Thue's equation and a conjecture of Siegel

by

J. MUELLER(¹) and W. M. SCHMIDT(²)

Institute for Advanced Study
Princeton, NJ, U.S.A.
and
Fordham University
Bronx, NY, U.S.A.

Institute for Advanced Study
Princeton, NJ, U.S.A.
and
University of Colorado
Boulder, CO, U.S.A.

1. Introduction

In his fundamental work on diophantine equations $f(x,y)=0$, Siegel [18] remarks that in the case when the curve defined by this equation is irreducible and of positive genus, it is possible to find an explicit bound for the number of integer solutions. He then says "Man kann nun vermuten, dass sich sogar eine Schranke finden lässt, die nur von der Anzahl der Koeffizienten abhängt", i.e. one may conjecture that a bound may be derived which depends only on the number of coefficients.

In this form, the conjecture is not true. Chowla [2] showed that for given $k \neq 0$, the number $N_k(h)$ of solutions of $x^3+ky^3=h$ has $N_k(h)=\Omega_k(\log\log h)$, i.e. there is a number $\gamma_k>0$ such that $N_k(h)>\gamma_k \log\log h$ for infinitely many positive values of h. More generally, consider cubic Thue equations, i.e. equations

$$F(x,y)=h, \tag{1.1}$$

where F is a cubic form with integer coefficients which is irreducible over the rationals. According to Mahler [11], the number $N_F(h)$ of solutions of such an equation has $N_F(h)=\Omega_F(\log^{1/4} h)$, and this was improved to $N_F(h)=\Omega_F(\log^{1/3} h)$ by Silverman [19]. In fact, Silverman shows the existence of infinitely many (non-equivalent) cubic forms F as above with $N_F(h)=\Omega_F(\log^{2/3} h)$. On the other hand it is conceivable that the number $N'_F(h)$ of *primitive* solutions (i.e. solutions with coprime x, y) of a cubic Thue equation (1.1) is under some absolute bound. Also, it is possible that Siegel's conjecture is true for curves of genus >1.

(¹) Supported in part by NSF grant DMS-8604568.

(²) Supported in part by NSF grant DMS-8603093.

Siegel in his paper goes on to show that the number of solutions of the cubic Thue equation (1.1) is bounded by a function of h only, i.e. a function independent of F.

In what follows, let $F(x, y)$ be a form of degree $r \geq 3$ with integer coefficients which is irreducible over $\mathbf{Q}$. Evertse [5] was the first to give a bound for the number of solutions of (1.1) which depends on r and h only. More recently, Bombieri and Schmidt [1] have shown that when h has ν distinct prime factors, the equation (1.1) has

$$\ll r^{1+\nu} \tag{1.2}$$

primitive solutions. Here and below, unless indicated otherwise, the constants implicit in $\ll$ will be absolute and effectively computable.

Suppose now that F is a form as above of degree $r \geq 3$, which has not more than $s+1$ nonzero coefficients, so that(1)

$$F(x, y) = \sum_{i=0}^{s} a_i x^{r_i} y^{r-r_i} \tag{1.3}$$

with $0 = r_0 < r_1 < \ldots < r_{s-1} < r_s = r$. A modified version of Siegel's conjecture for the special case of Thue equations, first mentioned to one of the present authors by Bombieri, would be that the number of solutions of (1.1) may be bounded in terms of s and h only. Then the number of solutions of the inequality

$$|F(x, y)| \leq h \tag{1.4}$$

may also be bounded in terms of s and h. Going somewhat in this direction, the second of the present authors [17] has established that (1.4) has

$$\ll (rs)^{1/2} h^{2/r} (1 + \log h^{1/r}) \tag{1.5}$$

solutions for $h \geq 1$. When $h=1$ this gives $\ll (rs)^{1/2}$, which contains the bound $\ll r$ coming from (1.2) with $\nu = 0$.

The modified version of Siegel's conjecture for $s=1$, i.e. for binomial forms $F = ax^r - by^r$, was proved in different ways by Hyyrö [6], then by Evertse [4], then by Mueller [13]. For such forms, the number of solutions of (1.1) or of (1.4) may be bounded in terms of h. The case $s=2$, i.e. the case of trinomial forms F, was settled by the present authors in [14]. Here we will establish the conjecture in general. In fact we have

(1) Our notation here differs from the one in [14] or [17].

THEOREM 1. *The number of solutions of* (1.4) *is*

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

In view of (1.5) it is natural to conjecture that the theorem would remain true with the factor s^2 in (1.6) replaced by s. Such a stronger version would imply the bound (1.5). But such a version may still not be optimal, since the coefficient c_2 in (1.7) below is in fact $\ll 1$ when $r \geq s \log s$.

It is easily seen that the factor $h^{2/r}$ in (1.6) is necessary, but the logarithmic factor is probably not. We can prove this when $r \geq 4s$:

THEOREM 2. *Suppose that* $r \geq 4s$. *Then the number of solutions of* (1.4) *is*

$$\leq c_1(r, s)\, h^{2/r}.$$

More precisely, suppose that $r \geq 2(\varepsilon^{-1}+1)s$ *where* $0<\varepsilon\leq 1$. *Then the number of solutions of* (1.4) *is*

$$\ll c_2\, h^{2/r} + (c_3/\varepsilon)\, h^{(1+\varepsilon)/r}, \tag{1.7}$$

where

$$c_2 = c_2(r, s) = (rs^2)^{2s/r}, \quad c_3 = c_3(r, s) = s^2\, e^{(3300 s \log^3 r)/r}. \tag{1.8}$$

Note that $c_2 \ll 1$ and $c_3 \ll s^2$ when $r \geq s \log^3 s$. Thus when

$$r \geq \max(4s, s\log^3 s), \tag{1.9}$$

the number of solutions of (1.4) is

$$\ll s^2 h^{2/r}.$$

Since (1.5) implies (1.6) when (1.9) is violated, Theorem 1 will follow once we have established Theorem 2.

The set of $(\xi, \eta) \in \mathbf{R}^2$ with $|F(\xi, \eta)| \leq 1$ was shown by Mahler [10] to have finite area A_F. Then the set $\mathfrak{D}$ of (ξ, η) with

$$|F(\xi, \eta)| \leq h \tag{1.10}$$

has area $A_F(h) = A_F h^{2/r}$. One should expect that the number $Z_F(h)$ of solutions of (1.4) in integers is close to $A_F(h)$. In fact Mahler [10] has shown that for fixed F we have the asymptotic relation

$$Z_F(h) = A_F h^{2/r} + O(h^{1/(r-1)})$$

as $h \to \infty$, and this was generalized by Ramachandra [15] to norm form equations. But here the "constant" implicit in O may depend on F.

THEOREM 3. *When $r \geq 4s$ we have*

$$A_F \ll (rs^2)^{2s/r}, \tag{1.11}$$

so that in particular $A_F \ll 1$ when $r \geq s \log s$. We further have

$$|Z_F(h) - A_F h^{2/r}| \leq c_4(r, s)\,(h^{1/(r-2s)} + h^{1/r} \log h^{1/r}) \tag{1.12}$$

with

$$c_4(r, s) = \min(e^{3400 \log^3 r}, c_5(s)),$$

where $c_5(s)$ depends on s only.

We have $h^{1/r} \log h^{1/r} \ll r h^{1/(r-2s)}$, which introduces a factor r, but one of the main features of the theorem is that $c_4(r, s) \leq c_5(s)$, so that it again implies the modified Siegel's conjecture. The function $c_5(s)$ depends on a quantity arising in work of Khovansky [7], and at the present state of knowledge may apparently be taken to be γ^{s^2} with a suitable absolute constant $\gamma > 1$.

Our method is not a straightforward generalization of the cases $s=1$ and $s=2$ treated earlier, and we cannot prove the analogues of certain results in [13] or [14]. We are unable to generalize Lemma 4.3 of [14] and we have to substitute the "operation flip" (in section 8) which is not quite as effective.

Let H be the height of F, so that

$$H = \max_{i=0}^{s} |a_i|.$$

Mueller [13] showed that for a binomial form F and for $h \leq H^{1-(1/r)-\varrho}$ with $\varrho > 0$, the number of primitive solutions of (1.4) is $\leq c_6(\varrho)$. The present authors [14] proved that for a trinomial form F and for $h \leq H^{1-(2/r)-\varrho}$, the number of primitive solutions is $\leq c_6'(\varrho)$. We conjecture that with F of the type (1.3) and with

$$h \leq H^{1-(s/r)-\varrho},$$

the number of primitive solutions of (1.4) is $\leq c_6(s, \varrho)$.

2. Outline of the paper

Throughout, write

$$f(z)=F(z,1)=\sum_{i=0}^{s} a_i z^{r_i}. \tag{2.1}$$

Section 3 contains a discussion of the "Newton polygon" and its application to the distribution of the roots of polynomials such as $f(z)$ with only $s+1$ coefficients. It turns out that the roots are located in not more than s fairly narrow annuli centered at 0. This fact, which may be of independent interest, is essential for deriving our estimates.

Put

$$\mathbf{x}=(x,y), \quad |\mathbf{x}|=\max(|x|,|y|), \quad \langle \mathbf{x} \rangle = \min(|x|,|y|).$$

Relative to two quantities Y_L, Y_S depending on F and h which will be defined below, a solution $\mathbf{x}=(x,y)$ of (1.4) will be called

large if $|\mathbf{x}|>Y_L$,
medium if $|\mathbf{x}|\leqslant Y_L$ and $\langle \mathbf{x}\rangle \geqslant Y_S$,
small if $\langle \mathbf{x}\rangle < Y_S$.

PROPOSITION 1. *The number of primitive large solutions is* $\ll s$.

PROPOSITION 2. *When* $r>2s$, *the number of primitive medium solutions is* $\ll s^2(1+\log h^{1/r})$.

PROPOSITION 3. *When* $r\geqslant 4s$, *the number of small solutions is*

$$\ll c_2 h^{2/r}+c_3 h^{1/(r-2s)}$$

where c_2, c_3 *are given by* (1.8).

One will find a proof of Proposition 1 in § 4, of Proposition 2 in §§ 5–9, and of Proposition 3 in §§ 10–11. Finally the proof of Theorem 3 will be contained in § 12.

We remark here that Theorem 2 follows from the above three propositions by a partial summation argument. First we note that when $0<\varepsilon\leqslant 1$ and $r\geqslant 2s(\varepsilon^{-1}+1)$, then $r-2s\geqslant r/(1+\varepsilon)$, so that the total number $P(h)$ of primitive solutions of (1.4) has

$$P(h)\ll c_2 h^{2/r}+c_3 h^{(1+\varepsilon)/r}. \tag{2.2}$$

The number $\pi(h)$ of primitive solutions of (1.1) is given by $\pi(h)=P(h)-P(h-1)$ (where we set $P(0)=0$). With [ ] denoting the integer part we have

$$\begin{aligned} Z_F(h) &= \sum_{n=1}^{h} \pi(n)\,[(h/n)^{1/r}] \\ &\leq h^{1/r}\sum_{n=1}^{h} \pi(n)\,n^{-1/r} \\ &= h^{1/r}\sum_{n=1}^{h} (P(n)-P(n-1))\,n^{-1/r} \\ &= P(h)+h^{1/r}\sum_{n=1}^{h-1} P(n)\,(n^{-1/r}-(n+1)^{-1/r}) \\ &\ll P(h)+h^{1/r}r^{-1}\sum_{n=1}^{h-1} P(n)\,n^{-1-(1/r)}. \end{aligned}$$

Substituting (2.2) we get

$$Z_F(h) \ll c_2\,h^{2/r}+(c_3/\varepsilon)\,h^{(1+\varepsilon)/r},$$

i.e. Theorem 2.

In both Proposition 1 and 2 we have to deal with integer points in the long "spidery legs" of the domain (1.10). This is intrinsically the harder part and will be dealt with by diophantine approximation methods. However, because of the results of a long development begun by Thue and by Siegel which we are able to use here, our proof of Proposition 1 may appear to be fairly easy. For both Proposition 1 and 2 we will proceed by three steps, as follows.

Step 1. To show that when $\mathbf{x}=(x, y)$ is a solution of (1.4) with $y\neq 0$, then x/y is a good rational approximation to some root α_i of the polynomial

$$f(z) = F(z, 1) = a_s\,(z-\alpha_1)\ldots(z-\alpha_r) \tag{2.3}$$

of (2.1).

Step 1a. To show that we may restrict ourselves to just a few of the roots α_i. In other words, there is a set S of roots of a cardinality which is bounded in terms of s only, such that when $\mathbf{x}$ is a solution of (1.4) with $y\neq 0$, then x/y is a good([1]) rational approximation to some α_i lying in S.

