

ANNALI DELLA
SCUOLA NORMALE SUPERIORE DI PISA
Classe di Scienze

ENRICO BOMBIERI

Effective diophantine approximation on \mathbb{G}_m

Annali della Scuola Normale Superiore di Pisa, Classe di Scienze 4^e série, tome 20, n° 1 (1993), p. 61-89

http://www.numdam.org/item?id=ASNSP_1993_4_20_1_61_0

© Scuola Normale Superiore, Pisa, 1993, tous droits réservés.

L'accès aux archives de la revue « Annali della Scuola Normale Superiore di Pisa, Classe di Scienze » (<http://www.sns.it/it/edizioni/riviste/annaliscienze/>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

Effective Diophantine Approximation on \mathbb{G}_m

ENRICO BOMBIERI

1. - Introduction

In this paper we introduce a new method for studying the problem of obtaining effective irrationality measures for division points of high order in the multiplicative group $\mathbb{G}_m(K)$, with K a number field, or in other words for roots of high order of algebraic numbers. There are applications to effective diophantine approximation in a finitely generated subgroup Γ of $\mathbb{G}_m(K)$, and our results provide a new effective solution of Thue's equation in number fields and a new proof of the Baker-Feldman theorem [F] giving an effective improvement in the exponent in Liouville's theorem on rational approximations to algebraic numbers. Until now, these questions could be attacked in full generality only by means of Baker's theory of linear forms in logarithms [Ba].

Our main tool is the Thue-Siegel Principle, which, roughly speaking, asserts that, if β is a "sufficiently good" rational approximation to an algebraic number α , then α admits a non-trivial effective irrationality measure determined by the pair (α, β) , which we call an "anchor pair". The first result in this direction is already in Thue's work ([T], Theorem III, p. 249), but the first completely explicit formulation of this principle, also in the p -adic case, is due to Mahler ([M], Hilfssatz 3, p. 709). A stronger, but less explicit, result is in Gelfond ([G], Theorem 1, p. 22); Hyvärinen [H] also obtained a quantitative form explicit in all constants.

The main difficulty with these early results was that the conditions required for (α, β) to be an anchor pair were so strong that no examples of pairs could be found satisfying them (indeed, it is unlikely that there are any), so that the Thue-Siegel Principle remained without applications. In 1982 the author [Bo], by exploiting a forgotten result of Dyson [Dy], obtained a new form of the Thue-Siegel Principle with the first explicit examples of anchor pairs. This was further improved in [Bo-M] and [Bo2]. As in the previous work, this paper is still based on a very refined construction of anchor pairs.

Some of the ideas here originate from a long-time collaboration with Jeff Vaaler and Alfred Van der Poorten. In particular, this joint work led to

an improved Thue-Siegel Principle, which may be described as an equivariant form of it with respect to a Galois group action, and will be treated in detail in a future joint paper. Here we use in a fundamental way the special case of a cyclic group action, and in fact Theorem 1 is to some extent the outcome of an attempt to optimize the Equivariant Thue-Siegel Principle in the case of roots of algebraic numbers.

Our results in this paper are completely explicit, although no attempt has been made here to get the sharpest constants, hoping to return to the subject in the future.

Finally, we wish to thank the Eidgenössische Technische Hochschule of Zürich for providing a stimulating atmosphere and financial support during the preparation of this paper, and W.M. Schmidt for pointing out some inaccuracies in a first version.

2. - The method of Siegel

In 1929 Siegel proved the finiteness of the number of integral points on affine curves of positive genus or of genus 0 with at least three distinct points at infinity. Some of the new ideas he introduced are essential ingredients for the application of our results to Thue's equation and more generally to effective diophantine approximation in finitely generated multiplicative groups. Consider for example Thue's equation over the rational integers. A first easy reduction shows that it suffices to deal with the equation $F(X, Y) = b$, where F is an irreducible form over \mathbb{Z} , of degree at least 3. We decompose F into linear factors, obtaining

$$F(X, Y) = a_0 \prod_{i=1}^n (X - \alpha_i Y) = b$$

and now it is easy to see, from the fact that the α_i 's are distinct, that if X, Y are large, then one factor is small and the others are large, of the same order as X or Y . Thus, for some i , we obtain an exceptionally good approximation to α_i , namely

$$(2.1) \quad \left| \alpha_i - \frac{X}{Y} \right| \leq c(F) \frac{1}{|Y|^n}$$

for an explicitly computable positive constant $c(F)$. Now α_i has degree $n \geq 3$, and if we had infinitely many solutions we would contradict Thue's theorem that an algebraic number of degree n has irrationality type at most $n/2 + 1$, or for that matter Roth's celebrated theorem that the irrationality type of an irrational real algebraic number is precisely 2.

Thue's equation over number fields may be treated by similar methods but the reduction to a diophantine approximation problem is more difficult. This

was done by Mahler [M], who obtained at the same time an important p -adic generalization of the Thue equation, nowadays called the Thue-Mahler equation.

The new ideas of Siegel in [S] deal with the reduction of the diophantine equation to a problem in the diophantine approximation of algebraic numbers.

Let C be an affine algebraic curve defined over K and let P_i be an infinite sequence of integral points on C . The sequence P_i cannot have a finite accumulation point on C and therefore any accumulation point P_∞ on the projective closure \overline{C} of C must be a component of the divisor at infinity of \overline{C} . Thus P_∞ is algebraic over K . Now, using again the fact that the P_i 's are integral points, it is easy to see that in a suitable embedding of \overline{C} in complex projective space we have

$$(2.2) \quad d(P_{i_\nu}, P_\infty) \ll H(P_{i_\nu})^{-\kappa}$$

for a subsequence $\{i_\nu\}$ and some fixed $\kappa > 0$. Here $d(P, Q)$ is the distance in a complex neighborhood of the point P_∞ of \overline{C} , and $H(P)$ is the height of the point P .

This inequality may be considered the analogue of (2.1) for C , but now we cannot apply the diophantine approximation immediately because in general the exponent κ is too small; for example, the direct application of Roth's theorem needs $\kappa > 2$.

The new idea is to exploit the fact that the points P_i lie on a curve of positive genus. To this end, we note that by extending K we may suppose that P_∞ is defined over K , and then embed \overline{C} into its Jacobian variety J by means of the map $P \mapsto cl(P - P_\infty)$, so that P_∞ goes to the origin O of the abelian variety J . Let $J(K)$ be the Mordell-Weil group of rational points of J ; by the Mordell-Weil theorem, the group $J(K)/rJ(K)$ is finite. Let us choose elements U_1, \dots, U_s in $J(K)$, one for each coset of $rJ(K)$. Then for every $P \in J(K)$ we can write $P = rQ - U_i$ for some U_i and some $Q \in J(K)$.

The inequality (2.2) can be written, in a suitable complex embedding of J , as $d(P, O) \ll H(P)^{-\kappa}$ and therefore we get $d(rQ, U_i) \ll H(P)^{-\kappa}$. Since multiplication by r on J is étale, we may divide by r obtaining

$$d\left(Q, \frac{1}{r}U_i\right) \ll H(P)^{-\kappa}$$

with $\frac{1}{r}U_i$ a suitable r -division point of U_i . The theory of heights on abelian varieties shows that

$$H(P) \gg H(Q)r^2$$

with an implied constant depending on r and U_i . If we combine the last two displayed inequalities, we deduce that

$$(2.3) \quad d\left(Q, \frac{1}{r}U_i\right) \ll H(Q)^{-\kappa r^2},$$

thus replacing κ with κr^2 . The point $\frac{1}{r} U_i$ is algebraic of degree at most r^{2g} over K , where g is the genus of C . If r is large enough, this yields a reinforcement of (2.2) bringing it into the range where we can use our knowledge about diophantine approximation. In fact, Roth's theorem can be used as soon as $\kappa r^2 > 2$.

Since Roth's theorem was not available in Siegel's time, Siegel had to overcome considerable difficulties with the diophantine approximation part of the argument. If $g = 1$, it turned out that the earlier improvement by Siegel of Thue's result was enough to conclude the proof. If instead $g > 1$, Siegel noticed that (2.3) could be viewed as a statement in simultaneous diophantine approximation rather than approximation to a single algebraic number, and then concluded by an extension of the Thue-Siegel theorem to the case of simultaneous approximations.

We can summarize Siegel's proof as follows:

- (i) embedding C into a group variety G , in such a way that the points P we are studying are a subset of a finitely generated subgroup Γ of G and converge to the identity of G at least as fast as $H(P)^{-\kappa}$, for some $\kappa > 0$;
- (ii) using the isogeny $G \xrightarrow{r} G$ and the appropriate cosets in $\Gamma/r\Gamma$ to obtain approximations Q in $G(K)$ to suitable r -division points of Γ , which converge as fast as $H(Q)^{-\kappa'}$, where κ' is much larger than κ ;
- (iii) applying diophantine approximation techniques to the last situation.

Siegel himself used this scheme also in the case $g = 0$. In this case, the Jacobian variety is trivial. On the other hand, if the support of the divisor at infinity has at least three distinct points, we can choose two such points P'_∞, P''_∞ distinct from the accumulation point P_∞ . It suffices to deal with the situation where everything is non-singular and now $G = \overline{C} - P'_\infty - P''_\infty$ can be given the structure of a linear torus with identity P_∞ . The group Γ is the group $G(R)$ where R is the ring of integers of K , and κ' is now $\kappa' = \kappa r$. The finite generation of the group $G(R)$ follows from Dirichlet's unit theorem, and the rest of the argument is as before.

