Compositio Mathematica **120:** 193–225, 2000.

The Number of Solutions of Polynomial-Exponential Equations

H. P. SCHLICKEWEI[1] and W. M. SCHMIDT[2*]
[1]*Fachbereich Mathematik, Universität Marburg, Lahnberge, 35032 Marburg, Germany*
[2]*Department of Mathematics, University of Colorado, Boulder, CO 80309-0395, U.S.A.*
e-mail: Schmidt@euclid.colorado.edu

(Received: 20 January 1998; accepted in final form: 18 September 1998)

Abstract. We will give explicit bounds for the number of solutions of polynomial-exponential equations. In contrast to earlier work, the bounds are independent of the coefficients of the equations, and they are of only single exponential growth in the number of coefficients.

Mathematics Subject Classifications (1991): 11D61, 11B37.

Key words: exponential diophantine equations, linear recurrence sequences.

I. GENERAL OUTLINE

1. Introduction

We will be concerned with the number of solutions of polynomial-exponential equations. Our equations will be of the type

$$\sum_{\ell=1}^{k} P_\ell(\mathbf{x})\boldsymbol{\alpha}_\ell^{\mathbf{x}} = 0 \tag{1.1}$$

in variables $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}^n$, where the P_ℓ are polynomials with coefficients in an algebraic number field K, and the $\boldsymbol{\alpha}_\ell^{\mathbf{x}}$ are characters $\mathbb{Z}^n \to K^\times$, i.e., $\boldsymbol{\alpha}_\ell^{\mathbf{x}} = \alpha_{\ell 1}^{x_1} \ldots \alpha_{\ell n}^{x_n}$, with given $\alpha_{\ell j} \in K^\times$ $(1 \leqslant \ell \leqslant k, 1 \leqslant j \leqslant n)$.

Very roughly speaking, we will show that subject to certain conditions, the number of solutions is less than $2^{35A^3} d^{6A^2}$, where A is the total number of coefficients of the polynomials $P_1, \ldots, P_k$, and d is the degree of K. As compared to our earlier work [16], our new bound incorporates two improvements. Firstly, it no longer depends on arithmetic properties of the $\alpha_{\ell j}$, except on the degree d of the number field K they lie in. This improvement was made possible by Schlickewei's new method, introduced in [14]. Secondly, our bound is only singly exponential in the number A of coefficients, whereas formerly it was triply exponential. One saving of exponentiation stems from Evertse's version [4] of the Subspace Theorem, which in turn

* The second author supported in part by NSF frant DMS-9401426.

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

rests on Faltings' Product Theorem [8]. Due to these authors' works, a saving of one exponentiation was almost automatic and would not have warranted a lengthy exposition. Most of the novelty of our present work is a new method to save another exponentiation. In most work up to now, and we will mention only a few instances, e.g., work of Evertse, Györy, Stewart and Tijdeman [5], Schlickewei and Schmidt [16], Bombieri and Mueller [1], dependency on the coefficients was eliminated by a determinant argument. But this argument changes an equation with A summands into an equation with $A!$ summands. In contrast, our new argument hinges on an idea from the Geometry of Numbers, which might see further applications.

Before giving a precise formulation of our result, let us briefly recall how partitions $\mathcal{P}$ of the set $\{1,\dots,k\}$ in (1.1) come into play. The equation $2^x - x^y + 3^z - 3^w = 0$ in $(x,y,z,w) \in \mathbb{Z}^4$ is of the type (1.1) with $k = 4$ and constant polynomials. This equation has infinitely many solutions, namely solutions with $x = y$, $z = w$. The point is that if $\mathcal{P}$ is the 'partition of the equation' into the two equations $2^x - 2^y = 0$, $3^z - 3^w = 0$, this system of equations has infinitely many solutions. A more detailed motivation for the partitions is given in [16].

Now let us give precise definitions. Let $\mathcal{P}$ be a partition of the set $\Lambda = \{1,\dots,k\}$. The sets $\lambda \subset \Lambda$ occurring in the partition $\mathcal{P}$ will be considered elements of $\mathcal{P}$:$\lambda \in \mathcal{P}$. Given $\mathcal{P}$, the system of equations

$$\sum_{\ell\in\lambda} P_\ell(\mathbf{x})\boldsymbol{\alpha}_\ell^{\mathbf{x}} = 0 \quad (\lambda \in \mathcal{P}), \tag{1.1 $\mathcal{P}$}$$

is a refinement of (1.1). When $\mathcal{Q}$ is a refinement of $\mathcal{P}$, then (1.1 $\mathcal{Q}$) implies (1.1 $\mathcal{P}$). As in [16], let $\mathcal{S}(\mathcal{P})$ consist of solutions of (1.1 $\mathcal{P}$) which are not solutions of (1.1 $\mathcal{Q}$) where $\mathcal{Q}$ is a proper refinement of $\mathcal{P}$. Every solution of (1.1) lies in $\mathcal{S}(\mathcal{P})$ for some $\mathcal{P}$, but the sets $\mathcal{S}(\mathcal{P})$ for various partitions $\mathcal{P}$ need not be disjoint.

Set $\ell \overset{\mathcal{P}}{\sim} m$ is ℓ, m lie in the same subset λ of $\mathcal{P}$. Let $G(\mathcal{P})$ be the subgroup of $\mathbb{Z}^n$ consisting of $\mathbf{z}$ with $\boldsymbol{\alpha}_\ell^{\mathbf{z}} = \boldsymbol{\alpha}_m^{\mathbf{z}}$ for any ℓ, m with $\ell \overset{\mathcal{P}}{\sim} m$.

Laurent [9] had shown that $\mathcal{S}(\mathcal{P})$ is finite if $G(\mathcal{P}) = \{\mathbf{0}\}$. Write

$$A = \sum_{\ell\in\Lambda} \binom{n+\delta_\ell}{n}, \tag{1.2}$$

where δ_ℓ is the total degree of the polynomial P_ℓ. Note that A is the potential number of nonzero coefficients of the polynomials $P_1,\dots,P_k$. Set

$$B = \max(n, A), \tag{1.3}$$

so that $B = \max(n,k)$ if all the polynomials $P_1,\dots,P_k$ are constants, and $B = A$ otherwise. Denote the cardinality of a set $\mathcal{S}$ by $|\mathcal{S}|$.

THEOREM 1. *Suppose $G(\mathcal{P}) = \{\mathbf{0}\}$. Then*

$$|\mathcal{S}(\mathcal{P})| < N(d,B) = 2^{35B^3} d^{6B^2}. \tag{1.4}$$

If the polynomials P_ℓ are constants, i.e., when we are dealing with a purely exponential equation, the dependence on the degree d can now be avoided (cf. the forthcoming paper by Evertse, Schlickewei and Schmidt [7]).

Another formulation of our Theorem is as follows. Consider a system of equations

$$\sum_{\ell=1}^{k_j} P_{j\ell}(\mathbf{x})\boldsymbol{\alpha}_{j\ell}^{\mathbf{x}} = 0 \quad (j = 1, \dots, m). \tag{1.5}$$

A solution $\mathbf{x}$ will be called *degenerate* if a subsum of one of the m sums in (1.5) vanishes, i.e., if there is a j in $1 \leqslant j \leqslant m$ and a nonempty, proper subset I of $\{1, \dots, k_j\}$ with $\sum_{\ell \in I} P_{j\ell}(\mathbf{x})\boldsymbol{\alpha}_{j\ell}^{\mathbf{x}} = 0$. Let G be the subgroup of $\mathbb{Z}^n$ consisting of vectors $\mathbf{z}$ with $\boldsymbol{\alpha}_{j1}^{\mathbf{z}} = \cdots = \boldsymbol{\alpha}_{j,k_j}^{\mathbf{z}} \quad (j = 1, \dots, m)$.

Write

$$A = \sum_{j=1}^{m} \sum_{\ell=1}^{k_j} \binom{n + \delta_{j\ell}}{n}, \qquad B = \max(n, A),$$

where $\delta_{j\ell}$ is the total degree of the polynomial $P_{j\ell}$. *Then* when $G = \{\mathbf{0}\}$, (1.5) *has at most* $N(d, B)$ *nondegenerate solutions.*

In a forthcoming paper S. Ahlgren will give a quantitative version of a more general theorem of Laurent [9] which describes the set of solutions when the group $G(\mathcal{P})$ is not necessarily $\{\mathbf{0}\}$.

Before commencing with the proof of Theorem 1 in Section 3, we will now give some applications.

2. Applications of Theorem 1 to Linear Recurrence Sequences

Let $\{u_m\}_{m \in \mathbb{Z}}$ be a linear recurrence sequence of order t, i.e., a not identically vanishing sequence satisfying a relation

$$u_{m+t} = v_{t-1}u_{m+t-1} + \cdots + v_1 u_{m+1} + v_0 u_m \quad (m \in \mathbb{Z}), \tag{2.1}$$

with $t > 0$ and fixed coefficients $v_0, \dots, v_{t-1}$, but no such relation with $0 < t' < t$. Then $v_0 \neq 0$. We will suppose that all members of the sequence lie in a number field K, and this (by the minimality of t) easily implies that $v_0, \dots, v_{t-1}$ lie in K. Let

$$F(z) = z^t - v_{t-1}z^{t-1} - \cdots - v_0 = \prod_{\ell=1}^{k} (z - \alpha_\ell)^{\sigma_\ell} \tag{2.2}$$

be the *companion polynomial* of the relation (2.1), with $\alpha_1, \dots, \alpha_k$ being the distinct roots. As $v_0 \neq 0$, these roots are nonzero. The sequence will be called

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

nondegenerate if no quotient α_ℓ/α_n with $\ell \neq n$, $1 \leqslant \ell, n \leqslant k$ is a root of unity. It will be called *strictly nondegenerate* if, with $\alpha_0 = 1$, no quotient α_ℓ/α_n with $\ell \neq n$, $0 \leqslant \ell, n \leqslant k$ is a root of unity. The *a-multiplicity* of $\{u_m\}$, denoted $\mathcal{U}(a)$, is the number of $m \in \mathbb{Z}$ with $u_m = a$.

THEOREM 2.1. *Let $\{u_m\}$ be of order t, and with elements in a number field K of degree d. When $\{u_m\}$ is nondegenerate, then*

$$\mathcal{U}(0) < (2t)^{35t^3} d^{6t^2}. \tag{2.3}$$

When $\{u_m\}$ is strictly nondegenerate, then for every $a \in K$,

$$\mathcal{U}(a) < (2t)^{36(t+1)^3} d^{6(t+1)^2}. \tag{2.4}$$

The bounds for $\mathcal{U}(0)$ and $\mathcal{U}(a)$ derived in [16] also depended only on d and t, but the dependence on t was triply exponential. When the companion polynomial has only simple roots, we are reduced to a purely exponential equation, so that there is a bound independent of d (cf. [7]). But this bound is doubly exponential in t. Recently Schmidt (in work in progress) obtained in the one variable case of Theorem 1 an estimate independent of d, which is however triply exponential in t, and this entails a version of Theorem 2.1 independent of d, which is triply exponential in t. His work depends on Proposition A formulated below, as well as on our Lemma 15.1.

Proof of Theorem 2.1. It is well known that u_m has a representation

$$u_m = \sum_{\ell=1}^{k} P_\ell(m)\alpha_\ell^m, \tag{2.5}$$

where P_ℓ is a nonzero polynomial of degree $\sigma_\ell - 1$ with coefficients in the field $L = K(\alpha_1, \ldots, \alpha_k)$. Since K has degree d, (2.2) yields $\deg L \leqslant dt!$. Now $\mathcal{U}(0)$ is the number of solutions of the equation

$$\sum_{\ell=1}^{k} P_\ell(m)\alpha_\ell^m = 0. \tag{2.7}$$

This equation is of the type (1.1) with $n = 1$. The quantity A from (1.2) becomes $\sigma_1 + \cdots + \sigma_k = t$, and thus also the quantity B from (1.3) equals t. It will suffice to study equations (2.7 $\mathcal{P}$) for every partition $\mathcal{P}$ of $\{1, \ldots, k\}$.

When $\mathcal{P}$ contains a singleton, then $|\mathcal{S}(\mathcal{P})| < t$, since our polynomials P_ℓ have degree $\sigma_\ell - 1 < t$. Otherwise $\mathcal{P}$ contains a set λ with $|\lambda| \geqslant 2$, and when our sequence is nondegenerate, we may conclude that $G(\mathcal{P}) = \{0\}$. Theorem 1 in conjunction with (2.6) gives $|\mathcal{S}(\mathcal{P})| < 2^{35t^3}(dt!)^{6t^2}$. This estimate therefore holds

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

for every partition $\mathcal{P}$. Using the bound k^k for the number of partitions $\mathcal{P}$, we obtain $\mathcal{U}(0) < k^k \cdot 2^{35t^3}(dt!)^{6t^2} < (2t)^{35t^3}d^{6t^2}$.

We next note that $\mathcal{U}(a)$ is the zero-multiplicity of the sequence $u'_m = u_m - a$. When $\{u_m\}$ is strictly nondegenerate of order t, then $\{u'_m\}$ is nondegenerate of order $t+1$. Therefore the argument given above may be applied with $t+1$ in place of t. We now have $k \leqslant t+1$. However, (2.6) is still valid as before. Hence

$$\begin{aligned}\mathcal{U}(a) &< k^k 2^{35(t+1)^3}(dt!)^{6(t+1)^2}\\ &< (t+1)^{t+1} \cdot 2^{35(t+1)^3} t^{6t(t+1)^2} d^{6(t+1)^2}\\ &< (2t)^{36(t+1)^3} d^{6(t+1)^2}.\end{aligned}$$

Remark. Since we suppose that $\{u_m\}$ is (strictly) of order t, so that the polynomials P_ℓ are nonzero, the hypothesis for (2.3) that $\{u_m\}$ be nondegenerate may be replaced by the weaker hypothesis that for some α_n, no quotient α_ℓ/α_n with $1 \leqslant \ell \leqslant k$ and $\ell \neq n$ is a root of 1. Similarly for (2.4) with $0 \leqslant \ell \leqslant k$ and $\ell \neq n$.

