TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 317, Number 1, January 1990

THE NUMBER OF SOLUTIONS OF NORM FORM EQUATIONS

WOLFGANG M. SCHMIDT

ABSTRACT. A norm form is a form $F(X_1, \dots, X_n)$ with rational coefficients which factors into linear forms over $\mathbf{C}$ but is irreducible or a power of an irreducible form over $\mathbf{Q}$. It is known that a nondegenerate norm form equation $F(x_1, \dots, x_n) = m$ has only finitely many solutions $(x_1, \dots, x_n) \in \mathbf{Z}^n$. We derive explicit bounds for the number of solutions. When F has coefficients in $\mathbf{Z}$, these bounds depend only on n, m and the degree of F, but are independent of the size of the coefficients of F.

1. INTRODUCTION

A *norm form* $F(\underline{\underline{X}}) = F(X_1, \dots, X_n)$ is a form

$$F(\underline{\underline{X}}) = a\mathfrak{N}(\alpha_1 X_1 + \cdots + \alpha_n X_n)$$

where a is a nonzero rational number, where $\alpha_1, \dots, \alpha_n$ lie in an algebraic number field K of some degree r, and where $\mathfrak{N}$ is the norm from K to $\mathbf{Q}$. Thus if $\sigma_1, \dots, \sigma_r$ are the isomorphic embeddings of K into $\mathbf{C}$, and if we set $\alpha^{(i)} = \sigma_i(\alpha)$ for $\alpha \in K$, then

$$\mathfrak{N}(\alpha_1 X_1 + \cdots + \alpha_n X_n) = \prod_{i=1}^{r} (\alpha_1^{(i)} X_1 + \cdots + \alpha_n^{(i)} X_n).$$

With $L(\underline{\underline{X}}) = \alpha_1 X_1 + \cdots + \alpha_n X_n$ and $L^{(i)}(\underline{\underline{X}}) = \alpha_1^{(i)} X_1 + \cdots + \alpha_n^{(i)} X_n$ we have

$$F(\underline{\underline{X}}) = aL^{(1)}(\underline{\underline{X}}) \cdots L^{(r)}(\underline{\underline{X}}).$$

In particular, such a norm form is a form of degree r with rational coefficients. We shall assume throughout that $\alpha_1, \dots, \alpha_n$ are linearly independent over $\mathbf{Q}$.

As $\underline{\underline{x}} = (x_1, \dots, x_n)$ runs through $\mathbf{Z}^n$, the values of $\alpha_1 x_1 + \cdots + \alpha_n x_n$ will run through a $\mathbf{Z}$-module $\mathfrak{M}$ contained in K. Let $\mathbf{Q}\mathfrak{M}$ consist of the values of $\alpha_1 x_1 + \cdots + \alpha_n x_n$ as $\underline{\underline{x}}$ runs through $\mathbf{Q}^n$. Given a subfield E of K, let $\mathfrak{M}^E$ consist of the elements μ of $\mathfrak{M}$ such that $\lambda\mu \in \mathbf{Q}\mathfrak{M}$ for every $\lambda \in E$. Then $\mathfrak{M}^E$ is a submodule of $\mathfrak{M}$. We call $\mathfrak{M}$ *degenerate* if there is a subfield E of K which is neither $\mathbf{Q}$ nor imaginary quadratic, such that $\mathfrak{M}^E \neq \{0\}$. Otherwise, $\mathfrak{M}$ is *nondegenerate*. We will call the norm form F degenerate

Received by the editors May 31, 1988.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 11D57, 11D41.
Written with partial support from NSF grant DMS-8603093.

or nondegenerate when $\mathfrak{M}$ is. Since $\mathfrak{M}$ is determined by F up to a factor of proportionality, this definition is legitimate. In [8] (and again in [9]) we showed that *when F is nondegenerate and m is a given integer, then the norm form equation*

$$F(\underline{\underline{x}}) = m \tag{1.1}$$

has at most finitely many solutions $\underline{\underline{x}} \in \mathbf{Z}^n$.

In this paper we will give upper bounds for the number of solutions. It will be convenient to begin with the case $m = 1$.

Theorem 1. *Suppose F is a nondegenerate norm form with coefficients in $\mathbf{Z}$. The number of integer solutions of*

$$F(\underline{\underline{x}}) = 1 \tag{1.2}$$

is under some bound $c_1(n, r)$. In particular we may take

$$c_1(n, r) = \min(r^{2^{30n} r^2}, r^{c_2(n)}) \tag{1.3}$$

with $c_2(n) = (2n)^{n2^{n+4}}$.

The most interesting aspect of our theorem is that the bound does not depend on the coefficients of F, and in particular it does not depend on K, except for its degree r. The second bound implicit in (1.3) grows rather fast as a function of n, but it grows only like a polynomial in r when n is given.

For a given norm form, the estimate of Theorem 1 can often be improved by determining the ranks of certain matrices. A particular such improvement is as follows. Let N be a normal extension of $\mathbf{Q}$ containing the conjugate fields $K^{(1)}, \dots, K^{(r)}$, and let G be the Galois group of N over $\mathbf{Q}$. Then G acts on the set $\{L^{(1)}, \dots, L^{(r)}\}$ by acting on the coefficients of the forms. We say that *G acts t times transitively* if $L^{(1)}, \dots, L^{(r)}$ are distinct and if for any distinct $i_1, \dots, i_t$ there is a $\gamma \in G$ with $\gamma(L^{(j)}) = L^{(i_j)}$ $(j = 1, \dots, t)$.

Theorem 2. *Let F be a norm form with coefficients in $\mathbf{Z}$. Assume that $r > n$ and any n of the forms $L^{(1)}, \dots, L^{(r)}$ are linearly independent, and that G acts $n - 1$ times transitively. Then the number of solutions of* (1.2) *is* $\leq r^{2^{32n}}$.

A Thue equation is a norm form equation in two variables: $n = 2$. For Thue equations a bound independent of the coefficients of F was first established by Evertse [5]. In [1] a bound $c_1(2, r) = c_3 r$ was established, where c_3 is absolute. Whereas the Thue equation is connected with rational approximation to an irrational algebraic number, norm form equations with $n > 2$ are connected with simultaneous approximations. As our tool from simultaneous approximations we will use a quantitative version of the "Subspace Theorem" as proved in [11]. One of the reasons why our estimates for general n are worse than for Thue equations is that for approximations to a single algebraic number α, the more classical "Roth's Lemma" may be replaced by a recent theorem due to Esnault and Viehweg [4], or even by the more special "Dyson's Lemma". A generalized version of the theorem of Esnault and Viehweg which is applicable to simultaneous approximations would lead to better values of $c_1(n, r)$; but as of now, such a version has not been established.

An integer point $\underline{\underline{x}} = (x_1, \dots, x_n)$ is *primitive* if g.c.d.$(x_1, \dots, x_n) = 1$.

Theorem 3. *Suppose F is a nondegenerate norm form with coefficients in* $\mathbf{Z}$. *The number of primitive solutions of* (1.1) *where $m > 0$ does not exceed $c_1(n, r)\cdot c_4(n, r, m)$, where we may take*

$$c_4(n, r, m) = \binom{r}{n-1}^{\omega} d_{n-1}(m^r),$$

with ω the number of distinct prime factors of m, and the function $d_{n-1}(k)$ denoting the number of factorizations of k into $n-1$ positive factors: $k = k_1 k_2 \cdots k_{n-1}$.

For forms as in Theorem 2, the number of solutions of (1.1) *does not exceed $r^{2^{32n}} c_4(n, r, m)$.*

In particular, this leads to a bound which depends only on n, r, Ω, where Ω is the total number of prime factors of m. It does not depend on the size of the prime factors of m.

Now $d_1(k) = 1$, so that for $n = 2$, we may take $c_4(2, r, m) = r^{\omega}$, as had already been shown in [1], where the bound $c_3 r^{1+\omega}$ had been established for the number of primitive solutions of (1.1). In general, $d_{n-1}(k) \leq d(k)^{n-2}$, where $d(k) = d_2(k)$ is the number of positive divisors of k. Moreover, when m has the prime factorization $m = p_1^{u_1} \cdots p_{\omega}^{u_{\omega}}$, then

$$d(m^r) = (u_1 r + 1) \cdots (u_{\omega} r + 1) < r^{\omega}(u_1 + 1) \cdots (u_{\omega} + 1) = r^{\omega} d(m).$$

Therefore

$$c_4(n, r, m) = \binom{r}{n-1}^{\omega} d_{n-1}(m^r) < r^{\omega}(r^{\omega} d(m))^{n-2} = r^{\omega(n-1)} d(m)^{n-2}.$$

Now $\omega(m) < (1+\delta)\log m/\log\log m$ and $d(m) < 2^{(1+\delta)\log m/\log\log m}$ when $\delta > 0$ and $m > m_0(\delta)$ (see e.g., [6, §22.10, and Theorem 317]), so that for $m > c_5(n)$

$$c_4(n, r, m) < (2r)^{n\log m/\log\log m} = m^{n\log(2r)/\log\log m}.$$

We may conclude that given $\varepsilon > 0$ the number of primitive solutions of (1.1) is below $c_6(n, r, \varepsilon)m^{\varepsilon}$. It is easily seen that the same type of bound holds for the number of all integer solutions of (1.1).

In the Thue case, i.e., for $n = 2$, I showed in [10] that the number of solutions of the inequality

$$|F(\underline{\underline{x}})| \leq m \tag{1.4}$$

is below $c_7 r m^{2/r}(1 + \log m^{1/r})$. The logarithmic term here is probably unnecessary. In general, I conjecture that the number of solutions of inequality (1.4) is below $c_8(n,r)m^{n/r}$, or perhaps $c_8(n,r)m^{n/r}$ times a logarithmic factor. The present method would give a much weaker estimate.

The proof of Theorems 1 and 2 will be by induction on n. Since the case $n=1$ is trivial, we will suppose throughout that $n \geq 2$. In fact, since for $n=2$ we are dealing with Thue equations which were treated in [1], we could suppose that $n \geq 3$. Using the Subspace Theorem from Diophantine approximations, we will show that the solutions of (1.2) lie in not more than $c_9(n,r)$ proper subspaces. The integer points in such a subspace may be parametrized by $\mathbf{Z}^{n-1}$, so that we are reduced to dealing with norm form equations in $n-1$ variables. Crucial for the induction argument in Theorem 1 is the rather obvious fact that a submodule of a nondegenerate module is again nondegenerate. We remark that in contrast to norm form equations, we have no way at present to make an inductive counting argument work for simultaneous approximations. We do know that when $\alpha_1, \dots, \alpha_n$ are algebraic, with $1, \alpha_1, \dots, \alpha_n$ linearly independent over $\mathbf{Q}$, then there are only finitely many simultaneous rational approximations $(p_1/q, \dots, p_n/q)$ with

$$\left|\alpha_i - \frac{p_i}{q}\right| < \frac{1}{q^{1+(1/n)+\varepsilon}} \qquad (i = 1, \dots, n)$$

where $\varepsilon > 0$ is given. But when $n > 1$, we are unable to provide any upper bounds whatsoever for the number of such approximations.

As in the work on Thue equations [1, 7, 10], we will distinguish "small" and "large" solutions. The treatment of the small solutions is perhaps the most interesting part of our work. Again, as for Thue equations, we shall employ the initial step of "jacking up the height".

The second bound implicit in Theorem 1, with $c_1(n,r) \leq r^{c_2(n)}$, depends on estimates of the ranks of subsets of the conjugate linear forms $L^{(1)}, \dots, L^{(r)}$. These estimates can be formulated in terms of rather general sets of algebraic points in n-space, and will be derived in an Appendix which is independent of the rest of the paper.

Theorem 3 will be derived from Theorems 1 and 2 by a p-adic method.

Symbols $X, Y, \dots$ will denote variables, and $x, y, \dots$ will denote rational integers or elements of a given field.

I wish to thank the referee, and also H. P. Schlickewei, for pointing out a number of inaccuracies in my original manuscript.

2. Heights and discriminants

Let K be an algebraic number field. By an absolute value of K we will always understand an absolute value which is normalized so that it extends either the standard absolute value or a p-adic absolute value of $\mathbf{Q}$. Given such an absolute value $|\cdot|_w$ of K, let n_w be its local degree. Let $M(K)$ be a set of symbols v, such that with every $v \in M(K)$ there is associated an absolute value $|\cdot|_v$ of K, and moreover every absolute value $|\cdot|_w$ of K is obtained for precisely n_w elements v of $M(K)$. In other words, one could say that $M(K)$ is the set of absolute values of K with multiplicities, so that a given $|\cdot|_w$ occurs n_w times. With this convention we have the product formula

$$\prod_{v\in M(K)} |\alpha|_v = 1 \quad \text{for } \alpha \in K,\ \alpha \neq 0.$$

Given $\underline{\underline{\alpha}} = (\alpha_1, \ldots, \alpha_n) \in K^n$ and given $v \in M(K)$, put

$$|\underline{\underline{\alpha}}|_v = \begin{cases} (|\alpha_1|_v^2 + \cdots + |\alpha_n|_v^2)^{1/2} & \text{if } v \text{ is archimedean},\\ \max(|\alpha_1|_v, \ldots, |\alpha_n|_v) & \text{if } v \text{ is nonarchimedean}.\end{cases}$$

Here we call v archimedean or nonarchimedean when $|\cdot|_v$ is. Thus v is archimedean precisely if $|\cdot|_v$ extends the standard absolute value of $\mathbf{Q}$. When $\underline{\underline{\alpha}} \neq \underline{\underline{0}}$ we define its field height $H_K(\underline{\underline{\alpha}})$ by

$$H_K(\underline{\underline{\alpha}}) = \prod_{v\in M(K)} |\underline{\underline{\alpha}}|_v .$$

By the product formula $H_K(\lambda\underline{\underline{\alpha}}) = H_K(\underline{\underline{\alpha}})$ for $\lambda \neq 0$ in K. The absolute height is $H(\underline{\underline{\alpha}}) = H_K(\underline{\underline{\alpha}})^{1/r}$ where r is the degree of K.