3. - Statement of results

Our notation is as follows. By $H(\cdot)$ and $h(\cdot)$ we shall denote the absolute Weil height and the absolute logarithmic Weil height. Absolute values $|\cdot|_v$ in a field K will be normalized, as in our earlier papers, by requiring that

$$|x|_v = \|x\|_v^{[K_v:\mathbb{Q}_v]/[K:\mathbb{Q}]}$$

with $\|\cdot\|_v$ the unique extension to the complete field K_v of the ordinary real or p -adic absolute value in \mathbb{Q}_v . With this normalization we have the useful

formula

$$h(x) = \sum_v \log^+ |x|_v$$

where $\log^+(a) = \max(\log(a), 0)$ and, for any extension K' of K , we have

$$\sum_{w|v} \log |x|_w = \log |x|_v$$

for every $x \in K$, $x \neq 0$, where the sum runs over all places w in K' lying over v .

In some cases it will be convenient to work in an extension K' of K with a normalized absolute value $|\cdot|_v$ of K , suitably extended to K' . Such an extension of $|\cdot|_v$ will be denoted by $|\cdot|_{\bar{v}}$.

The first part of this paper deals with the problem of effective irrationality types for an algebraic number $\alpha = \sqrt[r]{a}$, where a is an element of an algebraic number field K , $a \neq 0$.

This problem was considered first by Baker [Ba2] using Padé approximation techniques (see however [T2], for some early work in a closely related direction). This method succeeds only on the assumption that $|\alpha - 1|$ is also rather small, although it has been successful with several interesting special numbers. For example, Baker [Ba2] used $\frac{4}{5} \sqrt[3]{2} = \sqrt[3]{128/125} \cong 1.008$ to obtain for the first time a non-trivial irrationality type for $\sqrt[3]{2}$. See also [C] for further refinements of the Padé method.

Using the method of linear forms in logarithms, Baker [Ba3] obtained an effective bound

$$\mu_{\text{eff}}(\sqrt[r]{a}; \mathbb{Q}, \infty) \ll_a \log r,$$

and the explicit bound

$$\mu_{\text{eff}}(\sqrt[r]{a}; \mathbb{Q}, \infty) \leq c_1 h(a) \log r,$$

for an explicit constant c_1 , follows from the recent work of Baker and Wüstholz [Ba-W]. This result is very good if a is fixed and r very large, and is non-trivial as soon as $r \gg h(a) \log h(a)$.

Our main result on effective approximations to roots, given here in a form explicit in all constants, is non-trivial if $r \gg h(a)$. This improved range of r is essential for applications to diophantine approximation on a finitely generated subgroup Γ of $\mathbb{G}_m(K)$.

THEOREM 1. *Let K be a number field of degree d over the rational field \mathbb{Q} and let $a \in K$, $a \neq 0$. Let r be a positive integer and let $\alpha = \sqrt[r]{a}$ be an r -th root of a .*

Suppose that $\kappa > 0$ and

$$r \geq e^{100d/\kappa^2} \max(h(a), 1/d).$$

Let v be an infinite place of K , with associated normalized absolute value $|\cdot|_v$. Then for every extension $|\cdot|_{\bar{v}}$ of $|\cdot|_v$ to $K(\alpha)$ and every $\gamma \in K$, $\gamma \neq 0$ with

$$h(\alpha\gamma) \geq 4900\kappa^{-4}$$

we have

$$|1 - \alpha\gamma|_{\bar{v}} \geq (2H(\alpha\gamma))^{-\kappa r}.$$

Theorem 1 is not the best obtainable by our method. For example, if $r > 2$ is a prime and $a > 0$ is a rational number not an r -th power of another rational number, one can obtain the effective irrationality type

$$\mu = \frac{1}{800} h(a) \log^8 \left(\frac{r}{h(a)} \right)$$

for the positive r -th root of a , again as soon as $r \gg h(a)$. This result and other substantial improvements of Theorem 1 will be contained in a future paper with Jeff Vaaler and Alfred Van der Poorten.

It may be worthwhile to describe the various steps in the proof of Theorem 1. Let us fix $a \in K$, $a \neq 0$. If r is large, we can choose the branch of the root so that $\alpha = \sqrt[r]{a}$ is close to 1 and $|\alpha - 1|$ is small, where $|\cdot|$ is the usual archimedean absolute value. If $\gamma^{-1} \in K$ is an approximation to α , then $\alpha' = \alpha\gamma$ is also very close to 1 and $|\alpha' - 1|$ is very small. We want to apply the Thue-Siegel Principle using $(\alpha, 1)$ as an anchor pair and conclude that $|\alpha' - 1|$ cannot be too small, thus obtaining an irrationality type for α . Unfortunately, this method succeeds only if $|\alpha - 1|$ is rather small to start with.

The new idea, found with Vaaler and Van der Poorten, is to note that not only α and α' are close to 1 but also the conjugates α_σ and $\alpha'_\sigma = \alpha_\sigma\gamma$ are equally close to a suitable r -th root of unity, depending on σ . In general, we may have an action by a group G on the pair (α, α') , and the orbit of (α, α') by G may be close to a set of approximations (β_g, β'_g) , $g \in G$, leading to an Equivariant Thue-Siegel Principle.

The fundamental construction of Thue and Siegel begins with the construction of an auxiliary polynomial vanishing at (α, α') to high order, and then proceeds by showing that it cannot vanish too much at the approximation (β, β') ; this will show that (β, β') cannot be too close to (α, α') . The non-vanishing, or small vanishing, of the polynomial is dealt with in our preceding papers by appealing to Dyson's Lemma [Dy].

Here it turns out that it is better to reverse the role of the algebraic and rational points. Thus we construct a polynomial vanishing to high order at (β, β') , and show that it cannot vanish too much at (α, α') . The main reason for doing this is that, since the polynomial is defined over K , if it vanishes a lot at (α, α') then it vanishes in the same way at $(\alpha_\sigma, \alpha'_\sigma)$, and then Dyson's Lemma shows that we gain a factor $1/\sqrt{[K(\alpha):K]}$ in the order of vanishing (the index) of the polynomial at (α, α') . Thus it appears that in our case we

would need to construct an auxiliary polynomial defined over K vanishing to high order at all points $(\varepsilon, \varepsilon)$ with ε an r -th root of unity.

On the other hand, for certain applications it is a severe limitation to treat only certain branches of $\sqrt[r]{a}$. One way out of this difficulty is replacing α by a power α^b , so to make the argument of α^b sufficiently small. As a result, the auxiliary construction is now done at the points $(\varepsilon^b, \varepsilon)$ rather than $(\varepsilon, \varepsilon)$.

This seemingly innocuous change in the auxiliary construction creates new problems, connected with the application of Dyson's Lemma. In fact one usually requires in Dyson's Lemma that the set of points should be admissible, that is all first co-ordinates should be distinct and all second co-ordinates also should be distinct, and now admissibility may fail for the set of first co-ordinates. Another problem with the use of Dyson's Lemma is that it contains an error term proportional to the number of points considered. Since the number of points here is rather large, application of Dyson's Lemma in its standard form would lead to relatively weak results and for example the Baker-Feldman theorem could not be obtained. Fortunately all these difficulties with the standard Dyson's Lemma have been completely resolved by Viola [V]. In his Main Theorem, Viola proves a refined form of Dyson's Lemma, which shows how to resolve the difficulty with multiplicities in the sets of co-ordinates and also gives the more precise error term we were seeking.

The second part of this paper deals with applications. Theorem 2 gives a result on the general problem of diophantine approximation in a number field by means of a finitely generated multiplicative subgroup, in a form suitable for applications. This is obtained by means of a reduction to Theorem 1 which uses a new variant (Lemma 4) of an argument which goes back to Stark ([St], p. 262).

THEOREM 2. *Let K be a number field of degree d over the rational field \mathbb{Q} , let Γ be a finitely generated subgroup of K^\times and let ξ_1, \dots, ξ_t be a set of generators of $\Gamma/\text{tors}(\Gamma)$.*

Let $A \in K^\times$, let v be an archimedean absolute value of K , and let $\xi \in \Gamma$ and $\kappa > 0$ be such that

$$0 < |1 - A\xi|_v \leq H(A\xi)^{-\kappa}.$$

Let us define $Q = 1$ if $t = 0$ and

$$Q = (e^{115d/\kappa^2 t})^{t+1} \prod_{i=1}^t h(\xi_i)$$

if $t \geq 1$. Then we have

$$h(A\xi) \leq \max(Qh(A), [Q]!).$$

The effective solution of the generalized unit equation in a number field,

and thereby the effective solution of Thue's equation and the proof of the Baker-Feldman theorem, follows easily from Theorem 2.

4. - The auxiliary polynomial

In this section we construct polynomials P which vanish to high order at certain points. Let D^I stand for the partial derivative of order $I = (i_1, i_2)$ with respect to (x_1, x_2) .

