

On the Thue–Mahler equation II

by

ENRICO BOMBIERI (Princeton, N.J.)

I. Introduction. In this paper we investigate the number of solutions of the so-called Thue–Mahler equation

$$F(x, y) = p_1^{a_1} \cdots p_s^{a_s}$$

where $F(x, y)$ is a binary form of degree $r \geq 3$ without multiple factors and with rational integral coefficients, to be solved in rational integers x, y . More generally, we shall consider the Thue–Mahler equation in number fields, as follows.

Let k be a number field and let S be a finite set of places of k which includes all places at ∞ . Let $[k : \mathbb{Q}]$ be the degree of k and for each absolute value v let $[k_v : \mathbb{Q}_v]$ be the corresponding local degree. We normalize the absolute value $|\cdot|_v$ so that

(i) if $v|p$, then

$$|p|_v = p^{-[k_v : \mathbb{Q}_v]/[k : \mathbb{Q}]},$$

(ii) if $v|\infty$ then

$$|\cdot|_v = \|\cdot\|^{[k_v : \mathbb{Q}_v]/[k : \mathbb{Q}]}$$

with $\|\cdot\|$ the euclidean absolute value.

We also define

$$\varepsilon(v) = \begin{cases} [k_v : \mathbb{Q}_v]/[k : \mathbb{Q}] & \text{if } v|\infty, \\ 0 & \text{otherwise.} \end{cases}$$

This quantity is useful in handling inequalities, because of

$$|a_1 + a_2 + \cdots + a_n|_v \leq n^{\varepsilon(v)} \max_i |a_i|_v.$$

We have

$$\sum_{v|\infty} \varepsilon(v) = 1,$$

therefore for any set of places S we have

$$\prod_{v \in S} |a_1 + a_2 + \cdots + a_n|_v \leq n \prod_{v \in S} \max_i |a_i|_v.$$

We denote by \mathcal{O}_S the ring of S -integers of k and by \mathcal{O}_S^* the group of units of \mathcal{O}_S , thus

$$\begin{aligned}\mathcal{O}_S &= \{x \in k : |x|_v \leq 1 \text{ for } v \notin S\}, \\ \mathcal{O}_S^* &= \{x \in k : |x|_v = 1 \text{ for } v \notin S\}.\end{aligned}$$

We define the S -absolute value $| \cdot |_S$ and the projective absolute S -height $H_S(x, y)$ to be

$$|x|_S = \prod_{v \in S} |x|_v, \quad H_S(x, y) = \prod_{v \in S} \max(|x|_v, |y|_v).$$

By the product formula we see that if x is a non-zero S -integer then $|x|_S \geq 1$ with equality if and only if x is an S -unit, and similarly $H_S(x, y) \geq 1$ if $(x, y) \in \mathcal{O}_S^2$ is not $(0, 0)$. Then the Thue–Mahler equation in the field k relative to the set S can be written as

$$|F(x, y)|_S = 1$$

with $F \in \mathcal{O}_S[x, y]$ homogeneous of degree r , to be solved in S -integers x, y . Solutions to the Thue–Mahler equation fall naturally into equivalence classes $\{(ux, uy)\}$ for $u \in \mathcal{O}_S^*$, and we are interested in obtaining bounds for the number of equivalence classes of solutions. Each equivalence class determines uniquely a projective solution x/y in the projective line $\mathbb{P}^1(k) = k \cup \infty$, and we shall denote by $N(F)$ the set of projective solutions so determined.

From now on, we assume that F has an irreducible factor over k of degree at least 3. Thus it suffices to consider the case in which F itself is irreducible of degree at least 3. For if $F = F'F''$ is a factorization of F over k then by Gauss’s Lemma we also have a factorization over \mathcal{O}_S and now $F(x, y)$ for $x, y \in \mathcal{O}_S$ is an S -unit if and only if $F'(x, y)$ and $F''(x, y)$ are both S -units.

Let T be the torsion subgroup of \mathcal{O}_S^* , that is, the roots of unity in k ; by Dirichlet’s Unit Theorem, the group \mathcal{O}_S^*/T is free abelian of rank $|S| - 1$, therefore $\mathcal{O}_S^*/T(\mathcal{O}_S^*)^r$ is finite abelian of order $r^{|S|-1}$. Thus every S -unit u can be written as $u = \zeta\eta\theta^r$ with $\zeta \in T$ and $\theta \in \mathcal{O}_S^*$ and η a suitable representative in \mathcal{O}_S^* of an element of $\mathcal{O}_S^*/T(\mathcal{O}_S^*)^r$, so that the equation $F(x, y) = u$ becomes

$$\eta^{-1}F(\theta^{-1}x, \theta^{-1}y) = \zeta.$$

We deduce that

$$N(F) = \bigcup N_0(F_i)$$

where F_i runs over the $r^{|S|-1}$ forms $\eta^{-1}F(\theta^{-1}x, \theta^{-1}y)$ and where $N_0(G)$, for a form $G \in \mathcal{O}_S[x, y]$, denotes the set of projective solutions of the equation

$$G(x, y) \in T.$$

In 1984 Evertse [E] obtained uniform bounds for the number of solutions

of the generalized S -unit equation in number fields, thereby obtaining for the first time a bound independent of the height for the number of solutions of the Thue–Mahler equation. From his results it follows that $c^{r^3|S|}$, for some absolute constant $c > 1$, is such a bound.

In 1987 Bombieri and Schmidt [BS] obtained the bound $c_6 r$ for the number of integral solutions of a Thue equation $F(x, y) = 1$ of degree $r \geq 3$, without multiple factors. This has been further improved and generalized to the Thue equation $F(x, y) = m$ by C. L. Stewart [S]. The interested reader will find in [E] and [S] an ample bibliography on the history of the Thue–Mahler equation.

In 1987, Bombieri [B1] claimed the bound $(4|S|)^{2[k:\mathbb{Q}]}(4r)^{26|S|}$ for the number of equivalence classes of solutions of the Thue–Mahler equation in number fields. Unfortunately, the proof of the Gap Principle in Section V of [B1] contains an error which invalidates the counting of the number of small solutions of the Thue–Mahler equation, and therefore the main result of [B1] cannot be considered as proved there.

In this paper we shall give an updated and complete treatment of the Thue–Mahler equation in number fields, showing how to establish a modified form of the Gap Principle in [B1] and recovering the Main Theorem of [B1], in a rather stronger form. However, Lemma 5 and Theorem 1 of [B1] should be considered as not being established.

We shall prove the following result.

MAIN THEOREM. *Let k be an algebraic number field, let S be a finite set of places of k containing all the infinite places and let \mathcal{O}_S and \mathcal{O}_S^* be the ring of S -integers and the group of S -units of k . Let $F(x, y) \in \mathcal{O}_S[x, y]$ be a form of degree r with an irreducible factor over k of degree at least 6. Then the number of equivalence classes of solutions $x, y \in \mathcal{O}_S$ of the Thue–Mahler equation $F(x, y) \in \mathcal{O}_S^*$ does not exceed $(12r)^{12|S|}$.*

We note that the constants 12 and 12 in this bound can be brought down somewhat and they are stated here as round numbers merely for convenience.

Finally, we remark that if F does not have an irreducible factor over k of degree at least 6 but still has at least three inequivalent linear factors over a finite extension k' of k , then we obtain a bound $c^{|S|}$ for the number of solutions of the Thue–Mahler equation $F(x, y) \in \mathcal{O}_S^*$, where c is an absolute constant. In fact, in this case F has a factor $F' \in \mathcal{O}_S[x, y]$ of degree at most 5 with at least three inequivalent linear factors over some extension k' of k , and we may apply Evertse’s results in [E] to the Thue–Mahler equation $F'(x, y) \in \mathcal{O}_S^*$.

Another proof of our Main Theorem, with somewhat better constants, has also been found by Evertse, using a rather different method.

The author is grateful to the referee for pointing out some inaccuracies in an earlier version of this paper.

II. Plan of the paper. In view of the complicated proof, it seems to be worthwhile to summarize its structure and more importantly to explain the motivation behind certain definitions and constructions.

The basic idea behind the proof goes back to Thue. Suppose that

$$F(x, y) = \prod_{i=1}^r (x - \alpha_i y) = 1$$

and that F is irreducible over k . Let k' be an extension of k containing all roots α_i , and let us fix an extension of the absolute value $|\cdot|_v$ to the field k' . This allows us to speak about the quantities $|x - \alpha_i y|_v$ for $i = 1, \dots, r$.

In an equivalent way, we can take $k'' = k(\alpha)$, the extension of degree r of k obtained by adding a root α of F , and note that in view of our normalizations we have

$$|F(x, y)|_v = \prod_{w|v} |x - \alpha y|_w^r.$$

Now we see that the set $|x - \alpha_1 y|_v, \dots, |x - \alpha_r y|_v$ for a fixed extension of v to k' coincides with the set consisting of $|x - \alpha y|_v = |x - \alpha y|_w^{r/[k''_w:k_v]}$, taken with multiplicity $[k''_w:k_v]$, for a fixed α and all extensions $w|v$ of v to k'' .

We expect exactly one factor $|x - \alpha_i y|_v$ to be small, and the others to be of order of magnitude of $\max(|x|_v, |y|_v)$. Since the product of the factors is 1, we get an inequality

$$\min_i |x - \alpha_i y|_v \leq c(F, v) \max(|x|_v, |y|_v)^{-(r-1)}$$

for a suitable coefficient $c(F, v)$ independent of the point (x, y) ; for example, one may take

$$c(F, v) = (2^{\varepsilon(v)}) \max_i \max(1, |\alpha_i|_v)^{r-1} \max_{i,j} |\alpha_i - \alpha_j|_v^{(r-1)(r-2)/2} |D(F)|_v^{-1/2}$$

with $D(F)$ the discriminant of F . If one takes the product of these inequalities for $v \in S$ one gets

$$(2.1) \quad \Lambda(x, y) = \prod_{v \in S} \min_i |x - \alpha_i y|_v \leq c(F) H_S(x, y)^{-(r-1)}$$

where $c(F)$ can be bounded by a power of the height of F .

This inequality is useful if the right-hand side is small, i.e. if $H_S(x, y)$ is larger than a certain power of the height of F , say M ; in this case, we speak of large solutions.