When

$$L(\underline{\underline{X}}) = \alpha_1 X_1 + \cdots + \alpha_n X_n = \underline{\underline{\alpha}}\,\underline{\underline{X}} \tag{2.1}$$

is a nonzero linear form with coefficients in K, we put

$$H_K(L) = H_K(\underline{\underline{\alpha}}),\ H(L) = H(\underline{\underline{\alpha}}).$$

When L has complex coefficients we put

$$|L| = (|\alpha_1|^2 + \cdots + |\alpha_n|^2)^{1/2},$$

where the absolute values on the right indicate the standard absolute value of $\mathbf{C}$. In particular, when L has coefficients in K, the forms $L^{(i)}$ $(i = 1, \ldots, r)$ as defined in the Introduction have coefficients in $\mathbf{C}$, so that $|L^{(i)}|$ is well defined.

Let $L(\underline{\underline{X}})$ and

$$F(\underline{\underline{X}}) = aL^{(1)}(\underline{\underline{X}}) \cdots L^{(r)}(\underline{\underline{X}}) \tag{2.2}$$

be as in the Introduction, so that in particular $F \neq 0$ and F has coefficients in $\mathbf{Z}$. Let $\operatorname{Cont} F$ be the greatest common divisor of the coefficients of F. We define a height $H^*(F)$ by

$$H^*(F) = |a||L^{(1)}| \cdots |L^{(r)}|. \tag{2.3}$$

Even though a and L are determined by F only up to factors, the height $H^*(F)$ clearly depends on F only.

Lemma 1. $H^*(F) = (\operatorname{Cont} F) \cdot H_K(L)$.

Proof. Let N be the compositum of $K^{(1)}, \dots, K^{(r)}$. We shall suppose throughout that K is embedded in $\mathbf{C}$ and that $\alpha^{(1)} = \alpha$ for $\alpha \in K$, so that $K^{(1)} = K$. Write $k = [N : K]$. There is a $k : 1$ map $\varphi \colon M(N) \to M(K)$ such that for $u \in M(N)$ and $v = \varphi(u)$, the restriction of $|\cdot|_u$ to K is the absolute value $|\cdot|_v$ of K. Write $u|p$ if $|\cdot|_u$ extends the p-adic absolute value of $\mathbf{Q}$, and $u|\infty$ if it extends the standard absolute value of $\mathbf{Q}$. Now

$$|a|^{-k} H_N(L) = |a|^{-k} \prod_{u \in M(N)} |L|_u, \tag{2.4}$$

where $|L|_u = |\underline{\underline{\alpha}}|_u$ if $L(\underline{\underline{X}}) = \underline{\underline{\alpha X}}$. Note that

$$\prod_{\substack{u \in M(N) \\ u|\infty}} |L|_u = \left(\prod_{\substack{u \in M(K) \\ v|\infty}} |L|_v \right)^k = (|L^{(1)}| \cdots |L^{(r)}|)^k. \tag{2.5}$$

On the other hand for given p,

$$\prod_{\substack{u \in M(N) \\ u|p}} |L|_u^r = \prod_{\substack{u \in M(N) \\ u|p}} (|L^{(1)}|_u \cdots |L^{(r)}|_u),$$

so that

$$\begin{aligned} |a|^{-kr} \prod_{\substack{u \in M(N) \\ u \nmid \infty}} |L|_u^r &= \prod_{\substack{u \in M(N) \\ u \nmid \infty}} (|a|_u |L^{(1)}|_u \cdots |L^{(r)}|_u) \\ &= \prod_{\substack{u \in M(N) \\ u \nmid \infty}} |F|_u = \left(\prod_p |F|_p \right)^{kr} = (\operatorname{Cont} F)^{-kr} \end{aligned} \tag{2.6}$$

by Gauss' Lemma, where $|F|_u$ is the maximum value of $|f|_u$ over the coefficients f of F. Substituting (2.5) and the rth root of (2.6) into (2.4), we obtain

$$|a|^{-k} H_N(L) = (|L^{(1)}| \cdots |L^{(r)}|)^k (\operatorname{Cont} F)^{-k}.$$

The lemma follows upon taking kth roots.

We suppose throughout that the coefficients $\alpha_1, \dots, \alpha_n$ of L are linearly independent over $\mathbf{Q}$. Then the matrix $\alpha_j^{(i)}$ $(1 \le i \le r,\ 1 \le j \le n)$ has rank n, so that there are n linearly independent forms among $L^{(1)}, \dots, L^{(r)}$. Let I be the set of n-tuples of integers $i_1, \dots, i_n$ in $1 \le i \le r$ such that $L^{(i_1)}, \dots, L^{(i_n)}$ are linearly independent. Suppose 1 occurs exactly in q of the n-tuples of I. Since the Galois group G of compositum N acts (one time) transitively on $L^{(1)}, \dots, L^{(r)}$, each integer i in $1 \le i \le r$ occurs in precisely

q of the n-tuples of I. Thus $rq = n|I|$ where $|I|$ is the cardinality of I, and $|I|n/r$ is an integer. Put

$$D = D(F) = |a|^{|I|n/r} \prod_{(i_1, \dots, i_n) \in I} |\det(L^{(i_1)}, \dots, L^{(i_n)})|. \tag{2.7}$$

The determinant of n linear forms here is the determinant of their coefficient matrix. Again, even though a and the forms $L^{(i)}$ are determined by F only up to factors, the *discriminant* $D(F)$ depends on F only.

An element γ of the Galois group G of N permutes $L^{(1)}, \dots, L^{(r)}$, say $\gamma(L^{(i)}) = L^{(\hat{\gamma}(i))}$ where $\hat{\gamma}$ is a permutation of $1, \dots, r$. It is clear that when $(i_1, \dots, i_n)$ is in I, then so is $(\hat{\gamma}(i_1), \dots, \hat{\gamma}(i_n))$. From this it follows that the product $\prod(\det(L^{(i_1)}, \dots, L^{(i_n)}))$ over $(i_1, \dots, i_n) \in I$ is invariant under G and lies in $\mathbf{Q}$. Since $|I|n/r$ lies in $\mathbf{Z}$, the discriminant D lies in $\mathbf{Q}$.

Now let $u \in M(N)$ be nonarchimedean. Then

$$|a|_u |L^{(1)}|_u \cdots |L^{(r)}|_u = |F|_u \leq 1$$

by Gauss' Lemma and since F has coefficients in $\mathbf{Z}$. Observe that

$$|\det(L^{(i_1)}, \dots, L^{(i_n)})|_u \leq |L^{(i_1)}|_u \cdots |L^{(i_n)}|_u. \tag{2.8}$$

Now each $L^{(i)}$ occurs $q = |I|n/r$ times in some determinant on the right-hand side of (2.7), so that

$$|D|_u \leq (|a|_u |L^{(1)}|_u \cdots |L^{(r)}|_u)^{|I|n/r} \leq 1.$$

Since this holds for every nonarchimedean $u \in M(N)$, we may conclude that D is an integer, in fact a nonzero rational integer, so that

$$|D| \geq 1. \tag{2.9}$$

On the other hand, since (2.8) holds for archimedean absolute values as well,

$$|D| \leq (|a||L^{(1)}| \cdots |L^{(r)}|)^{|I|n/r} = H^*(F)^{|I|n/r}. \tag{2.10}$$

Given a linear map $T: \mathbf{R}^n \to \mathbf{R}^n$ mapping integer points into integer points, and given a polynomial $P(\underline{X})$, put $P^T(\underline{X}) = P(T(\underline{X}))$. We have

$$\det(L^{(i_1)T}, \dots, L^{(i_n)T}) = (\det T) \cdot (\det(L^{(i_1)}, \dots, L^{(i_n)})),$$

so that

$$D(F^T) = |\det T|^{|I|} D(F). \tag{2.11}$$

3. Comparison of invariants

Write $F \sim G$ and call F, G *equivalent* if there is a $T \in SL(n, \mathbf{Z})$ with $F^T = G$. The height $H^*(F)$ is not an invariant; equivalent forms in general will not have the same height H^*. We therefore now introduce the invariant height

$$\mathfrak{H}(F) = \min_{G \sim F} H^*(G).$$

Note that the value set of H^* is discrete ([12, Theorem 5.11 of Chapter VIII], but note that the height in [12] is slightly different from ours), so that the minimum does exist.

Given a linear form $L(\underline{\underline{X}}) = \alpha_1 X_1 + \cdots + \alpha_n X_n$ with coefficients in K, introduce the column vectors

$$(3.1) \qquad \underline{\underline{a}}_j = \begin{pmatrix} \alpha_j^{(1)} \\ \vdots \\ \alpha_j^{(r)} \end{pmatrix} \qquad (j = 1, \ldots, n)$$

with r components. Let $\underline{\underline{A}}$ be the exterior product:

$$\underline{\underline{A}} = \underline{\underline{a}}_1 \wedge \cdots \wedge \underline{\underline{a}}_n,$$

so that $\underline{\underline{A}}$ has $l = \binom{r}{n}$ components. Put $\Delta(L) = |\underline{\underline{A}}|$, where, as always, we set $|\underline{\underline{A}}| = (|\gamma_1|^2 + \cdots + |\gamma_l|^2)^{1/2}$ when $\underline{\underline{A}} = (\gamma_1, \ldots, \gamma_l)$. Then $\Delta(L^T) = \Delta(L)$ for $T \in SL(n, \mathbf{Z})$, so that $\Delta(L)$ is invariant.

Lemma 2. *Suppose the norm form F is given by* (2.2). *Then*

$$(3.2) \qquad |a|^{n/r}\Delta(L) \geq 2^{-n} n^{-3/2} V(n) r^{n/2} (\operatorname{Cont} F)^{(n-1)/r} \mathfrak{H}(F)^{1/r},$$

where $V(n)$ is the volume of the unit ball in $\mathbf{R}^n$.

The right-hand side here depends on F only, but the left-hand side depends on the particular representation of F as in (2.2). Given $\lambda \neq 0$ in K we may set $L' = \lambda L$, so that $L'^{(i)} = \lambda^{(i)} L^{(i)}$, and we may write $F = a' L'^{(i)} \cdots L'^{(r)}$ with $a' = a(\lambda^{(1)} \cdots \lambda^{(r)})^{-1}$. There is then no simple relation between $\Delta(L)$ and $\Delta(L')$.

Proof. Since $\alpha_1, \ldots, \alpha_n$ are linearly independent over $\mathbf{Q}$, the matrix $(\alpha_j^{(i)})$ $(1 \leq i \leq r,\ 1 \leq j \leq n)$ has rank n, and $\underline{\underline{a}}_1, \ldots, \underline{\underline{a}}_n$ are linearly independent vectors in $\mathbf{C}^r$. In particular, $\underline{\underline{A}} \neq \underline{\underline{0}}$ and $\Delta(L) > 0$. Now $\underline{\underline{a}}_1, \ldots, \underline{\underline{a}}_n$ generate a $\mathbf{Z}$-module Λ^* of rank n. Even though Λ^* is not a point lattice in $\mathbf{R}^n$, we may define successive minima $\mu_1^*, \ldots, \mu_n^*$: here μ_j^* is least such that there are j linearly independent points $\underline{\underline{g}}$ in Λ^* with $|\underline{\underline{g}}| \leq \mu_j^*$.

Every point $\underline{\underline{z}} = u_1\underline{\underline{a}}_1 + \cdots + u_n\underline{\underline{a}}_n$ of Λ^*, say with components $z_1, \ldots, z_r$, has

$$|a||z_1 \cdots z_r| = \left| a \prod_{i=1}^{r} L^{(i)}(u_1, \ldots, u_n) \right| = |F(u_1, \ldots, u_n)|.$$

Now when $\underline{\underline{z}} \neq \underline{\underline{0}}$, so that $(u_1, \ldots, u_n) \in \mathbf{Z}^n \backslash \underline{\underline{0}}$, then $F(u_1, \ldots, u_n) \neq 0$ by the linear independence of $\alpha_1, \ldots, \alpha_n$. Thus $|a||z_1 \cdots z_r| = |F(u_1, \ldots, u_n)| \geq$

$\mathrm{Cont}\,F$, and $|z_1\cdots z_r| \geq C/|a|$, where $C = \mathrm{Cont}\,F$. By the arithmetic-geometric inequality,

$$|\underline{\underline{z}}| = (|z_1|^2 + \cdots + |z_r|^2)^{1/2} \geq \left(r\sqrt[r]{|z_1|^2\cdots|z_r|^2}\right)^{1/2} \geq r^{1/2}(C/|a|)^{1/r}.$$

Therefore the successive minima $\mu_j^* \geq r^{1/2}(C/|a|)^{1/r}$ $(j = 1, \ldots, n)$.