DEFINITION. Let K be a field of characteristic 0 and let (β_1, β_2) be a point with co-ordinates in K . Let $M_1 > 0$, $M_2 > 0$ and let $P(x_1, x_2) \in K[x_1, x_2]$. The index of P at (β_1, β_2) relative to (M_1, M_2) is by definition

$$\text{ind}_{(\beta_1, \beta_2)}(P; M_1, M_2) = \min \left\{ \frac{i_1}{M_1} + \frac{i_2}{M_2} \mid D^I P(\beta_1, \beta_2) \neq 0 \right\}.$$

LEMMA 1. Let $0 < \theta_i \leq 1$, $i = 1, 2$ and let $T = \theta_1 \theta_2 / 2$ satisfy $rT < 1$. Let also b be an integer. Then there is a polynomial P , with rational integral coefficients, not identically 0, of degree $\deg_{x_i}(P) \leq N_i$, with index

$$\text{ind}_{(\varepsilon^b, \varepsilon)}(P; \theta_1 N_1, \theta_2 N_2) \geq 1$$

at every point $(\varepsilon^b, \varepsilon)$ with ε an r -th root of unity, and with height at most

$$h(P) \lesssim \frac{rT \log 2}{1 - rT} (N_1 + N_2).$$

The above asymptotic inequality holds as $N_1 \rightarrow \infty$ and $N_2 \rightarrow \infty$, keeping θ_1 and θ_2 fixed.

PROOF. We need to solve the linear system

$$\sum_{j_1=0}^{N_1} \sum_{j_2=0}^{N_2} \binom{j_1}{i_1} \binom{j_2}{i_2} \varepsilon^{bj_1+j_2-bi_1-i_2} p_{j_1 j_2} = 0$$

in rational integers $p_{j_1 j_2}$ not all 0, for all pairs (i_1, i_2) with

$$(4.1) \quad \frac{i_1}{\theta_1 N_1} + \frac{i_2}{\theta_2 N_2} < 1$$

and all r -th roots of unity ε .

We transform this system in a system with rational integral coefficients by multiplying it by $\varepsilon^{-a+bi_1+i_2}$ and summing over all roots of unity ε , for

$a = 0, \dots, r - 1$. After division by r , the system becomes

$$\sum_{j_1=0}^{N_1} \sum_{j_2=0}^{N_2} \binom{j_1}{i_1} \binom{j_2}{i_2} P_{j_1 j_2}$$

$b_{j_1+j_2} \equiv a \pmod{r}$

for all pairs (i_1, i_2) as in (4.1) and $a = 0, \dots, r - 1$. The conclusion of Lemma 1 follows from $\binom{n}{i} \leq 2^n$ and Siegel's Lemma, [S], p. 213. A much better estimate, which turns out to have interesting consequences, follows from Struppeck and Vaaler [S-V].

5. - Application of Viola's Theorem

Let P be as in the preceding section. We are interested in the vanishing of P at another point.

Let K be a number field, let $a \in K$, $a \neq 0$ and let $\gamma \in K$, $\gamma \neq 0$. Let α , α' be algebraic numbers

$$\alpha = \sqrt[r]{a}, \quad \alpha' = \alpha\gamma$$

and let $K' = K(\alpha)$.

LEMMA 2. *Let θ_i , T , b and P be as in Lemma 1 and suppose that α' is not an r -th root of unity. Suppose further that $b \neq 0$, $\theta_1 \geq 2|b|\theta_2$ and*

$$(5.1) \quad 1 - rT + r \frac{\theta_2 N_2}{\theta_1 N_1} < [K' : K]T.$$

Then the polynomial P in Lemma 1 has index at (α^b, α') bounded by

$$\text{ind}_{(\alpha^b, \alpha')}(P; \theta_1 N_1, \theta_2 N_2) \leq \sqrt{\left(1 - rT + r \frac{\theta_2 N_2}{\theta_1 N_1}\right) / ([K' : K]T)}.$$

PROOF. The polynomial P has index at least 1 at $(\varepsilon^b, \varepsilon)$ by construction. Now let c be the index of P at (α^b, α') , let $T(c)$ be the area of the region

$$\left\{ 0 < x_1 < 1, \quad 0 < x_2 < 1, \quad \frac{x_1}{\theta_1} + \frac{x_2}{\theta_2} < c \right\}$$

and let

$$\tilde{c} = \min(c, \theta_1^{-1}).$$

The polynomial P again has index c at all points $(\alpha_\sigma^b, \alpha'_\sigma)$ where $\alpha'_\sigma = \alpha_\sigma \gamma$ and α_σ runs over the $[K' : K]$ conjugates of α over K .

We apply Viola's Main Theorem ([V], p. 109). For the reader's convenience, we restate a special case of Viola's result in our notation, in a slightly weaker form which is amply sufficient for our purposes.

VIOLA'S THEOREM. *Let P be a polynomial with complex coefficients of bidegree at most (N_1, N_2) , not identically 0, with index at least c_h with respect to $(\theta_1 N_1, \theta_2 N_2)$ at distinct points (x_h, x'_h) , $h = 1, \dots, m$. Let ρ_h be the number of occurrences of x'_h among x'_1, \dots, x'_m and let σ_h be the number of occurrences of x_h among x_1, \dots, x_m . Suppose further that*

$$(C) \quad c_h \leq \min(\rho_h^{-1} \theta_1^{-1}, \sigma_h^{-1} \theta_2^{-1})$$

for $h = 1, \dots, m$.

Then we have

$$\sum_{h=1}^m T(c_h) \leq 1 + \max\left(\frac{m}{2} - 1, 0\right) \min(1, \theta_2/\theta_1) \frac{N_2}{N_1}.$$

In our case, we choose $c_h = 1$ for the points $(\varepsilon^b, \varepsilon)$ and $c_h = \bar{c}$ for the points $(\alpha_\sigma^b, \alpha'_\sigma)$. Then the quantities ρ_h and σ_h appearing in Viola's Theorem satisfy $\rho_h = 1$ for all h because the numbers ε and α'_σ are all distinct, while $\sigma_h \leq 2|b|$ because the numbers ε^b and α_σ^b cannot be repeated more than $2|b|$ times. Now the hypothesis $\theta_1 \geq 2|b|\theta_2$ shows that condition (C) of Viola's Theorem is satisfied if $\bar{c} \leq \theta_1^{-1}$, which is the case by definition of \bar{c} . Thus we can apply Viola's Theorem, obtaining

$$rT + [K' : K]T(\bar{c}) \leq 1 + r \frac{\theta_2 N_2}{\theta_1 N_1},$$

that is

$$(5.2) \quad T(\bar{c}) \leq \frac{1 - rT}{[K' : K]} + \frac{r}{[K' : K]} \frac{\theta_2 N_2}{\theta_1 N_1}.$$

From (5.2), $\theta_2 \leq \theta_1$, $\theta_2/\theta_1 = 2T\theta_1^{-2} > T$ and the hypothesis (5.1) we find

$$T(\bar{c}) \leq \frac{1}{2} \min(\theta_1/\theta_2, \theta_2/\theta_1)$$

and this implies that $T(\bar{c}) = T\bar{c}^2$. Now (5.2) shows that

$$\bar{c} \leq \sqrt{\left(1 - rT + r \frac{\theta_2 N_2}{\theta_1 N_1}\right) / ([K' : K]T)}$$

and therefore (5.1) yields $\bar{c} < 1 \leq \theta_1^{-1}$. Thus $\bar{c} = c$, completing the proof of Lemma 2.

6. - Application of the equivariant Thue-Siegel principle

Let P be the polynomial constructed in the preceding section and let (i_1^*, i_2^*) be such that

$$(6.1) \quad \frac{1}{i_1^*! i_2^*!} D^{I^*} P(\alpha^b, \alpha') \neq 0$$

with

$$(6.2) \quad \frac{i_1^*}{\theta_1 N_1} + \frac{i_2^*}{\theta_2 N_2} = \text{ind}_{(\alpha^b, \alpha')} (P; \theta_1 N_1, \theta_2 N_2) = c$$

and let us abbreviate P^* for the polynomial

$$(6.3) \quad P^*(x_1, x_2) = \frac{1}{i_1^*! i_2^*!} D^{I^*} P(x_1, x_2).$$

Differentiation in (6.1) multiplies the coefficients of P by a product of appropriate binomial coefficients, majorized by $2^{N_1+N_2}$. Hence by Lemma 1 we get

$$(6.4) \quad h(P^*) \leq h(P) + (\log 2)(N_1 + N_2) \lesssim \frac{\log 2}{1 - \tau T} (N_1 + N_2).$$

Suppose that $c < 1$. The polynomial $D^I P^*$ vanishes at each point $(\varepsilon^b, \varepsilon)$ for I in the set Δ defined by

$$\Delta = \left\{ (i_1, i_2) \left| \frac{i_1}{\theta_1 N_1} + \frac{i_2}{\theta_2 N_2} < 1 - c \right. \right\}.$$

Since P^* does not vanish at (α^b, α') by construction, the product formula gives

$$(6.5) \quad \sum_w \log |P^*(\alpha^b, \alpha')|_w = 0$$

where the sum is over all places w of the field $K' = K(\alpha)$.

