and $\tilde{\boldsymbol{\eta}}$ of $\tilde{\Gamma}(\boldsymbol{\delta}) = \tilde{\Gamma}(\mathbf{1})$ by Lemma 6.3 and write $\boldsymbol{z}^{1-\sigma} = \pm\tilde{\boldsymbol{\xi}}^m\tilde{\boldsymbol{\eta}}^n$ with integers m and n. Then, we have

$$\Lambda = m\log\left|\tilde{\xi}_1\right| + n\log|\tilde{\eta}_1|\,.$$

We specify parameters for Theorem 7.4. Put $\gamma_1 = |\tilde{\xi}_1|$, $\gamma_2 = |\tilde{\eta}_1|$, $b_1 = m$, $b_2 = n$ and $d = [\tilde{\mathfrak{K}} : \mathbf{Q}]$. We set $\lambda = 1.7$ or 2.1 according as $d = 3$ or 6. Let $c = \sqrt{6}/9 \cdot (3(e^\lambda - 1) + 2d)$, $A_1 = c\|\log\tilde{\boldsymbol{\xi}}\|$ and $A_2 = c\|\log\tilde{\boldsymbol{\eta}}\|$. We have $A_i \geq (e^\lambda - 1)|\log\gamma_i| + 2dh(\gamma_i)$ by Lemma 6.3. We also have

$$A_i \geq \begin{cases} 4.68 & \text{if } d = 3; \\ 6.72 & \text{if } d = 6 \end{cases} \tag{38}$$

by Lemma 6.5 and hence $A_i \geq 2\lambda \geq 2$. Thus, A_1 and A_2 are suitable for Theorem 7.4. We set $B = 2.02\sqrt{3}\,ct/A_1A_2$, which is also suitable since we have $2\sqrt{3}\,\|\phi(\boldsymbol{z})\| = 2\|(1-\sigma)\phi(\boldsymbol{z})\| = 2\|m\log\tilde{\boldsymbol{\xi}} + n\log\tilde{\boldsymbol{\eta}}\| \geq \|m\log\tilde{\boldsymbol{\xi}}\| + \|n\log\tilde{\boldsymbol{\eta}}\|$ by Lemma 6.3 and $\|\phi(\boldsymbol{z})\| \leq 1.01t$ by Theorem 5.2.

We have $\|\log\tilde{\boldsymbol{\xi}}\| \cdot \|\log\tilde{\boldsymbol{\eta}}\| \leq (2/\sqrt{3})\operatorname{disc}\tilde{\mathfrak{G}} \leq 2\sqrt{3}\operatorname{disc}(\mathbf{Z}\phi(\boldsymbol{\delta}) + \mathfrak{E}(\mathfrak{K}))$ since $(1-\sigma)(\mathbf{Z}\phi(\boldsymbol{\delta}) + \mathfrak{E}(\mathfrak{K})) \subset \tilde{\mathfrak{G}}$.Thus,

$$B = \frac{2.02\sqrt{3}\,t}{c\|\log\tilde{\boldsymbol{\xi}}\| \cdot \|\log\tilde{\boldsymbol{\eta}}\|} \geq \frac{1.01\,t}{c\operatorname{disc}(\mathbf{Z}\phi(\boldsymbol{\delta}) + \mathfrak{E}(\mathfrak{K}))} \geq 10^3. \tag{39}$$

Therefore, $b = d(\log B + \log\lambda + 1.56)$ is also suitable.

Theorem 7.4 and the inequality (37) now imply

$$\frac{B/\sqrt{2}}{2.02cb'^2} \leq \frac{16}{9\lambda^3} + \frac{8(A_1+A_2)+6\lambda}{3\lambda A_1A_2b'} + \frac{16\sqrt{2}}{3\sqrt{\lambda^3A_1A_2b'}} + \frac{\log\left(\lambda^{-3}A_1A_2b'^2\right)}{A_1A_2b'^2}.$$