Suppose now that

$$\underline{\underline{b}}_j = \begin{pmatrix} \beta_j^{(1)} \\ \vdots \\ \beta_j^{(r)} \end{pmatrix} \qquad (j = 1, \ldots, n)$$

is another basis of Λ^*. Say $\underline{\underline{b}}_j = u_{j1}\underline{\underline{a}}_1 + \cdots + u_{jn}\underline{\underline{a}}_n$ $(j = 1, \ldots, n)$, with a matrix $U = (u_{jk}) \in SL(n, \mathbf{Z})$. Then the matrices $A = (\alpha_j^{(i)})$, $B = (\beta_j^{(i)})$ (where i denotes the row and j the column) have $B = AV$ where V is the transpose of U. In particular, the rows $\underline{\underline{\alpha}}^{(i)} = (\alpha_1^{(i)}, \ldots, \alpha_n^{(i)})$, $\underline{\underline{\beta}}^{(i)} = (\beta_1^{(i)}, \ldots, \beta_n^{(i)})$ have $\underline{\underline{\beta}}^{(i)} = \underline{\underline{\alpha}}^{(i)}V$ $(i = 1, \ldots, r)$. The form

$$F^V(\underline{\underline{X}}) = F(V\underline{\underline{X}}) = a\prod_{i=1}^{r}(\underline{\underline{\alpha}}^{(i)}V\underline{\underline{X}}) = a\prod_{i=1}^{r}(\underline{\underline{\beta}}^{(i)}\underline{\underline{X}})$$

is equivalent to F. Therefore, by (2.3)

$$H^*(F^V) = |a|\prod_{i=1}^{r}|\underline{\underline{\beta}}^{(i)}| = |a|\prod_{i=1}^{r}(|\beta_1^{(i)}|^2 + \cdots + |\beta_n^{(i)}|^2)^{1/2} \geq \mathfrak{H}(F).$$

By the arithmetic-geometric inequality,

$$\sum_{i=1}^{r}(|\beta_1^{(i)}|^2 + \cdots + |\beta_n^{(i)}|^2) \geq r(\mathfrak{H}(F)/|a|)^{2/r},$$

so that

$$|\underline{\underline{b}}_1|^2 + \cdots + |\underline{\underline{b}}_n|^2 \geq r(\mathfrak{H}(F)/|a|)^{2/r}. \tag{3.3}$$

There is a basis $\underline{\underline{b}}_1, \ldots, \underline{\underline{b}}_n$ of Λ^* with $|\underline{\underline{b}}_j| \leq j\mu_j^*$ $(j = 1, \ldots, n)$ (see, e.g. Cassels [3, Lemma 8, p. 135]. Thus

$$|\underline{\underline{b}}_1|^2 + \cdots + |\underline{\underline{b}}_n|^2 \leq (1^2 + 2^2 + \cdots + n^2)\mu_n^{*2} \leq n^3\mu_n^{*2},$$

and

$$\mu_n^* \geq n^{-3/2}r^{1/2}(\mathfrak{H}(F)/|a|)^{1/r}.$$

In conjunction with the estimate for μ_j^* given above this yields

$$\mu_1^*\cdots\mu_n^* \geq n^{-3/2}r^{n/2}C^{(n-1)/r}|a|^{-n/r}\mathfrak{H}(F)^{1/r}. \tag{3.4}$$

Rather than do geometry of numbers in complex space such as, e.g. in [2], we now proceed as follows. Say $\sigma_1, \ldots, \sigma_r$ are the embeddings of K into $\mathbf{C}$, with $\sigma_1, \dots, \sigma_s$ real and $\sigma_{s+1}, \sigma_{s+2}, \dots, \sigma_{r-1}, \sigma_r$ complex conjugate in pairs. Then Λ^* lies in the space S of vectors

$$\underline{\underline{z}} = \begin{pmatrix} z_1 \\ \vdots \\ z_r \end{pmatrix}$$

where $z_1, \dots, z_s$ are real and $\overline{z_{s+1}} = z_{s+2}, \dots, \overline{z_{r-1}} = z_r$. Let $(\underline{\underline{z}}, \underline{\underline{z}}')$ be the inner product $z_1 z_1' + \cdots + z_s z_s' + z_{s+1}\overline{z_{s+1}'} + z_{s+2}\overline{z_{s+2}'} + \cdots + z_r \overline{z_r'}$. Let $\underline{\underline{f}}$ be the map $S \to \mathbf{R}^r$ with

$$\underline{\underline{f}}(\underline{\underline{z}}) = \begin{pmatrix} w_1 \\ \vdots \\ w_r \end{pmatrix}$$

where $w_1 = z_1, \dots, w_s = z_s$, $w_{s+1} = \sqrt{2}\operatorname{Re} z_{s+1} = (\sqrt{2})^{-1}(z_{s+1} + z_{s+2})$, $w_{s+2} = \sqrt{2}\operatorname{Im} z_{s+1} = (i\sqrt{2})^{-1}(z_{s+1} - z_{s+2}), \dots, w_{r-1} = \sqrt{2}\operatorname{Re} z_{r-1}$, $w_r = \sqrt{2}\operatorname{Im} z_{r-1}$. Then the inner product

$$\begin{aligned} \underline{\underline{f}}(\underline{\underline{z}}) \cdot \underline{\underline{f}}(\underline{\underline{z}}') &= z_1 z_1' + \cdots + z_s z_s' + 2(\operatorname{Re} z_{s+1})(\operatorname{Re} z_{s+1}') + 2(\operatorname{Im} z_{s+1})(\operatorname{Im} z_{s+1}') \\ &\quad + \cdots + 2(\operatorname{Im} z_{r-1})(\operatorname{Im} z_{r-1}') \\ &= z_1 z_1' + \cdots + z_s z_s' + z_{s+1}\overline{z_{s+1}'} + z_{s+2}\overline{z_{s+2}'} + \cdots + z_r \overline{z_r'} \\ &= (\underline{\underline{z}}, \underline{\underline{z}}'). \end{aligned}$$

Thus $\underline{\underline{f}}$ preserves inner products. For any $\underline{\underline{z}}_1, \dots, \underline{\underline{z}}_n$ in S, Laplace's identity yields

$$\text{(3.5)} \quad |\underline{\underline{f}}(\underline{\underline{z}}_1) \wedge \cdots \wedge \underline{\underline{f}}(\underline{\underline{z}}_n)|^2 = \det(\underline{\underline{f}}(\underline{\underline{z}}_i) \cdot \underline{\underline{f}}(\underline{\underline{z}}_j)) = \det((\underline{\underline{z}}_i, \underline{\underline{z}}_j)) = |\underline{\underline{z}}_1 \wedge \cdots \wedge \underline{\underline{z}}_n|^2 .$$

Now $\underline{\underline{f}}$ maps Λ^* into an n-dimensional lattice Λ contained in $\mathbf{R}^r$. Since $\underline{\underline{f}}$ preserves inner products, the successive minima $\mu_1, \dots, \mu_n$ of Λ have $\mu_j = \mu_j^*$ $(i = 1, \dots, n)$, so that

$$\mu_1 \cdots \mu_n \geq n^{-3/2} r^{n/2} C^{(n-1)/r} |a|^{-n/r} \mathfrak{H}(F)^{1/r}$$

by (3.4). On the other hand by Minkowski,

$$\mu_1 \cdots \mu_n V(n) \leq 2^n \det \Lambda ,$$

so that

$$\text{(3.6)} \qquad \det \Lambda \geq \frac{V(n)}{n^{3/2} 2^n} r^{n/2} C^{(n-1)/r} |a|^{-n/r} \mathfrak{H}(F)^{1/r} .$$

Since $\underline{\underline{f}}(\underline{\underline{a}}_1), \dots, \underline{\underline{f}}(\underline{\underline{a}}_n)$ are a basis of Λ, we have

$$(\det \Lambda)^2 = |\underline{\underline{f}}(\underline{\underline{a}}_1) \wedge \cdots \wedge \underline{\underline{f}}(\underline{\underline{a}}_n)|^2 = |\underline{\underline{a}}_1 \wedge \cdots \wedge \underline{\underline{a}}_n|^2 = |\underline{\underline{A}}|^2$$

by (3.5). Substitution into (3.6) yields

$$|a|^{n/r} |\underline{\underline{A}}| \geq \frac{V(n)}{n^{3/2} 2^n} r^{n/2} C^{(n-1)/r} \mathfrak{H}(F)^{1/r} .$$

4. Products of n linear forms

Lemma 3. *Suppose $F(\underline{\underline{X}})$ is a norm form with coefficients in $\mathbf{Z}$, and written as (2.2). Suppose $\underline{\underline{x}}$ is an integer point with (1.2). Then there are $i_1, \dots, i_n$ with $1 \le i_1 < \cdots < i_n \le r$ and*

$$(4.1) \qquad |L^{(i_1)}(\underline{\underline{x}}) \cdots L^{(i_n)}(\underline{\underline{x}})| \le \frac{2^n \cdot n^{3/2}}{(n!)^{1/2} V(n)} |\det(L^{(i_1)}, \dots, L^{(i_n)})| \mathfrak{H}(F)^{-1/r}.$$

Here $\det(L^{(i_1)}, \dots, L^{(i_n)})$ *is the determinant of the coefficient matrix.*

Proof. Since $F(\underline{\underline{x}}) = aL^{(1)}(\underline{\underline{x}}) \cdots L^{(r)}(\underline{\underline{x}}) = 1$ we have

$$(4.2) \qquad F(\underline{\underline{X}}) = V^{(1)}(\underline{\underline{X}}) \cdots V^{(r)}(\underline{\underline{X}})$$

with $V(\underline{\underline{X}}) = L(\underline{\underline{x}})^{-1} L(\underline{\underline{X}})$. We will apply Lemma 2 to F written as (4.2) (rather than as (2.2)). Then $a = 1$, and $\operatorname{Cont} F = 1$ since $F(\underline{\underline{X}}) \in \mathbf{Z}[\underline{\underline{X}}]$ has $F(\underline{\underline{x}}) = 1$. The vectors (3.1) now become

$$\underline{\underline{a}}_j = \begin{pmatrix} \alpha_j^{(1)}/L^{(1)}(\underline{\underline{x}}) \\ \vdots \\ \alpha_j^{(r)}/L^{(r)}(\underline{\underline{x}}) \end{pmatrix} \qquad (j = 1, \dots, n).$$

Lemma 2 yields

$$(4.3) \qquad |\underline{\underline{a}}_1 \wedge \cdots \wedge \underline{\underline{a}}_n|^2 = |\underline{\underline{A}}|^2 = \Delta(V)^2 \ge 4^{-n} n^{-3} V(n)^2 r^n \mathfrak{H}(F)^{2/r}.$$

Let $D(i_1, \dots, i_n)$ be the $(n \times n)$-determinant formed from the rows $i_1, \dots, i_n$ of the matrix with columns $\underline{\underline{a}}_1, \dots, \underline{\underline{a}}_n$. The left-hand side of (4.3) is

$$\sum_{1 \le i_1 < \cdots < i_n \le r} |D(i_1, \dots, i_n)|^2.$$

There are $\binom{r}{n} \le r^n/n!$ summands, so that for some $i_1 < \cdots < i_n$ we have

$$|D(i_1, \dots, i_n)| > (4^{-n} n^{-3} \cdot n! V(n)^2 \mathfrak{H}(F)^{2/r})^{1/2}.$$

After multiplication by $|L^{(i_1)}(\underline{\underline{x}}) \cdots L^{(i_n)}(\underline{\underline{x}})|$ this becomes

$$\det(L^{(i_1)}, \dots, L^{(i_n)}) \ge \frac{(n!)^{1/2} V(n)}{2^n \cdot n^{3/2}} |L^{(i_1)}(\underline{\underline{x}}) \cdots L^{(i_n)}(\underline{\underline{x}})| \mathfrak{H}(F)^{1/r}.$$

5. A reduction

Given a prime p, consider the matrices

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

and identify them with the linear maps they induce. We have

$$\mathbf{Z}^2 = \bigcup_{j=0}^{p} A_j \mathbf{Z}^2.$$

More generally, when $n \geq 2$, consider the $(n \times n)$-matrices

$$B_j = \begin{pmatrix} A_j & 0 \\ 0 & E \end{pmatrix} \qquad (j = 0, \dots, p),$$

where E is the identity matrix with $n-2$ rows. We have

$$\mathbf{Z}^n = \bigcup_{j=0}^{p} B_j \mathbf{Z}^n .$$

Therefore the number of $\underline{\underline{x}} \in \mathbf{Z}^n$ with $F(\underline{\underline{x}}) = 1$ does not exceed $z_0 + z_1 + \cdots + z_p$, where z_j is the number of $\underline{\underline{x}} \in \mathbf{Z}^n$ with $F^{B_j}(\underline{\underline{x}}) = 1$. Note that $\det B_j = p$, so that

$$D(F^{B_j}) \geq p^{|I|} D(F) \geq p^{|I|}$$

by (2.9), (2.11). Therefore if $N(r, n)$ is the maximum number of solutions of (1.2) for forms F such as in Theorem 1, and $N(r, n; p)$ is the maximum number when F is restricted to forms with

$$D(F) \geq p^{|I|}, \tag{5.1}$$

then

$$N(r, n) \leq (p+1) N(r, n; p). \tag{5.2}$$

By the invariance of $D(F)$, (2.10) implies that $D(F) \leq \mathfrak{H}(F)^{|I|n/r}$, which together with (5.1) gives

$$\mathfrak{H}(F) \geq p^{r/n}. \tag{5.3}$$

Proposition 1. *Suppose that*

$$p > n^{10n^2}. \tag{5.4}$$

Then

$$N(r, n; p) < \min(r^{2^{30n} r^2 - 10n^2}, r^{c_2(n) - 10n^2}).$$

We may apply the proposition with a prime $p \leq 2n^{10n^2} < r^{10n^2}$. Then by (5.2), Theorem 1 follows. It remains for us to prove the proposition.

The number of solutions of F is unchanged if F is replaced by an equivalent form. Call F *reduced* if $\mathfrak{H}(F) = H^*(F)$; every form is equivalent to at least one reduced form. In proving the proposition we may thus suppose that F is reduced.