The idea for bounding the number of large solutions goes back to Mahler. First of all, with each solution one associates a vector $\{i(v) : v \in S\}$ deter-

mined by $|x - \alpha_{i(v)}y|_v = \min_i |x - \alpha_i y|_v$; the number of such vectors is at most $r^{|S|}$. Now consider solutions associated with the same vector. We have

$$0 \leq \log |x - \alpha_{i(v)}y|_v / \log A(x, y) \leq 1$$

for each $v \in S$, so that we can view $\{\log |x - \alpha_{i(v)}y|_v / \log A(x, y) : v \in S\}$ as a point in the unit $|S|$ -dimensional cube, lying on the hyperplane where the sum of the coordinates is 1. We partition this cube into $N^{|S|}$ smaller cubes by means of a grid of side $1/N$, and put solutions into the same class if they belong to the same cube. Thus we obtain approximation classes determined by a vector $\{i(v) : v \in S\}$ and by the cube containing the point $\{\log |x - \alpha_{i(v)}y|_v / \log A(x, y) : v \in S\}$; the quantity $1/N$ is the size of the approximation class. Obviously the number of classes does not exceed $N^{|S|}$, but since the points involved are restricted to a special hyperplane one finds the better bound $|S| \binom{N+|S|-1}{|S|-1}$ for the number of approximation classes. This improved bound is important for our estimates.

From (2.1) one can prove that if \mathbf{x}' and \mathbf{x} are two distinct solutions in the same approximation class and $H_S(\mathbf{x}') \geq H_S(\mathbf{x})$ then

$$H_S(\mathbf{x}') > H_S(\mathbf{x})^{(1-\delta)(r-1)}$$

where $\delta < 1/2$ is an absolute constant. This implies that the number of large solutions in a fixed approximation class with $H_S(\mathbf{x}) < M^{((1-\delta)(r-1))^K}$ does not exceed K ; this gives a total of $K|S| \binom{N+|S|-1}{|S|-1} r^{|S|}$. The remaining solutions with $H_S(\mathbf{x}) \geq M^{((1-\delta)(r-1))^K}$ are called very large solutions. Very large solutions give rise to exceptionally good diophantine approximations to the α_i 's, and one can use the fundamental method of Thue to prove that in each approximation class there are at most 8 very large solutions if K is larger than some absolute constant. This gives a bound for the number of large and very large solutions of the Thue–Mahler equation.

This part of the proof follows classical lines and numerical improvements originate from more careful estimates and a sharper form of the Thue Principle used in dealing with very large solutions.

There remains the problem of estimating the number of small solutions. The paper [BS] of Bombieri and Schmidt deals directly with the problem of obtaining a good counting of small solutions. In case $k = \mathbb{Q}$ and $|S| = 1$, Bombieri and Schmidt were able to show that one could replace (2.1) by another inequality which could be used to prove again a gap principle; in a somewhat weaker but simpler way, their argument can be summarized as follows.

The first step consists in proving an inequality

$$(2.2) \quad \min_i |x - \alpha_i y| \leq \frac{2}{\max_{i,j} |\beta_i - \beta_j|} \cdot \frac{1}{|y|}$$

where the β_i 's are obtained from the α_i 's by means of a fractional linear transformation associated with a suitable element $A \in \mathrm{SL}(2, \mathbb{Z})$. If $G(\mathbf{x}) = F(A\mathbf{x})$ then $\prod \max(1, |\beta_i|)$ is the Mahler height of G . Note that the β_i 's are determined only up to a translation. By choosing this translation accurately one verifies that we may also suppose

$$(2.3) \quad \max_{i,j} |\beta_i - \beta_j| \geq 2^{-1} M(G)^{1/r}$$

where $M(G)$ is the Mahler measure of G . If we start with a form F of smallest Mahler height in its $\mathrm{SL}(2, \mathbb{Z})$ equivalence class then $M(G) \geq M(F)$ and thus

$$(2.4) \quad \min_i |x - \alpha_i y| \leq \frac{4M(F)^{-1/r}}{|y|}.$$

Now suppose x/y and x'/y' are two distinct projective solutions with \mathbf{x} and \mathbf{x}' in the same approximation class. Then by (2.4) we have

$$(2.5) \quad \begin{aligned} 1 &\leq |xy' - x'y| = |(x - \alpha_i y)y' - (x' - \alpha_i y')y| \\ &\leq 8M(F)^{-1/r} \max(|y'/y|, |y/y'|), \end{aligned}$$

therefore, ordering solutions by increasing values of $|y|$, we find that the sequence of second coordinates of inequivalent solutions in a given approximation class satisfies the gap principle

$$|y_{i+1}| \geq \frac{1}{8} M(F)^{1/r} |y_i|.$$

After K steps we reach the lower bound $(\frac{1}{8} M(F)^{1/r})^K$ and thus there are at most Kr solutions with $y \neq 0$ and $|y| < (\frac{1}{8} M(F)^{1/r})^K$. If K is large enough (a power of r suffices) we reach $(\frac{1}{8} M(F)^{1/r})^K > M$, i.e. the level of large and very large solutions. This of course requires $M(F)$ to be not too small, for example $M(F) > 16^r$. The actual argument used by Bombieri and Schmidt is more refined, leading to sharp bounds.

The elimination of the condition $M(F) > 16^r$ requires one final idea, namely transforming the equation $F(x, y) = 1$ into finitely many equations $F_i(x, y) = 1$ where F_i is obtained from F by means of a transformation of determinant greater than 1. It is easily seen that such transformations necessarily increase the height of the F_i 's, thus removing the condition of a lower bound for the height of F .

It is tempting to follow the same argument in analyzing the general Thue–Mahler equation, but we encounter new problems in dealing with small solutions. Inequality (2.2) carries over to the general case without difficulty, but (2.5) creates a problem. First of all, it is not true that $|xy' - x'y|_v \geq 1$ for every v , and what we have at our disposal is only $|xy' - x'y|_S \geq 1$. A more serious problem, which was overlooked in [B1], is that we cannot use

(2.2) directly in estimating (2.5), since this would lead to

$$|xy' - x'y|_v \leq 4^{\varepsilon(v)} \max \left(\frac{|y'/y|_v}{\max_{i,j} |\beta_i - \beta_j|_v}, \frac{|y/y'|_v}{\max_{i,j} |\beta'_i - \beta'_j|_v} \right)$$

and the difficulty is that $\max_{i,j} |\beta'_i - \beta'_j|_v$ and $\max_{i,j} |\beta_i - \beta_j|_v$ are not related to each other in any obvious way; the error in [B1] was to forget the prime $'$ in the expression $\max_{i,j} |\beta'_i - \beta'_j|_v$.

A way out of this difficulty can be seen if we assume (for the time being as a working hypothesis) that $|x' - \alpha_{i(v)}y'|_v \leq |x - \alpha_{i(v)}y|_v$ for each $v \in S$. Now we can use (2.2) in estimating (2.5) and find

$$|xy' - x'y|_v \leq \frac{4^{\varepsilon(v)}}{\max_{i,j} |\beta_i - \beta_j|_v} \max \left(1, \frac{|y'|_v}{|y|_v} \right)$$

and hence

$$\max(1, \max_{i,j} |\beta_i - \beta_j|_v |xy' - x'y|_v) \leq 4^{\varepsilon(v)} \max \left(1, \frac{|y'|_v}{|y|_v} \right).$$

The analogue of (2.3) would be a lower bound for

$$\prod_{v \in S} \max(1, \max_{i,j} |\beta_i - \beta_j|_v |xy' - x'y|_v)$$

or equivalently, since the β_i 's are determined only up to translation, a lower bound for

$$\prod_{v \in S} \max(1, \max_i |\beta_i - \bar{\beta}|_v |xy' - x'y|_v)$$

where $\bar{\beta} \in k$ is the average of the β_i 's. Although this expression is reminiscent of the Mahler height, because of the presence of the term $|xy' - x'y|_v$ it does not seem to be immediately comparable with the Mahler height of G . So in [B1] we introduced a new height

$$m_S(F) = \inf \prod_{v \in S} \max(1, \max_i |\alpha_i - \bar{\alpha}|_v |y|_v)$$

where the infimum is taken over $y \in \mathcal{O}_S$, $y \neq 0$, and asked that F be reduced with respect to this height, i.e. that $m_S(F)$ be a minimum (actually the infimum in [B1] is taken over a smaller set). Now the analogue of (2.5) becomes

$$m_S(F) \leq 4r \prod_{v \in S} \max \left(1, \frac{|y'|_v}{|y|_v} \right).$$

The working hypothesis $|x' - \alpha_{i(v)}y'|_v \leq |x - \alpha_{i(v)}y|_v$ can be handled by restricting solutions to the same approximation class with sufficiently fine mesh, provided $\Lambda(x, y)$ and $\Lambda(x', y')$ are larger than $m_S(F)^{-K}$; the Gap Principle so obtained leads essentially to the same estimate as that claimed in [B1], V, p. 227, but since it needs solutions to belong to the same

approximation class the bound stated in [B1], Lemma 5, p. 231 must be modified.

Suppose this done. There is one last difficulty, already dealt with in [B1].

The classical definition of large and very large solutions is given in terms of the Mahler height $M(F)$ of F , and we could not find an immediate comparison between the height $m_S(F)$ and $M(F)$. This means that the diophantine approximation properties of solutions must be stated in such a way as to involve only the new height, and this was done in [B1] by introducing a modified height for solutions, called here the reduced height.

It remains to remove the technical hypothesis of a lower bound for the height $m_S(F)$. The method used in [BS] and [B1] depends on choosing transformations in $\mathrm{GL}(2, \mathcal{O}_S)$ with large determinant. This works well if $[k : \mathbb{Q}]$ is kept fixed or if k has a prime not in S with small norm, but it leads to estimates which are no longer exponential in $|S|$ if primes not in S have large norm. In this paper we deal directly with the case of small height $m_S(F)$, at the cost of additional complications.

III. Equivalent and reduced forms. We say that two forms F and G are *equivalent*, and write $F \sim G$, if there is $A \in \mathrm{GL}(2, \mathcal{O}_S)$ with $\det A \in \mathcal{O}_S^*$ and a root of unity $\zeta \in T$ such that

$$G(\mathbf{x}) = \zeta F(A\mathbf{x})$$

where $\mathbf{x} = (x, y)$. It is clear that if $F \sim G$ then $|N_0(F)| = |N_0(G)|$ and more precisely

$$N_0(F) = AN_0(G)$$

where A acts on $N_0(G)$ by the obvious fractional linear transformation.