Now let $\{u_m\}_{m\in\mathbb{Z}}$ and $\{v_n\}_{n\in\mathbb{Z}}$ be nondegenerate recurrence sequences of order $\leqslant t$, and consider the equation

$$u_m = v_n \tag{2.8}$$

in integers m, n. In view of the special rôle played by roots of unity, we will change the notation (2.2) for the polynomial $F(z)$ associated with $\{u_m\}$. Let $\alpha_1, \dots, \alpha_{k_1}$ be the roots of F which are not roots of unity. We will write $F(z) = \prod_{\ell=0}^{k_1}(z-\alpha_\ell)^{\rho_\ell}$, where either α_0 is a root of F which is a root of unity (such α_0 then is unique), or $\alpha_0 = 1$, $\rho_0 = 0$. Then

$$u_m = \sum_{\ell=0}^{k_1} P_\ell(m)\alpha_\ell^m, \tag{2.9}$$

where P_ℓ is a polynomial of degree $\rho_\ell - 1$. (A polynomial of degree 0 is a nonzero constant, and a polynomial of degree -1 is zero.) Similarly,

$$v_n = \sum_{\ell=0}^{k_2} Q_\ell(n)\beta_\ell^n. \tag{2.10}$$

The sequences $\{u_m\}$, $\{v_n\}$ are said to be *related* if $k_1 = k_2$ $(= k$, say), and after a suitable reordering of $\beta_1, \dots, \beta_k$,

$$\alpha_\ell^p = \beta_\ell^q \quad (\ell = 1, \dots, k), \tag{2.11}$$

with nonzero integers p, q. They are *doubly related* if there is a second reordering of $\beta_1, \dots, \beta_k$ with this property, i.e., if there is a nontrivial permutation π of

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

$\{1, \dots, k\}$ such that we have both (2.11) and $\alpha_\ell^{p'} = \beta_{\pi(\ell)}^{q'}$ $(\ell = 1, \dots, k)$ with nonzero integers p', q'. Then it was shown in [17] that k is even, that $p'/q' = -p/q$, and after a suitable reordering of $\alpha_1, \dots, \alpha_k$ and $\beta_1, \dots, \beta_k$ we have both (2.11) and

$$\alpha_\ell^{p'} = \beta_{\ell+1}^{q'}, \qquad \alpha_{\ell+1}^{p'} = \beta_\ell^{q'} \quad \text{for } \ell \text{ odd}, \ 1 \leqslant \ell \leqslant k. \tag{2.12}$$

The sequences $\{u_m\}$ and $\{v_n\}$ are called *simply related* if they are not doubly related. A sequence $\{u_m\}$ is always related to itself; it is called *symmetric* if it is doubly related to itself.

THEOREM 2.2 *Suppose the members of* $\{u_m\}, \{v_n\}$ *lie in a number field* K *of degree* d. *Suppose* $k_1 > 0$, $k_2 > 0$ *in* (2.9), (2.10). *Then*

(a) *the Equation* (2.8) *has at most*

$$Z = 2^{310t^6} d^{24t^4} \tag{2.13}$$

solutions when $\{u_m\}, \{v_n\}$ *are not related.*

(b) *When* $\{u_m\}, \{v_n\}$ *are simply related with* (2.11), *then all but at most* Z *solutions of* (2.8) *have*

$$P_\ell(m)\alpha_\ell^m = Q_\ell(n)\beta_\ell^n \quad (\ell = 0, \dots, k). \tag{2.14}$$

(c) *When* $\{u_m\}, \{v_n\}$ *are doubly related with* (2.11), (2.12), *then all but at most* Z *solutions satisfy* (2.14) *or the system*

$$P_\ell(m)\alpha_\ell^m = Q_{\ell+1}(n)\beta_{\ell+1}^n, \qquad P_\ell(m)\alpha_{\ell+1}^m = Q_\ell(n)\beta_\ell^n$$

$$(\ell \text{ odd}, 1 \leqslant \ell \leqslant k), \tag{2.15i}$$

$$P_0(m)\alpha_0^m = Q_0(n)\beta_0^n. \tag{2.15ii}$$

It may easily be deduced that when $\{u_m\}$ is not symmetric, the equation $u_m = u_n$ has at most Z solutions with $m \neq n$.

An estimate given in [18] was weaker in its dependence on t and d, and moreover it involved the number of prime ideal factors of the roots α_ℓ and β_ℓ. The order or magnitude of our estimates can be further reduced when the α_ℓ and β_ℓ are simple roots.

Proof of Theorem 2.2. We rewrite (2.8) as

$$\sum_{\ell=0}^{k_1} P_\ell(x)\alpha_\ell^x - \sum_{\ell=0}^{k_2} Q_\ell(y)\beta_\ell^y = 0, \tag{2.16}$$

to be solved in integers x, y. We symbolize the summands in (2.16) by

$$(0_x), 1_x, \dots, k_{1x}, (0_y), 1_y, \dots, k_{2y}.$$

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

The parentheses indicate that, e.g., 0_x occurs only when α_0 is a root of $F(x)$, i.e., only if $P_0 \neq 0$. Let $\mathscr{P}$ be a partition of this set, and $G(\mathscr{P})$ the associated group. Similarly to (1.1 $\mathscr{P}$), let (2.16 $\mathscr{P}$) denote the system obtained by splitting (2.16) into vanishing subsums, the summands of each subsum parametrized by a set $\lambda \in \mathscr{P}$. Suppose at first that $\mathscr{P}$ contains a singleton, say ℓ_y. Then (2.16 $\mathscr{P}$) yields $Q_\ell(y)\beta_\ell^y = 0$, and since Q_ℓ is of degree $< t$, there are fewer than t choices for y. Given y, (2.16) becomes an equation in x of the type considered in Theorem 2.1. We therefore can estimate the number of choices for x by (2.4). Thus when $\mathscr{P}$ contains a singleton, $|\mathscr{S}(\mathscr{P})| < t \cdot (2t)^{36(t+1)^3} d^{6(t+1)^2}$. Now suppose that $\mathscr{P}$ does not contain a singleton. Then, as shown in [17], $G(\mathscr{P}) = \{\mathbf{0}\}$ unless $\{u_m\}$ and $\{v_n\}$ are related. Note that the field $L = K(\alpha_1, \ldots, \alpha_{k_1}, \beta_1, \ldots, \beta_{k_2})$ has degree $\leqslant (t!)^2 d < 2^{(3/4)t^2} d$, since $t! < 2^{(3/8)t^2}$. We apply Theorem 1 with $n = 2$ and observe that (with $\delta_t = \deg P_\ell = \rho_\ell - 1$, $\delta'_\ell = \deg Q_\ell = \rho'_\ell - 1$, say),

$$
\begin{aligned}
A &= \sum_{\ell=0}^{k_1} \binom{\rho_\ell + 1}{2} + \sum_{\ell=0}^{k_2} \binom{\rho'_\ell + 1}{2} \\
&= \frac{1}{2}\left(\sum_{\ell=0}^{k_1} (\rho_\ell^2 + \rho_\ell) + \sum_{\ell=0}^{k_2} (\rho_\ell'^2 + \rho'_\ell) \right) \\
&\leqslant \frac{1}{2}((t^2 + t) + (t^2 + t)) \leqslant 2t^2,
\end{aligned}
$$

so that $B = A \leqslant 2t^2$. Therefore

$$
\begin{aligned}
|\mathscr{S}(\mathscr{P})| &< 2^{36B^3} (2^{(3/4)t^2} d)^{6B^2} \\
&\leqslant 2^{306t^6} d^{24t^4}.
\end{aligned}
$$

Since the number of partitions $\mathscr{P}$ is at most $(2t)^{2t} < 2^{2t^6}$, the first assertion of Theorem 2.2 follows.

Now if $\{u_m\}$ and $\{v_n\}$ are simply related, it was shown in [17] that the only partition $\mathscr{P}$ which does not contain a singleton and has $G(\mathscr{P}) \neq \{\mathbf{0}\}$ is $\{0_x, 0_y\}$, $\{1_x, 1_y\}, \ldots, \{k_x, k_y\}$. (Here $\{0_x, 0_y\}$ occurs only if P_0, Q_0 are nonzero.) So (b) follows as well. As for (c), in addition to the exceptional partition from (b), again by [17], we need only consider partitions containing the sets $\{1_x, 2_y\}$, $\{2_x, 1_y\}, \ldots$, $\{(k-1)_x, k_y\}$, $\{k_x, (k-1)_y\}$.

3. A Proposition on Linear Equations

We will formulate a proposition which may be of independent interest.

Consider the multiplicative group $(\mathbb{C}^\times)^m = \mathbb{C}^\times \times \cdots \times \mathbb{C}^\times$, and a subgroup Γ of finite rank r. In [7] we had studied the equation

$$z_1 + \cdots + z_m = 1 \tag{3.1}$$

in variables $\mathbf{z} = (z_1, \ldots, z_m) \in \Gamma$. Here we will need this equation in variable $\mathbf{z}$ which lies 'almost' in Γ.

We cannot go further without introducing heights. We define the height $H(\boldsymbol{\alpha})$ of a point $(\alpha_0 : \cdots : \alpha_m)$ in projective space $\mathbb{P}^m(\overline{\mathbb{Q}})$ as usual. Suppose $\alpha_0, \ldots, \alpha_m$ lie in a number field K, and let $V = V(K)$ be the set of places of K. With each $v \in V$ we associate the absolute value $|\ \ |_v$, normalized so that it extends the standard or a p-adic absolute value of $\mathbb{Q}$, and we further set $\|\alpha\|_v = |\alpha|_v^{d_v/d}$, where d is the degree of K, and d_v the local degree. We then define $H(\boldsymbol{\alpha}) = \prod_{v \in V(K)} \|\boldsymbol{\alpha}\|_v$, where $\|\boldsymbol{\alpha}\|_v = \max\{\|\alpha_0\|_v, \ldots, \|\alpha_m\|_v\}$. By the product formula $H(\boldsymbol{\alpha})$ depends only on the projective point $\boldsymbol{\alpha} = (\alpha_0 : \cdots : \alpha_m)$. It is independent of the field K with $\alpha_i \in K$ $(i = 0, \ldots, m)$ and is usually called the *absolute multiplicative height*. We will also use the *absolute logarithmic height* $h(\boldsymbol{\alpha}) = \log H(\boldsymbol{\alpha})$.

When $\mathbf{x} = (x_1, \ldots, x_m)$ is in affine space $\overline{\mathbb{Q}}^m$, we set

$$H(\mathbf{x}) = H(1 : x_1 : \cdots : x_m), \qquad h(\mathbf{x}) = h(1 : x_1 : \cdots : x_m) = \log H(\mathbf{x}).$$

In particular, when $m = 1$, we have $H(x) = H(1 : x)$, $h(x) = h(1 : x)$.

Now let K be a number field of degree d. When $\mathbf{x} = (x_1, \ldots, x_m)$, $\mathbf{y} = (y_1, \ldots, y_m)$ are in K^m, we set $\mathbf{x} * \mathbf{y} = (x_1 y_1, \ldots, x_m y_m)$.

PROPOSITION A. *Let $m \geqslant 1$ and let Γ be a finitely generated subgroup of $(K^\times)^m$ of rank $r \geqslant 0$. Then the solutions $\mathbf{z}$ of* (3.1) *of the type $\mathbf{z} = \mathbf{x} * \mathbf{y}$ where $\mathbf{x} \in \Gamma$, $\mathbf{y} \in \mathbb{Q}^m$ and*

$$h(\mathbf{y}) \leqslant \frac{1}{4m^2} h(\mathbf{x}) \tag{3.2}$$

are contained in the union of at most $f(m, r, d) = 2^{30m^2}(32m^2)^r d^{3r+2m}$ proper linear subspaces of K^m.

4. The Germ of the Proof of Theorem 1

Let $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k$ be as in the theorem. As $\mathbf{x}$ runs through $\mathbb{Z}^n$, the vector

$$(\boldsymbol{\alpha}_1^{\mathbf{x}}, \ldots, \boldsymbol{\alpha}_k^{\mathbf{x}}) \tag{4.1}$$

runs through a subgroup Γ of $(K^\times)^k$ of rank $\leqslant n$. If in (1.1) the polynomials P_ℓ are all identically equal to 1, we obtain an equation

$$z_1 + \cdots + z_k = 0, \tag{4.2}$$

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

with $\mathbf{z} = (z_1, \ldots, z_k) \in \Gamma$. This is a homogeneous version of Equation (3.1). Now (4.2) defines a subspace T of K^k of codimension 1, and it is known (cf. [7]) that the solutions $\mathbf{z} \in \Gamma$ lie in a finite number (and this number may be effectively estimated) of proper subspaces of T (thus subspaces of K^k of codimension $\geqslant 2$).

This gives us information on the equation $\sum_{\ell=1}^{k} \boldsymbol{\alpha}_\ell^{\mathbf{x}} = 0$. The situation is similar for

$$\sum_{\ell=1}^{k} a_\ell \boldsymbol{\alpha}_\ell^{\mathbf{x}} = 0 \tag{4.3}$$

with coefficients $a_\ell \in K^\times$: one could consider it of the type (4.2) with Γ the group of rank $\leqslant n+1$ generated by the points (4.1) and by $(a_1, \ldots, a_k)$.

Now in Equation (1.1), let $\mathbf{M}_\ell$ be the set of monomials of total degree $\leqslant \delta_\ell$. Write $P_\ell = \sum_{M \in \mathbf{M}_\ell} a_{\ell M} M \quad (1 \leqslant \ell \leqslant k)$.

Then Equation (1.1) may be rewritten as $\sum_{(\ell,M) \in \mathcal{A}} M(\mathbf{x}) a_{\ell M} \boldsymbol{\alpha}_\ell^{\mathbf{x}} = 0$, where $\mathcal{A}$ consists of the pairs (ℓ, M) with $1 \leqslant \ell \leqslant k$, $M \in \mathbf{M}_\ell$ and $a_{\ell M} \neq 0$. With the notation $\eta_{\ell M}(\mathbf{x}) = M(\mathbf{x}) a_{\ell M} \boldsymbol{\alpha}_\ell^{\mathbf{x}}$, the equation becomes

$$\sum_{(\ell,M) \in \mathcal{A}} \eta_{\ell M}(\mathbf{x}) = 0. \tag{4.4}$$

If it were not for the monomials $M(\mathbf{x})$, this would be the type (4.3). The vector $\boldsymbol{\eta}(\mathbf{x})$ with components $\eta_{\ell M}(\mathbf{x})$ lies in K^a where $a = |\mathcal{A}|$, and (4.4) says that $\boldsymbol{\eta}(\mathbf{x})$ lies in a certain subspace T of K^a of codimension 1. We wish to show that as $\mathbf{x} \in \mathbb{Z}^n$ ranges through the solutions of (4.4), then $\boldsymbol{\eta}(\mathbf{x})$ lies in a finite union of proper subspaces of T, and we want to estimate the number of required subspaces. This can in fact be done if the vector with components

$$M(\mathbf{x}) \quad (M \in \mathbf{M} = \mathbf{M}_1 \cup \cdots \cup \mathbf{M}_k) \tag{4.5}$$

is 'small' compared to the vector with components

$$a_{\ell m} \boldsymbol{\alpha}_\ell^{\mathbf{x}} \quad ((\ell, m) \in \mathcal{A}). \tag{4.6}$$

Let $h_M(\mathbf{x})$ be the logarithmic height of the vector (4.5), and $h_E(\mathbf{x})$ the height of the vector (4.6).

PROPOSITION B. *Suppose $a \geqslant 3$. Then as $\mathbf{x}$ ranges through solutions of* (4.4) *with*

$$h_M(\mathbf{x}) \leqslant \frac{1}{4a^2} h_E(\mathbf{x}), \tag{4.7}$$

the vector $\boldsymbol{\eta}(\mathbf{x})$ will be contained in a union of not more than

$$2^{30a^2} (32a^2)^n d^{3(n+a)} \tag{4.8}$$

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

proper subspaces of T.