Step 2. To give a bound for the number of good rational approximations x/y (with coprime x, y) to a given α_i.

([1]) but perhaps not quite as good as in Step 1.

In Proposition 3 we have to deal with integer points in the relatively fat domain of $(\xi,\eta)\in\mathbf{R}^2$ with

$$|F(\xi,\eta)|\leq h,\quad \min(|\xi|,|\eta|)\leq Y_S.$$

Here the basic idea is that the number of integer points differs from the area by not much more than Y_S.

Finally we have to specify the quanities Y_L, Y_S introduced above. Unfortunately their definitions are complicated. Put

$$R=e^{800\log^3 r}, \tag{2.4}$$

$$C=(2r^{1/2}M)^r Rh, \tag{2.5}$$

where M is the Mahler height of F. Pick numbers a, b with $0<a<b<1$ and set

$$t=\sqrt{2/(r+a^2)},\qquad \tau=bt,$$

$$\lambda=2/(t-\tau), \tag{2.6}$$

$$A=\frac{t^2}{2-rt^2}(\log M+\frac{r}{2})=\frac{1}{a^2}(\log M+\frac{r}{2}), \tag{2.7}$$

$$\delta=(rt^2+\tau^2-2)/(r-1)=\frac{2}{r-1}\,\frac{b^2-a^2}{r+a^2}. \tag{2.8}$$

If a, b were chosen sufficiently small, then

$$r-\lambda=r-(\sqrt{2r+2a^2}/(1-b))>0$$

for $r\geq 3$. We now set

$$Y_L=(2C)^{1/(r-\lambda)}(4e^A)^{\lambda/(r-\lambda)}. \tag{2.9}$$

We further set

$$Y_S=Y_0^{1/(r-2s)} \tag{2.10}$$

with

$$Y_0=(e^6 s)^r R^{2s} h. \tag{2.11}$$

It follows that

$$Y_S>Y_0^{1/r}=e^6 s R^{2s/r} h^{1/r}=e^6 s(rs)^{2s/r} h^{1/r}. \tag{2.12}$$

3. The location of roots in annuli

Let

$$f(z) = a_0 + a_1 z^{r_1} + \ldots + a_{s-1} z^{r_{s-1}} + a_s z^{r_s} \tag{3.1}$$

be a polynomial with $0=r_0<r_1<\ldots<r_s=r$, and with all the coefficients $\neq 0$. Construct the points

$$P_i = (r_i, -\log|a_i|) \quad (i=0,\ldots,s)$$

and their convex hull C. The *Newton polygon* of f consists of the "lower boundary" of C, i.e. elements $(x,y)\in C$ having $(x,y^*)\notin C$ for every $y^*<y$. The Newton polygon will then consist of certain vertices

$$P_0 = P_{i(0)}, P_{i(1)}, \ldots, P_{i(l)} = P_s \tag{3.2}$$

with $0=i(0)<i(1)<\ldots<i(l)=s$, and the segments between adjacent vertices. We define a function $\sigma(i)$ for $i\in\{i(1),i(2),\ldots,i(l)\}$: namely, when $i=i(k)$, we define $\sigma(i)$ to be the slope of the line segment $P_{i(k-1)}, P_{i(k)}$. Similarly we define a function $\sigma^+(i)$ for $i\in\{i(0),i(1),\ldots,i(l-1)\}$: namely, when $i=i(k)$, then $\sigma^+(i)$ will be the slope of the line segment $P_{i(k)}, P_{i(k+1)}$. Thus $\sigma^+(i(k-1))=\sigma(i(k))$ for $0<k<l$, and the Newton polygon will have slopes

$$\sigma^+(i(0)) = \sigma(i(1)) < \sigma^+(i(1)) = \sigma(i(2)) < \ldots < \sigma^+(i(l-1)) = \sigma(i(l)).$$

The inequalities come from the fact that the Newton polygon is convex. See Figure 1. It is "known" that $r_{i(k)} - r_{i(k-1)}$ of the roots of f will have modulus approximately equal to $e^{\sigma(i(k))}$ $(1\leq k\leq l)$. (In the p-adic case this holds with equality.) In order to make this precise, and for technical expediency, we will introduce "strong vertices" and a "modified Newton polygon".

The point is that, e.g., the vertices P_2 and P_4 in the picture are not useful, since the slopes of the segments to their left and right differ by little. So let us define a vertex $P_{i(k)}$ to be *strong* if

$$\sigma^+(i(k)) > \sigma(i(k)) + 2\theta \quad (0<k<l), \tag{3.3}$$

where

$$\theta = \log 3. \tag{3.4}$$

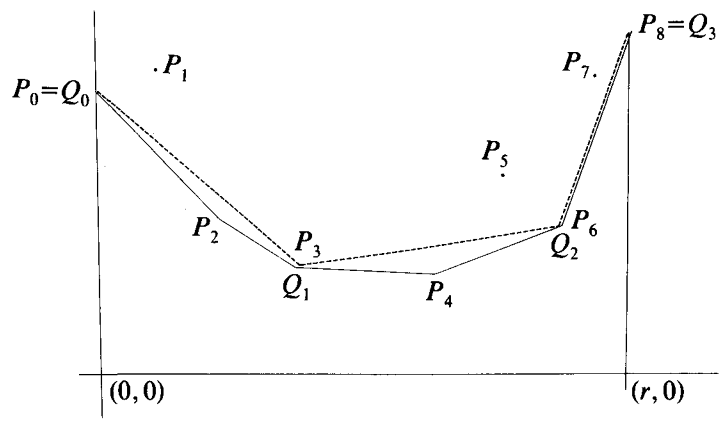

Fig. 1. Example with $s=8$, $l=5$, $i(0)=0$, $i(1)=2$, $i(2)=3$, $i(3)=4$, $i(4)=6$, $i(5)=8$. The modified polygon is drawn with a dashed line and has $p=3$, $i[0]=0$, $i[1]=3$, $i[2]=6$, $i[3]=8$.

Also, by definition, P_0 and P_s will be called *strong*.

Let

$$Q_0=P_0=P_{i[0]},\, Q_1=P_{i[1]},\ldots,\, Q_p=P_{i[p]}=P_s$$

with $0=i[0]<i[1]<\ldots<i[p]=s$ be the sequence of strong vertices. It is a subsequence of (3.2). The number p will be in $1\leqslant p\leqslant s$. The *modified Newton polygon* is the polygon Q_0, $Q_1, \ldots, Q_p$.

Put $R_k=r_{i[k]}, A_k=|a_{i[k]}|\,(0\leqslant k\leqslant p)$; then $Q_k=(R_k,-\log A_k)$. Set

$$M_k=e^{\sigma(i[k])}\quad (k=1,\ldots,p), \tag{3.5}$$

$$N_k=e^{\sigma^+(i[k])}\quad (k=0,\ldots,p-1). \tag{3.6}$$

Then

$$N_0\leqslant M_1<N_1\leqslant M_2<\ldots\leqslant M_{p-1}<N_{p-1}\leqslant M_p,$$

and in fact

$$N_k>e^{2\theta}M_k=9M_k\quad (0<k<p), \tag{3.7}$$

since Q_k is a strong vertex.

LEMMA 1. (i) *Every root* α *of* f *has* $\frac{1}{2}N_0<|\alpha|<2M_p$.

(ii) *For* $0<k<p$, *there are exactly* R_k *roots with* $|\alpha|<3M_k$, *and exactly* $r-R_k$ *roots with* $|\alpha|>3^{-1}N_k$. *(Here roots are counted with multiplicities.)*

Proof. Let us do the upper bounds first. If $f(z)$ is given by (3.1), suppose we substitute z with $|z|=M_k$. Then the ith term in (3.1) will have modulus

$$V_{ki}=|a_i|M_k^{r_i}.$$

We will show that (as an easy consequence of the geometry of the Newton polygon)

$$V_{k,i[k]}=\max_i V_{ki}.$$

Namely, let $\sigma(i,j)$ for $0\leq i,j\leq s$ and $i\neq j$ be the slope of the segment P_i, P_j. Thus

$$\sigma(i,j)=-(\log|a_i|-\log|a_j|)/(r_i-r_j).$$

Then by the convexity of the Newton polygon

$$\sigma(i[k],i)\geq\sigma^+(i[k]) \quad \text{for} \quad i>i[k], \tag{3.8}$$

$$\sigma(i,i[k])\leq\sigma(i[k]) \quad \text{for} \quad i<i[k]. \tag{3.9}$$

Now

$$\begin{aligned}\log(V_{ki}/V_{k,i[k]})&=\log|a_i|-\log|a_{i[k]}|+(r_i-r_{i[k]})\,\sigma(i[k])\\&=(r_i-r_{i[k]})\,(\sigma(i[k])-\sigma(i,i[k])).\end{aligned}\tag{3.10}$$

When $i<i[k]$, the first factor is <0 and the second is ≥ 0 by (3.9), so that

$$V_{ki}/V_{k,i[k]}\leq 1 \quad (i<i[k]). \tag{3.11}$$

When $i>i[k]$, the first factor on the right hand side of (3.10) is >0, and the second is $\leq\sigma(i[k])-\sigma^+(i[k])<-2\theta$ by (3.8), (3.3), so that in fact

$$V_{ki}/V_{k,i[k]}<9^{-(r_i-r_{i[k]})} \quad (i>i[k]). \tag{3.12}$$

Thus when we substitute z with $|z|=M_k$ into (3.1), the $i[k]$th term will dominate the ith terms with $i<i[k]$, but it will *greatly* dominate the ith terms with $i>i[k]$. The idea now is that if we substitute z with $|z|$ somewhat larger than M_k, then the $i[k]$th term will greatly dominate each other term, and hence will exceed the sum of the other terms.

To begin with, when $z=M_p w$ with $|w|\geqslant 2$, then the $i[p]$th term (i.e., since $i[p]=s$, the last term) will dominate:

$$\begin{aligned}|f(z)| &\geqslant |a_s z^{r_s}|-|a_{s-1} z^{r_{s-1}}|-\ldots-|a_0|\\ &= V_{p,i[p]}|w|^{r_s}-V_{p,s-1}|w|^{r_{s-1}}-\ldots-V_{p0}\\ &\geqslant V_{p,i[p]}(|w|^{r_s}-|w|^{r_{s-1}}-\ldots-1)\\ &>0\end{aligned}$$

by (3.11) and since $|w|\geqslant 2$. Thus in fact every root α has $|\alpha|<2M_p$, which is the upper bound in assertion (i).

We now turn to the upper bound in (ii). By part (i), applied to the polynomial

$$f_k(z)=a_0+a_1 z^{r_1}+\ldots+a_{i[k]} z^{r_{i[k]}},$$

this polynomial has all its $r_{i[k]}=R_k$ roots in the disc $|z|<2M_k<3M_k$. In order to show that $f(z)$ itself has R_k roots in the disc $|z|<3M_k$, it will suffice, in view of Rouché's Theorem, to show that

$$|f(z)-f_k(z)|<|f_k(z)| \quad \text{for} \quad |z|=3M_k.$$

But

$$|f_k(z)|-|f(z)-f_k(z)|\geqslant |a_{i[k]} z^{r_{i[k]}}|-\sum_{i\neq i[k]} |a_i z^{r_i}|, \tag{3.13}$$

and we have to make good on our claim that the term $a_{i[k]} z^{r_{i[k]}}$ will dominate. But when $|z|=3M_k$, the right hand side of (3.13) is

$$3^{r_{i[k]}} V_{k,i[k]}-\sum_{i<i[k]} 3^{r_i} V_{ki}-\sum_{i>i[k]} 3^{r_i} V_{ki}\geqslant 3^{r_{i[k]}} V_{k,i[k]}-\sum_{i<i[k]} 3^{r_i} V_{k,i[k]}-\sum_{i>i[k]} 3^{r_{i[k]}-(r_i-r_{i[k]})} V_{k,i[k]}$$

by (3.11) and by (3.12). We get

$$|f_k(z)|-|f(z)-f_k(z)|>3^{r_{i[k]}} V_{k,i[k]}(1-2(3^{-1}+3^{-2}+\ldots))=0.$$

The lower bounds of the lemma may be proved similarly. Alternatively, one could use the reciprocal polynomial

$$\hat{f}(w)=w^r f(w^{-1})=a_s+a_{s-1} w^{r-r_{s-1}}+\ldots+a_0 w^r.$$

The Newton polygon (resp. modified Newton polygon) of $\hat{f}$ is obtained from the Newton polygon (or modified Newton polygon) of f by reflection on the line $x=r/2$. In this transformation, $R_0, R_1, \dots, R_p$ will change into $r-R_p, r-R_{p-1}, \dots, r-R_0$. The reflection changes slopes into minus themselves, so that by (3.5), (3.6), $M_1, \dots, M_p, N_0, \dots, N_{p-1}$ will change respectively into $N_{p-1}^{-1}, \dots, N_0^{-1}, M_p^{-1}, \dots, M_1^{-1}$. Thus by what we have already shown, every root β of $\hat{f}$ will have $|\beta|<2N_0^{-1}$, so that every root α of f has $|\alpha|>\frac{1}{2}N_0$. Similarly, for $0<l<p$, exactly $r-R_{p-l}$ roots of $\hat{f}$ have $|\beta|<3N_{p-l}^{-1}$. Thus for $0<k<p$, precisely $r-R_k$ roots β of $\hat{f}$ have $|\beta|<3N_k^{-1}$, so that $r-R_k$ roots α of f have $|\alpha|>3^{-1}N_k$.

