

## VOJTA'S REFINEMENT OF THE SUBSPACE THEOREM

WOLFGANG M. SCHMIDT

**ABSTRACT.** Vojta's refinement of the Subspace Theorem says that given linearly independent linear forms  $L_1, \dots, L_n$  in  $n$  variables with algebraic coefficients, there is a finite union  $U$  of proper subspaces of  $\mathbb{Q}^n$ , such that for any  $\varepsilon > 0$  the points  $\underline{x} \in \mathbb{Z}^n \setminus \{0\}$  with (1)  $|L_1(\underline{x}) \cdots L_n(\underline{x})| < |\underline{x}|^{-\varepsilon}$  lie in  $U$ , with finitely many exceptions which will depend on  $\varepsilon$ . Put differently, if  $X(\varepsilon)$  is the set of solutions of (1), if  $\overline{X}(\varepsilon)$  is its closure in the subspace topology (whose closed sets are finite unions of subspaces) and if  $\overline{X}'(\varepsilon)$  consists of components of dimension  $> 1$ , then  $\overline{X}'(\varepsilon) \subset U$ . In the present paper it is shown that  $\overline{X}'(\varepsilon)$  is in fact constant when  $\varepsilon$  lies outside a simply described finite set of rational numbers.

More generally, let  $k$  be an algebraic number field and  $S$  a finite set of absolute values of  $k$  containing the archimedean ones. For  $v \in S$  let  $L_1^v, \dots, L_m^v$  be linear forms with coefficients in  $k$ , and for  $\underline{x} \in K^n \setminus \{0\}$  with height  $H_k(\underline{x}) > 1$  define  $a_{vi}(\underline{x})$  by  $|L_i^v(\underline{x})|_v / |\underline{x}|_v = H_k(\underline{x})^{-a_{vi}(\underline{x})/d_v}$  where the  $d_v$  are the local degrees. The approximation set  $A$  consists of tuples  $\underline{a} = \{a_{vi}\}$  ( $v \in S$ ,  $1 \leq i \leq m$ ) such that for every neighborhood  $O$  of  $\underline{a}$  the points  $\underline{x}$  with  $\{a_{vi}(\underline{x})\} \in O$  are dense in the subspace topology. Then  $A$  is a polyhedron whose vertices are rational points.

### 1. INTRODUCTION

In its simplest form, the Subspace Theorem [2, 3] says that if  $L_1, \dots, L_n$  are linearly independent linear forms in  $n$  variables with algebraic coefficients, then for any  $\varepsilon > 0$ , the integer points  $\underline{x} \neq \underline{0}$  with

$$(1.1) \quad |L_1(\underline{x}) \cdots L_n(\underline{x})| < |\underline{x}|^{-\varepsilon}$$

(where  $|\underline{x}|$  is the maximum norm, say), lie in finitely many proper subspaces of  $\mathbb{Q}^n$ . In its simplest form, Vojta's refinement [5] says that there are proper subspaces  $T_1, \dots, T_t$  independent of  $\varepsilon$ , such that the integer points with (1.1) lie in the union of  $T_1, \dots, T_t$ , with the exception of finitely many points which may depend on  $\varepsilon$ .

More generally, let  $k \subset K$  be algebraic number fields. Let  $M(k)$  denote the set of absolute values of  $k$ , normalized such that they extend the standard or a  $p$ -adic absolute value of  $\mathbb{Q}$ , so that for  $\varphi \in k^*$  we have the product formula

---

Received by the editors September 30, 1991.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 11J68, 11J13.

*Key words and phrases.* Simultaneous approximation to algebraic numbers, Subspace Theorem, subspace topology.

