THUE EQUATIONS WITH FEW COEFFICIENTS

WOLFGANG M. SCHMIDT

ABSTRACT. Let $F(x,y)$ be a binary form of degree $r \geq 3$ with integer co-
efficients, and irreducible over the rationals. Suppose that only $s + 1$ of the
$r + 1$ coefficients of F are nonzero. Then the Thue equations $F(x,y) = 1$
has $\ll (rs)^{1/2}$ solutions. More generally, the inequality $|F(x,y)| \leq h$ has
$\ll (rs)^{1/2} h^{2/r} (1 + \log h^{1/r})$ solutions.

1. Introduction. Our goal is the proof of the following

THEOREM 1. *Let* $F(x,y) = a_0 x^r + a_1 x^{r-1} y + \cdots + a_r y^r$ *be a form of degree*
$r \geq 3$ *with integer coefficients which is irreducible over the rationals. Suppose that
not more than* $s + 1$ *of the coefficients* a_i *are nonzero. Then given* $h \geq 1$, *the
inequality*

$$(1.1) \qquad\qquad |F(x,y)| \leq h$$

has

$$(1.2) \qquad\qquad \ll (rs)^{1/2} h^{2/r} (1 + \log h^{1/r})$$

solutions in integers x, y.

Here and throughout, the constants implied by $\ll$ are absolute; they could be
explicitly given with some extra effort.

When working on this subject, I had initially only considered $F(x,y) = 1$ and I
showed that it has $\ll (rs)^{1/2}$ solutions. Since $s \leq r$, this bound yields the estimate
$\ll r$ which had recently been obtained by Bombieri and the author [1]. After being
repeatedly reminded by my colleagues that $F(x,y) = 1$ was a rather special Thue
equation, I turned to the general Thue equation $F(x,y) = h$ and I derived the
bound (1.2) for the number of solutions. On the other hand Bombieri and I had
obtained the bound

$$(1.3) \qquad\qquad \ll r^{1+\nu}$$

for the number of primitive solutions (i.e. solutions with g.c.d.$(x,y) = 1$), where ν
is the number of distinct prime factors of h. When r is given and h is large, then
(1.3) is better than (1.2), but when h is given and r is large, (1.2) is better. At
any rate, the bound (1.2), as a function of h, seems to be much too large for the
Thue *equation*. In fact it was not difficult to modify the proof (which is based on
the archimedean absolute value) to deal with the *inequality* (1.1) rather than the

equation $F(x, y) = h$. For the inequality, it is clear that the factor $h^{2/r}$ in (1.2) is needed. But the logarithmic factor is probably unnecessary.

It turns out that the number of solutions of (1.1) may be bounded in terms of s and h only. This had been shown for $s = 1$ (i.e. for binomial forms) by Mueller [3], for $s = 2$ (i.e. for trinomial forms) by Mueller and Schmidt [4], and in general very recently by Mueller and Schmidt [5]. However, the bound in [5] is not strong enough to imply the present theorem.

It would have been simpler in the formulation of Theorem 1 to suppose that F has not more than s nonzero coefficients (rather than $s + 1$), but for technical reasons, and to conform with other work, the notation as given above is preferable.

It is easily seen (e.g. in Lemma 1 below) that with F as in Theorem 1, the polynomial

$$(1.4) \qquad f(z) = F(z, 1)$$

has $\leq 2s$ real roots. The form $F = x^{2k} + c(x - y)^2(2x - y)^2 \cdots (kx - y)^2$ with $c > 0$ has no real roots but $F(x, y) = 1$ has the solutions $\pm(1, 1), \pm(1, 2), \ldots, \pm(1, k)$, so that the number of solutions may not be estimated in terms of the number of real roots.[1] What we will actually need in our present argument is that with F, f as above, f has few roots near the real axis.

The condition that "$f'(z)$ has few real roots" is stronger than the condition that $f(z)$ has few real roots. But it lacks symmetry; it is not even invariant under replacing $F(x, y)$ by $F(y, x)$. We will formulate a condition which is invariant under substitutions in $GL(2, \mathbf{Z})$. For $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ let us write F_A for the binary form

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

we also write $\mathbf{x}$ for (x, y) and $A\mathbf{x}$ for $(ax + by, cx + dy)$.

The zero set in $\mathbf{C}^2$ of a form $K(x, y)$ of positive degree consists of certain 1-dimensional subspaces of $\mathbf{C}^2$. The "number of real zeros of K" is now defined as the number of these subspaces which are defined over the reals, i.e. defined by equations $\alpha x + \beta y = 0$ where α, β are real and not both zero. (We do not count multiplicities.) This number is invariant under substitutions $A \in GL(2, \mathbf{R})$.

We now define a class $C(t)$ of forms of degree r as follows. It is the set of forms $F(x, y)$ of degree r with coefficients in $\mathbf{Z}$, and irreducible over $\mathbf{Q}$, such that for any real $u, v \neq 0, 0$ the form

$$(1.5) \qquad uF_x + vF_y$$

has at most t real zeros. Note that for $r > 0$, the irreducibility of F implies that the form (1.5) of degree $r - 1$ is not identically zero. Note also that for $F \in C(t)$, the derivative $f'(z) = F_x(z, 1)$ has $\leq t$ real zeros. The class $C(t)$ is closed under substitutions $A \in GL(2, \mathbf{Z})$. For when $G = F_A$, then

$$uG_x(\mathbf{x}) + vG_y(\mathbf{x}) = (ua + vb)F_x(A\mathbf{x}) + (uc + vd)F_y(A\mathbf{x})$$
$$= u_1 F_x(A\mathbf{x}) + v_1 F_y(A\mathbf{x}),$$

say.

We will show in Lemma 2 below that F as in Theorem 1 lies in $C(4s - 2)$. (But it may be seen that $C(t)$ in general also contains forms which are not "lacunary" such as the forms of Theorem 1.) At any rate, it follows that Theorem 1 is a consequence of the following theorem.

---

[1] This example was told to me by M. Waldschmidt.