Lemma 1 has been established.

For $k=1,\dots,p$, let $\mathfrak{A}_k$ be the annulus

$$\mathfrak{A}_k: 3^{-1}N_{k-1}<|z|<3M_k. \tag{3.14}$$

By (3.7), these annuli are disjoint. Each root α of f has $|\alpha|>3^{-1}N_0$, and R_1 of the roots have $|\alpha|<3M_1$. Therefore $\mathfrak{A}_1$ contains exactly R_1 roots. All the other roots have $|\alpha|>3^{-1}N_1$. Since there are R_2 roots with $|\alpha|<3M_2$, there must be precisely R_2-R_1 roots in $\mathfrak{A}_2$. And so forth. Thus we have

LEMMA 2. *The annulus $\mathfrak{A}_k$ (where $1\leq k\leq p$) contains exactly R_k-R_{k-1} roots of f.*

Suppose now that $\alpha\in\mathfrak{A}_k$; then $\zeta=\log|\alpha|$ lies in the interval

$$\sigma^+(i[k-1])-\theta<\zeta<\sigma(i[k])+\theta. \tag{3.15}$$

There may be vertices $P_{i(t)}$ of the Newton polygon between $P_{i[k-1]}$ and $P_{i[k]}$; say

$$P_{i[k-1]}=P_{i(a)}, P_{i(a+1)}, \dots, P_{i(b)}=P_{i[k]}$$

are adjacent vertices. Then

$$\begin{aligned}\sigma^+(i[k-1])=\sigma^+(i(a))=\sigma(i(a+1))<\sigma^+(i(a+1))=\sigma(i(a+2))<\dots\\ =\sigma(i(b-1))<\sigma^+(i(b-1))=\sigma(i(b))=\sigma(i[k]).\end{aligned}$$

Since there are no strong vertices between $P_{i[k-1]}, P_{i[k]}$, we have

$$\sigma^+(i(t))\leq\sigma(i(t))+2\theta \quad (a<t<b).$$

Thus for every ζ in the interval (3.15) there is a t in $a<t\leq b$ with $|\zeta-\sigma(i(t))|\leq\theta$. We have proved

LEMMA 3. *For every root α of f there is a t in $1\leq t\leq l$ with*

$$|\log|\alpha|-\sigma(i(t))|\leq\theta. \tag{3.16}$$

4. Large solutions

As we remarked in Section 2, the proof of Proposition 1 will be in three steps. For Step 1 we simply quote Lemma 1 of [1]:

LEMMA 4. *For any* (x, y) *with* (1.4) *and* $y \neq 0$, *there is a root* α_i *of* f *with*

$$\min(1, |\alpha_i - \frac{x}{y}|) \leq (2r^{1/2}M)^r h|\mathbf{x}|^{-r},$$

where M *is the Mahler height of* F.

For Step 1a, we begin by quoting Lemma 9 of [17]:

LEMMA 5. *Let* $f(z)$ *be a polynomial of degree* r *with real coefficients. Let* $\sigma = A + Bi$ *be a root of* f *with* $B > 0$ *and suppose that* $f(x)f'(x) < 0$ *for real* x *in the interior of*

$$A \leq x < A + (9r)^h B, \tag{4.1}$$

where h *is a positive integer. Then* f *has at least* $e^{\sqrt{h}/16}$ *roots in the square consisting of* $z = x + iy$ *with* (4.1) *and* $0 < y \leq (9r)^h B$.

Put

$$R_1 = e^{790 \log^3 r}, \quad R = e^{800 \log^3 r}. \tag{4.2}$$

LEMMA 6. *Let* $f(z)$ *be a polynomial of degree* r *with real coefficients. Suppose that* $f(x)f'(x) \neq 0$ *for real* x *in* I, *where* I *is an interval* $X_1 < x < X_2$, *or a half line* $x < X_2$, *or* $x > X_1$. *Suppose there are* $m \geq 1$ *roots* $\alpha_j = x_j + iy_j$ $(j = 1, \ldots, m)$ *with real parts* $x_j \in I$. *Then there is a root* α_l *among these* m *roots such that for every real* ξ,

$$|\xi - \alpha_l| < R \min_{1 \leq i \leq m} |\xi - \alpha_i|.$$

Proof. Replacing $f(x)$ by $f(-x)$ if necessary, we may suppose that $f(x)f'(x) < 0$ in I, so that $|f(x)|^2$ decreases in I. Then I cannot contain arbitrarily large x, hence must be of the type $X_1 < x < X_2$ or $x < X_2$. We claim that

$$|y_i| \geq R_1^{-1}(X_2 - x_i) \quad (i = 1, \ldots, m). \tag{4.3}$$

For otherwise there is a root $z = A + Bi$ with $A < X_2$ and

$$|B| < R_1^{-1}(X_2 - A).$$

There will be such a root with $B>0$. We set

$$h=[260\log^2 r].$$

Then

$$A+(9r)^h B<A+R_1 B<X_2.$$

The hypotheses of the preceding lemma are satisfied, so that f has at least $e^{\sqrt{h}/16}$ roots. But $h>256\log^2 r$, so that we would get more than r roots, which is impossible. Therefore (4.3) is indeed true.

Now let α_l be among $\alpha_1,\ldots,\alpha_m$ with a minimum value of $|y_l|$. Then for $1\leqslant i\leqslant m$, by (4.3),

$$\begin{aligned}|\alpha_l-\alpha_i| &\leqslant |x_l-x_i|+|y_l-y_i|\\ &\leqslant |X_2-x_i|+|X_2-x_l|+2|y_i|\\ &\leqslant R_1|y_i|+R_1|y_l|+2|y_i|\\ &<4R_1|y_i|.\end{aligned}$$

Thus for real ξ,

$$\begin{aligned}|\xi-\alpha_l| &\leqslant |\xi-\alpha_i|+|\alpha_l-\alpha_i|<|\xi-\alpha_i|+4R_1|y_i|\\ &\leqslant |\xi-\alpha_i|+4R_1|\xi-\alpha_i|\\ &<R|\xi-\alpha_i|.\end{aligned}$$

LEMMA 7. *Let $f(z)$ be a polynomial of degree r with real coefficients and having $\leqslant s+1$ nonzero coefficients. Then there is a set S of roots α_l of f of cardinality $|S|\leqslant 6s+4$ such that for any real ξ*

$$\min_{\alpha_l\in S}|\xi-\alpha_l|\leqslant R\min_{1\leqslant i\leqslant r}|\xi-\alpha_i|. \tag{4.4}$$

Proof. It is easily seen that f has $\leqslant 2s+1$ real zeros, and also f' has $\leqslant 2s+1$ real zeros, so that ff' has $\leqslant 4s+2$ real zeros. Thus the real numbers x with $ff'(x)\neq 0$ fall into $\leqslant 4s+3$ intervals (or half lines) I. Let S consist on the one hand of the real zeros of f, and on the other hand for each interval I as above for which there are roots of f with real part in I, pick an α_l according to Lemma 6. The set S so attained will have $|S|\leqslant 2s+1+4s+3=6s+4$, and clearly (4.4) is valid.

Remark. In our work [14] we showed that for $s=2$, the arguments of the roots of f are well distributed. More generally, Khovansky [8] has shown (see also [9], [16]) that for polynomials with $\leq s+1$ coefficients, the arguments of the roots are fairly well distributed. This fact, together with information on the moduli of the roots as given in Lemma 2, could also be used to obtain a lemma like our Lemma 7. But our present knowledge about certain bounds of Khovansky is poor, so that in this way the set S could only be shown to have cardinality $<\gamma_1^{s^2}$ with some constant $\gamma_1>1$.

Combining Lemmas 4 and 7, and recalling (2.5), we get

LEMMA 8. *There is a set S of roots of f of cardinality $\leq 6s+4$, such that for any (x, y) with* (1.4) *and $y\neq 0$, there is a root $\alpha_i \in S$ with*

$$\min(1, |\alpha_i - \frac{x}{y}|) \leq C|\mathbf{x}|^{-r}. \tag{4.5}$$

This completes Step 1a for large solutions. Step 2 will be accomplished via the Thue–Siegel principle.

LEMMA 9. *Given a root α_i of f, the number of reduced fractions with* (4.5) *and with* $\max(|x|, |y|) > Y_L$ *is* $\ll 1$.

Proof. With the notations (2.6), (2.8), the number in question is

$$<2+\frac{\log(\delta^{-1}(\lambda-2)^{-1})}{\log(r-1)} \tag{4.6}$$

by [1, § III]. In [1] we had $C=(2r^{1/2}M)^r$ rather than (2.5), both in (4.5) and in the definition [1, (2.9)] of Y_L, but clearly this does not matter. Now by (2.6),

$$\lambda = \frac{2}{t(1-b)} = \frac{\sqrt{2r+2a^2}}{1-b} > \sqrt{r}$$

if b is sufficiently small, and

$$\delta^{-1} < \frac{r^2}{2(b^2-a^2)}.$$

Hence the quantity in (4.6) is bounded.

Proposition 1 is an immediate consequence of Lemmas 8 and 9.

5. Medium solutions

In contrast to large solutions, the bulk of the work on medium solutions will be in Step 1, i.e. to show that each medium solution of (1.4) is such that x/y is a good rational approximation to some α_i.

LEMMA 10. *Let x, y satisfy* (1.4), *and let α be an element of $\{\alpha_1, \dots, \alpha_r\}$ with*

$$|x-\alpha y| = \min_{1 \leq j \leq r} |x-\alpha_j y|. \tag{5.1}$$

Suppose that $y \neq 0$ and that $f^{(u)}(\alpha) \neq 0$ with u in $1 \leq u \leq r$. Then

$$\left|\alpha - \frac{x}{y}\right| \leq \frac{r}{2}\left(\frac{2^r h}{|f^{(u)}(\alpha) y^r|}\right)^{1/u}. \tag{5.2}$$

Proof. The symbol $\mathfrak{S}$ will denote an ordered subset of $\{1, \dots, r\}$ or cardinality $|\mathfrak{S}| = u$. There are $r(r-1)\dots(r-u+1) \leq r^u$ such subsets. Introducing

$$f_{\mathfrak{S}}(z) = a_s \prod_{j \notin \mathfrak{S}} (z-\alpha_j)$$

where the product is over j in $1 \leq j \leq r$ with $j \notin \mathfrak{S}$, we have

$$f^{(u)}(z) = \sum_{\mathfrak{S}} f_{\mathfrak{S}}(z).$$

Hence there is a set $\mathfrak{S}$ with

$$|f^{(u)}(\alpha)| \leq r^u |f_{\mathfrak{S}}(\alpha)|. \tag{5.3}$$

Such a set $\mathfrak{S}$ will be fixed in what follows.

For any j,

$$|y(\alpha-\alpha_j)| \leq |x-\alpha y| + |x-\alpha_j y| \leq 2|x-\alpha_j y|,$$

so that

$$|y^{r-u} f_{\mathfrak{S}}(\alpha)| = |a_s| \prod_{j \notin \mathfrak{S}} |y(\alpha-\alpha_j)| \leq 2^{r-u} |a_s| \prod_{j \notin \mathfrak{S}} |x-\alpha_j y|,$$

and further by (5.3),

$$|y^{r-u}f^{(u)}(\alpha)| \leq r^u 2^{r-u} |a_s| \prod_{j \notin \mathfrak{S}} |x-\alpha_j y|.$$

If we multiply by $\Pi_{j\in\mathfrak{S}}|x-\alpha_j y|$, we obtain

$$\left(\prod_{j\in\mathfrak{S}} |x-\alpha_j y|\right) |y^{r-u}f^{(u)}(\alpha)| \leq r^u 2^{r-u} |F(x,y)| \leq r^u 2^{r-u} h.$$

But this yields

$$|x-\alpha y|^u |y^{r-u} f^{(u)}(\alpha)| \leq r^u 2^{r-u} h$$

and (5.2).