We will now deduce Proposition B from Proposition A. Let (ℓ_0, M_0) be a particular element of $\mathcal{A}$, and $\mathcal{A}'$ the complement of (ℓ_0, M_0) in $\mathcal{A}$. Define

$$\boldsymbol{\beta}_\ell = \boldsymbol{\alpha}_\ell/\boldsymbol{\alpha}_{\ell_0} = (\alpha_{\ell 1}/\alpha_{\ell_0 1}, \ldots, \alpha_{\ell n}/\alpha_{\ell_0 n}). \tag{4.9}$$

Then when $M_0(\mathbf{x}) \neq 0$, (4.4) may be rewritten as

$$\sum_{(\ell,M)\in\mathcal{A}'} Z_{\ell M} = 1 \tag{4.10}$$

where $Z_{\ell M} = X_{\ell M} Y_{\ell M}$ with

$$X_{\ell M} = -(a_{\ell M}/a_{\ell_0} M_0)\boldsymbol{\beta}_\ell^{\mathbf{x}}, \qquad Y_{\ell M} = M(\mathbf{x})/M_0(\mathbf{x}).$$

Let $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ respectively be the points in K^{a-1} with components $X_{\ell M}, Y_{\ell M}, Z_{\ell M}$ where $(\ell, M) \in \mathcal{A}'$. Then $\mathbf{X}$ lies in a group of rank $\leqslant n+1$, and $\mathbf{Y}$ lies in $\mathbb{Q}^{a-1}$. Further $h_M(\mathbf{x}) = h(\mathbf{Y})$, $h_E(\mathbf{x}) = h(\mathbf{X})$, so that

$$h(\mathbf{Y}) \leqslant \frac{1}{4a^2} h(\mathbf{X}) \tag{4.11}$$

by (4.7).

By Proposition A with $m = a-1$, the solutions $\mathbf{Z}$ of (4.10) with (4.11) lie in the union of

$$\begin{aligned} f(a-1, n+1, d) &= 2^{30(a-1)^2}(32(a-1)^2)^{n+1} d^{3(n+1)+2(a-1)} \\ &< 2^{30a^2}(32a^2)^n d^{3(n+a)} \end{aligned} \tag{4.12}$$

subspaces of K^{a-1}. Here the $n+1$ comes from the fact that $\mathbf{X}$ runs through a group of rank $\leqslant n+1$. When W is one of these subspaces, the solutions of (4.4) with $\mathbf{Z}(\mathbf{x}) \in W$ will have $\boldsymbol{\eta}(\mathbf{x})$ in a certain proper subspace W' of T. To these subspaces W' we have to add the subspace with $M_0(\mathbf{x}) = 0$, thus giving the bound (4.8).

5. Induction on the Dimension of Subspaces

The vectors $\boldsymbol{\xi}$ with components $\xi_{\ell M}$ where $\ell \in \Lambda$, $M \in \mathbf{M}_\ell$ lie in K^A with A given by (1.2).

When λ is any subset of Λ, let V_λ be the coordinate subspace of K^A consisting of vectors $\boldsymbol{\xi}$ with $\xi_{\ell M} = 0$ when $\ell \notin \lambda$. For any partition $\mathcal{Q}$ of Λ,

$$K^A = \bigoplus_{\lambda\in\mathcal{Q}} V_\lambda. \tag{5.1}$$

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

When W is a subspace of K^A, let $W(\mathcal{Q}) = \sum_{\lambda \in \mathcal{Q}} (W \cap V_\lambda)$, so that $W(\mathcal{Q})$ is a subspace of W. We have $W(\mathcal{Q}') \subseteq W(\mathcal{Q})$ if $\mathcal{Q}'$ is a refinement of $\mathcal{Q}$. We will say that $\mathcal{Q}$ is *agreeable* with W if $W(\mathcal{Q}) = W$. If $\mathcal{Q}'$ is agreeable with W where $\mathcal{Q}'$ is a refinement of $\mathcal{Q}$, then $\mathcal{Q}$ is agreeable with W. Write $\mathcal{Q} \prec W$ if $\mathcal{Q}$ is agreeable with W, but no proper refinement of $\mathcal{Q}$ is agreeable with W. For any W, there is a $\mathcal{Q}$ with $\mathcal{Q} \prec W$, but this $\mathcal{Q}$ is not necessarily unique.

Suppose for each $\ell \in \Lambda$ we are given a polynomial P_ℓ of degree $\leqslant \delta_\ell$. Thus

$$P_\ell = \sum_{M \in \mathbf{M}_\ell} a_{\ell M} M \quad (\ell \in \Lambda). \tag{5.2}$$

Given $\mathbf{x} \in \mathbb{Z}^n$, let $\boldsymbol{\xi} = \boldsymbol{\xi}(\mathbf{x}) = K^A$ have components $\xi_{\ell M} = \xi_{\ell M}(\mathbf{x}) = M(\mathbf{x})\boldsymbol{\alpha}_\ell^{\mathbf{x}}$. The equations (1.1 $\mathcal{P}$) mean that $\boldsymbol{\xi}(\mathbf{x})$ lies in the subspace W of K^A defined by

$$\sum_{\ell \in \lambda} \sum_{M \in \mathbf{M}_\ell} a_{\ell M} \xi_{\ell M} = 0 \quad (\lambda \in \mathcal{P}). \tag{5.3}$$

For any subspace T of K^A, let $\mathcal{X}(T)$ consist of $\mathbf{x} \in \mathbb{Z}^n$ with $\boldsymbol{\xi}(\mathbf{x}) \in T$. Let $\mathcal{X}(T, \mathcal{P})$ consist of $\mathbf{x}$ with $\boldsymbol{\xi}(\mathbf{x}) \in T(\mathcal{P})$, but $\boldsymbol{\xi}(\mathbf{x}) \notin T(\mathcal{Q})$ for any proper refinement $\mathcal{Q}$ of $\mathcal{P}$. In the notation of the Introduction, $\mathcal{S}(\mathcal{P}) = \mathcal{X}(W, \mathcal{P})$ where W is given by (5.3).

PROPOSITION C. *Recall the definition* (1.3) *of B and set*

$$C = 2^{34B^2} d^{6B}. \tag{5.4}$$

Let $\mathcal{P}$ be a partition of Λ with $G(\mathcal{P}) = \{\mathbf{0}\}$. Let $T \neq \{\mathbf{0}\}$ be a subspace of K^A with $\mathcal{P} \prec T$. Then there is a subspace $T' \subsetneqq T$ having $T'(\mathcal{P}) = T'$ and

$$|\mathcal{X}(T, \mathcal{P})| \leqslant C|\mathcal{X}(T, \mathcal{P}) \cap \mathcal{X}(T')| + C. \tag{5.5}$$

We are going to derive Theorem 1 from the proposition. First we claim that every subspace T with $\mathcal{P} \prec T$ and dimension t has

$$|\mathcal{X}(T, \mathcal{P})| \leqslant (2C)^t. \tag{5.6}$$

This is done by induction on t. When $t = 0$, then $\mathcal{X}(T)$ is empty, since $\boldsymbol{\xi}(\mathbf{x}) = \mathbf{0}$ is impossible because $\boldsymbol{\xi}(\mathbf{x})$ has the nonzero components $\xi_{\ell M}(\mathbf{x}) = \boldsymbol{\alpha}_\ell^{\mathbf{x}}$ when $M = 1$. Thus (5.6) is true in this case. When $t > 0$, let T' be the subspace of the proposition. There are two possibilities.

Either $\mathcal{P} \prec T'$ fails to hold. There is then a proper refinement $\mathcal{Q}$ of $\mathcal{P}$ with $T'(\mathcal{Q}) = T'$. Then $\mathcal{X}(T') = \mathcal{X}(T'(\mathcal{Q})) \subseteq \mathcal{X}(T(\mathcal{Q}))$ has empty intersection with $\mathcal{X}(T, \mathcal{P})$, so that $\mathcal{X}(T, \mathcal{P}) = \emptyset$ by (6.5). *Or* $\mathcal{P} \prec T'$. Then every $\mathbf{x} \in \mathcal{X}(T, \mathcal{P}) \cap \mathcal{X}(T')$ has $\boldsymbol{\xi}(\mathbf{x}) \in T' = T'(\mathcal{P})$, but in view of $\mathbf{x} \in \mathcal{X}(T, \mathcal{P})$ it cannot have $\boldsymbol{\xi}(\mathbf{x}) \in$

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

$T'(\mathcal{Q}) \subset T(\mathcal{Q})$ for a proper refinement $\mathcal{Q}$ of $\mathcal{P}$. Therefore $\mathbf{x} \in \mathcal{X}(T', \mathcal{P})$, i.e., $\mathcal{X}(T, \mathcal{P}) \cap \mathcal{X}(T') \subset \mathcal{X}(T', \mathcal{P})$. Now (5.5) together with the induction hypothesis gives

$$|\mathcal{X}(T, \mathcal{P})| \leqslant C|\mathcal{X}(T', \mathcal{P})| + C \leqslant C \cdot (2C)^{t-1} + C \leqslant (2C)^t.$$

The theorem is about $\mathcal{S}(\mathcal{P}) = \mathcal{X}(W, \mathcal{P})$ with W given by (5.3). Clearly $W(\mathcal{P}) = W$. Again there are two possibilities. Either $\mathcal{P} \prec W$ fails to hold (this could only happen if some polynomials P_ℓ are zero). Then $W = W(\mathcal{Q})$ where $\mathcal{Q}$ is a proper refinement of $\mathcal{P}$, so that $\mathcal{X}(W, \mathcal{P}) = \emptyset$. Or $\mathcal{P} \prec W$. Then we may apply (5.6) to $T = W$. Since $\dim T \leqslant A \leqslant B$, we obtain $|\mathcal{S}(\mathcal{P})| \leqslant (2C)^B < 2^{35B^3} d^{6B^2}$. The theorem follows.

It remains for us to prove Proposition A, and to show that Proposition B can be used to deduce Proposition C. The first of these tasks will be accomplished in Sections 6–11, the second in Sections 12–17. The second task is the more original one. The geometric idea alluded to above will occur in the proof of Lemma 15.1. Unfortunately, our arguments will be rather complicated.

II. PROOF OF PROPOSITION A

6. Small Solutions

We will initially only study solutions $\mathbf{z} = \mathbf{x} * \mathbf{y}$ of (3.1) with $\mathbf{x} \in \Gamma$, $\mathbf{y} \in (\mathbb{Q}^\times)^m$, so that all the components of $\mathbf{z}$ are nonzero. A solution will be called *small* if

$$h(\mathbf{x}) \leqslant 2m \log m. \tag{6.1}$$

A solution which is not small will be called *large*.

LEMMA 6.1. *The number of small solutions* $\mathbf{z}$ *occurring in Proposition A is*

$$< (4d^2)^m (86d^2 m \log m)^r. \tag{6.2}$$

Proof. According to Theorem 4 of Schmidt [19] the number of elements $\mathbf{x} \in \Gamma$ with $h(\mathbf{x}) \leqslant 2m \log m$ does not exceed

$$(2d^2)^m (86d^2 m \log m)^r. \tag{6.3}$$

Further, $h(\mathbf{y}) \leqslant (4m^2)^{-1} h(\mathbf{x}) \leqslant (2m)^{-1} \log m \leqslant 1/2$ by (3.2). Therefore each component y_i of $\mathbf{y}$ has $h(y_i) \leqslant 1/2$, hence $H(y_i) \leqslant \mathrm{e}^{1/2} < 2$, so that y_i, being rational, is 1 or -1. This gives 2^m choices for $\mathbf{y}$. Allowing a factor 2^m in (6.3) we get the assertion.

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

7. Remarks on Heights

For $x \in K^\times$ we note that

$$h(x) = h(1{:}x) = \sum_{v\in V(K)} \max\{0, \log \|x\|_v\}$$
$$= \frac{1}{2}\sum_{v\in V(K)} |\log \|x\|_v|. \tag{7.1}$$

We then have $h(1/x) = h(x), \quad h(xy) \leqslant h(x) + h(y)$.

As was pointed out in [17], it is an immediate consequence of work of Dobrowolski [3] that when x is of degree d, and not zero or a root of unity, then

$$h(x) > 1/21d^3. \tag{7.2}$$

When $\mathbf{x} = (x_1, \ldots, x_m) \in \overline{\mathbb{Q}}^m$, we will also use the logarithmic height $h_s(\mathbf{x}) = \sum_{i=1}^m h(x_i)$. We notice that

$$h(\mathbf{x}) \leqslant h_s(\mathbf{x}) \leqslant mh(\mathbf{x}), \tag{7.3}$$

$$h(\mathbf{x} * \mathbf{y}) \leqslant h(\mathbf{x}) + h(\mathbf{y}), \qquad h_s(\mathbf{x} * \mathbf{y}) \leqslant h_s(\mathbf{x}) + h_s(\mathbf{y}), \tag{7.4}$$

$$h_s(\mathbf{x}^{-1}) = h_s(\mathbf{x}), \tag{7.5}$$

where $\mathbf{x}^{-1}$ denotes the inverse of $\mathbf{x}$ in $(\overline{\mathbb{Q}}^\times)^m$.

Let $\Gamma \subseteq (K^\times)^m$ be a finitely generated group of rank $r > 0$. Let $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ be a set of generators of Γ, so that the elements of Γ may be written as

$$\mathbf{x} = \boldsymbol{\zeta} * \boldsymbol{\alpha}_1^{u_1} * \cdots * \boldsymbol{\alpha}_r^{u_r}, \tag{7.6}$$

where $(u_1, \ldots, u_r)$ runs through $\mathbb{Z}^r$, and $\boldsymbol{\zeta}$ runs through the torsion group $T(\Gamma) = \Gamma \cap U^m$ of Γ, with U the group of roots of unity of K.

For $\mathbf{u} = (u_1, \ldots, u_r) \in \mathbb{Z}^r$ we put

$$\psi(\mathbf{u}) = h_s(\boldsymbol{\alpha}_1^{u_1} * \cdots * \boldsymbol{\alpha}_r^{u_r}). \tag{7.7}$$

For $v \in V$ write

$$\beta_{ijv} = \log \|\alpha_{ij}\|_v \quad (1 \leqslant i \leqslant r, 1 \leqslant j \leqslant m),$$

where $\boldsymbol{\alpha}_i = (\alpha_{i1}, \ldots, \alpha_{im})$. By the product formula we get $\sum_{v\in V} \beta_{ijv} = 0$ $(1 \leqslant i \leqslant r, 1 \leqslant j \leqslant m)$. Let S be the subset of V consisting of the Archimedean places of K and of v's with $\beta_{ijv} \neq 0$ for some i, j $(1 \leqslant i \leqslant r, 1 \leqslant j \leqslant m)$. Then also $\sum_{v\in S} \beta_{ijv} = 0$.

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

For $\boldsymbol{\xi} \in \mathbb{R}^r$ we define

$$g_{jv}(\boldsymbol{\xi}) = \sum_{i=1}^{r} \beta_{ijv}\xi_i \quad (1 \leqslant j \leqslant m, v \in V). \tag{7.8}$$

Then again $\sum_{v\in S} g_{jv}(\boldsymbol{\xi}) = 0$ $(1 \leqslant j \leqslant m)$ and $g_{jv}(\boldsymbol{\xi}) = 0$ for $v \notin S$, $j = 1, \ldots, m$.