Let $F(x, y) = a_0x^r + a_1x^{r-1}y + \dots + a_ry^r$ be a form and suppose that $N_0(F)$ is not empty, so that we have a solution $F(x_0, y_0) = \zeta_0$. Then the transformation

$$A_0 = \begin{pmatrix} x_0 & -\zeta_0^{-1}a_ry_0^{r-1} \\ y_0 & \zeta_0^{-1}(a_0x_0^{r-1} + a_1x_0^{r-2}y_0 + \dots + a_{r-1}y_0^{r-1}) \end{pmatrix}$$

has $\det A_0 = 1$, $A_0(1, 0) = (x_0, y_0)$ and the form $G(\mathbf{x}) = \zeta_0^{-1}F(A_0\mathbf{x})$ is equivalent to F with $G(1, 0) = 1$, i.e. G has leading coefficient 1. We then say that the form G is *normalized* and denote by \mathcal{F} the set of normalized forms equivalent to F . Normalized forms split into linear factors as

$$F(x, y) = \prod_{i=1}^r (x - \alpha_i y)$$

with the α_i 's in a finite extension k' of k . We define

$$\bar{\alpha} = \frac{1}{r} \sum_{i=1}^r \alpha_i, \quad m_v(F, y) = \max(1, \max_i |\alpha_i - \bar{\alpha}|_v |y|_v),$$

$$m_S(F, y) = \prod_{v \in S} m_v(F, y) \quad \text{and} \quad m_S(F) = \inf m_S(F, y)$$

where the infimum is taken over $y \in \mathcal{O}_S^*$. The quantities $m_v(F, y)$ are independent of the extension of v to the field k' .

We consider all pairs (F, y) consisting of a normalized form $F \in \mathcal{F}$ and of an element $y \in \mathcal{O}_S$, $y \neq 0$.

DEFINITION. A pair (F, y_0) with $F \in \mathcal{F}$ and $y_0 \in \mathcal{O}_S$, $y_0 \neq 0$ is *reduced* if

$$m_S(F, y_0) = \inf m_S(G, y)$$

where the infimum is taken over all $G \in \mathcal{F}$ and all $y \in \mathcal{O}_S$, $y \neq 0$.

We verify that $\inf m_S(F, y)$ is attained as follows. Suppose that $m_S(F, y) \leq M$. We have

$$|y|_S \prod_{v \in S} \max_i |\alpha_i - \bar{\alpha}|_v = \prod_{v \in S} \max_i |\alpha_i - \bar{\alpha}|_v |y|_v \leq m_S(F, y) \leq M.$$

For every v we have $|\alpha_i - \alpha_j|_v \leq 2^{\varepsilon(v)} \max_i |\alpha_i - \bar{\alpha}|_v$, thus

$$\left(\prod_{v \in S} \max_i |\alpha_i - \bar{\alpha}|_v \right)^{r(r-1)} \geq \prod_{v \in S} \prod_{i \neq j} (2^{-\varepsilon(v)} |\alpha_i - \alpha_j|_v) = 2^{-r(r-1)} |D(F)|_S$$

where $D(F) \in k$ is the discriminant of F . Since $|D(F)|_S \geq 1$ it follows that

$$(3.1) \quad |y|_S \leq 2M$$

and therefore $|y|_S$ is bounded from above. Now we note that if $u \in \mathcal{O}_S^*$ then the transformation $A = \begin{pmatrix} 1 & 0 \\ 0 & u \end{pmatrix}$ yields $m_v(F \circ A, y/u) = m_v(F, y)$ for every v , while $F \circ A \sim F$ is still normalized. It follows that y may be replaced by any representative (mod \mathcal{O}_S^*), and therefore since $|y|_S$ is bounded we can take y from a finite set. Moreover, this shows that we may assume that $\max_i |\alpha_i - \bar{\alpha}|_v$ is bounded too. But if $a \in \mathcal{O}_S$ then the transformation $A = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$ yields $m_v(F \circ A, y) = m_v(F, y)$ for every v , while $F \circ A \sim F$ is still normalized. This means that we may translate α_i by any element in \mathcal{O}_S , and in particular by an element a such that $\bar{\alpha} - a$ is in a fundamental domain for $(\prod_{v \in S} k_v) / \mathcal{O}_S$, and shows that we may assume that each $|\alpha_i|_v$ is bounded, and hence the α_i 's can also be taken from a finite set. This proves that a reduced pair exists.

By (3.1), it is immediate that

LEMMA 1. *If (F, y_0) is a reduced pair then $|y_0|_S \leq 2m_S(F)$.*

If (F, y_0) is reduced and $A = \begin{pmatrix} 1 & 0 \\ 0 & u \end{pmatrix}$ with $u \in \mathcal{O}_S^*$ then $(F \circ A, y_0/u)$ is also reduced, and similarly if $A = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$ and $a \in \mathcal{O}_S$ then $(F \circ A, y_0)$ is reduced too. This shows that the subgroup $A = \begin{pmatrix} 1 & * \\ 0 & * \end{pmatrix}$ of $\text{GL}(2, \mathcal{O}_S)$ transforms reduced pairs into reduced pairs. The next definition introduces a height on solutions which is invariant by the action of this subgroup.

DEFINITION. Let (F, y_0) be reduced. The *reduced F -height* of a solution $x/y \in N_0(F)$ is

$$H_S^*(x, y) = H_S(x - \bar{\alpha}y, y/y_0) = \prod_{v \in S} \max(|x - \bar{\alpha}y|_v, |y/y_0|_v).$$

Now we consider the effect of equivalence transformations on forms in \mathcal{F} . We follow the argument in [BS].

LEMMA 2. *There is $G(u, w) = \prod(u - \beta_i w)$ with $G \sim F$ such that for each $v \in S$*

$$\min_i |x - \alpha_i y|_v \leq \frac{(2r)^{\varepsilon(v)} |r|_v^{-1}}{\max_i |\beta_i - \bar{\beta}|_v |y|_v}.$$

Here $r\bar{\beta} = \sum_i \beta_i$.

PROOF. Let $F(x, y) = x^r + a_1 x^{r-1} y + \dots + a_r y^r = \prod(x - \alpha_i y)$ be a normalized form and let $\mathbf{x} = (x, y)$ be a solution of the Thue–Mahler equation $F(\mathbf{x}) = \zeta$. If A is the matrix

$$A = \begin{pmatrix} x & -\zeta^{-1} a_r y^{r-1} \\ y & \zeta^{-1} (x^{r-1} + a_1 x^{r-2} y + \dots + a_{r-1} y^{r-1}) \end{pmatrix}$$

we have $\det A = 1$, thus if $\mathbf{z} = (-\zeta^{-1} a_r y^{r-1}, \zeta^{-1} (x^{r-1} + a_1 x^{r-2} y + \dots + a_{r-1} y^{r-1}))$ we have $\det(\mathbf{x}, \mathbf{z}) = 1$. Thus for \mathbf{x}' with S -integer coordinates we can write $\mathbf{x}' = a\mathbf{x} + b\mathbf{z}$ with S -integers a, b , and actually $b = \det(\mathbf{x}, \mathbf{x}') = xy' - x'y$. We abbreviate $L_i(\mathbf{x}) = x - \alpha_i y$ and deduce by linearity that

$$(3.2) \quad L_i(\mathbf{x}')/L_i(\mathbf{x}) = a - \beta_i b$$

where

$$\beta_i = \beta_i(A) = -L_i(\mathbf{z})/L_i(\mathbf{x}).$$

Let $G(u, w) = \zeta^{-1} F(u\mathbf{x} + w\mathbf{z})$. Then $G \sim F$ and we have, noting that $F(\mathbf{x}) = \zeta$,

$$G(u, w) = \zeta^{-1} F(A(u, w)) = \prod_{i=1}^r L_i(u\mathbf{x} + w\mathbf{z})/L_i(\mathbf{x}) = \prod_{i=1}^r (u - \beta_i w).$$

Now we choose $\mathbf{x}' = (1, 0)$ and obtain, by (3.2),

$$\frac{1}{L_i(\mathbf{x})} - \frac{1}{L_j(\mathbf{x})} = (\beta_i - \beta_j)y,$$

so that

$$(3.3) \quad \max_{i,j} |\beta_i - \beta_j|_v |y|_v \leq 2^{\varepsilon(v)} \max_i \frac{1}{|L_i(\mathbf{x})|_v}.$$

Finally,

$$|r(\beta_i - \bar{\beta})|_v = \left| \sum_{j=1}^r (\beta_i - \beta_j) \right|_v \leq r^{\varepsilon(v)} \max_{i,j} |\beta_i - \beta_j|_v$$

and the conclusion of Lemma 2 is an easy consequence of the two last displayed inequalities.

Lemma 2 is the key ingredient in showing that small solutions of the Thue–Mahler equation are sparse.

IV. Diophantine approximation properties of solutions. Let $F \in k[x, y]$ be a form

$$F(x, y) = a_0 x^r + a_1 x^{r-1} y + \dots + a_r y^r = a_0 \prod (x - \alpha_i y)$$

with coefficients in k . Intuitively, if $|F(x, y)|_v$ is small then we expect x/y to approximate one of the roots α_i associated with F . The following result, which holds for general forms F , is a quantitative form of this principle.

LEMMA 3. *For each absolute value $| \cdot |_v$ and $x, y \in k$ we have*

$$\begin{aligned} \min_i |x - \alpha_i y|_v \\ \leq 2^{(r-1)\varepsilon(v)} |a_0|_v^{r-2} |D(F)|_v^{-1/2} \max_{i,j} |\alpha_i - \alpha_j|_v^{(r-1)(r-2)/2} \frac{|F(\mathbf{x})|_v}{|y|_v^{r-1}}. \end{aligned}$$

Proof. We abbreviate $\xi = x/y$ and note that

$$a_0 \prod_{i=1}^r (\xi - \alpha_i) = F(\mathbf{x})/y^r,$$

thus

$$(4.1) \quad |\xi - \alpha_{i(v)}|_v = |a_0|_v^{-1} \left(\prod' |\xi - \alpha_i|_v \right)^{-1} |F(\mathbf{x})|_v |y|_v^{-r}$$

where $i(v)$ is such that $|\xi - \alpha_{i(v)}|_v = \min_i |\xi - \alpha_i|_v$ and where the prime $'$ means that the value $i = i(v)$ is omitted from the product. We have

$$|\alpha_{i(v)} - \alpha_i|_v \leq 2^{\varepsilon(v)} \max(|\xi - \alpha_{i(v)}|_v, |\xi - \alpha_i|_v) = 2^{\varepsilon(v)} |\xi - \alpha_i|_v,$$

therefore

$$(4.2) \quad \prod' |\xi - \alpha_i|_v \geq 2^{-(r-1)\varepsilon(v)} \prod' |\alpha_{i(v)} - \alpha_i|_v.$$

Also

$$|D(F)|_v^{1/2} = |a_0|_v^{r-1} \prod_{i < j} |\alpha_i - \alpha_j|_v,$$

which clearly implies

$$(4.3) \quad \prod' |\alpha_{i(v)} - \alpha_i|_v \geq |D(F)|_v^{1/2} |a_0|_v^{-r+1} \left(\prod'' |\alpha_i - \alpha_j|_v \right)^{-1}$$

where \prod'' runs over $i < j$, $i, j \neq i(v)$. The conclusion of Lemma 3 follows easily from (4.1), (4.2) and (4.3).

