ACTA ARITHMETICA
LXXIX.2 (1997)

# Lower bounds of heights of points on hypersurfaces

by

FRITS BEUKERS (Utrecht) and DON ZAGIER (Bonn)

*To J. W. S. Cassels with respect and admiration*

**1. Introduction.** Let us first recall Lehmer's conjecture [Le] on lower bounds for the height of an algebraic number which was stated in 1933. Let $K$ be an algebraic number field of degree $D$ over $\mathbb{Q}$. For any valuation $v$ we define $D_v = [K_v : \mathbb{Q}_v]$, where $K_v, \mathbb{Q}_v$ are the completions of $K, \mathbb{Q}$ with respect to $v$. For archimedean $v$ we normalise the valuation by $|x|_v = |x|^{D_v/D}$ where $|\cdot|$ is the ordinary complex absolute value. When $v$ is non-archimedean we take $|p|_v = p^{-D_v/D}$ where $p$ is the unique rational prime such that $|p|_v < 1$. The height of an algebraic number $\alpha \in K$ is defined by

$$H(\alpha) = \prod_v \max(1, |x|_v).$$

Because of our normalisation $H(\alpha)$ does not depend on the choice of the field $K$ in which $\alpha$ is contained. We can now state Lehmer's conjecture.

CONJECTURE 1.1. *There exists a number $c > 1$ such that for any algebraic number $\alpha$, not a root of unity and of degree $D$ we have*

$$H(\alpha)^D \geq c.$$

*Presumably $c = 1.1762808\ldots$, which is the larger real root of $x^{10} + x^9 - x^7 - x^6 - x^5 - x^4 - x^3 + x + 1$.*

The best unconditional result so far follows from work of Dobrowolski, Cantor & Straus and Louboutin [Lo], stating that there exists $\gamma > 0$ such that

$$H(\alpha)^D \geq 1 + \gamma\left(\frac{\log\log D}{\log D}\right)^3.$$

It came as a great surprise when S. Zhang [Zh] showed in 1992 that there does exist a number $c_1 > 1$ such that

$$H(\alpha)H(1-\alpha) \geq c_1$$

for all $\alpha \in \overline{\mathbb{Q}}$ such that $\alpha \neq 0, 1, \frac{1}{2} \pm \frac{1}{2}\sqrt{-3}$. This was proved by using Arakelov intersection theory on $\mathbb{P}^1$. It was almost immediately realised by one of us (see [Za]) that an elementary proof could be given which at the same time yields the best possible $c_1$, namely $\sqrt{\eta}$ where $\eta = (1+\sqrt{5})/2$, the golden ratio. The minimum is attained when $\alpha$ is minus a fifth root of unity. In [Za] there is also a generalisation of the following sort. For any $K$-rational point $P = (P_0 : P_1 : \ldots : P_n)$ in $n$-dimensional projective space $\mathbb{P}^n$ we define the height by

$$H(P) = \prod_v \max(|P_0|_v, \ldots, |P_n|_v).$$

In particular, the height of an algebraic number $\alpha$ is nothing but the projective height of $(1 : \alpha) \in \mathbb{P}^1(K)$. Then it is shown in [Za] that for any $(x_0 : x_1 : x_2) \in \mathbb{P}^2(\overline{\mathbb{Q}})$ such that $x_0 + x_1 + x_2 = 0$, $x_0x_1x_2 \neq 0$ and $(x_0 : x_1 : x_2) \neq (1 : \omega^{\pm 1} : \omega^{\mp 1})$ $(\omega^3 = 1)$, we have

$$H(x_0, x_1, x_2) \geq c_2$$

where $c_2$ is the larger real root of $x^6 - x^4 - 1$. The minimum is attained when the $x_i$ are the roots of $x^3 + x - 1$.

Inspired by [Za], H. P. Schlickewei and E. Wirsing [SW] showed the following result. Consider the line $L : \lambda x + \mu y + \nu z = 0$ in $\mathbb{P}^2$ with $\lambda\mu\nu \neq 0$. Suppose that $\lambda + \mu + \nu = 0$. Then, for any two points $P_1, P_2 \in L(\overline{\mathbb{Q}})$ with non-zero coordinates and such that $(1 : 1 : 1), P_1, P_2$ are distinct, we have

$$H(P_1)H(P_2) > \exp(1/2400) = 1.00041\ldots$$

This result was applied by Schlickewei [Schl] to estimating numbers of solutions of three term $S$-unit equations in a strikingly successful way. Although very useful, the derivation of the Schlickewei–Wirsing result did not look optimal. It is the goal of this paper to remedy this situation and also give a generalisation which encompasses the previous results. We finish the introduction by giving a description of our general setup and main result.