THEOREM 2. *For F in C(t), the number of solutions of* (1.1) *is*

(1.6) $$\ll (rt)^{1/2} h^{2/r} (1 + \log h^{1/r}).$$

With f given by (1.4) write

(1.7) $$f(z) = a_0 (z - \alpha_1) \cdots (z - \alpha_r).$$

The Mahler height is

$$M(F) = M(f) = |a_0| \prod_{i=1}^{r} \max(1, |\alpha_i|).$$

Let $\operatorname{Im} \alpha_i$ be the imaginary part of α_i. When $M(F) > 1$, define Φ_i $(i = 1, \ldots, r)$ by

(1.8) $$\begin{cases} \Phi_i = 0 \quad \text{when } |\operatorname{Im} \alpha_i| > 1, \\ M(F)^{-\Phi_i} = |\operatorname{Im} \alpha_i| \quad \text{when } 0 < |\operatorname{Im} \alpha_i| \le 1, \\ \Phi_i = +\infty \quad \text{when } \alpha_i \text{ is real.} \end{cases}$$

Put

(1.9) $$S(F) = 1 + \sum_{i=1}^{r} \min(1, \Phi_i).$$

THEOREM 3. *Let C be a class of irreducible forms of degree $r \ge 3$ which is closed under substitutions in $GL(2, \mathbf{Z})$, i.e. $F_A \in C$ when $F \in C$ and $A \in GL(2, \mathbf{Z})$. Set*

(1.10) $$S(C) = \max S(F),$$

where the maximum is over forms F in C with

(1.11) $$M(F) > e^{2r}.$$

Then for forms F in C, the number of solutions (1.1) *is* $\ll S(C) h^{2/r} (1 + \log h^{1/r})$.

THEOREM 4. $S(C(t)) \ll (rt)^{1/2}$.

Clearly, Theorem 2 is an immediate consequence of Theorems 3, 4.

Our proofs have many points in common with [1]. But the treatment of (1.1) for general h is new, and the analysis of the location of roots given in §§8, 9 has no precedent in the earlier work.

2. The number of real roots of certain polynomials.

LEMMA 1. *Suppose $g(z)$ is polynomial with $g(0) \ne 0$ and with real coefficients, of which precisely $s + 1$ are nonzero. Then $g(z)$ has $\le 2s$ distinct real zeros.*

PROOF. When $s = 0$, $g(z)$ is a nonzero constant, and the assertion is correct. When $s > 0$, then $g'(z)$ does not vanish identically. We may write $g'(z) = z^m h(z)$ where h is a polynomial with $h(0) \ne 0$, having precisely s nonzero coefficients. By induction, h has $\le 2s - 2$ real zeros, so that g' has $\le 2s - 1$ real zeros, and g itself has $\le 2s$ real zeros.

LEMMA 2. *Suppose $F(\mathbf{x}) = a_0 x^r + \cdots + a_r y^r$ is irreducible and has $s+1$ nonzero coefficients. Then F lies in $C(4s - 2)$.*

PROOF. Given $u, v \neq 0, 0$, write

$$(2.1) \qquad\qquad u F_x + v F_y = x^k y^l Q(x, y)$$

with $Q(1, 0) \neq 0$, $Q(0, 1) \neq 0$. The number of nonzero coefficients of (2.1) is $\leq 2s$, and in fact the number of nonzero coefficients other than the coefficients of x^{r-1} or y^{r-1} is $\leq 2s - 2$. The number of real zeros of $Q(x, y)$ is the same as the number of real zeros of $g(z) = Q(z, 1)$. When $k = l = 0$, the number of real zeros of $g(z)$ is $\leq 2 \cdot (2s - 1) = 4s - 2$ by Lemma 1, so that the number of real zeros of (2.1) is $\leq 4s - 2$. When $k > 0$, $l = 0$ (or when $k = 0$, $l > 0$), the number of coefficients of g is $\leq 2s - 1$, so that the number of real zeros of $g(z)$ is $\leq 2 \cdot (2s - 2) = 4s - 4$, and the number of real zeros of (2.1) is $\leq (4s - 4) + 1 < 4s - 2$. When $k > 0$, $l > 0$, the number of real zeros of $g(z)$ is $\leq 2 \cdot (2s - 3) = 4s - 6$, and the number of real zeros of (2.1) is $\leq (4s - 6) + 2 < 4s - 2$.

3. Preliminaries to the proof of Theorem 3. Let C be a class of binary forms which is closed under substitutions from $GL(2, \mathbf{Z})$. Given a prime p, let C_p be the subclass of forms F in C whose discriminant $D(F)$ has

$$(3.1) \qquad\qquad |D(F)| \geq p^{r(r-1)}.$$

Since

$$(3.2) \qquad\qquad D(F_A) = |A|^{r(r-1)} D(F)$$

where $|A|$ is the determinant of A, the subclass C_p is again closed under substitutions forms $GL(2, \mathbf{Z})$. Given a finite set $\mathbf{N}$ of natural numbers, write $P(C, \mathbf{N})$ for the maximum number of primitive $\mathbf{x}$ with $F(\mathbf{x}) \in \mathbf{N}$, the maximum being taken over $F \in C$. Similarly define $P(C_p, \mathbf{N})$.

LEMMA 3. $P(C, \mathbf{N}) \leq (p + 1) P(C_p, \mathbf{N})$.

PROOF. We use an argument already employed in [1]. Let

$$A_0 = \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}, \qquad A_j = \begin{pmatrix} 0 & -1 \\ p & j \end{pmatrix} \qquad (j = 1, \ldots, p).$$

It is easily seen that $\mathbf{Z}^2 = \bigcup_{j=0}^{p} A_j \mathbf{Z}^2$. Therefore the number of primitive $\mathbf{x}$ with $F(\mathbf{x}) \in \mathbf{N}$ does not exceed $n_0 + n_1 + \cdots + n_p$, where n_j is the number of primitive $\mathbf{x}$ with $F_{A_j}(\mathbf{x}) \in \mathbf{N}$. But when F lies in C, then F_{A_j} lies in C_p by (3.2).

Write $N(F, h)$ for the number of solutions of (1.1) and $P(F, h)$ for the number of primitive solutions of (1.1). Further write $P_1(F, h)$ for the number of primitive solutions of

$$(3.3) \qquad\qquad 2^{-r} h \leq F(\mathbf{x}) < h.$$

Now let C be a class of forms which is closed under substitutions from $GL(2, \mathbf{Z})$. Let $N(C, h)$, $P(C, h)$, $P_1(C, h)$ respectively be the maximum of $N(F, h)$, $P(F, h)$, $P_1(F, h)$ over forms $F \in C$.

