JOURNAL OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 4, Number 4, October 1991

ON THE NUMBER OF SOLUTIONS OF POLYNOMIAL CONGRUENCES AND THUE EQUATIONS

C. L. STEWART

1. INTRODUCTION

Let $F(x, y) = a_r x^r + a_{r-1} x^{r-1} y + \cdots + a_0 y^r$ be a binary form with rational integer coefficients and with $r \geq 3$. Let h be a nonzero integer. In 1909 Thue proved that if F is irreducible then the equation

$$F(x, y) = h \tag{1}$$

has only finitely many solutions in integers x and y. In the first part of this paper we shall establish upper bounds for the number of solutions of (1) in coprime integers x and y under the assumption that the discriminant $D(F)$ of F is nonzero. For most integers h these bounds improve upon those obtained by Bombieri and Schmidt in [5]. In the course of proving these bounds we shall establish a result on polynomial congruences that extends earlier work of Nagell [30], Ore [32], Sándor [33], and Huxley [19]. In fact we shall establish an upper bound for the number of solutions of a polynomial congruence that is, in general, best possible.

In the second part we shall address the problem of finding forms F for which (1) has many solutions for arbitrarily large integers h. Finally we shall obtain upper bounds for the number of solutions of certain Thue-Mahler and Ramanujan-Nagell equations by appealing to estimates of Evertse, Györy, Stewart, and Tijdeman [17] for the number of solutions of S-unit equations.

2. THE THUE AND THUE-MAHLER EQUATIONS

For any nonzero integer h let $\omega(h)$ denote the number of distinct prime factors of h. In 1933 Mahler [23] proved that if F is irreducible then (1) has at most $C_1^{1+\omega(h)}$ solutions in coprime integers x and y, where C_1 is a positive number that depends on F only. Let $p_1, \ldots, p_t$ be distinct prime numbers. The equation

$$F(x, y) = p_1^{k_1} \cdots p_t^{k_t} \tag{2}$$

Received by the editors March 21, 1991.
1991 *Mathematics Subject Classification*. Primary 11D41, 11C08, 11J25.
Key words and phrases. Thue equations, polynomial congruences, S-unit equations.
Research supported in part by a Killam Research Fellowship and by Grant A3528 from the Natural Sciences and Engineering Research Council of Canada.

in coprime integers x and y and integers $k_1, \ldots, k_t$ is known as a Thue-Mahler equation. In fact Mahler proved the stronger result that (2) has at most C_1^{1+t} such solutions. In 1938 Erdös and Mahler [9] proved that if F has nonzero discriminant, $h > C_2$ and g is a divisor of h with $g > h^{6/7}$ then the number of solutions of (1) in coprime integers x and y is at most $C_3^{1+\omega(g)}$, where C_2 and C_3 are positive numbers that depend on F only. In 1961 Lewis and Mahler [21] showed that the number of primitive solutions of (2), that is, solutions with x and y coprime, is at most

$$c_1(ar)^{c_2\sqrt{r}} + (c_3 r)^{1+t},$$

where c_1, c_2, and c_3 are absolute constants, provided F has nonzero discriminant, $a_r a_0 \neq 0$, and the coefficients of F have absolute values not exceeding a. In 1984 Evertse [13] gave

$$2 \cdot 7^{r^3(2t+3)} \tag{3}$$

as an upper bound for the number of primitive solutions of (2) under the assumption that F is divisible by at least three pairwise linearly independent linear forms in some algebraic number field. Evertse's result resolved a conjecture of Siegel since his upper bound for the number of primitive solutions of (2) depends only on r and t, and so for (1) depends only on r and $\omega(h)$, and does not depend on the coefficients of F. In 1987 Bombieri and Schmidt [5] refined the result of Evertse for the Thue equation. They proved that if F is irreducible then the number of solutions of (1) in coprime integers x and y is at most

$$c_4 r^{1+\omega(h)}, \tag{4}$$

where c_4 is an absolute constant. Further they showed that one may take c_4 to be 430 if r is sufficiently large.

Let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

and define the binary form F_A by

$$F_A(x, y) = F(ax + by, cx + dy).$$

Observe that if A is in $\mathrm{GL}(2, \mathbb{Z})$, in other words, A has integer entries and determinant ± 1, and (x, y) is a solution of (1) in coprime integers x and y then $A(x, y) = (ax+by, cx+dy)$ is a solution of $F_{A^{-1}}(X, Y) = h$ in coprime integers. For any $A \in \mathrm{GL}(2, \mathbb{Z})$ we say that F_A and $-F_A$ are equivalent to F. We remark that the number of solutions of (1) in coprime integers is the same for equivalent forms. For any polynomial G in $\mathbb{C}[z_1, \ldots, z_n]$ that is not identically zero the Mahler measure $M(G)$ is defined by

$$M(G) = \exp \int_0^1 dt_1 \cdots \int_0^1 dt_n \log |G(e^{2\pi i t_1}, \ldots, e^{2\pi i t_n})|.$$

Thus if $n = 1$ and $G(z) = a_r(z-\alpha_1)\cdots(z-\alpha_r)$ with $a_r \neq 0$, then, by Jensen's theorem,

$$M(G) = |a_r| \prod_{i=1}^{r} \max(1, |\alpha_i|).$$

Suppose that F is a binary form that factors as $\prod_{j=1}^{r}(\alpha_j x - \beta_j y)$. The discriminant $D = D(F)$ of F is given by

$$D(F) = \prod_{i<j} (\alpha_i \beta_j - \alpha_j \beta_j)^2 .$$

For any nonzero integer t we have

$$D(tF) = t^{2(r-1)} D(F), \tag{5}$$

and for any matrix A with integer entries

$$D(F_A) = (\det A)^{r(r-1)} D(F). \tag{6}$$

Thus for any $A \in \mathrm{GL}(2, \mathbb{Z})$ we have $D = D(F) = D(F_A)$. For any nonzero integer n and prime number p let $\mathrm{ord}_p\, n$ denote the exact power of p that divides n. For any real number x let $[x]$ denote the greatest integer less than or equal to x. Let p be a prime number, and let r, k, and D be integers with $r \geq 2$ and $D \neq 0$. We define $T = T(r, k, p, D)$ by

$$T = \min\left(\left[\left(\frac{r-1}{r}\right)k\right], \min_{j=0,\dots,r-2}\left(\left[\frac{\mathrm{ord}_p D}{(j+1)(j+2)} + \left(\frac{j}{j+2}\right)k\right]\right)\right) \tag{7}$$

and for any nonzero integer g we define $G(g, r, D)$ by

$$G(g, r, D) = \prod_{p|g} p^{T(r, \mathrm{ord}_p g, p, D)} .$$

Recall that the content of F is the greatest common divisor of the coefficients of F. We shall prove the following result.

Theorem 1. *Let F be a binary form with integer coefficients of degree r (≥ 3), content 1, and nonzero discriminant D. Let h be a nonzero integer, and let ε be a positive real number. Let g be any divisor of h with*

$$\frac{g^{1+\varepsilon}|D|^{1/r(r-1)}}{G(g, r, D)} \geq |h|^{2/r+\varepsilon} . \tag{8}$$

The number of pairs of coprime integers (x, y) for which $F(x, y) = h$ is at most

$$2800\left(1 + \frac{1}{8\varepsilon r}\right) r^{1+\omega(g)} . \tag{9}$$

Notice that if g or D is squarefree, or if g and D are coprime, then $G(g, r, D)$ is 1. Further $G(g, r, D)$ is always bounded from above by $(D, g^2)^{1/2}$; here (D, g^2) denotes the greatest common divisor of D and g^2. Thus Theorem 1 sharpens the result of Erdös and Mahler [9]. Furthermore if we take $g = |h|$ and ε any positive real number then condition (8) holds since on taking $j = r - 2$ in (7) we see that

$$T(r, k, p, D) \le \left[\frac{\operatorname{ord}_p D}{(r-1)r} + \left(\frac{r-2}{r}\right) k\right],$$

hence

$$G(|h|, r, D) \le |h|^{(r-2)/r} |D|^{1/r(r-1)}.$$

Thus if $D(F) \neq 0$ then the number of primitive solutions of (1) is at most $2800 r^{1+\omega(h)}$; in particular, we recover estimate (4) of Bombieri and Schmidt. In general this choice for g and ε is not optimal. Indeed the significant feature of estimate (9) is that the term $\omega(h)$ in estimate (4) has been replaced by the quantity $\omega(g)$. For almost all integers h in the sense of natural asymptotic density, and any $\delta > 0$, $\omega(h) = \log\log h + O(\log\log h^{1/2+\delta})$ (see [18]). On the other hand (see [6]), if $\varepsilon < (r-2)/r$ then for a positive proportion of integers h we may take g to be a prime, hence $\omega(g) = 1$, and estimate (9) becomes $C(\varepsilon) r^2$. In fact $\omega(h)$ may be as large as $\log h/(4 \log\log h)$ while $\omega(g) = 1$. No particular significance attaches to the constant 2800 in (9). It can certainly be improved. In particular, if either h or r is large, then (9) holds with a much smaller constant.

Our proof depends upon the Thue-Siegel principle as enunciated in Bombieri and Schmidt [5] and follows quite closely the proof given in [5]. (The author would like to thank Professor Evertse for his suggestion, in connection with an earlier version of this result, that he follow the approach of Bombieri and Schmidt [5] for dealing with the small solutions of (1). This allowed him to remove a factor involving $M(F)$ from his original estimates.)

Our argument differs from that of Bombieri and Schmidt in that they reduce the study of (1) to the case when $h = 1$ by splitting solutions according to congruence classes modulo h. On the other hand, we reduce h to h/g by splitting the solutions into congruence classes modulo g. Further we appeal to Theorem 2 to spread apart solutions in the same congruence class. Both arguments owe much to the work of Mahler [26].

Observe that if $|D|^{1/r(r-1)} \ge |h|^{2/r+\varepsilon}$, then we may apply Theorem 1 with $g = 1$ to deduce that the number of pairs of coprime integers (x, y) for which (1) holds is at most

$$2800\left(1 + \frac{1}{8\varepsilon r}\right) r.$$

Evertse and Győry [12, 16] have obtained a related result for the Thue inequality

$$(10) \qquad 0 < |F(x, y)| \le h.$$

Define $(N(r), \delta(r))$ by $(N(r), \delta(r)) = (6r7^{2\binom{r}{3}}, \frac{5}{6}r(r-1))$ for $3 \le r < 400$ and $(N(r), \delta(r)) = (6r, 120(r-1))$ for $r > 400$. They prove that if

$$|D| \ge h^{\delta(r)} \exp(80r(r-1)),$$

then the number of solutions of (10) in coprime integers x and y with y positive is at most $N(r)$.

Recall that the term in the denominator on the left-hand side of inequality (8) is at most $(D, g^2)^{1/2}$. If $g \geq |h|^{2/r+\varepsilon}$ then, since $|D|$ is at least 1, whenever $(D, g^2)^{1/2} \leq |h|^{\varepsilon/2}$ inequality (8) holds with ε replaced by $\varepsilon/2$. This gives immediately the following consequence of Theorem 1.

Corollary 1. *Let F be a binary form with integer coefficients of degree r (≥ 3), content 1, and nonzero discriminant D. Let h be a nonzero integer and let ε be a positive real number. Let g be any divisor of h with $g \geq |h|^{2/r+\varepsilon}$. If $|h| \geq (D, g^2)^{1/\varepsilon}$ then the number of pairs of coprime integers (x, y) for which $F(x, y) = h$ is at most*

$$2800\left(1+\frac{1}{4\varepsilon r}\right) r^{1+\omega(g)}.$$

If F has few nonzero coefficients, say s, then upper bounds for the number of primitive solutions of (1) have been given by Mueller and Schmidt [29] and Schmidt [34] that depend on s and h only. Further, the special case of binomial forms $F(x, y) = a_r x^r + a_0 y^r$ has been much studied by the hypergeometric method. This study was initiated by Siegel [35] in 1937 and refined by several authors, most recently Evertse [11] in 1982; see, in particular, Theorem 2 of [11], which is of a similar character to Corollary 1. Finally we mention that Silverman [36] in 1983 proved that if $D(F) \neq 0$ and h is r-powerfree and sufficiently large relative to F then the number of primitive solutions of (1) is at most

$$r^{2r^2}(8r^3)^{R_F(h)},$$

where $R_F(h)$ is the rank of the Mordell-Weil group of the Jacobian of the curve (1) over $\mathbb{Q}$.

3. On polynomial congruences

The results in this section were motivated by the reduction theory of §VI of Bombieri and Schmidt [5]. The author is grateful to Professor Bombieri for correspondence that clarified for him some aspects of their argument.

Let Ω_p be a completion of an algebraic closure of $\mathbb{Q}_p$, the field of p-adic numbers. Let $|\ |_p$ denote the usual p-adic value in $\mathbb{Q}_p$, so $|p|_p = p^{-1}$, as well as an extension of it to Ω_p. Let $\mathbb{Z}_p$ denote the ring of integers in $\mathbb{Q}_p$, and let R_p denote the ring of elements α in Ω_p with $|\alpha|_p \leq 1$. We define $\operatorname{ord}_p \gamma$ for $\gamma \in \Omega_p$ by $\operatorname{ord}_p \gamma = -(\log|\gamma|_p)/\log p$. Recall that the content of a polynomial f with integer coefficients is the greatest common divisor of its coefficients and the discriminant $D(f)$ of f is given by $D(F)$, where F is the binary form $F(x, y) = y^r f(x/y)$ and r is the degree of f.

For any prime p and nonzero integer D we define $l = l(p, D)$ by

$$l = \operatorname{ord}_p D.$$

Further for primes p and nonzero integers r, k, and D with $r \geq 2$, we have that $T = T(r, k, p, D)$, as defined in (7), satisfies

$$T = \begin{cases} \left[\frac{l}{2}\right] & \text{if } k \geq l, \\ \left[\frac{l}{(j+1)(j+2)} + \left(\frac{j}{j+2}\right)k\right] & \text{if } \frac{l}{j} \geq k \geq \frac{l}{j+1} \text{ for } j = 1, \ldots, r-2, \\ \left[\left(\frac{r-1}{r}\right)k\right] & \text{if } \frac{l}{r-1} \geq k \geq 1. \end{cases}$$

Theorem 2. *Let p be a prime number, and let f be a polynomial with integer coefficients, content coprime with p, degree r (≥ 2), and nonzero discriminant D. Put $\operatorname{ord}_p D = l$ and let s denote the number of zeros of f in R_p. For each positive integer k there is an integer t $(= t(k))$ with $0 \leq t \leq s$ $(\leq r)$ and there are nonnegative integers b_1 $(= b_1(k)), \ldots, b_t$ $(= b_t(k))$ and u_1 $(= u_1(k)), \ldots, u_t$ $(= u_t(k))$ such that the complete solution of the congruence*

$$f(x) \equiv 0 \pmod{p^k}, \tag{11}$$

is given by the t congruences

$$x \equiv b_i \pmod{p^{k-u_i}} \tag{12}$$

for $i = 1, \ldots, t$ and such that if $k > l$ then $t(k) = t(l+1)$ and $u_i(k) = u_i(l+1)$ for $i = 1, \ldots, t$, while if $j \geq k > l$ then

$$b_i(j) \equiv b_i(k) \pmod{p^{k-u_i(k)}} \tag{13}$$

for $i = 1, \ldots, t$. Further, for each positive integer k, p^k divides the content of $f(p^{k-u_i}x + b_i)$ for $i = 1, \ldots, t$,

$$0 \leq u_i(k) \leq T \tag{14}$$

for $i = 1, \ldots, t$, and

$$u_1 + \cdots + u_t \leq \min\left(2\left[\frac{l}{2}\right], r\left[\left(\frac{r-1}{r}\right)k\right]\right). \tag{15}$$

Furthermore, for each positive integer k, at most s_1 of the integers $b_1, \ldots, b_t$ are divisible by p, where s_1 is the number of roots α of f with $|\alpha|_p < 1$.

