ANNALI DELLA
SCUOLA NORMALE SUPERIORE DI PISA
Classe di Scienze

E. BOMBIERI
A. J. VAN DER POORTEN
J. D. VAALER

Effective measures of irrationality for cubic extensions of number fields

Annali della Scuola Normale Superiore di Pisa, Classe di Scienze 4e série, tome 23, nº 2 (1996), p. 211-248

<http://www.numdam.org/item?id=ASNSP_1996_4_23_2_211_0>

Effective Measures of Irrationality for Cubic Extensions of Number Fields

E. BOMBIERI * - A. J. VAN DER POORTEN ** - J. D. VAALER ***

1. – Introduction

A basic problem of Diophantine approximation is to give effective measures of irrationality for algebraic numbers. We formulate this problem as follows. Let k be an algebraic number field, v a place of k, and $|\ |_v$ the associated absolute value on the completion k_v. Here we normalize $|\ |_v$ exactly as in [5], [7] or [8]. Let K be a finite extension of k having degree $r \geq 2$. We assume that K has at least one embedding in the completion k_v and then we identify K with a fixed embedding of K in k_v. If β belongs to k we write $H(\boldsymbol{\beta})$ for the absolute height of the corresponding projective point $\boldsymbol{\beta} = \begin{bmatrix} \beta \\ 1 \end{bmatrix}$ in $\mathbb{P}^1(k)$.

We warn the reader here that this height $H(\ )$ differs from the Weil height $H_{\text{Weil}}(\ )$ because we use the ℓ^2-norm, rather than the ℓ^∞-norm, for the places at ∞. In any case, we have

$$H_{\text{Weil}}(\boldsymbol{\beta}) \leqq H(\boldsymbol{\beta}) \leqq 2H_{\text{Weil}}(\boldsymbol{\beta}),$$

so that the two heights are comparable. However, the use of the projective height $H(\ )$ leads to neater and sharper results and suggests itself naturally in this context.

If α belongs to k_v but not to k then Dirichlet's theorem asserts the existence of a positive constant C_1 depending on α and infinitely many distinct β in k such that

$$|\alpha - \beta|_v \leq C_1 H(\boldsymbol{\beta})^{-2}. \tag{1.1}$$

By a measure of irrationality for α with respect to v and k we understand a positive number μ for which an inequality of the type

$$|\alpha - \beta|_v \geq C_2 H(\boldsymbol{\beta})^{-\mu}$$

* Research of the first author was supported by Macquarie University.
** Research of the second author was supported partially by grants from the Australian Research Council and Macquarie University, and a research agreement with Digital Equipment Corporation.
*** Research of the third author was supported by a grant from the National Science Foundation, DMS-8701396.
Pervenuto alla Redazione il 26 giugno 1995.

is valid for all β in k with $H(\beta) \geq C_3$. Here C_2 and C_3 are positive constants independent of β. If the constants C_2 and C_3 can be computed from knowledge of α and μ then the measure of irrationality is said to be effective. In view of (1.1) any measure of irrationality μ must satisfy $\mu \geq 2$.

If α belongs to K but not to k then the fundamental theorem of Roth states that $2 + \varepsilon$ is a measure of irrationality for every $\varepsilon > 0$. Unfortunately, all known proofs of Roth's theorem result in constants C_2 and C_3 which are not effectively computable. Alternatively, Roth's theorem asserts that

$$- \log |\alpha - \beta|_v \leq 2 \log H(\beta) + o(\log H(\beta))$$

as $H(\beta) \to \infty$.

Our knowledge of effective measures of irrationality for algebraic numbers is still rather limited. To begin with there is the simple but important Liouville bound

$$|\alpha - \beta|_v \geq H(\alpha)^{-r} H(\beta)^{-r},$$

which is easily obtained from the product formula. This shows that r is an effective measure of irrationality for any α in K but not in k.

There are essentially three methods for obtaining improvements in the Liouville bound in which all constants can be effectively determined. Historically the first of these is in the work of Baker [1], [2], although a forerunner of it is implicit in the early work of Thue [17]. This method is based on the explicit construction of Padé approximations to algebraic functions. At present it provides good effective measures of irrationality for special numbers, in particular for certain cubic irrationalities and for r-th roots of certain rational numbers. See Chudnovsky [10].

The second method is based on Baker's theory of linear forms in logarithms of algebraic numbers. The first general improvement on Liouville type bounds was obtained by Baker [3]. This had far reaching consequences, allowing for example the effective solution of the general Thue equation $F(x, y) = m$, where $F(x, y) \in \mathbb{Z}[x, y]$ is a binary form with at least three distinct linear factors in $\mathbb{C}[x, y]$.

In the form obtained by Feldman [11], the method yields an effective measure of irrationality for α when $r \geq 3$ which is slightly smaller than r. Although the improvement on the Liouville bound is usually very small, this method applies to all algebraic numbers α. For special numbers such as cube roots of positive integers, very explicit bounds have been given recently by Baker and Stewart [4].

The third method has been developed by Bombieri [5] and by Bombieri and Mueller [7]. This approach combines elements of the original non-effective methods of Thue and Siegel with an important result of Dyson for establishing the non-vanishing of the auxiliary polynomial in two variables. This method is not unrelated to the Padé approximation method, and it may be considered as a two-variable Padé approximation method in which the approximants do not quite reach the maximum order of approximation one can achieve. In some favorable cases, this method leads to the construction of triples (K, k, v) for which an effective measure of irrationality much smaller than r can be established. An important feature of the method is that the measure of irrationality which is obtained holds for all α in K but not in k.

We define $\mu_v(\alpha)$ to be the greatest lower bound for effective measures of irrationality of α, relative to the absolute value v.

In the present paper we formulate and prove a new equivariant version of the Thue-Siegel principle which applies only to special numbers in the Galois closure of K/k. In order to define these special numbers it is necessary for the Galois group to have a faithful, projective representation in $\mathrm{PGL}(2, k)$. We study here in detail the representation of the symmetric group S_3 in $\mathrm{PGL}(2, \mathbb{Q})$ generated by the classical substitutions $z \mapsto 1 - z$ and $z \mapsto 1/z$. If K/k is a cubic extension we obtain by our methods very good effective measures of irrationality which approach the limiting value 2 in favorable cases. Suppose, for example, that m is a large positive integer and $K \subseteq \mathbb{R}$ is the cubic extension of $\mathbb{Q}$ formed by adjoining the unique real root of $x^3 + mx + 1$. Then we obtain an effective measure of irrationality $\mu_v(\alpha) \le 2 + 10(\log m)^{-1/3}$ for all generators α of $K/\mathbb{Q}$.

Another interesting application of the new equivariant form of the Thue-Siegel principle in this paper has been worked out in detail in Bombieri [6]. The group in question now is the cyclic group of order r, acting by multiplication by r-th roots of unity. The results obtained there are sufficiently strong to imply, by known reductions going back to Siegel and Baker, a new proof of the Baker-Feldman improvement of Liouville's inequality, valid for every algebraic number.

The method used in the present work originates in the paper of Bombieri [5] and continues to use Dyson's Lemma in an essential way. We also introduce some new technical devices. Let K be a cubic extension of k, identify K with an embedding of K in k_v, and then fix an embedding of the Galois closure L of K/k in $\bar{k}_v$. We determine a faithful, projective representation $\sigma \to P_\sigma$ of $G = \mathrm{Gal}(L/k)$ into $\mathrm{PGL}(2, \mathbb{Q})$ and then define a subset $\Lambda(P) \subseteq \mathbb{P}^1(L)$ by

$$\Lambda(P) = \left\{\boldsymbol{\lambda} \in \mathbb{P}^1(L) : \sigma^{-1}(\boldsymbol{\lambda}) = P_\sigma \boldsymbol{\lambda} \ \text{ for all } \sigma \text{ in } G\right\}.$$

Our next step is to consider how well the point $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ in $\mathbb{P}^1(L)$ can be approximated by points $\boldsymbol{\lambda}$ in $\Lambda(P)$. For this purpose we introduce a suitable metric δ_v on $\mathbb{P}^1(\bar{k}_v)$ which is induced by $|\ |_v$ on k_v. In Theorem 4 we show that $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ cannot have two excellent approximations $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ in $\Lambda(P)$ for which $H(\boldsymbol{\lambda}_1)$ and $H(\boldsymbol{\lambda}_2)$ are both large. This is our new form of the Thue-Siegel principle for special points. It is expressed as an inequality in which all constants are given explicitly and we refer to it now as the Thue-Siegel inequality. The proof of Theorem 4 is given in Sections 4 through 8. During the course of the proof we take advantage of the fact that the action of G on points in $\Lambda(P)$ is strictly controlled.

It remains then to demonstrate how effective measures of irrationality can be obtained from Theorem 4. It turns out to be very convenient to reformulate the basic problem of lower bounds for $|\alpha - \beta|_v$ in terms of the projective metric δ_v. Thus we let $\boldsymbol{\alpha} = \begin{bmatrix} \alpha \\ 1 \end{bmatrix}$ in $\mathbb{P}^1(K)$, $\boldsymbol{\beta} = \begin{bmatrix} \beta \\ 1 \end{bmatrix}$ in $\mathbb{P}^1(k)$, and provide lower bounds for $\delta_v(\boldsymbol{\alpha}, \boldsymbol{\beta})$. Since $\delta_v(\boldsymbol{\alpha}, \boldsymbol{\beta}) \leq |\alpha - \beta|_v$, our lower bounds apply automatically to $|\alpha - \beta|_v$.

In Theorem 5 we show that if $\boldsymbol{\lambda}_1$ in $\Lambda(P)$ is an excellent approximation to $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ then

$$\mu_v^*(\boldsymbol{\lambda}_1) = \left(- \log \delta_v \left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \boldsymbol{\lambda}_1 \right) \right)^{-1} \left(6 \log\{8H(\boldsymbol{\lambda}_1)\} + 19 \left(\log\{8H(\boldsymbol{\lambda}_1)\} \right)^{2/3} \right)$$

is an effective measure of irrationality for each generator α of the cubic extension K/k. Of course $\boldsymbol{\lambda}_1$ qualifies as an excellent approximation to $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ precisely when $\mu_v^*(\boldsymbol{\lambda}_1) < 3$.

In Section 9 we assume that $f(x) = x^3 + px + q$, $p \neq 0$, is irreducible in $k[x]$ and that the smallest root of f in $\bar{k}_v$ generates K/k. We use the roots of f to construct a point $\boldsymbol{\lambda}_1$ in $\Lambda(P)$. Then we estimate the measure of irrationality $\mu_v^*(\boldsymbol{\lambda}_1)$ in terms of p and q. This allows us to construct cubic extensions of k which have an effective measure of irrationality close to 2 by making a suitable choice of p and q. A precise upper bound for $\mu_v^*(\boldsymbol{\lambda}_1)$ with explicit constants is given in Theorem 17. Suppose, for example, that K is generated by the unique real root of $x^3 + px + q = 0$, where $p > 0$ and q are relatively prime integers. We get a non-trivial effective measure of irrationality for each generator of $K/\mathbb{Q}$ as soon as $p > C(\varepsilon)|q|^{2+\varepsilon}$, for a suitable effective $C(\varepsilon)$ and any $\varepsilon > 0$. The measure of irrationality is

$$\mu_\infty^*(\boldsymbol{\lambda}_1) = \frac{2 \log(p^3)}{\log(p^3/q^2)} + O\left((\log p)^{-1/3} \right) .$$

In our proof of the Thue-Siegel inequality it is necessary to construct an auxiliary polynomial with prescribed vanishing and small coefficients. We note that this step has now been disconnected from any conditions involving irrational algebraic numbers. We require that the polynomial be bihomogeneous with integer coefficients, and that it vanish to high order at the points $\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right)$, $\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)$, and $\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right)$ in $\mathbb{P}^1 \times \mathbb{P}^1$.

In this respect, our auxiliary polynomials are universal for the problem of effective approximation of cubic irrationals. Here we use a simple form of Siegel's lemma to construct them. We remark, however, that a direct construction may lead to a much better understanding of these universal polynomials and thereby drastically improve and modify the application of our method.

2. – Inequalities for the projective metric

Let k be an algebraic number field, v a place of k and k_v the completion of k with respect to v. We will use two absolute values $|\ |_v$ and $\|\ \|_v$ from v which are determined exactly as in our previous papers, such as [5], [7] or [8]. In particular we have $|x|_v = \|x\|_v^{d_v/d}$ for all x in k_v, where $d = [k : \mathbb{Q}]$ and $d_v = [k_v : \mathbb{Q}_v]$. These absolute values have unique extensions to Ω_v, the completion of an algebraic closure of k_v. We extend $|\ |_v$ to a norm on finite dimensional vector spaces over Ω_v as follows. If

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_N \end{pmatrix}$$

is a column vector in Ω_v^N we set

$$\|\mathbf{x}\|_v = \begin{cases} \max\{\|x_n\|_v : 1 \le n \le N\} & \text{if } v \nmid \infty \\ \left\{\sum_{n=1}^{N} \|x_n\|_v^2\right\}^{1/2} & \text{if } v|\infty \end{cases} \tag{2.1}$$

and $|\mathbf{x}|_v = \|\mathbf{x}\|_v^{d_v/d}$ in both cases. Let $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_N$ denote the standard basis vectors in Ω_v^N. For each subset $I \subseteq \{1, 2, \dots, N\}$ let $\mathbf{e}_I$ be the corresponding standard basis vector in the exterior algebra $\bigwedge(\Omega_v^N)$. We recall that this is a graded algebra

$$\bigwedge(\Omega_v^N) = \sum_{n=0}^{\infty} \bigwedge\nolimits_n(\Omega_v^N),$$

in which each subspace $\bigwedge_n(\Omega_v^N)$ has dimension $\binom{N}{n}$ and basis vectors $\{\mathbf{e}_I : |I| = n\}$. As usual we identify Ω_v^N with the subspace $\bigwedge_1(\Omega_v^N)$ so that

$$\mathbf{e}_I = \mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \cdots \wedge \mathbf{e}_{i_n}$$

whenever $I = \{i_1 < i_2 < \cdots < i_n\} \subseteq \{1, 2, \dots, N\}$ is not empty. Then $|\ |_v$ extends to $\bigwedge(\Omega_v^N)$ by applying (2.1) to the basis $\{\mathbf{e}_I : I \subseteq \{1, 2, \dots, N\}\}$.