Lemma 10 will be useful provided we can show that some derivative $f^{(u)}(\alpha)$ is large. This will be our next task.

6. Estimation of derivatives

The aim of this section is to show that for each root α of f there is a u, $1\leq u\leq r$, such that $|f^{(u)}(\alpha)|$ is large:

LEMMA 11. *Suppose* (1.4) *holds with* $y\neq 0$, *and let* α *be a root of* f *with* (5.1). *Let* $K=K(\alpha)$ *and* $k=k(\alpha)$ *be certain integers to be defined below. Then there is a* u *in* $1\leq u\leq i(K)$ *with*

$$|f^{(u)}(\alpha)| \geq \frac{1}{4s}(2s^2r)^{1-s} |a_{i(K)}| |\alpha|^{r_{i(K)}-u}. \tag{6.1}$$

Also, there is a v *in* $1\leq v\leq s-i(k)$ *with*

$$|f^{(v)}(\alpha)| \geq \frac{1}{4s}(2s^2r)^{1-s} |a_{i(k)}| |\alpha|^{r_{i(k)}-v}. \tag{6.2}$$

Recall the vertices (3.2) of the Newton polygon. We now define $K(\alpha), k(\alpha)$ as follows. We set $K(\alpha)=l$ if $\sigma^+(i(l-1))=\sigma(i(l))=\sigma(s)<\log|\alpha|+\log s+3$. When $\sigma(s)\geq \log|\alpha|+\log s+3$, we let $K=K(\alpha)$ be least in $0\leq K\leq l-1$ with $\sigma^+(i(K))\geq\log|\alpha|+\log s+3$. Similarly, we set $k(\alpha)=0$ if $\sigma^+(0)=\sigma^+(i(0))=\sigma(i(1))>\log|\alpha|-\log s-3$. When $\sigma^+(0)\leq\log|\alpha|-\log s-3$, we let $k(\alpha)$ be the largest integer in $1\leq k\leq l$ with $\sigma(i(k))\leq \log|\alpha|-\log s-3$. It is then clear that $k(\alpha)\leq K(\alpha)$. In fact we claim that

$$k(\alpha)<K(\alpha). \tag{6.3}$$

For it may not actually happen that $K(\alpha)=0$, for this would mean that $\sigma(i(1))=\sigma^+(i(0))\geq \log|\alpha|+\log s+3$, which is incompatible with Lemma 3. Similarly it may not happen that $k(\alpha)=l$. But if we had $0<k(\alpha)=K(\alpha)<l$, then $\sigma(i(k(\alpha)+1)))=\sigma^+(i(k(\alpha)))=\sigma^+(i(K(\alpha)))\geq \log|\alpha|+\log s+3$, $\sigma(i(k(\alpha)))\leq \log s-3$, and there could be no t with (3.16), in contradiction to Lemma 3.

In what follows, set $K=K(\alpha)$, $k=(\alpha)$.

Before giving a proof of Lemma 11, we will try to give motivation for (6.1), (6.2) and their proofs. Suppose

$$g(z)=\sum_{i=m}^{m+t} a_i z^{r_i}$$

is a polynomial involving $t+1$ powers $z^{r_m}, \ldots, z^{r_{m+t}}$ with $r_m<\ldots<r_{m+t}$. For given $\alpha\neq 0$, the system of relations

$$g(\alpha)=g'(\alpha)=\ldots=g^{(t)}(\alpha)=0$$

is equivalent to

$$g(\alpha)=\alpha g'(\alpha)=\ldots=\alpha^t g^{(t)}(\alpha)=0. \tag{6.4}$$

This is a system of $t+1$ homogeneous linear equations in the $t+1$ coefficients a_i. The matrix of this system of equations has columns $\alpha^{r_i}\mathbf{r}_i^t$ with

$$\mathbf{r}_i^t=\begin{pmatrix} 1 \\ r_i \\ r_i(r_i-1) \\ \vdots \\ r_i(r_i-1)\ldots(r_i-t+1) \end{pmatrix} \quad (m\leq i\leq m+t). \tag{6.5}$$

The determinant $\det(\mathbf{r}_m^t, \ldots, \mathbf{r}_{m+t}^t)$ with columns $\mathbf{r}_m^t, \ldots, \mathbf{r}_{m+t}^t$ is a certain obvious Van der Monde determinant, so that

$$\det(\mathbf{r}_m^t, \ldots, \mathbf{r}_{m+t}^t)=\prod_{m\leq i<j\leq m+t}(r_j-r_i)\neq 0. \tag{6.6}$$

Thus when $\alpha\neq 0$, the relations (6.4) are impossible unless g vanishes identically. In a similar manner one can see that if the coefficients of g are not small, then $|g(\alpha)|$, $|g'(\alpha)|, \ldots, |g^{(t)}(\alpha)|$ cannot all be small. In particular when α is a root of g, then $|g'(\alpha)|, \ldots, |g^{(t)}(\alpha)|$ cannot all be small.

Suppose now that α is a root of f. Suppose further that the term $a_i\alpha^{r_i}$ in $f(\alpha)$ is "dominant", i.e. that

$$|a_i\alpha^{r_i}| >> \max_k |a_k\alpha^{r_k}|.$$

Then if no cancellations arise, we can expect that

$$|f^{(u)}(\alpha)| >> |a_i||\alpha|^{r_i-u}.$$

Proof of Lemma 11. Set

$$g_1(z) = \sum_{i\leq i(K)} a_i z^{r_i}, \quad g_2(z) = \sum_{i\geq i(k)} a_i z^{r_i}.$$

If we consider g_1 as the main part of f and apply the ideas discussed above to g_1, we will be led to (6.1); on the other hand by considering g_2 we will be led to (6.2).

Put

$$t_1 = i(K), \quad t_2 = s - i(k). \tag{6.7}$$

For $\nu=1,2$, for $0\leq u\leq t_\nu$ and $0\leq i\leq s$, let $\mathbf{r}_{ui}^{(\nu)}$ be the column

$$\mathbf{r}_{ui}^{(\nu)} = \begin{pmatrix} 1 \\ r_i \\ \vdots \\ r_i(r_i-1)\dots(r_i-u+2) \\ r_i(r_i-1)\dots(r_i-u) \\ \vdots \\ r_i(r_i-1)\dots(r_i-t_\nu+1) \end{pmatrix}.$$

Note that $\mathbf{r}_{ui}^{(\nu)}$ is the column (6.5) with $t=t_\nu$ and with the $(u+1)$st row removed. Put

$$E_u^{(1)} = (-1)^{u+t_1}\det(\mathbf{r}_{u0}^{(1)}, \mathbf{r}_{u1}^{(1)}, \dots, \mathbf{r}_{u,i(K)-1}^{(1)}) \quad (0\leq u\leq t_1),$$
$$E_u^{(2)} = (-1)^{u+t_2}\det(\mathbf{r}_{u,i(k)+1}^{(2)}, \dots, \mathbf{r}_{u,s-1}^{(2)}, \mathbf{r}_{us}^{(2)}) \quad (0\leq u\leq t_2).$$

Then[1] for $0\leq j\leq s$,

---

[1] With $r_j(r_j-1)\dots(r_j-u+1)$ understood to be 1 when $u=0$.

$$\sum_{u=0}^{t_1} E_u^{(1)} r_j(r_j-1)\dots(r_j-u+1) = \det(\mathbf{r}_0^{t_1},\dots,\mathbf{r}_{i(K)-1}^{t_1},\mathbf{r}_j^{t_1}) = D_j^{(1)},$$

say, and

$$\sum_{u=0}^{t_2} E_u^{(2)} r_j(r_j-1)\dots(r_j-u+1) = \det(\mathbf{r}_{i(k)+1}^{t_2},\dots,\mathbf{r}_{s-1}^{t_2},\mathbf{r}_s^{t_2},\mathbf{r}_j^{t_2}) = D_j^{(2)},$$

say. As a consequence,

$$\sum_{u=0}^{t_\nu} E_u^{(\nu)} z^u f^{(u)}(z) = \sum_{j=0}^{s} a_j D_j^{(\nu)} z^{r_j} \quad (\nu = 1,2).$$

Now $D_j^{(1)}=0$ when $j<i(K)$ and $D_j^{(2)}=0$ when $j>i(k)$. Since $f(\alpha)=0$, we obtain

$$\sum_{u=0}^{t_1} E_u^{(1)} \alpha^u f^{(u)}(\alpha) = a_{i(K)} D_{i(K)}^{(1)} \alpha^{r_{i(K)}} + \sum_{j>i(K)} a_j D_j^{(1)} \alpha^{r_j}, \tag{6.8}$$

$$\sum_{u=1}^{t_2} E_u^{(2)} \alpha^u f^{(u)}(\alpha) = a_{i(k)} D_{i(k)}^{(2)} \alpha^{r_{i(k)}} + \sum_{j<i(k)} a_j D_j^{(2)} \alpha^{r_j}. \tag{6.9}$$

There are no terms $j>i(K)$ in (6.8) when $K=l$. When $K<l$ we wish to show that the first term on the right hand side of (6.8) dominates. To this end we introduce the quotients

$$W_j^{(1)} = |a_j D_j^{(1)} \alpha^{r_j} / a_{i(K)} D_{i(K)}^{(1)} \alpha^{r_{i(K)}}| \quad (j>i(K)).$$

Similarly, when $k>0$ we put

$$W_j^{(2)} = |a_j D_j^{(2)} \alpha^{r_j} / a_{i(k)} D_{i(k)}^{(2)} \alpha^{r_{i(k)}}| \quad (j<i(k)).$$

By the product formula (6.6) and in view of $|r_j-r_w|\leqslant|r_{i(K)}-r_w|+|r_j-r_{i(K)}|$ we have

$$|D_j^{(1)}/D_{i(K)}^{(1)}| = \prod_{w<i(K)} \left|\frac{r_j-r_w}{r_{i(K)} r_w}\right| \leqslant \prod_{w<i(K)} \left(1+\left|\frac{r_j-r_{i(K)}}{r_{i(K)}-r_w}\right|\right).$$

Note that $w<i(K)$ yields $r_{i(K)}-r_w\geqslant i(K)-w$, so that

$$|D_j^{(1)}/D_{i(K)}^{(1)}| \leqslant \prod_{m=1}^{i(K)} (1+m^{-1}|r_j-r_{i(K)}|)$$

and

$$\log |D_j^{(1)}/D_{i(K)}^{(1)}| \leq (r_j - r_{i(K)}) \sum_{m=1}^{i(K)} m^{-1} < (r_j - r_{i(K)})(1+\log s)$$

for $j>i(K)$. Similarly,

$$\log |D_j^{(2)}/D_{i(k)}^{(2)}| < (r_{i(K)} - r_j)(1+\log s)$$

for $j<i(k)$.

For $j>i(K)$ we have

$$\log W_j^{(1)} < (r_j - r_{i(K)})(\log|\alpha|+1+\log s) + \log|a_j| - \log|a_{i(K)}|$$
$$= (r_j - r_{i(K)})(\log|\alpha|+1+\log s - \sigma(j, i(K))),$$

where (as in Section 3), $\sigma(i,j)$ for $0\leq i,j\leq s$ and $i\neq j$ denotes the slope of the segment P_i, P_j. Since $j>i(K)$, and by the convexity of the Newton polygon,

$$\sigma(j, i(K)) \geq \sigma^+(i(K)) \geq \log|\alpha|+\log s+3,$$

so that

$$\log W_j^{(1)} < -2(r_j - r_{i(K)}),$$

i.e.

$$W_j^{(1)} < e^{-2(r_j - r_{i(K)})},$$

and therefore

$$\sum_{j>i(K)} W_j^{(1)} < e^{-2}+e^{-4}+\ldots<\frac{1}{2}.$$

Because of (6.8) and the definition of $W_j^{(1)}$ we may conclude that

$$\left|\sum_{u=1}^{t_1} E_u^{(1)} \alpha^u f^{(u)}(\alpha)\right| > \frac{1}{2}|a_{i(K)} D_{i(K)}^{(1)} \alpha^{r_{i(K)}}|. \tag{6.10}$$

In analogous manner, (6.9) yields

$$\left|\sum_{u=1}^{t_2} E_u^{(2)} \alpha^u f^{(u)}(\alpha)\right| > \frac{1}{2}|a_{i(k)} D_{i(k)}^{(2)} \alpha^{r_{i(k)}}|. \tag{6.11}$$

In the next section we will show that

$$|E_u^{(1)}| \leqslant 2^s(s^2r)^{s-1}|D_{i(K)}^{(1)}| \quad (1 \leqslant u \leqslant t_1), \tag{6.12}$$

$$|E_u^{(2)}| \leqslant 2^s(s^2r)^{s-1}|D_{i(k)}^{(2)}| \quad (1 \leqslant u \leqslant t_2). \tag{6.13}$$