Proof of Theorem 2. We first prove that the solutions of (11) are given by t such congruences and to this end our argument will follow initially that of Lemma 7 of [5] or Proposition 4.8.1 of [1]. Let $U = \{\sigma \in R_p \mid |f(\sigma)|_p \leq p^{-k}\}$. Then U can be written as the disjoint union of maximal discs in R_p. Let $aR_p + b$ be such a disc with $a, b \in R_p$, $a \neq 0$, and $|a|_p \leq |b|_p$. Since f has content coprime with p there exists a γ in R_p with $|f(\gamma)|_p = 1$. Thus $aR_p + b$ is properly contained in R_p hence $|a|_p < 1$. Now let $f(x) = a_r(x-\alpha_1)\cdots(x-\alpha_r)$ in Ω_p. Observe now that since $aR_p + b$ is maximal, it contains a root α_i of

f. For otherwise there exists an $a' \in R_p$ with $|a|_p < |a'|_p < |\alpha_i - b|_p$ for $i = 1, \ldots, r$ and so $a'R_p + b$ is a disc in U that properly contains $aR_p + b$, contradicting the assumption that $aR_p + b$ is maximal. Thus each maximal disc contains a root of f in R_p and so U is the disjoint union of at most s such discs. For each disc $aR_p + b$ we consider the disc $aR_p + b \cap \mathbb{Z}_p$. This disc is either empty or of the form $A\mathbb{Z}_p + B$ with $A, B \in \mathbb{Z}_p$, $|A|_p \le |B|_p$, and $|A|_p < 1$ since $|a|_p < 1$. If x in $\mathbb{Z}$ satisfies $f(x) \equiv 0 \pmod{p^k}$ then $|f(x)|_p \le p^{-k}$ and so x lies in one of the discs $A\mathbb{Z}_p + B$. Thus there exists an integer t with $0 \le t \le s$ and integers $b_1, \ldots, b_t$ and $u_1, \ldots, u_t$ with $0 \le u_i \le k$ for $i = 1, \ldots, t$ such that x satisfies one of the congruences $x \equiv b_i \pmod{p^{k-u_i}}$ with $1 \le i \le t$.

For each integer i for which $k - u_i$ is less then k we consider the integers $e_{i,j} = b_i + jp^{k-u_i}$ for $j = 1, \ldots, p$. If for some integer i and for each integer j from 1 to p there is a root of f, say α, for which $|\alpha - e_{i,j}|_p < |\alpha - e_{i,m}|_p$ for $1 \le m \le p$ with $m \ne j$ then we replace the single congruence $x \equiv b_i \pmod{p^{k-u_i}}$ by the p congruences $x \equiv e_{i,j} \pmod{p^{k-u_i+1}}$ for $j = 1, \ldots, p$. We now relabel the b_i's and u_i's to take into account the fact that we have p new congruences in place of the single congruence $x \equiv b_i \pmod{p^{k-u_i}}$. Since each maximal disc contains a root of f we see that to each of the original b_i's we may associate a root of f that is p-adically closer to it than to any of the other b_i's. This situation still applies after the above substitution. Each such root is one of the s roots of f from R_p and hence $0 \le t \le s$. Further, there can be only finitely many applications of the above procedure. Thus, we may assume that for each integer i for which $k - u_i$ is less than k there is an integer $j = j(i)$ with $1 \le j \le p$ such that for each root α of f there is an integer $m = m(\alpha, j)$, different from j, with $1 \le m \le p$ such that $|\alpha - e_{i,m}|_p \le |\alpha - e_{i,j}|_p$. But $|e_{i,j} - e_{i,m}|_p = p^{-k+u_i}$ and so by the triangle inequality $|\alpha - e_{i,j}|_p \ge p^{-k+u_i}$. Note that we may replace b_i by $e_{i,j}$ and so may suppose that $|\alpha - b_i|_p \ge p^{-k+u_i}$ for all roots α of f whenever $k - u_i < k$. Then we may suppose, without loss of generality, that $u_i > 0$ for $i = 1, \ldots, t_1$, and that $u_i = 0$ for $i = t_1 + 1, \ldots, t$, where t_1 is an integer with $0 \le t_1 \le t$. Further, again without loss of generality, we may suppose that the roots of f are ordered so that

$$|\alpha_i - b_i|_p \le |\alpha_j - b_i|_p \tag{16}$$

for $i = 1, \ldots, t$ and $j = 1, \ldots, r$. Put

$$\delta_{i,j} = \operatorname{ord}_p(b_i - \alpha_j)$$

for $i = 1, \ldots, t$ and $j = 1, \ldots, r$ and note that, by (16),

$$\delta_{i,i} \ge \delta_{i,j} \tag{17}$$

for $i=1,\ldots,t$ and $j=1,\ldots,r$. Since $|f(b_i)|_p \le p^{-k}$ we have

$$\operatorname{ord}_p a_r + \sum_{j=1}^{r} \delta_{i,j} \ge k \tag{18}$$

for $i=1,\ldots,t$. Also, since $|\alpha_j - b_i|_p \ge p^{-k+u_i}$ we have

$$\delta_{i,j} \le k - u_i \tag{19}$$

for $i=1,\ldots,t_1$ and $j=1,\ldots,r$. Since f has content coprime with p,

$$\operatorname{ord}_p a_r + \sum_{\operatorname{ord}_p \alpha_j < 0} \operatorname{ord}_p \alpha_j = 0. \tag{20}$$

Thus, for all integers b,

$$\operatorname{ord}_p a_r + \sum_{\operatorname{ord}_p \alpha_j < 0} \operatorname{ord}_p (b - \alpha_j) = 0. \tag{21}$$

Accordingly, by (18) and (21) with b replaced by b_i,

$$\sum_{1 \le j \le r,\ \delta_{i,j} \ge 0} \delta_{i,j} \ge k \tag{22}$$

for $i=1,\ldots,t$. Therefore, by (19) and (22),

$$\sum_{1 \le j \le r,\ j \ne i}^{*} \delta_{i,j} \ge u_i \tag{23}$$

for $i=1,\ldots,t$. Here $\sum^*$ indicates that the sum is taken over those terms $\delta_{i,j}$ that are nonnegative. Further, for integers i and j with $1 \le i < j \le r$, $\alpha_i - \alpha_j = (b_h - \alpha_j) - (b_h - \alpha_i)$ for $h = 1,\ldots,t$, hence

$$\operatorname{ord}_p(\alpha_i - \alpha_j) \ge \max_{1 \le h \le t} \{\min(\delta_{h,i}, \delta_{h,j})\}. \tag{24}$$

Recall that

$$\frac{l}{2} = \frac{1}{2} \operatorname{ord}_p D = (r-1) \operatorname{ord}_p a_r + \sum_{i<j} \operatorname{ord}_p(\alpha_i - \alpha_j). \tag{25}$$

By (20),

$$(r-1) \operatorname{ord}_p a_r + \sum_{i<j,\ \min(\operatorname{ord}_p \alpha_i, \operatorname{ord}_p \alpha_j) < 0} \operatorname{ord}_p(\alpha_i - \alpha_j) \ge 0$$

and so, by (25),

$$\frac{l}{2} \ge \sum_{i<j,\ \min(\operatorname{ord}_p \alpha_i, \operatorname{ord}_p \alpha_j) \ge 0} \operatorname{ord}_p(\alpha_i - \alpha_j). \tag{26}$$

Thus it follows from (24) that

$$\frac{l}{2} \ge \sum_{i<j}^{*} \max_{1 \le h \le t} \{\min(\delta_{h,i}, \delta_{h,j})\}. \tag{27}$$

By (17), $\min(\delta_{i,i}, \delta_{i,j}) = \delta_{i,j}$ for $i \neq j$ and so

$$\frac{l}{2} \geq \sum^{*}_{1\leq j\leq r,\, j\neq i} \delta_{i,j} + \sum^{*}_{j<m,\, j\neq i,\, m\neq i} \min(\delta_{i,j}, \delta_{i,m}) \tag{28}$$

for $i = 1, \ldots, t$.

Let us for the moment fix i with $1 \leq i \leq t$, and put $u_i = u$. Let n denote the number of terms $\delta_{i,j}$ with $i \neq j$ and $\delta_{i,j} \geq 0$; note that $0 \leq n \leq r-1$. Relabel these terms as $x_1, \ldots, x_n$ in such a way that $x_1 \geq x_2 \geq \cdots \geq x_n \geq 0$. Then, from (28),

$$\frac{l}{2} \geq x_1 + \cdots + x_n + \sum_{1\leq j<m\leq n} \min(x_j, x_m),$$

hence

$$\frac{l}{2} \geq x_1 + 2x_2 + \cdots + nx_n. \tag{29}$$

Further, by (23),

$$x_1 + \cdots + x_n \geq u \tag{30}$$

and, by (19),

$$-x_m \geq u - k \tag{31}$$

for $m = 1, \ldots, n$. Since $x_m \geq 0$ for $m = 1, \ldots, n$ we deduce from (29) that

$$\frac{l}{2} \geq j(x_1 + \cdots + x_n) - (j-1)x_1 - (j-2)x_2 - \cdots - x_{j-1}$$

for $j = 1, \ldots, n$. By (30) and (31),

$$\frac{l}{2} \geq ju + \frac{j(j-1)}{2}(u-k),$$

hence

$$u \leq \frac{l}{j(j+1)} + \left(\frac{j-1}{j+1}\right)k$$

for $j = 1, \ldots, n$. It also follows from (30) and (31) that $n(k-u) \geq u$, whence

$$u \leq \left(\frac{n}{n+1}\right)k.$$

Therefore

$$u \leq \min\left(\left(\frac{n}{n+1}\right)k, \min_{j=0,\ldots,n-1}\left(\frac{l}{(j+1)(j+2)} + \left(\frac{j}{j+2}\right)k\right)\right).$$

Since $n \leq r-1$, we certainly have

$$u \leq \min\left(\left(\frac{r-1}{r}\right)k, \min_{j=0,\ldots,r-2}\left(\frac{l}{(j+1)(j+2)} + \left(\frac{j}{j+2}\right)k\right)\right),$$

which establishes (14).

Note that for any pair of integers (i, j) with $1 \le i < j \le t$,

$$(32)\quad \max_{1\le h\le t}(\min(\delta_{h,i}, \delta_{h,j})) \ge \frac{\min(\delta_{i,i}, \delta_{i,j}) + \min(\delta_{j,i}, \delta_{j,j})}{2} \ge \frac{\delta_{i,j}+\delta_{j,i}}{2}.$$

Further, for any pair of integers (i, j) with $1 \le i \le t < j \le r$,

$$\max_{1\le h\le t}(\min(\delta_{h,i}, \delta_{h,j})) \ge \delta_{i,j}.$$

Thus, by (27),

$$(33)\qquad l \ge \sum_{i=1}^{t}\Bigl(\sideset{}{^*}\sum_{1\le j\le r,\ j\ne i}\delta_{i,j} + \sideset{}{^*}\sum_{t<j\le r}\delta_{i,j}\Bigr),$$

and, by (23),

$$(34)\qquad u_1 + \cdots + u_t \le l.$$

Observe that if (34) holds with equality then (23), (27), and (33) must also hold with equality. By (23), $\delta_{i,j} \le 0$ whenever $i > t_1$. By (33), if $1 \le i \le t_1$ and $t < j \le r$ then $\delta_{i,j} \le 0$. By (32), if $1 \le i \le t$, $i < j \le t$, and $\delta_{i,j} \ge 0$ then $\delta_{i,j} = \delta_{j,i}$. Therefore if $u_1 + \cdots + u_t = l$ then, by (33),

$$(35)\qquad l = \sum_{i=1}^{t_1}\Bigl(\sideset{}{^*}\sum_{1\le j\le t,\ j\ne i}\delta_{i,j}\Bigr) = \sum_{i=1}^{t_1} 2\Bigl(\sideset{}{^*}\sum_{i<j\le t_1}\delta_{i,j}\Bigr).$$

Further if (34) holds with equality then by (19) and (23) we see that $\delta_{i,i} = k - u_i$ for $i = 1, \dots, t_1$. Since

$$(\alpha_i - b_i) - (\alpha_i - b_j) = b_j - b_i$$

and

$$|\alpha_i - b_i|_p \le |\alpha_i - b_j|_p,$$

for $1 \le i \le t_1$ and $1 \le j \le r$ we deduce that

$$|\alpha_i - b_j|_p = \max(|\alpha_i - b_i|_p, |b_j - b_i|_p).$$

Since $\delta_{i,i}$ is an integer and $\operatorname{ord}_p(b_i - b_j)$ is also an integer, we conclude that $\delta_{i,j}$ is an integer for all pairs (i, j) with $1 \le i \le t_1$ and $1 \le j \le r$. Therefore, by (35), l is an even integer. Accordingly, when l is odd inequality (34) is not sharp and so we may replace l in (34) by $l - 1$. Therefore we have

$$u_1 + \cdots + u_t \le 2\left[\frac{l}{2}\right].$$

Further since $t \le r$ and $T \le [((r-1)/r)k]$ we also have

$$u_1 + \cdots + u_t \le r\left[\left(\frac{r-1}{r}\right)k\right].$$

Next consider $f(p^{k-u_i}x + b_i)$ for $i = 1, \dots, t$. If $t_1 + 1 \le i \le t$ then $u_i = 0$ and since $f(b_i) \equiv 0 \pmod{p^k}$ it is immediate that p^k divides the content of f. Suppose therefore that $1 \le i \le t_1$, hence that $u_i > 0$. We have $|p^{-k} f(p^{k-u_i}x + b_i)|_p \le 1$ for all x in $\mathbb{Z}_p$ and it suffices to prove that $|p^{-k} f(p^{k-u_i}x + b_i)|_p \le 1$ for all x in R_p. We have

$$|f(p^{k-u_i}x + b_i)|_p = |a_r|_p \prod_{j=1}^{r} |p^{k-u_i}x + b_i - \alpha_j|_p .$$

Recall that $|b_i - \alpha_j|_p \ge p^{-(k-u_i)}$ for $j = 1, \dots, r$. Thus for x in R_p, $|p^{k-u_i}x + b_i - \alpha_j|_p \le |b_i - \alpha_j|_p$ for $j = 1, \dots, r$ and so for all x in R_p, $|p^{-k} f(p^{k-u_i}x + b_i)|_p \le |p^{-k} f(b_i)|_p \le 1$, as required.

Next we take $k = l+1$ in (11) and apply the argument of Sándor [33] to lift the $t = t(l+1)$ congruences (12). This then gives that for $k > l$, $t(k) = t(l+1)$ and $u_i(k) = u_i(l+1)$ for $i = 1, \dots, t$ and also yields (13).

Finally, observe that if $p|b_i$ for some i with $1 \le i \le t$ then $b_i \in A\mathbb{Z}_p + B$ for one of the discs with $|B|_p \le p^{-1}$. Therefore the maximal disc $aR_p + b$ for which $aR_p + b \cap \mathbb{Z}_p = A\mathbb{Z}_p + B$ satisfies $|b|_p < 1$. Since $aR_p + b$ is maximal it contains a root α of f and since $|a|_p \le |b|_p$, $|\alpha|_p < 1$. Therefore at most s_1 of the integers $b_1, \dots, b_t$ are divisible by p. This complete the proof.

That we may take $t(k)$ and $u_i(k)$ for $i = 1, \dots, t$ to be constant for $k > l$ follows from an argument of Sándor [33] as does the fact that the condition $k > l$ cannot be weakened. It follows from Lemma 7 of Bombieri and Schmidt [5] that p^k divides the content of $f(p^{k-u_i} + b_i)$ for $i = 1, \dots, t$. However, they do not give an estimate for u_i. Indeed, the main novelty in the statement of Theorem 2 lies in the estimates (14) and (15). These estimates are, in general, best possible.

We shall show first that estimate (14) is best possible in the following sense. Let θ and ε be positive real numbers, r an integer with $r \ge 2$, and p a prime number larger than r. Then there exist positive integers k and l with $(\theta - \varepsilon)l \le k \le (\theta + \varepsilon)l$ and there exists a polynomial f of degree r and discriminant D with $l = \mathrm{ord}_p D$ and for which the solutions of (11) are given by t congruences (12) with

$$\max_{1 \le i \le t} \{u_i(k)\} = T .$$

Note that since $p > r$ and $t \le r$ the congruences (12) are uniquely determined.

Let $r, t, m,$ and n be integers with $r \ge t \ge 2$, $m > 0$, and $m \ge n \ge 0$, and let p be a prime number with $p > r$. We define $f(x)$ by

$$f(x) = (x + p^m)(x + 2p^m) \cdots (x + tp^m) \cdot (x + (t+1)p^n)(x + t + 2) \cdots (x + r) . \tag{36}$$

Let D denote the discriminant of f. Then $l = \mathrm{ord}_p D = t(t-1)m + 2tn$.