We estimate the various terms in (6.5) in different ways. If w is a finite place we have

$$\log |P^*(\alpha^b, \alpha')|_w \leq \log |P^*|_w + |b| \log^+ |\alpha|_w N_1 + \log^+ |\alpha'|_w N_2.$$

If instead w is an infinite place we proceed as follows. Let v be the infinite place of K with $w|v$. The associated absolute value $|\cdot|_v$ in K_v has a unique extension, again denoted by $|\cdot|_v$, to the complex field \mathbb{C} , and in view of our normalizations we have in K'_w :

$$|\cdot|_v = |\cdot|_w^{[K':K]/[K'_w:K'_v]}.$$

Let us fix an extension $|\cdot|_{\bar{v}}$ of $|\cdot|_v$ to K' , with a corresponding embedding of $K' \subset \mathbb{C}$, determining α and α' as complex numbers. Considering a different extension of v simply means taking a different embedding of K' into \mathbb{C} , and this replaces α and α' by $\varepsilon\alpha$ and $\varepsilon\alpha'$ for some r -th root of unity ε . Thus we obtain that for every w over v there is an r -th root of unity ε such that

$$|\alpha^b - \varepsilon^b|_w^{[K':K]/[K'_w:K_v]} = |(\varepsilon\alpha)^b - \varepsilon^b|_{\bar{v}} = |\alpha^b - 1|_{\bar{v}}.$$

A corresponding calculation holds with α' in place of α^b . This shows that if we choose an extension $|\cdot|_{\bar{v}}$ of $|\cdot|_v$ from K to K' then for every $w|v$ there is an r -th root of unity ε such that

$$(6.6) \quad |\alpha^b - \varepsilon^b|_w = |\alpha^b - 1|_{\bar{v}}^{[K'_w:K_v]/[K':K]}, \quad |\alpha' - \varepsilon|_w = |\alpha' - 1|_{\bar{v}}^{[K'_w:K_v]/[K':K]}.$$

At w , we expand P^* in Taylor series around $(\varepsilon^b, \varepsilon)$, obtaining:

$$P^*(\alpha^b, \alpha') = \sum_{J \notin \Delta} \frac{1}{j_1! j_2!} D^J P^*(\varepsilon^b, \varepsilon) (\alpha^b - \varepsilon^b)^{j_1} (\alpha' - \varepsilon)^{j_2}.$$

This gives, taking into account the differentiation:

$$\begin{aligned} \log |P^*(\alpha^b, \alpha')|_w &\lesssim \log |P^*|_w + \frac{[K'_w : \mathbb{Q}_w]}{[K' : \mathbb{Q}]} (\log 2)(N_1 + N_2) \\ &\quad + \max_{J \notin \Delta} (j_1 \log |\alpha^b - \varepsilon^b|_w + j_2 \log |\alpha' - \varepsilon|_w). \end{aligned}$$

Now we sum over all places w . We find, using (6.6):

$$\begin{aligned} 0 &\lesssim h(P^*) + (|b|h(\alpha) + \log 2)N_1 + (h(\alpha') + \log 2)N_2 \\ &\quad + \sum_{v|\infty} \max_{J \notin \Delta} (j_1 \log |\alpha^b - 1|_{\bar{v}} + j_2 \log |\alpha' - 1|_{\bar{v}}). \end{aligned}$$

In taking the maximum, we can replace (j_1, j_2) by a continuous variable. Let us abbreviate

$$x = \frac{j_1}{N_1}, \quad y = \frac{j_2}{N_2}, \quad a = |\alpha^b - 1|_{\bar{v}}, \quad a' = |\alpha' - 1|_{\bar{v}}.$$

We need to maximize $x(\log a)N_1 + y(\log a')N_2$ in the convex set $x/\theta_1 + y/\theta_2 \geq 1 - c$, $x \geq 0$, $y \geq 0$. Since $a < 1$ and $a' < 1$ the maximum occurs at one of the two extreme points $(x, y) = ((1 - c)\theta_1 N_1, 0)$ or $(x, y) = (0, (1 - c)\theta_2 N_2)$. This yields the asymptotic inequality

$$\begin{aligned} (1 - c) \sum_{v|\infty} \min \left(N_1 \theta_1 \log \frac{1}{|\alpha^b - 1|_{\bar{v}}}, N_2 \theta_2 \log \frac{1}{|\alpha' - 1|_{\bar{v}}} \right) &\lesssim (|b|h(\alpha) + V)N_1 \\ &\quad + (h(\alpha') + V)N_2 \end{aligned}$$

with $V = (1 + 1/(1 - rT)) \log 2$.

Finally we let $N_2 \rightarrow \infty$ with $N_1/N_2 \rightarrow z$, and we may assume that c converges to t , say. In the limit we find:

LEMMA 3. *Let $a, \gamma \in K$, $a, \gamma \neq 0$ and let $\alpha = \sqrt[r]{a}$, $\alpha' = \alpha\gamma$. Suppose that α' is not an r -th root of unity. Let $K' = K(\alpha)$, and for each archimedean absolute value v of K let $|\cdot|_{\bar{v}}$ be an extension of $|\cdot|_v$ to K' . Let also b be a non-zero integer.*

Suppose further that $0 < \theta_1 \leq 1$, $0 < \theta_2 \leq 1$, $T = \theta_1\theta_2/2$, $rT < 1$, $z > 0$, $\theta_1 \geq 2|b|\theta_2$ and

$$t = \sqrt{\left(1 - rT + \frac{\theta_2 r}{\theta_1 z}\right) / ([K' : K]T)} < 1.$$

Then we have, with $V = (1 + 1/(1 - rT)) \log 2$:

$$(1 - t) \sum_{v|\infty} \min \left(z\theta_1 \log \frac{1}{|\alpha^b - 1|_{\bar{v}}}, \theta_2 \log \frac{1}{|\alpha' - 1|_{\bar{v}}} \right) \leq (|b|h(\alpha) + V)z + h(\alpha') + V.$$

7. - Proof of Theorem 1

In this section we complete the proof of Theorem 1.

Let $|\cdot|_{\bar{v}}$, α , $\alpha' = \alpha\gamma$ be as in the preceding section and define $d = [K : \mathbb{Q}]$, $\delta = [K(\alpha) : K]/r$. Our proof is by contradiction, hence we assume the hypotheses

$$(H1) \quad r \geq e^{100d/\kappa^2} \max(h(a), 1/d)$$

and

$$(H2) \quad h(\alpha') \geq 4900\kappa^{-4}$$

of Theorem 1, and negate the conclusion, thus

$$(A1) \quad |\alpha' - 1|_{\bar{v}} < (2H(\alpha'))^{-\kappa r}.$$

Further, we have positive parameters B , θ_1 , θ_2 , T , z and an integer b with $1 \leq |b| \leq B$, to be chosen later, which satisfy the following conditions:

$$(A2) \quad B \geq 1, \quad 0 < 2B\theta_2 \leq \theta_1 \leq 1,$$

$$(A3) \quad T = \theta_1\theta_2/2, \quad rT < 1,$$

$$(A4) \quad z > 0, \quad 1 - rT + \frac{\theta_2 r}{\theta_1 z} < \frac{1}{9} \theta_1 r T.$$

We choose first $\theta_1 = \kappa$ and note that (A1) and the Liouville inequality, for example as formulated in [Bo2], p. 26, give

$$(7.1) \quad \delta \geq \theta_1.$$

Next, we note that (A4) and (7.1) yield $t = \sqrt{(1 - rT + \theta_2 r / \theta_1 z) / (\delta r T)} < 1/3$. Thus the conditions of Lemma 3 are verified and we get

$$\frac{2}{3} \min \left(z\theta_1 \log \frac{1}{|\alpha^b - 1|_{\bar{v}}}, \theta_2 \log \frac{1}{|\alpha' - 1|_{\bar{v}}} \right) \leq (|b|h(\alpha) + V)z + h(\alpha') + V,$$

which by (A1), $\kappa = \theta_1$ and $\theta_1\theta_2 = 2T$ implies

$$(7.2) \quad \frac{2}{3} \min \left(z\theta_1 \log \frac{1}{|\alpha^b - 1|_{\bar{v}}}, 2rTh(\alpha') \right) \leq (Bh(\alpha) + V)z + h(\alpha') + V.$$

Let $|\cdot|$ be the ordinary absolute value and let K be embedded in the complex field by means of the inclusion $K_v \subset \mathbb{C}$. The extension \bar{v} of v to K' determines an embedding of K'_v in \mathbb{C} and corresponding complex values for α , α' .

The conclusion of Theorem 1 is that α' cannot be too close to 1 in this embedding. Suppose this is not the case. We would like to be able to say that α , being a root of a very high order, is close to 1 and apply Lemma 3 to conclude the proof. However we cannot do this in general because the argument of α is not controlled *a priori*, unless for example $K(\alpha)$ is a real field. This difficulty can be overcome by replacing α by a suitable power α^b , chosen in such a way that the argument of α^b becomes small and $|\alpha^b| \leq 1$. The details are as follows.

Suppose first that $K'_v = \mathbb{C}$ and let $|\cdot|$ refer to the absolute value in \mathbb{C} and in K'_v , by means of this identification. We have

$$-\frac{d}{r} h(a) \leq \log |\alpha| \leq \frac{d}{r} h(a).$$

Next, let $\phi = \arg(\alpha)/2\pi$ and let $B \geq 1$ be a real number. By Dirichlet's Theorem there are integers l and $1 \leq |b| \leq B$ such that $|b\phi - l| \leq \frac{1}{B}$; note also that if (b, l) is a solution then $(-b, -l)$ is another.