Now (1.2) has no solution unless $\operatorname{Cont} F = 1$, which we shall assume from now on. Then Lemma 1 says that

$$\mathfrak{H}(F) = H^*(F) = H_K(L) = H(L)^r. \tag{5.5}$$

The relations (5.3), (5.4) yield

$$H(L) > n^{10n}. \tag{5.6}$$

We will distinguish *small* and *large* solutions. By definition, the small ones will be those with

$$|\underline{\underline{x}}| \le H^*(F)^{6nr^n} = H(L)^{6nr^{n+1}} . \tag{5.7}$$

6. The small solutions

The forms $L^{(i_1)}, \dots , L^{(i_n)}$ in (4.1) do not necessarily have real coefficients. However, by following the procedure of [11, §2], i.e. by replacing each of these forms by their real of imaginary parts, we obtain n linearly independent forms $L_1, \dots , L_n$ with real coefficients such that every $\underline{\underline{x}}$ with (4.1) satisfies

$$|L_1(\underline{\underline{x}}) \cdots L_n(\underline{\underline{x}})| < \frac{8^n}{(n!)^{1/2}V(n)} |\det(L_1, \dots , L_n)| \mathfrak{H}(F)^{-1/r} . \tag{6.1}$$

It is easily seen that $(n!)^{1/2}V(n) > (2\pi)^{-n/2}$, so that

$$\frac{8^n}{(n!)^{1/2}V(n)} < (8\sqrt{2\pi})^n < H(L)^{1/2} = \mathfrak{H}(F)^{1/2r}$$

by (5.6), (5.5). Therefore

$$|L_1(\underline{\underline{x}}) \cdots L_n(\underline{\underline{x}})| < \mathfrak{H}(F)^{-1/2r} |\det(L_1, \dots , L_n)| . \tag{6.2}$$

We now quote Lemma 3.1 of [11]:

Lemma 4. *Let $L_1, \dots , L_n$ be linearly independent forms with real coefficients in n variables. Suppose that $(n!)^4 \le P \le B$, and put $Q = (\log B)/(\log P)$. The integer points $\underline{\underline{x}}$ in the ball $|\underline{\underline{x}}| \le B$ with*

$$|L_1(\underline{\underline{x}}) \cdots L_n(\underline{\underline{x}})| < P^{-1} |\det(L_1, \dots , L_n)|$$

lie in the union of not more than $n^{3n}Q^{n-1}$ proper subspaces.

We will apply this with $P = \mathfrak{H}(F)^{1/2r}$; in view of (5.5), (5.6), the condition $P \ge (n!)^4$ is satisfied. The small solutions are the ones with $|\underline{\underline{x}}| \le \mathfrak{H}(F)^{6nr^n} = B$, say. Integer points $\underline{\underline{x}}$ with $|\underline{\underline{x}}| \le B$ and with (6.2) lie in the union of at most $n^{3n}(12nr^{n+1})^{n-1}$ subspaces. Taking into account the $\binom{r}{n} \le r^n$ possible choices for $i_1, \dots , i_n$, we obtain

Lemma 5. *When* (5.1), (5.4) *hold, the small solutions of* (1.2) *lie in the union of less than $12^n n^{4n} r^{n^2+n}$ proper subspaces.*

7. Ranks of linear forms

The *rank* of a set of linear forms is the rank of their coefficient matrix. Since the coefficients of L are linearly independent over $\mathbf{Q}$, the rank of $L^{(1)}, \dots , L^{(r)}$ is n.

Recall that we denoted the compositum of the fields $K^{(1)}, \dots, K^{(r)}$ by N, and the Galois group of N over $\mathbf{Q}$ by G. As before, $\sigma_1, \dots, \sigma_r$ are the isomorphic embeddings of K into $\mathbf{C}$, so that $\sigma_i(\alpha) = \alpha^{(i)}$. Let Φ be the set $\{\sigma_1, \dots, \sigma_r\}$. Given a subset A of Φ, let $|A|$ be its cardinality and $\rho(A)$ the rank of the set of linear forms $\sigma(L)$ (where σ acts coefficientwise, so that $\sigma(L)$ is some $L^{(i)}$) with $\sigma \in A$. In particular, $\rho(\Phi) = n$. Put

$$q(A) = \begin{cases} \rho(A)/|A| & \text{when } A = \varnothing, \\ n/r & \text{when } A = \varnothing. \end{cases}$$

We will suppose that $N \subset \mathbf{C}$. Given $\gamma \in G$, the maps $\gamma\sigma_1, \dots, \gamma\sigma_r$ will be a permutation of Φ, so that G acts on Φ; in fact it acts transitively on Φ. We have $|\gamma(A)| = |A|$ and $\rho(\gamma(A)) = \rho(A)$ for $\gamma \in G$.

Lemma 6. *Let q_0 be the minimum value of $q(A)$ for all subsets A of Φ. Then $q_0 = n/r$.*

Moreover, there are integers, l, m with $lm = r$, a set $T \subset \Phi$ with $|T| = m$, and elements $\gamma_1, \dots, \gamma_l$ of G such that Φ is the disjoint union of the sets

$$\gamma_1(T), \dots, \gamma_l(T), \tag{7.1}$$

and a set A has $q(A) = q_0$ precisely if it is the union of some of these sets.

This is Lemma 7B of Chapter VII of [9].

In other words, a set $A \subset \Phi$ has

$$|A| \le (r/n)\rho(A), \tag{7.2}$$

with equality precisely when A is the union of some of the sets (7.1).

We may suppose that $T = \{\sigma_1, \dots, \sigma_m\}$. As was shown in §VII.7 of [9], there is a field[1] $L \subset K$ of degree l with the following properties. *The conjugates of an element λ of L over $\mathbf{Q}$ are $\gamma_1(\lambda), \dots, \gamma_l(\lambda)$. Thus the restrictions of $\gamma_1, \dots, \gamma_l$ to L, call them $\phi_1, \dots, \phi_l$, are the isomorphic embeddings of L into $\mathbf{C}$. Moreover, the conjugates of an element κ of K over $\mathbf{Q}$ are $\gamma_i\sigma_j(\kappa)$ $(1 \le i \le l,\ 1 \le j \le m)$, and $\gamma_i\sigma_j(\lambda) = \gamma_i(\lambda) = \phi_i(\lambda)$ for $\lambda \in L$. Thus $\phi_{ij} = \gamma_i\sigma_j$ $(j = 1, \dots, m)$ are embeddings of K into $\mathbf{C}$ extending ϕ_i, and $\Phi = \{\phi_{11}, \dots, \phi_{lm}\}$. We shall write*

$$\begin{aligned} \lambda^{|i|} &= \phi_i(\lambda) \qquad (i = 1, \dots, l) \textit{ for } \lambda \in L, \\ \kappa^{|i,j|} &= \phi_{ij}(\kappa) \qquad (i = 1, \dots, l;\ j = 1, \dots, m) \textit{ for } \kappa \in K. \end{aligned} \tag{7.3}$$

By Lemma 7D of §VII.7 of [9] we have: *A subfield P of K has $\mathfrak{M}^P = \mathfrak{M}$ if and only if $P \subset L$.* Now when $\mathfrak{M}$ is nondegenerate, this may happen only if $P = \mathbf{Q}$ or is imaginary quadratic. It follows that $L = \mathbf{Q}$ or L is imaginary quadratic.

[1] There should be no confusion of the field L and the linear form L.

It is clear from Lemma 6 that there is some $\eta > 0$ such that for every $A \subset \Phi$ which is not the union of some of the sets (7.1), i.e. some of the sets

$$\Phi_i = \{\phi_{i1}, \ldots, \phi_{im}\} \qquad (i = 1, \ldots, l),$$

we have

$$|A| \leq (r/n)\rho(A) - \eta. \tag{7.4}$$

Call η *admissible* if (7.4) holds for any A which is not such a union and which is maximal in the sense that there is no $B \supsetneqq A$ with $\rho(B) \leq \rho(A)$.

Lemma 7. (i) $\eta = 1/n$ *is admissible.*

(ii) $\eta = r(2n)^{-n2^{n+1}}$ *is admissible.*

(iii) $\eta = r/(n(n+1))$ *is admissible under the hypothesis of Theorem* 2.

The values in (i), (ii) will respectively lead to the first and second estimate implicit in (1.3).

Proof. (i) follows from the fact that $|A|$, $\rho(A)$ are integers. The more difficult proof of (ii) will be given in an Appendix. As for (iii), we note that the hypotheses of Theorem 2 yield $|A| = \rho(A)$ for every A with $\rho(A) < n$. Since $r \geq n+1$ and thus $(r-n)/n \geq r/(n(n+1))$, such an A has

$$|A| = \rho(A) = (r/n)\rho(A) - ((r-n)/n)\rho(A) \leq (r/n)\rho(A) - r/(n(n+1))$$

when $A \neq \emptyset$. On the other hand when $\rho(A) = n$ and A is maximal, then $A = \Phi$, and A is the union of the sets (7.1).

8. Products of n linear forms, again

We will deal with large solutions of (1.2). Such solutions by definition (5.7) have

$$|\underline{\underline{x}}| > H^*(F)^{6nr^n} = H(L)^{6nr^{n+1}}. \tag{8.1}$$

In view of (5.6), such $\underline{\underline{x}}$ will certainly have

$$|\underline{\underline{x}}| > n^3 H(L)^{3(n+1)r^n}. \tag{8.2}$$

Write $M^{(i)}(\underline{\underline{X}}) = |L^{(i)}|^{-1} L^{(i)}(\underline{\underline{X}})$, so that $M^{(i)}$ is the "normalization" of $L^{(i)}$. The forms $M^{(1)}, \ldots, M^{(r)}$ need not be conjugates.

Put

$$\delta = \eta n/3r \tag{8.3}$$

where η is admissible as defined in the last section.

Lemma 8. *Let* $\underline{\underline{x}}$ *be a large solution of* (1.2). *Then there are distinct integers* $i_1, \ldots, i_n$ *in* $1 \leq i \leq r$ *such that*

$$|M^{(i_1)}(\underline{\underline{x}}) \cdots M^{(i_n)}(\underline{\underline{x}})| < |\underline{\underline{x}}|^{-\delta}. \tag{8.4}$$

This lemma will play the same role for large solutions that Lemma 3 played for small solutions.

Proof. Pick i_1 such that $|M^{(i_1)}(\underline{\underline{x}})|$ is the minimum of $|M^{(i)}(\underline{\underline{x}})|$ for $1 \leq i \leq r$. Let I_1 be the set of numbers i such that $M^{(i)}(\underline{\underline{X}})$ is a multiple of $M^{(i_1)}(\underline{\underline{X}})$. Pick $i_2 \notin I_1$ such that $|M^{(i_2)}(\underline{\underline{x}})|$ is the minimum of $|M^{(i)}(\underline{\underline{x}})|$ for $i \notin I_1$. Let I_2 be the set of numbers $i \notin I_1$ such that $M^{(i)}(\underline{\underline{X}})$ is a linear combination of $M^{(i_1)}(\underline{\underline{X}})$, $M^{(i_2)}(\underline{\underline{X}})$. Continuing in this way we obtain numbers $i_1, \dots, i_n$ and disjoint sets $I_1, \dots, I_n$ whose union is $\{1, \dots, r\} =: \Omega$. We claim that (8.4) holds with these particular numbers $i_1, \dots, i_n$. We shall assume to the contrary that

$$|M^{(i_1)}(\underline{\underline{x}}) \cdots M^{(i_n)}(\underline{\underline{x}})| \geq |\underline{\underline{x}}|^{-\delta}, \tag{8.5}$$

and this will lead to a contradiction.

By definition of the forms $M^{(i)}$, every solution of (1.2) has

$$|a||L^{(1)}| \cdots |L^{(r)}||M^{(1)}(\underline{\underline{x}}) \cdots M^{(r)}(\underline{\underline{x}})| = 1,$$

and therefore

$$|M^{(1)}(\underline{\underline{x}}) \cdots M^{(r)}(\underline{\underline{x}})| = 1/H^*(F) \leq 1 \tag{8.6}$$

by (2.3), and since $H^*(F) \geq 1$, which follows from Lemma 1 (as well as from (5.5), (5.6)).

Suppose now that certain forms $M^{(j_1)}, \dots, M^{(j_n)}$ are linearly independent. Then each X_i is a linear combination: $X_i = \gamma_{i1} M^{(j_1)} + \cdots + \gamma_{in} M^{(j_n)}$, and here $|\gamma_{ik}| \leq H(L)^{nd}$ by Lemma 5.6 of [11], where d is the degree of the field obtained by adjoining to $\mathbf{Q}$ the coefficients of the forms $L^{(j_1)}, \dots, L^{(j_n)}$. Clearly $d \leq r^n$. Thus with the notation

$$|\overline{M(\underline{\underline{x}})}| = \max(|M^{(1)}(\underline{\underline{x}})|, \dots, |M^{(r)}(\underline{\underline{x}})|),$$

we have

$$|\underline{\underline{x}}| \leq nH(L)^{nr^n} |\overline{M(\underline{\underline{x}})}| < |\underline{\underline{x}}|^{1/3} |\overline{M(\underline{\underline{x}})}| \tag{8.7}$$

by (8.2).