Let s be a positive integer with $s \le m$ and take $t = r$ and $n = 0$ in (36). Then $l = r(r-1)m$. Put $k = rs$. The complete solution of $f(x) \equiv 0 \pmod{p^k}$

is given by $x \equiv 0 \pmod{p^s}$. In this case $u_1 = k - s = ((r-1)/r)k$. Since $1 \le s \le m$, $rm = l/(r-1)$, and m is at our disposal, we see that the upper bound for $u_i(k)$ in (14) of $T = [((r-1)/r)k]$ is best possible for the range $1 \le k \le l/(r-1)$. Next let j be an integer with $1 \le j \le r-2$ and take $t = j+1$ in (36). Then $l = j(j+1)m + 2(j+1)n$. Put $k = (j+1)m + n$ and $v = k - (j+1)n$. The complete solution of $f(x) \equiv 0 \pmod{p^k}$ is given by $x \equiv 0 \pmod{p^m}$, or $x \equiv -(j+2)p^n \pmod{p^v}$, or $x \equiv -i \pmod{p^k}$ for $i = j+3, \dots, r$. In this case $u_1 = k - m = jm + n$, hence $u_1 = l/((j+1)(j+2)) + (j/(j+2))k$. Note that as n varies from 0 to m, k varies from l/j to $l/(j+1)$. Thus the upper bound for $u_i(k)$ in (14) is best possible for k with $l/j \ge k \ge l/(j+1)$ for $j = 1, \dots, r-2$. Finally, take $t = 2$ and $n = 0$ in (36). Let s be an integer larger than m and put $k = s + m = l + (s - m)$. Then the complete solution of $f(x) \equiv 0 \pmod{p^k}$ is given by $x \equiv -p^m \pmod{p^s}$, or $x \equiv -2p^m \pmod{p^s}$, or $x \equiv -i \pmod{p^k}$ for $i = 3, \dots, r$, and so $u_1 = u_2 = m = l/2$, whence the upper bound for $u_i(k)$ in (14) is best possible for the range $k \ge l$ and therefore for the range $k \ge 1$. Further we note that

$$u_1 + u_2 = l \tag{37}$$

and that the number of solutions modulo p^k of $f(x) \equiv 0 \pmod{p^k}$ is $2p^m + r - 2$ or equivalently

$$2p^{l/2} + r - 2. \tag{38}$$

We shall now show that estimate (15) is best possible for $k \ge l/(r-1)$. For the range $k \ge l$ it suffices to recall (37). Next, as before, let j be an integer with $1 \le j \le r-2$ and take $t = j+1$ in (36) so that $l = j(j+1)m + 2(j+1)n$. Put $k = (j+1)m + n + 1$ and $v = k - (j+1)n$. The complete solution of $f(x) \equiv 0 \pmod{p^k}$ is given by $x \equiv -ip^m \pmod{p^{m+1}}$ for $i = 1, \dots, j+1$, or $x \equiv -(j+2)p^n \pmod{p^v}$, or $x \equiv -i \pmod{p^k}$ for $i = j+3, \dots, r$. In this case

$$u_1 + \cdots + u_r = j(j+1)m + 2(j+1)n = l,$$

and if n is positive then $l/j > k > l/(j+1)$. Further for n positive the number of solutions modulo p^k of $f(x) \equiv 0 \pmod{p^k}$ is

$$\begin{aligned}(j+1)p^{jm+n} + p^{(j+1)n} + r - j - 2\\ = (j+1)p^T + p^{l-(j+1)T} + r - j - 2.\end{aligned} \tag{39}$$

Since m and n are still free to be chosen we see that estimate (15) is best possible for $l/j \ge k \ge l/(j+1)$ for $j = 1, \dots, r-2$ and so for the range $k \ge l/(r-1)$. For k satisfying $1 \le k < l/(r-1)$ the minimum of $2[l/2]$ and $r[((r-1)/r)k]$ is $r[((r-1)/r)k]$. In this case estimate (15) is close to optimal as the following example shows. Let r, w, and m be positive integers with $w \ge m+1$ and $r \ge 2$. Let p be a prime number with $p > r$. We define $g(x)$ by

$$g(x) = (x + p^w)(x + 2p^w)(x + 3p^m) \cdots (x + rp^m).$$

Let D denote the discriminant of g. Then $l = \operatorname{ord}_p D = 2(w-m)+r(r-1)m$. Put $k = 1+rm$. The complete solution of $g(x) \equiv 0 \pmod{p^k}$ is given by $x \equiv 0 \pmod{p^{m+1}}$ or $x \equiv -ip^m \pmod{p^{m+1}}$ for $i = 3, \ldots, r$. Thus $t = r-1$,

$$u_1 + \cdots + u_t = (r-1)^2 m = (r-1)\left[\left(\frac{r-1}{r}\right)k\right],$$

and the number of solutions modulo p^k of $g(x) \equiv 0 \pmod{p^k}$ is

$$(r-1)p^{[((r-1)/r)k]}. \tag{40}$$

For any prime p and nonzero integers r, k, and D with $r \geq 2$ and $k > 0$ we define $Q = Q(r, k, p, D)$ and $B = B(r, k, p, D)$ in the following way. We put $(Q, B) = (r, 0)$ except when $T \neq 0$ and $2[l/2]/T \leq r$, in which case we put $(Q, B) = (Q_1, B_1)$ where

$$2\left[\frac{l}{2}\right] = Q_1 T + B_1,$$

with $0 \leq Q_1$ and $0 \leq B_1 < T$. Now observe that the single congruence $x \equiv b_i \pmod{p^{k-u_i}}$ is equivalent to the p^{u_i} congruences $x \equiv a_j \pmod{p^k}$, where $a_j = b_i + jp^{k-u_i}$ for $j = 1, \ldots, p^{u_i}$. Thus the number of solutions modulo p^k of (11) is $p^{u_1} + \cdots + p^{u_t}$. Since for any positive integers u, v with $u \geq v$ we have

$$p^{u+1} + p^{v-1} > p^u + p^v,$$

it follows from (14) and (15) that

$$p^{u_1} + \cdots + p^{u_t} \leq Qp^T + p^B + r - Q - 1.$$

Since $T \leq [l/2]$ we see that

$$Qp^T + p^B + r - Q - 1 \leq 2p^{[l/2]} + r - 2.$$

Therefore we have proved the following result.

Corollary 2. *Let p be a prime number, k a positive integer, and f a polynomial with integer coefficients, content coprime with p, degree r (≥ 2), and nonzero discriminant D. The number of solutions modulo p^k of*

$$f(x) \equiv 0 \pmod{p^k} \tag{41}$$

is at most

$$Qp^T + p^B + r - Q - 1, \tag{42}$$

which in particular is at most

$$2p^{[l/2]} + r - 2. \tag{43}$$

In 1921 Nagell [30] and Ore [32] proved independently that the number of solutions modulo p^k of (41) is at most rp^{2l}. This was improved by Sándor

[33] in 1952 to $rp^{l/2}$ for $k > l$ and in 1981 Huxley [19] obtained the same bound for all positive integers k. Estimates (42) and (43) coincide when $k \geq l$ and, by (38), they are best possible for this range.

If $k < l$ then we may have $T < [l/2]$ in which case $Qp^T + p^B + r - Q - 1$ is smaller than $2p^{[l/2]} + r - 2$. It follows from (39) that (42) is best possible for the range $l/j \geq k \geq l/(j+1)$ for $j = 1, \ldots, r-2$. Finally since $T \leq [((r-1)/r)k]$ for all positive integers k we have

$$Qp^T + p^B + r - Q - 1 \leq rp^{[((r-1)/r)k]}. \tag{44}$$

Thus, by virtue of (40), estimate (42) is close to best possible for the range $1 \leq k \leq l/(r-1)$.

By Theorem 2 and the Chinese Remainder Theorem we obtain the following result.

Corollary 3. *Let m be a positive integer, and let f be a polynomial with integer coefficients, content coprime with m, degree r (≥ 2), and nonzero discriminant D. There is an integer t with $0 \leq t \leq r^{\omega(m)}$, nonnegative integers $b_1, \ldots, b_t$, and positive integers $d_1, \ldots, d_t$ satisfying*

$$\operatorname{ord}_p d_i \leq T(r, \operatorname{ord}_p m, p, D)$$

for $i = 1, \ldots, t$ and all prime numbers p, such that the complete solution of

$$f(x) \equiv 0 \pmod{m} \tag{45}$$

is given by the t mutually disjoint congruences $x \equiv b_i \pmod{m/d_i}$ for $i = 1, \ldots, t$.

By Corollary 2 and the Chinese Remainder Theorem the number of solutions modulo m of (45) is at most the product over all primes p dividing m of the upper bound given by (42) with $k = \operatorname{ord}_p m$. In particular, by (43) and (44), we see that the number of solutions modulo m of (45) is at most

$$\prod_{p \mid m} \min(2p^{[\operatorname{ord}_p D/2]} + r - 2, rp^{[((r-1)/r)\operatorname{ord}_p m]}). \tag{46}$$

We remark that the upper bound in (46) is again sharp. Let w, r, and $j_1, \ldots, j_w$ be positive integers with $r \geq 2$, and let $p_1, \ldots, p_w$ be prime numbers larger than r. Put

$$h(x) = x(x + p_1^{j_1} \cdots p_w^{j_w})(x+1) \cdots (x + r - 2), \tag{47}$$

let v be an integer with $v \geq w$, and let $p_{w+1}, \ldots, p_v$ be primes that are larger than $p_1^{j_1} \cdots p_w^{j_w}$. Finally let $k_1, \ldots, k_v$ be positive integers with $k_i > 2j_i$ for $i = 1, \ldots, w$ and put $m = p_1^{k_1} \cdots p_v^{k_v}$. Then by the Chinese Remainder Theorem and the discussion preceding (38), it follows that the number of solutions of (45) with h as in (47) is exactly

$$\prod_{p \mid m} (2p^{[\operatorname{ord}_p D/2]} + r - 2).$$

4. Preliminary lemmas

Let α be an algebraic number of degree n and define the height of α, denoted by $h(\alpha)$, by

$$h(\alpha) = (M(f))^{1/n},$$

where f is the minimal polynomial of α over the integers. Let t and τ be positive numbers such that $t < \sqrt{2/n}$ and $\sqrt{2-nt^2} < \tau < t$, and put $\lambda = 2/(t-\tau)$ and

$$A_1 = \frac{t^2}{2-nt^2}\left(n\log(h(\alpha)) + \frac{n}{2}\right).$$

Suppose that $\lambda < n$. A rational number x/y is said to be a very good approximation to α if

$$|\alpha - x/y| < (4e^{A_1}H(x, y))^{-\lambda},$$

where $H(x, y) = \max(|x|, |y|)$. Bombieri and Schmidt [5], building on the earlier work of Bombieri [2] and Bombieri and Mueller [4], and of course the classical work of Thue and Siegel, proved the following result.

Thue-Siegel principle. *If α is of degree n (≥ 3) and x/y and x'/y' are two very good approximations to α then*

$$\log(4e^{A_1}) + \log(H(x', y')) \leq \gamma^{-1}(\log(4e^{A_1}) + \log(H(x, y))),$$

where $\gamma = (nt^2 + \tau^2 - 2)/(n-1)$.

We must also deal with the possibility that α is of degree 1 or 2. In this case we appeal to the following simple result.

Lemma 1. *Let α be an algebraic number with minimal polynomial f over $\mathbb{Q}$, and let a be the leading coefficient of f and D the discriminant of f. Suppose that p/q is a rational number with $q \neq 0$. If α is a rational and $\alpha \neq p/q$ then*

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{1}{|aq|} \geq \frac{1}{M(f)|q|}.$$

If α is of degree 2 over $\mathbb{Q}$ then

$$\left|\alpha - \frac{p}{q}\right| \geq \min(M(f)^{-1}, (2|D|^{1/2}q^2)^{-1}).$$

Proof. First assume that α is of degree 1. Then $f(x) = ax - b$ with a and b coprime integers and with $a \neq 0$. Thus $\alpha = b/a$ and for any rational number p/q with $q \neq 0$ and $\alpha \neq p/q$,

$$\left|\alpha - \frac{p}{q}\right| = \left|\frac{b}{a} - \frac{p}{q}\right| \geq \frac{1}{|aq|}.$$

Since $M(f) \geq |a|$ the result holds.

Next assume that α is of degree 2. Then $f(x) = ax^2 + bx + c$ with $a \neq 0$. Let α' denote the other root of f. For any rational number p/q with $q \neq 0$,

$$\left|\alpha - \frac{p}{q}\right| = \frac{|ap^2 + bpq + cq^2|}{q^2|a|\,|\alpha' - p/q|} \geq \frac{1}{q^2|a|\,|\alpha' - p/q|}. \tag{48}$$

Now either

$$\left|\alpha - \frac{p}{q}\right| \geq |\alpha - \alpha'| = \frac{|D|^{1/2}}{|a|} \geq \frac{1}{|a|} \geq \frac{1}{M(f)}, \tag{49}$$

or

$$\left|\alpha' - \frac{p}{q}\right| = \left|(\alpha' - \alpha) + \left(\alpha - \frac{p}{q}\right)\right| \leq 2|\alpha' - \alpha| \leq 2\frac{|D|^{1/2}}{|a|}.$$

In the latter case, by (48),

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{1}{2|D|^{1/2}q^2}, \tag{50}$$

and so the result follows from (49) and (50).

From the proof of Lemma 1 of Lewis and Mahler [21] together with refinements (a) and (b) of Bombieri and Schmidt [5, p. 72] we obtain the following result.

Lemma 2. *Let f be a polynomial with coefficients from the complex numbers $\mathbb{C}$, degree n (≥ 2), and zeros $\alpha_1, \ldots, \alpha_n$ in $\mathbb{C}$. For every z in $\mathbb{C}$,*

$$|f(z)| \geq \frac{|D(f)|^{1/2}}{n^{(n-1)/2}2^{n-1}M(f)^{n-2}} \min_{1\leq i\leq n} |z - \alpha_i|.$$

We may apply Lemma 2 to obtain the following version of Lemma 1 of [5].

Lemma 3. *Let F be a binary form of degree r (≥ 3) with integer coefficients and nonzero discriminant $D(F)$. For every pair of integers (x, y) with $y \neq 0$*

$$\min_{\alpha} \left|\alpha - \frac{x}{y}\right| \leq \frac{2^{r-1}r^{(r-1)/2}(M(F))^{r-2}|F(x, y)|}{|D(F)|^{1/2}|y|^r},$$

where the minimum is taken over the zeros α of $F(z, 1)$.

Proof. Put $f(z) = F(z, 1)$ and denote the degree of f by n. Since F has degree at least 3 and $D(F) \neq 0$, n is at least 2 and so by Lemma 2

$$|F(x, y)|\,|y|^{-r} = \left|f\left(\frac{x}{y}\right)\right| \geq \frac{|D(f)|^{1/2}}{n^{(n-1)/2}2^{n-1}(M(f))^{n-2}} \min_{\alpha} \left|\alpha - \frac{x}{y}\right|.$$

Since F is homogeneous, $M(f) = M(F)$. Let

$$F(x, y) = a_r x^r + a_{r-1}x^{r-1}y + \cdots + a_0 y^r.$$

If $a_r \neq 0$ then $n = r$, $D(F) = D(f)$, and the result follows immediately. If $a_r = 0$ then $a_{r-1} \neq 0$, $n = r - 1$, and $|D(F)|^{1/2} = |a_{r-1}|\,|D(f)|^{1/2}$. But then $M(f) \geq |a_{r-1}|$ and the result again follows.

5. Proof of Theorem 1

Let p be a prime and suppose that p^k exactly divides h. If (x, y) is a primitive solution of (1) then certainly

$$F(x, y) \equiv 0 \pmod{p^k}.$$

If p does not divide y then y is invertible modulo p^k and so

$$F(xy^{-1}, 1) \equiv 0 \pmod{p^k}.$$

By Theorem 2 there is an integer t with $0 \le t \le s$, where s denotes the number of zeros of $F(z, 1)$ in R_p, and there are integers $b_1, \dots, b_t$ and $u_1, \dots, u_t$ with u_i satisfying (14) such that $xy^{-1} \equiv b_i \pmod{p^{k-u_i}}$ for some integer i with $1 \le i \le t$. We suppose, as we may without loss of generality, that for each integer j with $1 \le j \le t$ there is a primitive solution (x, y) of (1) for which $xy^{-1} \equiv b_j \pmod{p^{k-u_j}}$. Note also that since $D(F(z, 1))$ divides $D(F)$ $(= D(F(X, Y)))$ we may take $D(F)$ in place of $D(F(z, 1))$ in estimate (14).

Put $F_i(X, Y) = F(p^{k-u_i}X + b_iY, Y)$ for $i = 1, \dots, t$. By Theorem 2 the content of F_i is divisible by p^k. Since F has content 1 the content of F_i is a power of p and since p^k exactly divides h and there is a primitive solution (x, y) of (1) for which $xy^{-1} \equiv b_i \pmod{p^{k-u_i}}$ the content of F_i is p^k for $i = 1, \dots, t$. Put $\widetilde{F}_i(X, Y) = p^{-k}F_i(X, Y)$ for $i = 1, \dots, t$. Plainly $\widetilde{F}_i$ has content 1 and by (5) and (6)

$$D(\widetilde{F}_i) = \frac{p^{(k-u_i)r(r-1)}D(F)}{p^{k2(r-1)}} \tag{51}$$

for $i = 1, \dots, t$. Further since $xy^{-1} \equiv b_i \pmod{p^{k-u_i}}$, there exists an integer x_i for which $x = p^{k-u_i}x_i + b_iy$. Thus (x_i, y) is a primitive solution of $\widetilde{F}_i(X, Y) = hp^{-k}$.