Substituting this into (6.10) we see that there is a u in $1 \leqslant u \leqslant t_1$ with

$$|f^{(u)}(\alpha)| \geqslant \frac{1}{4s}(2s^2r)^{1-s}|a_{i(K)}||\alpha|^{r_{i(K)}-u}.$$

Similarly (6.11), (6.13) yield a v in $1 \leqslant v \leqslant t_2$ with

$$|f^{(v)}(\alpha)| \geqslant \frac{1}{4s}(2s^2r)^{1-s}|a_{i(k)}||\alpha|^{r_{i(k)}-v}.$$

7. Estimation of determinants

LEMMA 12. *Let $m_1, \dots, m_T$ be integers with $0 \leqslant m_j \leqslant j$ $(j=1, \dots, T)$ and $m_i - m_j \neq i-j$ for $1 \leqslant j < i \leqslant T$. Then*

$$m_1 + m_2 + \dots + m_T \leqslant T.$$

Proof. This is clear for $T=1$. When $T>1$ and $m_j \neq j$ for $j=1, \dots, T$, then $m_1=0$. Furthermore $m_2 = m_2 - m_1 \neq 2-1 = 1$, so that also $m_2=0$. Continuing in this way we get $m_j = 0$ $(j=1, \dots, T)$. Suppose then that $T>1$ and $m_i = i$ for some particular i, and that $m_j < j$ for $j<i$. It follows in the same manner as above that $m_1 = \dots = m_{i-1} = 0$. For $k=1,2,\dots,T-i$ and $t=0,1,\dots,i-1$ we have $m_{i+k} = m_{i+k} - m_t \neq i+k-t$, so that $0 \leqslant m_{i+k} \leqslant k$. Applying the lemma with $T-i$ in place of T we get

$$m_{i+1} + m_{i+2} + \dots + m_T \leqslant T-i.$$

Since $m_0 + m_1 + \dots + m_i = 0 + \dots + 0 + i$, the assertion follows.

Consider variables $r_0, r_1, \dots, r_J$ and vectors

$$\mathbf{a}_j = (r_0^j, r_1^j, \dots, r_J^j) \quad (j = 0, 1, \dots).$$

For $0 \leqslant v \leqslant J+1$ put

$$G_v = \det(\mathbf{a}_0, \mathbf{a}_1, \dots, \mathbf{a}_{v-1}, \mathbf{a}_{v+1}, \dots, \mathbf{a}_{J+1}).$$

In particular, G_{J+1} is the Van der Monde determinant

$$M = \prod_{0 \leqslant w < u \leqslant J} (r_u - r_w).$$

LEMMA 13.

$$G_v = S_{J+1-v} M \quad (0 \leq v \leq J)$$

where S_{J+1-v} *is the elementary symmetric function in* $r_0, \ldots, r_J$ *of degree* $J+1-v$.

Proof. G_v vanishes when $r_u = r_w$ for some $u \neq w$, so that G_v is divisible by M, say $G_v = G_v^* M$. Comparison of degrees shows that G_v^* is homogenous of total degree $J+1-v$ and of degree 1 in each r_i. Since G_v/M is invariant under permutations of $r_0, \ldots, r_J$, the quotient G_v^* is a constant multiple of S_{J+1-v}, so that $G_v = c_v S_{J+1-v} M$ with constant c_v. Since G_v, S_{J+1-v} and M are polynomials with integer coefficients having no nontrivial common divisors (in fact the coefficients are among $1, -1, 0$), we have $c_v = \pm 1$. Now when $r_0, \ldots, r_J$ is a rapidly increasing sequence, the main term in G_v will be $r_1 r_2^2 \ldots r_{v-1}^{v-1} r_v^{v+1} \ldots r_J^{J+1}$, and $G_v > 0$, $M > 0$, $S_{J+1-v} > 0$, so that $c_v > 0$ and therefore $c_v = 1$.

Consider the vectors

$$\mathbf{b}_j = (r_0(r_0-1) \ldots (r_0-j+1), \ldots, r_J(r_J-1) \ldots (r_J-j+1)) \quad (j = 0, 1, \ldots),$$

with $\mathbf{b}_0$ to be interpreted as $(1, 1, \ldots, 1)$. For $0 \leq u \leq J+1$, put

$$H_u = \det(\mathbf{b}_0, \mathbf{b}_1, \ldots, \mathbf{b}_{u-1}, \mathbf{b}_{u+1}, \ldots, \mathbf{b}_{J+1}).$$

LEMMA 14.

$$H_u = N_{J+1-u} M,$$

where N_{J+1-u} *is a polynomial in* $r_0, \ldots, r_J$ *of degree* $J+1-u$, *with integer coefficients whose moduli have a sum*

$$\leq 2^{J+1}(J+1)^{2(J+1-u)}.$$

Proof. It is clear that H_u is divisible by M, and comparison of degrees shows that H_u/M is of degree $J+1-u$. The hard part is to estimate the coefficients. Observe that

$$r(r-1)(r-2) \ldots (r-j+1) = \sum_{n=0}^{j} (-1)^{n-j} c_{jn} r^j$$

where $c_{jj} = 1$, and for $0 \leq n < j$,

$$c_{jn} = \sum_{0<i_1<\ldots<i_{j-n}<j} i_1 i_2 \ldots i_{j-n} \leqslant \left(\sum_{0<i<j} i \right)^{j-n} < (j^2/2)^{j-n}. \tag{7.1}$$

We have

$$\mathbf{b}_j = \sum_{n=0}^{j} (-1)^{n-j} c_{jn} \mathbf{a}_j$$

and therefore

$$H_u = \det(\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_{u-1}, \mathbf{b}_{u+1}, \ldots, \mathbf{b}_{J+1})$$
$$= \sum_{u \leqslant n_{u+1} \leqslant u+1} \cdots \sum_{u \leqslant n_{J+1} \leqslant J+1} (-1)^{\varrho} c_{u+1, n_{u+1}} \cdots c_{J+1, n_{J+1}} \det(\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_{u-1}, \mathbf{a}_{n_{u+1}}, \ldots, \mathbf{a}_{n_{J+1}})$$

with $\varrho=(u+1-n_{u+1})+\ldots+(J+1-n_{J+1})$. The determinant on the right hand side vanishes unless $n_{u+1}, \ldots, n_{J+1}$ are distinct, in which case it is a determinant G_v. By Lemma 13 the quotient $H_u/M=N_{J+1-u}$ is a polynomial whose "length" $L(N_{J+1-u})$, i.e. the sum of the moduli of its coefficients, is

$$\leqslant 2^{J+1} \sum_{u \leqslant n_{u+1} \leqslant u+1} \cdots \sum_{\substack{u \leqslant n_{J+1} \leqslant J+1 \\ n_{u+1}, \ldots, n_{J+1} \text{ distinct}}} c_{u+1, n_{u+1}} \cdots c_{J+1, n_{J+1}}$$

since $L(S_{J+1-v}) \leqslant 2^{J+1}$. We set $T=J+1-u$ and

$$m_j = u+j-n_{u+j} \quad (1 \leqslant j \leqslant T),$$

so that by (7.1)

$$c_{u+j, n_{u+j}} \leqslant ((u+j)^2/2)^{m_j} \leqslant ((J+1)^2/2)^{m_j}$$

and

$$L(N_{J+1-u}) \leqslant 2^{J+1} \sum\nolimits^* ((J+1)^2/2)^{m_j}$$

where the sum Σ^* is over $0 \leqslant m_j \leqslant j \ (=1, \ldots, T)$ with $m_i - m_j \neq i-j$ when $i \neq j$. The number of summands is 2^T, so that by Lemma 12,

$$L(N_{J+1-v}) \leqslant 2^{T+J+1}((J+1)^2/2)^T = 2^{J+1}(J+1)^{2(J+1-u)}.$$

We now turn to the proof of (6.12). Setting $J=i(K)-1$ we have

$$D^{(1)}_{i(K)} = \prod_{0<v<w\leq J+1} (r_w - r_v) = M \prod_{0<v\leq J} (r_{i(K)} - r_v).$$

Substituting integers with $0=r_0<r_1<\ldots<r_{i(K)}\leq r$ we get $|D^{(1)}_{i(K)}|\geq M>0$. On the other hand, $E^{(1)}_u$, except for the sign, is the same as H_u with $J=i(K)-1$, so that by Lemma 14

$$|E^{(1)}_u| \leq 2^{J+1}(J+1)^{2(J+1-u)} r^{J+1-u} M.$$

For $1\leq u\leq t_1=i(K)\leq s$ we have $J+1\leq s$ and $J+1-u\leq s-1$, so that

$$|E^{(1)}_u / D^{(1)}_{i(K)}| \leq 2^s s^{2(s-1)} r^{s-1},$$

which establishes (6.12). The relation (6.13) is shown similarly.

8. Good rational approximations

We will complete Step 1 for medium solutions. Let q be the smallest number with

$$|a_q| = \max_{0\leq j\leq s} |a_j| = H. \tag{8.1}$$

Then P_q is the lowest vertex (or the left one of two equally low vertices) of the Newton polygon.

LEMMA 15. *Suppose* (1.4) *holds with* $y\neq 0$, *and let* α *be a root of* f *with* (5.1). *Suppose that* $q<i(K)$ *where* $K=K(\alpha)$. *Pick* u *according to Lemma* 11. *Then*

$$\left|\alpha - \frac{x}{y}\right| < \frac{1}{H^{(1/u)-(1/r)}} \left(\frac{(rs)^{2s}(2e^3 s)^r h}{|y|^r} \right)^{1/u} \tag{8.2}$$

Proof. By Lemma 10 in conjunction with Lemma 11,

$$\left|\alpha - \frac{x}{y}\right| < \left(\frac{2^r (rs)^{2s} h}{|a_{i(K)}| |\alpha|^{r_{i(K)}-u} |y|^r} \right)^{1/u}.$$

We will show that

$$\Delta(\alpha, u) = |a_{i(K)}| |\alpha|^{r_{i(K)}-u}$$

has

$$\Delta(\alpha, u) \geqslant (e^3 s)^{-r} H^{1-(u/r)}, \tag{8.3}$$

so that (8.2) will follow.

By the construction of $K(\alpha)$,

$$\log|\alpha| \geqslant \sigma(i(K)) - \log s - 3.$$

Since $u \leqslant i(K) \leqslant r_{i(K)}$ we have

$$\begin{aligned}\log \Delta(\alpha, u) &= (r_{i(K)} - u)(\log|\alpha|) + \log|a_{i(K)}| \\ &\geqslant (r_{i(K)} - u)(\sigma(i(K)) - \log s - 3) + \log|a_{i(K)}|.\end{aligned}$$

In view of $q < i(K)$ and the convexity of the Newton polygon, $\sigma(i(K)) \geqslant \sigma(q, i(K))$ and thus

$$\log \Delta(\alpha, u) \geqslant (r_{i(K)} - u)\,\sigma(q, i(K)) + \log|a_{i(K)}| - r\log(e^3 s). \tag{8.4}$$

By the general identity

$$r_i \sigma(q, i) + \log|a_i| = r_q \sigma(q, i) + \log|a_q| \quad (i \neq q), \tag{8.5}$$

the right-hand side of (8.4) is

$$(r_q - u)\,\sigma(q, i(K)) + \log|a_q| - r\log(e^3 s).$$

Now since P_q is the lowest (or one of the lowest) vertices of the Newton polygon, and since $q < i(K)$, we have $\sigma(q, i(K)) \geqslant 0$. Therefore in the case when $r_q \geqslant u$ we get

$$\log \Delta(\alpha, u) \geqslant \log|a_q| - r\log(e^3 s) = \log H - r\log(e^3 s),$$

establishing (8.3). On the other hand when $r_q < u$, we note that $\sigma(q, i(K)) \leqslant \sigma(q, s)$ and we obtain

$$\begin{aligned}\log \Delta(\alpha, u) &\geqslant (r_q - u)\,\sigma(q, s) + \log|a_q| - r\log(e^3 s) \\ &= -(r_q - u)\frac{\log|a_s| - \log|a_q|}{r_s - r_q} + \log|a_q| - r\log(e^3 s) \\ &= \left(1 - \frac{u - r_q}{r_s - r_q}\right)\log|a_q| + \frac{u - r_q}{r_s - r_q}\log|a_s| - r\log(e^3 s) \\ &\geqslant \left(1 - \frac{u - r_q}{r - r_q}\right)\log|a_q| - r\log(e^3 s)\end{aligned}$$

since $r_s = r$. But $u \leqslant r$, so that $(u - r_q)/(r - r_q) \leqslant u/r$ and

$$\log \Delta(\alpha, u) \geqslant (1 - (u/r)) \log H - r \log(e^3 s),$$

i.e. (8.3).