LEMMA 4. *For each absolute value $| \cdot |_v$ and $x, y \in k$ we have*

$$\begin{aligned} & \min_i |x - \alpha_i y|_v \\ & \leq 2^{(r-1)\varepsilon(v)} |a_0|_v^{r-2} |D(F)|_v^{-1/2} \max_i |\alpha_i|_v^{r-1} \max_{i,j} |\alpha_i - \alpha_j|_v^{(r-1)(r-2)/2} \frac{|F(\mathbf{x})|_v}{|x|_v^{r-1}}. \end{aligned}$$

Proof. We have

$$F(x, y) = a_0 \prod_{i=1}^r (x - \alpha_i y) = (-1)^r a_r \prod_{i=1}^r \left(y - \frac{1}{\alpha_i} x \right).$$

Let $G(X, Y)$ be the form

$$G(X, Y) = (-1)^r a_r \prod_{i=1}^r \left(X - \frac{1}{\alpha_i} Y \right)$$

where $(X, Y) = (y, x)$. Then Lemma 4 follows from the analogue of (4.1), (4.2) and (4.3) for the form $G(X, Y)$, noting that $(-1)^r a_r = a_0 \alpha_1 \dots \alpha_r$.

DEFINITION.

$$\Lambda(\mathbf{x}) = \prod_{v \in S} \min_i |x - \alpha_i y|_v.$$

Now we are ready to prove the main result of this section.

LEMMA 5. *Suppose the form $F(x, y) \in \mathcal{O}_S[x, y]$ has no multiple factors and that (F, y_0) is reduced. Let*

$$H_S^*(\mathbf{x}) = \prod_{v \in S} \max(|x - \bar{\alpha} y|_v, |y/y_0|_v)$$

be the corresponding reduced F -height. Then

$$(2m_S(F))^{-(r-1)} \frac{|F(\mathbf{x})|_S}{H_S^*(\mathbf{x})^{r-1}} \leq \Lambda(\mathbf{x}) \leq \left(\frac{2m_S(F)}{|y_0|_S} \right)^{r(r-1)/2} \frac{|F(\mathbf{x})|_S}{H_S^*(\mathbf{x})^{r-1}}.$$

Proof. We begin by proving the upper bound. We apply Lemmas 3 and 4 to the form

$$G(X, Y) = \prod_{i=1}^r (X - \alpha_i^* Y)$$

where $\alpha_i^* = (\alpha_i - \bar{\alpha})y_0$, at the point

$$(X, Y) = (x - \bar{\alpha}y, y/y_0)$$

for which $G(X, Y) = F(x, y)$. We have

$$\begin{aligned} a_0 &= 1, \quad \max(1, \max_i |\alpha_i^*|_v) = m_v(F, y_0), \\ \max |\alpha_i^* - \alpha_j^*|_v &\leq 2^{\varepsilon(v)} m_v(F, y_0), \end{aligned}$$

therefore we get

$$\begin{aligned} & \min_i |x - \alpha_i y|_v \\ & \leq (2^{\varepsilon(v)} m_v(F, y_0))^{r(r-1)/2} \frac{|F(\mathbf{x})|_v}{|D(G)|_v^{1/2}} (\max(|x - \bar{\alpha}y|_v, |y/y_0|_v))^{-(r-1)}. \end{aligned}$$

The required inequality follows by taking the product of the last displayed inequality for $v \in S$, noting that $D(G) = y_0^{r(r-1)} D(F)$ and $|D(F)|_S \geq 1$.

The proof of the lower bound is even simpler. We have

$$|x - \alpha_i y|_v = |(x - \bar{\alpha}y) - \alpha_i^*(y/y_0)|_v \leq 2^{\varepsilon(v)} m_v(F, y_0) \max(|x - \bar{\alpha}y|_v, |y/y_0|_v),$$

therefore from $\prod (x - \alpha_i y) = F(\mathbf{x})$ we get

$$\min_i |x - \alpha_i y|_v \geq |F(\mathbf{x})|_v (2^{\varepsilon(v)} m_v(F, y_0) \max(|x - \bar{\alpha}y|_v, |y/y_0|_v))^{-(r-1)}$$

and the result follows by taking the product of the last displayed inequality for $v \in S$.

V. Classification of solutions. For a form $F \in \mathcal{O}_S[x, y]$ of degree r we consider the set $N_0(F) = \{x/y : F(x, y) \in T\}$ of projective solutions of the Thue–Mahler equation. We assume that F is normalized, so that

$$F(x, y) = \prod_{i=1}^r (x - \alpha_i y).$$

We begin by subdividing solutions into not more than $r^{|S|}$ classes, according to the vector $\{i(v) : v \in S\}$ of indices determined by the condition

$$|x - \alpha_{i(v)} y|_v = \min_i |x - \alpha_i y|_v.$$

A further subdivision into classes follows an idea of Mahler and is quite standard in this type of problem. Lemma 6 below replaces the more complicated Mahler’s Lemma of [B1] and Lemma 4 of [E].

We recall that by definition

$$\Lambda(\mathbf{x}) = \prod_{v \in S} \min_i |x - \alpha_i y|_v$$

and note that since $\min_i |x - \alpha_i y|_v \leq 1$ for $v \in S$ we have

$$0 \leq \log |x - \alpha_{i(v)} y|_v / \log \Lambda(\mathbf{x}) \leq 1$$

for each $v \in S$. Thus the vector $\{\log |x - \alpha_{i(v)} y|_v / \log \Lambda(\mathbf{x}) : v \in S\}$ is a point in the unit $|S|$ -dimensional cube and lies on the hyperplane where the sum of the coordinates is 1. We partition this cube by means of a grid of side $1/N$ with N a positive integer, and we further classify solutions into approximation classes according to the cube determined by them. The quantity $1/N$ is the size of the class.

LEMMA 6. *The number of approximation classes of size $1/N$ determined by the vectors*

$$\{\log |x - \alpha_{i(v)}y|_v / \log \Lambda(\mathbf{x}) : v \in S\}$$

does not exceed

$$|S|r^{|S|} \binom{N + |S| - 1}{|S| - 1} < r|S| \left(r \left(\frac{N}{|S|} + 1 \right) \right)^{|S|-1} \left(1 + \frac{|S|}{N} \right)^N.$$

PROOF. Let us fix the vector $\{i(v) : v \in S\}$ and define $\gamma_v = \log |x - \alpha_{i(v)}y|_v / \log \Lambda(\mathbf{x})$. In order to bound the number of cubes we associate with each vector $\{\gamma_v : v \in S\}$ the south-west corner

$$\left\{ \frac{1}{N} [N\gamma_v] : v \in S \right\}$$

of the cube containing the point $\{\gamma_v : v \in S\}$. Then since $\sum_{v \in S} \gamma_v = 1$ we see that

$$\sum_{v \in S} [N\gamma_v] \leq N.$$

We fix one place $v_0 \in S$ and suppose we are given the $|S| - 1$ integers $n_v = [N\gamma_v]$ for $v \neq v_0$, $v \in S$. Then $n_v \leq N\gamma_v < n_v + 1$, therefore using $\sum_{v \in S} \gamma_v = 1$ we find

$$N - \sum_{v \neq v_0} n_v - (|S| - 1) < N\gamma_{v_0} \leq N - \sum_{v \neq v_0} n_v$$

and deduce that $[N\gamma_{v_0}]$ has not more than $|S|$ possibilities. The number of solutions of $\sum_{v \neq v_0} n_v \leq N$ is precisely $\binom{N + |S| - 1}{|S| - 1}$, the number of choices for $\{i(v) : v \in S\}$ is $r^{|S|}$, and

$$\binom{N + |S| - 1}{|S| - 1} < \left(\frac{N}{|S|} + 1 \right)^{|S|-1} \left(1 + \frac{|S|}{N} \right)^N;$$

Lemma 6 follows.

Classification into approximation classes is done with respect to the linear forms $x - \alpha_i y$. We also need to compare the sizes of $|y|_v$, which leads to a further subdivision into classes which we call *magnitude classes*.

We want to compare y with a fixed non-zero S -integer y_0 . We then consider the subset $S' = S'(y) = \{v \in S : |y/y_0|_v \geq 1\}$. Now we have again

$$0 \leq \log |y/y_0|_v / \log |y/y_0|_{S'} \leq 1$$

and we may partition the $|S'|$ -dimensional unit cube into cubes of size $1/M$, analogously to what was done before. In the same way, let S'' be the subset $S'' = S''(y) = \{v \in S : |y/y_0|_v < 1\}$ and partition the $|S''|$ -dimensional unit cube into cubes of size $1/M$. A magnitude class thus consists of a partition

$S = S' \cup S''$ and of two cubes, with solutions falling into magnitude classes according to which cubes contain the points

$$\{\log |y/y_0|_v / \log |y/y_0|_{S'} : v \in S'\}, \quad \{\log |y/y_0|_v / \log |y/y_0|_{S''} : v \in S''\}.$$

LEMMA 7. *The number of magnitude classes of size $1/M$ within a given approximation class is at most*

$$\begin{aligned} 2|S| \binom{M+|S|-1}{|S|-1} + \sum_{s=1}^{|S|-1} s \binom{|S|-s}{s} \binom{M+s-1}{s-1} \binom{M+|S|-s-1}{|S|-s-1} \\ \leq |S|^2 \left(\frac{4M}{|S|} + 2 \right)^{|S|-2} \left(1 + \frac{|S|}{2M} \right)^{2M} \end{aligned}$$

if $|S| \geq 2$, and 2 if $|S| = 1$.

Proof. The proof is the same as in Lemma 6, noting that the number of subsets of S of cardinality s is $\binom{|S|}{s}$, and that

$$\sum s \binom{|S|-s}{s} = |S|(|S|-1)2^{|S|-2}.$$

VI. The Gap Principle. The Gap Principle in this section replaces the erroneous Gap Principle in [B1], and provides the correct extension of (2.5) to the general case.