We identify elements of $\mathrm{Hom}(\Omega_v^N, \Omega_v^M)$ with $M \times N$ matrices over Ω_v. Then we extend $|\ |_v$ to such matrices A by setting

$$|A|_v = \sup\{|A\mathbf{x}|_v : \mathbf{x} \in \Omega_v^N, |\mathbf{x}|_v \le 1\}.$$

If $v \nmid \infty$ and $A = (a_{mn})$ we find that

$$|A|_v = \max\{|a_{mn}|_v : 1 \le m \le M, 1 \le n \le N\}. \tag{2.2}$$

If $v \mid \infty$ let A^* denote the complex conjugate transpose of A and let

$$0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N$$

denote the eigenvalues of the positive semi-definite matrix A^*A. It follows that

$$|A|_v = \lambda_N^{d_v/2d}. \tag{2.3}$$

We define a map

$$\eta_v : \mathrm{GL}(N, \Omega_v) \to [1, \infty)$$

by $\eta_v(A) = |A|_v |A^{-1}|_v$. Obviously $\eta_v(aA) = \eta_v(A)$ for all $a \neq 0$ in Ω_v and therefore η_v is well defined as a map

$$\eta_v : \mathrm{PGL}(N, \Omega_v) \to [1, \infty).$$

We note that $\eta_v(A) = 1$ if and only if the identity $|A\mathbf{x}|_v = |A|_v |\mathbf{x}|_v$ holds for all $\mathbf{x}$ in Ω_v^N. In the special case $N = 2$ the identities (2.2) and (2.3) can be used to verify the formula

$$\eta_v(A) = |\det A|_v^{-1} |A|_v^2. \tag{2.4}$$

We define a second map

$$\delta_v : \Omega_v^N \times \Omega_v^N \to [0, 1]$$

by

$$\delta_v(\mathbf{x}, \mathbf{y}) = \frac{|\mathbf{x} \wedge \mathbf{y}|_v}{|\mathbf{x}|_v |\mathbf{y}|_v}. \tag{2.5}$$

Since $\delta_v(a\mathbf{x}, b\mathbf{y}) = \delta_v(\mathbf{x}, \mathbf{y})$ for all $a \neq 0$ and $b \neq 0$ in Ω_v, it follows that δ_v is well defined as a map

$$\delta_v : \mathbb{P}^{N-1}(\Omega_v) \times \mathbb{P}^{N-1}(\Omega_v) \to [0, 1].$$

It can be shown that δ_v is a metric on $\mathbb{P}^{N-1}(\Omega_v)$ and the induced metric topology coincides with the quotient topology determined by the norm $|\ |_v$ on Ω_v^N. This observation, but with a different definition for δ_v, is due to Néron [12]. The fact that our definition is equivalent to Néron's was established by Rumely [15]. If $N = 2$ and $v \mid \infty$ then $\Omega_v \cong \mathbb{C}$ and $\delta_v(\mathbf{x}, \mathbf{y})$ is the cordal distance between points $\mathbf{x}$ and $\mathbf{y}$ on the Riemann sphere $\mathbb{P}^1(\mathbb{C})$ normalized so that the diameter of the sphere is 1.

In case $N = 2$ the maps η_v and δ_v satisfy the basic inequality

$$\eta_v(A)^{-1} \delta_v(\mathbf{x}, \mathbf{y}) \leq \delta_v(A\mathbf{x}, A\mathbf{y}) \leq \eta_v(A) \delta_v(\mathbf{x}, \mathbf{y}) \tag{2.6}$$

for all A in $\mathrm{PGL}(2, \Omega_v)$ and $\mathbf{x}, \mathbf{y}$ in $\mathbb{P}^1(\Omega_v)$. To verify the inequality on the left of (2.6) we note that

$$\begin{aligned}\delta_v(A\mathbf{x}, A\mathbf{y}) &= |A\mathbf{x}|_v^{-1}|A\mathbf{y}|_v^{-1}|\det A|_v|\mathbf{x}\wedge\mathbf{y}|_v \\ &\geq |A|_v^{-2}|\det A|_v|\mathbf{x}|_v^{-1}|\mathbf{y}|_v^{-1}|\mathbf{x}\wedge\mathbf{y}|_v \\ &= \eta_v(A)^{-1}\delta_v(\mathbf{x},\mathbf{y})\,.\end{aligned}$$

Then the inequality on the right of (2.6) follows since $\eta_v(A^{-1}) = \eta_v(A)$. This result continues to be true if $N \geq 3$, as shown by K.-K. Choi [9]; a proof of the weaker inequality with $\eta_v^{\pm 2}(A)$ in place of $\eta_v^{\pm 1}(A)$ can be found in S. Tyler [18].

If β occurs in k^N we define its height by

$$H(\beta) = \prod_v |\beta|_v\,, \tag{2.7}$$

where the product extends over all places v of k. By the product formula this height is well defined on $\mathbb{P}^{N-1}(k)$.

We mention here, as an illustration of what we have done so far, two theorems which show how classical results can be formulated in our setting.

If α belongs to $\mathbb{P}^{N-1}(k_v)$ for some place v of k then we may ask how well α can be approximated by points β in $\mathbb{P}^{N-1}(k)$ with $H(\beta)$ bounded by a suitable parameter. In order to state a reasonably precise result let

$$c_k(N) = 2|\Delta_k|^{1/2d} \prod_{w\,|\,\infty} r_w(N)^{d_w/d}\,, \tag{2.8}$$

where Δ_k is the discriminant of k and

$$r_w(N) = \begin{cases} \pi^{-1/2}\left\{\Gamma\left(\frac{1}{2}N+1\right)\right\}^{1/N} & \text{if } w \text{ is real} \\ (2\pi)^{-1/2}\{\Gamma(N+1)\}^{1/2N} & \text{if } w \text{ is complex.}\end{cases} \tag{2.9}$$

Then we have the following projective form of Dirichlet's theorem.

THEOREM 1. *Let α belong to $\mathbb{P}^{N-1}(k_v)$, τ belong to k_v and assume that $1 \leq |\tau|_v$. Then there exists β in $\mathbb{P}^{N-1}(k)$ such that*

(i) $H(\beta) \leq c_k(N)|\tau|_v^{N-1}$,

(ii) $\delta_v(\alpha,\beta) \leq c_k(N)\{|\tau|_v H(\beta)\}^{-1}$.

If α belongs to $\mathbb{P}^{N-1}(k_v)$ but not to $\mathbb{P}^{N-1}(k)$ then (i) and (ii) can be combined to establish the existence of infinitely many distinct β in $\mathbb{P}^{N-1}(k)$ such that

$$\delta_v(\alpha,\beta) \leq c_k(N)^{N/(N-1)} H(\beta)^{-N/(N-1)}\,. \tag{2.10}$$

This formulation of Dirichlet's theorem is similar to, but slightly more precise than, that obtained by Schmidt [16], Theorem 1. A proof has been given by S. Tyler [18], Theorem 3.2.

Next we consider lower bounds for the projective distance $\delta_v(\boldsymbol{\alpha}, \boldsymbol{\beta})$ and in doing so we restrict our attention to the projective line $\mathbb{P}^1(k_v)$. Suppose then that

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha \\ 1 \end{bmatrix}$$

in homogeneous coordinates with α algebraic over k of degree $r \geq 2$. As $K = k(\alpha)$ is embedded in k_v there exists a place $\tilde{w}$ of K with $\tilde{w} \mid v$, $[K_{\tilde{w}} : k_v] = 1$, and $|\alpha|_{\tilde{w}}^r = |\alpha|_v$. It follows that

$$\delta_{\tilde{w}}(\mathbf{x}, \mathbf{y})^r = \delta_v(\mathbf{x}, \mathbf{y})$$

for all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{P}^1(K_{\tilde{w}}) = \mathbb{P}^1(k_v)$. If $\boldsymbol{\beta}$ occurs in $\mathbb{P}^1(k) \subseteq \mathbb{P}^1(K)$ then

$$\delta_v(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \delta_{\tilde{w}}(\boldsymbol{\alpha}, \boldsymbol{\beta})^r \geq \left\{ \prod_w \delta_w(\boldsymbol{\alpha}, \boldsymbol{\beta}) \right\}^r, \tag{2.11}$$

where the product on the right of (2.11) extends over all places w of K. Since $\boldsymbol{\beta} \neq \boldsymbol{\alpha}$ we conclude that

$$\begin{aligned} \delta_v(\boldsymbol{\alpha}, \boldsymbol{\beta}) &\geq \left\{ \prod_w |\boldsymbol{\alpha}|_w^{-1} |\boldsymbol{\beta}|_w^{-1} |\boldsymbol{\alpha} \wedge \boldsymbol{\beta}|_w \right\}^r \\ &= \{H(\boldsymbol{\alpha}) H(\boldsymbol{\beta})\}^{-r}. \end{aligned} \tag{2.12}$$

This is the Liouville lower bound for the projective metric. If $r = 2$ then (2.10) with $N = 2$ shows that this bound is essentially sharp.

3. – The Thue-Siegel inequality

We assume throughout this section that α in k_v is algebraic over k of degree 3, that $K = k(\alpha) \subseteq k_v$, and that α' and α'' are conjugates of α in $\bar{k}_v$. Let $L = k(\alpha, \alpha', \alpha'') \subseteq \bar{k}_v$ denote the Galois closure of the extension K/k. We write

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha \\ 1 \end{bmatrix}, \quad \boldsymbol{\alpha}' = \begin{bmatrix} \alpha' \\ 1 \end{bmatrix}, \quad \boldsymbol{\alpha}'' = \begin{bmatrix} \alpha'' \\ 1 \end{bmatrix}$$

for the corresponding points in $\mathbb{P}^1(L)$. For each σ in $G = \mathrm{Gal}(L/k)$ there exists a unique element Q_σ in $\mathrm{PGL}(2, L)$ such that

$$\sigma(\boldsymbol{\alpha}) = Q_\sigma \boldsymbol{\alpha}, \quad \sigma(\boldsymbol{\alpha}') = Q_\sigma \boldsymbol{\alpha}', \quad \sigma(\boldsymbol{\alpha}'') = Q_\sigma \boldsymbol{\alpha}''.$$

Clearly $\sigma \to Q_\sigma$ is a faithful, projective representation of G in PGL(2, L). Also, there exists a unique element $\boldsymbol{\Phi}$ in PGL(2, L) such that

$$\text{(3.1)} \qquad \boldsymbol{\Phi\alpha} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{\Phi\alpha}' = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{\Phi\alpha}'' = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

It follows that the conjugate representation $\sigma' \to P_\sigma = \boldsymbol{\Phi} Q_\sigma \boldsymbol{\Phi}^{-1}$ acts to permute the elements of the set $\left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$. If G is noncyclic then $\{P_\sigma : \sigma \in G\} = \mathcal{G}$, where

$$\text{(3.2)} \quad \mathcal{G} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \right\}$$

is exactly the subgroup of PGL(2, $\mathbb{Q}$) which permutes the elements of the set $\left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$. If G is cyclic then

$$\{P_\sigma : \sigma \in G\} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \right\} \subseteq \mathcal{G}.$$

In particular, $\sigma \to P_\sigma$ is a faithful, projective representation of G in PGL(2, $\mathbb{Q}$).

LEMMA 2. *Let $\boldsymbol{\lambda}$ belong to $\mathbb{P}^1(L)$. Then the identity*

$$\sigma^{-1}(\boldsymbol{\lambda}) = P_\sigma \boldsymbol{\lambda}$$

holds for all σ in G if and only if $\boldsymbol{\Phi}^{-1}\boldsymbol{\lambda}$ occurs in $\mathbb{P}^1(k)$.

PROOF. If σ is in G then

$$\begin{aligned} \sigma^{-1}(\boldsymbol{\Phi})\boldsymbol{\alpha} &= \sigma^{-1}(\boldsymbol{\Phi})\sigma^{-1}(\sigma(\boldsymbol{\alpha})) = \sigma^{-1}\{\boldsymbol{\Phi}\sigma(\boldsymbol{\alpha})\} \\ &= \sigma^{-1}\{\boldsymbol{\Phi} Q_\sigma \boldsymbol{\Phi}^{-1}\boldsymbol{\Phi\alpha}\} = \sigma^{-1}\{P_\sigma \boldsymbol{\Phi\alpha}\}. \end{aligned}$$

But $P_\sigma\boldsymbol{\Phi\alpha}$ occurs in $\left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$ and it follows that $\sigma^{-1}(\boldsymbol{\Phi})\boldsymbol{\alpha} = P_\sigma\boldsymbol{\Phi\alpha}$. In a similar manner we find that $\sigma^{-1}(\boldsymbol{\Phi})\boldsymbol{\alpha}' = P_\sigma\boldsymbol{\Phi\alpha}'$ and $\sigma^{-1}(\boldsymbol{\Phi})\boldsymbol{\alpha}'' = P_\sigma\boldsymbol{\Phi\alpha}''$. This shows that $\sigma^{-1}(\boldsymbol{\Phi}) = P_\sigma\boldsymbol{\Phi}$ in PGL(2, L) for each σ in G.

If $\sigma^{-1}(\boldsymbol{\lambda}) = P_\sigma\boldsymbol{\lambda}$ for all σ in G then

$$\sigma^{-1}(\boldsymbol{\Phi}^{-1}\boldsymbol{\lambda}) = (\sigma^{-1}(\boldsymbol{\Phi}))^{-1}\sigma^{-1}(\boldsymbol{\lambda}) = (P_\sigma\boldsymbol{\Phi})^{-1}P_\sigma\boldsymbol{\lambda} = \boldsymbol{\Phi}^{-1}\boldsymbol{\lambda}.$$

Hence $\boldsymbol{\Phi}^{-1}\boldsymbol{\lambda}$ belongs to $\mathbb{P}^1(k)$. On the other hand, if $\boldsymbol{\Phi}^{-1}\boldsymbol{\lambda}$ occurs in $\mathbb{P}^1(k)$ then $\boldsymbol{\lambda} = \boldsymbol{\Phi\beta}$ with $\boldsymbol{\beta}$ in $\mathbb{P}^1(k)$. Thus we have

$$\sigma^{-1}(\boldsymbol{\lambda}) = \sigma^{-1}(\boldsymbol{\Phi})\boldsymbol{\beta} = P_\sigma\boldsymbol{\Phi\beta} = P_\sigma\boldsymbol{\lambda}$$

for all σ in G. This proves the lemma.

We define the subset

$$\Lambda(P) = \{\boldsymbol{\lambda} \in \mathbb{P}^1(L) : \sigma^{-1}(\boldsymbol{\lambda}) = P_\sigma \boldsymbol{\lambda} \text{ for all } \sigma \text{ in } G\} = \{\boldsymbol{\Phi}\boldsymbol{\beta} : \boldsymbol{\beta} \in \mathbb{P}^1(k)\},$$

where the second equality follows from the previous lemma. Thus $\Lambda(P)$ is exactly the set of all cross ratios $[\boldsymbol{\alpha}, \boldsymbol{\alpha}', \boldsymbol{\alpha}''; \boldsymbol{\beta}]$ with $\boldsymbol{\beta}$ in $\mathbb{P}^1(k)$. We note that $\Lambda(P)$ does not depend on our choice of generator α for the extension K/k. For suppose that $K = k(\alpha_1)$ with $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1'$ and $\boldsymbol{\alpha}_1''$ the corresponding points in $\mathbb{P}^1(L)$. Then there exists a unique element $\boldsymbol{\Psi}$ in $\mathrm{PGL}(2, k)$ such that $\boldsymbol{\alpha} = \boldsymbol{\Psi}\boldsymbol{\alpha}_1$. Therefore

$$\boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{\alpha}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{\alpha}_1' = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{\alpha}_1'' = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

and the corresponding set of cross ratios is

$$\{\boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{\beta} : \boldsymbol{\beta} \in \mathbb{P}^1(k)\} = \Lambda(P).$$

Also, the only points in $\mathbb{P}^1(\bar{k}_v)$ which are fixed by each element of $\{P_\sigma : \sigma \in G\}$ are $\begin{bmatrix} \zeta_6 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} \zeta_6^5 \\ 1 \end{bmatrix}$, where ζ_6 is a primitive sixth root of 1, and these are fixed points only when G is cyclic. It follows that $\Lambda(P)$ does not intersect $\mathbb{P}^1(\mathbb{Q})$.