Forms F, G will be considered equivalent, with the notation $F \sim G$, if $G = F_A$ with $A \in SL(2, \mathbf{Z})$. Now $P_1(F, h)$ is not affected if we replace F by an equivalent

form; nor is the discriminant affected. When (3.3) has at least one primitive solution $\mathbf{x}^0$, there is an $A \in SL(2, \mathbf{Z})$ with $A^{-1}\mathbf{x}^0 = (1, 0)$, so that $2^{-r}h \leq F_A(1, 0) < h$. Therefore in order to estimate $P_1(C, h)$, we may restrict ourselves to forms $F \in C$ which are *normalized* in the sense that $2^{-r}h \leq F(1, 0) < h$, i.e. their leading coefficient a_0 has $2^{-r}h \leq a_0 < h$. We will say that a form F is *reduced* if it is normalized and has smallest Mahler height among all normalized forms equivalent to F. Every form with a primitive solution to (3.3) is equivalent to at least one reduced form.

PROPOSITION. *Suppose F lies in C, is reduced and has Mahler height*

(3.4) $$M(F) \geq 100^r h.$$

Then

$$P_1(F, h) \ll S(C)(1 + \log h^{1/r}).$$

This Proposition will be proved in §§4, 5, and 6. Here we will deduce Theorem 3. Pick a prime

(3.5) $$p > 10^6 h^{2/r}.$$

Then $F \in C_p$ has $|D(F)| \geq p^{r(r-1)}$. Now if $F_1 \sim F$ is reduced, we have

$$|D(F)| = |D(F_1)| \leq r^r M(F_1)^{2r-2} < (rM(F_1))^{2r-2}$$

by an inequality of Mahler [2], so that

$$M(F_1) > r^{-r}p^{r/2} > r^{-1}10^{3r}h > 100^r h.$$

By the Proposition,

$$P_1(C_p, h) \ll S(C)(1 + \log h^{1/r}).$$

Let p be the least prime with (3.5); then Lemma 3 gives

$$P_1(C, h) \ll h^{2/r}P_1(C_p, h) \ll S(C)h^{2/r}(1 + \log h^{1/r}).$$

Now, when u is the integer with $2^{ru} \leq h < 2^{r(u+1)}$, then

$$P(C, h) \leq P(C, 2^{r(u+1)} - 1) = \sum_{j=1}^{u+1} P_1(C, 2^{rj})$$

$$\ll S(C) \sum_{j=1}^{u+1} 2^{2j}(1 + \log 2^j)$$

$$\ll S(C)2^{2u}(1 + u)$$

(3.6) $$\leq S(C)h^{2/r}(1 + \log h^{1/r}).$$

Given $F \in C$, let $\pi(F, h)$ be the number of primitive solutions of $F(\mathbf{x}) = h$. Then

$$\pi(F, h) = P(F, h) - P(F, h - 1)$$

(with $P(F, 0) = 0$ and with $[\ ]$ denoting integer parts). We have

$$N(F, h) = \sum_{n=1}^{h} \pi(F, n) \left[\left(\frac{h}{n} \right)^{1/r} \right] \le h^{1/r} \sum_{n=1}^{h} \pi(F, n) n^{-1/r}$$

$$= h^{1/r} \sum_{n=1}^{h} (P(F, h) - P(F, n - 1)) n^{-1/r}$$

$$= h^{1/r} \left(\sum_{n=1}^{h-1} P(F, n)(n^{-1/r} - (n + 1)^{-1/r}) + P(F, h) h^{-1/r} \right)$$

$$\ll h^{1/r} \left(\sum_{n=1}^{h-1} P(F, n) r^{-1} n^{-(1/r)-1} + P(F, h) h^{-1/r} \right).$$

In view of (3.6) we obtain

$$N(F, h) \ll h^{1/r} S(C) \left(r^{-1} \sum_{n=1}^{h-1} n^{(1/r)-1} + h^{1/r} \right) (1 + \log h^{1/r})$$

$$\ll S(C) h^{2/r} (1 + \log h^{1/r}).$$

4. Large and small solutions. Let $|\mathbf{x}| = \max(|x|, |y|)$ be the *maximum norm* of points $\mathbf{x} = (x, y)$. Given F with $M(F) = M$, set

(4.1) $$K = (2r^{1/2} M)^r h.$$

With F normalized and of the form

$$F = a_0(x - \alpha_1 y) \cdots (x - \alpha_r y),$$

every solution of (1.1) or (3.3) with $y \ne 0$ has

(4.2) $$\min_{1 \le i \le r} \min \left(1, \left| \alpha_i - \frac{x}{y} \right| \right) \le \frac{K}{|\mathbf{x}|^r}$$

according to [1, Lemma 1].

Fix numbers a, b with $0 < a < b < 1$ and so small that

(4.3) $$\sqrt{2}\sqrt{3 + a^2}/(1 - b) < 3.$$

Put

(4.4) $$t = \sqrt{2/(r + a^2)}, \qquad \lambda = 2/((1 - b)t),$$

so that by (4.3)

$$\lambda = \sqrt{2}\sqrt{r + a^2}/(1 - b) < r$$

for $r \ge 3$. Set

(4.5) $$A = \frac{t^2}{2 - rt^2} \left(\log M + \frac{1}{2} r \right),$$

(4.6) $$Y = (2K)^{1/(r-\lambda)} (4e^A)^{\lambda/(r-\lambda)}.$$

Henceforth, solutions to (3.3) with $|\mathbf{x}| \ge Y$ will be called *large*, solutions with $|\mathbf{x}| < Y$ will be called *small*.

In view of (4.2), every large solution with $y \neq 0$ will have

$$(4.7) \qquad |\alpha_i - x/y| < (4e^A |\mathbf{x}|)^{-\lambda}$$

for some α_i. Furthermore, by [1, §3, upper bound for n], the number of primitive $\mathbf{x}$ with $|\mathbf{x}| \geq Y$ and (4.7) for some *given* α_i is bounded by an absolute constant. But (4.7) yields

$$|\operatorname{Im} \alpha_i| \leq |\alpha_i - x/y| < |\mathbf{x}|^{-\lambda} \leq y^{-\lambda} < M^{-\lambda r/(r-\lambda)} < M^{-1},$$

so that $\Phi_i \geq 1$ by the definition (1.8). The number of such i is bounded by $S(C)$. Therefore *the number of large solutions is* $\ll S(C)$.