Consider a hypersurface $S$ of multidegree $d_1, \ldots, d_r$ on $\mathbb{P}^{n_1} \times \ldots \times \mathbb{P}^{n_r}$ given by a polynomial $F$ with coefficients in $\mathbb{Z}$. Denote the coordinates of $\mathbb{P}^{n_i}$ by $\mathbf{x}_i = (x_{i0}, x_{i1}, \ldots, x_{in_i})$. The degree of $F$ in the variable $x_{ij}$ is denoted by $d_{ij}$. We define $\widetilde{d}_i = -d_i + \sum_j d_{ij}$.

Choose a subset $I$ of $\{i : n_i = 1\}$ and let $E$ be the set $\{(i,0) : i \in I\}$, to which we refer as *exceptional index pairs*. For any polynomial with coefficients in $\mathbb{Z}$ we denote by $\|P\|$ the sum of the absolute values of the coefficients. We define

$$c_{ij} = \left\| \frac{\partial F}{\partial x_{ij}} \right\|, \qquad c_F = \max_{(i,j) \notin E} c_{ij}.$$

The advantage of having the exceptional set $E$ is that the value of $c_F$ may be smaller than the one we would get by taking the maximum over all pairs $(i,j)$. In the first example in [Za] this enables us to get the optimal lower bound for the product $H(\alpha)H(1-\alpha)$. By $\delta$ we denote the maximum of the numbers $\max_{i\in I}(\widetilde{d}_i + d_{i,1})/2$ and $\max_{i\notin I}\widetilde{d}_i/(n_i+1)$.

THEOREM 1.2. *For each point* $(\mathbf{x}_1,\ldots,\mathbf{x}_r)\in\mathbb{P}^{n_1}(\overline{\mathbb{Q}})\times\ldots\times\mathbb{P}^{n_r}(\overline{\mathbb{Q}})$ *such that* $F(x_{ij})=0$, $x_{ij}\neq 0$ *for all* $i,j$ *and* $F(x_{ij}^{-1})\neq 0$ *we have*

$$H(\mathbf{x}_1)^{n_1+1}\ldots H(\mathbf{x}_r)^{n_r+1}\geq\varrho,$$

*where* $\varrho$ *is the unique real root larger than* 1 *of* $x^{-2}+c_F^{-1}x^{-\delta}=1$.

During the preparation of this paper W. M. Schmidt informed us that in [Schm] he had already proved a theorem very similar to ours in the case where one works in $(\mathbb{P}^1)^r$. The logarithm of the lower bound given in [Schm] is $1/(2^{4f+2r}H)$, where $f$ is the total degree of $F$ and $H$ the maximum of all coefficients. Although the basic starting point in this paper and [Schm] is the same, we nevertheless found that the principle of our approach and the better values of the constants have some interest.

## 2. Applications.

Before proving the theorem we describe a few consequences. First of all consider $r$ algebraic numbers $\alpha_1,\ldots,\alpha_r$ whose sum is a rational integer $N$. We like to interpret the $r$-tuple as a point $(1:\alpha_1)\times\ldots\times(1:\alpha_r)\in(\mathbb{P}^1)^r$. For the set $I$ of our theorem we choose $\{1,\ldots,r\}$. Letting $F$ be the homogeneous version of $x_1+\ldots+x_r-N$ one easily checks that $c_i=1$ for all $i$. Note that the coefficient $N$ in $F$ does not appear in the $c_i$ because of our choice of $I$. So we get $c_F=1$. Moreover, $n_i=1$ and $d_i=1$ for all $i$. Hence $\delta=1$. Thus we find