We now consider the problem of giving good effective lower bounds for the projective distance $\delta_v\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \boldsymbol{\lambda}\right)$ for all $\boldsymbol{\lambda}$ in $\Lambda(P)$. Toward this end it will be convenient to define the exponent of approximation $e_v(\boldsymbol{\lambda})$, for all $\boldsymbol{\lambda}$ in $\Lambda(P)$, to be

$$e_v(\boldsymbol{\lambda}) = \frac{-\log \delta_v\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \boldsymbol{\lambda}\right)}{\log\{8H(\boldsymbol{\lambda})\}}, \tag{3.3}$$

so that

$$\delta_v\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \boldsymbol{\lambda}\right) = \{8H(\boldsymbol{\lambda})\}^{-e_v(\boldsymbol{\lambda})}.$$

We also set

$$\eta_w(P) = \max\{\eta_w(P_\sigma) : \sigma \in G\}$$

at each place w of L and

$$\eta(P) = \prod_w \eta_w(P).$$

A simple calculation using (2.2) and (2.3) shows that $\eta(P) = \frac{1}{2}(3 + 5^{1/2})$. If $[L : k] = 6$ and α' and α'' occur in k_v then the Liouville bound (2.12) implies that $e_v(\boldsymbol{\lambda}) < 6$ for all $\boldsymbol{\lambda}$ in $\Lambda(P)$. In fact this estimate can be substantially improved simply by using the defining property of $\Lambda(P)$.

THEOREM 3. *For each point $\boldsymbol{\lambda}$ in $\Lambda(P)$ we have $0 \leq e_v(\boldsymbol{\lambda}) < 3$.*

PROOF. Our initial embedding of L into $\bar{k}_v$ determines a place $\tilde{w}$ of L such that $\tilde{w} \mid v$ and

$$\frac{[L:k]}{[L_{\tilde{w}}:k_v]} \log \delta_{\tilde{w}}(\mathbf{x}, \mathbf{y}) = \log \delta_v(\mathbf{x}, \mathbf{y}) .$$

If $[L:k][L_{\tilde{w}}:k_v]^{-1} = 3$ the result follows as in our derivation of the Liouville bound (2.12). Only the special case $[L:k][L_{\tilde{w}}:k_v]^{-1} = 6$ requires a separate argument. The group G acts on places w of L such that $w \mid v$, where v is a fixed place of k. If σ is in G then σw is the place of L determined by

$$|x|_{\sigma w} = |\sigma^{-1}(x)|_w \text{ for all } x \text{ in } L .$$

With respect to the projective metrics δ_w and points $\boldsymbol{\lambda}$ in $\Lambda(P)$ this action results in the identity

$$\delta_{\sigma w}\left(\begin{bmatrix}0\\1\end{bmatrix}, \boldsymbol{\lambda}\right) = \delta_w\left(\begin{bmatrix}0\\1\end{bmatrix}, \sigma^{-1}(\boldsymbol{\lambda})\right) = \delta_w\left(\begin{bmatrix}0\\1\end{bmatrix}, P_\sigma \boldsymbol{\lambda}\right)$$

for all σ in G. We have

$$\begin{aligned}
&\log \delta_v\left(\begin{bmatrix}0\\1\end{bmatrix}, \boldsymbol{\lambda}\right)\\
&\quad = \sum_{w \mid v} \log \delta_{\tilde{w}}\left(\begin{bmatrix}0\\1\end{bmatrix}, \boldsymbol{\lambda}\right)\\
&\quad \geq \sum_{w \mid v} \left(\min\left\{\log \delta_{\sigma w}\left(\begin{bmatrix}0\\1\end{bmatrix}, \boldsymbol{\lambda}\right) : \sigma \in G\right\}\right)\\
&\quad = \sum_{w \mid v} \left(\min\left\{\log \delta_w\left(\begin{bmatrix}0\\1\end{bmatrix}, P_\sigma\boldsymbol{\lambda}\right) : \sigma \in G\right\}\right)\\
&\quad = \sum_{w \mid v} \left(\min\left\{\log \delta_w\left(P_\sigma P_\sigma^{-1}\begin{bmatrix}0\\1\end{bmatrix}, P_\sigma\boldsymbol{\lambda}\right) : \sigma \in G\right\}\right)\\
&\quad \geq \sum_{w \mid v} \left(\min\left\{-\log \eta_w(P_\sigma) + \log \delta_w\left(P_\sigma^{-1}\begin{bmatrix}0\\1\end{bmatrix}, \boldsymbol{\lambda}\right) : \sigma \in G\right\}\right)\\
&\quad \geq \sum_{w \mid v} \left(-\log \eta_w(P) + \min\left\{\log \delta_w\left(P_\sigma^{-1}\begin{bmatrix}0\\1\end{bmatrix}, \boldsymbol{\lambda}\right) : \sigma \in G\right\}\right)\\
&\quad \geq -\log \eta(P)\\
&\qquad + \sum_{w \mid v} \min\left\{\log \delta_w\left(\begin{bmatrix}0\\1\end{bmatrix}, \boldsymbol{\lambda}\right), \log \delta_w\left(\begin{bmatrix}1\\1\end{bmatrix}, \boldsymbol{\lambda}\right), \log \delta_w\left(\begin{bmatrix}1\\0\end{bmatrix}, \boldsymbol{\lambda}\right)\right\}\\
&\quad \geq -\log \eta(P) + \sum_{w \mid v} \log \delta_w\left(\begin{bmatrix}0\\1\end{bmatrix}, \boldsymbol{\lambda}\right)\\
&\qquad + \sum_{w \mid v} \log \delta_w\left(\begin{bmatrix}1\\1\end{bmatrix}, \boldsymbol{\lambda}\right) + \sum_{w \mid v} \log \delta_w\left(\begin{bmatrix}1\\0\end{bmatrix}, \boldsymbol{\lambda}\right)\\
&\quad \geq -\log \eta(P) - \log H\left(\begin{bmatrix}0\\1\end{bmatrix}\right) - \log H\left(\begin{bmatrix}1\\1\end{bmatrix}\right) - \log H\left(\begin{bmatrix}1\\0\end{bmatrix}\right) - 3\log H(\boldsymbol{\lambda})\\
&\quad = -3\log\{(1.5470\ldots)H(\boldsymbol{\lambda})\} .
\end{aligned}$$

The desired inequality plainly follows from this.

For each pair of points λ_1 and λ_2 in $\Lambda(P)$ we define

$$r(\lambda_1, \lambda_2) = \min\left\{\frac{\log\{8H(\lambda_1)\}}{\log\{8H(\lambda_2)\}}, \frac{\log\{8H(\lambda_2)\}}{\log\{8H(\lambda_1)\}}\right\} \tag{3.4}$$

and

$$s(\lambda_1, \lambda_2) = \left(\log\{8H(\lambda_1)\}\right)^{-1} + \left(\log\{8H(\lambda_2)\}\right)^{-1}. \tag{3.5}$$

The main technical result of this paper is a formulation of the Thue-Siegel principle as an inequality which is explicit in all constants. The inequality bounds the product $e_v(\lambda_1)e_v(\lambda_2)$ whenever λ_1 and λ_2 occur in $\Lambda(P)$. By Theorem 3 we have $e_v(\lambda_1)e_v(\lambda_2) < 9$, but if $r(\lambda_1, \lambda_2)$ and $s(\lambda_1, \lambda_2)$ are small then our inequality shows that $e_v(\lambda_1)e_v(\lambda_2)$ is not much larger than 6. The precise result is as follows.

THEOREM 4. *If λ_1 and λ_2 are points in $\Lambda(P)$ then*

$$e_v(\lambda_1)e_v(\lambda_2) \leq 6 + 19s(\lambda_1, \lambda_2)^{1/3} + 24r(\lambda_1, \lambda_2)^{1/2}. \tag{3.6}$$

After several preliminary inequalities, our proof of Theorem 4 will be given in Section 8.

If $\varepsilon > 0$ then by a familiar argument (3.6) implies that

$$\{\lambda \in \Lambda(P) : 6^{1/2} + \varepsilon \leq e_v(\lambda)\} \tag{3.7}$$

is a finite set. Of course the argument which leads to this conclusion does not provide an upper bound for $H(\lambda)$ when λ belongs to the set (3.7). In order to obtain an effective result we must be able to determine a point λ_1 in $\Lambda(P)$ for which

$$\mu_v^*(\lambda_1) = e_v(\lambda_1)^{-1}\{6 + 19(\log\{8H(\lambda_1)\})^{1/3}\}. \tag{3.8}$$

is less than 3. Then $\mu_v^*(\lambda_1)$ will be an effective measure of irrationality for each generator α of the cubic extension $K = k(\alpha) \subseteq k_v$.

THEOREM 5. *Assume that λ_1 in $\Lambda(P)$ satisfies $\mu_v^*(\lambda_1) < 3$. Then we have*

$$-\log \delta_v(\alpha, \beta) \leq \mu_v^*(\lambda_1) \log\{8H(\beta)\} + O\left\{\left(\log\{8H(\beta)\}\right)^{1/2}\right\} \tag{3.9}$$

for all β in $\mathbb{P}^1(k)$. The implied constant in (3.9) *depends in λ_1 and α.*

PROOF. From the mean value theorem we have the elementary inequality

$$(x^{-1}+y^{-1})^{1/3}-x^{-1/3}\leq \frac{1}{3}x^{2/3}y^{-1}$$

whenever x and y are positive real numbers. If λ_2 belongs to $\Lambda(P)$ then

$$e_v(\lambda_1)^{-1}\{6+19s(\lambda_1,\lambda_2)^{1/3}\}-\mu_v^*(\lambda_1)\leq 4\big(\log\{8H(\lambda_1)\}\big)^{2/3}\big(\log\{8H(\lambda_2)\}\big)^{-1}$$

and, using (3.6), we find that

$$\begin{aligned}e_v(\lambda_2)\leq \mu_v^*(\lambda_1)&+4\big(\log\{8H(\lambda_1)\}\big)^{2/3}\big(\log\{8H(\lambda_2)\}\big)^{-1}\\&+12\big(\log\{8H(\lambda_1)\}\big)^{1/2}\big(\log\{8H(\lambda_2)\}\big)^{-1/2}.\end{aligned}\tag{3.10}$$

Let Φ be the unique element of PGL(2, L) which satisfies (3.1) and write

$$\eta(\Phi)=\prod_w \eta_w(\Phi)=\prod_w |\Phi|_w^2,$$

where the equality on the right follows from (2.4). If β occurs in $\mathbb{P}^1(k)$ then

$$H(\Phi\beta)\leq\prod_w |\Phi|_w|\beta|_w=\eta(\Phi)^{1/2}H(\beta).\tag{3.11}$$

And from (2.6) we conclude that

$$\begin{aligned}\delta_v(\alpha,\beta)&=\delta_v\left(\Phi^{-1}\begin{bmatrix}0\\1\end{bmatrix},\Phi^{-1}\Phi\beta\right)\\&\geq \eta_v(\Phi)^{-1}\delta_v\left(\begin{bmatrix}0\\1\end{bmatrix},\Phi\beta\right)\\&\geq \eta(\Phi)^{-1}\delta_v\left(\begin{bmatrix}0\\1\end{bmatrix},\Phi\beta\right).\end{aligned}\tag{3.12}$$

As $\Phi\beta$ occurs in $\Lambda(P)$ the inequalities (3.10) and (3.12) can be combined. In this way we obtain the bound

$$\begin{aligned}-\log\delta_v(\alpha,\beta)\leq \mu_v^*(\lambda_1)\log\{8H(\Phi\beta)\}&+12\big(\log\{8H(\lambda_1)\}\big)^{1/2}\big(\log\{8H(\Phi\beta)\}\big)^{1/2}\\&+4\big(\log\{8H(\lambda_1)\}\big)^{2/3}+\log\eta(\Phi).\end{aligned}$$

The result plainly follows by appealing to (3.11).

We now turn our attention to the proof of Theorem 4.

4. – Local estimates

Throughout this section M and N are nonnegative integers, w is a place of the number field L and Ω_w is the completion of an algebraic closure $\bar{L}_w$. We identify the vector space $E_w = \Omega_w^{(M+1)(N+1)}$ with bihomogeneous polynomials

$$F(\mathbf{x}, \mathbf{y}) = \sum_{m=0}^{M} \sum_{n=0}^{N} f_{m,n} x_0^m x_1^{M-m} y_0^n y_1^{N-n} \tag{4.1}$$

in $\Omega_w[\mathbf{x}, \mathbf{y}]$ having bidegree (M, N). We define two norms on E_w. The first of these we denote by $| \ |_w$ and define by setting

$$|F|_w = \sup \left\{ |F(\mathbf{x}, \mathbf{y})|_w : \mathbf{x} \in \Omega_w^2, \mathbf{y} \in \Omega_w^2, |\mathbf{x}|_w \le 1 \text{ and } |\mathbf{y}|_w \le 1 \right\}. \tag{4.2}$$

As $(\mathbf{x}, \mathbf{y}) \to |\mathbf{x}|_w^{-M} |\mathbf{y}|_w^{-N} |F(\mathbf{x}, \mathbf{y})|_w$ is a well defined map from $\mathbb{P}^1(\Omega_w) \times \mathbb{P}^1(\Omega_w)$ into $[0, \infty)$, it follows that

$$|F|_w = \sup \left\{ |\mathbf{x}|_w^{-M} |\mathbf{y}|_w^{-N} |F(\mathbf{x}, \mathbf{y})|_w : \mathbf{x} \in \mathbb{P}^1(\Omega_w) \text{ and } \mathbf{y} \in \mathbb{P}^1(\Omega_w) \right\} \tag{4.3}$$

is an alternative definition for this norm. The group $\mathrm{GL}(2, \Omega_w) \times \mathrm{GL}(2, \Omega_w)$ acts on E_w as follows, if $(A, B) \in \mathrm{GL}(2, \Omega_w) \times \mathrm{GL}(2, \Omega_w)$ we define $\rho_{(A,B)} : E_w \to E_w$ by

$$(\rho_{(A,B)} F)(\mathbf{x}, \mathbf{y}) = F(A^{-1}\mathbf{x}, B^{-1}\mathbf{y}).$$

Then $\rho_{(A,B)} \circ \rho_{(C,D)} = \rho_{(AC,BD)}$ and therefore $(A, B) \to \rho_{(A,B)}$ is a representation of $\mathrm{GL}(2, \Omega_w) \times \mathrm{GL}(2, \Omega_w)$ in $\mathrm{GL}(E_w)$.

LEMMA 6. *Let $F \in E_w$ and $(A, B) \in \mathrm{GL}(2, \Omega_w) \times \mathrm{GL}(2, \Omega_w)$. Then we have*

$$|A|_w^{-M} |B|_w^{-N} |F|_w \le |\rho_{(A,B)} F|_w \le |A^{-1}|_w^M |B^{-1}|_w^N |F|_w . \tag{4.4}$$

PROOF. For each $\mathbf{x}$ in $\mathbb{P}^1(\Omega_w)$ the inequality

$$|A|_w^{-M} |A^{-1}\mathbf{x}|_w^{-M} \le |\mathbf{x}|_w^{-M} \le |A^{-1}|_w^M |A^{-1}\mathbf{x}|_w^{-M}$$

holds, and similarly with A, $\mathbf{x}$ and M replaced by B, $\mathbf{y}$ and N. The lemma follows easily by using (4.3).