5. Estimation of linear forms. Let F be as in the Proposition, in particular reduced and with (3.4), so that

$$(5.1) \qquad Q := M/h \geq 100^r.$$

We have

$$(5.2) \qquad F(\mathbf{x}) = a_0 L_1(\mathbf{x}) \cdots L_r(\mathbf{x})$$

with $L_i(\mathbf{x}) = x - \alpha_i y$.

LEMMA 4. *Suppose* $G(\mathbf{x}) = b_0(x - \beta_1 y) \cdots (x - \beta_r y)$ *is normalized and equivalent to F, and let*

$$(5.3) \qquad \eta_i = |\beta_i - m| + 1 \qquad (i = 1, \ldots, r)$$

where m is an integer. Then

$$(5.4) \qquad \eta_1 \cdots \eta_r > Q.$$

PROOF.

$$\hat{G}(\mathbf{x}) = G(x + my, y) = b_0 \prod_i (x + (m - \beta_i)y)$$

is normalized and equivalent to G, hence to F. Since F is reduced, $M = M(F) \leq M(\hat{G})$. On the other hand it is clear that $M(\hat{G}) < |b_0|\eta_1 \cdots \eta_r < h\eta_1 \cdots \eta_r$. The lemma follows.

Given $\mathbf{x}_0 = (x_0, y_0)$, $\mathbf{x} = (x, y)$, put $D(\mathbf{x}_0, \mathbf{x}) = x_0 y - y_0 x$.

LEMMA 5. *Suppose F is as above, and $\mathbf{x}_0$, $\mathbf{x}$ are linearly independent primitive integer points with (3.3). Then there are numbers $\psi_1, \ldots, \psi_r$, where each*

$$(5.5) \qquad \psi_i = 0 \quad \text{or has} \quad 1/2r \leq \psi_i \leq 1,$$

with

$$(5.6) \qquad \sum_{i=1}^r \psi_i \geq \frac{1}{2},$$

such that

$$(5.7) \qquad |L_i(\mathbf{x}_0)/L_i(\mathbf{x})| \geq (Q^{\psi_i} - 7/2)|D(\mathbf{x}_0, \mathbf{x})| \qquad (i = 1, \ldots, r).$$

PROOF. Pick $\mathbf{x}' \in \mathbf{Z}^2$ with $D(\mathbf{x}', \mathbf{x}) = 1$, so that $\mathbf{x}'$, $\mathbf{x}$ is a basis of $\mathbf{Z}^2$. We may write $\mathbf{x}_0 = a\mathbf{x} + b\mathbf{x}'$. Then $D(\mathbf{x}_0, \mathbf{x}) = bD(\mathbf{x}', \mathbf{x}) = b$, so that $\mathbf{x}_0 = a\mathbf{x} + D(\mathbf{x}_0, \mathbf{x})\mathbf{x}'$. Therefore

$$(5.8) \qquad \frac{L_i(\mathbf{x}_0)}{L_i(\mathbf{x})} = a + D(\mathbf{x}_0, \mathbf{x})\frac{L_i(\mathbf{x}')}{L_i(\mathbf{x})} = a - D(\mathbf{x}_0, \mathbf{x})\beta_i,$$

say $(i = 1, \ldots, r)$.

Set $G(v, w) = F(v\mathbf{x} + w\mathbf{x}')$, so that $G \sim F$ and G is normalized. Moreover,

$$G(v, w) = a_0 \prod_{i=1}^{r} L_i(v\mathbf{x} + w\mathbf{x}') = a_0 \prod_{i=1}^{r} (vL_i(\mathbf{x}) + wL_i(\mathbf{x}'))$$

$$= b_0 \prod_{i=1}^{r} \left(v + \frac{L_i(\mathbf{x}')}{L_i(\mathbf{x})} w \right) = b_0 \prod_{i=1}^{r} (v - \beta_i w)$$

with $b_0 = F(\mathbf{x})$. Note that $0 < F(\mathbf{x}_0)/F(\mathbf{x}) \leq 2^r$. We may suppose that

$$|L_r(\mathbf{x}_0)/L_r(\mathbf{x})| = \min_i |L_i(\mathbf{x}_0)/L_i(\mathbf{x})|,$$

so that $|L_r(\mathbf{x}_0)/L_r(\mathbf{x})| \leq 2$, and $|a - D(\mathbf{x}_0, \mathbf{x})\beta_r| \leq 2$ by (5.8), which in turn yields

(5.9) $$|a - D(\mathbf{x}_0, \mathbf{x})\beta| \leq 2,$$

where β is the real part of β_r. Let m be an integer with $|m - \beta| \leq \frac{1}{2}$ and define $\eta_1, \ldots, \eta_r$ by (5.3), so that (5.4) holds by Lemma 4. Put

$$\eta_i' = \begin{cases} Q & \text{if } \eta_i \geq Q, \\ \eta_i & \text{if } Q^{1/2r} \leq \eta_i < Q, \\ 1 & \text{if } \eta_i < Q^{1/2r}, \end{cases}$$

so that $\eta_1' \cdots \eta_r' \geq Q^{1/2}$. The numbers ψ_i defined by $\eta_i' = Q^{\psi_i}$ $(i = 1, \ldots, r)$ satisfy (5.5), (5.6). Now from (5.8), (5.9),

$$|L_i(\mathbf{x}_0)/L_i(\mathbf{x})| = |(\beta - \beta_i)D(\mathbf{x}_0, \mathbf{x}) + a - D(\mathbf{x}_0, \mathbf{x})\beta|$$

$$\geq |\beta - \beta_i||D(\mathbf{x}_0, \mathbf{x})| - 2 \geq (|m - \beta_i| - \tfrac{1}{2})|D(\mathbf{x}_0, \mathbf{x})| - 2$$

$$= (\eta_i - \tfrac{3}{2})|D(\mathbf{x}_0, \mathbf{x})| - 2 \geq (\eta_i - \tfrac{7}{2})|D(\mathbf{x}_0, \mathbf{x})|.$$

Since $\eta_i \geq \eta_i' = Q^{\psi_i}$, the assertion follows.