LEMMA 16. *Suppose* (1.4) *holds with* $x \neq 0$ *and*

$$|y| \geqslant 2(rs)^{2s/r} h^{1/r}. \tag{8.6}$$

Let α *be a root of* f *with* (5.1), *and suppose that* $i(k) < q$. *Pick* v *according to Lemma* 11. *Then*

$$\left| \alpha^{-1} - \frac{y}{x} \right| < \frac{1}{H^{(1/v) - (1/r)}} \left(\frac{(rs)^{2s} (4e^3 s)^r h}{|x|^r} \right)^{1/v}. \tag{8.7}$$

Proof. By (6.2) in conjunction with Lemma 10,

$$\left| \alpha - \frac{x}{y} \right| < \left(\frac{2^r (rs)^{2s} h}{|a_{i(k)}| |\alpha|^{r_{i(k)} - v} |y|^r} \right)^{1/v} = \left(\frac{2^r (rs)^{2s} h}{\Delta^* |y|^r} \right)^{1/v} \tag{8.8}$$

with

$$\Delta^* = \Delta^*(\alpha, v) = |a_{i(k)}| |\alpha|^{r_{i(k)} - v}.$$

Here

$$\log(|\alpha|^v \Delta^*(\alpha, v)) = r_{i(k)} \log |\alpha| + \log |a_{i(k)}|.$$

When $k=0$ we get $r_{i(0)} = r_0 = 0$ and $\log |a_{i(k)}| = \log |a_0| \geqslant 0$. When $k > 0$ we have $\log |\alpha| \geqslant \sigma(i(k)) + \log s + 3$, so that

$$\begin{aligned} \log(|\alpha|^v \Delta^*(\alpha, v)) &> r_{i(k)} \sigma(i(k)) + \log |a_{i(k)}| \\ &\geqq r_{i(k)} \sigma(0, i(k)) + \log |a_{i(k)}| = \log |a_0| \geqslant 0. \end{aligned}$$

Inserting this into (8.8) we get

$$\left| \alpha - \frac{x}{y} \right| < |\alpha| \left(\frac{2^r (rs)^{2s} h}{|y|^r} \right)^{1/v} \leqslant |\alpha|$$

by (8.6), so that

$$|x| < |2\alpha y|. \tag{8.9}$$

Now comes "operation flip". From (8.8),

$$\left|\alpha^{-1}-\frac{y}{x}\right| = \left|\frac{y}{x\alpha}\right|\left|\alpha-\frac{x}{y}\right| < \left(\frac{2^r(rs)^{2s}h}{|a_{i(k)}||\alpha|^{r_{i(k)}}}\right)^{1/v}\frac{1}{|x||y|^{(r/v)-1}}.$$

Now by (8.9), and with the notation

$$\Gamma(\alpha, v) = |a_{i(k)}||\alpha|^{-(r-r_{i(k)}-v)},$$

we obtain

$$\left|\alpha^{-1}-\frac{y}{x}\right| < \left(\frac{4^r(rs)^{2s}h}{\Gamma(\alpha, v)|x|^r}\right)^{1/v}.$$

Lemma 16 will follow once we have shown that

$$\Gamma(\alpha, v) \geqslant (e^3s)^{-r}H^{1-(v/r)}. \tag{8.10}$$

The proof of this is "dual" to the proof of (8.3). The number v of Lemma 11 has $v \leqslant s-i(k) \leqslant r-r_{i(k)}$, and since $\sigma^+(i(k)) > \log|\alpha| - \log s - 3$ by the definition of k, we have

$$\log\Gamma(\alpha, v) \geqslant (r-r_{i(k)}-v)(-\sigma^+(i(k)) - \log(e^3s)) + \log|a_{i(k)}|.$$

But $q > i(k)$, so that $\sigma(i(k), q) \geqslant \sigma^+(i(k))$ and

$$\begin{aligned}\log\Gamma(\alpha, v) &\geqslant -(r-r_{i(k)}-v)\,\sigma(i(k), q) + \log|a_{i(k)}| - r\log(e^3s)\\ &= -(r-r_q-v)\,\sigma(i(k), q) + \log|a_q| - r\log(e^3s)\end{aligned}$$

by the identity (8.5).

In the case when $v \leqslant r-r_q$, this is

$$\geqslant \log|a_q| - r\log(e^3s) = \log H - r\log(e^3s),$$

since $\sigma(i(k), q) \leqslant 0$. Thus (8.10) follows in this case. On the other hand $\sigma(i(k), q) \geqslant \sigma(0, q)$, so that in the case when $v > r-r_q$, we obtain

$$\begin{aligned}\log\Gamma(\alpha, v) &\geqslant -(r-r_q-v)\,\sigma(0, q) + \log|a_q| - r\log(e^3s)\\ &= (r-r_q-v)((\log|a_q| - \log|a_0|)/r_q) + \log|a_q| - r\log(e^3s)\\ &\geqslant ((r-v)/r_q)\log|a_q| - r\log(e^3s).\end{aligned}$$

But

$$(r-v)/r_q \geq (r-v)/r = 1-(v/r),$$

so that

$$\log \Gamma(\alpha, v) \geq (1-(v/r)) \log H - r \log(e^3 s),$$

whence again (8.10).

Lemmas 15, 16 are essentially dual to each other. But the choice of α according to (5.1) is not symmetric in x, y which necessitated "operation flip" in the proof of Lemma 16. In the trinomial case (the case $s=2$) [14] we were able to pick α (resp. $1/\alpha$) in a way which was symmetric in x, y; this is more difficult to do in the general case.

9. The number of medium solutions

With Y_S satisfying (2.12), Lemmas 15, 16 show that when (1.4) holds with

$$\langle \mathbf{x} \rangle = \min(|x|, |y|) \geq Y_S,$$

there is either a root α of $F(x,1)$ with (8.2) or a root α^{-1} of $F(1,y)$ with (8.7). Moreover, since $\langle \mathbf{x} \rangle \geq Y_S$, the right hand sides increase with u resp. v, so that we may replace u, v by s.

Step 1a for medium solutions is accomplished just as for large solutions. With R given by (2.4) and with the aid of Lemma 7, we see the following.

LEMMA 17. *There is a set S of roots of $F(x,1)$ and a set S^* of roots of $F(1,y)$, both with cardinalities $\leq 6s+4$, such that any solution $\mathbf{x}$ of (1.4) with $\langle \mathbf{x} \rangle \geq Y_S$ either has*

$$\left| \alpha - \frac{x}{y} \right| < \frac{R}{H^{(1/s)-(1/r)}} \left(\frac{(rs)^{2s}(4e^3 s)^r h}{|y|^r} \right)^{1/s} \tag{9.1}$$

with some $\alpha \in S$, or has

$$\left| \alpha^* - \frac{y}{x} \right| < \frac{R}{H^{(1/s)-(1/r)}} \left(\frac{(rs)^{2s}(4e^3 s)^r h}{|x|^r} \right)^{1/s} \tag{9.2}$$

for some $\alpha^* \in S^*$.

For Step 2 for medium solutions it will suffice without loss of generality to estimate the number of solutions of (9.1) with

$$Y_S \leq y \leq Y_L. \tag{9.3}$$

Now (9.1) implies that

$$\left|\alpha-\frac{x}{y}\right|<\frac{K}{2y^{r/s}} \tag{9.4}$$

with

$$K=2R(rs)^2(4e^3s)^{r/s}h^{1/s}H^{(1/r)-(1/s)}<R^2(e^5s)^{r/s}h^{1/s}H^{-1/2s}, \tag{9.5}$$

since we suppose $r>2s$. Let $x_0/y_0,\dots,x_\nu/y_\nu$ be the solutions of (9.3), (9.4) with g.c.d.$(x_i,y_i)=1$ and ordered such that

$$Y_S\leq y_0\leq\dots\leq y_\nu\leq Y_L.$$

Then (9.4) gives

$$\frac{1}{y_iy_{i+1}}\leq\frac{K}{y_i^{r/s}},$$

so that we have the "gap principle",

$$y_{i+1}\geq K^{-1}y_i^{(r/s)-1}\geq K^{-1}Y_S^{(r/s)-2}y_i=K^{-1}Y_0^{1/s}y_i\geq e^{r/s}H^{1/2s}y_i$$

by (2.10), (2.12), (9.5) $(i=0,\dots,\nu-1)$. Thus $y_\nu\geq(e^{r/s}H^{1/2s})^\nu$ and

$$\log y_\nu\geq(\nu/2s)(r+\log H). \tag{9.6}$$

On the other hand by (2.9)

$$\log y_\nu\leq\log Y_L=\frac{1}{r-\lambda}\log 2C+\frac{\lambda}{r-\lambda}(A+\log 4).$$

Here $\lambda\ll\sqrt{r}$ and $r-\lambda\gg r$, so that in conjunction with (2.5), (2.7),

$$\begin{aligned}\log y_\nu&\ll r^{-1}\log 2C+r^{-1/2}(A+\log 4)\\&\ll\log M+\log r+\log h^{1/r}+r^{-1/2}(\log M+r)\\&\ll\log H+r^{1/2}+\log h^{1/r},\end{aligned}$$

since $M\leq(r+1)H$ (Mahler [12]). In conjunction with (9.6) this gives

$$\nu\ll s(1+\log h^{1/r}),$$

so that in fact the number of primitive solutions of (9.4) with (9.3) is $\ll s(1+\log h^{1/r})$. This is true for each $\alpha\in S$, and a similar situation pertains for each $\alpha^*\in S^*$ in (9.2). Proposition 2 has been proved.

10. Small solutions

We will suppose that $r \geqslant 4s$.

LEMMA 18. *Given* $Y \geqslant 1$, *the number of solutions of* (1.4) *with* $\langle \mathbf{x} \rangle < Y$ *is*

$$\ll c_2(r, s)\, h^{2/r} + sY.$$

From (2.10), (2.11) we have (on using $r \geqslant 4s$)

$$\begin{aligned} sY_S &\ll s(e^6 s)^{r/(r-2s)} R^{2s/(r-2s)} h^{1/(r-2s)} \\ &\ll s^2 \cdot s^{4s/r} R^{4s/r} h^{1/(r-2s)} < s^2 \cdot e^{(3300 s \log^3 r)/r} h^{1/(r-2s)} \\ &= c_3(r, s)\, h^{1/(r-2s)}. \end{aligned} \tag{10.1}$$

Thus Proposition 3 is a consequence of Lemma 18.

To prove the lemma, it will suffice by symmetry to estimate the number of solutions with

$$0 \leqslant y < Y. \tag{10.2}$$

For given y, the *real* numbers x with (1.4) fall into a finite number of intervals. Since there are not more than $2s$ numbers x with $F(x, y) = h$, and similarly with $F(x, y) = -h$, there are $\leqslant 4s$ such intervals. For given y, the number of *integers* x with (1.4) then is

$$\leqslant \mu(y) + 4s,$$

where $\mu(y)$ is the measure of the set of real numbers x with $|F(x, y)| \leqslant h$. Therefore the number of integer pairs with (1.4) and (10.2) is

$$\leqslant \sum_{y \text{ with } (10.2)} (\mu(y) + 4s) \leqslant \sum_{y=-\infty}^{\infty} \mu(y) + 8(Y+1)s.$$

In the next section we will show that

$$\sum_{y=-\infty}^{\infty} \mu(y) \ll c_2(r, s)\, h^{2/r}, \tag{10.3}$$

so that Lemma 18 will follow.

11. Estimation of measures

For $t \geqslant 1$ set

$$P_t(r_0, r_1, \ldots, r_t) = \sum_{i=0}^{t} [(r_t - r_0)(r_t - r_1) \ldots (r_t - r_{i-1})][(r_0 - r_{i+1})(r_0 - r_{i+2}) \ldots (r_0 - r_t)],$$

where the first bracket is understood to be 1 when $i=0$, and the second bracket stands for 1 when $i=t$. We claim that

$$P_t(r_0, r_1, \ldots, r_t) = 0. \tag{11.1}$$

This is seen by inspection when $t=1$. When $t>1$,

$$\begin{aligned}P_t(r_0, r_1, \ldots, r_t) &= \sum_{i=2}^{t} (r_t - r_0)(r_t - r_1) \ldots (r_t - r_{i-1})(r_0 - r_{i+1})(r_0 - r_{i+2}) \ldots (r_0 - r_t) \\ &\quad + (r_0 - r_1)(r_0 - r_2) \ldots (r_0 - r_t) + (r_t - r_0)(r_0 - r_2) \ldots (r_0 - r_t) \\ &= \sum_{i=2}^{t} (r_t - r_0)(r_t - r_1) \ldots (r_t - r_{i-1})(r_0 - r_{i+1})(r_0 - r_{i+2}) \ldots (r_0 - r_t) \\ &\quad + (r_t - r_1)(r_0 - r_2)(r_0 - r_3) \ldots (r_0 - r_t) \\ &= (r_t - r_1) P_{t-1}(r_0, r_2, \ldots, r_t).\end{aligned}$$

Hence (11.1) follows by induction.