As

$$\log \|\alpha_{1j}^{u_1} \ldots \alpha_{rj}^{u_r}\|_v = \sum_{i=1}^{r} \beta_{ijv} u_i = g_{jv}(\mathbf{u}),$$

we obtain from (7.1) and (7.7)

$$\psi(\mathbf{u}) = \frac{1}{2}\sum_{j=1}^{m}\sum_{v\in V} |g_{jv}(\mathbf{u})| = \frac{1}{2}\sum_{j=1}^{m}\sum_{v\in S} |g_{jv}(\mathbf{u})|. \tag{7.9}$$

More generally, for $\boldsymbol{\xi} \in \mathbb{R}^r$ we put

$$\psi(\boldsymbol{\xi}) = \frac{1}{2}\sum_{v\in V}\sum_{j=1}^{m} |g_{jv}(\boldsymbol{\xi})|. \tag{7.10}$$

It was shown in [19] (Section 3) that ψ is a distance function in the sense of Cassels [2] (Chapter IV) and that the set

$$\Psi = \{\boldsymbol{\xi} \in \mathbb{R}^r | \psi(\boldsymbol{\xi}) \leqslant 1\} \tag{7.11}$$

is a symmetric, convex body.

8. Special Solutions

Let $K, \Gamma, \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ be as in Section 7. Put

$$q = 8m + 4. \tag{8.1}$$

When $\mathbf{x} \in \Gamma$, set

$$h = h(\mathbf{x}), \qquad H = H(\mathbf{x}) = e^h. \tag{8.2}$$

Express $\mathbf{x}$ as in (7.6). Given $\boldsymbol{\rho} \in \mathbb{R}^r$, an element $\mathbf{x} \in \Gamma$ will be called $\boldsymbol{\rho}$-*special* if $h > 0$ and

$$\mathbf{u} \in (h/q)\Psi + h\boldsymbol{\rho}, \tag{8.3}$$

where Ψ is the set $\{\boldsymbol{\xi}|\psi(\boldsymbol{\xi}) \leqslant 1\}$ from (7.11); the right-hand side of (8.3) signifies $(h/q)\Psi$ translated by $h\boldsymbol{\rho}$.

Let Φ be a symmetric convex body in $\mathbb{R}^r$. Let $\Delta = \Delta(\Phi)$ be the least covering density of $\mathbb{R}^r$ by translates of Φ (not necessarily by points of a lattice). Thus Δ is least such that there are $\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \ldots$ in $\mathbb{R}^r$ such that the union of the translates $\Phi + \boldsymbol{\rho}_i$ $(i = 1, 2, \ldots)$ is $\mathbb{R}^r$, and if $\nu(\boldsymbol{\xi})$ is the number of translates which contain $\boldsymbol{\xi}$, then

$$\int_{|\boldsymbol{\xi}|\leqslant X} \nu(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} \Bigg/ \int_{|\boldsymbol{\xi}|\leqslant X} \mathrm{d}\boldsymbol{\xi} < \Delta(1+\varepsilon) \tag{8.4}$$

when $\varepsilon > 0$ and $X > X_0(\varepsilon)$. Here $|\boldsymbol{\xi}|$ denotes the maximum norm, say. Let $\Delta(r)$ be the supremum of the covering densities of symmetric convex bodies in $\mathbb{R}^r$. It is relatively easy ([11]) to show that

$$\Delta(r) \leqslant 2^r. \tag{8.5}$$

Better bounds are known (cf. [12]), but (8.5) will do for us.

LEMMA 8.1. *Let Φ be a symmetric convex body in $\mathbb{R}^r$. Suppose $\lambda > 0$. Then $\lambda\Phi$ can be covered by not more than $(\lambda+2)^r\Delta(r) \leqslant (2\lambda+4)^r$ translates of Φ.*

Proof. In view of (8.4), it is not hard to see that there is a translate of $(\lambda+2)\Phi$, say $(\lambda+2)\Phi + \boldsymbol{\tau}$, such that

$$\int_{(\lambda+2)\Phi+\boldsymbol{\tau}} \nu(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}/(\lambda+2)^r V(\Phi) < \Delta(r)(1+2\varepsilon), \tag{8.6}$$

where $V(\Phi)$ is the volume of Φ (so that $(\lambda+2)^r V(\Phi)$ is the volume of $(\lambda+2)\Phi + \boldsymbol{\tau}$). Then (replace the $\boldsymbol{\rho}_i$ by $\boldsymbol{\rho}_i - \boldsymbol{\tau}$ $(i = 1, 2, \ldots)$), there is also a covering such that (8.6) is true with $\boldsymbol{\tau} = \mathbf{0}$. Now if Z of the translates $\Phi + \boldsymbol{\rho}_i$ intersect $\lambda\Phi$, then these are contained in $(\lambda+2)\Phi$, so that

$$\int_{(\lambda+2)\Phi} \nu(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} \geqslant ZV(\Phi).$$

Comparison with (8.6) yields

$$Z < (\lambda+2)^r\Delta(r)(1+2\varepsilon). \tag{8.7}$$

For every $\varepsilon > 0$ there is a covering of $\lambda\Phi$ by Z translates of Φ with Z satisfying (8.7). The lemma is now obvious.

Applying Lemma 8.1 with $\Phi = m\Psi$, $\lambda = 1/(mq)$, we may conclude that $m\Psi$ may be covered by $Z = (2mq+4)^r$ translates of $q^{-1}\Psi$, say by $q^{-1}\Psi + \boldsymbol{\rho}_i$ $(i = 1, \ldots, Z)$. Then $hm\Psi$ is covered by $(h/q)m\Psi + h\boldsymbol{\rho}_i$ $(i = 1, \ldots, Z)$. When

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

(8.2) holds, then $\mathbf{u}$ as in (7.6) lies in $h_s(\mathbf{x})\Psi \subseteq mh\Psi$ (by (7.3)). Thus $\mathbf{x}$ is special for at least one of $\boldsymbol{\rho}_1, \ldots, \boldsymbol{\rho}_Z$. With our value q as in (8.7) we obtain

$$Z = (16m^2 + 8m + 4)^r \leqslant (21m^2)^r, \tag{8.8}$$

since $m \geqslant 2$. By (8.3), when $\mathbf{u} \in hm\Psi$, then $\boldsymbol{\rho} \in (m + q^{-1})\Psi$. Hence we may take

$$\boldsymbol{\rho}_1, \ldots, \boldsymbol{\rho}_Z \in (m + q^{-1})\Psi.$$

9. Properties of Special Solutions

Let $\boldsymbol{\rho} \in (m + q^{-1})\Psi$ be fixed, where $m \geqslant 2$. Set

$$m_{jv} = \begin{cases} g_{iv}(\boldsymbol{\rho}), & \text{if } v \in V, 1 \leqslant j \leqslant m, \\ 0, & \text{if } v \in V, j = 0. \end{cases} \tag{9.1}$$

Then, as was seen below (7.8), we have

$$\sum_{v \in V} m_{jv} = 0 \quad (j = 0, 1, \ldots, m). \tag{9.2}$$

By the definitions (7.10), (7.11) of ψ, Ψ and by (9.1),

$$\sum_{v \in V} \sum_{j=0}^{m} |m_{jv}| \leqslant 2\psi(\boldsymbol{\rho}) \leqslant 2(m + q^{-1}). \tag{9.3}$$

Let $L_0, \ldots, L_m$ be the linear forms in $\mathbf{X} = (X_1, \ldots, X_m)$ defined by

$$\begin{aligned} L_0(\mathbf{X}) &= X_1 + \cdots + X_m, \\ L_j(\mathbf{X}) &= X_j \quad (j = 1, \ldots, m). \end{aligned} \tag{9.4}$$

Suppose now we have a solution $\mathbf{z} = \mathbf{x} * \mathbf{y}$ of (3.1) where $\mathbf{x} \in \Gamma$, $\mathbf{y} \in (\mathbb{Q}^\times)^m$ and where (3.2) holds. Write $y_j = w_j/w_0$ with $w_0, \ldots, w_m \in \mathbb{Z}$ and g.c.d. $(w_0, \ldots, w_m) = 1$. Then (3.1) may be rewritten as

$$z'_1 + \cdots + z'_m = z'_0, \tag{9.5}$$

where $z'_0 = w_0$ and $z'_j = x_j \cdot w_j$ for $j = 1, \ldots, m$. We write $\mathbf{z}' = (z'_1, \ldots, z'_m)$. Recall the definition of S in Section 7.

LEMMA 9.1. *Let $\boldsymbol{\rho}$ be as above. Then there are m-element subsets $\mathcal{I}(v)$ of $\{0, 1, \ldots, m\}$ defined for $v \in V$, and there are numbers ℓ_{jv} $(v \in V, j \in \mathcal{I}(v))$ with the following properties.*

$$\mathcal{I}(v) = \{1, \ldots, m\} \quad \textit{for } v \notin S, \tag{9.6}$$

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

$$\ell_{jv} = 0 \quad \textit{for} \ \ v \notin S,\, j \in \mathcal{I}(v), \tag{9.7}$$

$$\sum_{v \in V} \sum_{j \in \mathcal{I}(v)} \ell_{jv} = 0, \quad \sum_{v \in V} \sum_{j \in \mathcal{I}(v)} |\ell_{jv}| \leqslant 1. \tag{9.8}$$

Moreover, if $\mathbf{z} = \mathbf{x} * \mathbf{y}$ *and* $\mathbf{z}'$ *are as above, where* $\mathbf{x}$ *is* $\boldsymbol{\rho}$*-special, then*

$$\prod_{v \in V} \max_{j \in \mathcal{I}(v)} \{\|L_j(\mathbf{z}')\|_v Q^{-\ell_{jv}}\} \leqslant Q^{-1/(4m^2+2m)}, \tag{9.9}$$

with $Q = H(\mathbf{x})^{2m+1}$.

Proof. We define $\mathcal{I}(v)$ as follows. For $v \in S$ we set $\mathcal{I}(v) = \{1, \ldots, m\}$ according to (9.6). For $v \in S$ we consider the elements m_{jv} from (9.1). Pick $j(v) \in \{0, \ldots, m\}$ such that

$$m_{j(v),v} = \max(m_{0v}, \ldots, m_{mv}). \tag{9.10}$$

and set

$$\mathcal{I}(v) = \{0, \ldots, m\} \backslash \{j(v)\} \quad (v \in S).$$

Now let $\mathbf{x} \in \Gamma$ be $\boldsymbol{\rho}$-special. Then with $\mathbf{u}$ as in (7.6) we have (8.3) with $h = h(\mathbf{x})$. So

$$g_{jv}(\mathbf{u}) = h(g_{jv}(\boldsymbol{\rho}) + q^{-1} g_{jv}(\boldsymbol{\xi})) = hm_{jv} + (h/q) g_{jv}(\boldsymbol{\xi}), \tag{9.11}$$

for a suitable $\boldsymbol{\xi} \in \Psi$ and for $v \in V$, $j = 1, \ldots, m$. If we put $g_{0v}(\boldsymbol{\xi}) = 0$ for $v \in V$ and $\boldsymbol{\xi} \in \mathbb{R}^r$, then (9.11) will be true for $j = 0$ as well. Since $\boldsymbol{\xi} \in \Psi$ we have $\sum_v |g_{jv}(\boldsymbol{\xi})| \leqslant 2$, and therefore

$$\sum_{v \in S} \sum_{j=0}^{m} |hm_{jv} - g_{jv}(\mathbf{u})| \leqslant 2h/q \quad (j = 0, \ldots, m). \tag{9.12}$$

By our definition of S in Section 7, and by (7.6), (7.8), any $\mathbf{x} \in \Gamma$ has

$$\begin{aligned} h(\mathbf{x}) &= \sum_{v \in S} \max(0, \log \|x_1\|_v, \ldots, \log \|x_m\|_v) \\ &= \sum_{v \in S} \max(g_{0v}(\mathbf{u}), \ldots, g_{mv}(\mathbf{u})). \end{aligned}$$

Thus by (9.10) and (9.12),

$$h \sum_{v \in S} m_{j(v),v} \geqslant h(\mathbf{x}) - (2h/q) = h(1 - 2/q),$$

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

so that

$$\sum_{v\in S} m_{j(v),v} \geqslant 1-2/q. \tag{9.13}$$

This estimate holds if there exists any $\boldsymbol{\rho}$-special point $\mathbf{x}\in\Gamma$.

Let s be the cardinality of S and write

$$\gamma = \frac{1}{ms}\sum_{v\in S} m_{j(v),v}. \tag{9.14}$$

We define numbers c_{jv} ($v\in V$, $j\in\mathcal{I}(v)$) by

$$c_{jv} = \begin{cases} m_{jv}+\gamma, & \text{if } v\in S,\ j\in\mathcal{I}(v),\\ m_{jv}(=0), & \text{if } v\notin S,\ j\in\mathcal{I}(v). \end{cases} \tag{9.15}$$

We infer from (9.2), (9.3) that

$$\sum_{v\in V}\sum_{j\in\mathcal{I}(v)} c_{jv} = 0, \qquad \sum_{v\in V}\sum_{j\in\mathcal{I}(v)} |c_{jv}| \leqslant 2(m+q^{-1}). \tag{9.16}$$

Observe that for $j=1,\dots,m$,

$$\begin{aligned}\log\|x_j\|_v &= g_{jv}(\mathbf{u}) = h(g_{jv}(\boldsymbol{\rho})+g_{jv}(\boldsymbol{\xi})/q)\\ &= h(m_{jv}+g_{jv}(\boldsymbol{\xi})/q),\end{aligned} \tag{9.17}$$

by (9.11), (9.1). But

$$\log|w_j| \leqslant h(w_0{:}\cdots{:}w_m) = h(\mathbf{y}) \leqslant h/4m^2 \quad (j=0,\dots,m)$$

by (3.2), so that by definition of z'_j and by (9.5), (9.17),

$$\log\|L_j(\mathbf{z}')\|_v = \log\|z'_j\|_v \leqslant h(m_{jv}+g_{jv}(\boldsymbol{\xi})/q+\delta_v/4m^2),$$

where $\delta_v = d_v/d$ when $v\in V_\infty$ (the set of Archimedean places), and $\delta_v=0$ otherwise. Since $z'_0=w_0$ and since $m_{0v}=g_{0v}(\boldsymbol{\xi})=0$, this inequality holds for $j=0,\dots,m$. When $j\in\mathcal{I}(v)$ we have by the definition (9.15) of the c_{jv} that

$$\log\|L_j(\mathbf{z}')\|_v - hc_{jv} \leqslant h(g_{jv}(\boldsymbol{\xi})/q+\delta_v/4m^2-\eta_v\gamma),$$

where $\eta_v=1$ when $v\in S$, and $\eta_v=0$ otherwise. We note that

$$\sum_{v\in V}\max_j |g_{jv}(\boldsymbol{\xi})| \leqslant 2\psi(\boldsymbol{\xi}) \leqslant 2,$$

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

since $\boldsymbol{\xi} \in \Psi$, and therefore

$$\sum_{v \in V} \max_{j \in \mathcal{I}(v)} (\log \|L_j(\mathbf{z}')\|_v - hc_{jv})$$
$$\leqslant h((2/q) + (1/4m^2) - \gamma s)$$
$$\leqslant h((1/4m) + (1/4m^2) - (1/m)(1 - 1/8m) \leqslant -h/2m,$$

by (8.1), (9.13), (9.14), and since $m \geqslant 2$. Exponentiating, we obtain

$$\prod_{v \in V} \max_{j \in \mathcal{I}(v)} \{\|L_j(\mathbf{z}')\|_v H^{-c_{jv}}\} \leqslant H^{-1/2m}. \tag{9.18}$$

We now renormalize using the quantity $Q = H(\mathbf{x})^{2m+1}$. We define

$$\ell_{jv} = c_{jv}/(2m+1) \quad (v \in V, j \in \mathcal{I}(v)).$$

Then (9.7), (9.8) hold as a consequence of (9.15), (9.16), and (9.9) holds by virtue of (9.18).