Next we introduce a second norm on E_w. In doing so it will be convenient to set

$$\varepsilon_w = \begin{cases} 0 & \text{if } w \nmid \infty \\ [L_w : \mathbb{Q}_w][L : \mathbb{Q}]^{-1} & \text{if } w \mid \infty . \end{cases}$$

Let $F(\mathbf{x}, \mathbf{y})$ be given by (4.1). If $w \nmid \infty$ we define

$$[F]_w = \max \left\{ |f_{m,n}|_w : 0 \le m \le M \text{ and } 0 \le n \le N \right\},$$

and if $w \mid \infty$ we set

$$[F]_w = \left\{ \sum_{m=0}^{M} \sum_{n=0}^{N} \binom{M}{m}^{-1} \binom{N}{n}^{-1} \|f_{m,n}\|_w^2 \right\}^{\varepsilon_w/2}.$$

LEMMA 7. *Let $F(\mathbf{x}, \mathbf{y})$ occur in E_w. If $w \nmid \infty$ then*

$$|F|_w = [F]_w, \tag{4.5}$$

and if $w \mid \infty$ then

$$\{(M+1)(N+1)\}^{-\varepsilon_w/2}[F]_w \leq |F|_w \leq [F]_w. \tag{4.6}$$

PROOF. If $w \nmid \infty$ the identity (4.5) can be established as in [13], Lemma 2. Assume then that $w \mid \infty$. Let

$$S = \{\mathbf{x} \in \Omega_w^2 : |\mathbf{x}|_w = 1\}$$

and write σ for the unique rotational invariant measure on the Borel subsets of S such that $\sigma(S) = 1$. It follows using [14], Proposition 1.4.8 and 1.4.9, that

$$\begin{aligned}(M+1)^{-1}(N+1)^{-1} \sum_{m=0}^{M} \sum_{n=0}^{N} \binom{M}{m}^{-1} \binom{N}{n}^{-1} \|f_{m,n}\|_w^2 \\ = \int_S \int_S \|F(\mathbf{x}, \mathbf{y})\|_w^2 d\sigma(\mathbf{y}) d\sigma(\mathbf{x}).\end{aligned}$$

We conclude that

$$\begin{aligned}\{(M+1)(N+1)\}^{-\varepsilon_w/2}[F]_w &\leq \sup\{|F(\mathbf{x}, \mathbf{y})|_w : \mathbf{x} \in S, \mathbf{y} \in S\} \\ &= |F|_w.\end{aligned}$$

To verify the remaining inequality in (4.6), let $\mathbf{x} \in S$ and $\mathbf{y} \in S$. Then by Cauchy's inequality

$$\begin{aligned}\|(F(\mathbf{x}, \mathbf{y})\|_w^2 &= \left\| \sum_{m=0}^{M} \sum_{n=0}^{N} f_{m,n} x_0^m x_1^{M-m} y_0^n y_1^{N-n} \right\|_w^2 \\ &\leq \left\{ \sum_{m=0}^{M} \sum_{n=0}^{N} \binom{M}{m}^{-1} \binom{N}{n}^{-1} \|f_{m,n}\|_w^2 \right\} \\ &\quad \times \left\{ \sum_{m=0}^{M} \sum_{n=0}^{N} \binom{M}{m} \binom{N}{n} \|x_0\|_w^{2m} \|x_1\|_w^{2M-2m} \|y_0\|_w^{2n} \|y_1\|_w^{2N-2n} \right\} \\ &= \left\{ \sum_{m=0}^{M} \sum_{n=0}^{N} \binom{M}{m}^{-1} \binom{N}{n}^{-1} \|f_{m,n}\|_w^2 \right\} \\ &\quad \times (\|x_0\|_w^2 + \|x_1\|_w^2)^M (\|y_0\|_w^2 + \|y_1\|_w^2)^N.\end{aligned}$$

This shows that

$$|F(\mathbf{x}, \mathbf{y})|_w \leq [F]_w$$

whenever $\mathbf{x} \in S$ and $\mathbf{y} \in S$. The inequality of the right of (4.6) follows immediately.

Suppose now that $N = 0$ and so F in E_w is simply a homogeneous polynomial in $\Omega_w[\mathbf{x}]$ which can be written as

$$F(\mathbf{x}) = \sum_{m=0}^{M} f_m x_0^m x_1^{M-m} . \tag{4.7}$$

If M is positive and F is not identically zero then there exist nonzero vectors $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \dots, \boldsymbol{\xi}_M$ in Ω_w^2 such that

$$F(\mathbf{x}) = \prod_{m=1}^{M} (\mathbf{x} \wedge \boldsymbol{\xi}_m) . \tag{4.8}$$

LEMMA 8. *Let $F(\mathbf{x})$ in E_w be given by* (4.7) *and* (4.8). *If $w \nmid \infty$ then*

$$\log |F|_w = \log[F]_w = \sum_{m=1}^{M} \log |\boldsymbol{\xi}_m|_w , \tag{4.9}$$

and if $w \mid \infty$ then

$$\begin{aligned} \log |F|_w \leq \log[F]_w &\leq \sum_{m=1}^{M} \log |\boldsymbol{\xi}_m|_w \\ &\leq \frac{\varepsilon_w}{2}\{M - \log(M+1)\} + \log[F]_w \\ &\leq \frac{\varepsilon_w}{2} M + \log |F|_w . \end{aligned} \tag{4.10}$$

PROOF. If $w \nmid \infty$ then (4.9) is Gauss' lemma for homogeneous polynomials. This can also be established as in [13], Lemma 2. We assume then that $w \mid \infty$. Let J denote a subset of $\{1, 2, \dots, M\}$ having cardinality $|J|$ and write (4.8) as

$$F(\mathbf{x}) = \prod_{m=1}^{M} (\xi_{1m} x_0 - \xi_{0m} x_1) .$$

Then we have

$$f_m = \sum_{\substack{J \\ |J|=m}} \left\{ \prod_{j \in J} \xi_{1j} \right\} \left\{ \prod_{j' \in \tilde{J}} (-\xi_{0j'}) \right\} ,$$

where $\tilde{J} = \{1, 2, \ldots, M\} \setminus J$. It follows that

$$\|f_m\|_w^2 \leq \binom{M}{m} \sum_{\substack{J \\ |J|=m}} \left\{\prod_{j\in J} \|\xi_{1j}\|_w^2\right\}\left\{\prod_{j'\in\tilde{J}} \|\xi_{0j'}\|_w^2\right\}$$

and

$$\begin{aligned}\sum_{m=0}^{M} \binom{M}{m}^{-1} \|f_m\|_w^2 &\leq \sum_{m=0}^{M} \sum_{\substack{J \\ |J|=m}} \left\{\prod_{j\in J} \|\xi_{1j}\|_w^2\right\}\left\{\prod_{j'\in\tilde{J}} \|\xi_{0j'}\|_w^2\right\} \\ &= \prod_{m=1}^{M} \{\|\xi_{0m}\|_w^2 + \|\xi_{1m}\|_w^2\} = \prod_{m=1}^{M} \|\boldsymbol{\xi}_m\|_w^2 .\end{aligned}$$

This shows that

$$\log[F]_w \leq \sum_{m=1}^{M} \log|\boldsymbol{\xi}_m|_w .$$

Next we define $I : \Omega_w^2 \setminus \{\mathbf{0}\} \to \mathbb{R}$ by

$$I(\boldsymbol{\xi}) = \int_S \log\|\mathbf{x} \wedge \boldsymbol{\xi}\|_w \, d\sigma(\mathbf{x}) , \tag{4.11}$$

with S and σ as in our proof of Lemma 7. As σ is rotationally invariant and the unitary group $U(2, \Omega_w)$ acts transitively on S, we find that I restricted to S is constant. The value of this constant is

$$\begin{aligned} I\left\{\begin{pmatrix}0\\1\end{pmatrix}\right\} &= \int_S \log\|x_0\|_w d\sigma(\mathbf{x}) \\ &= \frac{1}{\pi}\int_0^1\int_0^{2\pi} r\log\|re^{i\theta}\|_w \, d\theta\, dr = -\frac{1}{2} .\end{aligned} \tag{4.12}$$

For $\alpha \neq 0$ in Ω_w and $\boldsymbol{\xi} \neq \mathbf{0}$ in Ω_w^2 we have

$$I(\alpha\boldsymbol{\xi}) = \log\|\alpha\|_w + I(\boldsymbol{\xi}) . \tag{4.13}$$

Combining (4.12) and (4.13) leads to the identity

$$\int_S \log\|\mathbf{x} \wedge \boldsymbol{\xi}\|_w d\sigma(\mathbf{x}) = -\frac{1}{2} + \log\|\boldsymbol{\xi}\|_w . \tag{4.14}$$

Finally, we use (4.8), (4.14), and Jensen's inequality to conclude that

$$\begin{aligned} -\frac{1}{2}M + \sum_{m=1}^{M} \log\|\boldsymbol{\xi}_m\|_w &= \int_S \log\|F(\mathbf{x})\|_w d\sigma(\mathbf{x}) \\ &\leq \frac{1}{2}\log\left\{\int_S \|F(\mathbf{x})\|_w^2 d\sigma(\mathbf{x})\right\} \\ &= \frac{1}{2}\log\left\{(M+1)^{-1}\sum_{m=0}^{M}\binom{M}{m}^{-1}\|f_m\|_w^2\right\} .\end{aligned}$$

This verifies the bound

$$\sum_{m=1}^{M} \log |\xi_m|_w \leq \frac{\varepsilon_w}{2}\{M - \log(M+1)\} + \log[F]_w .$$

The remaining inequalities in (4.10) follow from (4.6).

We return to consideration of bihomogeneous polynomials and assume now that M and N are positive integers. We require two positive real parameters θ and τ. Then we define

$$\begin{aligned} \Gamma &= \Gamma(M, N; \theta, \tau) \\ &= \left\{(m, n) \in \mathbb{Z}^2 : 0 \leq m \leq M, 0 \leq n \leq N, \text{ and } 1 \leq \frac{m}{\theta M} + \frac{n}{\tau N}\right\}. \end{aligned} \tag{4.15}$$

We also define

$$\mathcal{Y}_w = \mathcal{Y}_w(\theta, \tau) \subseteq E_w$$

to be the subspace of all polynomials $F(\mathbf{x}, \mathbf{y})$ in E_w such that

$$f_{m,n} = 0 \text{ whenever } \frac{m}{\theta M} + \frac{n}{\tau N} < 1 .$$

Thus $F(\mathbf{x}, \mathbf{y})$ belongs to $\mathcal{Y}_w$ precisely when its vector of coefficients $(m, n) \to f_{m,n}$ is supported on Γ. In our next lemma we make use of the following elementary inequality: if $0 \leq s \leq 1$ and $0 \leq t \leq 1$ then

$$\max\left\{s^m t^n : (m, n) \in \Gamma\right\} \leq \max\left\{s^{\theta M}, t^{\tau N}\right\}. \tag{4.16}$$

LEMMA 9. *Let F be a polynomial in $\mathcal{Y}_w$, $\mathbf{x} \in \mathbb{P}^1(\Omega_w)$ and $\mathbf{y} \in \mathbb{P}^1(\Omega_w)$. Then we have*

$$\begin{aligned} &|\mathbf{x}|_w^{-M} |\mathbf{y}|_w^{-N} |F(\mathbf{x}, \mathbf{y})|_w \\ &\leq \left(2^{(M+N)/2}\right)^{\varepsilon_w} [F]_w \max\left\{\delta_w\left(\begin{bmatrix}0\\1\end{bmatrix}, \mathbf{x}\right)^{\theta M}, \delta_w\left(\begin{bmatrix}0\\1\end{bmatrix}, \mathbf{y}\right)^{\tau N}\right\}. \end{aligned} \tag{4.17}$$

PROOF. Suppose that $w \nmid \infty$. Then

$$\begin{aligned} &|\mathbf{x}|_w^{-M} |\mathbf{y}|_w^{-N} |F(\mathbf{x}, \mathbf{y})|_w \\ &\leq \max_{(m,n)\in\Gamma} \left\{ |f_{m,n}|_w \left(\frac{|x_0|_w}{|\mathbf{x}|_w}\right)^m \left(\frac{|x_1|_w}{|\mathbf{x}|_w}\right)^{M-m} \left(\frac{|y_0|_w}{|\mathbf{y}|_w}\right)^n \left(\frac{|y_1|_w}{|\mathbf{y}|_w}\right)^{N-n} \right\} \\ &\leq [F]_w \max_{(m,n)\in\Gamma} \left\{ \left(\frac{\left|\begin{bmatrix}0\\1\end{bmatrix} \wedge \mathbf{x}\right|_w}{|\mathbf{x}|_w}\right)^m \left(\frac{\left|\begin{bmatrix}0\\1\end{bmatrix} \wedge \mathbf{y}\right|_w}{|\mathbf{y}|_w}\right)^n \right\} \\ &\leq [F]_w \max\left\{\delta_w\left(\begin{bmatrix}0\\1\end{bmatrix}, \mathbf{x}\right)^{\theta M}, \delta_w\left(\begin{bmatrix}0\\1\end{bmatrix}, \mathbf{y}\right)^{\tau N}\right\} \end{aligned}$$

follows using (4.16).

If $w \mid \infty$ we apply Cauchy's inequality to deduce that

$$\begin{aligned}(4.18)\qquad & \|\mathbf{x}\|_w^{-2M}\|\mathbf{y}\|_w^{-2N}\|F(\mathbf{x},\mathbf{y})\|_w^2 \\ &\le \left\{\sum_{m=0}^{M}\sum_{n=0}^{N}\binom{M}{m}^{-1}\binom{N}{n}^{-1}\|f_{m,n}\|_w^2\right\} \\ &\times\left\{\sum_{\substack{m=0\\(m,n)\in\Gamma}}^{M}\sum_{n=0}^{N}\binom{M}{m}\binom{N}{n}\left(\frac{\|x_0\|_w^2}{\|\mathbf{x}\|_w^2}\right)^m\left(\frac{\|x_1\|_w^2}{\|\mathbf{x}\|_w^2}\right)^{M-m}\left(\frac{\|y_0\|_w^2}{\|\mathbf{y}\|_w^2}\right)^n\left(\frac{\|y_1\|_w^2}{\|\mathbf{y}\|_w^2}\right)^{N-n}\right\} \\ &\le\left\{2^{M+N}\sum_{m=0}^{M}\sum_{n=0}^{N}\binom{M}{m}^{-1}\binom{N}{n}^{-1}\|f_{m,n}\|_w^2\right\} \\ &\times\max\left\{\left(\frac{\left\|\begin{bmatrix}0\\1\end{bmatrix}\wedge\mathbf{x}\right\|_w^2}{\|\mathbf{x}\|_w^2}\right)^m\left(\frac{\left\|\begin{bmatrix}0\\1\end{bmatrix}\wedge\mathbf{y}\right\|_w^2}{\|\mathbf{y}\|_w^2}\right)^n : (m,n)\in\Gamma\right\}.\end{aligned}$$

We raise both sides of (4.18) to the power $\varepsilon_w/2$ and apply (4.16). The desired inequality (4.17) follows immediately.