We also note that if we have a relation of dependence: $M^{(u)} = c_1 M^{(u_1)} + \cdots + c_l M^{(u_l)}$ with $l \leq n$, then there is by Lemma 5.7 of [11] a relation of this type with $|c_i| \leq H(L)^{d(l+1)}$, hence with

$$|c_i| \leq H(L)^{(n+1)r^n} \qquad (i = 1, \dots, l). \tag{8.8}$$

Since $M^{(i)}$ with $i \in I_j$ is a linear combination of $M^{(i_1)}, \dots, M^{(i_j)}$, it is such a combination with coefficients satisfying (8.8), so that by (8.2),

$$|M^{(i_j)}(\underline{\underline{x}})| \leq |M^{(i)}(\underline{\underline{x}})| \leq nH(L)^{(n+1)r^n} |M^{(i_j)}(\underline{\underline{x}})| < |\underline{\underline{x}}|^{1/3} |M^{(i_j)}(\underline{\underline{x}})|. \tag{8.9}$$

We have

$$(8.10)\qquad \prod_{j=1}^{n}|M^{(i_j)}(\underline{\underline{x}})|^{|I_j|}\le\prod_{j=1}^{n}\prod_{i\in I_j}|M^{(i)}(\underline{\underline{x}})|=|M^{(1)}(\underline{\underline{x}})\cdots M^{(r)}(\underline{\underline{x}})|\le 1$$

by (8.6).

Define $c_1=c_1(\underline{\underline{x}}),\dots,c_n=c_n(\underline{\underline{x}})$ by

$$(8.11)\qquad |M^{(i_j)}(\underline{\underline{x}})|=|\underline{\underline{x}}|^{c_j}\qquad (j=1,\dots,n).$$

Then

$$(8.12)\qquad c_1\le\cdots\le c_n\le 1;$$

the last relation here is true since the forms $M^{(i)}$ are normalized. Furthermore,

$$(8.13)\qquad c_1+\cdots+c_n\ge-\delta$$

by (8.5) and

$$(8.14)\qquad c_1|I_1|+\cdots+c_n|I_n|\le 0$$

by (8.10). Since $|I_1|+\cdots+|I_n|=r$, the last inequality here may be rewritten as

$$(c_1+\cdots+c_n)(r/n)+(c_2-c_1)((r/n)-|I_1|)+(c_3-c_2)(2(r/n)-|I_1|-|I_2|)$$
$$+\cdots+(c_n-c_{n-1})((n-1)(r/n)-|I_1|-\cdots-|I_{n-1}|)\le 0,$$

so that in conjunction with (8.13),

$$(8.15)\qquad \sum_{k=1}^{n-1}(c_{k+1}-c_k)\left(k\left(\frac{r}{n}\right)-|I_1\cup\cdots\cup I_k|\right)\le\frac{r\delta}{n}.$$

The correspondence $i\leftrightarrow\sigma_i$ gives an identification of $\Omega=\{1,\dots,r\}$ and $\Phi=\{\sigma_1,\dots,\sigma_r\}$. Thus G acts on Ω, and the function ρ of §7 is defined on sets $A\subset\Omega$. With this notation, $\rho(I_1\cup\cdots\cup I_k)=k$ $(0\le k\le n)$. There are certain values of k such that $|I_1\cup\cdots\cup I_k|=(r/n)k$, and in fact $I_1\cup\cdots\cup I_k$ is the union of some of the sets $\gamma_1(T'),\dots,\gamma_l(T')$, where $T'\subset\Omega$ corresponds to $T\subset\Phi$. Suppose this to be the case for $k=k_1,\dots,k_g$ with $1\le k_1<\cdots<k_g=n$, and for no other values of k. When k is distinct from $k_1,\dots,k_g$, then

$$|I_1\cup\cdots\cup I_k|\le(r/n)k-\eta$$

by (7.4). Therefore (8.15) yields

$$\sum_{\substack{k=1\\k\ne k_1,\dots,k_g}}^{n-1}(c_{k+1}-c_k)\eta\le r\delta/n,$$

so that in particular (with the notation $k_0=0$)

$$c_{k_e}-c_{k_{e-1}+1}\le r\delta/\eta n=\tfrac{1}{3}\qquad (e=1,\dots,g).$$

Then by definition of the "exponents" c_j,

$$|M^{(i_h)}(\underline{\underline{x}})| \le |\underline{\underline{x}}|^{1/3}|M^{(i_j)}(\underline{\underline{x}})| \quad \text{if } k_{e-1}+1 \le h\,,\ j \le k_e\,.$$

Combining this with (8.9) we get

$$|M^{(u)}(\underline{\underline{x}})| \le |\underline{\underline{x}}|^{2/3}|M^{(v)}(\underline{\underline{x}})| \quad \text{if } u\,,v \in I_{k_{e-1}+1}\cup\cdots\cup I_{k_e} = V_e\,,$$

say.

We had seen that each set $I_1\cup\cdots\cup I_{k_e}$ is the union of some of the sets (7.1), and hence V_e $(1 \le e \le g)$ also is such a union. Therefore each set $\gamma_i(T')$ is contained in some V_e, so that with a notation in the spirit of (7.3),

$$(8.16) \qquad |M^{[i,h]}(\underline{\underline{x}})| \le |\underline{\underline{x}}|^{2/3}|M^{[i,j]}(\underline{\underline{x}})| \qquad (1 \le i \le l;\ 1 \le j\,,\ h \le m)\,.$$

Put

$$U^{[i]}(\underline{\underline{x}}) = M^{[i,1]}(\underline{\underline{X}})\cdots M^{[i,m]}(\underline{\underline{X}}) \qquad (1 \le i \le l)\,,$$
$$W^{[i]}(\underline{\underline{X}}) = L^{[i,1]}(\underline{\underline{X}})\cdots L^{[i,m]}(\underline{\underline{X}}) \qquad (1 \le i \le l)\,.$$

Then $W^{[i]}$ is a form of degree m with coefficients in $L^{(i)}$, and $W^{[1]}, \dots, W^{[l]}$ are "conjugates".

Now in Theorem 1, where the module $\mathfrak{M}$ is nondegenerate, we can only have $L = \mathbf{Q}$ or L is imaginary quadratic.[2] In the first case, $l = 1$, and in the second case, $l = 2$ and $W^{[1]}$, $W^{[2]}$ are complex conjugates (i.e., their respective coefficients are complex conjugates). For $l = 2$ it follows that $|L^{[1,1]}|\cdots|L^{[1,m]}| = |L^{[2,1]}|\cdots|L^{[2,m]}|$, and also $U^{[1]}$, $U^{[2]}$, which are proportional to $W^{[1]}$, $W^{[2]}$, are complex conjugates. Thus $|U^{[1]}(\underline{\underline{x}})| = |U^{[2]}(\underline{\underline{x}})|$.

Therefore both in the case $l = 1$ and the case $l = 2$, we have by (8.16), (8.6) that

$$\begin{aligned}|M^{[i,h]}(\underline{\underline{x}})| &\le |\underline{\underline{x}}|^{2/3}|U^{[i]}(\underline{\underline{x}})|^{1/m} = |\underline{\underline{x}}|^{2/3}|U^{[1]}(\underline{\underline{x}})\cdots U^{[l]}(\underline{\underline{x}})|^{1/r}\\ &= |\underline{\underline{x}}|^{2/3}|M^{(1)}(\underline{\underline{x}})\cdots M^{(r)}(\underline{\underline{x}})|^{1/r} \le |\underline{\underline{x}}|^{2/3}\,.\end{aligned}$$

Thus $|\overline{M(\underline{\underline{x}})}| \le |\underline{\underline{x}}|^{2/3}$. Together with (8.7) this gives $|\underline{\underline{x}}| < |\underline{\underline{x}}|$, and the desired contradiction to (8.5).

9. The large solutions

Set

$$(9.1) \qquad \eta = \max(1/n\,, r\cdot(2n)^{-n\cdot 2^{n+1}})\,,$$

so that η is admissible by Lemma 7. Define δ by (8.3). Let us first consider large solutions satisfying (8.4) with fixed $i_1, \dots, i_n$. Now

$$|\det(M^{(i_1)}, \dots, M^{(i_n)})| \ge H(L)^{-nd} \ge H(L)^{-nr^n}$$

[2] For Theorem 2 see §11.

by (5.3) of [11], so that

$$(9.2)\qquad \begin{aligned} |M^{(i_1)}(\underline{\underline{x}})\cdots M^{(i_n)}(\underline{\underline{x}})| &< |\det(M^{(i_1)},\dots,M^{(i_n)})|H(L)^{rn^n}|\underline{\underline{x}}|^{-\delta} \\ &< |\det(M^{(i_1)},\dots,M^{(i_n)})||\underline{\underline{x}}|^{-\delta/2} \end{aligned}$$

by (8.4), (8.1), (8.3), (9.1).

By the quantitative version of the Subspace Theorem [11], applied with $\delta/2$, the solutions $\underline{\underline{x}} \neq \underline{\underline{0}}$ of (9.2) either have

$$(9.3)\qquad |\underline{\underline{x}}| \le \max((n!)^{8/\delta}, H(L)),$$

or they lie in the union of t proper subspaces, where

$$t = [(2d)^{2^{26n}\cdot 4\delta^{-2}}] \le [(2r^n)^{2^{26n}\cdot 4\delta^{-2}}].$$

Now $(n!)^{8/\delta} \le n^{8n/\delta} \le n^{8n\cdot 3r}$ by (8.3), (9.1), and $n^{24nr} < H(L)^{24nr} < H(L)^{6nr^{n+1}}$ by (5.6). Hence there are no large solutions with (9.3).

Taking into account the possible choices for $i_1, \dots, i_n$, we see that the large solutions lie in at most $r^n t$ proper subspaces.

Lemma 9. *The large solutions of* (1.2) *lie in at most* $[(2r^n)^{2^{26n}\cdot 5\delta^{-2}}]$ *subspaces.*

10. Proof of Proposition 1 and Theorem 1

Suppose that condition (5.4) of the proposition is satisfied and that (5.1) holds. Combining Lemma 5 on small solutions and Lemma 9 on large solutions, we see that all the solutions lie in the union of not more than

$$12^n n^{4n} r^{n^2+n} + (2r^n)^{2^{26n}\cdot 5\delta^{-2}} < 2\cdot(2r^n)^{2^{26n}\cdot 5\delta^{-2}} < r^{2^{27n}\cdot 5\delta^{-2}}$$

subspaces, where δ is given by (8.3), (9.1). Thus

$$(10.1)\qquad \delta^{-2} = \frac{9r^2}{n^2}\min\left(n^2, \frac{1}{r^2}(2n)^{n\cdot 2^{n+2}}\right) \le 9\min(r^2, (2n)^{n\cdot 2^{n+2}}) = 9\lambda_n$$

with $\lambda_n = \min(r^2, (2n)^{n\cdot 2^{n+2}})$, and the solutions lie in not more than $r^{2^{27n}\cdot 45\lambda_n}$ subspaces.

The integer points in a proper rational subspace S may be parametrized as $\underline{\underline{x}} = T\underline{\underline{y}}$ where T is a linear map $\mathbf{R}^{n-1} \to S$ which sets up a 1-1 correspondence between $\mathbf{Z}^{n-1}$ and integer points on S. Restricted to S, the norm form equation (1.2) becomes

$$F^T(\underline{\underline{y}}) = F(T\underline{\underline{y}}) = 1$$

with $\underline{\underline{y}} \in \mathbf{Z}^{n-1}$. This is a norm form equation in $n-1$ variables. As $\underline{\underline{y}}$ runs through Z^{n-1}, then $L^T(\underline{\underline{y}})$ will run through a submodule of $\mathfrak{M}$, and such a submodule will again be nondegenerate.

We now proceed by induction on n. The case $n = 1$ is trivial. Suppose that $n > 1$ and Theorem 1 is true for $n-1$; this means exactly that a nondegenerate norm form equation in $n-1$ variables has not more than

$$c_1(n-1, r) = \min(r^{2^{30(n-1)}r^2}, r^{c_2(n-1)})$$

solutions. This is then a bound for the number of solutions in each of the proper subspaces S, and the given equation (1.2) in n variables has

$$\leq c_1(n-1, r) r^{2^{27n} \cdot 45\lambda_n} \leq r^{\min(\mu_1, \mu_2)}$$

solutions, where

$$\mu_1 = 2^{30(n-1)}r^2 + 2^{27n} \cdot 45r^2 < 2^{30n}r^2 - 10n^2$$

(since $n \geq 2$) and

$$\mu_2 = c_2(n-1) + 2^{27n} \cdot 45(2n)^{n \cdot 2^{n+2}} = (2n-2)^{(n-1) \cdot 2^{n+3}} + 2^{27n} \cdot 45(2n)^{n \cdot 2^{n+2}}$$
$$< (2n)^{n \cdot 2^{n+4}} - 10n^2 = c_2(n) - 10n^2.$$

This establishes Proposition 1, and hence Theorem 1, for the case of n variables.

11. Proof of Theorem 2

Let L_0 be the restriction of L to a rational subspace S of dimension m. Since $L^{(1)}, \ldots, L^{(r)}$ have rank n, the forms $L_0^{(1)}, \ldots, L_0^{(r)}$ have rank m. Now when $m \leq n-1$, we may infer from m-times transitivity that any m forms among $L_0^{(r)}, \ldots, L_0^{(r)}$ are of rank m. Thus the hypothesis of Theorem 2 remain true for the restrictions to subspaces.