Now we have, for the determination of the argument with $-\pi < \arg \leq \pi$, the inequality $|\arg(\alpha^b)| \leq 2\pi/B$, and we choose the sign of b so that $|\alpha^b| \leq 1$. We get

$$\begin{aligned} |\alpha^b - 1|^2 &\leq 1 - 2 \cos \left(\frac{2\pi}{B} \right) |\alpha^b| + |\alpha^b|^2 = (1 - |\alpha^b|)^2 + 4 \sin^2 \left(\frac{\pi}{B} \right) |\alpha^b| \\ &\leq (1 - e^{-Bdh(a)/r})^2 + \left(\frac{2\pi}{B} \right)^2 \leq \left(\frac{Bdh(a)}{r} \right)^2 + \left(\frac{2\pi}{B} \right)^2. \end{aligned}$$

We define a modified height $h'(\cdot)$ and a parameter R by

$$h'(a) = \max(h(a), 1/d), \quad R = \frac{r}{4\pi dh'(a)},$$

choose $B = \pi\sqrt{8R}$ and get, for the corresponding b :

$$|\alpha^b - 1| \leq \frac{1}{\sqrt{R}}.$$

This can be rewritten in terms of $|\cdot|_{\bar{v}}$ as

$$(7.3) \quad \log \frac{1}{|\alpha^b - 1|_{\bar{v}}} \geq \frac{1}{d} \log R.$$

If instead $K'_v = \mathbb{R}$ we choose $b = \pm 1$, $B = 1$ if $\alpha > 0$ and $b = \pm 2$, $B = 2$ if $\alpha < 0$ and we verify, after a similar but simpler calculation, that (7.3) still holds.

In conclusion, we have shown that there is a non-zero integer b , $1 \leq |b| \leq B$ such that (7.3) holds with

$$(7.4) \quad R = r/(4\pi dh'(a)), \quad B = \pi\sqrt{8R}.$$

By (7.2) and (7.3) we find

$$(7.5) \quad \frac{2}{3} \min \left(\frac{z\theta_1}{d} \log R, 2rTh(\alpha') \right) \leq (Bh(\alpha) + V)z + h(\alpha') + V.$$

Suppose $R > 1$. If we impose to the parameter z the further condition

$$(A5) \quad z \geq \frac{2dh(\alpha')}{\theta_1 \log R},$$

then, noting that $rT < 1$, we see that the minimum in (7.5) occurs with the term $2rTh(\alpha')$. We find, after dividing by $h(\alpha')$:

$$(7.6) \quad \frac{4}{3} rT \leq 1 + (Bh(\alpha) + V) \frac{z}{h(\alpha')} + \frac{V}{h(\alpha')}.$$

Further simplifications are obtained by noting that $h(\alpha) = h(a)/r$, hence by (7.4) we have $Bh(\alpha) \leq 1/(d\sqrt{2R})$, thus

$$(7.7) \quad Bh(\alpha) \leq V/10$$

if $R \geq 50d^{-2}V^{-2}$, that is

$$(A6) \quad r \geq 200\pi d^{-1}V^{-2}h'(a).$$

We also assume

$$(A7) \quad h(\alpha') > 450V$$

and deduce from (7.6), (7.7) and (A7) that

$$(7.8) \quad \frac{4}{3} rT < 1 + \frac{11}{10} V \frac{z}{h(\alpha')} + \frac{1}{450}.$$

Finally we choose

$$(7.9) \quad rT = 1 - \frac{\theta_1}{12} \geq \frac{11}{12}$$

and

$$(7.10) \quad z = (5V)^{-1} h(\alpha')$$

and deduce from (7.8), (7.9) and (7.10) that

$$\frac{11}{9} < 1 + \frac{11}{50} + \frac{1}{450} = \frac{11}{9}.$$

This is a contradiction, therefore not all assumptions (A1) to (A7) can hold if z and T are determined by (7.9) and (7.10).

The proof of Theorem 1 is complete if we show that (H1), (H2), (7.9) and (7.10) imply the validity of (A2) to (A7).

First of all, we have with this choice of rT :

$$(7.11) \quad V = \left(\frac{12}{\theta_1} + 1 \right) \log 2 \leq \frac{13}{\theta_1} \log 2 < \frac{9.02}{\theta_1}.$$

Verification of (A2). Since $\theta_2 < 2/(r\theta_1)$ it suffices $r \geq 4B\theta_1^{-2}$, which follows from $dh'(a) \geq 1$, (7.4) and (H1).

Verification of (A3). Clear from (7.9).

Verification of (A4). We want

$$1 - rT + \frac{\theta_2 r}{\theta_1 z} < \frac{1}{9} \theta_1 rT$$

and hence by (7.9) it suffices to have $\theta_2 r / (\theta_1 z) < \theta_1 / 54$. By (7.9), (7.10) and $r\theta_2 = 2rT/\theta_1$ we need

$$h(\alpha') \geq 540\theta_1^{-3}V,$$

which follows from (7.11) and (H2).

Verification of (A5). We need

$$(5V)^{-1} h(\alpha') \geq \frac{2dh(\alpha')}{\theta_1 \log R},$$

thus $\log R \geq 10d\theta_1^{-1}V$ and by (7.11) this is implied by $r \geq 4\pi de^{91d/\theta^2}h'(a)$, which follows from (H1).

Verification of (A6). Clear from (H1).

Verification of (A7). Clear from (H2).

8. - Application of the isogeny $G \xrightarrow{r} G$

In this section we exploit Siegel's idea ([S], §5, eq. (111)) to obtain good rational approximations to suitable roots of an element of K .

LEMMA 4. *Let n_i , $i = 1, \dots, t$ be rational integers, let λ_i , $i = 1, \dots, t$ be positive real numbers with $\lambda_1\lambda_2\dots\lambda_t = 1$ and let N , Q be positive integers with $Q > \max \lambda_i^t$. Then there are a natural integer r and rational integers p_i , $i = 1, \dots, t$ such that*

$$|n_i - rp_i| \leq r\lambda_i Q^{-1/t}, \quad i = 1, \dots, t$$

and

$$(Q - 1)!N \leq r \leq Q!N.$$

PROOF. Let us abbreviate $\phi_i = n_i/(Q!N)$ and consider the box \mathcal{B} given by

$$(8.1) \quad |x_0| \leq Q, \quad |\phi_1 x_0 - x_1| \leq \lambda_1 Q^{-1/t}, \quad \dots, \quad |\phi_t x_0 - x_t| \leq \lambda_t Q^{-1/t}$$

in $(t + 1)$ -dimensional euclidean space with co-ordinates (x_0, x_1, \dots, x_t) . The box \mathcal{B} is convex and symmetric about the origin and has volume 2^{t+1} , therefore by Minkowski's theorem there is an integral point (q, p_1, \dots, p_t) in the box \mathcal{B} , different from the origin $(0, 0, \dots, 0)$. By taking the point $(-q, -p_1, \dots, -p_t)$ if needed, we may assume $q \geq 0$. We cannot have $q = 0$, otherwise $|p_i| \leq \lambda_i Q^{-1/t} < 1$ and $p_i = 0$ for every i , contradicting the fact that our point is not the origin. Therefore $1 \leq q \leq Q$ and $r = Q!N/q$ is a natural integer satisfying $(Q - 1)!N \leq r \leq Q!N$. The lemma follows by evaluating (8.1) at (q, p_1, \dots, p_t) and multiplying by r .

LEMMA 5. *Let K be a number field, let Γ be a finitely generated multiplicative subgroup of K^\times and let ξ_1, \dots, ξ_t be a set of generators of $\Gamma/\text{tors}(\Gamma)$. Let Q and N be positive integers such that*

$$Q > (\min h(\xi_i))^{-t} \prod_{i=1}^t h(\xi_i).$$

Let $A \in K$, $A \neq 0$, let v be an archimedean absolute value of K , and let $\xi \in \Gamma$ be such that

$$0 < |1 - A\xi|_v \leq 2^{-(\kappa+2/[K:\mathbb{Q}])Q!N} H(A\xi)^{-\kappa}.$$

Then we can find an element $a \in A\Gamma$, an element $\eta \in \Gamma$, an r -th root $\alpha = \sqrt[r]{a}$ and an extension $|\cdot|_{\bar{v}}$ of $|\cdot|_v$ to $K' = K(\alpha)$ such that $(Q-1)!N \leq r \leq Q!N$ and

$$(\alpha\eta)^r = A\xi, \quad h(\alpha\eta) = \frac{1}{r} h(A\xi),$$

$$|h(a) - h(A)| \leq rt \left(Q^{-1} \prod_{i=1}^t h(\xi_i) \right)^{1/t},$$

$$|1 - \alpha\eta|_{\bar{v}} < (2H(\alpha\eta))^{-\kappa r}.$$

PROOF. We want to find r , η and ξ_0 such that $\xi = \eta^r \xi_0$ with a small ξ_0 . We write

$$\xi = \zeta \xi_1^{n_1} \cdots \xi_t^{n_t}$$

with $\zeta \in \text{tors}(\Gamma)$ a root of unity in K and apply Lemma 4 to the vector (n_1, \dots, n_t) , choosing

$$\lambda_i = h(\xi_i)^{-1} \left(\prod_{i=1}^t h(\xi_i) \right)^{1/t}.$$

If we write $q_i = n_i - rp_i$ and define

$$\eta = \xi_1^{p_1} \cdots \xi_t^{p_t}, \quad \xi_0 = \zeta \xi_1^{q_1} \cdots \xi_t^{q_t}$$

then we have

$$(Q-1)!N \leq r \leq Q!N, \quad |q_1| \leq r\lambda_1 Q^{-1/t}, \quad \dots, \quad |q_t| \leq r\lambda_t Q^{-1/t}$$

therefore

$$(8.2) \quad h(\xi_0) \leq \sum_{i=1}^t |q_i| h(\xi_i) \leq rt \left(Q^{-1} \prod_{i=1}^t h(\xi_i) \right)^{1/t}.$$