Let F in E_w be given by (4.1) and let (A, B) occur in $\mathrm{GL}(2,\Omega_w)\times\mathrm{GL}(2,\Omega_w)$. Then $\rho_{(A,B)}$ acts on F and so also on the vector of coefficients $(m, n)\to f_{m,n}$. We write

$$\begin{aligned}(\rho_{(A,B)}F)(\mathbf{x},\mathbf{y}) &= F(A^{-1}\mathbf{x}, B^{-1}\mathbf{y}) \\ &= \sum_{m=0}^{M}\sum_{n=0}^{N}(\rho_{(A,B)}f)_{m,n}x_0^m x_1^{M-m}y_0^n y_1^{N-n},\end{aligned}$$

so that $(m, n)\to(\rho_{(A,B)}f)_{m,n}$ is the vector of coefficients of $\rho_{(A,B)}F$. Then we define

$$\mathrm{Index}(F; A, B, \theta M, \tau N) = \min\left\{\frac{m}{\theta M}+\frac{n}{\tau N} : (\rho_{(A,B)}f)_{m,n}\ne 0\right\}.$$

(If F is identically zero we set its Index equal to ∞.) For $\kappa > 0$ we have

$$(4.19)\qquad \kappa^{-1}\mathrm{Index}(F; A, B, \theta M, \tau N) = \mathrm{Index}(F; A, B, \kappa\theta M, \kappa\tau N)$$

and so it suffices to consider the Index only when $\theta M\tau N = 1$. It will be notationally convenient, however, to use all four parameters in later applications. As ρ is a representation we have the simple identity

$$(4.20)\qquad \mathrm{Index}(\rho_{(C,D)}F; A, B, \theta M, \tau N) = \mathrm{Index}(F; AC, BD, \theta M, \tau N).$$

Also, if $a \neq 0$ and $b \neq 0$ belong to Ω_w then

$$\mathrm{Index}(F; aA, bB, \theta M, \tau N) = \mathrm{Index}(F; A, B, \theta M, \tau N).$$

The map

$$(A, B) \to \mathrm{Index}(F; A, B, \theta M, \tau N)$$

is therefore well defined for (A, B) in $\mathrm{PGL}(2, \Omega_w) \times \mathrm{PGL}(2, \Omega_w)$. Clearly the polynomial F occurs in $\mathcal{Y}_w$ if and only if

$$1 \leq \mathrm{Index}(F; 1_2, 1_2, \theta M, \tau N),$$

where $1_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

THEOREM 10. *Let F be a polynomial in E_w, $\mathbf{x} \in \mathbb{P}^1(\Omega_w)$ and $\mathbf{y} \in \mathbb{P}^1(\Omega_w)$. Assume that*

$$1 \leq \mathrm{Index}(F; A, B, \theta M, \tau N)$$

for some pair (A, B) in $\mathrm{PGL}(2, \Omega_w) \times \mathrm{PGL}(2, \Omega_w)$. Then we have

$$\begin{aligned}|\mathbf{x}|_w^{-M}|\mathbf{y}|_w^{-N}|F(\mathbf{x}, \mathbf{y})|_w &\leq \left\{(M+1)(N+1)2^{M+N}\right\}^{\varepsilon w/2}\eta_w(A)^M\eta_w(B)^N \\ &\quad |F|_w \max\left\{\delta_w\left(\begin{bmatrix}0\\1\end{bmatrix}, A\mathbf{x}\right)^{\theta M}, \delta_w\left(\begin{bmatrix}0\\1\end{bmatrix}, B\mathbf{y}\right)^{\tau N}\right\}.\end{aligned} \tag{4.21}$$

PROOF. Using (4.20) we find that

$$1 \leq \mathrm{Index}(\rho_{(A,B)}F; 1_2, 1_2, \theta M, \tau N).$$

Therefore we may apply Lemma 9 and obtain the inequality

$$\begin{aligned}&|\mathbf{x}|_w^{-M}|\mathbf{y}|_w^{-N}|F(\mathbf{x}, \mathbf{y})|_w \\ &\quad\leq |A|_w^M|B|_w^N|A\mathbf{x}|_w^{-M}|B\mathbf{y}|_w^{-N}|(\rho_{(A,B)}F)(A\mathbf{x}, B\mathbf{y})|_w \\ &\quad\leq (2^{M+N})^{\varepsilon w/2}|A|_w^M|B|_w^N[\rho_{(A,B)}F]_w \\ &\qquad \max\left\{\delta_w\left(\begin{bmatrix}0\\1\end{bmatrix}, A\mathbf{x}\right)^{\theta M}, \delta_w\left(\begin{bmatrix}0\\1\end{bmatrix}, B\mathbf{y}\right)^{\tau N}\right\}.\end{aligned} \tag{4.22}$$

Then Lemma 6 and Lemma 7 imply that

$$\begin{aligned}&|A|_w^M|B|_w^N[\rho_{(A,B)}F]_w \\ &\quad\leq \left\{(M+1)(N+1)\right\}^{\varepsilon w/2}|A|_w^M|B|_w^N|\rho_{(A,B)}F|_w \\ &\quad\leq \left\{(M+1)(N+1)\right\}^{\varepsilon w/2}|A|_w^M|A^{-1}|_w^M|B|_w^N|B^{-1}|_w^N|F|_w \\ &\quad= \left\{(M+1)(N+1)\right\}^{\varepsilon w/2}\eta_w(A)^M\eta_w(B)^N|F|_w.\end{aligned} \tag{4.23}$$

The result follows by combining (4.22) and (4.23).

5. – Global estimates

Let L/k be a Galois extension of algebraic number fields with $G = \mathrm{Gal}(L/k)$. We assume that $\sigma \to \psi_\sigma$ is a faithful, projective representation of G in $\mathrm{PGL}(2, k)$. At each place v of k we define

$$\eta_v(\psi) = \max\{\eta_v(\psi_\sigma) : \sigma \in G\}.$$

Because of the way we have normalized absolute values we have

$$\eta_v(\psi) = \prod_{w \mid v} \left(\max\{\eta_w(\psi_\sigma) : \sigma \in G\}\right),$$

where the product extends over all places w of L such that $w \mid v$. As $\eta_v(\psi) = 1$ for almost all places v of k we define

$$\eta(\psi) = \prod_v \eta_v(\psi).$$

Next we define

$$\Lambda(\psi) = \left\{\boldsymbol{\lambda} \in \mathbb{P}^1(L) : \sigma^{-1}(\boldsymbol{\lambda}) = \psi_\sigma \boldsymbol{\lambda} \text{ for all } \sigma \text{ in } G\right\}. \tag{5.1}$$

The group G acts on places w of L such that $w \mid v$ where v is a fixed place of k. If σ is in G then σw is the place of L determined by

$$|x|_{\sigma w} = |\sigma^{-1}(x)|_w \text{ for all } x \text{ in } L.$$

With respect to the projective metrics δ_w and points $\boldsymbol{\lambda}$ in $\Lambda(\psi)$ this action results in the identity

$$\delta_{\sigma w}\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \boldsymbol{\lambda}\right) = \delta_w\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \psi_\sigma \boldsymbol{\lambda}\right) \tag{5.2}$$

for all σ in G.

Let $E = L^{(M+1)(N+1)}$ denote the L-vector space of bihomogeneous polynomials

$$F(\mathbf{x}, \mathbf{y}) = \sum_{m=0}^{M} \sum_{n=0}^{N} f_{m,n} x_0^m x_1^{M-m} y_0^n y_1^{N-n}$$

in $L[\mathbf{x}, \mathbf{y}]$ having bidegree (M, N).

LEMMA 11. *Let $F(\mathbf{x}, \mathbf{y})$ be a polynomial in E and let $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ be points in $\Lambda(\psi)$. Assume that $F(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) \neq 0$ and that*

$$1 \leq \mathrm{Index}(F; \psi_\sigma, \psi_\sigma, \theta M, \tau N) \tag{5.3}$$

for all σ in G. Then for any place v of k and any embedding of L into $\bar{k}_v$ we have

$$\begin{aligned} 0 \leq\; & M \log\left\{2^{1/2}\eta(\psi)H(\boldsymbol{\lambda}_1)\right\} + N \log\left\{2^{1/2}\eta(\psi)H(\lambda_2)\right\} \\ & + \frac{1}{2}\log\left\{(M+1)(N+1)\right\} + \sum_w \log|F|_w \\ & + \max\left\{\theta M \log \delta_v\left(\begin{bmatrix}0\\1\end{bmatrix}, \boldsymbol{\lambda}_1\right), \tau N \log \delta_v\left(\begin{bmatrix}0\\1\end{bmatrix}, \boldsymbol{\lambda}_2\right)\right\}. \end{aligned} \tag{5.4}$$

PROOF. We apply (4.21) at each place w of L. In this way we obtain the estimate

$$\begin{aligned} & |\boldsymbol{\lambda}_1|_w^{-M} |\boldsymbol{\lambda}_2|_w^{-N} |F(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)|_w \\ & \leq \left\{(M+1)(N+1)2^{M+N}\right\}^{\varepsilon_w/2} \eta_w(\psi)^M \eta_w(\psi)^N \\ & \quad |F|_w \left(\min\left\{\max\left\{\delta_w\left(\begin{bmatrix}0\\1\end{bmatrix}, \psi_\sigma\boldsymbol{\lambda}_1\right)^{\theta M}, \delta_w\left(\begin{bmatrix}0\\1\end{bmatrix}, \psi_\sigma\boldsymbol{\lambda}_2\right)^{\tau N}\right\} : \sigma \in G\right\}\right). \end{aligned}$$

Taking the product over all places w leads to the bound

$$\begin{aligned} 0 \leq\; & M \log\left\{2^{1/2}\eta(\psi)H(\boldsymbol{\lambda}_1)\right\} + N \log\left\{2^{1/2}\eta(\psi)H(\lambda_2)\right\} \\ & + \frac{1}{2}\log\left\{(M+1)(N+1)\right\} + \sum_w \log|F|_w \\ & + \sum_{w\,|\,v}\left(\min\left\{\max\left\{\theta M \log \delta_w\left(\begin{bmatrix}0\\1\end{bmatrix}, \psi_\sigma\boldsymbol{\lambda}_1\right), \tau N \log \delta_w\left(\begin{bmatrix}0\\1\end{bmatrix}, \psi_\sigma\boldsymbol{\lambda}_2\right)\right\} : \sigma \in G\right\}\right). \end{aligned}$$

An embedding of L into $\bar{k}_v$ determines a place $\tilde{w}$ of L such that $\tilde{w}\,|\,v$ and

$$\frac{[L:k]}{[L_{\tilde{w}} : k_v]} \log \delta_{\tilde{w}}(\mathbf{x}, \mathbf{y}) = \log \delta_v(\mathbf{x}, \mathbf{y})$$

at all points $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{P}^1(\bar{k}_v)$. As L/k is a Galois extension, $w \to [L_w : k_v]$ is constant on places w with $w\,|\,v$ and therefore

$$\sum_{w\,|\,v} 1 = \frac{[L:k]}{[L_{\tilde{w}} : k_v]}. \tag{5.5}$$

Finally, we use (5.5) and the fact that G acts transitively to conclude that

$$\begin{aligned} & \sum_{w\,|\,v}\left(\min\left\{\max\left\{\theta M \log \delta_w\left(\begin{bmatrix}0\\1\end{bmatrix}, \psi_\sigma\boldsymbol{\lambda}_1\right), \tau N \log \delta_w\left(\begin{bmatrix}0\\1\end{bmatrix}, \psi_\sigma\boldsymbol{\lambda}_2\right)\right\} : \sigma \in G\right\}\right) \\ & \quad \leq \sum_{w\,|\,v}\left\{\max\left\{\theta M \log \delta_{\tilde{w}}\left(\begin{bmatrix}0\\1\end{bmatrix}, \boldsymbol{\lambda}_1\right), \tau N \log \delta_{\tilde{w}}\left(\begin{bmatrix}0\\1\end{bmatrix}, \boldsymbol{\lambda}_2\right)\right\}\right\}) \\ & \quad = \max\left\{\theta M \log \delta_v\left(\begin{bmatrix}0\\1\end{bmatrix}, \boldsymbol{\lambda}_1\right), \tau N \log \delta_v\left(\begin{bmatrix}0\\1\end{bmatrix}, \boldsymbol{\lambda}_2\right)\right\}. \end{aligned}$$

This proves the lemma.

6. – Dyson's lemma

In this section we assume that Ω is an algebraically closed field of characteristic zero and $F(\mathbf{x}, \mathbf{y})$ is a bihomogeneous polynomial of bidegree (M, N) having coefficients in Ω. Let

$$\operatorname{stab}\left\{\begin{bmatrix}0\\1\end{bmatrix}\right\} = \left\{U \in \mathrm{PGL}(2,\Omega) : U\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix}\right\}$$
$$= \left\{\begin{bmatrix}u_{11} & u_{12}\\ u_{21} & u_{22}\end{bmatrix} \in \mathrm{PGL}(2,\Omega) : u_{12} = 0\right\}$$

denote the stabilizer of $\begin{bmatrix}0\\1\end{bmatrix}$. Then

$$\mathrm{Index}(F; A, B, \theta M, \tau N) = \mathrm{Index}(F; UA, VB, \theta M, \tau N)$$

whenever U and V belong to $\operatorname{stab}\left\{\begin{bmatrix}0\\1\end{bmatrix}\right\}$. That is, the map

$$(A, B) \to \mathrm{Index}(F; A, B, \theta M, \tau N)$$

is constant on all right cosets of $\operatorname{stab}\left\{\begin{bmatrix}0\\1\end{bmatrix}\right\} \times \operatorname{stab}\left\{\begin{bmatrix}0\\1\end{bmatrix}\right\}$ in $\mathrm{PGL}(2,\Omega) \times$ $\times\mathrm{PGL}(2,\Omega)$. Now let (α, β) belong to $\mathbb{P}^1(\Omega) \times \mathbb{P}^1(\Omega)$ and select (A, B) so that $A^{-1}\begin{bmatrix}0\\1\end{bmatrix} = \alpha$ and $B^{-1}\begin{bmatrix}0\\1\end{bmatrix} = \beta$. We set

$$\mathrm{index}(F; \alpha, \beta, \theta M, \tau N) = \mathrm{Index}(F; A, B, \theta M, \tau N)\,.$$

By our previous remarks index is well defined. We note that the analogue of (4.12) is the identity

$$\mathrm{index}(F; \alpha, \beta, \theta M, \tau N) = \mathrm{index}(\rho_{(C,D)}F; C\alpha, D\beta, \theta M, \tau N)\,. \tag{6.1}$$

We now describe the projective form of Dyson's lemma. We require the following objects:

$$\alpha_1, \alpha_2, \dots, \alpha_J \text{ are distinct points in } \mathbb{P}^1(\Omega)\,,$$
$$\beta_1, \beta_2, \dots, \beta_J \text{ are distinct points in } \mathbb{P}^1(\Omega)\,,$$
$$0 < \theta_j < 1, 0 < \tau_j < 1\,, \text{ for } j = 1, 2, \dots, J.$$

DYSON'S LEMMA. *Let F be a bihomogeneous polynomial of bidegree (M, N) in $\Omega[\mathbf{x}, \mathbf{y}]$ which is not identically zero. If*

$$1 \le \mathrm{index}(F; \alpha_j, \beta_j, \theta_j M, \tau_j N)$$

for each $j = 1, 2, \dots, J$ then

$$\frac{1}{2}\sum_{j=1}^{J} \theta_j \tau_j \le 1 + \left(\frac{J-2}{2}\right) \min\left\{\frac{M}{N}, \frac{N}{M}\right\}.$$

This can be obtained from the affine formulation given in [5] with minor modifications. It can also be obtained, again with minor modifications, from the more general result of Vojta [19]. For our purposes it will be useful to have a second version which specifically bounds the index.