LEMMA 6. *Suppose F is as above, and $\mathbf{x}$ is primitive with $y > 0$ and (3.3). Then there are numbers $\psi_i = \psi_i(\mathbf{x})$ $(i = 1, \ldots, r)$ with (5.5), (5.6) such that*

(5.10) $$|L_i(\mathbf{x})| < 1/(Q^{\psi_i/2}y)$$

for each i with $\psi_i > 0$.

PROOF. We first note that for i with $\psi_i > 0$ we have $\psi_i \geq 1/2r$ and $Q^{\psi_i} \geq Q^{1/2r} > 7$ by (5.1), so that $Q^{\psi_i} - \frac{7}{2} \geq \frac{1}{2}Q^{\psi_i} \geq Q^{\psi_i/2}$. We now apply Lemma 5 with $\mathbf{x}_0 = (1, 0)$. Then $D(\mathbf{x}_0, \mathbf{x}) = y$ and $L_i(\mathbf{x}_0) = 1$ $(i = 1, \ldots, r)$, so that (5.10) is an immediate consequence of (5.7).

6. Proof of the Proposition and of Theorem 3.

LEMMA 7. *Make the same hypotheses as in the last section, including (5.1). Let $\mathfrak{X}$ be the set of primitive integer points with (3.3) and with $0 < y \leq Y$. Then for $i = 1, \ldots, r$,*

(6.1) $$\sum_{\mathbf{x} \in \mathfrak{X}} \psi_i(\mathbf{x}) \ll (1 + \log h^{1/r}) \min(1, \Phi_i).$$

PROOF. Given i, let $\mathbf{x}_1, \ldots, \mathbf{x}_\nu$ be the elements of $\mathfrak{X}$ with $\psi_i(\mathbf{x}) > 0$, ordered such that $y_1 \leq \cdots \leq y_\nu$. Now (5.10) yields

$$|\operatorname{Im} \alpha_i| \leq |\alpha_i - x_j/y_j| < 1/(Q^{\psi_i(\mathbf{x}_j)/2}y_j^2)$$

for $j = 1, \ldots, \nu$. When $|\operatorname{Im} \alpha_i| > 1$ this is impossible and the sum of the lemma is zero. When $|\operatorname{Im} \alpha_i| \leq 1$, so that $|\operatorname{Im} \alpha_i| = M^{-\Phi_i}$ (with $M^{-\infty} = 0$), we have

$$M^{-\Phi_i} < 1/(Q^{\psi_i(\mathbf{x}_j)/2} y_j^2),$$

and therefore

(6.2) $$y_j < M^{\Phi_i/2}, \quad \psi_i(\mathbf{x}_j) < 2\frac{\log M}{\log Q} \Phi_i.$$

In particular, $y_\nu \leq \min(Y, M^{\Phi_i/2}) = Y_i$, say.

When $\nu > 1$ and $1 \leq j < \nu$, we have

$$|L_i(\mathbf{x}_j)| = |x_j - \alpha_i y_j| < 1/(Q^{\psi_i(\mathbf{x}_j)/2} y_j),$$

$$|L_i(\mathbf{x}_{j+1})| = |x_{j+1} - \alpha_i y_{j+1}| < 1/(Q^{\psi_i(\mathbf{x}_{j+1})/2} y_{j+1}).$$

Since $D(\mathbf{x}_j, \mathbf{x}_{j+1}) \neq 0$, we may infer that

$$1 \leq |D(\mathbf{x}_j, \mathbf{x}_{j+1})| < \frac{y_{j+1}}{y_j Q^{\psi_i(\mathbf{x}_j)/2}} + \frac{y_j}{y_{j+1} Q^{\psi_i(\mathbf{x}_{j+1})/2}} < \frac{y_{j+1}}{y_j Q^{\psi_i(\mathbf{x}_j)/2}} + \frac{1}{3},$$

since $Q^{\psi_i(\mathbf{x}_{j+1})/2} \geq Q^{1/4r} > 3$ by (5.1). Therefore ("gap principle")

$$y_{j+1} > \tfrac{2}{3} Q^{\psi_i(\mathbf{x}_j)/2} y_j > Q^{\psi_i(\mathbf{x}_j)/4} y_j \quad (1 \leq j < \nu),$$

again using (5.1). We may conclude that

$$Q^{(1/4)(\psi_i(\mathbf{x}_1) + \cdots + \psi_i(\mathbf{x}_{\nu-1}))} < y_\nu \leq Y_i$$

and

(6.3) $$\sum_{j=1}^{\nu-1} \psi_i(\mathbf{x}_j) < 4\frac{\log Y_i}{\log Q}.$$

Now λ as given by (4.4) has $r^{1/2} \ll \lambda \ll r^{1/2}$, and A given by (4.5) is $A = a^{-2}(\log M + \tfrac{1}{2}r) \ll \log M$ by (5.1), and Y given by (4.6) has

$$\log Y \ll r^{-1}(\log 2C) + r^{-1/2}(A + \log 4).$$

But by (4.1), (5.1), $\log 2C \ll r(\log M + \log r) + \log h \ll r \log M$, so that in fact $\log Y \ll \log M$. From the definition of Y_i,

$$\log Y_i \ll \min(1, \Phi_i) \log M.$$

Note that by (5.1)

$$\frac{\log M}{\log Q} = 1 + \frac{\log h}{\log Q} \ll 1 + r^{-1} \log h = 1 + \log h^{1/r}.$$

Since by (5.5), (6.2),

$$\psi_i(\mathbf{x}_\nu) \ll \min\left(1, \frac{\log M}{\log Q}\Phi_i\right) \leq \min(1, \Phi_i)\frac{\log M}{\log Q},$$

we may infer from (6.3) that

$$\sum_{j=1}^{\nu} \psi_i(\mathbf{x}_j) \ll \frac{\log Y_i}{\log Q} + \frac{\log M}{\log Q} \min(1, \Phi_i) \ll \frac{\log M}{\log Q} \min(1, \Phi_i)$$

$$\ll (1 + \log h^{1/r}) \min(1, \Phi_i).$$

PROOF OF THE PROPOSITION. In view of what we have said in §4, we need to show that the number of small solutions is $\ll (1 + \log h^{1/r})S(C)$. But small solutions have $|y| \leq Y$. By (5.6), (6.1)

$$\sum_{\mathbf{x} \in \mathfrak{X}} 1 \leq 2 \sum_{i=1}^{r} \sum_{\mathbf{x} \in \mathfrak{X}} \psi_i(\mathbf{x})$$

$$\ll (1 + \log h^{1/r}) \sum_{i=1}^{r} \min(1, \Phi_i) \leq (1 + \log h^{1/r})S(C).$$

(We used that F has (3.4), hence (1.11).) Small solutions with $y < 0$ are counted similarly. The only other small solutions are $(1, 0)$ and possibly (when r is even) $(-1, 0)$.