Let

$$p(z) = A_s z^{r_s} + \ldots + A_1 z^{r_1} + A_0 \tag{11.2}$$

be a polynomial with $r=r_s>r_{s-1}>\ldots>r_1>r_0=0$. We have

$$p'(z) = p_1(z) z^{r_1 - 1}$$

with

$$p_1(z) = r_s A_s z^{r_s - r_1} + r_{s-1} A_{s-1} z^{r_{s-1} - r_1} + \ldots + r_1 A_1.$$

More generally, defining

$$p_i(z) = \sum_{j=i}^{s} r_j (r_j - r_1) \ldots (r_j - r_{i-1}) A_j z^{r_j - r_i} \quad (0 \leq i \leq s)$$

we have $p_0(z) = p(z)$ and

$$p_i'(z) = p_{i+1}(z) z^{r_{i+1} - r_i - 1} \quad (0 \leq i < s). \tag{11.3}$$

Now for $0 \leq l \leq s$,

$$\sum_{i=l}^{s} p_i(z) z^{r_i} (r_l - r_{i+1})(r_l - r_{i+2}) \ldots (r_l - r_s) = \sum_{j=l}^{s} A_j z^{r_j} T(s, l, j)$$

with

$$T(s,l,j)=\sum_{i=l}^{j} r_j(r_j-r_1)\dots(r_j-r_{i-1})(r_l-r_{i+1})(r_l-r_{i+2})\dots(r_l-r_s).$$

Here

$$T(s,l,l)=r_l(r_l-r_1)\dots(r_l-r_{l-1})(r_l-r_{l+1})\dots(r_l-r_s).$$

But for $j>l$

$$T(s,l,j)=T_1(s,l,j)\,T_2(s,l,j)$$

with

$$\begin{aligned}T_1(s,l,j)&=r_j(r_j-r_1)\dots(r_j-r_{l-1})(r_l-r_{j+1})(r_l-r_{j+2})\dots(r_l-r_s),\\ T_2(s,l,j)&=\sum_{i=l}^{j}(r_j-r_l)(r_j-r_{l+1})\dots(r_j-r_{i-1})(r_l-r_{i+1})(r_l-r_{i+2})\dots(r_l-r_j)\\ &=P_{j-l}(r_l,r_{l+1},\dots,r_j)\\ &=0.\end{aligned}$$

Therefore

$$\begin{aligned}&\sum_{i=l}^{s} p_i(z)\,z^{r_i}(r_l-r_{i+1})(r_l-r_{i+2})\dots(r_l-r_s)\\ &=A_l z^{r_l} r_l(r_l-r_1)\dots(r_l-r_{l-1})(r_l-r_{l+1})\dots(r_l-r_s).\end{aligned} \tag{11.4}$$

LEMMA 19. *Let $p(z)$ be a polynomial as above, with real coefficients and with $A_s \neq 0$. The real numbers x with*

$$|p(x)|\leq h \tag{11.5}$$

make up a set of measure

$$\mu<10(rs^2)^{s/r}(h/|A_s|)^{1/r}. \tag{11.6}$$

When $|A_0|\geq 2h$, we also have

$$\mu\leq 20rs^2|A_s|^{-1/r}|A_0|^{(1/r)-(1/s)}h^{1/s}. \tag{11.7}$$

Proof. We introduce a new parameter X. Numbers x with $|x| \leqslant X$ make up a set of measure $2X$. We will now concentrate on numbers x with

$$|x| > X. \tag{11.8}$$

Observe that $p'(x)$ assumes a given value for not more than $2s$ real numbers x; therefore the numbers x with $|p'(x)| > hs^2/X$ constitute not more than $4s$ intervals and half lines. The numbers x with (11.5) lying in an interval with $|p'(x)| > hs^2/X$ make up a set of measure $\leqslant 2X/s^2$. Therefore if we neglect a set of measure $\leqslant 4s \cdot 2X/s^2 = 8X/s$, we may concentrate on x with $|p'(x)| > hs^2/X$, i.e. with

$$|p_1(x) x^{r_1 - 1}| \leqslant hs^2/X. \tag{11.9}$$

We now repeat the argument with hs^2/X in place of h and with $q(x) = p_1(x) x^{r_1 - 1}$ in place of $p(x)$. If we neglect a further set of measure $\leqslant 8X/s$, we may concentrate on x with $|q'(x)| \leqslant hs^4/X^2$. But

$$q'(x) = p_1'(x) x^{r_1 - 1} + (r_1 - 1) p_1(x) x^{r_1 - 2}.$$

The second summand here is of modulus $\leqslant (r_1 - 1) hs^2/X^2$ by (11.8) and (11.9), so that we get $|p_1'(x) x^{r_1 - 1}| \leqslant r_1 hs^4/X^2$, i.e. by (11.3),

$$|p_2(x) x^{r_2 - 2}| \leqslant r_1 s^4 h/X^2. \tag{11.10}$$

We now deal with (11.10) the way we dealt with (11.9). To do so we replace p_1, r_1, h by $p_2, r_2 - 1, r_1 s^2 h/X$. So if we neglect another set of measure $\leqslant 8X/s$, we may suppose that

$$|p_3(x) x^{r_3 - 3}| \leqslant r_1 (r_2 - 1) s^6 h/X^3.$$

This argument may be continued. We see: except for a set of measure $\leqslant 2X + s(8X/s) = 10X$, the numbers x with (11.5) have

$$|p_i(x) x^{r_i - i}| \leqslant r_1 r_2 \dots r_{i-1} s^{2i} h/X^i \quad (i = 1, \dots, s),$$

(where $r_1 r_2 \dots r_{i-1} = 1$ when $i = 1$), and in fact

$$|x|^{-s} |p_i(x) x^{r_i}| \leqslant r_1 r_2 \dots r_{i-1} s^{2i} h/X^s \quad (i = 1, \dots, s)$$

by (11.8). Mindful of the identities (11.4), we may infer that for $l = 1, 2, \dots, s$,

$$|A_l r_l (r_l - r_1) \dots (r_l - r_{l-1})(r_l - r_{l+1}) \dots (r_l - r_s)|$$
$$\leq |x|^{s-r_l} \sum_{i=l}^{s} r_1 r_2 \dots r_{i-1} (r_l - r_{i+1})(r_l - r_{i+2}) \dots (r_l - r_s) s^{2i} h / X^s.$$

(When $l=s$, then $(r_l - r_{l+1}) \dots (r_l - r_s)$ means 1, etc.) Each summand on the right contains $s-1$ factors linear in the r_j, and each such factor has modulus $\leq r$. The left hand side contains the factors $|r_l(r_l - r_s)| \geq \frac{1}{2} r$ when $l<s$, and the factor $r_s = r$ when $l=s$. Therefore

$$|A_l| \leq |x|^{s-r_l} \cdot 2r^{s-2} s^{2s} h / X^s \quad (l = 1, \dots, s). \tag{11.11}$$

In particular,

$$|A_s| \leq 2 \cdot s^{2s} r^{s-2} h |x|^{s-r} X^{-s} \leq 2 s^{2s} r^{s-2} h X^{-r}.$$

This is impossible unless $X \leq X_0 = (2s^{2s} r^{s-2} / |A_s|)^{1/r}$. Therefore the numbers x with (11.5) make up a set of measure

$$\leq 10 X_0 < 10 s^{2s/r} r^{s/r} (h / |A_s|)^{1/r},$$

and (11.6) is established.

On the other hand when (11.5) holds and when $|A_0| \geq 2h$, there is an l in $1 \leq l \leq s$ with

$$|A_l x^{r_l}| \geq |A_0| / (2s) = Q, \tag{11.12}$$

say. We will set $X = X_1$ with

$$X_1 = 2s^2 r^{1-(2/s)} |A_s|^{-1/r} Q^{(1/r)-(1/s)} h^{1/s}.$$

In the case when (11.12) holds with $l=s$, we have $|A_s x^r| \geq Q$, so that $|x| \geq (Q/|A_s|)^{1/r}$ and

$$|A_s x^{r-s}| \geq |A_s| (Q/|A_s|)^{1-(s/r)} = |A_s|^{s/r} Q^{1-(s/r)}.$$

This is incompatible with (11.11) with $l=s$ and $X=X_1$. Suppose then that (11.12) fails to hold for $l=s$ but is true for some l in $1 \leq l < s$. Then $|A_l x^{r_l}| > |A_s x^{r_s}|$, so that $|x|^r < |A_l x^{r_l}| / |A_s|$. Therefore

$$|A_l x^{r_l - s}| = |A_l x^{r_l}| |x|^{-s} > |A_l x^{r_l}|^{1-(s/r)} |A_s|^{s/r} \geq Q^{1-(s/r)} |A_s|^{s/r}.$$

This is incompatible with (11.11) with $X = X_1$.

Thus when $X=X_1$, the inequalities (11.5), (11.8), (11.11) are incompatible, and the numbers x with (11.5) have measure

$$\leq 10X_1 < 20s^2r|A_s|^{-1/r}|A_0|^{(1/r)-(1/s)}h^{1/s}.$$

Proof of (10.3). Given y,

$$F(x, y) = \sum_{i=0}^{s} A_i x^{r_i} = p(x),$$

say, with $A_i = a_i y^{r-r_i}$. We note that $|A_s| = |a_s| \geq 1$ and that $|A_0| = |a_0 y^r| \geq |y|^r$. Therefore by Lemma 19,

$$\mu(y) < 10(rs^2)^{s/r}h^{1/r}, \tag{11.13}$$

and moreover

$$\mu(y) < 20rs^2h^{1/s}|y|^{1-(r/s)} \tag{11.14}$$

when $|y| \geq (2h)^{1/r}$. We have

$$\sum_{|y|<(rs^2)^{s/r}h^{1/r}} \mu(y) \ll (rs^2)^{2s/r}h^{2/r},$$

$$\begin{aligned}\sum_{|y|\geq(rs^2)^{s/r}h^{1/r}} \mu(y) &\ll rs^2h^{1/s} \sum_{y>(rs^2)^{s/r}h^{1/r}} y^{1-(r/s)} \\ &\ll rs^2h^{1/s}((r/s)-2)^{-1}\, h^{(2/r)-(1/s)}(rs^2)^{(s/r)(2-(r/s))} \\ &\ll \frac{r}{r-2s}s^3(rs^2)^{2s/r}(rs^2)^{-1}h^{2/r} \\ &\leq \frac{r}{r-2s}(rs^2)^{2s/r}h^{2/r}.\end{aligned}$$

When $r \geq 4s$, the assertion (10.3) follows.

In the same way one sees that for $r > 2s$

$$\int_{-\infty}^{\infty} \mu(y)\, dy \ll \frac{r}{r-2s}(rs^2)^{2s/r}h^{2/r},$$

and taking $h=1$ we obtain

$$A_F \ll \frac{r}{r-2s}(rs^2)^{2s/r}.$$

Assuming $r \geq 4s$ we get $A_F \ll (rs^2)^{2s/r}$, i.e. (1.11).

12. Proof of Theorem 3

We will suppose throughout that $r \geqslant 4s$. Put

$$Y = eY_S.$$

The domain $\mathfrak{D} = \mathfrak{D}(F, h)$ of $(\xi, \eta) \in \mathbf{R}^2$ with $|F(\xi, \eta)| \leqslant h$ is the disjoint union

$$\mathfrak{D}(F, h) = \mathfrak{D}_1 \cup \mathfrak{D}_2 \cup \mathfrak{D}_3,$$

where $\mathfrak{D}_1, \mathfrak{D}_2, \mathfrak{D}_3$ respectively consist of points in $\mathfrak{D}$ with

$$\max(|\xi|, |\eta|) \leqslant Y, \tag{12.1}$$

$$\min(|\xi|, |\eta|) \leqq Y < \max(|\xi|, |\eta|), \tag{12.2}$$

$$\min(|\xi|, |\eta|) > Y. \tag{12.3}$$

Since $Y > Y_S > (2sh)^{1/r}$, we may apply (11.14) to obtain

$$\int_Y^\infty \mu(\eta)\, d\eta \ll rs^2 h^{1/s} \int_Y^\infty \eta^{1-(r/s)}\, d\eta \ll s^3 h^{1/s} Y^{2-(r/s)}$$

$$< s^3 h^{1/s} Y_0^{-1/s} = s^3 (e^6 s)^{-r/s} R^{-2} < 1.$$

Thus the part of $\mathfrak{D}(F, h)$ with $\eta > Y$ has measure $\ll 1$, so that by symmetry the part with $\max(|\xi|, |\eta|) > Y$ has measure $\ll 1$. Therefore if A_i denotes the area of $\mathfrak{D}_i$, we have $A_2 + A_3 \ll 1$, and since $A_1 + A_2 + A_3 = A_F(h) = A_F h^{2/r}$, we get

$$A_1 = A_F h^{2/r} + O(1), \tag{12.4}$$

with an absolute constant implicit in O. We write $Z_F(h)$ as

$$Z_F(h) = Z_1 + Z_2 + Z_3,$$

where Z_i denotes the number of integer pairs in $\mathfrak{D}_i (i = 1, 2, 3)$. Not surprisingly, Z_1 will be the main term and Z_2, Z_3 will be error terms.