Corollary 12. *Let F be a bihomogeneous polynomial of bidegree (M, N) in $\Omega[\mathbf{x}, \mathbf{y}]$ which is not identically zero. Then we have*

$$\begin{aligned}\frac{1}{2}\sum_{j=1}^{J}\min\left\{\frac{\tau_j}{\theta_j}, \frac{\theta_j}{\tau_j}, \theta_j\tau_j\big(\text{index}(F; \boldsymbol{\alpha}_j, \boldsymbol{\beta}_j, \theta_j M, \tau_j N)\big)^2\right\}\\ \leq 1 + \left(\frac{J-2}{2}\right)\min\left\{\frac{M}{N}, \frac{N}{M}\right\}.\end{aligned} \tag{6.2}$$

Proof. Select κ_j, $j = 1, 2, \ldots, J$, so that

$$0 < \kappa_j < \min\left\{\theta_j^{-1}, \tau_j^{-1}, \text{index}(F; \boldsymbol{\alpha}_j, \boldsymbol{\beta}_j, \theta_j M, \tau_j N)\right\}.$$

It follows that $0 < \kappa_j\theta_j < 1$, $0 < \kappa_j\tau_j < 1$, and

$$1 \leq \kappa_j^{-1}\text{index}(F; \boldsymbol{\alpha}_j, \boldsymbol{\beta}_j, \theta_j M, \tau_j N) = \text{index}(F; \boldsymbol{\alpha}_j, \boldsymbol{\beta}_j, \kappa_j\theta_j M, \kappa_j\tau_j N).$$

By Dyson's lemma we conclude that

$$\frac{1}{2}\sum_{j=1}^{J}\theta_j\tau_j\kappa_j^2 \leq 1 + \left(\frac{J-2}{2}\right)\min\left\{\frac{M}{N}, \frac{N}{M}\right\}. \tag{6.3}$$

The corollary follows by taking the supremum on the left hand side of (6.3) over all values of the parameters κ_j, $j = 1, 2, \ldots, J$.

In applications of Dyson's lemma it is necessary to estimate how the index changes when a differential operator is applied to F. Let

$$T(\mathbf{x}, \mathbf{y}) = \sum_{r=1}^{R}\sum_{s=1}^{S} t_{r,s}x_0^r x_1^{R-r} x_1^{R-r} y_o^s y_1^{S-s}$$

be a second bihomogeneous polynomial of bidegree (R, S) in $\Omega[\mathbf{x}, \mathbf{y}]$ and let

$$T(\mathcal{D}) = \sum_{r=1}^{R}\sum_{s=1}^{S} t_{r,s}\frac{1}{r!}\left(\frac{\partial}{\partial x_0}\right)^r \frac{1}{(R-r)!}\left(\frac{\partial}{\partial x_1}\right)^{R-r}\frac{1}{s!}\left(\frac{\partial}{\partial y_0}\right)^s \frac{1}{(S-s)!}\left(\frac{\partial}{\partial y_1}\right)^{S-s}$$

be the corresponding linear partial differential operator. If $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ belongs to $\mathbb{P}^1(\Omega) \times \mathbb{P}^1(\Omega)$ we find that

$$\begin{aligned}\text{index}\big(T(\mathcal{D})F; \boldsymbol{\alpha}, \boldsymbol{\beta}, \theta M, \tau N\big) &\geq \text{index}(F; \boldsymbol{\alpha}, \boldsymbol{\beta}, \theta M, \tau N)\\ &+ \text{index}(T; I_2\boldsymbol{\alpha}, I_2\boldsymbol{\beta}, \theta M, \tau N) - \left(\frac{R}{\theta M} + \frac{S}{\tau N}\right),\end{aligned} \tag{6.4}$$

where I_2 is the involution

$$I_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

In the special case $\alpha = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $\beta = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ the inequality (6.4) follows from the expansion

$$\begin{aligned}\{T(\mathcal{D})F\}(\mathbf{x}, \mathbf{y}) = \sum_{m=0}^{M-R} \sum_{n=0}^{N-S} \Bigg\{ \sum_{r=0}^{R} \sum_{s=0}^{S} f_{m+r,n+s} t_{r,s} \binom{m+r}{r} \\ \times \binom{M-m-r}{R-r} \binom{n+s}{s} \binom{N-n-s}{S-s} \Bigg\} x_0^m x_1^{M-R-m} y_0^n y_1^{N-S-n} .\end{aligned} \tag{6.5}$$

The general case can then be established by selecting A and B so that $A^{-1}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \alpha$, $B^{-1}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \beta$ and using the identity

$$\rho_{(A,B)}\{(\rho_{(A',B')}T)(\mathcal{D})F\} = T(\mathcal{D})(\rho_{(A,B)}F) , \tag{6.6}$$

where A' and B' are the transpose of A and B respectively. (Note also that $A'I_2A = I_2$ and $B'I_2B = I_2$ in PGL(2, Ω).)

Again let A and B satisfy $A^{-1}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \alpha$ and $B^{-1}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \beta$. We say that the pair of nonnegative integers (R, S) is *critical* with respect to the index at (α, β) if

$$\begin{aligned}\text{index}(F; \alpha, \beta, \theta M, \tau N) &= \min\left\{ \frac{m}{\theta M} + \frac{n}{\tau N} : (\rho_{(A,B)} f)_{m,n} \neq 0 \right\} \\ &= \frac{R}{\theta M} + \frac{S}{\tau N}\end{aligned}$$

and

$$(\rho_{(A,B)} f)_{R,S} \neq 0 .$$

If (R, S) is critical at (α, β) then (6.4) and (6.5) can be used to show that

$$\text{index}(T(\mathcal{D})F; \alpha, \beta, \theta M, \tau N) = 0$$

if and only if

$$\text{index}(T; I_2\alpha, I_2\beta, \theta M, \tau N) = 0 .$$

Alternatively, if (R, S) is critical at (α, β) then

$$\{T(\mathcal{D})F\}(\alpha, \beta) \neq 0$$

if and only if

$$T(I_2\alpha, I_2\beta) \neq 0 .$$

7. – The auxiliary polynomial

Here we assume that $0 < \theta < 1$, $0 < \tau < 1$ and also that $\theta\tau < 2/3$. Then we define

$$\begin{aligned}\Gamma' &= \Gamma'(M, N, \theta, \tau)\\ &= \left\{(m, n) \in \mathbb{Z}^2 : 0 \le m \le M, 0 \le n \le N, \text{ and } 1 \le \frac{m}{\theta M} + \frac{n}{\tau N} \le \frac{1}{\theta} + \frac{1}{\tau} - 1\right\}\end{aligned}$$

and

$$\begin{aligned}\Gamma'' &= \Gamma''(M, N, \theta, \tau)\\ &= \left\{(m, n) \in \mathbb{Z}^2 : 0 \le m \le M, 0 \le n \le N, \text{ and } \frac{m}{\theta M} + \frac{n}{\tau N} < 1\right\}.\end{aligned}$$

Let $0 < c_1 < c_2 < \infty$. Then

$$\lim M^{-1}N^{-1}|\Gamma'| = 1 - \theta\tau,$$
$$\lim M^{-1}N^{-1}|\Gamma''| = \frac{1}{2}\theta\tau,$$

where $M \to \infty$, $N \to \infty$, in such a way that

$$c_1 \le MN^{-1} \le c_2. \tag{7.1}$$

Therefore

$$\lim M^{-1}N^{-1}(|\Gamma'| - |\Gamma''|) = 1 - \frac{3}{2}\theta\tau \tag{7.2}$$

is positive. In particular, $|\Gamma'| - |\Gamma''| > 0$ if M and N satisfy (7.1) and are sufficiently large.

THEOREM 13. *Let M and N be positive integers such that* (7.1) *holds and* $|\Gamma'| - |\Gamma''| > 0$. *Then there exists a bihomogeneous polynomial*

$$F(\mathbf{x}, \mathbf{y}) = \sum_{m=0}^{M}\sum_{n=0}^{N} f_{m,n} x_0^m x_1^{M-m} y_0^n y_1^{N-n}$$

of bidegree (M, N) in $\mathbb{Z}[\mathbf{x}, \mathbf{y}]$ such that

(i) *F is not identically zero,*

(ii) $1 \le \operatorname{index}\left(F; \begin{bmatrix}0\\1\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}, \theta M, \tau N\right)$,

(iii) $1 \le \operatorname{index}\left(F; \begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}, \theta M, \tau N\right)$,

(iv) $1 \le \operatorname{index}\left(F; \begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}1\\0\end{bmatrix}, \theta M, \tau N\right)$,

(v) *the coefficients* $f_{m,n}$ *satisfy*

$$\log|f_{m,n}| \leq \frac{1}{4}\left(1-\frac{3}{2}\theta\tau\right)^{-1}(\theta M+\tau N)+o(M)+o(N), \tag{7.3}$$

(vi) *either*

$$F(\mathbf{x},\mathbf{y}) = F\left(\begin{pmatrix}0&1\\1&0\end{pmatrix}\mathbf{x},\begin{pmatrix}0&1\\1&0\end{pmatrix}\mathbf{y}\right)$$

or

$$F(\mathbf{x},\mathbf{y}) = -F\left(\begin{pmatrix}0&1\\1&0\end{pmatrix}\mathbf{x},\begin{pmatrix}0&1\\1&0\end{pmatrix}\mathbf{y}\right).$$

PROOF. We will construct F so that its coefficients $f_{m,n}$ are supported on Γ'. This insures that (ii) and (iv) hold. We also require that

$$\sum_{(m,n)\in\Gamma'}\binom{m}{i}\binom{n}{j}f_{m,n}=0$$

for all pairs (i,j) in Γ'' and this implies that (iii) is satisfied. Thus we need a nontrivial solution in integers to $|\Gamma''|$ homogeneous linear equations in $|\Gamma'|$ variables. By the simplest form of Siegel's lemma there exists such a solution with

$$\log|f_{m,n}| \leq \{|\Gamma'|-|\Gamma''|\}^{-1}\sum_{(i,j)\in\Gamma''}\log\left\{|\Gamma'|\binom{M}{i}\binom{N}{j}\right\}. \tag{7.4}$$

It remains then to estimate the right hand side of (7.4) as $M\to\infty$, $N\to\infty$, in such a way that (7.1) holds.

Let

$$\psi(x) = -x\log x-(1-x)\log(1-x)$$

for $0<x<1$, $\psi(0)=\psi(1)=0$, and define

$$\gamma'' = \left\{(x,y)\in\mathbb{R}^2 : 0\leq x\leq 1, 0\leq y\leq 1, \text{ and } \frac{x}{\theta}+\frac{y}{\tau}<1\right\}.$$

Then we have

$$\log\binom{M}{i}+\log\binom{N}{j} \leq M\psi\left(\frac{i}{M}\right)+N\psi\left(\frac{j}{N}\right), \tag{7.5}$$

$$\lim M^{-1}N^{-1}\sum_{(i,j)\in\Gamma''}\psi\left(\frac{i}{M}\right) = \iint_{\gamma''}\psi(x)dx\,dy, \tag{7.6}$$

and

$$\lim M^{-1}N^{-1} \sum_{(i,j)\in\Gamma''} \psi\left(\frac{j}{N}\right) = \iint_{\gamma''} \psi(y)dx\,dy\,. \tag{7.7}$$

By combining (7.5), (7.6) and (7.7) we find that

$$\begin{aligned}&\{|\Gamma'|-|\Gamma''|\}^{-1}\left(|\Gamma''|\log|\Gamma'| + \sum_{(i,j)\in\Gamma''}\left\{\log\binom{M}{i} + \log\binom{N}{j}\right\}\right)\\ &\le \left(1-\frac{3}{2}\theta\tau\right)^{-1}\left\{M\iint_{\gamma''}\psi(x)dx\,dy + N\iint_{\gamma''}\psi(y)dx\,dy\right\}+o(M)+o(N)\,.\end{aligned} \tag{7.8}$$

Next we observe that

$$\begin{aligned}\iint_{\gamma''}\psi(x)dx\,dy &= \theta\tau\int_0^1 \psi(\theta x)(1-x)dx\\ &= \theta\tau\left\{-\frac{1}{6}\theta\log\theta + \frac{11}{36}\theta - \sum_{\ell=4}^{\infty}\frac{\theta^{\ell-2}}{\ell(\ell-1)(\ell-2)(\ell-3)}\right\}\\ &\le -\frac{1}{6}\theta^2\tau\log\theta + \frac{11}{36}\theta^2\tau\,,\end{aligned}$$

and in a similar manner

$$\iint_{\gamma''}\psi(y)dx\,dy \le -\frac{1}{6}\theta\tau^2\log\tau + \frac{11}{36}\theta\tau^2\,.$$

It follows that

$$\begin{aligned}&M\iint_{\gamma''}\psi(x)dx\,dy + N\iint_{\gamma''}\psi(y)dx\,dy\\ &\le \theta M\left(-\frac{1}{6}\theta\tau\log\theta + \frac{11}{36}\theta\tau\right) + \tau N\left(-\frac{1}{6}\theta\tau\log\tau + \frac{11}{36}\theta\tau\right)\\ &\le \{\theta M + \tau N\}\max\left\{-\frac{1}{6}\theta\tau\log\theta + \frac{11}{36}\theta\tau, -\frac{1}{6}\theta\tau\log\tau + \frac{11}{36}\theta\tau\right\}\,.\end{aligned} \tag{7.9}$$

Finally, it can be shown that

$$\begin{aligned}&\max\left\{-\frac{1}{6}\theta\tau\log\theta + \frac{11}{36}\theta\tau, -\frac{1}{6}\theta\tau\log\tau + \frac{11}{36}\theta\tau\right\}\\ &\le \frac{1}{9}\log\left(\frac{3}{2}\right) + \frac{11}{54}\\ &\le \frac{1}{4}\,,\end{aligned} \tag{7.10}$$

in the region $0 < \theta < 1$, $0 < \tau < 1$, and $\theta\tau < 2/3$. The bound (v) plainly follows from (7.4), (7.8), (7.9) and (7.10).

In the polynomial F *fails* to satisfy (vi) then we may replace F with

$$F(\mathbf{x}, \mathbf{y}) + F\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \mathbf{x}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \mathbf{y}\right). \tag{7.11}$$

It is easy to verify that (7.11) continues to satisfy (i)-(v).

8. – Proof of the Thue-Siegel inequality

We now prove Theorem 4 in the slightly stronger form

$$e_v(\boldsymbol{\lambda}_1)e_v(\boldsymbol{\lambda}_2) \leq 6\{1 + C_s s(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)^{1/3} + C_r r(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)^{1/2}\}, \tag{8.1}$$

where $C_s = 2^{-3}3^{7/3} + 2^{-3}3^{4/3} = 3.152267442\ldots$ and $C_r = (3/4)7^{1/2} + 2^{-1/2}7^{1/2} = 3.855142176\ldots$. In doing so we may assume that $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ belong to $\Lambda(P)$ and satisfy

$$2 < e_v(\boldsymbol{\lambda}_1) < 3, \qquad 2 < e_v(\boldsymbol{\lambda}_2) < 3,$$

and

$$C_s s(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)^{1/3} + C_r r(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)^{1/2} < 1/2. \tag{8.2}$$

For if any of these conditions fail to hold then the basic inequality (8.1) follows from Theorem 3. Let $\mathcal{G} \subseteq \mathrm{PGL}(2, \mathbb{Q})$ denote the subgroup on the right hand side of (3.2) which is isomorphic to S_3. Then $\mathcal{G}$ acts on $\mathbb{P}^1(\bar{k}_v)$ and every orbit has six elements with exactly three exceptions. The exceptional orbits are

$$\left\{\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}, \left\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}\right\}, \text{and } \left\{\begin{bmatrix} \zeta_6 \\ 1 \end{bmatrix}, \begin{bmatrix} \zeta_6^5 \\ 1 \end{bmatrix}\right\},$$

where ζ_6 is a primitive sixth root of 1. If $\boldsymbol{\lambda}$ in $\Lambda(P)$ belongs to any of these exceptional orbits then $e_v(\boldsymbol{\lambda}) \leq 2$ follows easily. Hence we may also assume that $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ each have orbits containing six points under the action of $\mathcal{G}$.