$\prod_{v \in M(k)} |\varphi|_v^{d_v} = 1$ , where the  $d_v$  are the local degrees of  $k$  over  $\mathbb{Q}$ . Let  $S$  be a finite subset of  $M(k)$  containing all the archimedean absolute values, and suppose that for each  $v \in S$ , the absolute value  $|\cdot|_v$  of  $k$  is extended in some way to an absolute value of  $K$ . Further suppose that for each  $v \in S$ , we are given linearly independent linear forms  $L_1^v, \dots, L_{a(v)}^v$  in  $n$  variables with coefficients in  $K$ . ( $v$  in raised position will always mean a superscript, not exponentiation.) For  $\underline{x} \in k^n \setminus \underline{0}$  put

$$(1.2) \quad \lambda_i^v(\underline{x}) = |L_i^v(\underline{x})|_v / |\underline{x}|_v \quad (v \in S, 1 \leq i \leq a(v)),$$

where  $|\underline{x}|_v$  is the maximum norm with respect to  $|\cdot|_v$ . Set

$$(1.3) \quad f(\underline{x}) = \prod_v \prod_i (\lambda_i^v(\underline{x}))^{d_v};$$

here and in the sequel,  $v$  will range through  $S$ , and  $i$  over  $1 \leq i \leq a(v)$ . For  $\underline{x} \in k^n \setminus \underline{0}$ , let  $H_k(\underline{x})$  be the multiplicative field height, as defined, e.g., in [4, §VIII.5]. A more general version of the Subspace Theorem due to Schlickewei [1], which is easily seen to contain the assertion stated at the beginning, says that for any  $\varepsilon > 0$  the points  $\underline{x} \in k^n \setminus \underline{0}$  with

$$(1.4) \quad 0 < f(\underline{x}) < H_k(\underline{x})^{-n-\varepsilon}$$

lie in finitely many proper subspaces of  $k^n$ . (We could have allowed  $f(\underline{x}) = 0$ , since such points lie in finitely many proper subspaces, but in what follows it will be more convenient to neglect such points.) Again there is a refinement due to Vojta. It says that the solutions of (1.4) lie in the union of finitely many proper subspaces  $T_1, \dots, T_l$  of  $k^n$  independent of  $\varepsilon$ , and a finite number of 1-dimensional subspaces which may depend on  $\varepsilon$ .

The *subspace topology* of  $k^n \setminus \underline{0}$  is the topology whose closed sets are the finite unions of sets  $T \setminus \underline{0}$ , where  $T$  is a subspace. (It is a linear analogue of the Zariski topology.) Thus the closure of a single point  $\underline{x} \in k^n \setminus \underline{0}$  is  $T \setminus \underline{0}$ , where  $T$  is the 1-dimensional subspace spanned by  $\underline{x}$ . A subset of  $k^n \setminus \underline{0}$  is *dense* if it has nonempty intersection with any nonempty open subset, and this means precisely that it is not contained in a finite union of proper subspaces. The Subspace Theorem says that the solutions of (1.4) do not lie dense. Put differently, if  $X(n+\varepsilon)$  is the set of solutions of (1.4) and  $\bar{X}(n+\varepsilon)$  is its closure, then  $\bar{X}(n+\varepsilon) \neq k^n \setminus \underline{0}$ .

The  $\lambda_i^v$ ,  $f$  and  $H_k$  are invariant under replacing  $\underline{x}$  by  $\varphi \underline{x}$  when  $\varphi \in k^*$ . Therefore (1.4) may be interpreted as a relation involving a point of projective space  $\mathbb{P}^{n-1}(k)$ . There is an obvious subspace topology on  $\mathbb{P}^{n-1}(k)$ . Nevertheless, it will be more convenient to deal with  $k^n \setminus \underline{0}$  rather than  $\mathbb{P}^{n-1}(k)$ .