The left hand side is an increasing function of B. (b/b' and $\sqrt{B}/b$ are increasing function of $\log B$.) The right hand side is a decreasing function in A_1, A_2 and b' since $A_1 A_2 b'^2 \geq 4.6^2 \cdot 3\log^2 10^3 > \lambda^3 e$. Therefore, we can replace A_1 and A_2 with one of the lower bounds (38) depending on d. Then, the left hand side is larger than the right hand side at $B = 9.63 \cdot 10^3$ when $d = 3$ or at $B = 3.92 \cdot 10^4$ when $d = 6$. Therefore, B is less than the mentioned values. The theorem now follows after substituting these values in (39). □

8. Proof of main results

We prove our main results. We use Proposition 5.1 to associate $\phi(\boldsymbol{z}) = \phi(\boldsymbol{\delta}(x\boldsymbol{\alpha} + y\boldsymbol{\beta}))$ with the representation (x, y) and apply results of the previous sections. The main task is to estimate $t = t(\phi(\boldsymbol{z}))$ and $r = \|\phi(\boldsymbol{z})\|$. Since the upper bound for t we want to establish is larger than 8, we will be working on points with $t \geq 8$. Hence, we can assume $t < r < 1.001t$ by Theorem 5.2.

PROOF of Theorem 1.1. Let $\boldsymbol{w}$ be the element of $\mathfrak{L}^\natural \cap \mathcal{H}$ such that $\|\phi(\boldsymbol{w})\|$ is minimal.

We suppose $\mathfrak{L}^\natural \cap \mathcal{H}_k$ with some k has three distinct points $\boldsymbol{z}$, $\boldsymbol{z}'$ and $\boldsymbol{z}''$ other than $\boldsymbol{w}$. Without loss of generality, we assume $t'' = t(\phi(\boldsymbol{z}'')) \geq t' = t(\phi(\boldsymbol{z}')) \geq t = t(\phi(\boldsymbol{z}))$. Put $r = \|\phi(\boldsymbol{z})\|$. We firstly show

$$r \leq 30.9. \tag{40}$$

Without loss of generality, we also assume $t \geq 10$.

We can apply Theorem 7.2 or Theorem 7.1 according as $\boldsymbol{\delta}^{1-\sigma}$ is a unit or not. Since the latter gives a larger upper bound, we get

$$\frac{t''}{\operatorname{disc}(\mathfrak{E}(\mathfrak{K}))} < 3.418 \cdot 10^{11} \cdot L \log^2(1.01 \cdot 10^{10} \cdot L \| \log \boldsymbol{\xi} \|), \tag{41}$$

where $\log \boldsymbol{\xi}$ is the first minimum of $\mathfrak{E}(\mathfrak{K})$ and $L = \sqrt{\|\phi(\boldsymbol{\delta})\|^2 + (1/3)\log^2 D}$.

Lemma 5.4 implies $t' - t > t$ and $t'' - t' > t' > t \geq \sqrt{6}\log 2$. By substituting the lower bound of Theorem 5.5 (with $\mathfrak{M} = \mathfrak{E}(\mathfrak{K})$) in (41), we

get

$$\frac{\sqrt{2}\exp\left(\sqrt{6}\,t/2\right)}{1+\exp\left(-2t/\sqrt{6}\log 2\right)} \tag{42}$$

$$< 3.418\cdot 10^{11}\cdot L\log^2(1.01\cdot 10^{10}\cdot L\|\log\boldsymbol{\xi}\|).$$

We estimate the quantities appearing in (42). The denominator of the left hand side is smaller than 1.001. We have $\|\log\boldsymbol{\xi}\| \le \|\phi(\boldsymbol{z})-\phi(\boldsymbol{w})\| \le 2r$. By Lemma 5.8, we have $\log D \le 4\sqrt{2}\,r+\log 4$.

Therefore, we get

$$\exp\left(\sqrt{6}\,r/2.002\right) < 2.42\cdot 10^{11}\cdot L\log^2(2.02\cdot 10^{10}\cdot Lr)$$

and $L \le \sqrt{r^2+(4\sqrt{2}\,r+\log 4)^2/3}$. These imply the inequality (40).