COROLLARY 2.1. *Let* $\alpha_1,\ldots,\alpha_r\in\overline{\mathbb{Q}}^*$ *and* $N\in\mathbb{Z}$ *be such that* $\alpha_1+\ldots+\alpha_r=N$ *and* $\alpha_1^{-1}+\ldots+\alpha_r^{-1}\neq N$. *Then*

$$H(\alpha_1)\ldots H(\alpha_r)\geq\sqrt{\eta}$$

*where* $\eta$ *is the golden ratio.*

Note that when $r\geq 4$ the lower bound is actually attained for the $r$-tuple $-\zeta_5, 1+\zeta_5, 1, \zeta_{r-2},\ldots,\zeta_{r-2}^{r-3}$ where $\zeta_k$ denotes a primitive $k$th root of unity. When we take for the $\alpha_i$ the conjugates of an algebraic number $\alpha$ of degree $D$ we get the following consequence.

COROLLARY 2.2. *Let* $\alpha\in\overline{\mathbb{Q}}^*$ *be such that* $\operatorname{trace}(\alpha)$ *is integral and* $\operatorname{trace}(\alpha)\neq\operatorname{trace}(\alpha^{-1})$. *Then* $H(\alpha)^D\geq\sqrt{\eta}$.

However, this result is already contained in a result of C. Smyth [Sm] which states that $H(\alpha)^D\geq\theta$ for every non-reciprocal $\alpha\in\overline{\mathbb{Q}}^*$. Here $\theta$ is the real root of $x^3-x-1$.

We now consider $r$ algebraic numbers $\alpha_i$ whose sum is 1 and give a lower bound for $H(1,\alpha_1,\ldots,\alpha_r)$. The polynomial $F$ can be written as $x_1+\ldots+x_r-x_0$ and we have $c_F=1$, $d_1=1$. Furthermore, $\delta=r/(r+1)$.

Corollary 2.3. *For any $\alpha_1,\ldots,\alpha_r\in\overline{\mathbb{Q}}^*$ such that $\alpha_1+\ldots+\alpha_r=1$ and $\alpha_1^{-1}+\ldots+\alpha_r^{-1}\neq 1$ we have*

$$H(1,\alpha_1,\ldots,\alpha_r)\geq \varrho$$

*where $\varrho$ is the real root larger than 1 of $1=x^{-2r-2}+x^{-r}$.*

As pointed out in the introduction, this result is optimal when $r=2$. For $r>2$ this is not true any more. When $r=3$ for example we find the lower bound $1.14613\ldots$ (which improves the bound $\exp(1/402)=1.00249\ldots$ from [SW]). However, the lowest height we could find was $H=1.15096\ldots$ when the $\alpha_i$ are the zeros of $x^3-x^2+1$. On the other hand, the asymptotic behaviour of $\varrho$ as a function of $r$ looks optimal. It is not hard to show that $\varrho^{r+1}\to\eta$ as $r\to\infty$ while on the other hand the zeros $\alpha_0,\ldots,\alpha_r$ of $x^{r+1}-x-1$ satisfy $H(\alpha_0,\ldots,\alpha_r)^{r+1}\to 2$ as $r\to\infty$.

We now consider the Schlickewei–Wirsing result. Suppose we have a line $L:\ \lambda x+\mu y+\nu z=0$ in $\mathbb{P}^2$ with $\lambda\mu\nu\neq 0$. Let $P_1,P_2,P_3\in L(\overline{\mathbb{Q}})$ be three distinct points with non-zero coordinates. Letting $P_i=(x_i:y_i:z_i)$ $(i=1,2,3)$ we get the relation

$$\Delta:=\begin{vmatrix} x_1 & y_1 & z_1\\ x_2 & y_2 & z_2\\ x_3 & y_3 & z_3\end{vmatrix}=0.$$

We want a lower bound of $H(P_1)H(P_2)H(P_3)$. Our polynomial $F$ is now the determinant form $\Delta$. First we point out that

$$\widetilde{\Delta}:=\begin{vmatrix} x_1^{-1} & y_1^{-1} & z_1^{-1}\\ x_2^{-1} & y_2^{-1} & z_2^{-1}\\ x_3^{-1} & y_3^{-1} & z_3^{-1}\end{vmatrix}\neq 0.$$