Similarly if (x, y) is a primitive solution of (1) and p divides y then p does not divide x and so x is invertible modulo p^k. In this case $F(1, yx^{-1}) \equiv 0 \pmod{p^k}$. By Theorem 2 applied to $F(1, z)$ there is an integer w with $t \le w$ and there are integers $b_{t+1}, \dots, b_w$ and $u_{t+1}, \dots, u_w$, with u_i satisfying (14) with $D(F)$ in place of $D(F(1, z))$, such that $yx^{-1} \equiv b_i \pmod{p^{k-u_i}}$ for some integer i with $t+1 \le i \le w$. We choose w to be minimal. Since p divides y it also divides b_i for $i = t+1, \dots, w$ and thus by Theorem 2, $w - t$ is at most s_1, where s_1 is the number of roots α of $F(1, z)$ with $|\alpha|_p < 1$. Since each nonzero root of $F(1, z)$ is the inverse of a nonzero root of $F(z, 1)$, $w \le r$. Arguing as before, but with the roles of x and y reversed, we determine binary forms $\widetilde{F}_i$ of content 1 that satisfy (51) for $i = t+1, \dots, w$.

Therefore if (x, y) is a primitive solution of (1) then it determines a triple (i, x', y'), where $1 \le i \le w$ and (x', y') is a pair of coprime integers for which $\widetilde{F}_i(x', y') = hp^{-k}$. Further, distinct primitive solutions of (1) determine distinct triples. We may assume, without loss of generality, that $\operatorname{ord}_p g = \operatorname{ord}_p h$ for all primes p that divide g. Then, by repeating the above construction for each prime p that divides g we obtain a set W of at most $r^{\omega(g)}$ binary forms with the property that distinct primitive solutions (x, y) of (1) correspond to distinct triples $(\widehat{F}, x', y')$, where $\widehat{F}$ is in W and (x', y') is a pair of coprime integers for which $\widehat{F}(x', y') = h/g$. Further if $\widehat{F}$ is in W then $\widehat{F}$ has content 1 and, by (51) and Theorem 2,

$$|D(\widehat{F})| \geq \frac{g^{(r-2)(r-1)}|D(F)|}{G(g, r, D)^{r(r-1)}},$$

hence, by (8),

$$(52) \qquad |D(\widehat{F})| \geq \left(\left|\frac{h}{g}\right|^{2/r+\varepsilon}\right)^{r(r-1)}.$$

Let $\widehat{F}$ be a form in W, let p a prime number, and let

$$B_0 = \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}, \qquad B_j = \begin{pmatrix} 0 & -1 \\ p & j \end{pmatrix}$$

for $j = 1, \dots, p$. Then, as in [5], we have

$$\mathbb{Z}^2 = \bigcup_{j=0}^{p} B_j \mathbb{Z}^2,$$

and the number of primitive solutions of $\widehat{F}(x, y) = h/g$ is at most $n_0 + n_1 + \cdots + n_p$, where n_j is the number of primitive solutions of

$$\widehat{F}_{B_j}(x, y) = h/g.$$

By (6),

$$(53) \qquad |D(\widehat{F}_{B_j})| = p^{r(r-1)}|D(\widehat{F})|.$$

Put $n = h/g$ and take $p = 41$ in (53). Then, by (52) and (53), the number of primitive solutions of (1) is at most $42r^{\omega(g)}$ times the maximum number of primitive solutions (x, y) of

$$(54) \qquad G(x, y) = n$$

for all binary forms G of degree r and for which

$$(55) \qquad |D(G)| \geq (41|n|^{2/r+\varepsilon})^{r(r-1)}.$$

Suppose that G is such a form and that (x_0, y_0) is a primitive solution of (54). Then there is an A in $\mathrm{GL}(2, \mathbb{Z})$ for which $A^{-1}(x_0, y_0)$ is $(1, 0)$ and so $(1, 0)$ is a solution of $G_A(x, y) = n$. Note that G_A has leading coefficient n. Thus we may suppose that G has leading coefficient n and that $M(G)$ is smallest among all equivalent forms that have n as their leading coefficient.

Let Y_0 be a positive real number. We shall now estimate the primitive solutions (x, y) of (54) for which $0 < y \leq Y_0$, and here we shall repeat the argument of Bombieri and Schmidt with some minor changes. We have

$$G(x, y) = n(x - \alpha_1 y) \cdots (x - \alpha_r y),$$

where $\alpha_1, \dots, \alpha_r$ are distinct complex numbers. Put $L_i(x, y) = x - \alpha_i y$ for $i = 1, \dots, r$. Then by the same argument given for the proof of Lemma 3 of [5] we obtain the next result.

Lemma 4. *Suppose* (x, y) *and* (x_0, y_0) *are primitive solutions of* (54). *Then for* $1 \le i, j \le r$,

$$\frac{L_i(x_0, y_0)}{L_i(x, y)} - \frac{L_j(x_0, y_0)}{L_j(x, y)} = (\beta_i - \beta_j)(xy_0 - x_0 y), \tag{56}$$

where $\beta_1, \dots, \beta_r$ *depend on* (x, y) *and are such that the form*

$$J(u, w) = n(u - \beta_1 w)\cdots(u - \beta_r w)$$

is equivalent to G.

We may take $(x_0, y_0) = (1, 0)$ in which case, by (56),

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

For every primitive solution (x, y) of (54) we choose $j = j(x, y)$ with $|L_j(x, y)| \ge 1$. Then

$$\frac{1}{|L_i(x, y)|} \ge |\beta_j - \beta_i|\,|y| - 1. \tag{57}$$

Since $|\overline{L_j(x, y)}| \ge 1$, (57) holds with $\overline{\beta_j}$ in place of β_j and so

$$\frac{1}{|L_i(x, y)|} \ge |\operatorname{Re}(\beta_j) - \beta_i|\,|y| - 1.$$

We now choose an integer $m = m(x, y)$ with $|m - \operatorname{Re}(\beta_j)| \le 1/2$ and we obtain

$$\frac{1}{|L_i(x, y)|} \ge \left(|m - \beta_i| - \frac{1}{2}\right)|y| - 1 \tag{58}$$

for $i = 1, \dots, r$.

For $1 \le i \le r$, let X_i be the set of primitive solutions of (54) with $1 \le y \le Y_0$ and $|L_i(x, y)| \le 1/2y$.

Lemma 5. *Suppose* $(x, y) \ne (x', y')$ *are in* X_i *with* $y \le y'$. *Then*

$$\frac{y'}{y} \ge \frac{2}{7}\max(1, |\beta_i - m|),$$

where $\beta_i = \beta_i(x, y)$ *and* $m = m(x, y)$.

Proof. This is Lemma 4 of [5] and the proof goes through unchanged.

Similarly we obtain the following version of Lemma 5 of [5].

Lemma 6. *Suppose* (x, y) *is a primitive solution of* (54) *with* $y > 0$ *and* $|L_i(x, y)| > 1/2y$. *Then*

$$|m - \beta_i| \le \tfrac{7}{2},$$

where again $\beta_i = \beta_i(x, y)$ *and* $m = m(x, y)$.

For each set X_i that is not empty let $(x^{(i)}, y^{(i)})$ be the element with the largest value of y. Let X be the set of solutions of (54) with $1 \le y \le Y_0$ minus the elements $(x^{(1)}, y^{(1)}), \dots, (x^{(r)}, y^{(r)})$.

Let i be an integer with $1 \leq i \leq r$ and, when X_i is nonempty, let $(x_1^{(i)}, y_1^{(i)}), \dots, (x_\nu^{(i)}, y_\nu^{(i)})$ be the elements of X_i with $y_1^{(i)} \leq \cdots \leq y_\nu^{(i)}$. Thus $(x_\nu^{(i)}, y_\nu^{(i)}) = (x^{(i)}, y^{(i)})$. By Lemma 5

$$\frac{2}{7}\max(1, |\beta_i(x_k^{(i)}, y_k^{(i)}) - m(x_k^{(i)}, y_k^{(i)})|) \leq \frac{y_{k+1}^{(i)}}{y_k^{(i)}}$$

for $k = 1, \dots, \nu - 1$, hence

$$\prod_{(x, y)\in X\cap X_i} (\tfrac{2}{7}\max(1, |\beta_i(x, y) - m(x, y)|)) \leq Y_0.$$

For (x, y) in X but not in X_i we have

$$\tfrac{2}{7}\max(1, |\beta_i(x, y) - m(x, y)|) \leq 1$$

by Lemma 6. Thus

$$\prod_{(x, y)\in X} (\tfrac{2}{7}\max(1, |\beta_i(x, y) - m(x, y)|)) \leq Y_0. \tag{59}$$

By Lemma 4 the form

$$J(u, w) = n\prod_{i=1}^{r}(u - \beta_i w)$$

is equivalent to G and thus so also is the form

$$\widehat{J}(u, w) = n\prod_{i=1}^{r}(u - (\beta_i - m)w).$$

Therefore

$$\prod_{i=1}^{r}\max(1, |\beta_i(x, y) - m(x, y)|) = \frac{M(\widehat{J})}{|n|} \geq \frac{M(G)}{|n|}.$$

Taking the product of (59) for $i = 1, \dots, r$ we find that

$$\left(\left(\frac{2}{7}\right)^r \frac{M(G)}{|n|}\right)^{|X|} \leq Y_0^r; \tag{60}$$

here $|X|$ denotes the cardinality of X. By a result of Mahler [25],

$$M(G) \geq \left(\frac{|D(G)|}{r^r}\right)^{1/(2r-2)},$$

and thus, by (55),

$$M(G) \geq \frac{(41|n|^{2/r+\varepsilon})^{r/2}}{r^{r/(2r-2)}}.$$

Since $r^{1/(2r-2)} \leq 3^{1/4}$ we find that

$$M(G) \geq \left(\frac{41^{1/2}}{3^{1/4}}\right)^r |n|^{1+\varepsilon r/2}. \tag{61}$$

For any positive real numbers a, b, c, d we have $(a+b)/(c+d) \le \max(a/c, b/d)$ and thus

$$\frac{\log|n| + r\log(7/2)}{(1+(\varepsilon r/2))\log|n| + r\log(41^{1/2}/3^{1/4})} \le \max\left(\frac{2}{2+\varepsilon r}, \frac{\log(7/2)}{\log(41^{1/2}/3^{1/4})}\right). \tag{62}$$

Therefore, by (61) and (62),

$$\left(\frac{2}{7}\right)^r \frac{M(G)}{|n|} \ge M(G)^\theta, \tag{63}$$

where $\theta = \min(\varepsilon r/(2+\varepsilon r), \theta_1)$ and $\theta_1 = 1 - (\log(7/2))/(\log(41^{1/2}/3^{1/4}))$. Accordingly, by (60) and (63),

$$|X| \le \frac{r\log Y_0}{\theta \log M(G)}.$$

We now take $Y_0 = M(G)^2$ so that $|X| \le 2r/\theta$. Thus the number of primitive solutions (x, y) of (54) with $1 \le y \le M(G)^2$ is at most $(2r/\theta) + r$ and therefore, since $|D(G)| = |D(-G)|$ and $M(G) = M(-G)$, the number with $|y| \le M(G)^2$ is at most $2((2r/\theta) + r + 1)$.

We shall now estimate the number of primitive solutions (x, y) of (54) with $|y| \ge M(G)^2$. To each such primitive solution (x, y) we associate a root α_i of $G(x, 1)$ for which

$$\left|\alpha_i - \frac{x}{y}\right| \le \left|\alpha_j - \frac{x}{y}\right|$$

for $j = 1, \dots, r$. For $i = 1, \dots, r$ let $I^{(i)}$ denote the set of such solutions associated to α_i.

Now fix i and let $(x_1, y_1), (x_2, y_2), \dots$ denote the elements of $I^{(i)}$ with $y_j > 0$ for $j = 1, 2, \dots$, ordered so that $y_1 \le y_2 \le \cdots$. By Lemma 3,

$$\left|\alpha_i - \frac{x_j}{y_j}\right| \le \frac{2^{r-1} r^{(r-1)/2} M(G)^{r-2} |n|}{|D(G)|^{1/2} y_j^r} \tag{64}$$

for $j = 1, 2, \dots$, and so

$$\left|\frac{x_{j+1}}{y_{j+1}} - \frac{x_j}{y_j}\right| \le \frac{2^r r^{(r-1)/2} |n|}{|D(G)|^{1/2}} \frac{M(G)^{r-2}}{y_j^r}.$$

Since $|x_{j+1}y_j - x_j y_{j+1}| \ge 1$, we have, by (55),

$$\frac{y_j^{r-1}}{M(G)^{r-2}} \le y_{j+1}. \tag{65}$$

We define the positive real number δ_j for each integer j for which there exists an element (x_j, y_j) in $I^{(i)}$ by

$$y_j = M(G)^{1+\delta_j}. \tag{66}$$

By (61) $M(G) > 1$ and so, by (65), $(r-1)\delta_j \le \delta_{j+1}$ for $j = 1, 2, \dots$. Note that $\delta \ge 1$ since $y_1 \ge M(G)^2$. Therefore

$$(r-1)^{j-1} \le \delta_j \tag{67}$$

for $j = 1, 2, \dots$, and for any positive integers k and l with (x_{k+l}, y_{k+l}) in $I^{(i)}$,

$$\delta_k (r-1)^l \le \delta_{k+l}. \tag{68}$$

Let f_i denote the minimal polynomial of α_i over the rationals. Then f_i divides $G(x, 1)$ in $\mathbb{Z}[x]$, $M(f_i) \le M(G)$, and $|D(f_i)| \le |D(G)|$. We suppose first that α_i is a rational number. Then, by (55), (64), and Lemma 1,

$$y_j^{r-1} \le M(G)^{r-1},$$

which is impossible since $y_1 \ge M(G)^2$. Next suppose that α_i is of degree 2 over the rationals. Then by (55), (61), (64), and Lemma 1,

$$y_j^{r-2} \le (2^r r^{(r-1)/2}|n|)M(G)^{r-2} < M(G)^{2(r-2)},$$

and again this is impossible since $y_1 \ge M(G)^2$.

Finally suppose that α_i is of degree d over the rationals with d at least 3. We shall apply the Thue-Siegel principle with

$$t = \sqrt{2/(d + a^2)}, \qquad a = .1,$$

and

$$\tau = 1.2\sqrt{2 - dt^2} = 1.2at = .12t.$$

Then $\lambda = 2/(t-\tau) = 2/(.88t)$ and so $\lambda < .93d \le .93r$. Further, $t^2/(2 - dt^2) = a^{-2} = 100$ so

$$A_1 = 100(d\log(h(\alpha_i)) + d/2) = 100(\log(M(f_i)) + d/2)$$

and

$$\gamma = (dt^2 + \tau^2 - 2)/(d-1) = ((.12)^2 - (.01))2/((d + .01)(d-1)),$$

hence

$$\gamma^{-1} < 172(d-1)^2 \le 172(r-1)^2. \tag{69}$$

Now observe that

$$4e^{A_1} \le 4M(f_i)^{100}e^{50d} \le 4M(G)^{100}e^{50r}.$$

By (61), $M(G) \ge (41^{1/2}/3^{1/4})^r$ and so

$$8e^{50r} \le (8^{1/3}e^{50})^r < M(G)^{33}.$$

Therefore

$$8e^{A_1} < M(G)^{133}. \tag{70}$$

Note that by (55) and (64), $|\alpha_i - x_j/y_j| < 1$, hence $|x_j| < |y_j|(|\alpha_i| + 1) \leq 2M(G)y_j$. Thus $H(x_j, y_j) < 2M(G)y_j$ and so

$$(4e^{A_1}H(x_j, y_j))^{\lambda} < M(G)^{(135+\delta_j)\lambda} < M(G)^{(135+\delta_j)(.93r)}. \tag{71}$$

By (55), (64), and (66),

$$|\alpha_i - x_j/y_j| < M(G)^{-\delta_j r}. \tag{72}$$

It follows from (71) and (72) that x_j/y_j is a very good approximation to α_i whenever

$$\delta_j r \geq (135 + \delta_j)(.93r),$$

hence for $\delta_j \geq 1794$. But $\delta_j \geq (r-1)^{j-1}$ and so if we put

$$k = 1 + \left[\frac{\log 1794}{\log(r-1)}\right],$$

then $x_{k+1}/y_{k+1}, x_{k+2}/y_{k+2}, \ldots, x_{k+l}/y_{k+l}$ are all very good approximations to α_i whenever (x_{k+l}, y_{k+l}) is in $I^{(i)}$. Suppose that there exists an integer l with $l \geq 2$ for which (x_{k+l}, y_{k+l}) is in $I^{(i)}$. Then by the Thue-Siegel principle and (69),

$$\log(4e^{A_1}) + \log y_{k+l} \leq 172(r-1)^2(\log(4e^{A_1}) + \log(2M(G)y_{k+1})),$$

and so, by (70),

$$\log y_{k+l} \leq 172(r-1)^2(134 \log M(G) + \log y_{k+1}).$$

Thus, recall (66),

$$\delta_{k+l} \leq 172(r-1)^2(135 + \delta_{k+1}).$$

Since $\delta_{k+1} \geq 1794$, we find that $\delta_{k+l}/\delta_{k+1} \leq 185(r-1)^2$. Thus, by (68), $(r-1)^{l-3} \leq 185$, whence

$$l \leq 3 + \frac{\log 185}{\log(r-1)}.$$

Therefore the number of primitive solutions in $I^{(i)}$ is at most

$$2(4 + \log 331890/\log(r-1))$$

for $i = 1, \ldots, r$. Consequently the number of primitive solutions of (1) is at most

$$42r^{\omega(g)}(2((2r/\theta) + r + 1) + 2r(4 + \log 331890/\log(r-1))).$$

This in turn is at most

$$84r^{1+\omega(g)}\left(\max\left(\frac{2}{\theta_1}, 2 + \frac{4}{\varepsilon r}\right) + 1 + \frac{1}{r} + 4 + \frac{\log 331890}{\log(r-1)}\right)$$
$$\leq \max\left(2800, 2160 + \frac{336}{\varepsilon r}\right) r^{1+\omega(g)} \leq 2800\left(1 + \frac{1}{8\varepsilon r}\right) r^{1+\omega(g)},$$

as required.