Now suppose $\#\mathcal{R}\ge 8$. Then, Proposition 3.3 and 5.1 implies $\mathfrak{L}^\natural\cap\mathcal{H}\ge 8$. Thus, three points of $\mathfrak{L}^\natural\cap\mathcal{H}$ other than $\boldsymbol{w}$ concentrate on $\mathfrak{L}^\natural\cap\mathcal{H}_k$ with some k. Therefore, the supposition at the beginning of the proof must hold. Now, (40) and Lemma 5.8 with $\boldsymbol{z}\ne\boldsymbol{w}$ imply $D\le 5.65\cdot 10^{65}$. Hence, the first assertion is established.

Assume $f(X,Y)$ is reduced and $f(1,0)\ne\pm 1$. The basic difference from the previous situation consists of two points: we can use Lemma 5.3 for estimating r; we cannot assume $\boldsymbol{\delta}\ne\boldsymbol{z}$ since we do not have an extra point $\boldsymbol{w}$ at our disposal. Therefore, we must estimate $\|\log\boldsymbol{\xi}\|$ by another method. To this end, we use an idea of Siegel.

Recall the analytic class number formula (see e.g. page 38 of [27]):

$$4h(\mathfrak{K})R(\mathfrak{K}) = \sqrt{D(\mathfrak{K})}\,\mathrm{Res}_{s=1}(\zeta_{\mathfrak{K}}),$$

where $h(\mathfrak{K})$ is a positive integer called the class number of $\mathfrak{K}$ and $R(\mathfrak{K}) = \mathrm{disc}(\mathfrak{E}(\mathfrak{K}))/\sqrt{3}$ is called the regulator of $\mathfrak{K}$. The right hand side is estimated by Louboutin's upper bound [14, Theorem 2]:

$$\mathrm{Res}_{s=1}(\zeta_{\mathfrak{K}}) \le (1/8)\log^2 D(\mathfrak{K}).$$

On the other hand, we have $\mathrm{disc}(\mathfrak{K}) \ge (\sqrt{3}/2)\|\log\boldsymbol{\xi}\|^2$ since $\log\boldsymbol{\xi}$ is the first minimum of $\mathfrak{E}(\mathfrak{K})$. Therefore, we get

$$\|\log\boldsymbol{\xi}\| \le (1/4)\,D(\mathfrak{K})^{1/4}\log D(\mathfrak{K}) \le (1/4)\,D^{1/4}\log D.$$

Since $\log D \leq 2\sqrt{6}\, r + \log 2.5$ by Lemma 5.3, we get an estimate of $\| \log \boldsymbol{\xi} \|$ in terms of r. By substituting it in (42), we get

$$r < 32.4.$$

Thus, Lemma 5.3 implies $D \leq 2.15 \cdot 10^{69}$. We can now deduce the second assertion by adding up the number of points in $\mathfrak{L}^\natural \cap \mathcal{H}_k$. □

PROOF of Theorem 1.2. Let $\boldsymbol{\delta}\boldsymbol{\varepsilon}$ and $\boldsymbol{\delta}\boldsymbol{\varepsilon}'$ be points in $\mathfrak{L}^\natural \cap \mathcal{H}$. By Theorem 4.1, $\boldsymbol{\varepsilon}$ and $\boldsymbol{\varepsilon}'$ belong $\mathfrak{O}(f)^\times$. Hence, they are totally positive by the assumption of the Theorem. Thus, $\boldsymbol{\delta}\boldsymbol{\varepsilon}$ and $\boldsymbol{\delta}\boldsymbol{\varepsilon}'$ have the same signature distribution. Therefore, they are on the same branch, say $\mathcal{H}^J$. We now see $\phi(\boldsymbol{\delta}\boldsymbol{\varepsilon}), \phi(\boldsymbol{\delta}\boldsymbol{\varepsilon}') \in \mathcal{C}^J_{J+1} \cup \mathcal{C}^J_{J+2}$. Since J is determined by one particular point $\boldsymbol{\delta}\boldsymbol{\varepsilon}$, all points of $\phi(\mathfrak{L}^\natural \cap \mathcal{H})$ lie on the two pieces $\mathcal{C}^J_{J+1}$ and $\mathcal{C}^J_{J+2}$.