Suppose $\widetilde{\Delta}=0$. Then there exist $\alpha,\beta,\gamma$, not all zero, such that $\alpha x_i^{-1}+\beta y_i^{-1}+\gamma z_i^{-1}=0$ for $i=1,2,3$. Hence $\alpha y_iz_i+\beta z_ix_i+\gamma x_iy_i=0$ $(i=1,2,3)$. The conic $C:\ \alpha yz+\beta zx+\gamma xy=0$ is reducible if and only if $\alpha\beta\gamma=0$. So, if $\gamma=0$ for example, we get $\alpha x_i+\beta y_i=0$ for $i=1,2,3$. But this contradicts $\nu\neq 0$. So $C$ is an irreducible conic. But then $P_1,P_2,P_3$ lie both on $C$ and $L$, which is impossible since $|C\cap L|\leq 2$. We conclude that $\widetilde{\Delta}\neq 0$. We can now apply our theorem with $r=3$, $n_1=n_2=n_3=2$, $d_1=d_2=d_3=1$, $c_F=2$ and $I=\emptyset$.

Corollary 2.4. *Consider the line $L:\ \lambda x+\mu y+\nu z=0$ in $\mathbb{P}^2$ with $\lambda\mu\nu\neq 0$. Let $P_1,P_2,P_3\in L(\overline{\mathbb{Q}})$ be three distinct points with non-zero coordinates.*

*Then*

$$H(P_1)H(P_2)H(P_3) \geq \varrho,$$

*where* $\varrho$ *is the real root larger than* 1 *of* $1 = \varrho^{-6} + (1/2)\varrho^{-2}$.

The numerical value of $\varrho$ is $1.09427\ldots$, which compares favourably with the value $1.00041\ldots$ from [SW] or $1.019\ldots$ from [Schm]. Moreover, this result was applied successfully to equations of the form $x + y = 1$ with $x, y$ unknowns in a finitely generated multiplicative group and to multiplicity estimates for binary recurrences in [BS].

**3. Proof of Theorem 1.2.** The proof is based on the following observation, which is a direct generalisation of [Za]. Let $X$ be a closed subvariety of $\mathbb{P}^{n_1} \times \ldots \times \mathbb{P}^{n_r}$ defined over $\mathbb{Q}$. We denote the coordinates by $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_r)$ with $\mathbf{x}_i = (x_{i0}, \ldots, x_{in_i})$. Denote by $X(\mathbb{C})_1$ the intersection of $X(\mathbb{C})$ with the polydisc $\{\mathbf{x} : |x_{ij}| \leq 1 \ \forall i, j\}$. We also give ourselves a collection of multihomogeneous polynomials $G_k(\mathbf{x}) \in \mathbb{Z}[\mathbf{x}]$ of multidegrees $(d_{k1}, \ldots, d_{kr})$.

Lemma 3.1. *Let* $\nu_k \geq 0$ *for all* $k$ *and set*

$$w_i = \sum_k \nu_k d_{ki}, \qquad \lambda = -\max_{\mathbf{x} \in X(\mathbb{C})_1} \Big\{ \sum_k \nu_k \log |G_k(\mathbf{x})| \Big\}. \tag{1}$$

*Then for any point* $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_r) \in X(\overline{\mathbb{Q}})$ *with* $\prod_k G_k(\mathbf{x}) \neq 0$ *we have*

$$\prod_{i=1}^{r} H(\mathbf{x}_i)^{w_i} \geq e^{\lambda}. \tag{2}$$

Proof. Suppose that $\mathbf{x} \in X(K)$ with $G_k(\mathbf{x}) \neq 0$ for all $k$. Here $K$ is an algebraic number field of degree $D$, say. For any valuation $v$ of $K$ we let $D_v = [K_v : \mathbb{Q}_v]$. Then the inequality

$$\sum_{i=1}^{r} w_i \log(\max_j |x_{ij}|_v) \geq \sum_k \nu_k \log |G_k(\mathbf{x})|_v + \begin{cases} \frac{D_v}{D}\lambda & \text{if } v \mid \infty, \\ 0 & \text{if } v \nmid \infty, \end{cases}$$

holds for all places $v$ of $K$, because by the homogeneity condition (1) we may assume that $\max_j |x_{ij}|_v = 1$ for all $i$ and the inequality follows from the definition of $\lambda$ if $v$ is infinite and is straightforward if $v$ is finite. The lemma follows by summing over all $v$ and using the product formula. ∎

The following lemma saves us a considerable amount of effort in the determination of $\lambda$ for the sake of the previous lemma.