10. Large Solutions

We quote a very special case of a theorem of Evertse and Schlickewei [6].

PROPOSITION D. *Suppose* $0 < \delta < 1$, *and let* L_j *be the linear forms of* (9.4). *For* $v \in V$ *let* $\mathcal{I}(v)$ *be as in Lemma* 9.1, *and let* ℓ_{jv} *(*$v \in V, j \in \mathcal{I}(v)$*) be as in* (9.7), (9.8). *Then there are proper linear subspaces* $T_1, \ldots, T_t$ *of* K^m *with*

$$t \leqslant 2^{2(m+5)^2} \delta^{-m-4}, \tag{10.1}$$

such that every $\mathbf{z} \in K^m$ *having*

$$\prod_{v \in V} \max_{j \in \mathcal{I}(v)} \{\|L_j(\mathbf{z})\|_v Q^{-\ell_{jv}}\} \leqslant Q^{-\delta/m}, \tag{10.2}$$

for some

$$Q > m^{m/\delta} \tag{10.3}$$

lies in the union of $T_1, \ldots, T_t$.

When we are dealing with a large solution of (3.1), then $h > 2m \log m$ by the definition (6.1), so that $Q = H^{2m+1}$ satisfies (10.3) with $\delta = 1/(4m+2)$. By Lemma 9.1, a point $\mathbf{z}'$ arising from a large special solution satisfies the hypotheses

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

of Proposition D. With our value of δ, we see that the large $\boldsymbol{\rho}$-special solutions will have $\mathbf{z}'$ contained in not more than

$$2^{2(m+5)^2}(4m+2)^{m+4} < 2^{49m^2/2}(5m)^{3m} < 2^{30m^2}$$

proper linear subspaces of K^m, since $m \geqslant 2$. As $\mathbf{z}'$ is proportional to $\mathbf{z}$ in (3.1), also $\mathbf{z}$ will be in the union of these subspaces.

Allowing a factor $(21m^2)^r$ from (8.8) for the number of points $\boldsymbol{\rho}$ needed, we may conclude that the set of large solutions of (3.1) may be covered by

$$< 2^{30m^2}(21m^2)^r \tag{10.4}$$

proper subspaces.

11. Proof of Proposition A

Let us recall that in the preceding sections, according to the convention adopted at the beginning of Section 6, we had restricted ourselves to solutions $\mathbf{z} = (z_1, \ldots, z_m)$ with $z_1 \ldots z_m \neq 0$. Clearly all the other solutions may be covered by the m coordinate subspaces $z_i = 0$ $(i = 1, \ldots, m)$. It will suffice to combined this bound with the bounds from Lemma 6.1 for the small solutions and the bound (10.4) for the large solutions. Altogether we need fewer than

$$m + (4d^2)^m(86d^3 m \log m)^r + 2^{30m^2}(21m^2)^r < 2^{30m^2}(32m^2)^r d^{3r+2m}$$

III. PROOF OF PROPOSITION C

It remains for us to deduce Proposition C from Proposition B. The main difficulty will be to satisfy condition (4.7) of Proposition B. A priori, it would seem that the height $h_M(\mathbf{x})$ of the vector (4.5) of monomials should be much smaller than the height $h_E(\mathbf{x})$ of the vector (4.6) of exponentials. But lacking information on the bases $\boldsymbol{\alpha}_\ell$ of these exponentials, condition (4.7) is difficult to enforce.

12. Minimal Forms

Recall that K^A is the space of vectors $\boldsymbol{\xi} = (\xi_{\ell M})$ where $\ell \in \Lambda = \{1, \ldots, k\}$ and $M \in \mathbf{M}_\ell$, i.e., the set of monomials of total degree $\leqslant \delta_\ell$. Every linear form L on K^A may be written as

$$L(\boldsymbol{\xi}) = L^1(\boldsymbol{\xi}_1) + \cdots + L^k(\boldsymbol{\xi}_k), \tag{12.1}$$

where $\boldsymbol{\xi} = (\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_k)$ and $\boldsymbol{\xi}_\ell = (\xi_{\ell M})$ with $M \in \mathbf{M}_\ell$ and where L^ℓ is a linear form on a space of dimension $|\mathbf{M}_\ell| = \operatorname{card} \mathbf{M}_\ell$ $(\ell = 1, \ldots, k)$. In fact $L^\ell(\boldsymbol{\xi}_\ell) = \sum_{M \in \mathbf{M}_\ell} b_{\ell M} \xi_{\ell M}$ with coefficients $b_{\ell M} \in K$. Write $\mathcal{B}(L)$ for the set of $\ell \in \Lambda$ with

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

$L^\ell \neq 0$. Write $\mathcal{A}(L)$ for the set of pairs (ℓ, M) with $b_{\ell M} \neq 0$. Thus $\mathcal{B}(L)$ consists of $\ell \in \Lambda$ for which there is an $M \in \mathbf{M}_\ell$ with $(\ell, M) \in \mathcal{A}(L)$. We call $\mathcal{A}(L)$ the *support* of L.

Let T be the subspace of K^A of Proposition C and $\mathcal{L}(T)$ the space of linear forms vanishing on T. If we had $\mathcal{L}(T) = \{0\}$, then $T = K^A$, so that $\mathcal{P} \prec T$ would imply that $\mathcal{P}$ is the partition into singletons $\{1\}, \ldots, \{k\}$, which is incompatible with $G(\mathcal{P}) = \{\mathbf{0}\}$. Therefore $\mathcal{L}(T) \neq \{0\}$.

A form $L \neq 0$ in $\mathcal{L}(T)$ will be called a *minimal form* if there is no nonzero form L' in $\mathcal{L}(T)$ with $\mathcal{A}(L')$ a proper subset of $\mathcal{A}(L)$. Since $\mathcal{P} \prec T$, a minimal form L has $\mathcal{B}(L) \subset \lambda$ for some $\lambda \in \mathcal{P}$. Say the minimal form is

$$L = \sum_{(\ell, M) \in \mathcal{A}(L)} b_{\ell M} \xi_{\ell M}. \tag{12.2}$$

When $\mathbf{x} \in \mathcal{X}(T)$, then $\mathcal{L}(\boldsymbol{\xi}(\mathbf{x})) = 0$ where $\boldsymbol{\xi}(\mathbf{x})$ is the vector having components $\xi_{\ell M}(\mathbf{x}) = M(\mathbf{x})\boldsymbol{\alpha}_\ell^{\mathbf{x}}$ with $\ell \in \Lambda$, $M \in \mathbf{M}_\ell$. Let us restrict to the vector $\boldsymbol{\xi}_L(\mathbf{x})$ with components $\xi_{\ell M}(\mathbf{x})$ where $(\ell, M) \in \mathcal{A} = \mathcal{A}(L)$. By a slight abuse of notation

$$L(\boldsymbol{\xi}_L(\mathbf{x})) = 0. \tag{12.3}$$

Here $\boldsymbol{\xi}_L(\mathbf{x}) \in K^a$ with $a = a_L = |\mathcal{A}(L)|$, and (12.3) says that $\boldsymbol{\xi}(\mathbf{x})$ lies in a subspace $U_L \subset K^a$ of codimension 1. The idea will be to show via Proposition B that when $\mathbf{x}$ lies outside an exceptional set of C elements, the set of solutions $\boldsymbol{\xi}_L(\mathbf{x})$ lies in a number of proper subspaces of U_L, say $U_{L1}, \ldots, U_{LC}$. Now U_{Li} is given by $L_i(\boldsymbol{\xi}_L) = 0$ for a linear form L_i which is, of course, not proportional to L. Since $\mathcal{A}(L_i) \subseteq \mathcal{A}(L)$, we may replace L_i by $L_i' = L_i - \alpha_i L$ with suitable α_i in such a way that $\mathcal{A}(L_i')$ is a proper subset of $\mathcal{A}(L)$. In other words, we may suppose that $\mathcal{A}(L_i)$ is a proper subset of $\mathcal{A}(L)$. By the minimality property of L, we have $L_i \notin \mathcal{L}(T)$, and therefore when $\xi_L(\mathbf{x}) \in U_{Li}$, then $\boldsymbol{\xi}_L(\mathbf{x})$ lies in a proper subspace T_i of T. Moreover, since $\mathcal{B}(L_i) \subset \mathcal{B}(L) \subset \lambda$ for some $\lambda \in \mathcal{P}$, we have $\mathcal{P} \prec T_i$.

The plan, then, will be to apply Proposition B to a minimal form L. At least one of the resulting subspaces T_i among $T_1, \ldots, T_C$ will have $|\mathcal{X}(T, \mathcal{P})| - C \leqslant C|\mathcal{X}(T, \mathcal{P}) \cap \mathcal{X}(T_i)|$, and Proposition C will follow with $T' = T_i$.

Note that the minimality of forms may be destroyed by the transformations of Sections 13 and 15 below, and that useful minimal forms will only be constructed in Section 16.

13. The Initial Transformation

Given a vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in (K^\times)^n$, set $\ell_{jv} = \log \|\alpha_j\|_v$ $(1 \leqslant j \leqslant n, v \in V = V(K))$. Then $\sum_v \ell_{jv} = 0$ $(1 \leqslant j \leqslant n)$ by the product formula; here and below, a sum over v, unless indicated otherwise, is over $v \in V$. For $\boldsymbol{\xi} \in \mathbb{R}^n$ set

$$g_v(\boldsymbol{\xi}) = \sum_{j=1}^{n} \ell_{jv} \xi_j, \tag{13.1}$$

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

so that g_v is a linear form. We have $\sum_v g_v(\boldsymbol{\xi}) = 0$. Put

$$\psi(\boldsymbol{\xi}) = \sum_v \max(0, g_v(\boldsymbol{\xi})) = \frac{1}{2}\sum_v |g_v(\boldsymbol{\xi})|. \tag{13.2}$$

Note that

$$\psi(\boldsymbol{\xi}+\boldsymbol{\eta}) \leqslant \psi(\boldsymbol{\xi}) + \psi(\boldsymbol{\eta}), \qquad \psi(\gamma\boldsymbol{\xi}) = |\gamma|\psi(\boldsymbol{\xi}), \tag{13.3}$$

for $\gamma \in \mathbb{R}$. Since $\log \|\boldsymbol{\alpha}^{\mathbf{x}}\|_v = g_v(\mathbf{x})$, we see from (7.1) that

$$\psi(\mathbf{x}) = h(\boldsymbol{\alpha}^{\mathbf{x}}), \tag{13.4}$$

for $\mathbf{x} \in \mathbb{Z}^n$.

Given $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_k$ as in our theorem, define $\boldsymbol{\alpha}_\ell/\boldsymbol{\alpha}_m$ in analogy to (4.9). Define $\psi_{\ell m}(\boldsymbol{\xi})$ as $\psi(\boldsymbol{\xi})$ above, but with $\boldsymbol{\alpha} = \boldsymbol{\alpha}_\ell/\boldsymbol{\alpha}_m$. Then

$$\psi_{\ell m}(\mathbf{x}) = h((\boldsymbol{\alpha}_\ell/\boldsymbol{\alpha}_m)^{\mathbf{x}}) = h(\boldsymbol{\alpha}_\ell^{\mathbf{x}}{:}\boldsymbol{\alpha}_m^{\mathbf{x}}),$$

for $\mathbf{x} \in \mathbb{Z}^n$. Given a subset λ of $\Lambda = \{1, \dots, k\}$, put

$$h^\lambda(\mathbf{x}) = \max_{\ell,m\in\lambda} h(\boldsymbol{\alpha}_\ell^{\mathbf{x}}{:}\boldsymbol{\alpha}_m^{\mathbf{x}}), \qquad \omega^\lambda(\boldsymbol{\xi}) = \max_{\ell,m\in\lambda} \psi_{\ell m}(\boldsymbol{\xi}).$$

Given a partition $\mathcal{P}$ of Λ, write

$$h^{\mathcal{P}}(\mathbf{x}) = \max_{\lambda\in\mathcal{P}} h^\lambda(\mathbf{x}), \qquad \omega^{\mathcal{P}}(\boldsymbol{\xi}) = \max_{\lambda\in\mathcal{P}} \omega^\lambda(\boldsymbol{\xi}).$$

The maximum of several functions with (13.3) still has this property, and therefore

$$\omega^{\mathcal{P}}(\boldsymbol{\xi}+\boldsymbol{\eta}) \leqslant \omega^{\mathcal{P}}(\boldsymbol{\xi}) + \omega^{\mathcal{P}}(\boldsymbol{\eta}), \qquad \omega^{\mathcal{P}}(\gamma\boldsymbol{\xi}) = |\gamma|\omega^{\mathcal{P}}(\boldsymbol{\xi}). \tag{13.5}$$

Clearly

$$\omega^{\mathcal{P}}(\mathbf{x}) = h^{\mathcal{P}}(\mathbf{x}), \tag{13.6}$$

for $\mathbf{x} \in \mathbb{Z}^n$.

Now suppose that the group $G(\mathcal{P}) = \{\mathbf{0}\}$. Then when $\mathbf{x} \in \mathbb{Z}^n \backslash \{\mathbf{0}\}$, there are ℓ, m with $\ell \overset{\mathcal{P}}{\sim} m$ and $\boldsymbol{\alpha}_\ell^{\mathbf{x}} \neq \boldsymbol{\alpha}_m^{\mathbf{x}}$, hence with $(\boldsymbol{\alpha}_\ell/\boldsymbol{\alpha}_m)^{\mathbf{x}} \neq 1$. In fact there is such a pair ℓ, m for which $(\boldsymbol{\alpha}_\ell/\boldsymbol{\alpha}_m)^{\mathbf{x}}$ is not a root of 1. Then according to (7.2), $h((\boldsymbol{\alpha}_\ell/\boldsymbol{\alpha}_m)^{\mathbf{x}}) > 1/(21d^3)$. We may conclude that for $\mathbf{x} \in \mathbb{Z}^n \backslash \{\mathbf{0}\}$,

$$\omega^{\mathcal{P}}(\mathbf{x}) > 1/(21d^3). \tag{13.7}$$

In view of (13.5), (13.7), the function $\omega^{\mathcal{P}}$ is a Minkowski distance in $\mathbb{R}^n$ (see [19, Lemma 3]), i.e., the set Ω of $\boldsymbol{\xi} \in \mathbb{R}^n$ with $\omega^{\mathcal{P}}(\boldsymbol{\xi}) \leqslant 1$ is convex, symmetric (that is, $\boldsymbol{\xi} \in \Omega$ implies $-\boldsymbol{\xi} \in \Omega$), compact, and contains $\mathbf{0}$ in its interior.