We set $a = A\xi_0$, thus $a \in A\Gamma$ and $a\eta^r = A\xi$, and note that

$$(8.3) \quad |h(a) - h(A)| \leq h(\xi_0).$$

We choose $\alpha = \sqrt[r]{a}$ such that the complex number α so determined satisfies

$$|1 - \alpha\eta| \leq |1 - \varepsilon\alpha\eta|$$

for every r -th root of unity ε . Then

$$|1 - \varepsilon\alpha\eta| = |(1 - \varepsilon) + \varepsilon(1 - \alpha\eta)| \geq |1 - \varepsilon| - |1 - \alpha\eta| \geq |1 - \varepsilon| - |1 - \varepsilon\alpha\eta|,$$

whence

$$|1 - \varepsilon\alpha\eta| \geq \frac{1}{2} |1 - \varepsilon|$$

for every ε . Since $\prod_{\varepsilon \neq 1} |1 - \varepsilon| = r$ we get

$$(8.4) \quad |1 - \alpha\eta| \leq \frac{2^{r-1}}{r} |1 - a\eta^r| < 2^r |1 - A\xi|.$$

Since

$$(\alpha\eta)^r = a\eta^r = A\xi, \quad h(\alpha\eta) = \frac{1}{r} h(A\xi),$$

we obtain the statement of Lemma 5 from (8.2), (8.3) and (8.4).

9. - Proof of Theorem 2

Notation is as in the preceding section with $d = [K : \mathbb{Q}]$.

Suppose that the hypotheses of Lemma 5 hold:

$$(B1) \quad 0 < |1 - A\xi|_v \leq 2^{-(\kappa+2/d)Q!N} H(A\xi)^{-\kappa},$$

$$(B2) \quad Q > (\min h(\xi_i))^{-t} \prod_{i=1}^t h(\xi_i).$$

Note also that (B1) implies $\kappa \leq 1$, otherwise we would contradict the Liouville inequality. The conclusion of Lemma 5 is the statement

$$|1 - \alpha\eta|_v < (2H(\alpha\eta))^{-\kappa r}$$

and the conclusion of Theorem 1, with α as in Lemma 5 and $\gamma = \eta$, is

$$|1 - \alpha\eta|_v \geq (2H(\alpha\eta))^{-\kappa r}.$$

This is a contradiction and it follows that the hypotheses of Theorem 1 cannot be all satisfied. Therefore if we assume

$$(B3) \quad r \geq e^{100d/\kappa^2} h'(a),$$

which is the first hypothesis of Theorem 1, the second hypothesis cannot hold. This proves

$$h(\alpha\eta) \leq 4900\kappa^{-4}$$

and since $h(\alpha\eta) = h(A\xi)/r$ we deduce

$$(9.1) \quad h(A\xi) \leq 4900\kappa^{-4}r.$$

Thus we have shown that (B1), (B2) and (B3) imply (9.1).

Let us abbreviate

$$M = e^{100d/\kappa^2}$$

and suppose that

$$(B4) \quad Q \geq (2Mt)^t \prod_{i=1}^t h(\xi_i).$$

Then Lemma 5 shows that

$$h'(a) \leq h'(A) + \frac{1}{2M} r$$

and therefore (B3) follows from

$$r \geq M \left(h'(A) + \frac{1}{2M} r \right) = Mh'(A) + \frac{1}{2} r,$$

that is from

$$(B3)' \quad r \geq 2Mh'(A).$$

We choose

$$N = [2Mh'(A)/(Q-1)] + 1$$

and recall that by Lemma 5 we have $(Q-1)!N \leq r \leq Q!N$. Then it is clear that (B3)' holds. We also have

$$(9.2) \quad r \leq Q!N \leq 2MQh'(A) + Q!.$$

We have thus shown that (B1), (B2) and (B4) imply (9.1), with r bounded by (9.2).

Suppose first that

$$h'(A) \geq \frac{1}{2M} (Q-1)!.$$

Then (9.2) gives $r \leq Q!N \leq 4MQh'(A)$ and we get from (9.1):

$$(9.3) \quad h(A\xi) \leq 4900\kappa^{-4} 4MQh'(A).$$

We claim that in this case (9.3) also follows from (B2), (B4) and the simpler

$$(B1)' \quad |1 - A\xi|_{\bar{v}} < H(A\xi)^{-(1+\lambda)\kappa}$$

with $\lambda = \log 2/1600$, instead of (B1). In fact, suppose that (9.3) does not hold. Then we must have

$$\lambda\kappa h(A\xi) \geq 3(\log 2)\kappa^{-3} 4MQh'(A) \geq (\kappa + 2/d)(\log 2)Q!N$$

and this shows that

$$H(A\xi)^{-(1+\lambda)\kappa} \leq 2^{-(\kappa+2/d)Q!N} H(A\xi)^{-\kappa},$$

so that condition (B1), and hence (9.3), holds.

Suppose instead that

$$(9.4) \quad h'(A) < \frac{1}{2M} (Q-1)!.$$

Then $N = 1$, $r \leq Q!$, so that (9.1) yields

$$(9.5) \quad h(A\xi) \leq 4900\kappa^{-4}Q!.$$

Again, we claim that in this case (9.5) follows from (B2), (B4) and the simpler (B1)' instead of (B1). In fact, suppose that (9.5) does not hold. Then we must have

$$\lambda\kappa h(A\xi) \geq 3(\log 2)Q! \geq (\kappa + 2/d)(\log 2)Q!N$$

and we conclude as in the preceding case.

Condition (B2) follows from (B4) because for any $c \in K^\times$ we have $h(c) \geq (9d^3)^{-1}$ unless c is a root of unity (see Dobrowolski, [Do]). Thus we may choose

$$Q = \left[(2Mt)^t \prod_{i=1}^t h(\xi_i) \right] + 1.$$

Then (B2) and (B4) hold, and therefore either (9.3) or (9.5) follow as a consequence of (B1)' alone. Theorem 2 follows replacing κ with $\kappa/(1+\lambda)$ and replacing Q with the larger constant used in Theorem 2.

10. - Thue's equation and the Baker-Feldman theorem

Let us consider the Thue equation

$$(10.1) \quad a_0X^n + a_1X^{n-1}Y + \cdots + a_nY^n = b$$

with coefficients in the ring of integers R of a number field K of degree d over \mathbb{Q} , to be solved in algebraic integers $X, Y \in R$. We may assume that $a_0 \neq 0$ and then, after multiplying by a_0^{n-1} and replacing X by $X' = a_0X$, we may also assume $a_0 = 1$. We decompose the resulting binary form into linear factors and obtain the equation

$$(10.2) \quad F(X, Y) = X^n + a_1X^{n-1}Y + \cdots + a_nY^n = \prod_{i=1}^n (X - \alpha_i Y) = m,$$

again to be solved in integers of the field K . As usual, we assume that $n \geq 3$ and that at least three of the α_i 's are distinct, say α_i, α_j and α_k . We shall also assume, in case that at least one α_l does not lie in K , that α_i and α_k are conjugates over K .

Let $K' = K(\alpha_i, \alpha_j, \alpha_k)$. It is well-known that equation (10.2) can be reduced to solving a finite set of equations of special type, namely the so-called unit equation

$$(10.3) \quad A\xi + B\xi' = 1$$

with $A, B \in K'$, to be solved in units ξ, ξ' of the field K' . The method of reduction goes back to Siegel [S2], who applied it to the hyperelliptic equation $y^2 = ax^n + bx^{n-1} + \dots + k$ as well as to the so-called superelliptic equation and Thue's equation ([S2], p. 68).

It turns out that in this reduction it is more efficient to work with a finite index subgroup of the group of units, rather than the full unit group. The details are as follows.

Let K' be as before and let $K_l = K(\alpha_l)$.