Now, the proof of Theorem 1.1 implies the assertion. □

PROOF of Theorem 1.3. By Theorem 4.2, the number $\#(\mathfrak{L}^\natural \cap \mathcal{H}_k)$ is independent of k. Hence, the assertion of the theorem is reduced to the assertion $\#(\mathfrak{L}^\natural \cap \mathcal{H}_3) \leq 1$.

For contradiction, suppose $\mathfrak{L}^\natural \cap \mathcal{H}_3$ have two distinct points $\boldsymbol{z}$ and $\boldsymbol{z}'$. Set $r = \|\phi(\boldsymbol{z})\|$, $t = t(\phi(\boldsymbol{z}))$ and $t' = t(\phi(\boldsymbol{z}'))$. Without loss of generality, we assume $t' \geq t$.

By Theorem 4.2 and Lemma 6.2, $\boldsymbol{\delta}^{1-\sigma}$ is a unit. Hence, the second assertion

$$\frac{t'}{\operatorname{disc}(\mathfrak{M})} < 5.04 \cdot 10^4$$

of Theorem 7.2 is valid, where we set $\mathfrak{M} = \mathbf{Z}\phi(\boldsymbol{\delta}) + \mathfrak{E}(\mathfrak{K})$.

Substituting the lower bound of Theorem 5.6, we get

$$\frac{\sqrt{2}\exp\left(\sqrt{6}\, t/2\right)}{1 + \exp\left(-2(t'-t)/\sqrt{6}\log 2\right)} \leq 5.04 \cdot 10^4.$$

The difference $t' - t$ in the denominator is estimated from below by $0.3e^{\sqrt{6}\, t/2} - t$. To see this, substitute $t' - t \geq 0$ in Theorem 5.6 and estimate $\operatorname{disc}(\mathfrak{M})$ by Lemmata 5.7 and 6.2.

Therefore, we get $t \leq 8.56$ and hence $r \leq 1.01 \cdot 8.56 < 8.65$. Thus, Lemma 5.8 implies $D \leq 2.56 \cdot 10^{18}$. The first assertion is established.

If the signature rank of $\mathfrak{O}(f)^\times$ is less than 3, at most 2 branches of $\mathcal{H}$, $\mathcal{H}^J$ and $\mathcal{H}^{J+1}$, say, can have a lattice point of $\mathfrak{L}^\natural$ (as it is proved in

the same way as Theorem 1.2). Hence, one set $\mathfrak{L}^\natural \cap \mathcal{H}^{J+2}$ is empty. This implies $\#\left(\mathfrak{L}^\natural \cap \mathcal{H}^j\right) = 0$ since the number $\#\left(\mathfrak{L}^\natural \cap \mathcal{H}^j\right)$ is independent of j. The second assertion of the theorem now follows immediately. □

PROOF of Theorem 1.4. Let $\varepsilon \in \mathcal{T}$. Let α and β form a $\mathbf{Z}$-basis of the module $\{\gamma \in \mathfrak{O} \mid \operatorname{tr}\gamma = 0\}$ Then, ε is written as $x\alpha + y\beta$ with a suitable pair of integers x and y. Let $\gamma \longmapsto \gamma_i$ $(i = 1, 2, 3)$ be the three real embeddings of $\mathfrak{O}$ in $\mathbf{R}$. Consider the cubic form $f(X, Y) = \prod_{i=1}^{3}(\alpha_i X + \beta_i Y)$. Then, the pair (x, y) is a representation of 1 by f. We now identify $\mathfrak{O}$ with its image in $\mathbf{R}^3$.

The covolume of $\mathbf{Z}\boldsymbol{\alpha} + \mathbf{Z}\boldsymbol{\beta}$ on the plane Π is $\sqrt{3^c D(\mathfrak{O})}$ with $c = 1$ or -1 according as $\mathfrak{O}$ contains an element of trace 1 or not. Hence, $\boldsymbol{\alpha} \times \boldsymbol{\beta} = \sqrt{3^c D(\mathfrak{O})}\,\mathbf{1}$. Thus, we have $\boldsymbol{\delta} = \boldsymbol{\delta}(f) = \boldsymbol{\alpha} \times \boldsymbol{\beta}$ and $D = D(f) = 3^{3c} D(\mathfrak{O})^3$.