Since $\mathcal{P}$ will be fixed, we will write more briefly ω for $\omega^{\mathcal{P}}$. By a theorem of Schlickewei [15], there is a basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ of $\mathbb{Z}^n$ such that

$$\omega(\xi_1\mathbf{b}_1 + \cdots + \xi_n\mathbf{b}_n) \geqslant 4^{-n} \max_{1\leqslant i\leqslant n} |\xi_i|\omega(\mathbf{b}_i),$$

and in view of (13.7) this is

$$\geqslant (4^n \cdot 21d^3)^{-1}|\boldsymbol{\xi}|,$$

where $|\boldsymbol{\xi}|$ denotes the maximum norm. In other words, there is a transformation $\tau \in \mathrm{GL}(n, \mathbb{Z})$ such that

$$\omega(\tau(\boldsymbol{\xi})) \geqslant c_1|\boldsymbol{\xi}|,$$

with

$$c_1 = (4^n \cdot 21d^3)^{-1}. \tag{13.8}$$

Now

$$\sum_{\ell=1}^{k} P_\ell(\tau(\mathbf{x}))\boldsymbol{\alpha}_\ell^{\tau(\mathbf{x})} = \sum_{\ell=1}^{k} \widehat{P}_\ell(\mathbf{x})\boldsymbol{\beta}^{\mathbf{x}}, \tag{13.9}$$

where $\widehat{P}_\ell(\mathbf{x}) = P_\ell(\tau(\mathbf{x}))$ is a polynomial of the same total degree as P_ℓ, and where $\boldsymbol{\beta}_\ell = (\boldsymbol{\alpha}_\ell^{\tau(\mathbf{e}_1)}, \ldots, \boldsymbol{\alpha}_\ell^{\tau(\mathbf{e}_n)})$ with $\mathbf{e}_1, \ldots, \mathbf{e}_n$ the standard basis of $\mathbb{Z}^n$. Our ω was defined in terms of $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k$; write $\omega = \omega_\alpha$. Similarly define ω_β in terms of $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_k$. Then

$$\omega_\beta(\boldsymbol{\xi}) = \omega_\alpha(\tau(\boldsymbol{\xi})) \geqslant c_1|\boldsymbol{\xi}|.$$

As is suggested by (13.9), and as was explained in detail in [16, §7], we may apply a substitution τ. Therefore we may suppose from now on that

$$\omega(\boldsymbol{\xi}) \geqslant c_1|\boldsymbol{\xi}|. \tag{13.10}$$

This is essentially [16, (7.8)], except that we went to the logarithm, and that we have a better value for c_1.

14. Producing Large Heights (i)

Let $L \in \mathcal{L}(T)$ be minimal, and write it as in (12.2). Set $\eta_{\ell M}(\mathbf{x}) = b_{\ell M} M(\mathbf{x})\boldsymbol{\alpha}_\ell^{\mathbf{x}} = b_{\ell M}\xi_{\ell M}(\mathbf{x})$, and let $\boldsymbol{\eta}(\mathbf{x})$ be the vector in K^a (where $a = |\mathcal{A}(L)|$) having components $\eta_{\ell M}(\mathbf{x})$ with $(\ell, M) \in \mathcal{A}(L)$. Then (12.3) is the same as (4.4), and $\boldsymbol{\eta}_L(\mathbf{x})$ lies in a subspace $U_L' \subseteq K^a$ of codimension 1. It will suffice to show that $\boldsymbol{\eta}_L(\mathbf{x})$ lies in

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

the union of proper subspaces $U'_{L1}, \dots, U'_{LC}$ of U'_L, for then $\boldsymbol{\xi}_L(\mathbf{x})$ will lie in the union of proper subspaces $U_{L1}, \dots, U_{LC}$ of U_L.

To apply Proposition B, we will need (4.7). So let $h_{LE}(\mathbf{x})$ be defined as in Section 4, i.e., as the height $h(\boldsymbol{\varepsilon}_L(\mathbf{x}))$ of the vector

$$\boldsymbol{\varepsilon}_L(\mathbf{x}) = \{b_{\ell M}\boldsymbol{\alpha}_\ell^{\mathbf{x}}\}_{(\ell,M)\in\mathcal{A}(L)}. \tag{14.1}$$

We need forms L with $h_{LE}(\mathbf{x})$ large. This we cannot do at once; we will first have to deal with the height $h_{LD}(\mathbf{x}) = h(\boldsymbol{\delta}_L(\mathbf{x}))$ where

$$\boldsymbol{\delta}_L(\mathbf{x}) = \{\boldsymbol{\alpha}_\ell^{\mathbf{x}}\}_{\ell\in\mathcal{B}(L)}. \tag{14.2}$$

LEMMA 14.1. *Suppose for each* $\lambda \in \mathcal{P}$ *we have forms* $L_{\lambda j}$ $(j = 1, \dots, t(\lambda)$ *where* $t(\lambda) \leqslant |\lambda| \leqslant k)$ *with* $\mathcal{B}(L_{\lambda j}) \subseteq \lambda$, *and such that for any* ℓ, m *in* λ, *there is a chain of forms* $L_{\lambda,j(1)}, \dots, L_{\lambda,j(q)}$ *with* $q \leqslant t(\lambda)$ *and* $\ell \in \mathcal{B}(L_{\lambda,j(1)})$, $m \in \mathcal{B}(L_{\lambda,j(q)})$ *having*

$$\mathcal{B}(L_{\lambda,j(i)}) \cap \mathcal{B}(L_{\lambda,j(i+1)}) \neq \emptyset, \tag{14.3}$$

for $1 \leqslant i < q$. *Then*

$$\max_{\lambda,j} h_{L_{\lambda,j}D}(\mathbf{x}) \geqslant c_2|\mathbf{x}| \tag{14.4}$$

with

$$c_2 = c_1/k = (21kd^3 \cdot 4^n)^{-1}. \tag{14.5}$$

Proof. By (13.6), (13.10) we have $h^{\mathcal{P}}(\mathbf{x}) \geqslant c_1|\mathbf{x}|$, therefore $h^\lambda(\mathbf{x}) \geqslant c_1|\mathbf{x}|$ for some $\lambda \in \mathcal{P}$, and then $h(\boldsymbol{\alpha}_\ell^{\mathbf{x}}:\boldsymbol{\alpha}_m^{\mathbf{x}}) \geqslant c_1|\mathbf{x}|$, for some ℓ, m in λ. Let $L_{\lambda,j(1)}, \dots, L_{\lambda,j(q)}$ be as above, and let ℓ_i be in the set (14.3). Then (since generally $h(\alpha:\gamma) \leqslant h(\alpha:\beta) + h(\beta:\gamma)$),

$$\begin{aligned} c_1|\mathbf{x}| &\leqslant h(\boldsymbol{\alpha}_\ell^{\mathbf{x}}:\boldsymbol{\alpha}_m^{\mathbf{x}}) \\ &\leqslant h(\boldsymbol{\alpha}_\ell^{\mathbf{x}}:\boldsymbol{\alpha}_{\ell_1}^{\mathbf{x}}) + h(\boldsymbol{\alpha}_{\ell_1}^{\mathbf{x}}:\boldsymbol{\alpha}_{\ell_2}^{\mathbf{x}}) + \cdots + h(\boldsymbol{\alpha}_{q-1}^{\mathbf{x}}:\boldsymbol{\alpha}_m^{\mathbf{x}}) \\ &\leqslant h_{L_{\lambda,j(1)}}(\mathbf{x}) + \cdots + h_{L_{\lambda,j(q)}}(\mathbf{x}). \end{aligned}$$

The assertion follows.

15. Producing Large Heights (ii)

Given $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n) \in (K^\times)^n$, we have defined $g_v(\boldsymbol{\xi})$, $\psi\boldsymbol{\xi})$ by (13.1), (13.2). Now let vectors $\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_r$ in $(K^\times)^n$ be given, and define $g_{iv}(\boldsymbol{\xi})$, $\psi_i(\boldsymbol{\xi})$ as above

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

but with $\boldsymbol{\alpha} = \boldsymbol{\beta}_i$ $(i = 1, \dots, r)$. Set $\chi(\boldsymbol{\xi}) = \max_{1 \leqslant i \leqslant r} \psi_i(\boldsymbol{\xi})$. We make the extra hypothesis that

$$\chi(\boldsymbol{\xi}) \geqslant c_2 |\boldsymbol{\xi}|, \tag{15.1}$$

for $\boldsymbol{\xi} \in \mathbb{R}^n$. Then χ is a Minkowski distance on $\mathbb{R}^n$. Let $\mathcal{X}$ consist of $\boldsymbol{\xi}$ with $\chi(\boldsymbol{\xi}) \leqslant 1$.

We have (in analogy to (13.4)) $\psi_i(\mathbf{x}) = h(\boldsymbol{\beta}_i^{\mathbf{x}})$ $(i = 1, \dots, r)$ for $\mathbf{x} \in \mathbb{Z}^n$, hence

$$\chi(\mathbf{x}) = \max(h(\boldsymbol{\beta}_1^{\mathbf{x}}), \dots, h(\boldsymbol{\beta}_r^{\mathbf{x}})). \tag{15.2}$$

LEMMA 15.1 *Let $\gamma_1, \dots, \gamma_r$ in $K^\times$ be given, and set*

$$\tilde{\chi}(\mathbf{x}) = \max(h(\gamma_1 \boldsymbol{\beta}_1^{\mathbf{x}}), \dots, h(\gamma_r \boldsymbol{\beta}_r^{\mathbf{x}})). \tag{15.3}$$

Then there is a $\mathbf{u} \in \mathbb{Z}^n$ *such that*

$$\tilde{\chi}(\mathbf{x} - \mathbf{u}) \geqslant \tfrac{1}{4} c_2 |\mathbf{x}|,$$

for $\mathbf{x} \in \mathbb{Z}^n$.

Proof. We have $g_{iv}(\boldsymbol{\xi}) = \sum_{j=1}^n \ell_{ijv} \xi_j$ with $\ell_{ijv} = \log \|\beta_{ij}\|_v$, and

$$\psi_i(\boldsymbol{\xi}) = \sum_v \max(0, g_{iv}(\boldsymbol{\xi})) = \frac{1}{2} \sum_v |g_{iv}(\boldsymbol{\xi})|,$$

for $1 \leqslant i \leqslant r$. Set $c_{iv} = \log \|\gamma_i\|_v$ and

$$\tilde{g}_{iv}(\boldsymbol{\xi}, \zeta) = g_{iv}(\boldsymbol{\xi}) + c_{iv}\zeta$$

for $(\boldsymbol{\xi}, \zeta) \in \mathbb{R}^n \times \mathbb{R}^1 = \mathbb{R}^{n+1}$. Further set

$$\tilde{\psi}_i(\boldsymbol{\xi}, \zeta) = \sum_v \max(0, \tilde{g}_{iv}(\boldsymbol{\xi}, \zeta)) = \frac{1}{2} \sum_v |\tilde{g}_{iv}(\boldsymbol{\xi}, \zeta)|,$$

$$\tilde{\chi}(\boldsymbol{\xi}, \zeta) = \max_{1 \leqslant i \leqslant r} \tilde{\psi}_i(\boldsymbol{\xi}, \zeta).$$

Let $\widetilde{\mathcal{X}} \subset \mathbb{R}^{n+1}$ consist of $(\boldsymbol{\xi}, \zeta)$ with $\tilde{\chi}(\boldsymbol{\xi}, \zeta) \leqslant 1$. Then $\widetilde{\mathcal{X}}$ is convex, symmetric, closed, and it contains $\mathbf{0}$ in its interior. But it may be unbounded. The intersection of $\widetilde{\mathcal{X}}$ with the coordinate hyperplane $\zeta = 0$ is $\mathcal{X}$. We have $\tilde{\chi}(\mathbf{x}, 1) = \tilde{\chi}(\mathbf{x})$.

When $\widetilde{\mathcal{X}}$ is unbounded, there is some $(\boldsymbol{\xi}_0, \zeta_0) \neq (\mathbf{0}, 0)$ with $\tilde{\chi}(\boldsymbol{\xi}_0, \zeta_0) = 0$. Since $\mathcal{X}$ is bounded, $\zeta_0 \neq 0$. By homogeneity, there is some $(\boldsymbol{\xi}_1, 1)$ with $\tilde{\chi}(\boldsymbol{\xi}_1, 1) = 0$. On the other hand, when $\widetilde{\mathcal{X}}$ is bounded, hence compact, pick $(\boldsymbol{\xi}_0, \zeta_0)$ in $\widetilde{\mathcal{X}}$ with ζ_0 maximal. We rewrite $\boldsymbol{\xi}_0 = \zeta_0 \boldsymbol{\xi}_1$, so that $\zeta_0(\boldsymbol{\xi}_1, 1) \in \widetilde{\mathcal{X}}$.

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

Now suppose that $(\boldsymbol{\xi}, \zeta) \in \widetilde{\mathcal{X}}$. When $\widetilde{\mathcal{X}}$ is unbounded, $\zeta(\boldsymbol{\xi}_1, 1) \in \widetilde{\mathcal{X}}$; but this is also true when $\widetilde{\mathcal{X}}$ is bounded, since $|\zeta| \leqslant \zeta_0$ in that case. Taking the difference, we see that $(\boldsymbol{\xi} - \zeta\boldsymbol{\xi}_1, 0) \in 2\widetilde{\mathcal{X}}$, which yields $\boldsymbol{\xi} - \zeta\boldsymbol{\xi}_1 \in 2\mathcal{X}$. Thus $(\boldsymbol{\xi}, \zeta) \in \widetilde{\mathcal{X}}$ implies $\boldsymbol{\xi} - \zeta\boldsymbol{\xi}_1 \in 2\mathcal{X}$. Therefore, by reason of homogeneity, $\chi(\boldsymbol{\xi} - \zeta\boldsymbol{\xi}_1) \leqslant 2\tilde{\chi}(\boldsymbol{\xi}, \zeta)$. Put differently,

$$\chi(\boldsymbol{\xi}) \leqslant 2\tilde{\chi}(\boldsymbol{\xi} + \zeta\boldsymbol{\xi}_1, \zeta), \tag{15.4}$$

for any $(\boldsymbol{\xi}, \zeta) \in \mathbb{R}^{n+1}$.