Now $\mathfrak{D}_1$ lies in the square (12.1) and is bounded in such a way that a line ξ=const. or a line η=const., with the exception of the lines $\xi = \pm Y$, $\eta = \pm Y$, contains $\ll s$ boundary points of $\mathfrak{D}_1$. This is true because of the special type of F. It follows, e.g. from work of Davenport [3], that

$$|Z_1 - A_1| \ll sY \ll c_3(r, s)\, h^{1/(r-2s)} \tag{12.5}$$

by (10.1).

We next turn to Z_3. Propositions 1, 2 give an estimate for the number of primitive integer points in $\mathfrak{D}_3$ but unfortunately not for the number of all integer points. A point $\mathbf{x}\in\mathfrak{D}_3$ may be written as $\mathbf{x}=l\mathbf{x}_0$ where $\mathbf{x}_0$ is primitive and l a positive integer. Clearly $1\leqslant l\leqslant h^{1/r}$. Let $Z_3(u)$ be the number of $\mathbf{x}\in\mathfrak{D}_3$ where $e^u\leqslant l<e^{u+1}$; then

$$Z_3=\sum_{u=0}^{[\log h^{1/r}]+1} Z_3(u).$$

For such $\mathbf{x}$ we have

$$|F(\mathbf{x}_0)|=l^{-r}|F(\mathbf{x})|\leqslant e^{-ur}h=h(u), \tag{12.6}$$

say. Further

$$\begin{aligned}\min(|x_0|,|y_0|)&=l^{-1}\min(|x|,|y|)\geqslant l^{-1}Y>e^{-u-1}Y \\ &=e^{-u}Y_S>e^{-ur/(r-2s)}\,Y_S=Y_S(u),\end{aligned} \tag{12.7}$$

where $Y_S(u)$ is as Y_S in (2.10), (2.11), but with $h(u)$ in place of h. By Propositions 1, 2, with $h(u)$ in place of h, the number of primitive $\mathbf{x}_0$ with (12.6), (12.7) is

$$\ll s^2(1+\log h^{1/r}(u))\leqslant s^2(1+\log h^{1/r}).$$

Given $\mathbf{x}_0$ the number of possibilities for l is $<e^{u+1}$. Thus

$$Z_3(u)\ll s^2e^u(1+\log h^{1/r})$$

and

$$Z_3\ll s^2h^{1/r}(1+\log h^{1/r}) \tag{12.8}$$

We next turn to Z_2; by reasons of symmetry we may restrict ourselves to the domain

$$0\leqslant\xi\leqslant Y,\quad \eta\geqslant Y,\quad |F(\xi,\eta)|\leqslant h. \tag{12.9}$$

When $h\neq0$, the polynomial $F(x,y)-h$ is irreducible over $\mathbf{C}$, so that $F(x,y)-h$ and $F_y(x,y)$ (where F_y denotes the partial derivative) have no nontrivial common factor. Therefore the system $F(\xi,\eta)=h$, $F_y(\xi,\eta)=0$ has only finitely many common solutions (ξ,η), henceforth called *critical points*. Now on the one hand, by Bezout's Theorem, the number of critical points is $\leqslant r^2$, and on the other hand, by work of Khovansky [7] (see also [9], [16]), since each of $F-h$, F_y has $\leq s+1$ terms, the number of these points $(\xi,\eta)\in\mathbf{R}^2$ is $\leq c_7(s)$. When (ξ^*,η^*) lies on the curve $F(\xi,\eta)=h$ but is not a critical point, there is a unique function $\eta=\eta(\xi)$, regular in a neighborhood of ξ^*, with $\eta(\xi^*)=\eta^*$ and $F(\xi,\eta(\xi))=h$. Similar remarks may be made with regard to $F(\xi,\eta)=-h$. Thus the domain (12.9) is the union of m domains (with disjoint interiors)

$$\mathfrak{E}_i\colon a_i\leq\xi\leq b_i, \eta_{i1}(\xi)\leq\eta\leq\eta_{i2}(\xi) \quad (i=1,\ldots,m),$$

where

$$m\ll c_8(r,s)=\min(r^2,c_7(s)),$$

where $\eta_{i1}(\xi)<\eta_{i2}(\xi)$ in $a_i<\xi<b_i$, and where each $\eta_{ik}(\xi)$ in $a_i<\xi<b_i$ is either constant and equal to Y, or is a regular solution of $F(\xi,\eta)=h$ or $-h$.

For each domain $\mathfrak{E}_i$, either $\eta_{i2}(\xi)-\eta_{i1}(\xi)$ is constant in $a_i<\xi<b_i$ (e.g., when $\eta_{i2}(\xi)=Y$, $\eta_{i1}(\xi)=-Y$), or it is not. Consider now domains $\mathfrak{E}_i$ where this difference is not constant, so that not identically $\eta'_{i1}(\xi)=\eta'_{i2}(\xi)$. We want to look at point pairs (ξ_0,η_1), (ξ_0,η_2) where $\eta_1=\eta_{i1}(\xi_0)$, $\eta_2=\eta_{i2}(\xi_0)$ for $a_i<\xi_0<b_i$ and where $\eta'_{i1}(\xi_0)=\eta'_{i2}(\xi_0)$. For such pairs, the numbers ξ_0, η_1, η_2 satisfy the three equations

$$F(\xi_0,\eta_1)=\pm h, \quad F(\xi_0,\eta_2)=\pm h, \quad \begin{vmatrix} F_x(\xi_0,\eta_1) & F_y(\xi_0,\eta_1) \\ F_x(\xi_0,\eta_2) & F_y(\xi_0,\eta_2)\end{vmatrix}=0.$$

By what we said above, the solutions (ξ_0,η_1,η_2) of these equations of the type indicated above are isolated. By Bezout's Theorem, there are $\leq r^3$ such solutions, and by Khovansky's Theorem, there are $\leq c_9(s)$ solutions. Therefore the domains $\mathfrak{E}_i$ may be further split up such that (12.9) is the union of n domains (with disjoint interiors)

$$\mathfrak{F}_i\colon a_i\leq\xi\leq b_i, \eta_{i1}(\xi)\leq\eta\leq\eta_{i2}(\xi) \quad (i=1,\ldots,n),$$

where

$$n\ll c_{10}(r,s)=\min(r^3,c_7(s)+c_9(s)),$$

where the functions $\eta_{i1}(\xi)$, $\eta_{i2}(\xi)$ have the properties enunciated above, and where moreover in each $\mathfrak{F}_i$ either $\eta_{i2}(\xi)-\eta_{i1}(\xi)$ is constant, or $\eta'_{i1}(\xi)\neq\eta'_{i2}(\xi)$ in $a_i<\xi<b_i$.

The number of integer points in $\mathfrak{F}_i$ is

$$\leq \sum_{a_i\leq x\leq b_i}(\eta_{i2}(x)-\eta_{i1}(x)+1).$$

Now $\eta_{i2}-\eta_{i1}$ is monotonic in $a_i \leq x \leq b_i$; say it is increasing. Then

$$\sum_{a_i \leq x \leq b_i} (\eta_{i2}(x)-\eta_{i1}(x)) \leq \int_{a_i}^{b_i} (\eta_{i2}(\xi)-\eta_{i1}(\xi))\, d\xi + \eta_{i2}(b_i)-\eta_{i1}(b_i).$$

Here $\eta_{i2}(b_i)-\eta_{i1}(b_i) \ll (rs^2)^{s/r} h^{1/r}$ by (11.13), so that the number of integer points in $\mathfrak{F}_i$ is

$$\ll u_i + v_i + w$$

with

$$u_i = \int_{a_i}^{b_i} (\eta_{i2}(\xi)-\eta_{i1}(\xi))\, d\xi, \quad v_i = b_i - a_i + 1, \quad w = (rs^2)^{s/r} h^{1/r}.$$

Here $\Sigma_{i=1}^{n} u_i$ is the area of (12.9) (which is contained in $\mathfrak{D}_2$), hence is $\leq A_2 \ll 1$. Since for given ξ there are $\ll s$ real numbers of η with $F(\xi, \eta) = \pm h$, we have

$$\sum_{i=1}^{n} v_i = \left(\sum_{i=1}^{n} (b_i - a_i) \right) + n \ll sY + n.$$

Thus the number of integer points in (12.9) is

$$\begin{aligned} &\ll sY + nw \ll sY_S + c_{10}(r, s)\,(rs^2)^{s/r} h^{1/r} \\ &\ll c_3\, h^{1/(r-2s)} + c_{10}(r, s)\,(rs^2)^{s/r} h^{1/r} \\ &\leq c_{11}(r, s)\, h^{1/(r-2s)}, \end{aligned}$$

where

$$c_{11}(r, s) = \min(e^{3400 \log^3 r}, c_{12}(s)).$$

By symmetry reasons mentioned above we have

$$Z_2 \ll c_{11}(r, s)\, h^{1/(r-2s)}.$$

Combining this with (12.4), (12.5), (12.8) we obtain

$$|Z_F(h) - A_F h^{2/r}| \ll c_4(r, s)\,(h^{1/(r-2s)} + h^{1/r} \log h^{1/r}),$$

i.e. Theorem 3.

References

[1] Bombieri, E., & Schmidt, W. M., On Thue's equation. *Invent. Math.*, 88 (1987), 69–81.
[2] Chowla, S., Contributions to the analytic theory of numbers II. *J. Indian Math. Soc.*, 20 (1933), 120–128.
[3] Davenport, H., Note on a principle of Lipschitz. *J. London Math. Soc.*, 26 (1951), 179–183.
[4] Evertse, J. H., On the equation $ax^n-by^n=c$. *Compositio Math.*, 47 (1982), 288–315.
[5] — Upper bounds for the number of solutions of diophantine equations. *Math. Centrum Amsterdam*, 1983, 1–127.
[6] Hyyrö, S., Über die Gleichung $ax^n-by^n=c$ und das Catalansche Problem. *Ann. Acad. Sci Fenn., Ser. AI*, 355 (1964), 1–50.
[7] Khovansky, A. G., On a class of transcendental equations. *Soviet Math. Dokl.*, 22 (1980), 762–765.
[8] — Sur les racines complexes des systèmes d'équations algébriques comportant peu de termes. *C.R. Acad. Sci. Paris*, 292 (1981), 937–940.
[9] — Fewnomials and Pfaff manifolds. *Proc. Int. Cong. Math.*, Warsaw 1983, pp. 549–564.
[10] Mahler, K., Zur Approximation algebraischer Zahlen. III. *Acta Math.*, 62 (1934), 91–166.
[11] — On the lattice points on curves of genus 1. *Proc. London Math. Soc.* (2), 39 (1935), 431–466.
[12] — An application of Jensen's formula for polynomials. *Mathematika*, 7 (1960), 98–100.
[13] Mueller, J., Counting solutions of $|ax^r-by^r|\leq h$. *Quart. J. Math. Oxford* (to appear).
[14] Mueller, J. & Schmidt, W. M., Trinomial Thue equations and inequalities. *J. Reine Angew. Math.*, 379 (1987), 76–99.
[15] Ramachandra, K., A lattice point problem for norm forms in several variables, *J. Number Theory*, 1 (1969), 534–555.
[16] Risler, J. J., Complexité et Géométrie Réelle. *Sém. Bourbaki* (1984–85), no. 637, pp. 89–100.
[17] Schmidt, W. M., Thue equations with few coefficients. *Trans. Amer. Math. Soc.*, 303 (1987), 241–255.
[18] Siegel, C. L., Über einige Anwendungen diophantischer Approximationen. *Abh. Preuss. Akad. Wiss. Phys.-math. Kl.*, 1929, Nr. 1.
[19] Silverman, J., Integer points on curves of genus 1. *J. London Math. Soc.* (2), 28 (1983), 1–7.

Received February 9, 1987
Received in revised form August 24, 1987