Every nonempty closed set  $\bar{X}$  may be written as a finite union of subspaces:  $\bar{X} = \bigcup_{l=1}^m T_l$ . (We ignore, for convenience, the fact that  $\underline{0}$  should be removed.) This union is unique if we ask for  $T_l \not\subseteq T_q$  if  $l \neq q$ . The  $T_l$  are the *components* of  $\bar{X}$ . Let  $\bar{X}'$  be the union of components of dimension  $> 1$ . Vojta's Theorem says that

$$(1.5) \quad \bar{X}'(n+\varepsilon) \subset Y,$$

where  $Y$  is a closed proper subset of  $k^n$ , independent of  $\varepsilon$ . (Again, more precisely, we should take  $k^n \setminus \underline{0}$ .) In fact Vojta constructs  $Y$  to be independent

of  $S$ , so that it depends on the set of linear forms  $L_i^v$ , but it does not matter to which absolute values they belong.

Again let  $S$  and linear forms  $L_i^v$  be given ( $v \in S$ ,  $1 \leq i \leq a(v)$ ). We suppose the forms  $L_i^v$  to be nonzero and to have coefficients in  $K$ , but we now do not make any linear independence assumption. Let  $f(\underline{x})$  be as above, and  $X(\rho)$  the set of points  $\underline{x} \in k^n \setminus \underline{0}$  with

$$(1.6) \quad 0 < f(\underline{x}) < H_k(\underline{x})^{-\rho}.$$

Write  $s = \text{card } S$  and

$$a = \sum_v a(v), \quad a^* = \max_v a(v).$$

Let  $\mathfrak{R}$  be the finite set of rational numbers  $\rho$  with  $0 \leq \rho \leq na^*$  and with denominator

$$(1.7) \quad \leq (n\sqrt{s})^a.$$

**Theorem 1.** *For  $\rho$  in any interval disjoint from  $\mathfrak{R}$ , the set  $\bar{X}'(\rho)$  is constant. For  $\rho \notin \mathfrak{R}$ , the set  $\bar{X}'(\rho)$  has at most*

$$(1.8) \quad (ns)^{2a^2}$$

*components.*

There is an  $\varepsilon_0$  such that  $n + \varepsilon \notin \mathfrak{R}$  for  $0 < \varepsilon < 2\varepsilon_0$ , so that  $\bar{X}'(n + \varepsilon) = \bar{X}'(n + \varepsilon_0)$  for  $\varepsilon$  in this range. Since  $X(n + \varepsilon)$  can only decrease as  $\varepsilon$  increases, (1.5) is true with  $Y = \bar{X}'(n + \varepsilon_0)$ , which is  $\neq k^n$  by the Subspace Theorem if  $L_1^v, \dots, L_{a(v)}^v$  are linearly independent for  $v \in S$ . Thus the main assertion of Vojta follows.

For  $n = 1$ ,  $\bar{X}'(\rho) = \emptyset$ , and the theorem is trivial. For  $n = 2$ ,  $\bar{X}'(\rho)$  is  $\emptyset$  or  $k^2$ . Since it cannot increase with  $\rho$ , there is an  $r$  such that  $\bar{X}'(\rho) = k^2$  for  $\rho < r$ , and  $\bar{X}'(\rho) = \emptyset$  for  $\rho > r$ . By the theorem,  $0 \leq r \leq 2a^*$ .

**Theorem 1'.** (i) *When  $n = 2$ ,  $r \in \mathbb{Z}$ .*

(ii) *Suppose  $n = 3$ , each  $a(v) = 3$ , and  $L_1^v, L_2^v, L_3^v$  are linearly independent. Then for  $\varepsilon > 0$ , the set  $\bar{X}'(3 + \varepsilon)$  has at most 3 components when  $s = 1$ , 3s components if no form  $L_i^v$  has a zero in  $k^3 \setminus \underline{0}$ , and  $9\binom{s}{2}$  components in general when  $s > 1$ . These bounds are best possible.*

The question of dependency of  $\bar{X}'(\rho)$  on  $S$  and on the fields  $k, K$  will be taken up in §8. What we do not seem to be able to do, in contrast to Vojta, is to construct  $Y$  by a finite process of taking conjugates, intersections and joins of the subspaces  $L_i^v = 0$ .