Pick $\mathbf{u} \in \mathbb{Z}^n$ such that $\mathbf{u} = -\boldsymbol{\xi}_1 + \boldsymbol{\mu}$, where the coordinates $|\mu_i| \leqslant \frac{1}{2}$ $(i = 1, \ldots, n)$. Then by (15.4) with $\zeta = 1$,

$$\begin{aligned} \tilde{\chi}(\mathbf{x} - \mathbf{u}) &= \tilde{\chi}(\mathbf{x} - \mathbf{u}, 1) = \tilde{\chi}(\mathbf{x} - \boldsymbol{\mu} + \boldsymbol{\xi}_1, 1) \\ &\geqslant \tfrac{1}{2}\chi(\mathbf{x} - \boldsymbol{\mu}) \geqslant \tfrac{1}{2}c_2|\mathbf{x} - \boldsymbol{\mu}| \geqslant \tfrac{1}{4}c_2|\mathbf{x}|, \end{aligned}$$

in view of (15.1).

As was pointed out above, we want $h_{LE}(\mathbf{x})$ large for certain forms L, and not $h_{LD}(\mathbf{x})$ as in Lemma 14.1. One might try to write $b_{\ell M}\boldsymbol{\alpha}_\ell^{\mathbf{x}} = b_{\ell M}^{y}\boldsymbol{\alpha}_\ell^{\mathbf{x}}$ with $y = 1$, i.e., to add a variable y, so that $b_{\ell M}^{y}\boldsymbol{\alpha}_\ell^{\mathbf{x}} = (\tilde{\boldsymbol{\alpha}}_\ell)^{\tilde{\mathbf{x}}}$ with $\tilde{\boldsymbol{\alpha}}_\ell = (b_{\ell M}, \alpha_{\ell 1}, \ldots, \alpha_{\ell n})$ and $\tilde{\mathbf{x}} = (y, \mathbf{x})$. This way the coefficients seem to disappear miraculously. However, then the initial transformation of Section 13, which now is in $\mathrm{GL}(n+1, \mathbb{Z})$, will transform a polynomial having all its coefficients equal to 1 into a polynomial whose coefficients are not necessarily 1, thus reintroducing coefficients. For this reason this simple idea does not seem to work. We will take recourse to Lemma 15.1 instead.

We order the monomials lexicographically: write $M > N$ when $M = X_1^{i_1} \ldots X_n^{i_n}$, $N = X_1^{j_1} \ldots X_n^{j_n}$ with $i_s > j_s$, $i_{s+1} = j_{s+1}, \ldots, i_n = j_n$ for some s. We also introduce a 'pseudomonomial' $\square$ and write $M > \square$ for every genuine monomial M. Let

$$L = L^1 + \cdots + L^k = \sum_{\ell \in \Lambda} \sum_{M \in \mathbf{M}_\ell} b_{\ell M}\xi_{\ell M}$$

be a form in the notation (12.1). When $\ell \in \mathcal{B}(L)$, so that $L^\ell \neq 0$, let $M_\ell(L)$ be the monomial which is largest with respect to the ordering $>$ among the monomials with nonzero coefficients $b_{\ell M}$, and let $b_\ell(L)$ be the corresponding coefficient. When $\ell \notin \mathcal{B}(L)$, so that $L^\ell = 0$, we set $M_\ell(L) = \square$, $b_\ell(L) = 0$, $b_\ell(L)M_\ell(L) = \square$. We call $M_\ell(L)$, $b_\ell(L)$ and $b_\ell M_\ell(L)$ the *leading monomials, leading coefficients*, and *leading terms* of L, respectively. To every form L there belongs a k-tuple of leading terms $(b_1(L)M_1(L), \ldots, b_k(L)M_k(L))$, as well as k-tuples of leading monomials and leading coefficients.

Clearly

$$h_{LE}(\mathbf{x}) \geqslant h'_{LE}(\mathbf{x}), \tag{15.5}$$

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

where $h'_{LE}(\mathbf{x})$ is the height of the vector $\boldsymbol{\varepsilon}'_l(\mathbf{x})$ with components $b_\ell(L)\boldsymbol{\alpha}_\ell^{\mathbf{x}}$ where $\ell \in \mathcal{B}(L)$.

An examination of the proof of Lemma 14.1 shows that we have really proved that

$$\chi(\mathbf{x}) := \max_{\ell,m} h((\boldsymbol{\alpha}_\ell/\boldsymbol{\alpha}_m)^{\mathbf{x}}) = \max_{\ell,m} h(\boldsymbol{\alpha}_\ell^{\mathbf{x}}:\boldsymbol{\alpha}_m^{\mathbf{x}}) \geqslant c_2|\mathbf{x}|, \tag{15.6}$$

where the maximum is over all pairs ℓ, m having $\ell, m \in \mathcal{B}(L_{\lambda j})$ for some λ, j. Thus $\chi(\mathbf{x}) = \operatorname{Max} h(\boldsymbol{\alpha}_\ell^{\mathbf{x}}:\boldsymbol{\alpha}_m^{\mathbf{x}}) = \operatorname{Max} h((\boldsymbol{\alpha}_\ell/\boldsymbol{\alpha}_m)^{\mathbf{x}})$, where Max signifies the maximum over all quadruples λ, j, m, ℓ with $\lambda \in \mathcal{P}$, $1 \leqslant j \leqslant t(\lambda)$ and $\ell, m \in \mathcal{B}(L_{\lambda j})$. Consider

$$\begin{aligned}\widetilde{\chi}(\mathbf{x}) &= \operatorname{Max} h(b_\ell(L_{\lambda j})\boldsymbol{\alpha}_\ell^{\mathbf{x}}:b_m(L_{\lambda j})\boldsymbol{\alpha}_m^{\mathbf{x}}) \\ &= \operatorname{Max} h((b_\ell(L_{\lambda j})/b_m(L_{\lambda j}))(\boldsymbol{\alpha}_\ell/\boldsymbol{\alpha}_m)^{\mathbf{x}}).\end{aligned}$$

In view of (15.6) and Lemma 15.1, there is a $\mathbf{u} \in \mathbb{Z}^n$ with

$$\widetilde{\chi}(\mathbf{x}-\mathbf{u}) \geqslant c_3|\mathbf{x}|, \tag{15.7}$$

for $\mathbf{x} \in \mathbb{Z}^n$, where we set

$$c_3 = \tfrac{1}{4}c_2 = (84kd^3 \cdot 4^n)^{-1}. \tag{15.8}$$

The idea now is to apply the translation $\mathbf{x} \mapsto \mathbf{x}-\mathbf{u}$. Then $P_\ell(\mathbf{x})\boldsymbol{\alpha}_\ell^{\mathbf{x}}$ becomes $P_\ell(\mathbf{x}-\mathbf{u})\boldsymbol{\alpha}_\ell^{\mathbf{x}-\mathbf{u}} = \widehat{P}_\ell(\mathbf{x})\boldsymbol{\alpha}_\ell^{\mathbf{x}}$ with $\widehat{P}_\ell(\mathbf{x}) = \boldsymbol{\alpha}_\ell^{-\mathbf{u}}P_\ell(\mathbf{x}-\mathbf{u})$. We had $L = \sum_{\ell,M} b_{\ell M}\xi_{\ell M}$ and $P_\ell = \sum_M b_{\ell M}M$; now write $\widehat{P}_\ell = \sum_M \hat{b}_{\ell M}M$ and set $\widehat{L} = \sum_{\ell,M}\hat{b}_{\ell M}\xi_{\ell M}$. Then $L(\boldsymbol{\xi}(\mathbf{x}-\mathbf{u})) = \widehat{L}(\boldsymbol{\xi}(\mathbf{x}))$. The subspace T consists of $\boldsymbol{\xi}$ having $L(\boldsymbol{\xi}) = 0$ for $L \in \mathcal{L}(T)$. Let $\widehat{T}$ be the space of $\boldsymbol{\xi}$ having $\widehat{L}(\boldsymbol{\xi}) = 0$ for $L \in \mathcal{L}(T)$, so that $\mathcal{L}(\widehat{T})$ consists of forms $\widehat{L}$ with $L \in \mathcal{L}(T)$. Our transformation does not mess up $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k$, so that again $\mathcal{P} \prec \widehat{T}$.

In short, we may replace T by $\widehat{T}$, the forms L by $\widehat{L}$. We have $\mathcal{B}(\widehat{L}) = \mathcal{B}(L)$, and when forms $L_{\lambda j}$ have the property enunciated in Lemma 14.1, then so do the forms $\widehat{L}_{\lambda j}$. The leading monomials are not changed by a substitution $\mathbf{x} \mapsto \mathbf{x}-\mathbf{u}$. Therefore when $b_\ell(L)$ was a leading coefficient of L, then $b_\ell(L)\boldsymbol{\alpha}_\ell^{-\mathbf{u}}$ is a leading coefficient of $\widehat{L}$. By (15.7) and the definition of $\widetilde{\chi}$,

$$\operatorname{Max} h(b_\ell(\widehat{L}_{\lambda j})\boldsymbol{\alpha}_\ell^{\mathbf{x}}:b_m(\widehat{L}_{\lambda j})\boldsymbol{\alpha}_m^{\mathbf{x}}) \geqslant c_3|\mathbf{x}|.$$

In other words, after performing the substitution $\mathbf{x} \mapsto \mathbf{x}-\mathbf{u}$, we may suppose that $\widetilde{\chi}(\mathbf{x}) \geqslant c_3|\mathbf{x}|$. In view of the definition of h'_{LE}, we have the following.

LEMMA 15.2. *After a suitable translation* $\mathbf{x} \mapsto \mathbf{x}-\mathbf{u}$, *the forms* $L_{\lambda j}$ *of Lemma* 14.1 *have*

$$\max_{\lambda,j} h'_{L_{\lambda j}E}(\mathbf{x}) \geqslant c_3|\mathbf{x}|, \tag{15.9}$$

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

where the maximum is over $\lambda \in \mathcal{P}$, $1 \leqslant j \leqslant t(\lambda)$.

16. Construction of Minimal Forms

As pointed out in Section 12, we need to apply Proposition B to a minimal form. Now it would be easy to construct forms $L_{\lambda j}$ ($\lambda \in \mathcal{P}, 1 \leqslant j \leqslant t(\lambda)$) as in Lemma 14.1 which are minimal. However, minimality may be destroyed by the substitution $\mathbf{x} \mapsto \mathbf{x} - \mathbf{u}$, i.e., changing $L_{\lambda j}$ to $\widehat{L}_{\lambda j}$. This difficulty necessitates a somewhat complicated construction.

Suppose

$$\mathcal{P} \prec T. \tag{16.1}$$

For $\lambda \in \mathcal{P}$, let $\mathcal{L}_\lambda(T)$ consist of forms $L \in \mathcal{L}(T)$ with $\mathcal{B}(L) \subset \lambda$. As a consequence of (16.1),

$$\mathcal{L}(T) = \bigoplus_{\lambda \in \mathcal{P}} \mathcal{L}_\lambda(T). \tag{16.2}$$

We now begin our construction. Let $\lambda \in \mathcal{P}$ with $|\lambda| > 1$ be given. To fix ideas, suppose that $\lambda = \{1, \ldots, r\}$. We will construct a partition of λ, $\lambda = \bigcup_{j=1}^t \mu_j$, into nonempty subsets $\mu_1, \ldots, \mu_t$, as well as forms $L_1, \ldots, L_t$ in $\mathcal{L}_\lambda(T)$.

A form $L \in \mathcal{L}_\lambda(T)$ will be called 1-*stable* if $|\mathcal{B}(L)| > 1$ and if L cannot be written as $L = L' + L''$ where L', L'' are nonzero, lie in $\mathcal{L}_\lambda(T)$, and have $\mathcal{B}(L') \cap \mathcal{B}(L'') = \emptyset$. There are 1-stable forms, for otherwise every form in $\mathcal{L}_\lambda(T)$ would be a sum of forms whose sets $\mathcal{B}$ are of cardinality 1, so that if $\mathcal{Q}$ is obtained from $\mathcal{P}$ by chopping up λ into the singletons $\{1\}, \ldots, \{r\}$, then $\mathcal{Q}$ would be agreeable with T, contradicting (16.1). Pick $\mu_1 \subset \lambda$ of minimal cardinality such that there is a 1-stable form L with $\mathcal{B}(L) = \mu_1$. Clearly $|\mu_1| > 1$.

Suppose $j > 1$, and subsets $\mu_1, \ldots, \mu_{j-1}$ of λ have been chosen. Set $\nu_{j-1} = \bigcup_{i=1}^{j-1} \mu_i$. We are finished if $\nu_{j-1} = \lambda$ (just set $t = j - 1$); otherwise let $\bar{\nu}_{j-1}$ be the complement of ν_{j-1} in λ. A form $L \in \mathcal{L}_\lambda(T)$ will be called j-*stable* if L cannot be written as $L = L' + L''$ where $\mathcal{B}(L') \subset \nu_{j-1}$, $\mathcal{B}(L'') \subset \bar{\nu}_{j-1}$. There are j-stable forms, for otherwise every form $L \in \mathcal{L}_\lambda(T)$ could be written as a sum: $L = L' + L''$ as above, so that if $\mathcal{Q}$ is obtained from $\mathcal{P}$ by dividing λ into ν_{j-1} and $\bar{\nu}_{j-1}$, then $\mathcal{Q}$ would be agreeable with T, contradicting (16.1). Pick $\mu_j \subset \bar{\nu}_{j-1}$ of minimal cardinality such that there is a j-stable form L with

$$\mathcal{B}(L) \cap \bar{\nu}_{j-1} = \mu_j. \tag{16.3}$$

Clearly $\mu_j \neq \emptyset$. Continuing in this way we finally get sets $\mu_1, \ldots, \mu_t$ which partition λ.

We may renumber the elements of λ such that $\mu_j = \{r_{j-1} + 1, \ldots, r_j\}$ ($j = 1, \ldots, t$) with $0 = r_0 < r_1 < \cdots < r_t = r$. Now recall that $M_1(L), \ldots,$

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

$M_r(L), \ldots, M_k(L)$ were the 'leading monomials' of L. Given forms L, L' in $\mathcal{L}_\lambda(T)$, write $L' < L$ if

$$M_s(L') < M_s(L), M_{s+1}(L') = M_{s+1}(L), \ldots, M_r(L') = M_r(L),$$

for some s.

Our construction was such that for each j, $1 \leqslant j \leqslant t$, there are j-stable forms L_j with (16.3) (where we set $\nu_0 = \emptyset$, $\bar{\nu}_0 = \lambda$). A form L_j will be called *proper* if it is minimal (with respect to $<$) among j-stable forms with (16.3). Since $<$ induces only a partial ordering of the forms (only the leading monomials matter for $<$), j-proper forms are not uniquely determined. However, if both L_j, L'_j are j-proper, then

$$(M_1(L_j), \ldots, M_{r_j}(L_j)) = (M_1(L'_j), \ldots, M_{r_j}(L'_j)).$$

LEMMA 16.1. *Suppose L_j, L'_j are j-proper. Then the r_j-tuples of leading coefficients*

$$(a_1(L_j), \ldots, a_{r_j}(L_j)) \quad \text{and} \quad (a_1(L'_j), \ldots, a_{r_j}(L'_j)) \tag{16.4}$$

are proportional.