We modify the proof of Theorem 1 as follows. Instead of (9.1) we now set

$$\eta = r/(n(n+1)). \tag{11.1}$$

Then $\eta \geq 1/n$. In §8 we either may proceed as before, by noting that our present hypotheses imply nondegeneracy.[3] Or, a simpler argument rests on the fact that in Lemma 7(iii) we have actually shown (7.4) to hold with (11.1) for every $A \neq \varnothing$, Φ. Now Lemmas 8, 9 apply. (10.1) is replaced by

$$\delta^{-2} = (9r^2/n^2)\min(n^2, r^{-2}n^2(n+1)^2) = 9(n+1)^2.$$

The solutions therefore lie in not more than $r^{2^{27n} \cdot 45(n+1)^2}$ subspaces. Induction on n may be carried out since the restrictions of our linear forms to subspaces satisfy the hypotheses. We obtain a modified version of Proposition 1, with the conclusion that

$$N_2(r, n; p) < r^{2^{32(n-1)}} \cdot r^{2^{27n} \cdot 45(n+1)^2} < r^{2^{32n} - 10n^2};$$

[3] For when $\mathfrak{M}^E \neq 0$, then $\mathfrak{M}$ contains a submodule with basis $\mu\varepsilon_1, \ldots, \mu\varepsilon_m$ where $\mu \neq 0$ and $\varepsilon_1, \ldots, \varepsilon_m$ is a basis of $E/\mathbf{Q}$. By what we said at the beginning, any m conjugates of this basis are linearly independent, so that for $m \geq 2$, any 2 of the r conjugates are nonproportional, and $m \geq r > n$, a contradiction. Thus $m = 1$, and $\mathfrak{M}$ is nondegenerate.

here $N_2(r,n;p)$ is the maximum number of solutions of (1.2) for forms F such as in Theorem 2 with (5.1). Theorem 2 follows.

12. Proof of Theorem 3

The arguments of this section work not only for norm forms, but more generally for *decomposable* forms. A form $F(\underline{X})$ of degree r is called decomposable if $F = L_1 \cdots L_r$, where $L_1, \dots, L_r$ are linear forms with coefficients which are algebraic over $\mathbf{Q}$. We will only consider decomposable forms F with coefficients in $\mathbf{Z}$.

Let C be a class of decomposable forms of degree r in n variables and with coefficients in $\mathbf{Z}$ which has the following closure property. When $F \in C$, and when $R(\underline{X}) = cF^T(\underline{X})$ with $T \in GL(n,\mathbf{Z})$ and $c \neq 0$ in $\mathbf{Q}$ has coefficients in $\mathbf{Z}$, then also $R \in C$. For example the class of nondegenerate norm forms with integer coefficients has this property. The class of forms considered in Theorem 2 also has this property. Let $N(C)$ be the maximum number (if there is such a number) of integer solutions of (1.2) for all forms $F \in C$. More generally, for $m \neq 0$ let $N(C,m)$ be the maximum number of primitive solutions of (1.1) for all forms $F \in C$. We have $N(C,-m) = N(C,m)$.

Theorem 3 is an immediate consequence of Theorems 1, 2 and of

Proposition 2. *For $m > 0$ we have*

$$N(C,m) \leq \binom{r}{n-1}^{\omega} d_{n-1}(m^r)N(C);$$

here $\omega = \omega(m)$ and d_{n-1} are defined as in the Introduction.

Given n, r, the quantity $g(m) = \binom{r}{n-1}^{\omega} d_{n-1}(m^r)$ is multiplicative in m: $g(m_1 m_2) = g(m_1)g(m_2)$ when g.c.d.$(m_1,m_2) = 1$. Therefore Proposition 2 follows from

Proposition 2a. *For $k > 0$ and a prime power p^u with $p \nmid k$ we have*

$$N(C,kp^u) \leq \binom{r}{n-1} d_{n-1}(p^{ru})N(C,k).$$

Let $|\cdot|$ denote a nonarchimedean absolute value on a field E. For $\underline{x} = (x_1, \dots, x_n)$ in E^n put $|\underline{x}| = \max(|x_1|, \dots, |x_n|)$, and for a polynomial $P(X_1, \dots, X_n)$ with coefficients in E, let $|P|$ be the maximum of $|c|$ over the coefficients c of P.

Lemma 10. *Suppose E is algebraically closed. Then for $P \in E[X_1, \dots, X_n]$ we have*

$$\max_{\substack{\underline{x}\in E^n \\ |\underline{x}|\leq 1}} |P(\underline{x})| = |P|. \tag{12.1}$$

Proof. Let us begin with the case $n = 1$, so that

$$P(X) = c_0 + c_1 X + \cdots + c_r X^r.$$

It is clear that the left side of (12.1) is bounded by the rignt-hand side. To prove the equality it will suffice to find $x \in E$ with $|x| = 1$ and $|P(x)| = |P|$. We may suppose that $P \neq 0$. Write $P(X) = P_1(X) + P_2(X)$, where every coefficient of P_1 has absolute value $|P|$, and every coefficient of P_2 has absolute value $< |P|$. Say

$$P_1(X) = c_{i_1}X^{i_1} + \cdots + c_{i_t}X^{i_t}$$

with $i_1 < \cdots < i_t$. Pick $x \in E$ with $P_1(x) = c_{i_t}x^{i_t+1}$. Then necessarily $|x| = 1$ and $|P_1(x)| = |c_{i_t}| = |P|$, $|P_2(x)| < |P|$, so that $|P(x)| = |P|$.

The general case of the lemma follows by induction on n.

Lemma 11. *Let $L_1(\underline{\underline{X}}), \dots, L_m(\underline{\underline{X}})$ be linearly dependent forms with coefficients in E. Given nonnegative reals $\lambda_1, \dots, \lambda_m$, let $\mathfrak{R}$ be the set of $\underline{\underline{x}}$ with components in E having*

$$(12.2) \qquad |L_i(\underline{\underline{x}})| \leq \lambda_i \qquad (i = 1, \dots, m).$$

Then $\mathfrak{R}$ may already be defined by $m - 1$ of these inequalities, i.e. there is an i_0 in $1 \leq i_0 \leq m$ such that $\mathfrak{R}$ is the set of points $\underline{\underline{x}}$ satisfying (12.2) *for $i = 1, \dots, i_0 - 1, i_0 + 1, \dots, m$.*

Proof. There is a nontrivial relation $\gamma_1 L_1 + \cdots + \gamma_m L_m = 0$ with coefficients $\gamma_i \in E$. We may suppose without loss of generality that $\gamma_1 \neq 0, \dots, \gamma_l \neq 0$, but $\gamma_{l+1} = \cdots = \gamma_m = 0$. We further may suppose that

$$|\gamma_1|\lambda_1 = \max(|\gamma_1|\lambda_1, \dots, |\gamma_l|\lambda_l).$$

In this case we set $i_0 = 1$. When (12.2) holds for $i = 2, \dots, m$, then

$$|\gamma_1 L_1(\underline{\underline{x}})| \leq \max_{2 \leq i \leq l} |\gamma_i L_i(\underline{\underline{x}})| \leq \max_{2 \leq i \leq l} |\gamma_i|\lambda_i \leq |\gamma_1|\lambda_1,$$

so that $|L_1(\underline{\underline{x}})| \leq \lambda_1$ and (12.2) holds for $i = 1$ also.

Lemma 12. *Let $L_1(\underline{\underline{X}}), \dots, L_n(\underline{\underline{X}})$ be linear forms in n variables and with coefficients in E. Suppose there exists an $\underline{\underline{x}}' \in E^n$ with*

$$|\underline{\underline{x}}'| = 1 \quad \text{and} \quad |L_i(\underline{\underline{x}}')| \leq \lambda_i \qquad (i = 1, \dots, n),$$

where $\lambda_1, \dots, \lambda_n$ are given. Then there is an i_0 in $1 \leq i_0 \leq n$ such that the relations

$$(12.3) \qquad |\underline{\underline{x}}| \leq 1 \quad \text{and} \quad |L_i(\underline{\underline{x}})| \leq \lambda_i \qquad (i = 1, \dots, i_0 - 1, i_0 + 1, \dots, n)$$

imply that

$$(12.4) \qquad |L_i(\underline{\underline{x}})| \leq \lambda_i \qquad (i = 1, \dots, n).$$

Proof. In view of Lemma 11 we may suppose that $L_1, \dots, L_n$ are linearly independent. There are relations

$$(12.5) \qquad X_i = \gamma_{i1}L_1(\underline{\underline{X}}) + \cdots + \gamma_{in}L_n(\underline{\underline{X}}) \quad (i = 1, \dots, n).$$

We may suppose without loss of generality that

$$(12.6) \qquad |\gamma_{11}|\lambda_1 = \max_{1 \leq i,j \leq n} |\gamma_{ij}|\lambda_j.$$

The given $\underline{\underline{x}}' = (x_1', \dots, x_n')$ has

$$|x_i'| \leq \max_j |\gamma_{ij}|\lambda_j \leq |\gamma_{11}|\lambda_1 .$$

Since this holds for $i = 1, \dots, n$ and since $|\underline{\underline{x}}'| = 1$, we have $1 \leq |\gamma_{11}|\lambda_1$. In particular, $\gamma_{11} \neq 0$.

We set $i_0 = 1$. Then (12.5) with $i = 1$ and (12.3) yield

$$|\gamma_{11}||L_1(\underline{\underline{x}})| \leq \max(|\gamma_{12}|\lambda_2, \dots, |\gamma_{1n}|\lambda_n, |x_1|) \leq |\gamma_{11}|\lambda_1 ,$$

in view of (12.6) and $|x_1| \leq 1 \leq |\gamma_{11}|\lambda_1$. Thus $|L_1(\underline{\underline{x}})| \leq \lambda_1$ and (12.4) is true.

Lemma 13. *Let $L_1, \dots, L_r$ with $r \geq n$ be linear forms in n variables and with coefficients in E. Let nonnegative reals $\lambda_1, \dots, \lambda_r$ be given, and suppose there is an $\underline{\underline{x}}' \in E^n$ with*

$$(12.7) \qquad |\underline{\underline{x}}'| = 1 \quad \text{and} \quad |L_i(\underline{\underline{x}}')| \leq \lambda_i \qquad (i = 1, \dots, r).$$

Then there are $n - 1$ among these forms, say $L_{i_1}, \dots, L_{i_{n-1}}$, such that any $\underline{\underline{x}} \in E^n$ with

$$(12.8) \qquad |\underline{\underline{x}}| \leq 1 \quad \text{and} \quad |L_{i_j}(\underline{\underline{x}})| \leq \lambda_{i_j} \qquad (j = 1, \dots, n-1)$$

has $|L_i(\underline{\underline{x}})| \leq \lambda_i$ $(i = 1, \dots, r)$.

Proof. This follows from $r - n$ applications of Lemma 11, followed by an application of Lemma 12.

Now let F be a form in our class C. We wish to estimate the number of primitive solutions of

$$(12.9) \qquad F(\underline{\underline{x}}) = kp^u$$

where k, p^u are as in Proposition 2a. Let E be the algebraic closure of $\mathbf{Q}$ and let $|\cdot|$ be an extension of the p-adic absolute value to E. We may write F as $F = L_1 \cdots L_r$, where L_i is a linear form with coefficients in a field $K_i \subset E$ of degree $\leq r$. With every primitive solution $\underline{\underline{x}}$ of (12.9) we associate numbers $\lambda_1, \dots, \lambda_r$ by setting

$$(12.10) \qquad |L_i(\underline{\underline{x}})| = \lambda_i \qquad (i = 1, \dots, r).$$

Since there is by construction a primitive $\underline{\underline{x}}$ with (12.10), we may apply Lemma 13 and we get forms $L_{i_1}, \dots, L_{i_{n-1}}$. We will call $(L_{i_1}, \dots, L_{i_{n-1}}; \lambda_{i_1}, \dots, \lambda_{i_{n-1}})$ an *anchor* of the solution.

Let us first count the solutions with a given anchor. Without loss of generality let us suppose that the anchor is $(L_1, \dots, L_{n-1}, \lambda_1, \dots, \lambda_{n-1})$. Solutions with this anchor will have

$$(12.11) \qquad |L_i(\underline{\underline{x}})| \leq \lambda_i \qquad (i = 1, \dots, n-1).$$

Points $\underline{\underline{x}} \in \mathbf{Z}^n$ with (12.11) make up a sublattice of $\mathbf{Z}^n$. Let $\underline{\underline{a}}_1, \dots, \underline{\underline{a}}_n$ be a basis of this lattice. The elements of this lattice are $\underline{\underline{x}} = y_1\underline{\underline{a}}_1 + \cdots + y_n\underline{\underline{a}}_n$

with $\underline{\underline{y}} = (y_1, \dots, y_n) \in \mathbf{Z}^n$. Thus writing $M_i(\underline{\underline{Y}}) = L_i(Y_1\underline{\underline{a}}_1 + \cdots + Y_n\underline{\underline{a}}_n)$ $(i = 1, \dots, r)$ and $G(\underline{\underline{Y}}) = M_1(\underline{\underline{Y}}) \cdots M_r(\underline{\underline{Y}})$, it will suffice to count the primitive solutions $\underline{\underline{y}}$ of the equation

$$G(\underline{\underline{y}}) = kp^u. \tag{12.12}$$

By the construction of anchors, and if our anchor $(L_1, \dots, L_{n-1}, \lambda_1, \dots, \lambda_{n-1})$ comes from an r-tuple $\lambda_1, \dots, \lambda_{n-1}, \dots, \lambda_r$, then every $\underline{\underline{x}} \in E^n$ with $|\underline{\underline{x}}| \le 1$ and (12.11) will in fact have $|L_i(\underline{\underline{x}})| \le \lambda_i$ for $i = 1, \dots, r$. When $\underline{\underline{x}} = y_1\underline{\underline{a}}_1 + \cdots + y_n\underline{\underline{a}}_n$ with $\underline{\underline{y}} = (y_1, \dots, y_n) \in E^n$ having $|\underline{\underline{y}}| \le 1$, then we do in fact have $|\underline{\underline{x}}| \le 1$ and (12.11). Therefore $|M_i(\underline{\underline{y}})| \le \lambda_i$ $(i = 1, \dots, r)$ for every $\underline{\underline{y}} \in E^n$ with $|\underline{\underline{y}}| \le 1$. Now $\lambda_1, \dots, \lambda_r$ come from a solution $\underline{\underline{x}}$ of (12.8), so that

$$\lambda_1 \cdots \lambda_r = |L_1(\underline{\underline{x}})| \cdots |L_r(\underline{\underline{x}})| = |kp^u| = p^{-u}. \tag{12.13}$$

Thus also $|G(\underline{\underline{y}})| = |M_1(\underline{\underline{y}})| \cdots |M_r(\underline{\underline{y}})| \le \lambda_1 \cdots \lambda_r = p^{-u}$; this holds for every $\underline{\underline{y}}$ with $|\underline{\underline{y}}| \le 1$. Lemma 10 yields $|G| \le p^{-u}$. Therefore every coefficient of G is divisible by p^u, and $G = p^u R$ where R has integer coefficients. Here $R(\underline{\underline{Y}}) = p^{-u}G(\underline{\underline{Y}}) = p^{-u}F(Y_1\underline{\underline{a}}_1 + \cdots + Y_n\underline{\underline{a}}_n)$ lies in C, and therefore the number of primitive solutions of (12.12), which is the same as the number of primitive solutions of $R(\underline{\underline{y}}) = k$, is bounded by $N(C, k)$. Therefore: *the number of solutions with a given anchor does not exceed* $N(C, k)$.