Therefore, Lemma 5.7 implies

$$\|\phi(\boldsymbol{\delta}\boldsymbol{\varepsilon})\| = \|\log \boldsymbol{\varepsilon}\| \geq \frac{1}{2\sqrt{2}} \log \frac{3^{-3} D(\mathfrak{O})^3}{P(D)} \tag{43}$$

Let $\boldsymbol{\varepsilon}$ and $\boldsymbol{\varepsilon}'$ be distinct elements of $\mathcal{T}$. Suppose their image under log lie on the same branch $\mathcal{C}_k$ of $\mathcal{C}$. Put $r = \|\log \boldsymbol{\varepsilon}\|$, $t = t(\log \boldsymbol{\varepsilon})$ and $t' = t(\log \boldsymbol{\varepsilon}')$. Without loss of generality, we assume $t' \geq t$.

Then, we have

$$\frac{t'}{\operatorname{disc}(\mathfrak{E}(\mathfrak{K}))} < 3.54 \cdot 10^5$$

by Theorem 7.2, where we set $\mathfrak{K} = \mathfrak{K}(f)$. On the other hand, we have

$$t' \geq \frac{\sqrt{2}\operatorname{disc}(\mathfrak{E}(\mathfrak{K}))\exp\left(\sqrt{6}\,t/2\right)}{1 + \exp\left(-2(t'-t)/\sqrt{6}\log 2\right)}$$

by Theorem 5.6. Here, the difference $t' - t$ in the denominator is estimated from below by $0.3e^{\sqrt{6}\,t/2} - t$. To see this, substitute $t' - t \geq 0$ in Theorem 5.6 and estimate $\operatorname{disc}(\mathfrak{E}(\mathfrak{K}))$ by Lemma 5.7.

Therefore, the two inequalities imply $t \leq 10.2$ and hence $r \leq 1.01 \cdot 10.2 < 10.4$.

Substituting this in (43), we get $3^{-3} D(\mathfrak{O})^3 = D < 1.36 \cdot 10^{22}$ and hence $D \leq 7.17 \cdot 10^7$. This contradicts the assumption of the theorem on D.

The contradiction proves $\#\left(\mathfrak{L}^\natural \cap \mathcal{H}_k\right) = \#\left(\phi(\mathfrak{L}^\natural) \cap \mathcal{C}_k\right) \leq 1$ for $k = 1, 2, 3$. The first assertion is now obvious.

If $\boldsymbol{\varepsilon} \in \mathcal{T}$, there is an index j such that $\varepsilon_j > 0 > \varepsilon_{j+1}, \varepsilon_{j+2}$. Thus, $\log \boldsymbol{\varepsilon} \in \mathcal{C}^j$. This means that the branch $\mathcal{C}^j$ is determined by the signature of $\boldsymbol{\varepsilon}$. Each signature except $(+,+,+)$ corresponds to one branch of the curve $\mathcal{C}$.

If the signature rank of $\mathfrak{O}^\times$ is 2, the image of $\boldsymbol{\varepsilon} \in \mathcal{T}$ under log lie on a branch $\mathcal{C}^J$ determined by $\mathfrak{O}^\times$. Hence, those images lie on the two pieces $\mathcal{C}^{J+1}$ and $\mathcal{C}^{J+2}$. The second assertion of the theorem now follows immediately.

The third assertion of the theorem is trivial.

We now show the last assertion. Assume $\boldsymbol{\varepsilon} \in \mathcal{T}$. Let $\boldsymbol{\omega} = \boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^\sigma$. Then, we have $\boldsymbol{\omega} - \boldsymbol{\omega}^\sigma = -3\boldsymbol{\varepsilon}$. Hence, $D(\mathbf{Z}[\boldsymbol{\omega}]) = N_{\mathfrak{K}/\mathbf{Q}}(-3\boldsymbol{\varepsilon})^2 = 3^6$. Thus, $\mathfrak{K}$ is the cyclic cubic field of discriminant 81. Let $\boldsymbol{\theta}$ be the root of $X^3 + 3X + 1$ and write $\boldsymbol{\varepsilon} = x\boldsymbol{\theta} + y\boldsymbol{\theta}^\sigma$ with $x, y \in \mathbf{Z}$. By taking norm, we get $f_4(x,y) = x^3 + 3x^2y - 6xy^2 + y^3 = 1$. Its solutions are $(x,y) = (1,0), (0,1), (-1,-1)$ and correspond to $\boldsymbol{\theta}$, $\boldsymbol{\theta}^\sigma$ and $\boldsymbol{\theta}^{\sigma^2}$ (see [17, 24]). Therefore, $D(\mathfrak{O}) = 81$ and $\#\mathcal{T} = 3$. □