Let $0 < \theta < 1$, $0 < \tau < 1$, with $\frac{1}{2} < \theta\tau < \frac{2}{3}$. Then let M and N be positive integers such that

$$\frac{1}{2} r(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) \leq \min\left\{\frac{M}{N}, \frac{N}{M}\right\} \leq 2r(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2).$$

We note that $r(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) \leq (50)^{-1}$ by (8.2) and therefore

$$1 + \frac{7}{2}\min\left\{\frac{M}{N}, \frac{N}{M}\right\} < \frac{9}{2}\theta\tau. \tag{8.3}$$

We assume now that M and N are so large that the hypotheses of Theorem 13 are satisfied. Then let F be a bihomogeneous polynomial in $\mathbb{Z}[\mathbf{x}, \mathbf{y}]$ having bidegree (M, N) for which the conclusions (i)-(vi) of that result hold. It is easy to verify that

$$A \to \text{index}(F; A\boldsymbol{\lambda}_1, A\boldsymbol{\lambda}_2, \theta M, \tau N)$$

is constant for A in $\mathcal{G}$. This is obvious if $G = \text{Gal}(L/k)$ is noncyclic since $\mathcal{G} = \{P_\sigma : \sigma \in G\}$ and $P_\sigma \boldsymbol{\lambda}_i = \sigma^{-1}(\boldsymbol{\lambda}_i)$ for $i = 1, 2$ in that case. If G is cyclic it follows also because

$$\mathcal{G} = \{P_\sigma : \sigma \in G\} \cup \left\{ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} P_\sigma : \sigma \in G \right\}$$

and using (6.1) and (vi) of Theorem 13 we have

$$\begin{aligned} &\text{index}(F; P_\sigma\boldsymbol{\lambda}_1, P_\sigma\boldsymbol{\lambda}_2, \theta M, \tau N) \\ &\quad = \text{index}\left(F; \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} P_\sigma\boldsymbol{\lambda}_1, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} P_\sigma\boldsymbol{\lambda}_2, \theta M, \tau N\right). \end{aligned}$$

To simplify notation we write

$$\kappa = \text{index}(F; \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \theta M, \tau N).$$

If $1 \le \kappa$ then we may apply Dyson's lemma at the nine points

$$\{(A\boldsymbol{\lambda}_1, A\boldsymbol{\lambda}_2) : A \in \mathcal{G}\} \cup \left\{ \left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right), \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right), \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) \right\}$$

in $\mathbb{P}^1(\bar{k}_v) \times \mathbb{P}^1(\bar{k}_v)$. We conclude that

$$\frac{9}{2}\theta\tau \le 1 + \frac{7}{2}\min\left\{\frac{M}{N}, \frac{N}{M}\right\},$$

which is impossible by (8.3). Hence we must have $0 \le \kappa < 1$. Then a second application of Dyson's lemma in the form (6.2) implies that

$$\frac{3}{2}\theta\tau + 3\theta\tau\kappa^2 \le 1 + \frac{7}{2}\min\left\{\frac{M}{N}, \frac{N}{M}\right\}.$$

Next we suppose that the pair of nonnegative integers (R, S) is critical with respect to the index at $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$. Then we select a bihomogeneous polynomial $T(\mathbf{x}, \mathbf{y})$ in $\mathbb{Z}[\mathbf{x}, \mathbf{y}]$ of bidegree (R, S) such that $T(I_2\boldsymbol{\lambda}_1, I_2\boldsymbol{\lambda}_2) \ne 0$. An advantageous choice for T is not apparent so we simply take T to be a monomial. Thus we let $T(\mathbf{x}, \mathbf{y}) = x_0^r x_1^{R-r} y_0^s y_1^{S-s}$, where $0 \le r \le R$ and $0 \le s \le S$. It follows from our remarks at the end of Section 7 that

$$\{T(\mathcal{D})F\}(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) \ne 0.$$

From (6.4) we conclude that

$$1-\kappa \le \operatorname{index}\left(T(\mathcal{D})F;\ \begin{bmatrix}0\\1\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}, \theta M, \tau N\right)$$

and similarly at $\left(\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}\right)$ and $\left(\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}1\\0\end{bmatrix}\right)$. Thus we may apply the global estimate (5.4) to $\{T(\mathcal{D})F\}(\mathbf{x}, \mathbf{y})$ with (M, N) replaced by $(M-R, N-S)$, with θ replaced by $(1-\kappa)(M-R)^{-1}\theta M$ and with τ replaced by $(1-\kappa)(N-S)^{-1}\tau N$. In this way we obtain the inequality

$$\begin{aligned}0 \le\,&(M-R)\log\left\{2^{1/2}\eta(P)H(\boldsymbol{\lambda}_1)\right\} + (N-S)\log\left\{2^{1/2}\eta(P)H(\boldsymbol{\lambda}_2)\right\}\\ &+\frac{1}{2}\log\left\{(M+1)(N+1)\right\} + \sum_w \log|T(\mathcal{D})F|_w\\ &+(1-\kappa)\max\left\{\theta M\log\delta_v\left(\begin{bmatrix}0\\1\end{bmatrix}, \boldsymbol{\lambda}_1\right), \tau N\log\delta_v\left(\begin{bmatrix}0\\1\end{bmatrix}, \boldsymbol{\lambda}_2\right)\right\}.\end{aligned} \tag{8.4}$$

We may assume without loss of generality that $T(\mathcal{D})F$ has relatively prime integer coefficients. From the estimate in (v) in Theorem 13 and the expansion (6.5), we find that a generic coefficient satisfies the bound

$$\begin{aligned}&\log\left|f_{m+r,n+s}\binom{m+r}{r}\binom{M-m-r}{R-r}\binom{n+s}{s}\binom{N-n-s}{S-s}\right|\\ &\le \log|f_{m+r,n+s}| + (m+r)\psi\left(\frac{r}{m+r}\right) + (M-m-r)\psi\left(\frac{R-r}{M-m-r}\right)\\ &\quad+(n+s)\psi\left(\frac{s}{n+s}\right) + (N-n-s)\psi\left(\frac{S-s}{N-n-s}\right)\\ &\le \frac{1}{4}\left(1-\frac{3}{2}\theta\tau\right)^{-1}(\theta M+\tau N) + o(M) + o(N) + M\psi\left(\frac{R}{M}\right) + N\psi\left(\frac{S}{N}\right).\end{aligned} \tag{8.5}$$

Using (8.5) we conclude that

$$\begin{aligned}\sum_w \log|T(\mathcal{D})F|_w &\le \sum_w \log[T(\mathcal{D})F]_w\\ &\le \frac{1}{4}\left(1-\frac{3}{2}\theta\tau\right)^{-1}(\theta M+\tau N) + (M+N)\log 2 + o(M) + o(N).\end{aligned} \tag{8.6}$$

After minor simplifications (8.4) and (8.6) imply that

$$\begin{aligned}0 \le\,& M\log\{8H(\boldsymbol{\lambda}_1)\} + N\log\{8H(\boldsymbol{\lambda}_2)\} + \frac{1}{4}\left(1-\frac{3}{2}\theta\tau\right)^{-1}(\theta M+\tau N)\\ &+(1-\kappa)\max\left\{\theta M\log\delta_v\left(\begin{bmatrix}0\\1\end{bmatrix}, \boldsymbol{\lambda}_1\right), \tau N\log\delta_v\left(\begin{bmatrix}0\\1\end{bmatrix}, \boldsymbol{\lambda}_2\right)\right\}\\ &+o(M)+o(N).\end{aligned} \tag{8.7}$$

We select

$$\theta = \left\{\frac{2(1-t)e_v(\boldsymbol{\lambda}_2)}{3e_v(\boldsymbol{\lambda}_1)}\right\}^{1/2}, \quad \tau = \left\{\frac{2(1-t)e_v(\boldsymbol{\lambda}_1)}{3e_v(\boldsymbol{\lambda}_2)}\right\}^{1/2},$$

where $0 < t < \frac{1}{4}$. Then $0 < \theta < 1$, $0 < \tau < 1$, and $\frac{1}{2} < \frac{2}{3}(1-t) = \theta\tau < \frac{2}{3}$. And we let $M \to \infty$, $N \to \infty$ in such a way that

$$\frac{M}{N} \to \frac{\log\{8H(\boldsymbol{\lambda}_2)\}}{\log\{8H(\boldsymbol{\lambda}_1)\}}.$$

By compactness we can restrict (M, N) to a suitable subsequence along which $\kappa \to \kappa^*$ with

$$(1-t) + 2(1-t)(\kappa^*)^2 \le 1 + \frac{7}{2}r(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2). \tag{8.8}$$

With these choices for θ, τ, M and N the inequality (8.7) implies that

$$(1-\kappa^*)\left\{\frac{2}{3}(1-t)e_v(\boldsymbol{\lambda}_1)e_v(\boldsymbol{\lambda}_2)\right\}^{1/2} \le 2 + (4t)^{-1}s(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2). \tag{8.9}$$

By combining (8.8) and (8.9) we obtain the bound

$$\begin{aligned}\{e_v(\boldsymbol{\lambda}_1)e_v(\boldsymbol{\lambda}_2)\}^{1/2} &\le 6^{1/2}\left\{\frac{1 + (8t)^{-1}s(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)}{(1-t)^{1/2} - \left(\frac{1}{2}t + \frac{7}{4}r(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)\right)^{1/2}}\right\} \\ &\le 6^{1/2}\left\{\frac{1 + (8t)^{-1}s(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)}{1 - \left(\frac{3}{2}t + \frac{7}{4}r(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)\right)^{1/2}}\right\}.\end{aligned} \tag{8.10}$$

Finally, we select $t = (24)^{-1/3}s(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)^{2/3}$ and use the simple inequality

$$\left(\frac{1+x}{1-y}\right)^2 \le 1 + 2(1-b)^{-2}(x+y), \tag{8.11}$$

which holds on the square

$$0 \le x \le b < 1, \quad 0 \le y \le b < 1.$$

From (8.2) we find that (8.11) can be applied with $b = 3 - 2^{3/2}$. After a brief calculation we conclude that

$$\begin{aligned} &e_v(\boldsymbol{\lambda}_1)e_v(\boldsymbol{\lambda}_2) \\ &\le 6\left\{1 + 2(1-b)^{-2}\left\{\frac{1}{4}(3s(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2))^{1/3} + \left\{\frac{1}{4}(3s(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2))^{2/3} + \frac{7}{4}r(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)\right\}^{1/2}\right\}\right\} \\ &\le 6\left\{1 + (1-b)^{-2}\left\{\frac{3}{2}(3s(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2))^{1/3} + (7r(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2))^{1/2}\right\}\right\} \\ &= 6\left\{1 + C_s s(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)^{1/3} + C_r r(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)^{1/2}\right\}.\end{aligned}$$

9. – Some effective measures of irrationality

We now show how to construct special cubic extensions to which Theorem 5 can be applied. Let

$$f(x) = x^3 + px + q\,, \qquad p \neq 0\,, \tag{9.1}$$

be an irreducible polynomial in $k[x]$. Let L be the splitting field of f, write

$$f(x) = (x-\alpha)(x-\alpha')(x-\alpha'')$$

in $L[x]$, and set

$$\Delta = (\alpha-\alpha')(\alpha'-\alpha'')(\alpha''-\alpha)\,.$$

Then the discriminant of f is $\Delta^2 = -4p^3 - 27q^2$. As in Section 3, let

$$\boldsymbol{\alpha} = \begin{bmatrix}\alpha\\1\end{bmatrix}, \quad \boldsymbol{\alpha}' = \begin{bmatrix}\alpha'\\1\end{bmatrix}, \quad \boldsymbol{\alpha}'' = \begin{bmatrix}\alpha''\\1\end{bmatrix}$$

be points in $\mathbb{P}^1(L)$. Then the unique element $\mathbf{\Phi}$ in $\mathrm{PGL}(2, L)$ which satisfies (3.1) is given explicitly by

$$\mathbf{\Phi} = \begin{bmatrix}(\alpha'-\alpha)^{-1} & -\alpha(\alpha'-\alpha)^{-1}\\ -(\alpha''-\alpha')^{-1} & \alpha''(\alpha''-\alpha')^{-1}\end{bmatrix}. \tag{9.2}$$

We find that $\mathbf{\Phi}^{-1}\begin{bmatrix}-\alpha\\ \alpha''\end{bmatrix} = \begin{bmatrix}3q\\ -2p\end{bmatrix}$ and therefore

$$\mathbf{\Phi}\begin{bmatrix}3q\\ -2p\end{bmatrix} = \begin{bmatrix}-\alpha\\ \alpha''\end{bmatrix} = \boldsymbol{\lambda}_1 \tag{9.3}$$

belongs to $\Lambda(P)$.

Next we introduce the polynomials

$$\begin{aligned} r_1(x) &= \left(x+\frac{\alpha'}{\alpha}\right)\left(x+\frac{\alpha''}{\alpha'}\right)\left(x+\frac{\alpha}{\alpha''}\right)\\ &= \left(x^3 - \frac{3}{2}x^2 - \frac{3}{2}x + 1\right) + \frac{\Delta}{2q}(x^2 - x)\,,\end{aligned}$$

$$\begin{aligned} r_2(x) &= \left(x+\frac{\alpha}{\alpha'}\right)\left(x+\frac{\alpha'}{\alpha''}\right)\left(x+\frac{\alpha''}{\alpha}\right)\\ &= \left(x^3 - \frac{3}{2}x^2 - \frac{3}{2}x + 1\right) - \frac{\Delta}{2q}(x^2 - x)\,,\end{aligned}$$

and their homogenizations

$$R_1(\mathbf{x}) = x_1^3 r_1\left(\frac{x_0}{x_1}\right), \quad R_2(\mathbf{x}) = x_1^3 r_2\left(\frac{x_0}{x_1}\right).$$

LEMMA 14. *Let w be a place of L. If $w \nmid \infty$ then*

$$\log[R_1]_w = \log[R_2]_w = \frac{1}{2}\log^+\left|\frac{p^3}{q^2}\right|_w, \tag{9.4}$$

if $w \mid \infty$ then

$$\begin{aligned}\log[R_1]_w &= \log[R_2]_w \\ &= \frac{\varepsilon_w}{2}\log\left\{\frac{7}{2}+\frac{2}{3}\left\|\frac{p^3}{q^2}+\frac{27}{4}\right\|_w\right\}.\end{aligned} \tag{9.5}$$

Moreover, the polynomials $r_1(x)$ and $r_2(x)$ are irreducible in $k(\Delta)[x]$.