It remains for us to estimate the number of possible anchors. The number of possibilities for $i_1 < \cdots < i_{n-1}$ is $\binom{r}{n-1}$. It will suffice to show that for given $i_1, \dots, i_{n-1}$, the number of possibilities for $\lambda_{i_1}, \dots, \lambda_{i_{n-1}}$ is $\le d_{n-1}(p^{ru})$. So let us estimate the number of anchors $(L_1, \dots, L_{n-1}, \lambda_1, \dots, \lambda_{n-1})$. We have to estimate the number of possibilities for $\lambda_1, \dots, \lambda_{n-1}$. We begin with three observations.

(i) Write $|L_i|$ for the maximum $|\cdot|$-value of the coefficients of L_i. By Gauss's Lemma,

$$|L_1| \cdots |L_r| = |F| \le 1.$$

On the other hand when $\lambda_i = |L_i(\underline{\underline{x}})|$ $(i = 1, \dots, r)$ where $\underline{\underline{x}}$ is a solution of (12.9), then (12.13) holds. It follows that

$$\lambda_1 \cdots \lambda_r \ge p^{-u}|L_1| \cdots |L_r|.$$

Note that $\lambda_i \le |L_i|$ $(i = 1, \dots, r)$. Therefore the quantities $\mu_i = \lambda_i/|L_i|$ have $\mu_i \le 1$ $(i = 1, \dots, r)$ and $\mu_1 \cdots \mu_r \ge p^{-u}$. We may conclude that

$$\mu_i \le 1 \ (i = 1, \dots, n-1) \quad \text{and} \quad \mu_1 \cdots \mu_{n-1} \ge p^{-u}. \tag{12.14}$$

(ii) When both $\lambda_1, \dots, \lambda_{n-1}$ and $\lambda'_1, \dots, \lambda'_{n-1}$ occur in anchors, and when $\lambda_i \le \lambda'_i$ $(i = 1, \dots, n-1)$, then in fact $\lambda_i = \lambda'_i$ $(i = 1, \dots, n-1)$: for suppose that $\lambda_1, \dots, \lambda_{n-1}$ comes from $\lambda_1, \dots, \lambda_{n-1}, \dots, \lambda_r$, and $\lambda'_1, \dots, \lambda'_{n-1}$ from $\lambda'_1, \dots, \lambda'_{n-1}, \dots, \lambda'_r$. Say $\underline{\underline{x}}$, $\underline{\underline{x}}'$ are solutions with $|L_i(\underline{\underline{x}})| = \lambda_i$, $|L_i(\underline{\underline{x}}')| = \lambda'_i$

$(i = 1, \dots, r)$. Then $|L_i(\underline{\underline{x}})| \leq \lambda_i'$ $(i = 1, \dots, n-1)$, and therefore $|L_i(\underline{\underline{x}})| \leq \lambda_i'$ $(i = 1, \dots, r)$, since $(L_1, \dots, L_{n-1}, \lambda_1', \dots, \lambda_{n-1}')$ is an anchor. Thus $\lambda_i \leq \lambda_i'$ $(i = 1, \dots, r)$, and since $\lambda_1 \cdots \lambda_r = \lambda_1' \cdots \lambda_r' = p^{-u}$, the assertion follows.

(iii) The restriction of $|\cdot|$ to the field K_i containing the coefficients of L_i has ramification index $e_i \leq r$ over the p-adic absolute value of $\mathbf{Q}$. Thus for $\alpha \neq 0$ in K_i we have $|\alpha| = p^{v/e_i}$ with $v \in \mathbf{Z}$.

In particular $|L_i(\underline{\underline{x}})|$ and $|L_i|$ are of this type, so that

$$\mu_i = |L_i(\underline{\underline{x}})|/|L_i| = p^{-v_i/e_i} \qquad (i = 1, \dots, n-1)$$

with $v_i \in \mathbf{Z}$. By (12.14),

$$v_i \geq 0 \ (i = 1, \dots, n-1) \quad \text{and} \quad \frac{v_1}{e_1} + \cdots + \frac{v_{n-1}}{e_{n-1}} \leq u.$$

By (ii), when both $v_1, \dots, v_{n-1}$ and $v_1', \dots, v_{n-1}'$ come from anchors and when $v_i' \leq v_i$ $(i = 1, \dots, n-1)$, then in fact $v_i' = v_i$ $(i = 1, \dots, n-1)$. This shows that for given $v_1, \dots, v_{n-2}$, there is at most one possibility for v_{n-1}. It will be enough for us to estimate the number of integer $(n-2)$-tuples $v_1, \dots, v_{n-2}$ with

$$v_i \geq 0 \ (i = 1, \dots, n-2) \quad \text{and} \quad \frac{v_1}{e_1} + \cdots + \frac{v_{n-2}}{e_{n-2}} \leq u.$$

Since $e_i \leq r$, the number of such tuples is bounded by the number of $(n-2)$-tuples of nonnegative integers $v_1, \dots, v_{n-2}$ with $v_1 + \cdots + v_{n-2} \leq ur$. This in turn is the number of $(n-1)$-tuples of nonnegative integers $v_1, \dots, v_{n-1}$ with $v_1 + \cdots + v_{n-1} = ur$, and is equal to $d_{n-1}(p^{ur})$.

Appendix. Sets of conjugate vectors

(a) Let K be an algebraic number field of degree r, and let $\sigma_1, \dots, \sigma_r$ be the isomorphic embeddings of K into $\mathbf{C}$. Let Φ be the set of these embeddings: $\Phi = \{\sigma_1, \dots, \sigma_r\}$. Further let $\underline{\underline{\alpha}} = (\alpha_1, \dots, \alpha_n)$ be a vector with components in K, and for $\sigma \in \Phi$ put $\sigma(\underline{\underline{\alpha}}) = (\sigma(\alpha_1), \dots, \sigma(\alpha_n))$. Given a subset B of Φ, let $|B|$ be its cardinality and $\rho(B)$ the dimension of the subspace of $\mathbf{C}^n$ spanned by the vectors $\sigma(\underline{\underline{\alpha}})$ with $\sigma \in B$. We shall assume throughout that $\alpha_1, \dots, \alpha_n$ are linearly independent over $\mathbf{Q}$; then $\sigma_1(\underline{\underline{\alpha}}), \dots, \sigma_r(\underline{\underline{\alpha}})$ span $\mathbf{C}^n$ and $\rho(\Phi) = n$.

Let N be a normal extension of $\mathbf{Q}$ containing the conjugate fields $\sigma_1(K)$, $\dots, \sigma_r(K)$. Let G be the Galois group of N over $\mathbf{Q}$. For $\gamma \in G$, the set $\gamma\sigma_1, \dots, \gamma\sigma_r$ is a permutation of $\sigma_1, \dots, \sigma_r$; thus G acts on Φ, and in fact it acts transitively on Φ. For a subset $B \subset \Phi$ we have $|\gamma(B)| = |B|$ and $\rho(\gamma(B)) = \rho(B)$.

Now let L be a subfield of K of degree l. Then $r = lm$ with integral m. There are l embeddings $\phi_1, \dots, \phi_l$ of L into $\mathbf{C}$. Each such embedding ϕ_i

can be extended to embeddings $\phi_{i1}, \dots, \phi_{im}$ of K into $\mathbf{C}$. Thus Φ consists of the embeddings ϕ_{ij} with $1 \le i \le l$, $1 \le j \le m$. Let Φ_i be the set

$$\Phi_i = \{\phi_{i1}, \dots, \phi_{im}\} \qquad (i = 1, \dots, l). \tag{A.1}$$

Then Φ is the disjoint union of $\Phi_1, \dots, \Phi_l$.

By applying the ideas of §7 to the linear form $L(\underline{X}) = \underline{\underline{\alpha}}\underline{X} = \alpha_1 X_1 + \dots + \alpha_n X_n$ we may restate and summarize results from §7 as follows.

Lemma 14. *Every subset* $B \subset \Phi$ *has*

$$|B| \le (r/n)\rho(B). \tag{A.2}$$

There is a subfield L *of* K *of some degree* l *such that equality holds in* (A.2) *if and only if* B *is the union of some of the sets* (A.1).

A set B will be called a *k-set* if $\rho(B) \le k$. It will be called a *maximal* k-set if there is no set $C \supsetneqq B$ with $\rho(C) \le k$.

Lemma 15. *Suppose* S *is a maximal k-set and is not the union of some of the sets* (A.1), *i.e.* $|S| < (r/n)k$. *Then*

$$|S| \le (k/n - (4kn)^{-k\cdot 2^k})r. \tag{A.3}$$

In particular, since $(4kn)^{k\cdot 2^k} \le (4n^2)^{n2^n} = (2n)^{n\cdot 2^{n+1}}$, this proves part (iii) of Lemma 7. It is likely that Lemma 15 is far from the full truth and that the exponent $k \cdot 2^k$ in (A.3) can be replaced by a linear function in k. The same remark applies to Proposition 3 below.

(b) The task of this Appendix will be a proof of Lemma 15. It will be advantageous to axiomatize. We will deal with a set Φ of cardinality $|\Phi| = r$ and an integer-valued function $\rho(B)$ defined on the subsets B of Φ with the following properties:

(i) $\rho(\varnothing) = 0$, $\rho(B) = 1$ when $|B| = 1$.

(ii) ρ is nondecreasing, i.e. $A \subset B$ implies $\rho(A) \le \rho(B)$.

(iii) If $A_1 \subset A_2$ and $\rho(A_1) = \rho(A_2)$, then $\rho(A_1 \cup B) = \rho(A_2 \cup B)$ for every B.

(iv) There is a group G which acts transitively on Φ, and such that $\rho(\gamma(B)) = \rho(B)$ for $\gamma \in G$.

In this context we again define a k-set as a set B with $\rho(B) \le k$, and a maximal k-set as a k-set which is not properly contained in another k-set.

Proposition 3. *Suppose* $k \ge 1$ *and* S^k *is a maximal k-set with* $|S^k| < ur/v$ *where* u/v *is rational. Then in fact*

$$|S^k| \le (u/v - (4kv)^{-k\cdot 2^k})r. \tag{A.4}$$

If we apply this with $u/v = k/n$ we get Lemma 15.

(c) A set $S \subset \Phi$ will be called *stable* if $\gamma(S) = S$ or $\gamma(S) \cap S = \varnothing$ for every $\gamma \in G$. If S is a nonempty stable set, then there are $\gamma_1, \dots, \gamma_t$ in G such that

Φ is the disjoint union of $\gamma_1(S), \dots, \gamma_t(S)$. A set of cardinality 1 is necessarily stable. A set is *unstable* if it is not stable, i.e. if there is a $\gamma \in G$ with $\gamma(S) \neq S$, $\gamma(S) \cap S \neq \varnothing$.

Lemma 16. *Suppose that* $q \geq 1$ *and* S *is a maximal* q*-set. Suppose that* $\rho(S) = q$ *and that* S *is stable. Say* Φ *is the disjoint union of* $S_1, \dots, S_t$ *where* $S_i = \gamma_i(S)$ *with* $\gamma_i \in G$ $(i = 1, \dots, t)$.

Let Ψ *be the set* $\{1, \dots, t\}$ *and given a subset* $C \subseteq \Psi$, *say* $C = \{i_1, \dots, i_p\}$, *put* $C^* = S_{i_1} \cup \dots \cup S_{i_p}$. *Define a function* σ *on the subsets* C *of* Ψ *by*

$$\sigma(C) = \begin{cases} \rho(C^*) - q + 1 & \text{when } C \neq \varnothing, \\ 0 & \text{when } C = \varnothing. \end{cases}$$

Then Ψ, σ *satisfy conditions* (i)–(iv).

We look upon Ψ as the "factor set" of Φ over S.

Proof. (i) When $|C| = 1$, then C^* consists of a single set S_i, so that $\sigma(C) = \rho(S_i) - q + 1 = q - q + 1 = 1$.

(ii) We have to show that $C_1 \subset C_2$ implies $\sigma(C_1) \leq \sigma(C_2)$. This is obvious when $C_1 = \varnothing$. When $C_1 \neq \varnothing$, we have $\sigma(C_1) = \rho(C_1^*) - q + 1 \leq \rho(C_2^*) - q + 1 = \sigma(C_2)$.