## 2. THE APPROXIMATION POLYHEDRON

We will assume from now on that  $n > 1$ . Let  $S$  be as above, and for  $v \in S$  let  $L_1^v, \dots, L_{a(v)}^v$  be nonzero linear forms with coefficients in  $K$ , and let  $\lambda_1^v, \dots, \lambda_{a(v)}^v$  be given by (1.2). Let  $\mathfrak{L}$  be the set of points  $\underline{x} \in k^n$  with  $H_k(\underline{x}) < e$  or with  $L_i^v(\underline{x}) = 0$  for some  $v, i$ . Thus  $\mathfrak{L}$  is the union of a finite set of 1-dimensional subspaces, and at most  $a$  further proper subspaces of  $k^n$ . For  $\underline{x} \in k^n \setminus \mathfrak{L}$  put

$$(2.1) \quad \alpha_i^v(\underline{x}) = -d_v(\log \lambda_i^v(\underline{x}))/\log H_k(\underline{x}).$$

Let  $\underline{\underline{\alpha}}(\underline{\underline{x}})$  be the point in  $\mathbb{R}^a$  with components  $\alpha_i^v(\underline{\underline{x}})$  ( $v \in S$ ,  $1 \leq i \leq a(v)$ ).

A point  $\underline{\underline{\alpha}} = \{\alpha_i^v\} \in \mathbb{R}^a$  will be called an *approximation point* if for every open neighborhood  $O$  of  $\underline{\underline{\alpha}}$  (where we use the standard topology of  $\mathbb{R}^a$ ), the points  $\underline{\underline{x}} \in k^n \setminus \mathcal{L}$  with  $\underline{\underline{\alpha}}(\underline{\underline{x}}) \in O$  are dense in  $k^n$ . The *approximation set*  $A$  is the set of all approximation points.

**Lemma 1.** *The approximation set  $A$  is closed. If  $C \subset \mathbb{R}^a$  is a compact set disjoint from  $A$ , then the points  $\underline{\underline{x}}$  with  $\underline{\underline{\alpha}}(\underline{\underline{x}}) \in C$  are not dense.*

*Proof.* Let  $\underline{\underline{\beta}}$  be in the closure of  $A$ , and  $O$  any open neighborhood of  $\underline{\underline{\beta}}$ . Then  $O$  contains a point  $\underline{\underline{\alpha}} \in A$ , and therefore points  $\underline{\underline{x}}$  with  $\underline{\underline{\alpha}}(\underline{\underline{x}}) \in O$  are dense. Therefore  $\underline{\underline{\beta}} \in A$ , and  $A$  is closed.

If  $C$  is as indicated and  $\underline{\underline{\beta}} \in C$ , there is an open neighborhood  $O$  of  $\underline{\underline{\beta}}$  such that the points  $\underline{\underline{x}}$  with  $\underline{\underline{\alpha}}(\underline{\underline{x}}) \in O$  are not dense. But  $C$  may be covered by a finite number of such open sets  $O_1, \dots, O_l$ , and points  $\underline{\underline{x}}$  with  $\underline{\underline{\alpha}}(\underline{\underline{x}})$  in the union of  $O_1, \dots, O_l$  are not dense.

Lemma 1 remains true under quite general assumptions. For instance,  $L_1^v, \dots, L_{a(v)}^v$  for  $v \in S$  may have coefficients in the completion  $K_v$  of  $K$  with respect to  $v$ . In the situation of the present paper, where our linear forms have coefficients in  $K$ , we will be able to prove much more.