6. Lower bounds for the number of solutions of Thue equations

Silverman [37], extending earlier work of Mahler [24] and Chowla [8], has shown that there exist cubic binary forms F, with nonzero discriminant, for which the number of solutions of the Thue equation (1) exceeds $c(\log|h|)^{2/3}$ for infinitely many integers h, where c is some positive constant. However, the solutions constructed are generally not primitive solutions and as we remarked with Erdös and Tijdeman [10] it may be that there exists a number $C_1(r)$, which depends on r only, such that (1) has at most $C_1(r)$ primitive solutions whenever F has nonzero discriminant and degree r at least three. Bombieri and Schmidt [5] showed that we may have at least r distinct primitive solutions of (1). They gave the example

$$F(x, y) = x^r + a(x-y)(2x-y)\cdots(rx-y),$$

where a is a nonzero integer. Then $(1, 1), (1, 2), \ldots, (1, r)$ are primitive solutions of $F(x, y) = 1$. We do not believe that for a fixed form F there are infinitely many integers h for which (1) has this many primitive solutions if r is large. Indeed we conjecture that there exists an absolute constant c_0 such that for any binary form $F \in \mathbb{Z}[x, y]$ with nonzero discriminant and degree at least three there exists a number C, which depends on F, such that if h is an integer larger than C then the Thue equation (1) has at most c_0 solutions in coprime integers x and y. For each binary form F let $\nu(F)$ denote the largest integer k such that (1) has at least k primitive solutions for arbitrarily large integers h; if k does not exist put $\nu(F) = \infty$. Next, for each integer r let $\nu^*(r)$ be the supremum of $\nu(F)$ over those binary forms F with integer coefficients, nonzero discriminant, and degree r. Of course if the above conjecture is valid then $\nu^*(r) \leq c_0$ for $r = 3, 4, \ldots$. In this section we shall prove the following result.

Theorem 3. *We have*

$$\nu^*(3) \geq 18, \qquad \nu^*(4) \geq 16, \qquad \nu^*(5) \geq 6, \tag{73}$$

and

$$\nu^*(6k) \geq 12, \qquad \nu^*(6k+1) \geq 2, \qquad \nu^*(6k+2) \geq 12,$$
$$\nu^*(6k+3) \geq 6, \qquad \nu^*(6k+4) \geq 8, \qquad \nu^*(6k+5) \geq 6$$

for $k = 1, 2, \ldots$.

Thus c_0 is at least 18. To prove Theorem 3 we shall determine various binary forms that are invariant under subgroups of $\mathrm{GL}(2, \mathbb{Z})$. Further, for (73) we shall also make use of parametric solutions of equations of the form $F(u, v) = F(r, s)$.

Let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

be in $\mathrm{GL}(2, \mathbb{Z})$. Recall that if x, y is a coprime solution of $F_A(x, y) = h$ for some integer h, then $(ax+by, cx+dy)$ is a coprime solution of $F(X, Y) = h$. We remark that if F is a form such that $F_A = F$ and (x, y) is a primitive solution of (1) then also $A(x, y) = (ax+by, cx+dy)$, $A^2(x, y)$, $A^3(x, y)$, ... are primitive solutions and so we obtain many primitive solutions of (1). Plainly we may restrict our attention to those elements A of finite order in $\mathrm{GL}(2, \mathbb{Z})$. In fact we shall look for forms F that are invariant under the action of a finite subgroup of $\mathrm{GL}(2, \mathbb{Z})$. Here again we may restrict our attention, this time to equivalence classes of subgroups of $\mathrm{GL}(2, \mathbb{Z})$ under conjugation. For let G be a finite subgroup of $\mathrm{GL}(2, \mathbb{Z})$, and let F be a binary form that is invariant under G, that is, $F_A = F$ for all A in G. Then, for any element T in $\mathrm{GL}(2, \mathbb{Z})$, F_T is invariant under TGT^{-1}. There are in total 13 mutually nonconjugate finite subgroups of $\mathrm{GL}(2, \mathbb{Z})$ and they are given in Table 1 (see p. 179 of [31]).

We shall now determine those homogeneous binary forms of small degrees that are invariant under the above 13 groups.

Plainly every binary form is invariant under C_1 and every form of even

TABLE 1

Group	Generators	Group	Generators
C_1	$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$	D_2	$\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}$
C_2	$\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}$	D_2^*	$\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}$
C_3	$\begin{pmatrix}0 & 1\\ -1 & -1\end{pmatrix}$	D_3	$\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\ -1 & -1\end{pmatrix}$
C_4	$\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}$	D_3^*	$\begin{pmatrix}0 & -1\\ -1 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\ -1 & -1\end{pmatrix}$
C_6	$\begin{pmatrix}0 & -1\\ 1 & 1\end{pmatrix}$	D_4	$\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}$
D_1	$\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$	D_6	$\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\ -1 & 1\end{pmatrix}$
D_1^*	$\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$		

degree is invariant under C_2. Forms F invariant under C_3 satisfy

$$F(x, y) = F(y, -x-y). \tag{74}$$

Thus if F is of degree 3 and we put $F(x, y) = ax^3 + bx^2y + cxy^2 + dy^2$ then, by (74), $d = -a$ and $c = b - 3a$ and so the forms of degree 3 invariant under C_3 are $F(x, y) = ax^3 + bx^2y + (b-3a)xy^2 - ay^3$ with a and b not both zero. If F is invariant under C_4 then $F(x, y) = F(y, -x)$ and so F must be of even degree and symmetric up to alternating signs. Thus, if F is of degree 4 it has the form $F(x, y) = ax^4 + bx^3y + cx^2y^2 - bxy^3 + ay^4$ with a, b, and c not all zero. Next if F is invariant under C_6 then $F(x, y) = F(-y, x+y)$ and F is not of degree 3 or 5 while if it is of degree 4 it has the form $a(x^2 + xy + y^2)^2$. The forms of degree 6 are

$$\begin{aligned} F(x, y) = ax^6 + bx^5y + cx^4y^2 + (2c + 10a - 5b)x^3y^3 \\ + (c + 15a - 5b)x^2y^4 + (6a - b)xy^5 + ay^6 \end{aligned}$$

with a, b, and c not all zero.

F is invariant under D_1 whenever the coefficients attached to odd powers of y are zero, and is invariant under D_1^* whenever F is reciprocal. Further F is invariant under D_2 whenever F is of even degree and the coefficients attached to odd powers of y are zero, while F is invariant under D_2^* whenever F is of even degree and reciprocal. The forms invariant under D_3 are for degree 3, $axy(x+y)$, degree 4, $a(x^2 + xy + y^2)^2$, degree 5, $axy(x+y)(x^2 + xy + y^2)$, all with $a \neq 0$, and for degree 6,

$$F(x, y) = ax^6 + 3ax^5y + cx^4y^2 + (2c - 5a)x^3y^3 + cx^2y^4 + 3axy^5 + ay^6, \tag{75}$$

with a and c not both zero. The forms invariant under D_3^* are for degree 3, $\frac{a}{2}(x+2y)(x-y)(2x+y)$, degree 4, $a(x^2 + xy + y^2)^2$, degree 5, $\frac{a}{2}(x+2y)(x-y)(2x+y)(x^2 + xy + y^2)$, all with $a \neq 0$, and for degree 6 they are of the form (75) with a and c not both zero. The forms invariant under D_4 are reciprocal, of even degree, and the coefficients attached to odd powers of y are zero. The forms of degree 4 are

$$F(x, y) = ax^4 + cx^2y^2 + ay^4,$$

with a and c not both zero. Finally we consider forms F invariant under D_6. Then $F(x, y) = F(y, x) = F(y, -x+y)$. There are no such forms of degree 3 or 5 and the only forms of degree 4 are $a(x^2 - xy + y^2)^2$ with $a \neq 0$. The forms of degree 6 invariant under D_6 are

$$F(x, y) = ax^6 - 3ax^5y + cx^4y^2 + (5a - 2c)x^3y^3 + cx^2y^4 - 3axy^5 + ay^6, \tag{76}$$

with a and c not both zero. For such a form F, if (x, y) is a solution of (1) then so also are $(y, -x+y)$, $(-x+y, -x)$, $(-x, -y)$, $(-y, x-y)$, $(x-y, x)$, (y, x), $(-x+y, y)$, $(-x, -x+y)$, $(-y, -x)$, $(x-y, -y)$, $(x, x-y)$. Further observe that these 12 solutions are distinct and primitive whenever (x, y)

is primitive and (x, y) is different from $(1, 0)$, $(-1, 0)$, $(0, 1)$, $(0, -1)$, $(1, 1)$, $(-1, -1)$, $(1, -1)$, $(-1, 1)$, $(1,2)$, $(-1, -2)$, $(2, 1)$, and $(-2, 1)$.

Proof of Theorem 3. We shall first prove that $\nu^*(3) \geq 18$. Consider $F(x, y) = xy(x+y)$. By the above discussion F is invariant under D_3 and so whenever (x, y) is a primitive solution of (1), $(y, -x-y)$, $(-x-y, x)$, (y, x), $(-x-y, y)$, and $(x, -x-y)$ are also primitive solutions. If x and y are coprime integers and, as is readily checked, (x, y) is not one of $(1, 1)$, $(-1, -1)$, $(1, -2)$, $(-1, 2)$, $(2, -1)$, or $(-2, 1)$ then the orbit of (x, y) under D_3 consists of six distinct pairs.

For integers a and b we define $\varepsilon = \varepsilon(a, b)$ by

$$\varepsilon = \begin{cases} 1 & \text{if } 3 \mid 2a+b, \\ 0 & \text{otherwise,} \end{cases}$$

and we put

$$f(a, b) = -a^4 - 2a^3b + 5a^2b^2 + 6ab^3 + b^4.$$

Next we put

$$x(a, b) = \frac{a(a-b)}{2 \cdot 3^{\varepsilon}}, \qquad y(a, b) = -\frac{(a+2b)(a+b)}{2 \cdot 3^{\varepsilon}},$$

$$u(a, b) = \frac{(2a+b)(a+b)}{2 \cdot 3^{\varepsilon}}, \qquad v(a, b) = \frac{b(a-b)}{2 \cdot 3^{\varepsilon}},$$

$$r(a, b) = \frac{b(2a+b) + \sqrt{f(a, b)}}{2 \cdot 3^{\varepsilon}}, \qquad s(a, b) = \frac{b(2a+b) - \sqrt{f(a, b)}}{2 \cdot 3^{\varepsilon}}.$$

We observe that

$$\begin{aligned} F(x(a, b), y(a, b)) &= F(u(a, b), v(a, b)) = F(r(a, b), s(a, b)) \\ &= \frac{ab(a-b)(a+b)(a+2b)(2a+b)}{4 \cdot 3^{3\varepsilon}}. \end{aligned}$$

We shall prove that if a and b are coprime odd integers for which $f(a, b)$ is the square of an integer then $(x(a, b), y(a, b))$, $(u(a, b), v(a, b))$, and $(r(a, b), s(a, b))$ are pairs of coprime integers. Further we shall show that there is a finite set of pairs such that if (a, b) is not from that set then the orbits of (x, y), (u, v), and (r, s) under D_3 are disjoint. This will then establish that $\nu^*(3) \geq 18$ provided that we prove there are infinitely many pairs of coprime odd integers (a, b) satisfying

$$z^2 = f(a, b) \tag{77}$$

for some integer z, since, as is easily verified, for any pair of integers (k, l) there are only finitely many pairs of coprime integers (a, b) with $(x(a, b), y(a, b))$, $(u(a, b), v(a, b))$, or $(r(a, b), s(a, b))$ equal to (k, l).

Put $f(w) = f(1, w)$. Corresponding to the curve $Z^2 = f(w)$ is the curve $t^2 = 4s^3 - g_2 s - g_3$, where g_2 and g_3 are the invariants of the quartic f (see, for instance, Chapter 16 of Mordell [27]). In particular, the map from the set of rational points (s, t) on the curve

$$t^2 = 4s^3 - \frac{49}{12}s + \frac{143}{216}, \tag{78}$$

minus the points $(17/12, \pm 5/2)$, to the set of rational points (w, Z) on the curve $Z^2 = f(w)$ given by

$$w = \frac{t - 5/2}{2s - 17/6} - \frac{3}{2} \quad \text{and} \quad Z = -\left(w + \frac{3}{2}\right)^2 + 2s + \frac{17}{12}, \tag{79}$$

is injective. It follows that there are infinitely many rational solutions (Z, w) of $Z^2 = f(w)$ whenever there are infinitely many rational solutions (s, t) of (78), or, equivalently, infinitely many rational solutions (S, T) of

$$T^2 = S^3 - 1323S + 7722. \tag{80}$$

Observe that $P = (1057/16, 29233/64)$ is a point on (80) ($P = 4P_1$, where $P_1 = (-21, 162)$), so that by the theorem of Lutz and Nagell (see Corollary 7.2 of [38]), P is a point of infinite order in the group of rational points of the elliptic curve given by (80). This shows that there are infinitely many rational solutions (Z, w) of $Z^2 = f(w)$. For each solution we write $w = b/a$, where a and b are coprime integers and then clear denominators by multiplying through by a^4 to give a solution (a^2Z, a, b) of (77) with a and b coprime integers. Thus there exist infinitely many pairs of coprime integers (a, b) satisfying (77) and it remains to check that infinitely many of these pairs have a and b odd. First note that if a and b are coprime integers that give a solution of (77) and a is odd and b is even then

$$z^2 \equiv -a^4 \equiv -1 \pmod 4,$$

which is impossible. On the other hand, if a and b are coprime integers that give a solution of (77) and a is even and b is odd then, since

$$f(a, b) = f(-a - b, b),$$

$-a-b$ and b are coprime odd integers that give a solution of (77). Thus there are infinitely many pairs (a, b) of coprime odd integers that give a solution of (77).

We shall assume for the balance of the proof that a and b are coprime odd integers for which $f(a, b)$ is the square of an integer. We first check that then $x(a, b)$, $y(a, b)$, $u(a, b)$, $v(a, b)$, $r(a, b)$, and $s(a, b)$ are all integers. We remark that since a and b are odd, $a + b$ and $a - b$ are even. Further if 3 divides $2a + b$ then 3 divides $a - b$ and $a + 2b$. Thus $x(a, b)$, $y(a, b)$, $u(a, b)$, and $v(a, b)$ are integers. Since a and b are odd, $f(a, b)$ is odd and, since

$$f(a, b) = (2a + b)(4a^3 - 3a^2b + 4ab^2 + b^3) - 9a^4, \tag{81}$$

$r(a, b)$ and $s(a, b)$ are integers.

Now we shall show that $x(a, b)$ and $y(a, b)$ are coprime. We remark that since a and b are coprime, a and $(a-b)/(2\cdot 3^{\varepsilon})$ are coprime and $(a+2b)/3^{\varepsilon}$ and $(a+b)/2$ are coprime. Thus if p is a prime that divides both x and y then either (i) $p|a$ and $p|(a+2b)/3^{\varepsilon}$, or (ii) $p|a$ and $p|(a+b)/2$, or (iii) $p|(a-b)/(2\cdot 3^{\varepsilon})$ and $p|(a+2b)/3^{\varepsilon}$, or (iv) $p|(a-b)/(2\cdot 3^{\varepsilon})$ and $p|(a+b)/2$. In case (i) $p|a$ and $p|2b$ so $p=2$, but a is odd, which is a contradiction. In case (ii) $p|a$ and $p|b$, which is impossible. In case (iii) p divides $3a$ and $3b$ so $p=3$. But one at least of $(a-b)/(2\cdot 3^{\varepsilon})$ and $(a+2b)/3^{\varepsilon}$ is not divisible by 3 and so case (iii) does not apply. Finally, in case (iv), $p|(a-b)/2$ and $p|(a+b)/2$, hence $p|(a, b)$, which is impossible. Thus $x(a, b)$ and $y(a, b)$ are coprime.