This proves the Proposition and thus Theorem 3.

7. A general lemma on roots of polynomials.

LEMMA 8. *Let* $f(z) = a_0(z - \alpha_1) \cdots (z - \alpha_r)$ *be a polynomial with integer coefficients and without multiple roots; let* M *be its Mahler height. Given a set* S *of pairs* (i, j) *with* $1 \leq i < j \leq r$, *put*

$$P(S) = \prod_{(i,j) \in S} (\alpha_i - \alpha_j).$$

Then

(7.1) $$|P(S)| > 2^{-r^2} M^{1-r}.$$

PROOF. Let T be the complement of S in the set of pairs and put

$$P^*(T) = a_0^{r-1} \prod_{(i,j) \in T} (\alpha_i - \alpha_j).$$

Then $P^*(T)$ is the sum of 2^t terms of the type

(7.2) $$\pm a_0^{r-1} \alpha_{i_1} \cdots \alpha_{i_t}$$

where t is the cardinality of T. Any particular α_i can occur at most $r - 1$ times in the product (7.2), so that by the definition of Mahler height, the product (7.2) has modulus $\leq M^{r-1}$. Therefore

(7.3) $$|P^*(T)| \leq 2^t M^{r-1} \leq 2^{\binom{r}{2}} M^{r-1} < 2^{r^2} M^{r-1}.$$

But $|P(S)||P^*(T)| = |D|^{1/2} \geq 1$ where D is the discriminant, and (7.1) follows from (7.3).

8. The clustering of roots with small imaginary parts. The purpose of this section is to show that if $f(x)f'(x) \neq 0$ for real x in some interval $X_1 < x < X_2$, and if there are many roots with real parts in this interval, and with small imaginary parts, then there are many such roots which are clustered together, i.e. whose mutual distances are small.

LEMMA 9. *Let $f(z)$ be a polynomial of degree r with real coefficients. Let $z_j = x_j + iy_j$ $(j = 1, \ldots, w)$ be roots of f lying in a square $a - \varepsilon \le x \le a$, $0 < y \le \varepsilon$. Suppose $f(x)f'(x) < 0$ for every real x in the interval $a + \varepsilon \le x < a + 8r\varepsilon$. Then there are $\ge w/120(\log w + 1)$ roots in the square*

(8.1)
$$a < x < a + 8r\varepsilon, \qquad 0 < y < 8r\varepsilon.$$

PROOF. Let q be a real number in $a + \varepsilon \le q \le a + 2\varepsilon$, which will initially be fixed. Let $z_j = x_j + iy_j$ $(j = 1, \ldots, r)$ be all the roots of f, ordered such that $z_{j+w} = \bar{z}_j$ $(j = 1, \ldots, w)$, where the bar denotes complex conjugation. We further may suppose that $x_j < q$ for $j = 1, \ldots, u$ and $x_j \ge q$ for $j = u + 1, \ldots, r$; then clearly $u \ge 2w$.

Our hypothesis implies that

(8.2)
$$f'(q)/f(q) < 0.$$

Here

$$\frac{f'(q)}{f(q)} = \sum_{j=1}^{r} \frac{1}{q - z_j} = \frac{1}{2} \sum_{j=1}^{r} \left(\frac{1}{q - z_j} + \frac{1}{q - \bar{z}_j} \right) = \sum_{j=1}^{r} \frac{q - x_j}{(q - x_j)^2 + y_j^2}.$$

For $j = 1, \ldots, 2w$ we have $|y_j| \le \varepsilon \le q - a \le q - x_j \le 3\varepsilon$, so that

$$\frac{q - x_j}{(q - x_j)^2 + y_j^2} \ge \frac{q - x_j}{2(q - x_j)^2} \ge \frac{1}{6\varepsilon} \qquad (j = 1, \ldots, 2w).$$

Since the summands with $j \le u$ are positive,

$$\sum_{j=1}^{u} \frac{q - x_j}{(q - x_j)^2 + y_j^2} \ge \sum_{j=1}^{2w} \frac{q - x_j}{(q - x_j)^2 + y_j^2} \ge \frac{2w}{6\varepsilon} = \frac{w}{3\varepsilon}.$$

In conjunction with (8.2) this yields

$$\sum_{j=u+1}^{r} \frac{x_j - q}{(x_j - q)^2 + y_j^2} > \frac{w}{3\varepsilon},$$

so that

$$\sum_{j=u+1}^{r} \frac{1}{|z_j - q|} > \frac{w}{3\varepsilon},$$

where (of course) $|z_j - q| = ((x_j - q)^2 + y_j^2)^{1/2}$. Since there are fewer than r summands, the summands with $|z_j - q| > 6\varepsilon r/w$ contribute less than $w/6\varepsilon$, so that

(8.3)
$$\sum_{\substack{j=u+1 \\ |z_j - q| \le 6\varepsilon r/w}}^{r} \frac{1}{|z_j - q|} > \frac{w}{6\varepsilon}.$$

Put

(8.4)
$$l = [\log w] + 1.$$

Let I_0 be the interval $0 < \xi < \varepsilon/w$; for $k = 1, \ldots, l$ let I_k be the interval $(\varepsilon/w)e^{k-1} \le \xi < (\varepsilon/w)e^k$; and finally let I_{l+1} be the interval $(\varepsilon/w)e^l \le \xi < 6r\varepsilon/w$.