Lemma 3.2. *With notations as above, the function* $\Psi := \sum \nu_k \log |G_k(\mathbf{x})|$ *assumes a maximum in* $\mathbf{x} \in X(\mathbb{C})_1$ *and it is attained at a point all of whose coordinates have absolute value* 1 *with at most one exception.*

Proof. Since the $\nu_k$ are positive, $\Psi$ is bounded from above in $X(\mathbb{C})_1$. For $\varepsilon > 0$ sufficiently small the set $\mathbf{x} \in X(\mathbb{C})_1$ such that $\Psi(\mathbf{x}) \geq \log(\varepsilon)$ is compact and not empty. Hence it is clear that $\Psi$, being continuous, assumes a maximum.

Now suppose that $\Psi$ assumes a maximum at a point $P$ where at least two coordinates have absolute value $< 1$. Call these coordinates $\xi, \eta$ and denote the values of these coordinates at $P$ by $\xi_0, \eta_0$. Substitute in $F = 0$ the values of all coordinates of $P$ except $\xi, \eta$. The equation $F = 0$ reduces to the equation of a curve $f(\xi, \eta) = 0$ containing the point $\xi_0, \eta_0$. By choosing a branch of $f = 0$ at the point $\xi_0, \eta_0$ we find locally analytic functions $\xi(t), \eta(t)$ such that $\xi(0) = \xi_0, \eta(0) = \eta_0$ and $f(\xi(t), \eta(t)) = 0$ identically in a neighbourhood of $t = 0$. When $f$ was identically zero anyway, we can choose $\xi(t), \eta(t)$ arbitrarily. Choose a disk $D$ in the complex $t$-plane around $0$ such that $|\xi(t)|, |\eta(t)| \le 1$ for all $t \in D$. Specialise the arguments in $\Psi$ to the values of the point $P$ except for $\xi$ and $\eta$ where we substitute $\xi(t)$ and $\eta(t)$. In this way we obtain a function $\psi(t)$ in $t \in D$ which assumes a maximum at $t = 0$. Notice that $\psi(t)$ is harmonic in the real and imaginary part of $t$. A harmonic function assuming a maximum in the interior of its domain is necessarily constant. Hence $\psi(t)$ is constant. But in that case we can continue $\xi(t)$ and $\eta(t)$ analytically until either one of them hits the unit circle. At that new point the value of $\Psi$ is again $\psi(0)$, i.e. maximal. We continue this procedure for other coordinates, if necessary, until we have found an optimal point all of whose coordinates have absolute value one with at most one exception. ■

LEMMA 3.3. *Let $\alpha, \beta, \gamma > 0$. Let $m$ be the unique minimum of the function*

$$u \log \frac{\gamma u}{u+v} + v \log \frac{v}{u+v}$$

*under the constraints $u, v \ge 0$, $\alpha u + \beta v = 1$. Then $e^{-m}$ is the unique real root larger than $1$ of $\gamma^{-1} x^{-\alpha} + x^{-\beta} = 1$.*

Proof. Put $x = v/(u+v)$ and $1 - x = u/(u+v)$. Then

$$u = \frac{1-x}{\beta x + \alpha(1-x)}, \qquad v = \frac{x}{\beta x + \alpha(1-x)}$$

and $x \in [0, 1]$. We must minimize

$$f(x) = \frac{(1-x)\log(\gamma(1-x)) + x \log x}{\beta x + \alpha(1-x)}$$

on $[0, 1]$. Differentiate with respect to $x$,

$$f'(x) = \frac{-\beta \log(\gamma(1-x)) + \alpha \log x}{(\beta x + \alpha(1-x))^2}.$$

This vanishes if $(\gamma(1-x))^\beta = x^\alpha$. Since $x$ is strictly increasing and $1 - x$ strictly decreasing there is a unique solution $x_0 \in ]0, 1[$. Choose $\varrho > 0$ such that $x_0 = \varrho^{-\beta}$. Then $\gamma(1-x) = \varrho^{-\alpha}$ and thus we see that $\varrho$ satisfies

$1-\varrho^{-\beta}=\gamma^{-1}\varrho^{-\alpha}$. It remains to verify that $f(x_0)=-\log\varrho$, which is straightforward. ∎