Next we show that $u(a, b)$ and $v(a, b)$ are coprime. Observe that $(2a+b)/3^{\varepsilon}$ and $(a+b)/2$ are coprime and that b and $(a-b)/(2\cdot 3^{\varepsilon})$ are coprime. Thus if p is a prime that divides both u and v then either (i) $p|(2a+b)/3^{\varepsilon}$ and $p|b$, (ii) $p|(2a+b)/3^{\varepsilon}$ and $p|(a-b)/(2\cdot 3^{\varepsilon})$, (iii) $p|(a+b)/2$ and $p|b$, or (iv) $p|(a+b)/2$ and $p|(a-b)/(2\cdot 3^{\varepsilon})$. In case (i) $p|b$ and $p|2a$ so $p=2$, which contradicts the fact that b is odd. In case (ii) $p|3a$ and $p|3b$, hence $p=3$. But one of $(2a+b)/3^{\varepsilon}$ and $(a-b)/(2\cdot 3^{\varepsilon})$ is not divisible by 3 and so (ii) does not hold. In case (iii) $p|a$ and $p|b$, which is impossible. Finally, in case (iv) $p|(a+b)/2$ and $p|(a-b)/2$, hence $p|a$ and $p|b$, which is a contradiction. Therefore $u(a, b)$ and $v(a, b)$ are coprime.

Finally we shall show that $r(a, b)$ and $s(a, b)$ are coprime. Suppose that p is a prime that divides both r and s. Then $p|r+s$ so $p|b(2a+b)/3^{\varepsilon}$. Since b is odd, b and $(2a+b)/3^{\varepsilon}$ are coprime. If $p|b$ then, since $p|r$, we see that $p|f(a, b)$ and hence that $p|a$, which is a contradiction. On the other hand, if $p|(2a+b)/3^{\varepsilon}$ then, since $p|r$, $p|f(a, b)$ and so by (81) $p|9a^4$. Thus $p=3$. If $3|(2a+b)/3^{\varepsilon}$ then $3^2|2a+b$ and, since a and b are coprime, 3 does not divide a. From (81) we find that

$$f(a, b) = (2a+b)((2a+b)(-7a^2+2ab+b^2)+18a^3)-9a^4$$

and so 9 divides $f(a, b)$ but 27 does not divide $f(a, b)$. Since 9 divides $2a+b$ we conclude that 3 exactly divides $b(2a+b)+\sqrt{f(a, b)}$, hence 3 does not divide r. Therefore $r(a, b)$ and $s(a, b)$ are coprime.

To complete our proof that $\nu(F)\geq 18$, and hence that $\nu^*(3)\geq 18$, it suffices to show that apart from a finite set of pairs (a, b) the orbits of (x, y), (u, v), and (r, s) under D_3 are disjoint. And a case by case analysis reveals that if a and b are odd and coprime and (a, b) is different from $(1, -1)$, $(-1, 1)$, $(1, 1)$, $(-1, -1)$, $(3, -1)$, $(-3, 1)$, $(1, -5)$, $(-1, 5)$ then indeed the orbits are distinct as required.

Next we shall prove that $\nu^*(4)\geq 16$. We consider $F(x, y)=x^4+y^4$, which is invariant under D_4. Thus whenever (x, y) is a solution of (1), it follows that $(-x, y)$, $(-x, -y)$, $(x, -y)$, (y, x), $(y, -x)$, $(-y, -x)$, and $(-y, x)$ are also solutions. We now appeal to the parametric solution due to Euler of the equation $x^4+y^4=u^4+v^4$. He showed (see [18, p. 201]), that if

$$x(t)=t^7+t^5-2t^3+3t^2+t,$$
$$y(t)=t^6-3t^5-2t^4+t^2+1,$$
$$u(t)=t^7+t^5-2t^3-3t^2+t,$$
$$v(t)=t^6+3t^5-2t^4+t^2+1,$$

then $x(t)^4+y(t)^4=u(t)^4+v(t)^4$. Put

$$S(t)=-11t^5+19t^4+68t^3+15t^2-18t-8$$

and

$$T(t)=11t^6+14t^5+7t^4+15t^3-24t^2+8t+66.$$

Then

(82) $$S(t)x(t)+T(t)y(t)=66.$$

If $t\equiv 0 \pmod{66}$ then $T(t)\equiv 0 \pmod{66}$, $x(t)\equiv 0 \pmod{66}$, and $y(t)\equiv 1 \pmod{66}$, hence by (82), $x(t)$ and $y(t)$ are coprime. Next we put $M(t)=-S(-t)$ and $N(t)=T(-t)$. Then, since $u(t)=-x(-t)$ and $v(t)=y(-t)$, we have

(83) $$M(t)u(t)+N(t)v(t)=66.$$

Again if $t\equiv 0 \pmod{66}$ then $N(t)\equiv 0 \pmod{66}$, $u(t)\equiv 0 \pmod{66}$, and $v(t)\equiv 1 \pmod{66}$, hence by (83), $u(t)$ and $v(t)$ are coprime. Plainly the orbit of $(x(t), y(t))$ under D_4 does not contain $(u(t), v(t))$ for t sufficiently large and thus $\nu^*(4)\geq 16$.

To prove that $\nu^*(5)\geq 6$ we merely note that $F(x, y)=xy(x+y)(x^2+xy+y^2)$ is a form of degree 5 with nonzero discriminant that is invariant under D_3.

To prove that $\nu^*(6k)\geq 12$ and $\nu^*(6k+2)\geq 12$ for $k=1, 2, \dots$, it suffices to show that there exists a binary form with nonzero discriminant that is invariant under D_6 for these degrees. Let

$$F_c(x, y)=x^6-3x^5y+cx^4y^2+(5-2c)x^3y^3+cx^2y^4-3xy^5+y^6$$

and put $f_c(x)=F_c(x, 1)$. Then the discriminant of F_c, and of f_c, is $-(4c+3)^3(c-6)^4$. Thus the roots of f_c are distinct provided that c is an integer different from 6. Further if c_1 and c_2 are distinct integers then

$$\begin{aligned}F_{c_1}(x, y)-F_{c_2}(x, y)&=(c_1-c_2)(x^4y^2-2x^3y^3+x^2y^4)\\&=(c_1-c_2)x^2y^2(x-y)^2.\end{aligned}$$

Since $f_c(1)=1$, f_{c_1} and f_{c_2} have no roots in common. Further, by our earlier discussion $F_c(x, y)$ is invariant under D_6. Thus, for $k=1, 2, \dots$,

$$\prod_{j=1}^{k} F_{6+j}(x, y)$$

is a binary form of degree $6k$ and nonzero discriminant that is invariant under D_6. Since $x^2 - xy + y^2$ is invariant under D_6 and $f_c(e^{2\pi i/6}) = f_c(e^{-2\pi i/6}) = c - 6$,

$$(x^2 - xy + y^2)\prod_{j=1}^{k} F_{6+j}(x, y)$$

is a binary form of nonzero discriminant of degree $6k + 2$ for $k = 1, 2, \ldots$ that is invariant under D_6. Thus $\nu^*(6k) \geq 12$ and $\nu^*(6k + 2) \geq 12$ for $k = 1, 2, \ldots$.

Next we put $G_c(x, y) = F_c(-x, y)$ so that

$$G_c(x, y) = x^6 + 3x^5y + cx^4y^2 + (2c - 5)x^3y^3 + cx^2y^4 + 3xy^5 + y^6.$$

Then, recall (75), G_c is invariant under D_3. Put $g_c(x) = G_c(x, 1)$ so that $g_c(x) = f_c(-x)$. Then, as before, the roots of g_c are distinct provided c is an integer different from 6. Further if c_1 and c_2 are distinct integers then

$$G_{c_1}(x, y) - G_{c_2}(x, y) = (c_1 - c_2)x^2y^2(x + y)^2.$$

Since $g_c(-1) = 1$, g_{c_1} and g_{c_2} have no roots in common. Now both $xy(x + y)$ and $xy(x + y)(x^2 + xy + y^2)$ are invariant under D_3 and $g_c(e^{2\pi i/3}) = g_c(e^{-2\pi i/3}) = c - 6$. Thus, for $k = 1, 2, \ldots$,

$$xy(x + y)\prod_{j=1}^{k} G_{6+j}(x, y)$$

is a binary form of nonzero discriminant of degree $6k + 3$ that is invariant under D_3, and

$$xy(x + y)(x^2 + xy + y^2)\prod_{j=1}^{k} G_{6+j}(x, y)$$

is a binary form of nonzero discriminant of degree $6k+5$ that is invariant under D_3. Thus $\nu^*(6k + 3) \geq 6$ and $\nu^*(6k + 5) \geq 6$ for $k = 1, 2, \ldots$. Finally, observe that the binary form $x^{2j} + y^{2j}$ is invariant under D_4 for $j = 1, 2, \ldots$ and that the binary form $x^j + y^j$ is invariant under D_1^* for $j = 1, 2, \ldots$. Thus $\nu^*(6k + 4) \geq 8$ and $\nu^*(6k + 1) \geq 2$ for $k = 1, 2, \ldots$.

7. On S-unit equations

Let K be an algebraic number field of degree d, with discriminant D_K, and ring of integers $\mathcal{O}_K$. Let M_K be the set of places (i.e., equivalence classes of multiplicative valuations) on K. A place v is called finite if v contains only non-Archimedean valuations and infinite otherwise. K has only finitely many infinite places. Let S be a finite subset of M_K, containing all infinite places. A number $\alpha \in K$ is called an S-unit if $|\alpha|_v = 1$ for every valuation $| \ |_v$ from a place $v \in M_K \backslash S$. The S-units form a multiplicative group. Put $K\backslash\{0\} = K^*$ and let $(\alpha_1, \alpha_2, \alpha_3)$ be in $(K^*)^3$. The number of solutions of the equation

$$\alpha_1 u_1 + \alpha_2 u_2 = \alpha_3 \tag{84}$$

in S-units u_1 and u_2 is finite (see Lang [20]). We say that two triples $(\alpha_1, \alpha_2, \alpha_3)$ and $(\beta_1, \beta_2, \beta_3)$ in $(K^*)^3$ are S-equivalent if there exist a permutation σ of $(1, 2, 3)$, a $\mu \in K^*$, and S-units $\varepsilon_1, \varepsilon_2, \varepsilon_3$ such that

$$\beta_i = \mu \varepsilon_i \alpha_{\sigma(i)} \quad \text{for } i = 1, 2, 3.$$

It is easy to check that if $(\alpha_1, \alpha_2, \alpha_3)$ and $(\beta_1, \beta_2, \beta_3)$ are S-equivalent then the equation $\beta_1 u_1 + \beta_2 u_2 = \beta_3$ in S-units u_1 and u_2 has the same number of solutions as (84). Next let $\mathfrak{p}_1, \dots, \mathfrak{p}_t$ be the prime ideals corresponding to the finite places in S. For any $\alpha \in K^*$ the principal ideal (α) can be written uniquely as a product of two (not necessarily principal) ideals $\mathfrak{a}'$ and $\mathfrak{a}''$, where $\mathfrak{a}'$ is composed of $\mathfrak{p}_1, \dots, \mathfrak{p}_t$ and $\mathfrak{a}''$ is composed solely of prime ideals different from $\mathfrak{p}_1, \dots, \mathfrak{p}_t$. We put

$$N_S(\alpha) = N_{K/\mathbb{Q}}(\mathfrak{a}'').$$

Recently, Evertse, Györy, Stewart, and Tijdeman [17] proved that almost all equivalence classes of S-unit equations of the form (84) have very few solutions and their result is our next lemma.

Lemma 7. *Let S be a finite subset of M_K containing all infinite places. There exists a finite set A of triples in $(\mathscr{O}_K\backslash\{0\})^3$ with the following property: for each triple $(\alpha_1, \alpha_2, \alpha_3) \in (K^*)^3$ that is not S-equivalent to any of the triples from A, the number of solutions of* (84) *is at most two.*

Proof. This is Theorem 1 of [17] together with the observation that we may take the triples in A from $(\mathscr{O}_K\backslash\{0\})^3$.

S-unit equations are of great interest since the study of many Diophantine equations can be reduced to the study of certain associated S-unit equations. In the next section we shall make use of such a reduction to study the Thue-Mahler equation and the generalized Ramanujan-Nagell equation. We shall also appeal to an effective version of Lemma 7 established in [17].

Lemma 8. *Let S be a finite subset of M_K of cardinality s, containing all infinite places. Suppose that the rational primes corresponding to the finite places in S do not exceed P (≥ 2). Let B denote the set of triples $(\beta_1, \beta_2, \beta_3)$ in $(\mathscr{O}_K\backslash\{0\})^3$ with*

$$\max(N_S(\beta_1), N_S(\beta_2), N_S(\beta_3)) \leq \exp((C_1 s)^{C_2 s} P^{d+1}), \tag{85}$$

where C_1 and C_2 are certain explicitly computed numbers depending only on d and $|D_K|$. Then for each triple $(\alpha_1, \alpha_2, \alpha_3) \in (K^*)^3$ that is not S-equivalent to any of the triples in B, the number of solutions of (84) *is at most $s + 1$.*

Proof. If in inequality (85) we replace $\max(N_S(\beta_1), N_S(\beta_2), N_S(\beta_3))$ by $\max(h(\beta_1), h(\beta_2), h(\beta_3))$ we have Theorem 2 of [17]. The result now follows from the observation that for each β in $\mathscr{O}_K \backslash \{0\}$,

$$1 \le N_S(\beta) \le |N_{K/\mathbb{Q}}(\beta)| \le h(\beta)^d .$$

8. Thue-Mahler and Ramanujan-Nagell equations

Bombieri [3] has obtained an estimate for the number of primitive solutions of the Thue-Mahler equation (2) that is better with respect to the dependence on the degree r than the estimate (3) of Evertse and yet is still independent of the coefficients of F. It follows from his result that, if r is at least 6 and the discriminant of F is nonzero then there are at most

$$(4(t+1))^2(4r)^{26(t+1)}$$

primitive solutions of (2).

Let h be a nonzero integer, let t be a nonnegative integer, and let $p_1, \dots, p_t$ be prime numbers. In this section we shall estimate the number of primitive solutions of the equation

$$F(x, y) = hp_1^{k_1} \cdots p_t^{k_t} ; \tag{86}$$

of course if $h = 1$ we again obtain (2). We shall establish bounds for the number of solutions of (86) in coprime integers x and y and integers $k_1, \dots, k_t$ under the assumption that h is coprime with p_i for $i = 1, \dots, t$ and sufficiently large. Our bounds are much sharper with respect to the parameter t than the exponential dependence on t of previous results.

Theorem 4. *Let F be a binary form with integer coefficients, content* 1, *degree r (≥ 3), and nonzero discriminant D. Let t be a nonnegative integer and let $p_1, \dots, p_t$ be prime numbers of size at most P (≥ 2). Let h be a positive integer that is coprime with p_i for $i = 1, \dots, t$. For h sufficiently large the number of solutions of equation* (86) *in coprime integers x and y and integers $k_1, \dots, k_t$ is at most*

$$4r^{\omega(h)} . \tag{87}$$

Further there exists a number C that is effectively computable in terms of r and D such that if

$$h > \exp((t+2)^{C(t+1)} P^{r^3}) , \tag{88}$$

then the number of solutions of (86) *in coprime integers x and y and integers $k_1, \dots, k_t$ is at most*

$$2(t+1)r^{3+\omega(h)} . \tag{89}$$

The most significant aspect of Theorem 4 is the dependence of the upper bounds (87) and (89) on the parameter t. Estimate (87), which is independent of t, applies for h sufficiently large. However, since the proof of (87) depends upon Lemma 7 and hence upon the Thue-Siegel-Roth-Schmidt theorem it does not yield an effective estimate for how large h must be. For this reason we have also given the slightly weaker estimate (89) that is linear in t and holds subject to h satisfying the effective estimate (88).

In fact, estimates as sharp as (87) and (89) do not apply for general h. Let ε be a positive number and let $2 = p_1, p_2, \ldots$ be the sequence of prime numbers. In [10] Erdös, Stewart, and Tijdeman proved that for every integer r with $r \geq 2$ there exists a number $t_0(\varepsilon, r)$, which is effectively computable in terms of ε and r such that if t is an integer with $t \geq t_0(\varepsilon, r)$ then there exists a monic polynomial f, with integer coefficients, degree r, and nonzero discriminant for which the equation

$$f(x) = p_1^{k_1} \cdots p_t^{k_t} \tag{90}$$

has at least

$$\exp((r^2 - \varepsilon)t^{1/r}(\log t)^{-(r-1)/r}) \tag{91}$$

solutions in nonnegative integers $x, k_1, \ldots, k_t$. Recently Moree and Stewart [28] proved that, provided we replace $r^2 - \varepsilon$ by $r - \varepsilon$ in (91), we may also suppose that f is irreducible.

We remark that when $t = 0$ estimate (87) gives a slight improvement, for h sufficiently large, of the estimate (4) of Bombieri and Schmidt [5]. Further, if r is odd then the proof of Theorem 4 allows one to replace $4r^{\omega(h)}$ in (87) by $2r^{\omega(h)}$ and similarly to eliminate the factor 2 in estimate (89).