(Since $e^l \leq ew < 6r$, the last interval is not empty.) There is by (8.3) a k in $0 \leq k \leq l+1$ with

$$\sum_{\substack{j=u+1 \\ |z_j-q| \in I_k}}^{r} \frac{1}{|z_j - q|} > \frac{w}{6(l+2)\varepsilon} \geq \frac{w}{18l\varepsilon}.$$

When $k = 0$, there is a z_j with $x_j \geq q$ and $|z_j - q| < \varepsilon/w$. When $1 \leq k \leq l$, then each summand is $< (w/\varepsilon)e^{1-k}$, so that the number of summands is $> e^{k-1}/18l > e^k/60l$. Thus in general when $0 \leq k \leq l$, there are

(8.5)   *more than $e^k/60l$ roots z_j with $x_j \geq q$ and $|z_j - q| < (\varepsilon/w)e^k$.*

When $k = l+1$, there are $> e^l/60l \geq w/60(\log w + 1)$ roots with $x_j \geq q$ and $|z_j - q| < 6r\varepsilon/w \leq 6r\varepsilon$.

In each case, half of the roots in question will have positive imaginary parts. Thus if $k = l+1$ for some q in $a + \varepsilon \leq q \leq a + 2\varepsilon$, there will be $\geq w/120(\log w + 1)$ roots in the square $a < x < a + 8r\varepsilon$, $0 < y < 8r\varepsilon$. We may therefore suppose that for each q in $a + \varepsilon \leq q \leq a + 2\varepsilon$, there is a $k = k(q)$ in $0 \leq k \leq l$ with (8.5).

We set $q_0 = a + \varepsilon$, and when q_t lies in $a + \varepsilon \leq q_t \leq a + 2\varepsilon$, we set

$$q_{t+1} = q_t + (\varepsilon/w)e^{k(q_t)}.$$

We obtain $q_0 < q_1 < \cdots < q_\nu \leq a + 2\varepsilon < q_{\nu+1}$, say. For $t = 0, 1, \ldots, \nu$, the square

$$Q_t: q_t \leq x < q_t + (\varepsilon/w)e^{k(q_t)} = q_{t+1}, \qquad 0 < y < q_{t+1} - q_t,$$

contains

$$\geq e^{k(q_t)}/120l = (w/120l\varepsilon)(q_{t+1} - q_t)$$

roots. The squares $Q_0, \ldots, Q_\nu$ are pairwise disjoint, and they lie in the square (8.1), since

$$q_t + (\varepsilon/w)e^{k(q_t)} \leq a + 2\varepsilon + (\varepsilon/w)e^l \leq a + 2\varepsilon + (\varepsilon/w)ew < a + 8r\varepsilon.$$

Thus (8.1) contains

$$\geq \frac{w}{120l\varepsilon} \sum_{t=0}^{\nu}(q_{t+1} - q_t) \geq \frac{w}{120l} \geq \frac{w}{120(\log w + 1)}$$

roots of f.

LEMMA 10. *Let $f(z)$ be a polynomial of degree r with real coefficients. Let $\alpha = A + Bi$ be a root with $B > 0$. Suppose that $f(x)f'(x) < 0$ for every real x in the interior of*

(8.6) $A \leq x < A + (9r)^h B,$

where h is a natural number. Then f has at least $e^{\sqrt{h}/16}$ roots in the square consisting of numbers $z = x + iy$ with (8.6) and with $0 < y \leq (9r)^h B$.

PROOF. For $m = 0, 1, \ldots, h$ let P_m be the square

$$P_m: A \leq x < A + (9r)^m B, \qquad 0 < y \leq (9r)^m B,$$

and let w_m be the number of roots in this square. Then $w_0 \geq 1$ by hypothesis. Given $m < h$ we apply Lemma 9 with $a = A + (9r)^m B$, $w = w_m$, and with $(9r)^m B$ in place of ε. The conclusion is that f has $\geq w_m/(120 \log ew_m)$ roots in the square

$$A + (9r)^m B < x < A + (9r)^m B + 8r(9r)^m B, \qquad 0 < y < 8r(9r)^m B.$$

This square is disjoint from P_m and is contained in P_{m+1}. Thus

$$w_{m+1} \geq w_m + \frac{w_m}{120 \log ew_m}.$$

We have $w_0 \geq 1$, $w_1 \geq 2$, $w_2 \geq 3$. For $m \geq 2$,

$$(\log w_{m+1})^2 \geq \left(\log w_m + \log \left(1 + \frac{1}{120 \log ew_m} \right) \right)^2$$

$$> \left(\log w_m + \frac{1}{240 \log ew_m} \right)^2$$

$$> (\log w_m)^2 + \frac{1}{120} \frac{\log w_m}{\log ew_m}$$

$$> (\log w_m)^2 + \frac{1}{240}.$$

Therefore $(\log w_m)^2 \geq m/240$ $(m = 0, 1, \ldots, h)$, and $w_h > e^{\sqrt{h}/16}$.
   Put

$$(8.7) \qquad\qquad\qquad R = e^{800(\log r)^3}.$$

LEMMA 11. *Suppose $f(z)$ is a polynomial of degree $r \geq 3$ with real coefficients. Suppose there are $u \geq 2$ roots $z = x + iy$ with $X_1 \leq x \leq X_2$, $0 < y < \varepsilon$. Finally suppose that $f(x)f'(x) \neq 0$ in $X_1 < x < X_2$. Then there are u roots with $X_1 \leq x \leq X_2$, $0 < y < R\varepsilon$, and with mutual distances $< 2R\varepsilon$.*

PROOF. Let the given roots be $z_j = x_j + iy_j$ $(j = 1, \ldots, u)$ with $x_1 \leq \cdots \leq x_u$. We may suppose that $X_1 = x_1$, $X_2 = x_u$. The assertion is obvious if $x_u - x_1 \leq R\varepsilon$; suppose then that $x_u - x_1 > R\varepsilon$. Let h be the smallest integer with $e^{\sqrt{h}/16} \geq u$; then

$$256 \log^2 u \leq h < 256 \log^2 u + 1 \leq 256 \log^2 r + 1 < 257 \log^2 r$$

and

$$(9r)^h \varepsilon < (9r)^{257 \log^2 r} \varepsilon < R\varepsilon < x_u - x_1.$$

We may suppose without loss of generality that $f(x)f'(x) < 0$ in $x_1 < x < x_u$. By Lemma 10 with $\alpha = z_1$ and with $B < \varepsilon$, there are at least u roots in the square $x_1 \leq x < x_1 + R\varepsilon$, $0 < y < R\varepsilon$.