We have  $\lambda_i^v(\underline{\underline{x}}) \leq c_1$  (a constant independent of  $\underline{\underline{x}}$ ), therefore  $-d_v \log \lambda_i^v(\underline{\underline{x}}) \geq -c_2$  and

$$(2.2) \quad \alpha_i^v(\underline{\underline{x}}) \geq -c_2 / \log H_k(\underline{\underline{x}}) \geq -c_2.$$

(We ignore points  $\underline{\underline{x}} \in \mathcal{L}$ , which are not dense.) Choose  $i = i(v) = i(v, \underline{\underline{x}})$  in  $1 \leq i(v) \leq a(v)$  for which  $(\lambda_i^v(\underline{\underline{x}}))^{d_v}$  is minimal. By the Subspace Theorem,  $\prod_v (\lambda_{i(v)}^v(\underline{\underline{x}}))^{d_v} > H_k(\underline{\underline{x}})^{-n-\varepsilon}$ , except for a set which is not dense. Therefore

$$(2.3) \quad \sum_v \left( \max_{1 \leq i \leq a(v)} \alpha_i^v(\underline{\underline{x}}) \right) \leq n + \varepsilon,$$

except for a set which is not dense. We may conclude that  $A \subseteq \mathbb{R}^a$  is contained in the compact set

$$(2.4) \quad \alpha_i^v \geq 0, \quad \sum_v \left( \max_i \alpha_i^v \right) \leq n,$$

and therefore  $A$  itself is compact. From (2.2), (2.3) we gather that  $|\underline{\underline{\alpha}}(\underline{\underline{x}})| \leq c_3$  for  $\underline{\underline{x}}$  outside some proper closed subset of  $k^n$ , where  $|\cdot|$  is the maximum norm on  $\mathbb{R}^a$ . Lemma 1 remains true for every *closed* set  $C$ , for  $\underline{\underline{\alpha}}(\underline{\underline{x}}) \in C$  is the same as  $\underline{\underline{\alpha}}(\underline{\underline{x}}) \in C_0$ , where  $C_0$  is the intersection of  $C$  and the cube  $|\underline{\underline{\alpha}}| \leq c_3$ .

By a *convex polyhedron* in  $\mathbb{R}^a$  we will understand a nonempty compact set which is the intersection of a finite number of closed half-spaces, i.e., sets  $g(\underline{\underline{\alpha}}) \leq 0$  where  $g$  is a linear form plus a constant. It is the convex hull of a finite number of vertices. By a *polyhedron* we will understand a finite union of convex polyhedra.

**Theorem 2.** *The approximation set  $A$  is a polyhedron contained in the set (2.4). It is a finite union of convex polyhedra whose vertices have rational coordinates, such that the components of each vertex  $\underline{\alpha}$  have a common denominator*

$$(2.5) \quad q(\underline{\alpha}) \leq (n\sqrt{s})^a.$$

*Given an open set  $O$  intersecting  $A$ , given  $\varepsilon > 0$  and  $H > H_0$  (where  $H_0$  may depend on all the data), there are  $n$  linearly independent points  $\underline{x}_1, \dots, \underline{x}_n$  in  $k^n \setminus \mathcal{L}$  with*

$$(2.6) \quad \underline{\alpha}(\underline{x}_l) \in O \quad \text{and} \quad H < H_k(\underline{x}_l) < H^{1+\varepsilon} \quad (l = 1, \dots, n).$$

**Theorem 2'.** *When  $n = 2$ , the vertices of  $A$  are integer points.*

I like to call the last assertion of Theorem 2 the *constancy principle*, since it asserts the existence of certain points of height about  $H$ , for all large values of  $H$ . The approximation set  $A$  will from now on be called the *approximation polyhedron*.