Equation (90) is an example of a Ramanujan-Nagell equation. In [13] Evertse proved that if f is a quadratic polynomial with integer coefficients and nonzero discriminant and $p_1, \ldots, p_t$ are distinct prime numbers then equation (90) has at most $3 \cdot 7^{6+4t}$ solutions in integers $x, k_1, \ldots, k_t$. Let h be a positive integer. Next we shall establish estimates for the number of solutions in integers $x, k_1, \ldots, k_t$ of the generalized Ramanujan-Nagell equation

$$f(x) = hp_1^{k_1} \cdots p_t^{k_t}. \tag{92}$$

Theorem 5. *Let f be a polynomial with integer coefficients, content* 1, *leading coefficient a, degree r (≥ 2), and nonzero discriminant D. Let t be a nonnegative integer and let $p_1, \ldots, p_t$ be prime numbers of size at most P (≥ 2). Let h be a positive integer that is coprime with p_i for $i = 1, \ldots, t$. For h sufficiently large the number of solutions of* (92) *in integers x and $k_1, \ldots, k_t$ is at most*

$$2r^{\omega(h)} \tag{93}$$

Further there exists a number C, which is effectively computable in terms of a, r, and D, such that if

$$h > \exp((t+2)^{C(t+1)}P^{r^2})$$

then the number of solutions of (92) *in integers* x *and* $k_1, \ldots, k_t$ *is at most* $(t+1)r^{2+\omega(h)}$.

Finally we mention that Evertse and Györy [14] have also applied the estimates for the number of solutions of S-unit equations from [17] to bound the number of solutions of equations such as (86). Let $S = \{p_1, \ldots, p_t\}$ be a set of primes. Two binary forms F and G are S-equivalent if $G(x, y) = ef^{-1}F(ax+by, cx+dy)$ for certain integers a, b, c, d, e, f with $|ad-bc|$, $|e|$, and $|f|$ composed of primes from S. For any algebraic number field L and integer $r \geq 3$ let $A(r, L)$ be the set of binary forms of degree r with integer coefficients that factorize into linear forms in $L[x, y]$ and whose factorization contains at least three pairwise linearly independent linear forms. Evertse and Györy show, for instance, that the set of forms in $A(r, L)$ for which (86) has more than $2(r, 2)$ solutions is contained in the union of a finite collection of S-equivalence classes.

9. Proof of Theorem 4

Since $D \neq 0$, $F(x, y)$ has at most a single power of x and at most a single power of y in any factorization in $\mathbb{C}[x, y]$. Thus we may factor F as

$$F(x, y) = ax^{\delta_1}(x - \alpha_{1+\delta_1}y)\cdots(x - \alpha_{r-\delta_2}y)y^{\delta_2} \tag{94}$$

and

$$F(x, y) = by^{\delta_2}(y - \gamma_{1+\delta_2}x)\cdots(y - \gamma_{r-\delta_1}x)x^{\delta_1}, \tag{95}$$

where a and b are nonzero integers, δ_1 and δ_2 are from $\{0, 1\}$, and $\gamma_{r+1-j} = \alpha_j^{-1}$ for $j = 1+\delta_1, \ldots, r-\delta_2$. Put $K = \mathbb{Q}(\alpha_{1+\delta_1}, \ldots, \alpha_{r-\delta_2})$ and let $\mathscr{O}_K$ denote the ring of algebraic integers of K. Let q be a prime number and let $\mathfrak{q}$ be a prime ideal in $\mathscr{O}_K$ lying above q. For each $\alpha \in K^*$ we define $\operatorname{ord}_\mathfrak{q} \alpha$ to be the exponent of $\mathfrak{q}$ in the prime ideal decomposition of the fractional ideal of K generated by α. We shall suppose that

$$\operatorname{ord}_\mathfrak{q} \alpha_{1+\delta_1} \geq \cdots \geq \operatorname{ord}_\mathfrak{q} \alpha_w \geq 0 > \operatorname{ord}_\mathfrak{q} \alpha_{w+1} \geq \cdots \geq \operatorname{ord}_\mathfrak{q} \alpha_{r-\delta_2},$$

where $\delta_1 \leq w \leq r - \delta_2$. Since F has content 1,

$$\operatorname{ord}_\mathfrak{q} a + \operatorname{ord}_\mathfrak{q} \alpha_{w+1} + \cdots + \operatorname{ord}_\mathfrak{q} \alpha_{r-\delta_2} = 0. \tag{97}$$

Put $u = r - \delta_2$ and if $\delta_1 = 1$ put $\alpha_1 = 0$. Then, from (94) we have

$$F(x, y) = a(x - \alpha_1 y)\cdots(x - \alpha_u y)y^{\delta_2}.$$

Similarly put $v = r - \delta_1$ and if $\delta_2 = 1$ put $\gamma_1 = 0$ so that, by (95),

$$F(x, y) = b(y - \gamma_1 x)\cdots(y - \gamma_v x)x^{\delta_1}.$$

We shall now consider the tuples of the form

$$(\operatorname{ord}_\mathfrak{q} x^{\delta_1}, \operatorname{ord}_\mathfrak{q}(x - \alpha_{1+\delta_1}y), \ldots, \operatorname{ord}_\mathfrak{q}(x - \alpha_{r-\delta_2}y), \operatorname{ord}_\mathfrak{q} y^{\delta_2}), \tag{98}$$

where (x, y) yields a primitive solution of (86). If q does not divide $hp_1 \cdots p_t$ then (98) is determined independently of x and y. For if $q \nmid a$ then by (86) it is $(0, 0, \ldots, 0)$ whereas if $q|a$ then by (86), (96), and (97) it is

$$(0, \ldots, 0, \operatorname{ord}_q \alpha_{w+1}, \ldots, \operatorname{ord}_q \alpha_{r-\delta_2}, 0).$$

We shall now show that if $q|h$ then there are at most r positive tuples of the form (98) whenever x and y give coprime solutions of (86).

We first suppose that (x, y) yields a solution of (86) in coprime integers and that $q|h$ and $q \nmid y$. We now choose an integer l, with $1 \le l \le u$, for which

$$\operatorname{ord}_q(x - \alpha_l y) = \max_{1 \le i \le u} \operatorname{ord}_q(x - \alpha_i y).$$

Since $q \nmid y$, $\operatorname{ord}_q \alpha_j = \operatorname{ord}_q \alpha_j y < 0 \le \operatorname{ord}_q x$ and hence $\operatorname{ord}_q(x - \alpha_j y) = \operatorname{ord}_q \alpha_j$ for $j = w+1, \ldots, u$. Since $q|h$ we conclude from (86) that $1 \le l \le w$. Observe that, for $j = 1, \ldots, u$,

$$x - \alpha_j y = x - \alpha_l y + (\alpha_l - \alpha_j)y,$$

hence, since $q \nmid y$,

$$\operatorname{ord}_q(x - \alpha_j y) = \min(\operatorname{ord}_q(x - \alpha_l y), \operatorname{ord}_q(\alpha_l - \alpha_j)). \tag{99}$$

Further, by (86) we have

$$\operatorname{ord}_q a + \operatorname{ord}_q(x - \alpha_1 y) + \cdots + \operatorname{ord}_q(x - \alpha_u y) = \operatorname{ord}_q h. \tag{100}$$

Equations (99) and (100) determine $\operatorname{ord}_q(x - \alpha_l y)$ and hence also $\operatorname{ord}_q(x - \alpha_j y)$ for $j = 1, \ldots, u$. Thus there are at most w possible tuples (98) that can arise from primitive solutions (x, y) of (86) for which $q \nmid y$ and $q|h$.

Suppose now that (x, y) is a primitive solution of (86) for which $q|y$ and $q|h$. Then, since x and y are coprime, $q \nmid x$. Since $\gamma_{r+1-j} = \alpha_j^{-1}$ for $j = 1 + \delta_1, \ldots, r - \delta_2$ we have, by (96),

$$\operatorname{ord}_q \gamma_{1+\delta_2} \ge \cdots \ge \operatorname{ord}_q \gamma_{r-w} > 0 \ge \operatorname{ord}_q \gamma_{r-w+1} \ge \cdots \ge \operatorname{ord}_q \gamma_{r-\delta_1}.$$

Further, since F has content 1,

$$\operatorname{ord}_q b + \operatorname{ord}_q \gamma_{r-w+1} + \cdots + \operatorname{ord}_q \gamma_{r-\delta_1} = 0. \tag{101}$$

Now $q \nmid x$ and $q|y$ so $\operatorname{ord}_q \gamma_j = \operatorname{ord}_q \gamma_j x \le 0 < \operatorname{ord}_q y$ and hence $\operatorname{ord}_q(y - \gamma_j x) = \operatorname{ord}_q \gamma_j$ for $j = r - w + 1, \ldots, r - \delta_1$. We now choose k so that

$$\operatorname{ord}_q(y - \gamma_k x) = \max_{1 \le i \le v} \operatorname{ord}_q(y - \gamma_i x).$$

Since $q|h$ we see from (86) and (101) that $1 \le k \le r - w$. Further, for $j = 1, \ldots, v$,

$$\operatorname{ord}_q(y - \gamma_j x) = \min(\operatorname{ord}_q(y - \gamma_k x), \operatorname{ord}_q(\gamma_k - \gamma_j)). \tag{102}$$

We also have, from (86),

$$\operatorname{ord}_{\mathfrak{q}} b + \operatorname{ord}_{\mathfrak{q}}(y - \gamma_1 x) + \cdots + \operatorname{ord}_{\mathfrak{q}}(y - \gamma_v x) = \operatorname{ord}_{\mathfrak{q}} h. \tag{103}$$

As before, (102) and (103) determine $\operatorname{ord}_{\mathfrak{q}}(y - \gamma_k x)$, hence $\operatorname{ord}_{\mathfrak{q}}(y - \gamma_j x)$ for $j = 1, \ldots, v$, and thus in turn they determine (98).

Therefore if (x, y) gives a primitive solution of (86) and $q|h$ then (98) is one of at most $(r-w)+w = r$ possible tuples, whereas if $q \nmid hp_1 \cdots p_t$ the tuple (98) is uniquely determined. Since K is Galois over $\mathbb{Q}$, all prime ideals of $\mathcal{O}_K$ lying over q are conjugate. Every automorphism of K induces a permutation of $(\alpha_{1+\delta_1}, \ldots, \alpha_{r-\delta_2})$ and thus corresponding to each tuple (98) there is, for each prime ideal $\mathfrak{q}'$ lying over q in $\mathcal{O}_K$, a unique tuple of the form (98) but with $\mathfrak{q}$ replaced by $\mathfrak{q}'$.

For notational ease we write

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

where $\theta_i = 1$ for $i = 1, \ldots, r$ except when $\delta_2 = 1$ in which case $\theta_r = 0$ and $\alpha_r = -1$. Let S denote the set of infinite places in K together with those finite places that correspond to a prime ideal in $\mathcal{O}_K$ that divides an ideal generated by p_j for $j = 1, \ldots, t$. Let $\{(x_1, y_1), \ldots, (x_n, y_n)\}$ be a set of pairs of coprime integers that give solutions of (97) and suppose that the set is maximal subject to the constraint that whenever $j \neq i$ the tuple

$$\left(\frac{\theta_1 x_j - \alpha_1 y_j}{\theta_1 x_i - \alpha_1 y_i}, \ldots, \frac{\theta_r x_j - \alpha_r y_j}{\theta_r x_i - \alpha_r y_i} \right) \tag{104}$$

is not a tuple of S-units. Then, by the preceding discussion, $n \leq r^{\omega(h)}$.

Let x and y be coprime integers that give a solution of (86). Then there is an integer j with $1 \leq j \leq n$ such that the tuple (104) is a tuple of S-units when we replace x_i, y_i by x, y respectively. We may assume, without loss of generality, that

$$\begin{aligned} &N_S((\theta_1 x_j - \alpha_1 y_j)(\theta_2 x_j - \alpha_2 y_j)(\theta_3 x_j - \alpha_3 y_j)) \\ &\qquad \geq N_S((\theta_{i_1} x_j - \alpha_{i_1} y_j)(\theta_{i_2} x_j - \alpha_{i_2} y_j)(\theta_{i_3} x_j - \alpha_{i_3} y_j)) \end{aligned}$$

for all triples (i_1, i_2, i_3) with $1 \leq i_1 < i_2 < i_3 \leq r$. Thus, by (86),

$$N_S((\theta_1 x_j - \alpha_1 y_j)(\theta_2 x_j - \alpha_2 y_j)(\theta_3 x_j - \alpha_3 y_j)) \geq (N_S(h/a))^{3/r}. \tag{105}$$

Let $K_1 = \mathbb{Q}(\alpha_1, \alpha_2, \alpha_3)$ and let S_1 denote the set of infinite places in K_1 together with those finite places that correspond to a prime ideal in $\mathcal{O}_{K_1}$ that divides an ideal generated by p_j for j with $1 \leq j \leq t$. Let d_1 denote the degree of K_1 over $\mathbb{Q}$. It follows from (105) that

$$N_{S_1}((\theta_1 x_j - \alpha_1 y_j)(\theta_2 x_j - \alpha_2 y_j)(\theta_3 x_j - \alpha_3 y_j)) \geq (N_{S_1}(h/a))^{3/r}. \tag{106}$$

Put

$$\lambda = a(\theta_2\alpha_3 - \theta_3\alpha_2)(\theta_1 x_j - \alpha_1 y_j),$$
$$\eta = a(\theta_3\alpha_1 - \theta_1\alpha_3)(\theta_2 x_j - \alpha_2 y_j),$$

and

$$\tau = a(\theta_2\alpha_1 - \theta_1\alpha_2)(\theta_3 x_j - \alpha_3 y_j).$$

Plainly λ, η, and τ are in $\mathscr{O}_{K_1}\backslash\{0\}$. Further

$$a(\theta_2\alpha_3 - \theta_3\alpha_2)(\theta_1 x - \alpha_1 y) = \lambda u_1,$$
$$a(\theta_3\alpha_1 - \theta_1\alpha_3)(\theta_2 x - \alpha_2 y) = \eta u_2,$$

and

$$a(\theta_2\alpha_1 - \theta_1\alpha_2)(\theta_3 x - \alpha_3 y) = \tau u_3,$$

where u_1, u_2, and u_3 are S_1-units. But then

(107)
$$\frac{(\theta_2\alpha_3 - \theta_3\alpha_2)(\theta_1 x - \alpha_1 y)}{(\theta_2\alpha_1 - \theta_1\alpha_2)(\theta_3 x - \alpha_3 y)} = \frac{\lambda}{\tau}U_1,$$
$$\frac{(\theta_3\alpha_1 - \theta_1\alpha_3)(\theta_2 x - \alpha_2 y)}{(\theta_2\alpha_1 - \theta_1\alpha_2)(\theta_3 x - \alpha_3 y)} = \frac{\eta}{\tau}U_2,$$

where U_1 and U_2 are S_1-units, and so

(108)
$$\lambda U_1 + \eta U_2 = \tau.$$

Further, by (107), each pair of S_1-integers (U_1, U_2) determines at most two pairs of coprime integers (x, y), $((x, y), (-x, -y))$.