Proof. The coefficients $a_{r_j}(L_j)$, and $a_{r_j}(L'_j)$ are nonzero by (16.3). Set $J = a_{r_j}(L'_j)L_j - a_{r_j}(L_j)L'_j$. Then

$$J < L_j. \tag{16.5}$$

If the r_j-tuples (16.4) were not proportional, there would be a g with $M_g(J) = M_g(L_j) \neq \square$. There is an $\alpha \in K$ with

$$M_g(L_j - \alpha J) < M_g(L_j). \tag{16.6}$$

In the case $j = 1$ write

$$J = J^1 + \cdots + J^{r_1}, \tag{16.7}$$

$$L_1 = L_1^1 + \cdots + L_1^{r_1} \tag{16.8}$$

in the notation of (12.1). Now (16.5), i.e., $J < L_1$, and the hypothesis that L_1 is minimal imply that J is not 1-stable, and by the minimality of μ_1 it is easy to conclude that each $J^i \in \mathcal{L}_\lambda(T)$. We say that '$J$ splits completely.' Then

$$\widetilde{L}_1 = L_1 - \alpha J^g \in \mathcal{L}_\lambda(T) \tag{16.9}$$

and $\widetilde{L}_1 < L_1$ by (16.6). Therefore also $\widetilde{L}_1$ splits completely. But

$$\widetilde{L}_1 = L_1^1 + \cdots + L_1^{g-1} + (L_1^g - \alpha J^g) + L_1^{g+1} + \cdots + L_1^{r_1}.$$

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

Therefore L_1^i for $i \neq g$ is in $\mathcal{L}_\lambda(T)$, hence so is L_1^g, and L_1 splits completely, against the fact that it is 1-stable.

In the case $j > 1$ write

$$J = J^* + J^{**}, \qquad L_j = L_j^* + L_j^{**},$$

with $\mathcal{B}(J^*), \mathcal{B}(L_j^*) \subset \nu_{j-1}$ and $\mathcal{B}(J^{**}), \mathcal{B}(L_j^{**}) \subset \mu_j$. Now (16.5) implies that J is not j-stable, so that $J^*, J^{**} \in \mathcal{L}_\lambda(T)$. We say that '$J$ splits.' Set

$$\widetilde{L}_j = \begin{cases} L_j - \alpha J^*, & \text{if } g \in \nu_{j-1}, \\ L_j - \alpha J^{**}, & \text{if } g \in \mu_j. \end{cases} \tag{16.10}$$

Then $\widetilde{L}_j \in \mathcal{L}_\lambda(T)$, further $\widetilde{L}_j < L_j$ by (16.6). Therefore $\widetilde{L}_j$ also splits. E.g., in the case when $g \in \nu_{j-1}$,

$$\widetilde{L}_j = (L_j^* - \alpha J^*) + L_j^{**},$$

so that $L_j^{**} \in \mathcal{L}_\lambda(T)$, hence L_j splits, against the fact that it is j-stable. The situation is similar when $g \in \mu_j$.

LEMMA 16.2. *Let L_j be a j-proper form with $|\mathcal{A}(L_j)|$ as small as possible. Then L_j is a minimal form.*

Proof. Suppose to the contrary that there is a form $J \neq 0$ in $\mathcal{L}_\lambda(T)$ with $\mathcal{A}(J) \subsetneqq \mathcal{A}(L_j)$. By the special choice of L_j, the form J cannot be j-proper. But $J < L_j$ or $J \sim L_j$ (meaning that J, L_j have the same leading monomials), and hence J cannot be j-stable.

Write

$$J = J^1 + \cdots + J^{r_j}, \tag{16.11}$$

$$L_j = L_j^1 + \cdots + L_j^{r_j} \tag{16.12}$$

in the notation of (12.1). Some $J^g \neq 0$. Every monomial occurring with nonzero coefficient in J^g also occurs so in L_j^g. Therefore there is an $\alpha \in K$ with

$$\mathcal{A}(L_j^g - \alpha J^g) \subsetneqq \mathcal{A}(L_j^g). \tag{16.13}$$

In the case $j = 1$, J (being not 1-stable) splits completely, and we have (16.9) again. But $\mathcal{A}(\widetilde{L}_1) \subsetneqq \mathcal{A}(L_1)$ by (16.13). Therefore by the special property of L_1, the form $\widetilde{L}_1$ cannot be 1-proper. But $\widetilde{L}_1 < L_1$ or $\widetilde{L}_1 \sim L_1$, so that $\widetilde{L}_1$ is not 1-stable, hence splits completely. We get a contradiction as in the proof of Lemma 16.1.

In the case $j > 1$, J splits, and $\widetilde{L}_j$ as defined by (16.10) is in $\mathcal{L}_\lambda(T)$. We have $\mathcal{A}(\widetilde{L}_j) \subsetneqq \mathcal{A}(L_j)$ by (16.13). We may infer that $\widetilde{L}_j$ is not j-proper, further that it is not j-stable, and it splits. Again we get a contradiction as in Lemma 16.1.

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

17. End of Proof

For each $\lambda \in \mathcal{P}$, construct sets $\mu_{\lambda 1}, \ldots, \mu_{\lambda t}$ and linear forms $L_{\lambda 1}, \ldots, L_{\lambda t}$ with $t = t(\lambda) \leqslant |\lambda|$ as described in the preceding section, such that $L_{\lambda j}$ is j-proper.

LEMMA 17.1. *The forms* $L_{\lambda j}$ $(\lambda \in \mathcal{P}, 1 \leqslant j \leqslant t(\lambda))$ *satisfy the hypotheses of Lemma* 14.1.

Proof. We may suppose that λ is given and we write the corresponding sets and forms again as $\mu_1, \ldots, \mu_t$ and $L_1, \ldots, L_t$. We will show by induction on q that if $\ell, m \in \nu_q = \bigcup_{j=1}^{q} \mu_i$, then there are forms $L_{j(1)}, \ldots, L_{j(q)}$ among $L_1, \ldots, L_t$ with $\ell \in \mathcal{B}(L_{j(1)})$, $m \in \mathcal{B}(L_{j(q)})$ and $\mathcal{B}(L_{j(i)}) \cap \mathcal{B}(L_{j(i+1)}) \neq \emptyset$ for $1 \leqslant i < q$. This is trivial for $q = 1$, for then $\ell, m \in \mu_1 = \mathcal{B}(L_1)$. When $q > 1$, we may suppose that $\ell \in \nu_{q-1}$, $m \in \mu_q$ (for if both $\ell, m \in \mu_q$, then both are in $\mathcal{B}(L_q)$). Now $m \in \mathcal{B}(L_q)$. There is an m' in the nonempty set $\mathcal{B}(L_q) \cap \nu_{q-1}$. By induction, there are forms $L_{j(1)}, \ldots, L_{j(q-1)}$ with $\ell \in \mathcal{B}(L_{j(1)})$, $m' \in \mathcal{B}(L_{j(q-1)})$, and with any two successive L's having their $\mathcal{B}$'s with nonempty intersection. The assertion now holds with $L_{j(q)} = L_q$.

By Lemmas 14.1, 17.1 we have (14.4). Further by Lemma 15.2, we have (15.9) after a suitable translation $\mathbf{x} \mapsto \mathbf{x} - \mathbf{u}$.

Now a translation changes forms L into forms $\widehat{L}$. But $\mathcal{B}(L) = \mathcal{B}(\widehat{L})$. Therefore the new forms $\widehat{L}_{\lambda j}$ are again j-stable $(\lambda \in \mathcal{P}, 1 \leqslant j \leqslant t(\lambda))$. In fact the leading monomials are not changed, and therefore the new forms $\widehat{L}_{\lambda j}$ are again j-proper (with respect to the new space $\widehat{T}$). These new forms have leading coefficients such that (15.9) holds. We finally replace $\widehat{L}_{\lambda j}$ by a j-proper form $\widetilde{L}_{\lambda j}$ whose support $\mathcal{A}(\widetilde{L}_{\lambda j})$ is minimal. Then $\widetilde{L}_{\lambda j}$ is a minimal form by Lemma 16.2. In view of Lemma 16.1, the leading coefficients of $\widetilde{L}_{\lambda j}$ $(j = 1, \ldots, t(\lambda))$ are proportional to the leading coefficients of $\widehat{L}_{\lambda j}$ $(j = 1, \ldots, t(\lambda))$, so that again (15.9) holds. Therefore in view of (15.5), *we may suppose that we have minimal forms* $L_{\lambda j}$ $(\lambda \in \mathcal{P}, 1 \leqslant j \leqslant t(\lambda))$ *with*

$$\max_{\lambda, j} h_{L_{\lambda j} E}(\mathbf{x}) \geqslant c_3 |\mathbf{x}|.$$

We now divide the solutions $\mathbf{x} \in \mathcal{X}(T, \mathcal{P})$ into possibly overlapping classes $C_{\lambda j}$, with $\mathbf{x} \in C_{\lambda j}$ if

$$h_{L_{\lambda j} E}(\mathbf{x}) \geqslant c_3 |\mathbf{x}|. \tag{17.1}$$

Since $t(\lambda) \leqslant |\lambda|$, the number of classes is $\leqslant |\Lambda| = k \leqslant A$.

Now let λ, j be fixed and consider solutions $\mathbf{x} \in C_{\lambda j}$. Here $L_{\lambda j}(\mathbf{x}) = 0$, and this equation is as in Proposition B, i.e., (4.4) with $\mathcal{A} = \mathcal{A}(L_{\lambda j})$. Suppose initially that $a = |\mathcal{A}(L_{\lambda j})| \geqslant 3$. The monomials occurring in $L_{\lambda j}$ have total degree $\leqslant \max(\delta_1, \ldots, \delta_k) \leqslant A$, so that $H_M(\mathbf{x}) \leqslant |\mathbf{x}|^{\mathcal{A}}$. In view of (17.1), the condition (4.7) will be satisfied if $A \log |\mathbf{x}| \leqslant 1/(4a^2) c_3 |\mathbf{x}|$. Since $a \leqslant A$, this will certainly be

https://doi.org/10.1023/A:1001719425893 Published online by Cambridge University Press

true if $|\mathbf{x}| \leqslant \exp((4A^3)^{-1}c_3|\mathbf{x}|)$. Since $\exp t \geqslant t^2/2$, the condition will be amply satisfied if

$$|\mathbf{x}| \geqslant 32A^6 c_3^{-2}. \tag{17.2}$$

By Proposition B, the solutions with (17.2) yield at most (4.8) proper subspaces of T. (Here we used that $L_{\lambda j}$ was minimal – see the discussion in Section 12.)

Summing over the classes, of which there are at most A, and noting that each $a = |\mathcal{A}(L_{\lambda j})| \leqslant A$, we get a bound $A \cdot 2^{30A^2}(32A^2)^n d^{3(n+A)}$. Since $2 \leqslant A \leqslant B$ and $n \leqslant B$, the total number of subspaces is $< 2^{34B^2} d^{6B} = C$.

We are left with the solutions where (17.2) is violated, so that by (15.8),

$$|\mathbf{x}| < 32B^6(84kd^3 \cdot 4^n)^2 < 2^{4n+18}k^2B^6d^6. \tag{17.3}$$

Since $2 \leqslant k \leqslant B$ and $n \leqslant B$ we obtain $|\mathbf{x}| < 2^{19B}d^6$, and the number of such $\mathbf{x} \in \mathbb{Z}^n$ is

$$< 2^{20B^2} d^{6n} < C.$$

This establishes Proposition C when $a \geqslant 3$.

When $a = 2$, the equation $L_{\lambda j} = 0$ is of the type $a_{\ell M} M(\mathbf{x}) \boldsymbol{\alpha}_\ell^{\mathbf{x}} + a_{\ell' M'} M'(\mathbf{x}) \boldsymbol{\alpha}_{\ell'}^{\mathbf{x}} = 0$, so that

$$h_{L_{\lambda j}E}(\mathbf{x}) = h_{L_{\lambda j}M}(\mathbf{x}) \leqslant A \log |\mathbf{x}|.$$

Together with (17.1) this yields $c_3|\mathbf{x}| \leqslant A \log |\mathbf{x}|$, so that $|\mathbf{x}| \geqslant \exp(A^{-1}c_3|\mathbf{x}|) \geqslant \frac{1}{2}A^{-2}c_3^2|\mathbf{x}|^2$, i.e., $|\mathbf{x}| \leqslant 2A^2c_3^{-2}$. This gives (17.3) and hence leads again to fewer than C solutions. So when $a = 2$, then $\mathcal{X}(T, \mathcal{P})| < C$, and Proposition C follows.

References

1. Bombieri, E. and Mueller, J.: On a conjecture of Siegel, *Monatsh. Math.* **125** (1998), 293–308.
2. Cassels, J. W. S.: *An Introduction to the Geometry of Numbers*, Grundlehren Math. Wiss. 99, Springer, New York, 1959.
3. Dobrowolski, E.: On a question of Lehmer and the number of irreducible factors of a polynomial, *Acta Arith.* **34** (1979), 391–401.
4. Evertse, J. H.: An improvement of the quantitative subspace theorem, *Compositio Math.* **101** (1996), 225–311.
5. Evertse, J.H., Győry, K., Stewart, C. L. and Tijdeman. R.: On S-unit equations in two unknowns, *Invent. Math.* **92** (1988), 461–477.
6. Evertse, J. H. and Schlickewei, H. P.: A quantitative version of the absolute subspace theorem, (to appear).
7. Evertse, J. H., Schlickewei, H. P. and Schmidt, W. M.: Linear equations in variables which Lie in a multiplicative group, (to appear).
8. Faltings, G.: Diophantine approximation on abelian varieties, *Ann. of Math.* **113** (1991), 549–576.

9. Laurent, M.: Équations exponentielles-polynômes et suites récurrentes linéaires, A*stérisque* **147–148** (1987), 121–139.
10. Laurent, M.: Équations exponentielles-polynômes et suites récurrentes linéaires, II, *J. Number Theory* **31** (1989), 24–53.
11. Rogers, C. A.: A note on coverings and packings, *J. London Math. Soc.* **25** (1950), 327–331.
12. Rogers, C. A.: A note on coverings, *Mathematika* **4** (1957), 1–6.
13. Schlickewei, H. P.: S-unit equations over number fields, *Invent. Math.* **102** (1990), 95–107.
14. Schlickewei, H. P.: Multiplicities of recurrence sequences, *Acta Math.* **176** (1996), 171–243.
15. Schlickewei, H. P.: Lower bounds for heights on finitely generated groups, *Monatsh. Math.* **123** (1997), 171–178.
16. Schlickewei, H. P. and Schmidt, W. M.: On polynomial exponential equations, *Math. Ann.* **296** (1993), 339–361.
17. Schlickewei, H. P. and Schmidt, W. M.: Linear equations in members of recurrence sequences, *Ann. Scuola Norm. Sup. Pisa, Serie Fis. e Mat.* **20** (1993), 219–246.
18. Schlickewei, H. P. and Schmidt, W. M.: The intersection of recurrence sequences, *Acta Arith.* **72** (1995), 1–44.
19. Schmidt, W. M.: Heights of algebraic points lying on curves or hypersurfaces, *Proc. Amer. Math. Soc.* **124** (1996), 3003–3013.