Proof of Theorem 1.2. We apply Lemma 3.1 to the hypersurface $X$ given by the multihomogeneous polynomial $F(\mathbf{x})\in\mathbb{Z}[\mathbf{x}]$ with multidegrees $d_1,\ldots,d_r$. For the $G_k$ we take the coordinates $x_{ij}$ and the function

$$\widetilde{F}(\mathbf{x})=F(x_{ij}^{-1})\prod x_{ij}^{d_{ij}}$$

where $d_{ij}$ is the degree of $F$ in $x_{ij}$. Let $\mu,\nu_{ij}\ge 0$. Let $\Phi(\mathbf{x})$ be the function $\mu\log|\widetilde{F}(\mathbf{x})|+\sum_{i,j}\nu_{ij}\log|x_{ij}|$ on $X(\mathbb{C})$. Let $\widetilde{d}_i=-d_i+\sum_j d_{ij}$ be the degree of $\widetilde{F}$ in $\mathbf{x}_i$ and suppose

$$w_i=\mu\widetilde{d}_i+\sum_j\nu_{ij},\qquad \lambda=-\max_{\mathbf{x}\in X(\mathbb{C})_1}\Phi(\mathbf{x}).$$

Then Lemma 3.1 states that (2) holds for all $\mathbf{x}\in X(\overline{\mathbb{Q}})$ with $x_{ij}\ne 0$ and $F(x_{ij}^{-1})\ne 0$.

Let us take $w_i=n_i+1$ for all $i$. Although there are many other choices for the weights $w_i$, this choice gives us the particularly simple shape of our main theorem. It remains to choose $\mu,\nu_{ij}$ in such a way that $\lambda$ becomes positive and as large as possible. We choose

$$\nu_{ij}=1-\frac{\widetilde{d}_i}{n_i+1}\mu\qquad\text{if } i\notin I$$

and

$$\nu_{i,0}=1-\frac{\widetilde{d}_i-d_{i,1}}{2}\mu,\quad \nu_{i,1}=1-\frac{\widetilde{d}_i+d_{i,1}}{2}\mu\quad\text{if } i\in I.$$

Let us determine $\max_{\mathbf{x}\in X(\mathbb{C})_1}\Phi(\mathbf{x})$. By Lemma 3.2 this maximum is attained at a point all of whose coordinates, with possibly one exception, lie on the unit circle. Suppose that $|x_{i_0j_0}|\le 1$ and that $|x_{ij}|=1$ for all $(i,j)\ne(i_0,j_0)$. Suppose first that $(i_0,j_0)\notin E$. Then

$$\begin{aligned}|\widetilde{F}(x_{ij})|&=|F(x_{ij}^{-1})|\cdot\Big|\prod_{i,j}(x_{ij})^{d_{ij}}\Big|=|F(x_{i_0j_0}^{-1},\overline{x}_{ij})|\cdot|x_{i_0j_0}|^{d_{i_0j_0}}\\&=|F(x_{i_0j_0}^{-1},\overline{x}_{ij})-F(\overline{x}_{ij})|\cdot|x_{i_0j_0}|^{d_{i_0j_0}}\\&\le c_{i_0j_0}|x_{i_0j_0}^{-1}-\overline{x}_{i_0j_0}|\max(|x_{i_0j_0}|^{-1},|x_{i_0j_0}|)^{d_{i_0j_0}-1}\cdot|x_{i_0j_0}|^{d_{i_0j_0}}\\&=c_{i_0j_0}(1-|x_{i_0j_0}|^2).\end{aligned}$$

Put $|x_{i_0j_0}|^2=\xi$. We see that the maximum of $\Phi$ is

$$\max_{\xi\in[0,1]}[\mu\log(c_{i_0j_0}(1-\xi))+(\nu_{i_0j_0}/2)\log\xi].$$

This maximum is attained at $\xi = \nu_{i_0j_0}/(\nu_{i_0j_0} + 2\mu)$ and its value is

$$\mu \log \frac{2\mu c_{i_0j_0}}{\nu_{i_0j_0} + 2\mu} + \frac{\nu_{i_0j_0}}{2} \log \frac{\nu_{i_0j_0}}{\nu_{i_0j_0} + 2\mu}.$$