9. Proof of Theorem 4.

LEMMA 12. *Let $f(z)$ be a polynomial of degree r with rational integer coefficients, of Mahler height M and without multiple roots. Suppose that $f(x)f'(x)$ has not more than $q - 1$ real roots, where $q \geq 1$. Suppose f has u roots with imaginary parts in*

$$(9.1) \qquad\qquad\qquad 0 < y < 1/K$$

where

(9.2) $$K > 36R^2 = 36e^{1600 \log^3 r}.$$

Then

(9.3) $$u \leq \max\left(q, \left(\frac{2r^2q + 2rq \log M}{\log K}\right)^{1/2}\right).$$

PROOF. We may suppose that $u > q$. There are q intervals $J_k: X_k \leq x < Y_k$ $(k = 1, \ldots, q)$ with ff' not vanishing in $X_k < x < Y_k$, and such that the real parts of the u roots in question lie in the union of these intervals. There is a partition

(9.4) $$u = u(1) + \cdots + u(q)$$

such that for $k = 1, \ldots, q$ there are $u(k)$ roots with real parts in J_k and imaginary parts in (9.1). By Lemma 11 there are then $u(k)$ roots with real parts in J_k, with imaginary parts in $0 < y < R/K$ and mutual distances $< 2R/K$. If we consider their complex conjugates as well, we obtain $2u(k)$ roots $\alpha_{k1}, \ldots, \alpha_{k,2u(k)}$ with real parts in J_k and with mutual distances $< 6R/K < K^{-1/2}$ by (9.2). The number of pairs i, j with $1 \leq i < j \leq 2u(k)$ is $\binom{2u(k)}{2}$, so that

$$P := \prod_{k=1}^{q} \prod_{1 \leq i < j \leq 2u(k)} |\alpha_{ki} - \alpha_{kj}| < \prod_{k=1}^{q} K^{-\binom{2u(k)}{2}/2}.$$

But by (9.4), and the fact that $\binom{2x}{2} = 2x^2 - x$ has a positive second derivative,

$$\sum_{k=1}^{q} \binom{2u(k)}{2} \geq q\binom{2u/q}{2} = \frac{2u^2}{q} - u > \frac{u^2}{q},$$

since $u > q$. We obtain $P < K^{-u^2/2q}$. Comparison with (7.1) gives $e^{-r^2}M^{-r} < K^{-u^2/2q}$, so that

$$u^2 \log K < 2r^2q + 2rq \log M,$$

whence (9.3).

COROLLARY. *Make the same assumption on $f(z)$ as in Lemma 12. Suppose further that $M > e^{2r}$. Then for ϕ in*

(9.5) $$1700r^{-1}(\log r)^3 \leq \phi \leq 1,$$

the number of roots with imaginary parts in $0 < y \leq M^{-\phi}$ is $\leq (8rq/\phi)^{1/2}$.

PROOF. Apply Lemma 12 with

$$K = M^{\Phi/2} > e^{r\Phi} > e^{1700 \log^3 r} > 16R^2.$$

The number u of roots with imaginary part in (9.1) satisfies

$$u \leq \max(q, (4rq \log M/\log K)^{1/2}) = \max(q, (8rq/\Phi)^{1/2}).$$

Since the conclusion of the corollary is obvious when $q > r$, we may suppose that $q \leq r$, and then $q < (8rq/\Phi)^{1/2}$.

The proof of Theorem 4 is now accomplished as follows. Suppose $F \in C(t)$ with $M(F) > e^{2r}$. Then $f(z) = F(z, 1)$ has $f'(z) = 1 \cdot F_x(z, 1)$, so that $f'(z)$ has $\leq t$ real roots, and ff' has $\leq 2t + 1 = q - 1$ real roots with $q = 2t + 2$.

Let $\Phi_1, \ldots, \Phi_r$ be defined by (1.7); we may suppose that $\Phi_1 \geq \cdots \geq \Phi_r$. The number of summands in

$$(9.6) \qquad \sum_{i=1}^{r} \min(1, \Phi_i)$$

with $\Phi_i \geq 1$, i.e. with $|\operatorname{Im} \alpha_i| \leq M^{-1}$, is $\ll (rq)^{1/2} \ll (rt)^{1/2}$ by the corollary. These summands contribute $\ll (rt)^{1/2}$. Summands with $\Phi_i \leq 1700r^{-1}(\log r)^3$ contribute $\ll (\log r)^3 \ll (rt)^{1/2}$. (Since $C(0)$ is easily seen to be empty, we may suppose that $t \geq 1$.) The remaining summands have $1700r^{-1}(\log r)^3 < \Phi_i < 1$. Since $|\operatorname{Im} \alpha_j| = M^{-\Phi_j} \leq M^{-\Phi_i}$ for $j \leq i$, the corollary yields $i < (rt/\Phi_i)^{1/2}$, so that $\Phi_i \ll rt/i^2$. We may conclude that these terms contribute

$$\ll \sum_{i=1}^{\infty} \min(1, rt/i^2) \ll (rt)^{1/2}$$

to the sum (9.6).

REFERENCES

1. E. Bombieri and W. M. Schmidt, *On Thue's equation*, Invent. Math. (to appear).
2. K. Mahler, *An inequality for the discriminant of a polynomial*, Michigan Math. J. **11** (1964), 257–262.
3. J. Mueller, *Counting solutions of $|ax^r - by^r| \leq h$*, Quart. J. Math. Oxford (to appear).
4. J. Mueller and W. M. Schmidt, *The number of solutions of trinomial Thue equations and inequalities*, Crelle's J. (to appear).
5. _____, *Thue equations and a conjecture of Siegel*, Acta Math. (submitted).

SCHOOL OF MATHEMATICS, INSTITUTE FOR ADVANCED STUDY, PRINCETON, NEW JERSEY 08540

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF COLORADO, BOULDER, COLORADO 80309