PROOF. Assume that $w \nmid \infty$. Then $[R_1]_w = [R_2]_w$ is obvious and by Gauss' lemma we have

$$\log[R_1]_w = \frac{1}{2}\log[R_1R_2]_w . \tag{9.6}$$

Using the identity

$$\left(x^3-\frac{3}{2}x^2-\frac{3}{2}x+1\right)^2+\frac{27}{4}(x^2-x)^2=(x^2-x+1)^3$$

we find that

$$\begin{aligned}R_1(\mathbf{x})R_2(\mathbf{x}) &= x_0^6-3x_0^5x_1+\left(\frac{p^3}{q^2}+6\right)x_0^4x_1^2 \\ &\quad-\left(2\frac{p^3}{q^2}+7\right)x_0^3x_1^3+\left(\frac{p^3}{q^2}+6\right)x_0^2x_1^4-3x_0x_1^5+x_1^6 .\end{aligned} \tag{9.7}$$

The identity (9.4) follows from (9.6) and (9.7). If $w \mid \infty$ then

$$\begin{aligned}[R_1]_w &= \left\{1+\frac{1}{3}\left\|\frac{\Delta}{2q}-\frac{3}{2}\right\|_w^2+\frac{1}{3}\left\|\frac{\Delta}{2q}+\frac{3}{2}\right\|_w^2+1\right\}^{\varepsilon_w/2} \\ &= \left\{\frac{7}{2}+\frac{2}{3}\left\|\frac{\Delta}{2q}\right\|_w^2\right\}^{\varepsilon_w/2} \\ &= \left\{\frac{7}{2}+\frac{2}{3}\left\|\frac{p^3}{q^2}+\frac{27}{4}\right\|_w\right\}^{\varepsilon_w/2},\end{aligned}$$

and an identical calculation applies to R_2.

The subgroup $\mathrm{Gal}(L/k(\Delta)) \subseteq \mathrm{Gal}(L/k)$ is cyclic of order 3 and acts transitively on the roots of f. It follows that $\mathrm{Gal}(L/k(\Delta))$ acts transitively on the roots of r_1. If r_1 has a root in $k(\Delta)$ then $\mathrm{Gal}(L/k(\Delta))$ fixes this root and we conclude that r_1 has one root of multiplicity 3. But the discriminant of r_1 is $p^6q^{-4} \neq 0$ and the contradiction shows that r_1 (and similarly for r_2) is irreducible in $k(\Delta)[x]$.

LEMMA 15. *Let $\boldsymbol{\lambda}_1$ in $\Lambda(P)$ be defined by* (9.3). *Then we have*

$$\begin{aligned}\frac{1}{6}\sum_{u\nmid\infty}\log^+\left|\frac{p^3}{q^2}\right| + \sum_{u\mid\infty}\frac{d_u}{6d}\log\left\{\frac{7}{2}+\frac{2}{3}\left\|\frac{p^3}{q^2}+\frac{27}{4}\right\|_u\right\}\\ \le \log H(\boldsymbol{\lambda}_1)\\ \le \frac{1}{2}+\frac{1}{6}\sum_{u\nmid\infty}\log^+\left|\frac{p^3}{q^2}\right|_u + \sum_{u\mid\infty}\frac{d_u}{6d}\log\left\{\frac{7}{8}+\frac{1}{6}\left\|\frac{p^3}{q^2}+\frac{27}{4}\right\|_u\right\},\end{aligned}$$

where each sum is over places u of k.

PROOF. It follows from Lemma 14 that $r_1(x)$ is the minimal polynomial of $-\alpha/\alpha''$ over the field $k(\Delta)$. By using the transitive action of $\mathrm{Gal}(L/k(\Delta))$ on places of L which lie over a fixed place of $k(\Delta)$ we find that the three points $\begin{bmatrix}-\alpha\\ \alpha''\end{bmatrix}$, $\begin{bmatrix}-\alpha''\\ \alpha'\end{bmatrix}$, and $\begin{bmatrix}-\alpha'\\ \alpha\end{bmatrix}$, in $\mathbb{P}^1(L)$ have the same height. Then we apply (4.9) and (4.10) to $R_1(\mathbf{x})$ and sum over all places w of L. In this way we obtain the inequality

$$\sum_w \log[R_1]_w \le 3\log H(\boldsymbol{\lambda}_1) \le \left(\frac{3}{2}-\log 2\right) + \sum_w \log[R_1]_w . \tag{9.9}$$

If u is a place of k, $u\nmid\infty$, then Lemma 14 implies that

$$\begin{aligned}\sum_{w\mid u}\log[R_1]_w &= \frac{1}{2}\sum_{w\mid u}\log^+\left|\frac{p^3}{q^2}\right|_w\\ &= \frac{1}{2}\log^+\left|\frac{p^3}{q^2}\right|_u .\end{aligned}\tag{9.10}$$

If $u\mid\infty$ then

$$\begin{aligned}\sum_{w\mid u}\log[R_1]_w &= \frac{1}{2}\sum_{w\mid u}\varepsilon_w\log\left\{\frac{7}{2}+\frac{2}{3}\left\|\frac{p^3}{q^2}+\frac{27}{4}\right\|_w\right\}\\ &= \frac{d_u}{2d}\log\left\{\frac{7}{2}+\frac{2}{3}\left\|\frac{p^3}{q^2}+\frac{27}{4}\right\|_u\right\}.\end{aligned}\tag{9.11}$$

The lemma follows by combining (9.9), (9.10) and (9.11).

Let v be a place of k. We assume now that f has a root in k_v and we identify L with an embedding of L in $\bar{k}_v$ such that

$$|\alpha|_v \le |\alpha'|_v \le |\alpha''|_v . \tag{9.12}$$

LEMMA 16. *Let $\boldsymbol{\lambda}_1$ in $\Lambda(P)$ be defined by* (9.3). *If $v\nmid\infty$ then*

$$-\log\delta_v\left(\begin{bmatrix}0\\1\end{bmatrix},\boldsymbol{\lambda}_1\right) = \frac{1}{2}\log^+\left|\frac{p^3}{q^2}\right|_v , \tag{9.13}$$

and if $v \mid \infty$ *then*

$$\begin{aligned}\frac{d_v}{2d}\log\left\{\frac{7}{8}+\frac{1}{6}\left\|\frac{p^3}{q^2}+\frac{27}{4}\right\|_v\right\} &\le -\log\delta_v\left(\begin{bmatrix}0\\1\end{bmatrix},\boldsymbol{\lambda}_1\right)\\ &\le \frac{3d_v}{2d}+\frac{d_v}{2d}\log\left\{\frac{7}{8}+\frac{1}{6}\left\|\frac{p^3}{q^2}+\frac{27}{4}\right\|_v\right\}.\end{aligned} \tag{9.14}$$

PROOF. Our embedding of L into $\bar{k}_v$ determines a place $\tilde{w}$ of L such that $\tilde{w} \mid v$ and $|\ |_{\tilde{w}} = |\ |_v$ on $\bar{k}_v$. If $v \nmid \infty$ then (4.9) and (9.4) imply that

$$\frac{1}{2}\log^+\left|\frac{p^3}{q^2}\right|_{\tilde{w}} = \log[R_2]_{\tilde{w}} = \log^+\left|\frac{\alpha}{\alpha'}\right|_{\tilde{w}} + \log^+\left|\frac{\alpha'}{\alpha''}\right|_{\tilde{w}} + \log^+\left|\frac{\alpha''}{\alpha}\right|_{\tilde{w}}. \tag{9.15}$$

We renormalize (9.15) with respect to v and use (9.12). It follows that

$$\frac{1}{2}\log^+\left|\frac{p^3}{q^2}\right|_v = \log^+\left|\frac{\alpha''}{\alpha}\right|_v = -\log\delta_v\left(\begin{bmatrix}0\\1\end{bmatrix},\boldsymbol{\lambda}_1\right).$$

If $v \mid \infty$ then (4.10) and (9.5) show that

$$\begin{aligned}\frac{\varepsilon_{\tilde{w}}}{2}\log\left\{\frac{7}{2}+\frac{2}{3}\left\|\frac{p^3}{q^2}+\frac{27}{4}\right\|_{\tilde{w}}\right\} &= \log[R_2]_{\tilde{w}}\\ &\le \frac{\varepsilon_{\tilde{w}}}{2}\log\left\{1+\left\|\frac{\alpha}{\alpha'}\right\|_{\tilde{w}}^2\right\}+\frac{\varepsilon_{\tilde{w}}}{2}\log\left\{1+\left\|\frac{\alpha'}{\alpha''}\right\|_{\tilde{w}}^2\right\}\\ &\quad+\frac{\varepsilon_{\tilde{w}}}{2}\log\left\{1+\left\|\frac{\alpha''}{\alpha}\right\|_{\tilde{w}}^2\right\}\\ &\le \frac{3\varepsilon_{\tilde{w}}}{2}+\frac{\varepsilon_{\tilde{w}}}{2}\log\left\{\frac{7}{8}+\frac{1}{6}\left\|\frac{p^3}{q^2}+\frac{27}{4}\right\|_{\tilde{w}}\right\}.\end{aligned}$$

Again we renormalize with respect to v and use (9.12). The desired result (9.14) follows as before because

$$-\log\delta_v\left(\begin{bmatrix}0\\1\end{bmatrix},\boldsymbol{\lambda}_1\right) = \frac{d_v}{2d}\log\left\{1+\left\|\frac{\alpha''}{\alpha}\right\|_v^2\right\}.$$

Plainly Lemma 15 and Lemma 16 provide a rather precise upper bound for the measure of irrationality

$$\mu_v^*(\boldsymbol{\lambda}_1)=\left(-\log\delta_v\left(\begin{bmatrix}0\\1\end{bmatrix},\boldsymbol{\lambda}_1\right)\right)^{-1}\left(6\log\{8H(\boldsymbol{\lambda}_1)\}+19\left(\log\{8H(\boldsymbol{\lambda}_1)\}\right)^{2/3}\right), \tag{9.16}$$

which occurs in Theorem 5. In order to apply Theorem 5 we assume that among the roots of f, which are arranged so as to satisfy (9.12), α is a root in k_v.

Of course there are various ways to insure that this will happen. For example, if $v \nmid \infty$, $|q|_v < |p|_v$ and $1 \le |p|_v$, then it follows using Hensel's lemma that f has a unique root in $\{x \in k_v : |x|_v \le |p^{-1}q|_v\}$. As $\alpha + \alpha' + \alpha'' = 0$ we find that the remaining roots satisfy $|\alpha|_v < |\alpha'|_v = |\alpha''|_v$. If $v \mid \infty$ and exactly one root belongs to k_v then $k_v \cong \mathbb{R}$ and $\Delta^2 < 0$. In this case the condition $0 < p$ implies that the unique real root α satisfies $|\alpha|_v < |\alpha'|_v = |\alpha''|_v$.

To simplify our notation we define

$$h_u(p^3/q^2) = \log^+ \left|\frac{p^3}{q^2}\right|_u$$

at all places u of k. Then

$$h(p^3/q^2) = \sum_u h_u(p^3/q^2)$$

denotes the logarithmic Weil height of p^3/q^2.

THEOREM 17. *Let $f(x) = x^3 + px + q$, $p \ne 0$, be irreducible in $k[x]$ and let $K = k(\alpha) \subseteq k_v$ be a cubic extension of k generated by the root α of f. Assume that the conjugate roots α' and α'' of f in $\bar{k}_v$ satisfy $|\alpha|_v \le |\alpha'|_v \le |\alpha''|_v$. And assume that*

$$3000 \le h(p^3/q^2) \le \frac{3}{2} h_v(p^3/q^2). \tag{9.17}$$

Then the effective measure of irrationality $\mu_v^*(\lambda_1)$ satisfies

$$\mu_v^*(\lambda_1) \le \frac{2h(p^3/q^2)}{h_v(p^3/q^2)} + \frac{14}{h(p^3/q^2)^{1/3}}. \tag{9.18}$$

PROOF. From Lemma 15 we have

$$6\log\{8H(\lambda_1)\} \le 2h(p^3/q^2) + \frac{65}{4},$$

and Lemma 16 leads to the estimate

$$h_v(p^3/q^2) - \frac{1}{2}\log\left(\frac{54}{7}\right) \le -\log \delta_v\left(\begin{bmatrix}0\\1\end{bmatrix}, \lambda_1\right).$$

The theorem follows now by a straightforward calculation using (9.16) and (9.17).

REFERENCES

[1] A. BAKER, *Rational approximation to $\sqrt[3]{2}$ and other algebraic numbers*, Quart. J. Math. Oxford **15** (1964), 375-383.

[2] A. BAKER, *Simultaneous rational approximations to certain algebraic numbers*, Proc. Camb. Phil. Soc. **63** (1967), 693-702.

[3] A. BAKER, *Linear forms in the logarithms of algebraic numbers I, II, III, IV*, Mathematika **13** (1966), 204-16; **14** (1967), 102-107, 220-228; **15** (1968), 204-216.

[4] A. BAKER - C.L. STEWART, *On effective approximations to cubic irrationals*, New Advances in Transcendence Theorey, A. Baker, ed., Cambridge University Press, 1988, 1-24.

[5] E. BOMBIERI, *On the Thue-Siegel-Dyson Theorem*, Acta Math. **148** (1982), 255-296.

[6] E. BOMBIERI, *Effective Diophantine approximation on G_m*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **20** (1993), n. 1, 61-89.

[7] E. BOMBIERI - J. MUELLER, *On effective measures of irrationality for $\sqrt[r]{a/b}$ and related numbers*, J. Reine Angew. Math. **342** (1983), 173-196.

[8] E. BOMBIERI - J. VAALER, *On Siegel's Lemma*, Invent. Math. **73** (1983), 11-32.

[9] K.-K. CHOI, preprint.

[10] G.V. CHUDNOVSKY, *On the method of Thue-Siegel*, Ann. of Math. (2) **117** (1983), 325-382.

[11] N.I. FELDMAN, *An effective refinement of the exponent in Liouville's theorem*, (Russian), Izv. Akad. Nauk **35** (1971), 973-990. Also: Math. USSR Izv. **5** (1971), 985-1002.

[12] A. NÉRON, Modèles Minimaux des Variétés Abéliennes sur les Corps Locaux et Globaux, IHES Publications Mathématiques, n. 21, Presses Universitaires de France, 1964.

[13] C.G. PINNER - J.D. VAALER, *The Number of Irreducible Factors of a Polynomial, I*, Trans. Amer. Math. Soc. (1993), 809-834.

[14] R.S. RUMELY, Capacity Theory on Algebraic Curves, Lecture Notes in Mathematics 1378, Springer-Verlag, New York, 1989.

[15] W.M. SCHMIDT, *Simultaneous Approximation to Algebraic Numbers by Elements of a Number field*, Monatsh. Math. **79** (1975), 55-66.

[16] A. THUE, *Über Annäherungswerte algebraischer Zahlen*, J. Reine Angew. Math. **136** (1909), 284-305.

[17] S.M. TYLER, The Lagrange Spectrum in Projective Space over a Local Field, Ph.D. Dissertation, The University of Texas at Austin, 1994.

[18] P. VOJTA, *Dyson's lemma for products of two curves of arbitrary genus*, Invent. Math. **98** (1989), 107-113.

School of Mathematics
Institute for Advanced Study
Princeton, New Jersey 08540 USA

School of Mathematics, Physics,
Computing and Electronics
Macquarie University
NSW 2109 Australia

Department of Mathematics
The University of Texas at Austin
Austin, Texas 78712 USA