LEMMA 8. *Let (F, y_0) be a reduced pair and let \mathbf{x} be a solution of the Thue–Mahler equation $F(x, y) \in T$, belonging to an approximation class $\mathcal{C}(\mathbf{x})$ of size $1/N$. Let $\mathbf{x}' \in \mathcal{C}(\mathbf{x})$ be a solution projectively distinct from \mathbf{x} and satisfying $\Lambda(\mathbf{x}') \leq \Lambda(\mathbf{x})$. Then we have*

$$\prod_{v \in S} \max \left(1, \frac{|y'|_v}{|y|_v} \right) \geq \frac{1}{4r} \Lambda(\mathbf{x})^{|S|/N} m_S(F).$$

Moreover, if $U \geq 1$ then the number of solutions $\mathbf{x}' \in \mathcal{C}(\mathbf{x})$, projectively distinct from \mathbf{x} , such that $\Lambda(\mathbf{x}') \leq \Lambda(\mathbf{x})$ and

$$\prod_{v \in S} \max \left(1, \frac{|y'|_v}{|y|_v} \right) \leq U$$

does not exceed

$$\frac{25}{16} (8\Lambda(\mathbf{x})^{-|S|/N} U^{6/5})^{[k:\mathbb{Q}]} r^{|S|} |T| \binom{6|S|}{|S|}.$$

Proof. Let $\{i(v) : v \in S\}$ and $\{\gamma_v : v \in S\}$ be the approximation class and the south-west corner of the cube associated with \mathbf{x} and \mathbf{x}' . Then

$$\Lambda(\mathbf{x})^{\gamma_v+1/N} \leq |x - \alpha_{i(v)} y|_v \leq \Lambda(\mathbf{x})^{\gamma_v},$$

hence

$$|x' - \alpha_{i(v)}y'|_v \leq \Lambda(\mathbf{x}')^{\gamma_v} \leq \Lambda(\mathbf{x})^{\gamma_v} \leq \Lambda(\mathbf{x})^{-1/N} |x - \alpha_{i(v)}y|_v.$$

Let β_i and $\bar{\beta}$ be as in Lemma 2. This inequality and Lemma 2 give

$$(6.1) \quad \begin{aligned} |\det(\mathbf{x}, \mathbf{x}')|_v &= |(x - \alpha_{i(v)}y)y' - (x' - \alpha_{i(v)}y')y|_v \\ &\leq 2^{\varepsilon(v)} \Lambda(\mathbf{x})^{-1/N} |x - \alpha_{i(v)}y|_v \max(|y|_v, |y'|_v) \\ &\leq \frac{(4r)^{\varepsilon(v)} |r|_v^{-1} \Lambda(\mathbf{x})^{-1/N}}{\max_i |\beta_i - \bar{\beta}|_v} \max\left(1, \frac{|y'|_v}{|y|_v}\right) \end{aligned}$$

and we deduce

$$(6.2) \quad \begin{aligned} \max(1, \max_i |\beta_i - \bar{\beta}|_v |\det(\mathbf{x}, \mathbf{x}')|_v) \\ \leq (4r)^{\varepsilon(v)} |r|_v^{-1} \Lambda(\mathbf{x})^{-1/N} \max\left(1, \frac{|y'|_v}{|y|_v}\right). \end{aligned}$$

If $G(x, y) = \prod(x - \beta_i y)$ then $G \sim F$ and

$$\prod_{v \in S} \max(1, \max_i |\beta_i - \bar{\beta}|_v |\det(\mathbf{x}, \mathbf{x}')|_v) = m_S(G, \det(\mathbf{x}, \mathbf{x}')) \geq m_S(F)$$

because F is reduced. Now the first clause of Lemma 8 follows by taking the product of (6.2) for all $v \in S$.

Let $\mathcal{X}(\mathbf{x})$ be the set of solutions $\mathbf{x}' \in \mathcal{C}(\mathbf{x})$, projectively distinct from \mathbf{x} , satisfying $\Lambda(\mathbf{x}') \leq \Lambda(\mathbf{x})$ and

$$\prod_{v \in S} \max\left(1, \frac{|y'|_v}{|y|_v}\right) \leq U.$$

We have

$$0 \leq \log \max\left(1, \frac{|y'|_v}{|y|_v}\right) / \log U \leq 1$$

for every $v \in S$. We partition the $|S|$ -dimensional unit cube into cubes of size $1/(5|S|)$ and associate with each solution \mathbf{x}' the north-east corner $\{\theta_v : v \in S\}$ of the cube containing the vector

$$\left\{ \log \max\left(1, \frac{|y'|_v}{|y|_v}\right) / \log U : v \in S \right\}.$$

Then

$$(6.3) \quad \max\left(1, \frac{|y'|_v}{|y|_v}\right) \leq U^{\theta_v},$$

$$(6.4) \quad \sum_{v \in S} \theta_v \leq 6/5$$

and, by an argument similar to that used in Lemma 6, the number of cubes involved does not exceed $|S| \binom{6|S|}{|S|}$.

Consider now a subset $\mathcal{Y}(\mathbf{x}) \subseteq \mathcal{X}(\mathbf{x})$ of solutions \mathbf{x}' associated with the same cube $\{\theta_v : v \in S\}$. We proceed as in the proof of (6.1), using (3.3) in place of Lemma 2 (this allows us to save a factor $4r$), and estimating $\max(1, |y'|_v/|y|_v)$ by means of (6.3). We obtain

$$(6.5) \quad |\det(\mathbf{x}, \mathbf{x}')|_v \leq 4^{\varepsilon(v)} B_v^{-1} \Lambda(\mathbf{x})^{-1/N} U^{\theta_v}$$

where we have abbreviated $B_v = \max |\beta_i - \beta_j|_v$.

We subdivide the set $\mathcal{Y}(\mathbf{x})$ into equivalence classes by putting two elements \mathbf{x}' and \mathbf{x}'' into the same equivalence class if

$$\det(\mathbf{x}, \mathbf{x}' - \mathbf{x}'') = 0,$$

and denote by $\mathcal{Z}(\mathbf{x})$ a set of representatives in $\mathcal{Y}(\mathbf{x})$ for these equivalence classes; it follows that

$$(6.6) \quad \det(\mathbf{x}, \mathbf{x}' - \mathbf{x}'') \neq 0$$

whenever \mathbf{x}' and \mathbf{x}'' are two distinct elements of $\mathcal{Z}(\mathbf{x})$.

Let us fix an archimedean absolute value v_0 and define

$$\|\cdot\| = |\cdot|_{v_0}^{[k:\mathbb{Q}]/[k_{v_0}:\mathbb{R}]}$$

so that $\|\cdot\|$ is the ordinary euclidean absolute value. Let us write for simplicity

$$C_v = 4^{\varepsilon(v)} B_v^{-1} \Lambda(\mathbf{x})^{-1/N} U^{\theta_v}$$

and consider the interval

$$\|\xi\| \leq C_{v_0}^{[k:\mathbb{Q}]}$$

if v_0 is real, or the square

$$\|\operatorname{Re}(\xi)\| \leq C_{v_0}^{[k:\mathbb{Q}]/2}, \quad \|\operatorname{Im}(\xi)\| \leq C_{v_0}^{[k:\mathbb{Q}]/2}$$

if v_0 is complex. We subdivide each interval or square, as the case may be, into P or P^2 subintervals or subsquares of side $1/P$ the original side, where P is a positive integer. By (6.5) and Dirichlet's Box Principle, if $|\mathcal{Z}(\mathbf{x})| > P$ when v_0 is real or $|\mathcal{Z}(\mathbf{x})| > P^2$ when v_0 is complex, one of these subintervals or subsquares must contain two points $\det(\mathbf{x}, \mathbf{x}')$ and $\det(\mathbf{x}, \mathbf{x}'')$ with distinct elements \mathbf{x}' and \mathbf{x}'' of $\mathcal{Z}(\mathbf{x})$. It follows that

$$\|\det(\mathbf{x}, \mathbf{x}' - \mathbf{x}'')\| \leq \frac{2}{P} C_{v_0}^{[k:\mathbb{Q}]/[k_{v_0}:\mathbb{R}]},$$

hence

$$|\det(\mathbf{x}, \mathbf{x}' - \mathbf{x}'')|_{v_0} \leq \left(\frac{2}{P}\right)^{\varepsilon(v_0)} C_{v_0}.$$

It is also clear from (6.5) that

$$|\det(\mathbf{x}, \mathbf{x}' - \mathbf{x}'')|_v \leq 2^{\varepsilon(v)} C_v$$

for $v \neq v_0$. We take the product of these inequalities for all $v \in S$ and find

$$|\det(\mathbf{x}, \mathbf{x}' - \mathbf{x}'')|_S \leq 2P^{-1} \prod_{v \in S} C_v.$$

We choose P the largest integer such that $P < |\mathcal{Z}(\mathbf{x})|$ if v_0 is real and $P^2 < |\mathcal{Z}(\mathbf{x})|$ if v_0 is complex, and assume correspondingly that $|\mathcal{Z}(\mathbf{x})| > 2$ or $|\mathcal{Z}(\mathbf{x})| > 25$. Then $\frac{16}{25}|\mathcal{Z}(\mathbf{x})| \leq P$ if v_0 is real, and $\leq P^2$ if v_0 is complex.

In view of (6.4), the last displayed inequality implies

$$(6.7) \quad |\det(\mathbf{x}, \mathbf{x}' - \mathbf{x}'')|_S \left(\frac{16}{25} |\mathcal{Z}(\mathbf{x})| \right)^{1/[k:\mathbb{Q}]} \leq 8 \left(\prod_{v \in S} B_v \right)^{-1} \Lambda(\mathbf{x})^{-|S|/N} U^{6/5}.$$

We note that

$$\left(\prod_{v \in S} B_v \right)^{r(r-1)} \geq \prod_{v \in S} \prod_{i \neq j} |\beta_i - \beta_j|_v = |D(G)|_S \geq 1,$$

and $|\det(\mathbf{x}, \mathbf{x}' - \mathbf{x}'')|_S \geq 1$ by (6.6). Thus (6.7) yields

$$|\mathcal{Z}(\mathbf{x})| \leq \frac{25}{16} (8\Lambda(\mathbf{x})^{-|S|/N} U^{6/5})^{[k:\mathbb{Q}]}.$$

This inequality, obtained with the assumption $|\mathcal{Z}(\mathbf{x})| > 2$ if v_0 is real and $|\mathcal{Z}(\mathbf{x})| > 25$ if v_0 is complex, clearly continues to hold in the remaining cases.

It remains to give an upper bound for the number of elements in a given equivalence class. But if $\det(\mathbf{x}, \mathbf{x}' - \mathbf{x}'') = 0$ then $\mathbf{x}'' = \mathbf{x}' + \lambda\mathbf{x}$ with $\lambda \in k$, so that the number of elements in a given equivalence class does not exceed the number of solutions λ of

$$F(\mathbf{x}' + \lambda\mathbf{x}) \in T.$$

This is an equation of degree at most r in λ , and therefore the cardinality of an equivalence class is at most $r|T|$.

The proof of Lemma 8 is completed by noting that the total number of solutions does not exceed the number of cubes times the maximum number of elements in an equivalence class times the maximum of $|\mathcal{Z}(\mathbf{x})|$.

VII. The Strong Gap Principle. This section follows rather closely the corresponding section in [B1], with some simplifications.