Since we have $\nu_{i_0j_0} \geq 1 - \delta\mu$, this maximum is bounded above by

$$\text{(M)} \qquad \mu \log c_F \frac{2\mu}{(1-\delta\mu) + 2\mu} + \frac{1-\delta\mu}{2} \log \frac{1-\delta\mu}{(1-\delta\mu)+2\mu}.$$

We now determine the maximum when $(i_0, j_0) \in E$. In particular, $j_0 = 0$. So suppose we have $|x_{i_0,0}| \leq 1$ and $|x_{ij}| = 1$ for all other $i, j$. Writing down the dependence on $x_{i_0,0}, x_{i_0,1}$ explicitly and putting $z = x_{i_0,0}/x_{i_0,1}$, we find

$$\begin{aligned} |\widetilde{F}(x_{ij})| &= |F(x_{i_0,0}^{-1}, x_{i_0,1}^{-1}, x_{ij}^{-1})| \cdot |x_{i_0,0}|^{d_{i_0,0}} = |F(1, z, x_{ij}^{-1})| \\ &= |F(1, z, \overline{x}_{ij}) - F(1, 1/\overline{z}, \overline{x}_{ij})| \\ &\leq c_{i_0,1} |z - 1/\overline{z}| \max(|z|, |z|^{-1})^{d_{i_0,1}-1} \\ &= c_{i_0,1} |1 - |z|^2| \cdot |z|^{-d_{i_0,1}}. \end{aligned}$$

Put $\xi = |z|^2 = |x_{i_0,0}|^2$. We see that the maximum of $\Phi$ is

$$\max_{\xi \in [0,1]} [\mu \log(c_{i_0,1}(1-\xi)) - (d_{i_0,1}\mu/2) \log |\xi| + (\nu_{i_0,0}/2) \log |\xi|],$$

which equals

$$\mu \log \frac{2c_{i_0j_0}\mu}{\widetilde{\nu} + 2\mu} + \frac{\widetilde{\nu}}{2} \log \frac{\widetilde{\nu}}{\widetilde{\nu} + 2\mu}$$

where $\widetilde{\nu} = \nu_{i_0,0} - d_{i_0,1}\mu/2$. Note that by our choice of $\nu_{i_0,0}$,

$$\widetilde{\nu} = 1 - (\widetilde{d}_{i_0} + d_{i_0,1})\mu/2 \geq 1 - \delta\mu.$$

Hence our maximum is again bounded by (M). Now use Lemma 3.3 with $\alpha = \delta, \beta = 2, \gamma = c_F$ to minimize (M) by letting $\mu$ vary. The assertion of our theorem follows immediately. ∎

## References


[BS] F. Beukers and H. P. Schlickewei, *The equation $x + y = 1$ in finitely generated groups*, Acta Arith. 78 (1996), 189–199.

[Le] D. H. Lehmer, *Factorisation of certain cyclotomic functions*, Ann. of Math. 34 (1933), 461–479.

[Lo] R. Louboutin, *Sur la mesure de Mahler d'un nombre algébrique*, C. R. Acad. Sci. Paris 296 (1983), 707–708.

[Schl] H. P. Schlickewei, *Equations $ax + by = 1$*, Ann. of Math., to appear.

[SW] H. P. Schlickewei and E. Wirsing, *Lower bounds for the heights of solutions of linear equations*, Invent. Math., to appear.

[Schm] W. M. Schmidt, *Heights of algebraic points lying on curves or hypersurfaces*, Proc. Amer. Math. Soc., to appear.

[Sm] C. J. Smyth, *On the product of the conjugates outside the unit circle of an algebraic integer*, Bull. London Math. Soc. 3 (1971), 169–175.

[Za] D. Zagier, *Algebraic numbers close to both* 0 *and* 1, Math. Comp. 61 (1993), 485–491.

[Zh] S. Zhang, *Positive line bundles on arithmetic surfaces*, Ann. of Math. 136 (1992), 569–587.

Department of Mathematics
University of Utrecht
Budapestlaan 6
P.O. Box 80010
3508 TA Utrecht, The Netherlands
E-mail: beukers@math.ruu.nl

Max-Planck-Institut für Mathematik
Gottfried-Claren-Str. 26
53225 Bonn, Germany

*Received on 18.3.1996* (2947)