9. Figures

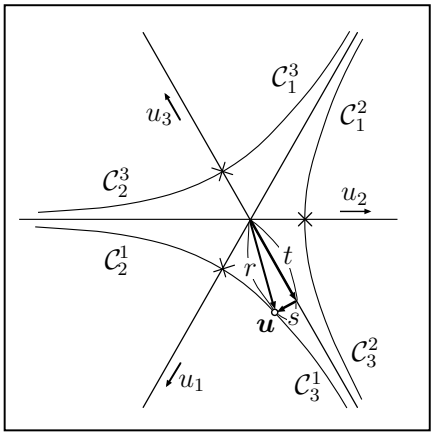

Figure 1: The Continuous Curve $\mathcal{C}$.

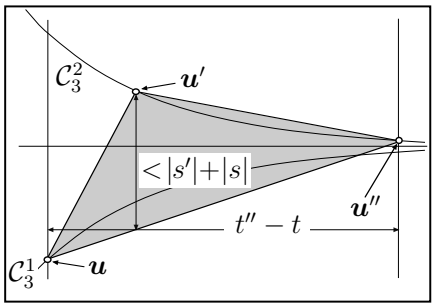

Figure 2: Area Estimate.

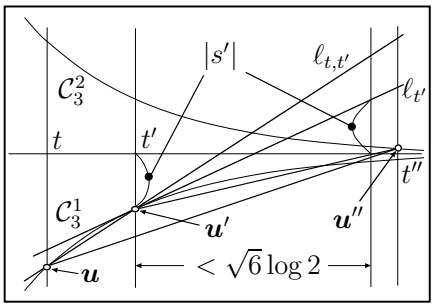

Figure 3: Triangle across Axis.

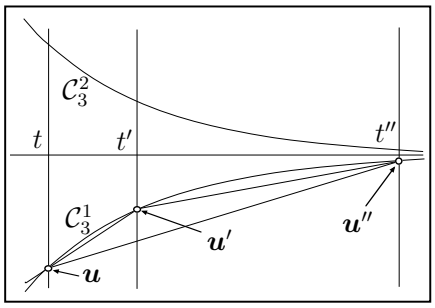

Figure 4: Triangle in One Side.

References

[1] M. A. Bennett, On the representation of unity by binary cubic forms, *Trans. Amer. Math. Soc.* **353** (2001), 1507–1534.

[2] Y. Bugeaud, Bornes effectives pour les solutions des équations en S-unités et des équations de Thue–Mahler, *J. Number Theory* **71** (1998), 227–244.

[3] Y. Bugeaud and K. Győry, Bounds for the solutions of Thue–Mahler equations and norm form equations, *Acta Arith.* **74** (1996), 273–292.

[4] J. W. S. Cassels, An Introduction to the Geometry of Numbers, *Springer Verlag*, 1997, (reprint of the 1971 edition).

[5] H. Cohen, Advanced Topics in Computational Number Theory, (G.T.M. 193), *Springer-Verlag*, 2000.

[6] B. N. Delone, On the number of representations of a number by a cubic binary form with negative discriminant, *Izv. Akad. Nauk SSSR* (6) **16** (1922), 253–272, (*in Russian*), translation in German: "Über die Darstellung der Zahlen durch die binären kubischen Formen von negativer Diskriminante", Math. Z. **31** (1930), 1–26.

[7] B. N. Delone and D. K. Faddeev, The Theory of Irrationalities of the Third Degree, Amer. Math. Soc. Transl. of Math. Monographs 10 Providence, *USA: Amer. Math. Soc.* (1964), (Bell & Howell Information and Learning Books on Demand, Ann Arbor http://www.umi.com).