Suppose that $(\beta_1, \beta_2, \beta_3)$ is a triple in $(\mathscr{O}_{K_1}\backslash\{0\})^3$ that is S_1-equivalent to (λ, η, τ). Then there exist a permutation σ of $(1, 2, 3)$, a μ in K_1^*, and S_1-units $\varepsilon_1, \varepsilon_2, \varepsilon_3$ such that $\mu\varepsilon_1\lambda = \beta_{\sigma(1)}$, $\mu\varepsilon_2\eta = \beta_{\sigma(2)}$, and $\mu\varepsilon_3\tau = \beta_{\sigma(3)}$. We shall now show that the maximum of $N_{S_1}(\beta_1)$, $N_{S_1}(\beta_2)$, and $N_{S_1}(\beta_3)$ is large. To this end, let $\mathfrak{p}$ be a prime ideal of $\mathscr{O}_{K_1}\backslash\{0\}$ and put $a_1 = \operatorname{ord}_{\mathfrak{p}} a$ and $b_1 = \operatorname{ord}_{\mathfrak{p}}(\theta_2\alpha_1 - \theta_1\alpha_2)$. Let $(i_1, \ldots, i_r)$ be a permutation of $(1, \ldots, r)$ for which

$$\operatorname{ord}_{\mathfrak{p}} \alpha_{i_1} \le \operatorname{ord}_{\mathfrak{p}} \alpha_{i_2} \le \cdots \le \operatorname{ord}_{\mathfrak{p}} \alpha_{i_r},$$

and let g be that integer with $0 \le g \le r$ for which

$$\operatorname{ord}_{\mathfrak{p}} \alpha_{i_1} \le \cdots \le \operatorname{ord}_{\mathfrak{p}} \alpha_{i_g} < 0 \le \operatorname{ord}_{\mathfrak{p}} \alpha_{i_{g+1}} \le \cdots \le \operatorname{ord}_{\mathfrak{p}} \alpha_{i_r}.$$

Since F has content 1 we have

(109)
$$-a_1 = \operatorname{ord}_{\mathfrak{p}} \alpha_{i_1} + \cdots + \operatorname{ord}_{\mathfrak{p}} \alpha_{i_g}.$$

Thus

$$\sum_{i<j,\, \operatorname{ord}_{\mathfrak{p}}(\theta_j\alpha_i - \theta_i\alpha_j)<0} \operatorname{ord}_{\mathfrak{p}}(\theta_j\alpha_i - \theta_i\alpha_j)$$
$$\ge (n-1)\operatorname{ord}_{\mathfrak{p}} \alpha_{i_1} + (n-2)\operatorname{ord}_{\mathfrak{p}} \alpha_{i_2} + \cdots + (n-g)\operatorname{ord}_{\mathfrak{p}} \alpha_{i_g}$$
$$\ge -(n-1)a_1.$$

Therefore, since

$$D = a^{2n-2}\prod_{i<j}(\theta_j\alpha_i - \theta_i\alpha_j)^2,$$

we find that

$$\operatorname{ord}_{\mathfrak{p}}(\theta_1\alpha_2 - \theta_2\alpha_1) + \max(\operatorname{ord}_{\mathfrak{p}}(\theta_2\alpha_3 - \theta_3\alpha_2), \operatorname{ord}_{\mathfrak{p}}(\theta_3\alpha_1 - \theta_1\alpha_3)) \le \frac{1}{2}\operatorname{ord}_{\mathfrak{p}} D. \tag{110}$$

We have

$$\theta_2(\theta_1 x_j - \alpha_1 y_j) - \theta_1(\theta_2 x_j - \alpha_2 y_j) = (\theta_1\alpha_2 - \theta_2\alpha_1)y_j \tag{111}$$

and

$$\alpha_2(\theta_1 x_j - \alpha_1 y_j) - \alpha_1(\theta_2 x_j - \alpha_2 y_j) = (\theta_1\alpha_2 - \theta_2\alpha_1)x_j. \tag{112}$$

Thus, by (109) and (111),

$$\min(\operatorname{ord}_{\mathfrak{p}}(\theta_1 x_j - \alpha_1 y_j), \operatorname{ord}_{\mathfrak{p}}(\theta_2 x_j - \alpha_2 y_j)) \le a_1 + b_1 + \operatorname{ord}_{\mathfrak{p}} y_j, \tag{113}$$

while, by (109) and (112),

$$\min(\operatorname{ord}_{\mathfrak{p}}(\theta_1 x_j - \alpha_1 y_j), \operatorname{ord}_{\mathfrak{p}}(\theta_2 x_j - \alpha_2 y_j)) \le a_1 + b_1 + \operatorname{ord}_{\mathfrak{p}} x_j. \tag{114}$$

Since x_j and y_j are coprime we conclude that

$$\min(\operatorname{ord}_{\mathfrak{p}}(\theta_1 x_j - \alpha_1 y_j), \operatorname{ord}_{\mathfrak{p}}(\theta_2 x_j - \alpha_2 y_j)) \le a_1 + b_1.$$

Therefore

$$\min(\operatorname{ord}_{\mathfrak{p}}\lambda, \operatorname{ord}_{\mathfrak{p}}\eta) \le 2a_1 + b_1 + \max(\operatorname{ord}_{\mathfrak{p}}(\theta_2\alpha_3 - \theta_3\alpha_2), \operatorname{ord}_{\mathfrak{p}}(\theta_3\alpha_1 - \theta_1\alpha_3)). \tag{115}$$

Thus, by (110) and (115),

$$\min(\operatorname{ord}_{\mathfrak{p}}\lambda, \operatorname{ord}_{\mathfrak{p}}\eta, \operatorname{ord}_{\mathfrak{p}}\tau) \le 2a_1 + \tfrac{1}{2}\operatorname{ord}_{\mathfrak{p}} D.$$

Accordingly,

$$N_{S_1}(\beta_1\beta_2\beta_3) \ge N_{S_1}(\lambda\eta\tau)\cdot(N_{S_1}(a^2D))^{-3}. \tag{116}$$

Since $a^3(\theta_2\alpha_3-\theta_3\alpha_2)(\theta_3\alpha_1-\theta_1\alpha_3)(\theta_1\alpha_2-\theta_2\alpha_1)$ is in $\mathscr{O}_{K_1}\backslash\{0\}$, it follows from (106) and (116) that

$$N_{S_1}(\beta_1\beta_2\beta_3) \ge (N_{S_1}(h/a))^{3/r}(N_{S_1}(a^2D))^{-3} \ge h^{3d_1/r}(N_{S_1}(a^3D))^{-3}.$$

Thus

$$\max(N_{S_1}(\beta_1), N_{S_1}(\beta_2), N_{S_1}(\beta_3)) \ge h^{d_1/r}|a^3D|^{-d_1}. \tag{117}$$

On the other hand, by Lemma 7, there is a finite set A of triples in $(\mathscr{O}_{K_1}\backslash\{0\})^3$ such that if $(\delta_1, \delta_2, \delta_3)$ is in $(K_1^*)^3$ and the S_1-unit equation $\delta_1 u_1 + \delta_2 u_2 = \delta_3$ has more than two solutions in S_1-units u_1 and u_2 then $(\delta_1, \delta_2, \delta_3)$ is S_1-equivalent to one of the triples from A. Thus, by (117), if h is sufficiently large then the equation (108) determined by the triple (λ, η, τ) has at most two solutions in S_1-units U_1 and U_2 and these solutions arise from at most four primitive solutions of (86), in fact at most two primitive solutions if r is odd. Since there are at most n such triples (λ, η, τ) the number of solutions of (86) in coprime integers x, y and integers $k_1, \ldots, k_t$ is at most $4n$, hence at most $4r^{w(h)}$ for h sufficiently large.

To prove (89) we apply Lemma 8 with $K = K_1$, $S = S_1$, and $d = d_1$. Note that

$$|S_1| + 1 \leq d_1 t + d_1 + 1 \leq r(r-1)(r-2)(t+1) + 1 < r^3(t+1).$$

Further, there exist numbers C_1 and C_2 that are explicitly computable in terms of d_1 and $|D_{K_1}|$ such that if $(\alpha_1, \alpha_2, \alpha_3)$ is a triple in $(K_1^*)^3$ that is not S_1-equivalent to a triple (ν_1, ν_2, ν_3) in $(\mathscr{O}_{K_1}\backslash\{0\})^3$ with

$$\max(N_{S_1}(\nu_1), N_{S_1}(\nu_2), N_{S_1}(\nu_3)) \leq \exp(C_1(t+1)^{C_2(t+1)} P^{d_1+1}),$$

then equation (84) has at most $|S_1|+1$ solutions in S_1-units u_1 and u_2. Thus, by (117), provided

$$h > (|a|^3|D|)^r \exp(rC_1(t+1)^{C_2(t+1)} P^{d_1+1}),$$

equation (108) has at most $|S_1| + 1$, hence at most $r^3(t+1)$, solutions in S_1-units U_1 and U_2 and these solutions arise from at most $2r^3(t+1)$ primitive solutions of (86). Since the number or primitive solutions of (86) is unchanged when we replace F by F_A for any A in $\mathrm{GL}(2, \mathbb{Z})$, we may suppose that $|a|$ is minimal. It follows from Theorem 1 of Evertse and Győry [15] that we may suppose that $|a|$ is less than a number that is effectively computable in terms of r and D only. Therefore there is a number C_3, which is effectively computable in terms of r and D, such that if

$$h > \exp((t+2)^{C_3(t+1)} P^{r^3}),$$

then (86) has at most $2(t+1)r^{3+w(h)}$ solutions in coprime integers x, y and integers $k_1, \ldots, k_t$.

10. Proof of Theorem 5

Let $f(x) = a(x-\alpha_1)\cdots(x-\alpha_r)$, let $K = \mathbb{Q}(\alpha_1, \ldots, \alpha_r)$, and let S denote the set of infinite places in K together with those finite places that correspond to a prime ideal in $\mathscr{O}_K$ that divides an ideal generated by p_j for $j = 1, \ldots, t$. Let $\{x_1, \ldots, x_n\}$ be integers that give solutions of (92) and suppose that the set is maximal subject to the constraint that whenever $j \neq i$ the tuple

$$\left(\frac{x_j - \alpha_1}{x_i - \alpha_1}, \ldots, \frac{x_j - \alpha_r}{x_i - \alpha_r}\right) \tag{118}$$

is not a tuple of S-units. Then, as in the proof of Theorem 4, $n \leq r^{w(h)}$.

Let x be an integer that gives a solution of (92). Then there is an integer j with $1 \leq j \leq n$ such that the tuple (118) is a tuple of S-units when we replace x_i by x. We may assume, without loss of generality, that

$$N_S((x_j - \alpha_1)(x_j - \alpha_2)) \geq N_S((x_j - \alpha_{i_1})(x_j - \alpha_{i_2}))$$

for all pairs (i_1, i_2) with $1 \leq i_1 < i_2 \leq r$. Thus, by (92),

$$N_S((x_j - \alpha_1)(x_j - \alpha_2)) \geq (N_S(h/a))^{2/r}. \tag{119}$$

Let $K_1 = \mathbb{Q}(\alpha_1, \alpha_2)$ and let S_1 be defined for K_1 in an analogous way to our definition of S for K. Let d_1 be the degree of K_1 over $\mathbb{Q}$. By (119),

$$N_{S_1}((x_j - \alpha_1)(x_j - \alpha_2)) \geq (N_{S_1}(h/a))^{2/r}. \tag{120}$$

Put $\lambda = a(x_j - \alpha_1)$, $\eta = a(x_j - \alpha_2)$, and $\tau = a(\alpha_2 - \alpha_1)$. Then λ, η, and τ are in $\mathscr{O}_{K_1} \backslash \{0\}$ and

$$a(x - \alpha_1) = \lambda u_1 \quad \text{and} \quad a(x - \alpha_2) = \eta u_2,$$

where u_1 and u_2 are S_1-units. Then

$$\lambda u_1 - \eta u_2 = \tau, \tag{121}$$

and each pair of S_1-units (u_1, u_2) determines a unique integer x.

If $(\beta_1, \beta_2, \beta_3)$ is a triple in $(\mathscr{O}_{K_1} \backslash \{0\})^3$ that is S_1-equivalent to (λ, η, τ) then, by (120),

$$\begin{aligned} N_{S_1}(\beta_1\beta_2\beta_3) &\geq N_{S_1}(\lambda\eta\tau)(N_{S_1}(a(\alpha_1 - \alpha_2)))^{-3} \\ &\geq (N_{S_1}(h))^{2/r}(N_{S_1}(a(\alpha_1 - \alpha_2)))^{-2}. \end{aligned}$$

Thus, as in the proof of (110),

$$\begin{aligned} N_{S_1}(\beta_1\beta_2\beta_3) &\geq h^{2d_1/r}(N_{S_1}(aD))^{-2} \\ &\geq h^{2d_1/r}|aD|^{-2d_1}. \end{aligned}$$

We now complete the proof by appealing to Lemmas 7 and 8 as we did for the proof of Theorem 4. Here we make use of the fact that $d_1 \leq r(r-1)$ and $|S_1| + 1 \leq d_1 t + d_1 + 1 < r^2(t+1)$.

References

1. Y. Amice, *Les nombres p-adiques*, Presses Univ. France, 1985.
2. E. Bombieri, *On the Thue-Siegel-Dyson theorem*, Acta Math. **148** (1982), 255–296.
3. ———, *On the Thue-Mahler equation*, Diophantine Approximation and Transcendence Theory (Seminar, Bonn 1985) (G. Wüstholz, ed.), Lecture Notes in Math., vol. 1290, Springer-Verlag, Berlin, Heidelberg, New York, 1987, pp. 213–243.
4. E. Bombieri and J. Mueller, *On effective measures of irrationality for* $\sqrt[r]{\frac{a}{b}}$ *and related numbers*, J. Reine Angew. Math. **342** (1983), 173–196.
5. E. Bombieri and W. M. Schmidt, *On Thue's equation*, Invent. Math. **88** (1987), 69–81.

6. N. G. de Bruijn, *On the number of positive integers* $\leq x$ *and free of prime factors* $> y$, Nederl. Akad. Wetensch. Proc. Ser. A **54** (1951), 50–60.
7. J. H. H. Chalk and R. A. Smith, *Sándor's theorem on polynomial congruences and Hensel's lemma*, C. R. Math. Rep. Acad. Sci. Canada **4** (1982), 49–54.
8. S. Chowla, *Contributions to the analytic theory of numbers*. II, J. Indian Math. Soc. **20** (1933), 120–128.
9. P. Erdös and K. Mahler, *On the number of integers which can be represented by a binary form*, J. London Math. Soc. **13** (1938), 134–139.
10. P. Erdös, C. L. Stewart, and R. Tijdeman, *Some diophantine equations with many solutions*, Compositio Math. **66** (1988), 37–56.
11. J. H. Evertse, *On the equation* $ax^n - by^n = c$, Compositio Math. **47** (1982), 289–315.
12. ——, *Upper bounds for the numbers of solutions of diophantine equations*, M.C.-Tract **168**, Centre of Mathematics and Computer Science, Amsterdam, 1983.
13. ——, *On equations in S-units and the Thue-Mahler equation*, Invent. Math. **75** (1984), 561–584.
14. J. H. Evertse and K. Györy, *Thue-Mahler equations with a small number of solutions*, J. Reine Angew. Math. **399** (1989), 60–80.
15. ——, *Effective finiteness results for binary forms with given discrminant*, Compositio Math. **79** (1991), 169–204.
16. ——, *Thue inequalities with a small number of solutions* (to appear).
17. J. H. Evertse, K. Györy, C. L. Stewart, and R. Tijdeman, *On S-unit equations in two unknowns*, Invent. Math. **92** (1988), 461–477.
18. G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 5th ed., Oxford Univ. Press, 1979.
19. M. N. Huxley, *A note on polynomial congruences*, Recent Progress in Analytic Number Theory, Vol. 1 (H. Halberstam and C. Hooley, eds.), Academic Press, London, 1981, pp. 193–196.
20. S. Lang, *Integral points on curves*, Inst. Hautes Études Sci. Publ. Math. **6** (1960), 27–43.
21. D. Lewis and K. Mahler, *Representation of integers by binary forms*, Acta Arith. **6** (1961), 333–363.
22. J. H. Loxton and R. A. Smith, *On Hua's estimate for exponential sums*, J. London Math. Soc. (2) **26** (1982), 15–20.
23. K. Mahler, *Zur Approximation algebraischer Zahlen*. II. *Uber die Anzahl der Darstellungen ganzer Zahlen durch Binärformen*, Math. Ann. **108** (1933), 37–55.
24. ——, *On the lattice points on curves of genus* 1, Proc. London Math. Soc. (2) **39** (1935), 431–466.
25. ——, *An inequality for the discriminant of a polynomial*, Michigan Math. J. **11** (1964), 257–262.
26. ——, *On Thue's theorem*, Math. Scand. **55** (1984), 188–200.
27. L. J. Mordell, *Diophantine equations*, Academic Press, London and New York, 1969.
28. P. Moree and C. L. Stewart, *Some Ramanujan-Nagell equations with many solutions*, Nederl. Akad. Wetensch. Proc. Ser. A (N.S.) **1** (1990), 465–472.
29. J. Mueller, and W. M. Schmidt, *Thue's equation and a conjecture of Siegel*, Acta Math. **160** (1988), 207–247.
30. T. Nagell, *Généralisation d'un theórème de Tchebicheff*, J. Math. **8** (1921), 343–356.
31. M. Newman, *Integral matrices*, Pure and Appl. Math. (S. Eilenberg and P. A. Smith, eds.), vol. 45, Academic Press, New York, 1972.
32. O. Ore, *Anzahl der Wurzeln höherer Kongruenzen.*, Norsk Matematisk Tidsskrift, 3 Aagang, Kristiana (1921), 343–356.
33. G. Sándor, *Uber die Anzahl der Lösungen einer Kongruenz*, Acta. Math. **87** (1952), 13–17.

34. W. M. Schmidt, *Thue equations with few coefficients*, Trans. Amer. Math. Soc. **303** (1987), 241–255.

35. C. L. Siegel, *Die Gleichung* $ax^n - by^n = c$, Math. Ann. **114** (1937), 57–68.

36. J. H. Silverman, *Representation of integers by binary forms and the rank of the Mordell-Weil group*, Invent. Math. **74** (1983), 281–292.

37. ——, *Integer points on curves of genus* 1, J. London Math. Soc. (2) **28** (1983), 1–7.

38. ——, *The arithmetic of elliptic curves*, Graduate Texts in Math., vol. 106, Springer-Verlag, New York, 1986.

Department of Pure Mathematics, The University of Waterloo, Waterloo, Ontario Canada N2L 3G1