LEMMA 9. *Let (F, y_0) be a reduced pair and let H_S^* be the corresponding reduced F -height. Then for any $S_1 \subseteq S$ and any solution \mathbf{x} of the Thue–Mahler equation $F(x, y) \in T$ we have*

$$(2m_S(F))^{-2} H_S^*(\mathbf{x})^{-1} \leq |y/y_0|_{S_1} \leq 2m_S(F) H_S^*(\mathbf{x}).$$

PROOF. We abbreviate $\alpha_i^* = (\alpha_i - \bar{\alpha})y_0$ and $(X, Y) = (x - \bar{\alpha}y, y/y_0)$. We have

$$\prod_{i=1}^r (X - \alpha_i^* Y) = F(\mathbf{x}) \in T,$$

thus $\max_i |X - \alpha_i^* Y|_v \geq 1$. It follows that

$$\begin{aligned} 1 &\leq 2^{\varepsilon(v)} \max(1, \max_i |\alpha_i^*|_v) \max(|X|_v, |Y|_v) \\ &= 2^{\varepsilon(v)} m_v(F, y_0) \max(|X|_v, |Y|_v). \end{aligned}$$

Let $S = S_1 \cup S_2$. We take the product of the last inequality over $v \in S_2$ and find

$$\prod_{v \in S_2} \max(|X|_v, |Y|_v) \geq \frac{1}{2m_S(F)}.$$

Finally,

$$(7.1) \quad H_S^*(\mathbf{x}) = \prod_{v \in S_1} \max(|X|_v, |Y|_v) \prod_{v \in S_2} \max(|X|_v, |Y|_v) \geq \frac{|y/y_0|_{S_1}}{2m_S(F)},$$

which gives the right-hand side inequality of Lemma 9.

We also note that since $|y|_S \geq 1$ we have

$$|y/y_0|_{S_1} = |y/y_0|_S / |y/y_0|_{S_2} \geq |y_0|_S^{-1} |y/y_0|_{S_2}^{-1}$$

and the left-hand side inequality of Lemma 9 follows from (7.1) applied to the set S_2 and from Lemma 1.

LEMMA 10. *Let (F, y_0) be a reduced pair, let H_S^* be the reduced F -height and let \mathbf{x} and \mathbf{x}' be two projectively distinct solutions of the Thue–Mahler equation $F(x, y) \in T$, belonging to the same approximation class of size $1/N$ with $N \geq \frac{r}{r-1}|S|$ and to the same magnitude class of size $1/M$. Suppose \mathbf{x} and \mathbf{x}' are so ordered that $H_S^*(\mathbf{x}') \geq H_S^*(\mathbf{x})$ and that $H_S^*(\mathbf{x}) > (2m_S(F))^{r/2}$. Then*

$$(2m_S(F)H_S^*(\mathbf{x}'))^{1+|S|/M} \geq \frac{1}{2}((2m_S(F))^{-r/2}H_S^*(\mathbf{x}))^{(1-|S|/N)(r-1)}.$$

PROOF. Let $\{\gamma_v : v \in S\}$ and $\{\delta_v : v \in S'\}$ be the south-west and north-east corners of the cubes determined by the approximation and magnitude class of \mathbf{x} and \mathbf{x}' . Note that

$$(7.2) \quad \sum_{v \in S} \gamma_v \geq 1 - \frac{|S|}{N},$$

$$(7.3) \quad \sum_{v \in S'} \delta_v \leq 1 + \frac{|S|}{M}.$$

We have

$$\begin{aligned}
(7.4) \quad |\det(\mathbf{x}, \mathbf{x}')|_v &= |(x - \alpha_{i(v)}y)y' - (x' - \alpha_{i(v)}y')y|_v \\
&\leq 2^{\varepsilon(v)} |y_0|_v \max(|x - \alpha_{i(v)}y|_v |y'/y_0|_v, |x' - \alpha_{i(v)}y'|_v |y/y_0|_v) \\
&\leq 2^{\varepsilon(v)} |y_0|_v \max(\Lambda(\mathbf{x})^{\gamma_v} |y'/y_0|_v, \Lambda(\mathbf{x}')^{\gamma_v} |y/y_0|_v).
\end{aligned}$$

We also have

$$\begin{aligned}
|y/y_0|_v &\leq |y/y_0|_{S'}^{\delta_v}, & |y'/y_0|_v &\leq |y'/y_0|_{S'}^{\delta_v} \text{ for } v \in S', \\
|y/y_0|_v &\leq 1, & |y'/y_0|_v &\leq 1 \text{ for } v \notin S'.
\end{aligned}$$

We substitute these inequalities in (7.4) as the case may be, substitute the upper bound of Lemma 5 for $\Lambda(\mathbf{x})$ and $\Lambda(\mathbf{x}')$, note that $H_S^*(\mathbf{x}) \leq H_S^*(\mathbf{x}')$, and take the product over $v \in S$. We get, using $|\det(\mathbf{x}, \mathbf{x}')|_S \geq 1$,

$$1 \leq 2|y_0|_S \left(\left(\frac{2m_S(F)}{|y_0|_S} \right)^{r(r-1)/2} H_S^*(\mathbf{x})^{-(r-1)} \right)^{\Sigma \gamma_v} \max(|y/y_0|_{S'}, |y'/y_0|_{S'})^{\Sigma \delta_v}.$$

In view of (7.2) and (7.3) and $|y_0|_S \geq 1$ this simplifies to

$$(7.5) \quad 1 \leq 2((2m_S(F))^{r/2} H_S^*(\mathbf{x})^{-1})^{(1-|S|/N)(r-1)} \max(|y/y_0|_{S'}, |y'/y_0|_{S'})^{1+|S|/M};$$

here we have used the condition $N \geq \frac{r}{r-1}|S|$ to absorb the term $|y_0|_S$ into the term $|y_0|_S^{-r(r-1)(\Sigma \gamma_v/2)}$.

By Lemma 9 and $H_S^*(\mathbf{x}) \leq H_S^*(\mathbf{x}')$ we have $\max(|y/y_0|_{S'}, |y'/y_0|_{S'}) \leq 2m_S(F)H_S^*(\mathbf{x}')$. We substitute this inequality into (7.5) and the conclusion of Lemma 10 follows.

VIII. Counting small solutions. In this section we apply the Gap Principle of Section V to obtain bounds for the number of solutions of bounded height in a given approximation and magnitude class.

LEMMA 11. *Let (F, y_0) be a reduced pair and let H_S^* be the associated reduced F -height. Let K, M and N be positive integers such that*

$$M \geq (4K + 8)|S|, \quad N \geq 4(r-1)(K+1)|S|.$$

Then if $2m_S(F) > (8r)^4$ then the number of projective solutions x/y of the Thue–Mahler equation $F(x, y) \in T$, in a given approximation class of size $1/N$ and magnitude class of size $1/M$ and with reduced F -height $H_S^(\mathbf{x}) < (2m_S(F))^K$, does not exceed $64(K+1)(K+2)$.*

If instead $2m_S(F) \leq (8r)^4$ then the number of projective solutions x/y of the Thue–Mahler equation $F(x, y) \in T$ in a given approximation class

of size $1/N$ and magnitude class of size $1/M$ and with reduced F -height $H_S^*(\mathbf{x}) < (2m_S(F))^K$ does not exceed

$$\frac{25}{16}(K+1)(K+2)r|S||T|(8r)^{\left(\frac{18}{5}-\frac{|S|}{N}\frac{4}{5}(r-1)(K+1)\right)[k:\mathbb{Q}]} 8^{[k:\mathbb{Q}]+2} \binom{6|S|}{|S|}.$$

Proof. Let \mathbf{x} and \mathbf{x}' be two solutions belonging to the same approximation and magnitude class of sizes $1/N$ and $1/M$. Let $\{\delta_v : v \in S'\}$ and $\{\eta_v : v \in S''\}$ be the north-east and south-west corners of the two cubes associated with the magnitude class of \mathbf{x} and \mathbf{x}' . Then we have

$$(8.1) \quad |y/y_0|_{S'}^{\delta_v-1/M} \leq |y/y_0|_v \leq |y/y_0|_{S'}^{\delta_v},$$

$$(8.2) \quad |y/y_0|_{S''}^{\eta_v+1/M} \leq |y/y_0|_v \leq |y/y_0|_{S''}^{\eta_v}$$

for $v \in S'$ and $v \in S''$ respectively; the same inequality holds if we replace y by y' , because \mathbf{x} and \mathbf{x}' have the same magnitude class.

We also have

$$(8.3) \quad \sum_{v \in S'} \delta_v \leq 1 + \frac{|S'|}{M},$$

$$(8.4) \quad \sum_{v \in S''} \eta_v \leq 1.$$

Suppose for example that $\Lambda(\mathbf{x}') \leq \Lambda(\mathbf{x})$.

Let us abbreviate $B = (2m_S(F))^{1/8} + \varepsilon$ where $\varepsilon > 0$ is small and let us say that a solution \mathbf{x} is of *type* (m, n) if

$$B^m \leq |y/y_0|_{S'} < B^{m+1}, \quad B^{-n-1} < |y/y_0|_{S''} \leq B^{-n}.$$

Then (8.1) and (8.2) show that if \mathbf{x} and \mathbf{x}' have the same type (m, n) we must have

$$(8.5) \quad \max\left(1, \frac{|y'|_v}{|y|_v}\right) \leq B^{\delta_v+(m+1)/M} \quad \text{if } v \in S',$$

$$(8.6) \quad \max\left(1, \frac{|y'|_v}{|y|_v}\right) \leq B^{\eta_v+(n+1)/M} \quad \text{if } v \in S''.$$

We apply Lemma 8 (the Gap Principle) to the pair \mathbf{x} and \mathbf{x}' , note that $(8r)^{-1} > (2m_S(F))^{-1/4}$ and use the lower bound of Lemma 5, obtaining

$$\prod_{v \in S} \max\left(1, \frac{|y'|_v}{|y|_v}\right) \geq \frac{1}{4r} \Lambda(\mathbf{x})^{|S|/N} m_S(F) > (2m_S(F))^{\frac{3}{4}-\frac{|S|}{N}(r-1)(K+1)}.$$

By (8.3)–(8.6), we now see that

$$(8.7) \quad B^{2+\max(m+1, n+1)|S|/M} > (2m_S(F))^{\frac{3}{4}-\frac{|S|}{N}(r-1)(K+1)}$$

whenever there are two distinct solutions \mathbf{x} and \mathbf{x}' with the same type (m, n) .

By choosing ε sufficiently small, we may interpret (8.7) as saying that as long as

$$(8.8) \quad \max(m+1, n+1) \leq \frac{8M}{|S|} \left(\frac{1}{2} - \frac{|S|}{N} (r-1)(K+1) \right)$$

there is at most one solution for each type (m, n) .

It remains to bound the number of types and to verify condition (8.8).