(iii) We have to show that $A_1 \subset A_2$ and $\sigma(A_1) = \sigma(A_2)$ implies $\sigma(A_1 \cup C) = \sigma(A_2 \cup C)$. The only nontrivial case is when A_1, A_2, C are nonempty. Then $A_1^* \subset A_2^*$ and $\rho(A_1^*) = \rho(A_2^*)$, so that $\sigma(A_1 \cup C) = \rho(A_1^* \cup C^*) - q + 1 = \rho(A_2^* \cup C^*) - q + 1 = \sigma(A_2 \cup C)$.

(iv) Since S was stable and Φ the disjoint union of $S_1, \dots, S_t$ with $S_i = \gamma_i(S)$, the group G acts transitively on the set with the t elements $S_1, \dots, S_t$. When $i \in \Psi$ and $\gamma(S_i) = S_j$, put $\gamma(i) = j$. Then G acts transitively on Ψ. For $C \subset \Psi$, $\gamma(C)^* = \gamma(C^*)$, and when C is nonempty, $\sigma(\gamma(C)) = \rho((\gamma(C))^*) - q + 1 = \rho(\gamma(C^*)) - q + 1 = \rho(C^*) - q + 1 = \sigma(C)$.

(d) **Lemma 17.** *Suppose that* $q \geq 2$ *and every* B *with* $\rho(B) < q$ *has* $|B| < 2(q-1)r^{1-2^{2-q}}$. *Let* S^q *be a maximal* q*-set which is unstable. Then* $|S^q| < 2qr^{1-2^{1-q}}$.

Proof. Let H be the subgroup of G consisting of elements γ with $\gamma(S^q) = S^q$. Then if G has coset decomposition $G = \gamma_1 H \cup \dots \cup \gamma_t H$, the sets $B_1 = \gamma_1(S^q), \dots, B_t = \gamma_t(S^q)$ will be all the distinct images of S^q under G. Say $B_1 = S^q$. Since S^q is unstable, there will be an $i \neq 1$ with $B_1 \cap B_i \neq \varnothing$. There is an element of Φ which lies in more than one of the sets $B_1, \dots, B_t$, say in $\nu \geq 2$ of these sets. Since G acts transitively on Φ, every element of Φ lies in exactly ν of the sets $B_1, \dots, B_t$, so that Φ is covered ν times by $B_1, \dots, B_t$. Writing $s = |S^q| = |B_1| = \dots = |B_t|$, we have

$$st = \nu r \geq 2r. \tag{A.5}$$

Since B_i is a maximal q-set and since $B_i \cup B_j$ strictly contains B_i when $i \neq j$, we have $\rho(B_i \cup B_j) > q$. Now if we had $\rho(B_i \cap B_j) = \rho(B_i)$, then

$\rho(B_j) = \rho((B_i \cap B_j) \cup B_j) = \rho(B_i \cup B_j)$ by property (iii), which is impossible. Therefore $\rho(B_i \cap B_j) < \rho(B_i) \leq q$, so that our hypothesis yields

$$|B_i \cap B_j| < 2(q-1)r^{1-2^{2-q}} \qquad (i \neq j).$$

Therefore for $j = 1, \dots, t$,

$$\text{(A.6)} \qquad \begin{aligned} |B_1 \cup \cdots \cup B_j| &\geq js - \binom{j}{2} \cdot 2(q-1)r^{1-2^{2-q}} \\ &> js - j^2(q-1)r^{1-2^{2-q}}. \end{aligned}$$

If it were true that

$$t < \tfrac{1}{2}(q-1)^{-1} s r^{-1+2^{2-q}} = \psi,$$

say, then (A.6) with $j = t$ yields

$$r = |B_1 \cup \cdots \cup B_t| > \tfrac{1}{2}ts,$$

contradicting (A.5). Thus $t > t_0$ where t_0 is the integer with $\psi - 1 \leq t_0 < \psi$.

The assertion to be proved is that

$$\text{(A.7)} \qquad s < 2qr^{1-2^{1-q}}.$$

If this were not true, then $\psi \geq q/(q-1)$ and $\psi - 1 \geq \psi/q$. By (A.6) with $j = t_0$,

$$\begin{aligned} r \geq |B_1 \cup \cdots \cup B_{t_0}| &\geq t_0 s - t_0^2(q-1)r^{1-2^{2-q}} \\ &> t_0(s - \psi(q-1)r^{1-2^{2-q}}) \\ &= \tfrac{1}{2}t_0 s \geq \tfrac{1}{2}(\psi - 1)s \geq \psi s/2q \\ &> (2q)^{-2} s^2 r^{-1+2^{2-q}}. \end{aligned}$$

But this does imply (A.7) after all.

(e) We now turn to the proof of Proposition 3. We begin with the case when S^k is stable. We have a disjoint union

$$\text{(A.8)} \qquad \Phi = S_1 \cup \cdots \cup S_t$$

where the S_i are the distinct images of S^k under G. Writing $s = |S^k|$ we have $st = r$ and $s < (u/v)r$, so that $t > v/u$, whence $t \geq (v+1)/u$ and

$$s = \frac{r}{t} \leq \frac{u}{v+1}r = \left(\frac{u}{v} - \frac{u}{v(v+1)}\right) r \leq \left(\frac{u}{v} - \frac{1}{2v^2}\right) r,$$

which is much stronger than the assertion of the proposition.

Now when $k = 1$, a maximal k-set must be stable. For if not, and if S^1, $\gamma(S^1)$ are neither identical nor disjoint, then

$$\rho(S^1) = \rho(\gamma(S^1)) = \rho(S^1 \cap \gamma(S^1)) = 1,$$

so that by (iii), $\rho(S^1 \cup \gamma(S^1)) = \rho((S^1 \cap \gamma(S^1)) \cup \gamma(S^1)) = \rho(\gamma(S^1)) = 1$, contradicting the maximality of S^1.

It follows that the proposition is true for $k = 1$. We may proceed by induction on k. In the step from $k-1$ to k we will initially suppose that *every* 1-*set has cardinality* ≤ 1.

We first consider the case when

$$s = |S^k| < 2kr^{1-2^{1-k}}. \tag{A.9}$$

Now if $r \geq (4kv/u)^{2^{k-1}}$, we may conclude that

$$s < 2kr(u/4kv) = (u/2v)r,$$

and (A.4) follows. When $r < (4kv/u)^{2^{k-1}}$, we use the trivial estimate

$$s \leq \left(\frac{u}{v}\right)r - \left(\frac{1}{v}\right) = \left(\frac{u}{v} - \frac{1}{vr}\right)r < \left(\frac{u}{v} - \frac{1}{v(4kv)^{2^{k-1}}}\right)r.$$

We may thus suppose that (A.9) is violated. There is a smallest $q \leq k$ such that there is a q-set S^q with

$$|S^q| \geq 2qr^{1-2^{1-q}}. \tag{A.10}$$

We may take S^q to be a maximal q-set. Under our assumption that 1-sets have cardinality ≤ 1, we may conclude that $q \geq 2$ so that $2 \leq q \leq k$. Since q was chosen smallest possible, S^q must be stable by Lemma 17. Since the case of stable S^k has already been dealt with, we may suppose that $2 \leq q < k$.

We have a decomposition (A.8), where $S_1, \dots, S_t$ are the distinct images of S^q under G. Writing $m = |S^q|$ we have

$$mt = r \tag{A.11}$$

so that by (A.10),

$$t \leq \frac{1}{2q}r^{2^{1-q}}. \tag{A.12}$$

Our maximal k-set S^k may be written as $S^k = X_1 \cup \cdots \cup X_t$ with $X_i \subset S_i$ $(i = 1, \dots, t)$. Say $\rho(X_1) = \cdots = \rho(X_w) = q$ and $\rho(X_{w+1}), \dots, \rho(X_t)$ are $< q$. By the minimality property of q,

$$|X_i| < 2qr^{1-2^{2-q}} \qquad (i = w+1, \dots, t),$$

and

$$|X_{w+1} \cup \cdots \cup X_t| < 2qtr^{1-2^{2-q}} \leq r^{1-2^{1-q}} \tag{A.13}$$

by (A.12).

On the other hand we have $\rho(X_i) = \rho(S_i) = q$ for $i = 1, \dots, w$. We claim that

$$\rho(X_1 \cup \cdots \cup X_i) = \rho(S_1 \cup \cdots \cup S_i) \qquad (i = 1, \dots, w).$$

This is true for $i = 1$, and the induction step from i to $i+1$ follows from two applications of (iii):

$$\rho(X_1 \cup \cdots \cup X_i \cup X_{i+1}) = \rho(S_1 \cup \cdots \cup S_i \cup X_{i+1}) = \rho(S_1 \cup \cdots \cup S_i \cup S_{i+1}).$$

Now $\rho(S_1 \cup \cdots \cup S_w) = \rho(X_1 \cup \cdots \cup X_w) \leq \rho(S^k) \leq k$.

We introduce $\Psi = \{1, \ldots, t\}$ and a function σ as in Lemma 16, with respect to the q-set S^q. We have just seen that $C = \{1, \ldots, w\}$ has $\sigma(C) = \rho(C^*) - q + 1 \leq k - q + 1$. Say $\sigma(C) = c$. We claim that C is a maximal c-set. Otherwise there is an $i \notin C$ with $\sigma(C \cup i) = c$ and $\rho(S_1 \cup \cdots \cup S_w \cup S_i) = \rho(S_1 \cup \cdots \cup S_w) = \rho(X_1 \cup \cdots \cup X_w)$. Thus $\rho(X_1 \cup \cdots \cup X_w) = \rho(X_1 \cup \cdots \cup X_w \cup S_i)$ and by (iii), $\rho(S^k) = \rho(S^k \cup S_i)$ where $i \notin C$, contradicting the maximality of S^k. Again, by (iii) we have

$$\rho(S^k) = \rho(S^k \cup X_1 \cup \cdots \cup X_w) = \rho(S^k \cup S_1 \cup \cdots \cup S_w),$$

and since S^k was a maximal k-set, S^k contains $S_1, \ldots, S_w$. Therefore $m|C| = |C^*| = |S_1 \cup \cdots \cup S_w| \leq |S^k| < (u/v)r = (u/v)mt$, and $|C| < (u/v)t$.

Since C was maximal, the case $c \leq k - q + 1 \leq k - 1$ of the proposition shows that

$$|C| < (u/v - (4kv)^{-(k-q+1)2^{k-q+1}})t.$$

The cardinality $|C^*| = m|C|$ and $r = mt$. Thus from (A.13) we get

$$|S^k| < (u/v - (4kv)^{-(k-q+1)2^{k-q+1}} + r^{-2^{1-q}})r.$$

In the case when

$$r \geq (4kv)^{(k-q+1)2^k} \cdot 2^{2^{q-1}} \tag{A.14}$$

we obtain

$$|S^k| < \left(\frac{u}{v} - \frac{1}{2}(4kv)^{-k \cdot 2^{k-q+1}}\right) r,$$

which gives (A.4) in view of $q \geq 2$. When (A.14) is violated we have $r < (4kv)^{k \cdot 2^k} v^{-1}$ and the trivial estimation

$$|S^k| \leq \left(\frac{u}{v} - \frac{1}{rv}\right) r < \left(\frac{u}{v} - (4kv)^{-k \cdot 2^k}\right) r.$$

We now turn to the case when there is a 1-set of cardinality > 1. In this case $q = 1$ in the above construction. In the factor set Ψ, a subset $D \subseteq \Psi$ with $|D| > 1$ has D^* made up of at least two sets S_i, and since S_i was a maximal 1-set we have $\rho(D^*) > 1$ and $\sigma(D) = \rho(D^*) - q + 1 = \rho(D^*) > 1$. Thus Ψ, σ satisfy the condition that every 1-set has cardinality ≤ 1. Thus by what we have already proved,

$$|C| < (u/v - (4kv)^{-k \cdot 2^k})t.$$

In the present situation, since the sets $X_{w+1}, \ldots, X_t$ have $\rho(X_i) < q = 1$, we have $X_{w+1} = \cdots = X_t = \emptyset$ and $S^k = C^*$, so that

$$|S^k| = |C^*| = m|C| < (u/v - (4kv)^{-k \cdot 2^k})r.$$

References

1. E. Bombieri and W. M. Schmidt, *On Thue's equation*, Invent. Math. **88** (1987), 69–81.
2. E. Bombieri and J. Vaaler, *On Siegel's lemma*, Invent. Math. **73** (1983), 11–32.
3. J. W. S. Cassels, *An introduction to the geometry of numbers*, Grundlehren Math. Wiss., vol. 99, Springer, 1959.
4. H. Esnault and E. Viehweg, *Dyson's Lemma for polynomials in several variables (and the theorem of Roth)*, Invent. Math. **78** (1984), 445–490.
5. J. H. Evertse, *Upper bounds for the numbers of solutions of Diophantine equations*, Math. Centrum, Amsterdam, 1983, pp. 1–127.
6. G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers* (3rd ed.), Oxford, Clarendon Press, 1954.
7. J. Mueller and W. M. Schmidt, *Thue's equation and a conjecture of Siegel*, Acta Math **160** (1988), 207–247.
8. W. M. Schmidt, *Linearformen mit algebraischen Koeffizienten.* II, Math. Ann. **191** (1971), 1–20.
9. ____, *Diophantine approximation*, Lecture Notes in Math., vol. 785, Springer-Verlag, Berlin and New York, 1980.
10. ____, *Thue equations with few coefficients*, Trans. Amer. Math. Soc. **303** (1987), 241–255.
11. ____, *The subspace theorem in Diophantine approximations*, Compositio Math. **69** (1989), 121–173.
12. J. Silverman, *The arithmetic of elliptic curves*, Graduate Texts in Math., vol. 106, Springer-Verlag, Berlin and New York, 1986.

Department of Mathematics, University of Colorado, Boulder, Colorado 80309-0426