[8] G. L. Dirichlet, Über die Reduktion der positiven quadratischen Formen mit drei unbestimmten ganzen Zahlen, Mathematische Werke I, *Chelsea Publishing Co.*, 1969.

[9] J.-H. Evertse, On the representation of integers by binary cubic forms of positive discriminant, *Invent. Math.* **73** (1983), 117–138.

[10] J.-H. Evertse, The number of solutions of the Thue–Mahler equation, *J. reine angew. Math.* **482** (1997), 121–149.

[11] D. Hilbert, Theory of Algebraic Invariants, *Cambridge Univ. Press*, 1993.

[12] M. Laurent, M. Mignotte and Y. Nesterenko, Formes linéaires en deux logarithmes et déterminants d'interpolation, *J. Number Theory* **55** (1995), 285–321.

[13] F. Lippok, On the representation of 1 by binary cubic forms of positive discriminant, *J. Symbolic Computation* **15** (1993), 297–313.

[14] S. Louboutin, Explicit upper bounds for residues of Dedekind zeta functions and values of L-functions at $s = 1$, and explicit lower bounds for relative class numbers of CM-fields, *Canad. Math. J.* **53** (2001), 1194–1222.

[15] E. M. Matveev, An explicit lower bound for a homogeneous rational linear form in logarithms of algebraic numbers, *Izv. Ross. Akad. Nauk Ser. Mat.* **62** (1998), 81–136, translation in Izv. Math. **62** (1998), 723–772 (in *Russian)*).

[16] E. M. Matveev, An explicit lower bound for a homogeneous rational linear form in logarithms of algebraic numbers. II, *Izv. Ross. Akad. Nauk Ser. Mat.* **64** (2000), 125–180, translation in Izv. Math. **64** (2000), 1217–1269 (in *Russian*).

[17] M. Mignotte, Verification of a conjecture of E. Thomas, *J. Number Theory* **44** (1993), 172–177.

[18] T. Nagell, Darstellung ganzer Zahlen durch binäre kubische Formen mit negativer Diskriminante, *Math. Z.* **28** (1928), 10–29.

[19] A. Pethő, On the representation of 1 by binary cubic forms with positive discriminant, (Springer LNM 1380), *Number Theory, Ulm*, 1987, 185–196.

[20] M. Pohst, Regulatorabschätzungen für total reelle algebraische Zahlkörper, *J. Number Theory* **9** (1977), 459–492.

[21] W. Scharlau and H. Opolka, From Fermat to Minkowski — Lectures on the Theory of Numbers and its Historical Development, *Springer Verlag*, 1985.

[22] C. L. SIEGEL, Über einige Anwendungen diophantischer Approximationen, *Abh. preuss. Akad. Wiss. Phys. Math. Kl.*, no. 1 (1929), (Gesammelte Abhandlungen I, 209–266, Springer Verlag 1966.).

[23] C. L. SIEGEL, Einige Erläuterungen zu Thues Untersuchungen über Annäherungswere algebraischer Zahlen und diophantische Gleichungen, *Nachr. Akad. Wiss. Göttingen II Math.-Phys. Kl.*, no. 8 (1970), (Gesammelte Abhandlungen IV, 140–146, Springer Verlag 1979.).

[24] E. THOMAS, Complete solutions to a family of cubic Diophantine equations, *J. Number Theory* **34** (1990), 235–250.

[25] A. THUE, Über Annäherungenswerte algebraischer Zahlen, *J. reine angew. Math.* **135** (1909), 284–305.

[26] B. VALLÉE, Une pprache géométrique des algorithmes de réduction en petite dimension, *Thesis, Univ. of Caen*, 1986.

[27] L. C. WASHINGTON, Introduction to Cyclotomic Fields, (2nd edn) (G.T.M. 83), *Springer-Verlag*, 1997.

RYOTARO OKAZAKI
DOSHISHA UNIVERSITY
DEPARTMENT MATHEMATICS
KYOTANABE, KYOTO 610-0394
JAPAN

E-mail: rokazaki@dd.iij4u.or.jp

(Received December 27, 2000; revised July 27, 2002)