We apply Lemma 9 to \mathbf{x} and obtain

$$\begin{aligned} |y/y_0|_{S'} &\leq (2m_S(F))^{K+1}, \\ |y/y_0|_{S''} &\geq (2m_S(F))^{-K-2}, \end{aligned}$$

therefore we must have $m < 8(K+1)$ and $n < 8(K+2)$ and the number of types does not exceed $64(K+1)(K+2)$. Moreover, $\max(m+1, n+1) \leq 8K+16$ and (8.8) becomes a consequence of the assumption

$$M \geq (4K+8)|S|, \quad N \geq 4(r-1)(K+1)|S|$$

of Lemma 11. This completes the proof of the first clause of Lemma 11.

Now suppose that $2m_S(F) \leq (8r)^4$. We apply the second clause of Lemma 8, again using the lower bound of Lemma 5, to infer that

$$(8.9) \quad \prod_{v \in S} \max \left(1, \frac{|y'|_v}{|y|_v} \right) \geq U$$

except for at most

$$(8.10) \quad \frac{25}{16} (8(2m_S(F))^{\frac{|S|}{N}(r-1)(K+1)} U^{6/5})^{[k:\mathbb{Q}]_r |S| |T|} \left(\frac{6|S|}{|S|} \right)$$

exceptions \mathbf{x}' for each choice of \mathbf{x} , including \mathbf{x} . If we choose

$$U = (2m_S(F))^{\frac{3}{4} - \frac{|S|}{N}(r-1)(K+1)}$$

then the counting proceeds exactly as before, except that the final bound has to be multiplied by the quantity in (8.10). An easy majorization now yields the last statement of Lemma 11.

IX. Counting large solutions. In this section we use the Strong Gap Principle to bound the number of large, but not very large, solutions of the Thue–Mahler equation in a fixed approximation and magnitude class of size $1/N$.

LEMMA 12. *Let (F, y_0) be a reduced pair. Let N be a positive integer such that $N > \frac{r}{r-2}|S|$, let*

$$\lambda = \frac{N - |S|}{N + |S|} (r - 1)$$

and let H satisfy $H \geq (2m_S(F))^{2\lambda r/(\lambda-1)}$. Then the number of projective solutions x/y of the Thue–Mahler equation $F(x, y) \in T$, in a given approximation class of size $1/N$ and magnitude class of size $1/N$ and with reduced F -height satisfying

$$\log H \leq \log H_S^*(\mathbf{x}) \leq \frac{\lambda^n}{2} \log H,$$

does not exceed n .

Proof. The hypothesis on N implies $\lambda > 1$. We order the solutions in the interval $H_S^*(\mathbf{x}) \geq H$ by increasing height, thus obtaining a sequence $\mathbf{x}_0, \mathbf{x}_1, \dots$. By Lemma 10 (the Strong Gap Principle) we have

$$H_S^*(\mathbf{x}_{n+1}) \geq D^{-1} H_S^*(\mathbf{x}_n)^\lambda$$

where

$$D = 2^{N/(|S|+N)} (2m_S(F))^{\lambda r/2} \quad \text{for } n = 0, 1, \dots$$

Now a classical iteration gives

$$H_S^*(\mathbf{x}_n) > (D^{-1/(\lambda-1)} H_S^*(\mathbf{x}_0))^{\lambda^n} \geq (D^{-1/(\lambda-1)} H)^{\lambda^n}$$

and Lemma 12 follows by noting that the lower bound on H implies $D^{-1/(\lambda-1)} H > H^{1/2}$.

X. Counting very large solutions. In this section we use the method of Thue to give an upper bound for the number of solutions of the Thue–Mahler equation in a given approximation and magnitude class and with sufficiently large height. The proof follows very closely the corresponding section of [B1], with some improvements in the numerical treatment.

We say that a solution \mathbf{x} of the Thue–Mahler equation $F(\mathbf{x}) \in T$ is *very large* if $H_S^*(\mathbf{x}) \geq (2m_S(F))^{10000r^3}$.

LEMMA 13. *Let $c(6) = 33$, $c(7) = 10$, $c(8) = 6$, $c(9) = 5$, $c(10) = c(11) = 4$, $c(r) = 3$ if $12 \leq r \leq 19$ and $c(r) = 2$ if $r \geq 20$. Suppose that $r \geq 6$ and $N \geq c(r)|S|$. Then the number of very large solutions \mathbf{x} of the Thue–Mahler equation $F(x, y) \in T$ in a given approximation and magnification class of size $1/N$ does not exceed 8.*

Proof. We recall the following result. Let k be a number field, let k' be an extension of k of degree $r \geq 2$ and let α', α'' be two elements of k' of degree r over k .

THUE PRINCIPLE. *Let $t, \tau, \theta, \delta', \delta''$ be positive real numbers such that*

$$\sqrt{\frac{2}{r+1}} < t < \sqrt{\frac{2}{r}}, \quad \tau < t, \quad t < \theta < 1/t$$

and let $\beta', \beta'' \in k$. Then either

$$\prod_{v \in S} \max(|\alpha' - \beta'|_v^{\theta \delta'}, |\alpha'' - \beta''|_v^{\theta^{-1} \delta''}) \\ > ((3H(\alpha'))^C H(\beta'))^{-\delta'/(t-\tau)} ((3H(\alpha''))^C H(\beta''))^{-\delta''/(t-\tau)}$$

with $C = 2/(2 - rt^2)$, or

$$\frac{r \delta''}{2 \delta'} > \frac{r}{2} t^2 + \frac{1}{2} \tau^2 - 1.$$

This is a rephrasing of Theorem 2 of [B2], with the following additional remarks. The conditions $|\alpha' - \beta'|_v \leq 1$ and $|\alpha'' - \beta''|_v \leq 1$ in [B2] are superfluous, because otherwise a stronger result can be obtained omitting the places $v \in S$ for which these inequalities fail; also the condition $k'_v = k_v$ in [B2] can be omitted, because it is used nowhere in the proof and it was stated only because otherwise we cannot get arbitrarily close approximations, a fact which is irrelevant for our purposes.

In our case we apply the Thue Principle as follows. Let $\mathbf{x}_1 = (x_1, y_1)$ and $\mathbf{x}_2 = (x_2, y_2)$ be two very large solutions of the Thue–Mahler equation belonging to the same approximation class and magnitude class of size $1/N$, and assume that

$$(2m_S(F))^{10000r^3} \leq H_S^*(\mathbf{x}_1) \leq H_S^*(\mathbf{x}_2).$$

We fix a root α of F and choose $\alpha' = \alpha'' = (\alpha - \bar{\alpha})y_0$, and $\beta' = (x_1 - \bar{\alpha}y_1)/(y_1/y_0)$, $\beta'' = (x_2 - \bar{\alpha}y_2)/(y_2/y_0)$. Since \mathbf{x}_1 and \mathbf{x}_2 are in the same approximation class, we can choose the extension of v to $k' = k(\alpha)$ for $v \in S$ so that

$$|x - \alpha_{i(v)}y|_v = |x - \alpha y|_v$$

for every $v \in S$.

We have for each $v \in S$ the bound

$$|\alpha' - \beta'|_v = |y_1/y_0|_v^{-1} |x_1 - \alpha y_1|_v \leq |y_1/y_0|_v^{-1} \Lambda(\mathbf{x}_1)^{\gamma_v}$$

where the γ_v 's are as in Section VII. Moreover, by Lemma 9 we have

$$|y_1/y_0|_v \geq 1 \quad \text{if } v \in S',$$

$$|y_1/y_0|_v \geq |y_1/y_0|_{S''}^{\eta_v+1/N} \geq ((2m_S(F))^2 H_S^*(\mathbf{x}_1))^{-\eta_v-1/N} \quad \text{if } v \in S''$$

and using Lemma 5 we obtain

$$(10.1) \quad |\alpha' - \beta'|_v \leq \left(\frac{(2m_S(F))^{r(r-1)/2}}{H_S^*(\mathbf{x}_1)^{r-1}} \right)^{\gamma_v} \quad \text{if } v \in S',$$

$$(10.2) \quad |\alpha' - \beta'|_v \leq ((2m_S(F))^2 H_S^*(\mathbf{x}_1))^{\eta_v+1/N} \left(\frac{(2m_S(F))^{r(r-1)/2}}{H_S^*(\mathbf{x}_1)^{r-1}} \right)^{\gamma_v} \\ \text{if } v \in S''.$$

Since we assume $r \geq 6$ and $(2m_S(F))^{10000r^3} \leq H_S^*(\mathbf{x}_1)$, inequalities (10.1) and (10.2) may be simplified to

$$(10.3) \quad |\alpha' - \beta'|_v \leq H_S^*(\mathbf{x}_1)^{-(r-1.001)\gamma_v} \quad \text{if } v \in S',$$

$$(10.4) \quad |\alpha' - \beta'|_v \leq H_S^*(\mathbf{x}_1)^{1.001(\eta_v+1/N)-(r-1.001)\gamma_v} \quad \text{if } v \in S'';$$

similar bounds also hold with $\alpha'', \beta'', \mathbf{x}_2$ in place of $\alpha', \beta', \mathbf{x}_1$, because \mathbf{x}_1 and \mathbf{x}_2 belong to the same approximation and magnitude classes.

Now we choose $\theta = 1$, $\delta' = 1/\log H_S^*(\mathbf{x}_1)$ and $\delta'' = 1/\log H_S^*(\mathbf{x}_2)$. By (10.3), (10.4) and $\sum \eta_v \leq 1$, $\sum \gamma_v \geq 1 - |S|/N$ we deduce

$$(10.5) \quad \prod_{v \in S} \max(|\alpha' - \beta'|_v^{\theta\delta'}, |\alpha'' - \beta''|_v^{\theta^{-1}\delta''}) \leq \exp\left(2.002 - r + (r + 2.002)\frac{|S|}{N}\right).$$

In order to apply the Thue Principle we need to compare this quantity with

$$((3H(\alpha'))^C H(\beta'))^{-\delta'/(t-\tau)} ((3H(\alpha''))^C H(\beta''))^{-\delta''/(t-\tau)}.$$

This is easily done. Let w, v run over all places of k', k respectively. Since $r\alpha' = (r\alpha - r\bar{\alpha})y_0$ is integral over \mathcal{O}_S we have $|(\alpha_i - \bar{\alpha})y_0|_v \leq |1/r|_v$ for every i and every $v \notin S$, hence

$$\begin{aligned} H(\alpha') &= \prod_{v \in S} \max(1, |\alpha'|_w) = \prod_{i=1}^r \left(\prod_{v \in S} \max(1, |(\alpha_i - \bar{\alpha})y_0|_v) \right)^{1/r} \\ &\leq \prod_{v \in S} \max(1, \max_i |(\alpha_i - \bar{\alpha})y_0|_v) \prod_{v \notin S} |1/r|_v. \end{aligned}$$