Let us abbreviate $Z_l = X - \alpha_l Y$, $l = 1, \dots, n$. Then each Z_l is an algebraic integer, therefore the product formula shows that

$$\sum_{w|\infty} \log |Z_l|_w \geq 0$$

for every l , where w runs over all infinite places of K' . Now the equation $Z_1 \cdots Z_n = m$ gives

$$\sum_{l=1}^n \sum_{w|\infty} \log |Z_l|_w = \sum_{w|\infty} \log |m|_w$$

and we infer that

$$\sum_{w|\infty} \log |Z_l|_w \leq \sum_{w|\infty} \log |m|_w.$$

Because of our normalizations for absolute values, we can descend the fields of definition in the last inequality to $K_l = K(\alpha_l)$ in the first sum and to K in the second sum. Expressing the sums in terms of norms we get

$$(10.4) \quad \frac{1}{[K_l : \mathbb{Q}]} \log |N_{K_l/\mathbb{Q}}(Z_l)| \leq \frac{1}{[K : \mathbb{Q}]} \log |N_{K/\mathbb{Q}}(m)| \leq h(m).$$

LEMMA 6. *Let K be a number field of degree d over the rational field \mathbb{Q} , and let R_K be its regulator. Let t be the rank of the group of units of K and suppose that $t \geq 1$.*

Then there are t multiplicatively independent units ξ_i , $i = 1, \dots, t$ in K such that

$$\prod_{i=1}^t h(\xi_i) \leq (t/2d)^t \sqrt{t} R_K.$$

Moreover, if $\Gamma \in K^\times$ is the group generated by the ξ_i 's, then for every $a \in K^\times$ there is an element $\xi \in \Gamma$ such that

$$|\log |a\xi|_v| \leq d^{-1} t^{1/2} (9td^2/2)^{t-1} R_K + \frac{1}{d} \log |N_{K/\mathbb{Q}}(a)|_v$$

for every infinite place v of K . In particular, if a is an integer in K we have

$$h(a\xi) \leq d^{-1} t^{3/2} (9td^2/2)^{t-1} R_K + \frac{1}{d} \log |N_{K/\mathbb{Q}}(a)|.$$

PROOF. We give here a simple proof using Geometry of Numbers. Let ρ_i , $i = 1, \dots, t$ be a basis of the group of units of K modulo torsion. Let Y be the matrix

$$Y = (\log |\rho_i|_v)$$

with rows indexed by the $t+1$ infinite places v of K and columns indexed by $i = 1, \dots, t$. Since the sum of the rows is the zero vector, for every $t \times t$ minor Y_i , $i = 1, \dots, t+1$ of Y we have

$$|\det(Y_i)| = d^{-t} R$$

(the factor d^{-t} comes from our normalization of $|\cdot|_v$).

Now let S be the symmetric convex set in \mathbb{R}^t defined by

$$S = \{\mathbf{x} \in \mathbb{R}^t \mid Y\mathbf{x} \in C\}$$

where C is the unit cube $|x_i| \leq 1/2$ in \mathbb{R}^{t+1} . Let $V(S)$ be the volume of S . Then

$$V(S) \geq t^{-1/2} d^t R_K^{-1},$$

as one sees applying a cube slicing theorem of Vaaler (see [Bo-V], (4.5), p. 24).

Let $\lambda_1, \dots, \lambda_t$ be the successive minima of S for the standard lattice \mathbb{Z}^t . It is immediate that there are linearly independent units ξ_1, \dots, ξ_t such that

$$(10.6) \quad |\log |\xi_i|_v| \leq \frac{1}{2} \lambda_i$$

for every infinite place v . By Minkowski's Second Theorem, we have

$$(10.7) \quad \lambda_1 \lambda_2 \cdots \lambda_t \leq 2^t V(S)^{-1} \leq (2/d)^t \sqrt{t} R_K.$$

It is clear from (10.6) that $h(\xi_i) = \frac{1}{2} \sum_{v|\infty} |\log |\xi_i|_v| \leq t\lambda_i/4$ and we conclude from (10.7) that

$$\prod_{i=1}^t h(\xi_i) \leq (t/2d)^t \sqrt{t} R_K.$$

This proves the first part of Lemma 6.

Next, we note that by the already quoted result by Dobrowolski [Do] we have

$$h(\xi_1) \geq 1/(9d^3).$$

It follows that $\lambda_1 \geq 4/(9td^3)$ and therefore (10.7) shows that

$$\lambda_i^{t-i+1} \leq (9td^3/4)^{i-1} (2/d)^t \sqrt{t} R_K$$

and in particular

$$\frac{1}{2} \lambda_t \leq d^{-1} (9td^2/2)^{t-1} \sqrt{t} R_K.$$

Let Γ be the multiplicative group generated by these units ξ_i . Then $\{\log |x_i|_v, v|\infty\}$ for $\xi \in \Gamma$ is a lattice Λ in the hyperplane $\sum_{v|\infty} x_v = 0$ of the euclidean space $\{x_v, v|\infty\}$. For any $a \in K^\times$, the point

$$\left\{ \log |a|_v - \frac{1}{d} \log |N_{K/\mathbb{Q}}(a)|_v, v|\infty \right\}$$

belongs to this hyperplane. We translate this point by an element of Λ so that it falls into a fundamental domain satisfying (10.6), and Lemma 6 follows from the bound we have obtained for λ_t . Sharper numerical estimates can be obtained using better lower bounds for $h(\xi_1)$ and the results in Siegel [S3].

Using (10.4), we apply Lemma 6 to the points Z_l , $l = i, j, k$ and the fields K_l and find units η_l in K_l such that the integers $\gamma_l = Z_l \eta_l$ have height at most

$$(10.8) \quad H(\gamma_l) \leq e^{c(K_l)} H(m),$$

where we have abbreviated, for a field K of degree d with unit group of rank t :

$$(10.9) \quad c(K) = d^{-1} t^{3/2} (9td^2/2)^{t-1} R_K.$$

On the other hand, Z_i, Z_j, Z_k satisfy the linear dependence condition

$$(\alpha_i - \alpha_j)Z_k + (\alpha_j - \alpha_k)Z_i + (\alpha_k - \alpha_i)Z_j = 0,$$

which we rewrite as

$$(10.3) \quad A\xi + B\xi^t = 1$$

with

$$(10.10) \quad A = \frac{\alpha_j - \alpha_k}{\alpha_j - \alpha_i} \frac{\gamma_i}{\gamma_k}, \quad B = \frac{\alpha_k - \alpha_i}{\alpha_j - \alpha_i} \frac{\gamma_j}{\gamma_k}, \quad \xi = \frac{\eta_k}{\eta_i}, \quad \xi' = \frac{\eta_k}{\eta_j}.$$

LEMMA 7. *Let A, B, ξ, ξ' be as before and let $H(F)$ denote the height of the vector $(1, a_1, \dots, a_n)$ of coefficients of the Thue equation. Let also (X, Y) be a solution with $Y \neq 0$.*

Then we have

$$h(A), h(B) \leq 2 \max_l c(K_l) + \log(8n) + 2h(F) + 2h(m)$$

and

$$h(X/Y) \leq h(A\xi) + 3h(F) + \frac{3}{2} \log(2n) + 5 \log 2.$$

PROOF. We estimate

$$\begin{aligned} h(A) &= h\left(\frac{\alpha_j - \alpha_k}{\alpha_j - \alpha_i} \frac{\gamma_i}{\gamma_k}\right) \leq h(\alpha_j - \alpha_k) + h(\alpha_j - \alpha_i) + h(\gamma_i) + h(\gamma_k) \\ &\leq h(\alpha_i) + 2h(\alpha_j) + h(\alpha_k) + 2 \log 2 + c(K_i) + c(K_k) + 2h(m). \end{aligned}$$

A similar estimate holds for $h(B)$. Thus the required upper bounds for $h(A), h(B)$ follow from the well-known estimate

$$\sum_{l=1}^n h(\alpha_l) \leq h(F) + \frac{1}{2} \log(2n),$$

see for instance [Bo2], Theorem 1 and eq. (14), (15), p. 21-22.

The upper bound for $h(X/Y)$ is obtained by rather similar considerations and the identity

$$X/Y = \alpha_k + (\alpha_i - \alpha_k) \left/ \left(1 - \frac{\alpha_j - \alpha_i}{\alpha_j - \alpha_k} A\xi \right) \right.$$

LEMMA 8. *Let A, B, ξ, ξ' be as in Lemma 7.*

Then there is an infinite place v of the field $K' = K(\alpha_i, \alpha_j, \alpha_k)$ such that

$$|1 - A\xi|_v \leq 2^{\varepsilon(v)} H(B)^{1+\varepsilon(v)} H(A\xi)^{-\varepsilon(v)}$$

with $\varepsilon(v) = [K'_v : \mathbb{Q}_v] / [K' : \mathbb{Q}]$.

PROOF. Since ξ' is a unit of K' we have

$$\sum_{v|\infty} \log |\xi'|_v = 0, \quad \sum_{v|\infty} \log^+ |\xi'|_v = h(\xi')$$

and it follows that there is a place $v|\infty$ such that

$$\log |\xi'|_v \leq -\varepsilon(v)h(\xi')$$

where as usual $\varepsilon(v) = [K'_v : \mathbb{Q}_v]/[K' : \mathbb{Q}]$. This gives

$$|1 - A\xi|_v = |B\xi'|_v = |B|_v|\xi'|_v \leq H(B)H(\xi')^{-\varepsilon(v)}.$$

Finally, we have

$$H(A\xi) = H(1 - B\xi') \leq 2H(B)H(\xi')$$

and Lemma 8 follows from the last two displayed inequalities.

THEOREM 3. *Let*

$$F(X, Y) = X^n + a_1X^{n-1}Y + \dots + a_nY^n = m$$

be a Thue equation with integral coefficients in a number field K of degree d , to be solved in integers of K . Let $\alpha_i, \alpha_j, \alpha_k$ be three distinct roots of the polynomial $F(X, 1)$.