It follows that

$$(10.6) \quad H(\alpha') \leq rm_S(F).$$

We also have

$$\begin{aligned} H(\beta') &= \prod_v \max(1, |y_0(x_1 - \bar{\alpha}y_1)/y_1|_v) = \prod_v \max(|x_1 - \bar{\alpha}y_1|_v, |y_1/y_0|_v) \\ &= H_S^*(\mathbf{x}_1) \prod_{v \notin S} \max(1, |y_0(x_1 - \bar{\alpha}y_1)/y_1|_v). \end{aligned}$$

Since $r(x_1 - \bar{\alpha}y_1)$ and $y_0(y_1/y_0)$ are S -integers we have

$$|x_1 - \bar{\alpha}y_1|_v \leq |1/r|_v, \quad |y_1/y_0|_v \leq |1/y_0|_v \quad \text{for } v \notin S,$$

and by Lemma 1 we deduce

$$(10.7) \quad H(\beta') \leq r|y_0|_S H_S^*(\mathbf{x}_1) \leq 2rm_S(F) H_S^*(\mathbf{x}_1).$$

The same inequalities as (10.6) and (10.7) hold with α'' , β'' and \mathbf{x}_2 in place of α' , β' and \mathbf{x}_1 . It is now clear that

$$\begin{aligned} & ((3H(\alpha'))^C H(\beta'))^{-\delta'/(t-\tau)} ((3H(\alpha''))^C H(\beta''))^{-\delta''/(t-\tau)} \\ & \geq ((3rm_S(F))^{C+1} H_S^*(\mathbf{x}_1))^{-\delta'/(t-\tau)} ((3rm_S(F))^{C+1} H_S^*(\mathbf{x}_2))^{-\delta''/(t-\tau)} \end{aligned}$$

and using $(2m_S(F))^{10000r^3} \leq H_S^*(\mathbf{x}_i)$ we may simplify the last displayed inequality to

$$(10.8) \quad \begin{aligned} & ((3H(\alpha'))^C H(\beta'))^{-\delta'/(t-\tau)} ((3H(\alpha''))^C H(\beta''))^{-\delta''/(t-\tau)} \\ & \geq \exp\left(-\left(2 + \frac{C+1}{5000r^2}\right) \frac{1}{t-\tau}\right). \end{aligned}$$

Now we can apply the Thue Principle. In view of (10.5) and (10.8) the first alternative becomes

$$(10.9) \quad r - 2.002 - (r + 2.002) \frac{|S|}{N} \leq \left(2 + \frac{4 - rt^2}{5000r^2(2 - rt^2)}\right) \frac{1}{t - \tau}$$

while the second alternative yields

$$(10.10) \quad \log H_S^*(\mathbf{x}_1) \leq \log H_S^*(\mathbf{x}_2) < \frac{r}{rt^2 + \tau^2 - 2} \log H_S^*(\mathbf{x}_1).$$

We choose for example $t^2 = 2/(r + 0.005)$ and $\tau^2 = 2 - rt^2 + 0.0001$. Let $c(6) = 33$, $c(7) = 10$, $c(8) = 6$, $c(9) = 5$, $c(10) = c(11) = 4$, $c(r) = 3$ if $12 \leq r \leq 19$ and $c(r) = 2$ if $r \geq 20$. Then a numerical calculation shows that the first alternative (10.9) cannot hold if $N \geq c(r)|S|$, hence the second alternative (10.10) must hold and we get

$$(10.11) \quad \log H_S^*(\mathbf{x}_1) \leq \log H_S^*(\mathbf{x}_2) < 10000r \log H_S^*(\mathbf{x}_1).$$

We choose for \mathbf{x}_1 the smallest solution (assuming it exists) with $H_S^*(\mathbf{x}_1) \geq (2m_S(F))^{10000r^3}$. Now by Lemma 12 and

$$\lambda = \frac{N - |S|}{N + |S|} (r - 1) \geq \frac{c(r) - 1}{c(r) + 1} (r - 1)$$

we see that the number of solutions of (10.11) does not exceed

$$[\log(20000r) / \log \lambda] + 1 \leq 8.$$

This proves Lemma 13.

XI. A bound for $N_0(F)$. It is now easy to put together the information obtained so far and get a bound for $N_0(\bar{F})$. Let $c(r)$ be as in Lemma 13.

LEMMA 14. *Suppose $r \geq 6$ and let $\lambda = \frac{c(r)-1}{c(r)+1}(r-1)$. Let $a(r)$ be given by $a(6) = 32921$, $a(7) = 3180$, $a(8) = 1606$, $a(9) = 852$, $a(10) = a(11) = 564$, $a(r) = 335$ if $12 \leq r \leq 19$ and $a(r) = 165$ if $r \geq 20$. Let (F, y_0) be*

a reduced pair. Then the number of projective solutions x/y of the Thue–Mahler equation $F(x, y) \in T$ with reduced F -height

$$H_S^*(\mathbf{x}) \geq (2m_S(F))^{2\lambda r/(\lambda-1)}$$

does not exceed $|S|^3(a(r)r)^{|S|}$.

PROOF. Let $N = c(r)|S|$. We divide solutions into approximation and magnitude classes of size $1/N$, and divide the range $H_S^*(\mathbf{x}) \geq (2m_S(F))^{2\lambda r/(\lambda-1)}$ into the two ranges

$$\begin{aligned} (2m_S(F))^{2\lambda r/(\lambda-1)} &\leq H_S^*(\mathbf{x}) < (2m_S(F))^{10000r^3}, \\ (2m_S(F))^{10000r^3} &\leq H_S^*(\mathbf{x}). \end{aligned}$$

In the first interval we use Lemma 12, obtaining not more than

$$[\log(20000r^3)/\log \lambda] \leq 9$$

solutions in each class. In the second interval we use Lemma 13, obtaining at most 8 solutions in each class. By Lemmas 6 and 7 the number of classes does not exceed

$$\frac{|S|^6}{(N+|S|)(4N+2|S|)^2} \left(r \left(\frac{N}{|S|} + 1 \right) \left(\frac{4N}{|S|} + 2 \right) \right)^{|S|} \left(1 + \frac{|S|}{N} \right)^N \left(1 + \frac{|S|}{2N} \right)^{2N},$$

and Lemma 14 follows after a straightforward numerical calculation.

LEMMA 15. Suppose (F, y_0) is a reduced pair. Then the number of projective solutions x/y of the Thue–Mahler equation $F(x, y) \in T$ with reduced F -height

$$H_S^*(\mathbf{x}) < (2m_S(F))^{2\lambda r/(\lambda-1)}$$

does not exceed $0.01|S|^3|T|e^{29.13|S|}r^{10.8|S|-1}$.

PROOF. We take $K = \lfloor \frac{2\lambda r}{\lambda-1} \rfloor + 1$ and apply Lemma 11 with $M = (4K + 8)|S|$, $N = 4(r-1)(K+1)|S|$. The bound in Lemma 15 will be the number of approximation and magnitude classes of sizes $1/N$, $1/M$ times the bound appearing in Lemma 11. By Lemmas 6 and 7 the number of approximation classes of size $1/N$ and of magnitude classes of size $1/M$ does not exceed

$$\binom{N+|S|-1}{|S|-1} \frac{|S|^5}{(4M+2|S|)^2} (2r)^{|S|} \left(\frac{2M}{|S|} + 1 \right)^{|S|} \left(1 + \frac{|S|}{2M} \right)^{2M}.$$

We note that the choices of K , M , N imply

$$\binom{N+|S|-1}{|S|-1} \leq \frac{2}{5} \sqrt{\frac{N+|S|}{N|S|}} \left(\frac{N}{|S|} + 1 \right)^{|S|-1} \left(1 + \frac{|S|}{N} \right)^N$$

and

$$2r \left(\frac{N}{|S|} + 1 \right) \left(\frac{2M}{|S|} + 1 \right) \left(1 + \frac{|S|}{N} \right)^{N/|S|} \left(1 + \frac{|S|}{2M} \right)^{2M/|S|} < 3367r^4,$$

therefore the number of approximation and magnitude classes does not exceed

$$(11.1) \quad \frac{2}{5} \sqrt{\frac{N+|S|}{N|S|}} \frac{|S|^6}{(N+|S|)(4M+2|S|)^2} (3367r^4)^{|S|}.$$

We also have

$$(K+1)(K+2) \leq \frac{17}{2}r^2, \quad \left(\frac{6|S|}{|S|}\right) \leq \frac{4}{9}|S|^{-1/2}(6^6/5^5)^{|S|}$$

hence the bound in Lemma 11 becomes

$$(11.2) \quad \frac{1700}{9}r^3|S|^{1/2}|T|2^{\frac{66}{5}[k:\mathbb{Q}]}r^{\frac{17}{5}[k:\mathbb{Q}]}(6^6/5^5)^{|S|}.$$

The proof of Lemma 15 follows from multiplying (11.1) and (11.2), and using the trivial estimate $[k:\mathbb{Q}] \leq 2|S|$.

XII. Proof of the Main Theorem. The proof of the Main Theorem is now immediate. Suppose $r \geq 6$. By Lemmas 14 and 15 we infer

$$\begin{aligned} |N(F)| &\leq r^{|S|-1}|N_0(F)| \\ &\leq r^{|S|-1}(0.01|S|^3|T|e^{29.13|S|}r^{10.8|S|-1} + |S|^3(a(r)r)^{|S|}). \end{aligned}$$

After dividing by $(12r)^{12|S|}$ we get a quantity which decreases with r . Therefore, taking $r = 6$ and using the bounds $|T| \leq 4|S|^2$ if $|S| \geq 15$, $|T| \leq 7.5|S|$ if $|S| \leq 15$, we see that this ratio does not exceed 1. This completes the proof.

References

- [B1] E. Bombieri, *On the Thue–Mahler equation*, in: Diophantine Approximation and Transcendence Theory (Seminar, Bonn, 1985), G. Wüstholz (ed.), Lecture Notes in Math. 1290, Springer, 1987, 213–243.
- [B2] —, *On the Thue–Siegel–Dyson theorem*, Acta Math. 148 (1982), 255–296.
- [BS] E. Bombieri and W. M. Schmidt, *On Thue’s equation*, Invent. Math. 88 (1987), 69–81; correction, ibid. 97 (1989), 445.
- [E] J.-H. Evertse, *On equations in S -units and the Thue–Mahler equation*, ibid. 75 (1984), 561–584.
- [S] C. L. Stewart, *On the number of solutions of polynomial congruences and Thue equations*, J. Amer. Math. Soc. 4 (1991), 793–835.

INSTITUTE FOR ADVANCED STUDY
PRINCETON, NEW JERSEY 08540
U.S.A.

*Received on 9.8.1993
and in revised form on 23.12.1993*

(2471)