Let R be the maximum of the regulators of the fields $K(\alpha_i), K(\alpha_j), K(\alpha_k)$ and let

$$Q = e^{500d^4n^{12}}R.$$

Then every solution (X, Y) of the Thue equation satisfies

$$h(X/Y) \leq Q(h(m) + h(F)) + [Q]!.$$

PROOF. By Lemma 8, we have

$$(10.11) \quad |1 - A\xi|_v \leq 2^{\varepsilon(v)}H(B)^{1+\varepsilon(v)}H(A\xi)^{-\varepsilon(v)} \leq H(A\xi)^{-\varepsilon(v)/2}$$

provided

$$(10.12) \quad H(A\xi) \geq 4H(B)^{2+2/\varepsilon(v)}.$$

We apply Theorem 2 to (10.11) with $\kappa = \varepsilon(v)/2$ and Γ a subgroup of the units of $K' = K(\alpha_i, \alpha_j, \alpha_k)$, with rank t and small generators.

Suppose first that $[K' : K] \geq 2$. Then we may suppose that $[K(\alpha_i) : K] \geq 2$ and α_k is a conjugate of α_i . Let Γ' be the subgroup of units of $K_i = K(\alpha_i)$ constructed in Lemma 6, with generators ξ'_1, \dots, ξ'_t . Thus Z_k is a conjugate of Z_i , which shows that in (10.8) we can take the unit η_k to be the corresponding conjugate of η_i . This implies that, if ξ''_1, \dots, ξ''_t are the corresponding conjugates of ξ'_1, \dots, ξ'_t , then $\xi = \eta_k/\eta_i$ belongs to the group Γ generated by $\xi_i = \xi'_i/\xi''_i$, $i = 1, \dots, t$. We also have $h(\xi'_i) = h(\xi''_i)$, therefore $h(\xi_i) \leq 2h(\xi'_i)$ and we obtain

$$(10.13) \quad \prod_{i=1}^t h(\xi_i) \leq \sqrt{t}R_{K(\alpha_i)}$$

in this case. If instead $K' = K$, we take Γ to be the subgroup of units of K constructed in Lemma 6, and (10.13) holds *a fortiori*.

Now (10.11), $[K' : \mathbb{Q}] \leq dn^3$, $\kappa \geq 1/(2dn^3)$ and Theorem 2 show that, if (10.12) holds, then

$$h(A\xi) \leq \max(Qh(A), [Q]!)$$

with $Q = 1$ if $t = 0$ and

$$Q = (e^{460d^3n^9}t)^{t+1} \prod_{i=1}^t h(\xi_i)$$

otherwise, where ξ_i the basis of the group Γ satisfying (10.13). If (10.12) does not hold, then we still get an estimate for $h(A\xi)$ and in any case we have

$$(10.14) \quad h(A\xi) \leq \max(Qh(A), [Q]!, (2d+2)h(B) + \log 4).$$

Now Theorem 3 follows from (10.13), (10.14) and Lemma 8, replacing Q by the larger constant $e^{500d^4n^{12}}R$ and performing some simplifications.

The following corollary is a quantitative form of the Baker-Feldman theorem. A much more explicit result is in Györy and Papp [G-P].

COROLLARY. *Let α be real algebraic of degree $n \geq 3$ and let R_K be the regulator of the field $K = \mathbb{Q}(\alpha)$.*

Then there is an effectively computable constant $c(\alpha) > 0$ such that, for all integers p, q with $q \geq 1$, we have

$$\left| \alpha - \frac{p}{q} \right| \geq c(\alpha)q^{\delta-n}$$

where

$$\delta = e^{-500n^{12}}R_K^{-1}.$$

For the proof, it suffices to apply Theorem 3 to the form $F(X, Y) = N_{\mathbb{Q}(\alpha)/\mathbb{Q}}(X - \alpha Y)$ and note that all fields $K(\alpha_i)$ have the same regulator. We leave the details to the reader.

Finally, it should be noted that the constant $c(\alpha)$ in the Corollary is very poor compared with what can be obtained using the theory of linear forms in logarithms, because of the term $[Q]!$ appearing in Theorem 3. Also the constant $e^{-500n^{12}}$ is too small, and it can be substantially improved, although as yet it is problematic to obtain anything better than n^{-Cn^3} for an absolute constant C . However the dependence of δ on the regulator is the same as what has been obtained using the sharpest known results from Baker's theory of linear forms in logarithms.

REFERENCES

- [Ba] A. BAKER, *Transcendental Number Theory*, Cambridge Univ. Press, Cambridge 1975.
- [Ba2] A. BAKER, *Rational approximations to certain algebraic numbers*, Quart. J. Math. Oxford, **15** (1964), 375-383.
- [Ba3] A. BAKER, *Recent advances in transcendence theory*, Proceedings of the International Conference in Number Theory, Moscow 14-18 September 1971, Moscow 1973, 67-69.
- [Ba-W] A. BAKER - G. WÜSTHOLZ, *Logarithmic Forms and Group Varieties*, submitted for publication.
- [Bo] E. BOMBIERI, *On the Thue-Siegel theorem*, Acta Math., **148** (1982), 255-296.
- [Bo2] E. BOMBIERI, *Lectures on the Thue Principle*, Analytic Number Theory and Diophantine Problems, Progr. Math. **70**, A.C. Adolphson, J.B. Conrey, A. Ghosh, R.I. Yager ed., Birkhäuser 1987, 15-52.
- [Bo-M] E. BOMBIERI - J. MUELLER, *On effective measures of irrationality for $\sqrt[n]{a/b}$ and related numbers*, J. Reine Angew. Math., **342** (1983), 173-196.
- [Bo-V] E. BOMBIERI - J. VAALER, *On Siegel's Lemma*, Invent. Math., **73** (1983), 11-32. Addendum, ibidem, **75** (1984), 177.
- [C] G.V. CHUDNOVSKY, *On the method of Thue-Siegel*, Ann. of Math., **117** (1983), 325-382.
- [Do] E. DOBROWOLSKI, *On the maximum modulus of conjugates of an algebraic integer*, Bull. Acad. Polon. Sci., **26** (1978), 291-292.
- [Dy] F. DYSON, *The approximation to algebraic numbers by rationals*, Acta Math., **79** (1947), 225-240.
- [F] N.I. FELDMAN, *An effective refinement in the exponent in Liouville's theorem* (Russian), Izv. Akad. Nauk SSSR Ser. Mat., **35** (1971), 985-1002.
- [G] A.O. GELFOND, *Transcendental and algebraic numbers*. English translation by L.F. Boron, Dover Publications Inc., New York 1960.
- [G-P] K. GYÖRY - Z.Z. PAPP, *Norm form equations and explicit lower bounds for linear forms with algebraic coefficients*, Studies in Pure Mathematics (to the memory of Paul Turán), Akadémiai Kiadó, Budapest and Birkhäuser Verlag, Basel 1983, 245-257.
- [H] S. HYRÖ, *Über rationale Näherungswerte algebraischer Zahlen*, Ann. Acad. Sci. Fenn. Ser. A I Math., **376** (1965), 1-15.
- [M] K. MAHLER, *Zur Approximation algebraischer Zahlen (I). Über den grössten Primteiler binärer Formen*, Math. Ann., **107** (1933), 691-730.
- [S] C.L. SIEGEL, *Über einige Anwendungen diophantischer Approximationen*, Abh. Preuß. Akad. Wissen. Phys.-Math. Kl. (1929) **1**, 1-70. Also Gesammelte Abhandlungen, Springer-Verlag, Berlin-Heidelberg-New York 1966, Bd. I, 209-274.
- [S2] C.L. SIEGEL, *The integer solutions of the equation $y^2 = ax^n + bx^{n-1} + \dots + k$* , J. London Math. Soc., **1** (1926), 66-68.
- [S3] C.L. SIEGEL, *Abschätzung von Einheiten*, Nachr. Akad. Wiss. Göttingen, Math. - Phys. Kl. (1969), **9**, 71-86.

- [St] H.M. STARK, *Further advances in the theory of linear forms in logarithms*, Diophantine Approximation and Its Applications, Charles F. Osgood ed., Academic Press, New York and London 1973, 255-293.
- [S-V] T. STRUPPECK - J.D. VAALER, *Inequalities for heights of algebraic subspaces and the Thue-Siegel Principle*, Analytic Number Theory, Proceedings of a Conference in Honor of Paul T. Bateman, B.C. Berndt, H.G. Diamond, H. Halberstam, A. Hildebrand ed., Progr. Math. **85**, Birkhäuser Boston, Boston 1990, 493-528.
- [T] A. THUE, *Über Annäherungswerte algebraischer Zahlen*, J. Reine Angew. Math., **135** (1909), 284-305.
- [T2] A. THUE, *Berechnung aller Lösungen gewisser Gleichungen von der Form $ax^r - by^r = f$* , Videnskabs-Selskabets Skrifter, I. Math.-Naturv. Kl., **4**, 9S, Christiania 1918.
- [V] C. VIOLA, *On Dyson's Lemma*, Ann. Scuola Norm. Sup. Pisa, Cl. Sci., (4) **12** (1985), 105-135.

Institute for Advanced Study
Princeton, N.J. 08540