**Corollary.** *Define  $r$  as the supremum of the numbers  $\rho$  with  $\bar{X}(\rho) = k^n$ . Then  $r \in \mathfrak{R}$ . When  $n = 2$ , then  $r \in \mathbb{Z}$ ,  $0 \leq r \leq 2a^*$ .*

*Proof.* Let  $V(\rho)$  be the open half-space in  $\mathbb{R}^a$  given by

$$\sum_v \sum_i \alpha_i^v > \rho,$$

and let  $\overline{V(\rho)}$  be its closure. Let  $r'$  be the largest number  $\rho$  such that  $\overline{V(\rho)}$  meets  $A$ . We claim that

$$(2.7) \quad r = r'.$$

For  $\underline{x} \in k^n \setminus \mathcal{L}$ , i.e., when  $f(\underline{x}) > 0$  and  $\underline{x}$  lies outside the finitely many 1-dimensional subspaces where  $H_k(\underline{x}) < e$ , the condition (1.6) means that

$$(\log f(\underline{x})) / \log H_k(\underline{x}) < -\rho.$$

This says precisely that  $\underline{\alpha}(\underline{x}) \in V(\rho)$ , and implies that  $\underline{\alpha}(\underline{x}) \in \overline{V(\rho)}$ . When  $\rho > r'$ , then  $\overline{V(\rho)}$  does not meet  $A$ , and the set of  $\underline{x}$  with  $\underline{\alpha}(\underline{x}) \in \overline{V(\rho)}$  is not dense. When  $\rho < r'$ , the open half-space  $V(\rho)$  does meet  $A$ , and therefore the points  $\underline{x} \in k^n \setminus \mathcal{L}$  with  $\underline{\alpha}(\underline{x}) \in V(\rho)$  lie dense in  $k^n$ . Thus (2.7) is true.

It is clear that there is a vertex  $\underline{\alpha}$  of  $A$  (a vertex of one of the convex polyhedra making up  $A$ ) with  $\sum_v \sum_i \alpha_i^v = r$ . Since the coordinates of  $\underline{\alpha}$  have a common denominator  $q(\underline{\alpha})$  satisfying (2.5), and in view of (2.4), indeed  $r \in \mathfrak{R}$ . When  $r = 2$ , the extra conclusion,  $r \in \mathbb{Z}$  also follows. In particular Theorem 1' (i) follows.

The Corollary also follows from Theorems 1, 1'. However, we shall derive Theorems 1, 1' from Theorems 2, 2' using ideas which went into the Corollary.

### 3. DEDUCTION OF THEOREM 1 FROM THEOREM 2

Suppose  $T$  is a subspace of  $k^n$  of positive dimension. Then we can again define a subspace topology on  $T$  (more precisely on  $T \setminus \underline{0}$ ).

**Lemma 2.** Let  $X \subset k^n \setminus \underline{0}$ , and let  $T$  be a component of  $\bar{X}$ . Then  $T = \overline{T \cap \bar{X}}$ , where it does not matter whether on the right-hand side we take the closure with respect to the subspace topology on  $k^n$  or on  $T$ .

*Proof.* Initially let the bar denote closure with respect to the topology on  $k^n$ . Let  $T_1, \dots, T_l$  be the components of  $\bar{X}$ , and let  $T'_j = \overline{T_j \cap \bar{X}}$  ( $j = 1, \dots, l$ ). Then  $X \subseteq \bigcup_{j=1}^l T'_j \subseteq \bar{X}$ , and  $\bigcup_{j=1}^l T'_j$  is closed, so that in fact it equals  $\bar{X}$ . Now  $T'_j \subseteq T_j$  yields  $T'_j = T_j$ , so that  $T_j = \overline{T_j \cap \bar{X}}$  ( $j = 1, \dots, l$ ), and in particular  $T = \overline{T \cap \bar{X}}$ .

A subset of  $T$  is closed with respect to the topology of  $k^n$  if and only if it is a finite union of subspaces, therefore if and only if it is closed with respect to the topology on  $T$ . Whence the last assertion of the lemma.

Let  $T$  be a subspace of  $k^n$  with  $\dim T > 1$  where none of the forms  $L_i^v$  vanish identically. We can look at the restrictions of our forms  $L_i^v$  to  $T$  and we can define  $A = A(T)$  and  $r = r(T)$  in terms of the subspace topology of  $T$ . The following remark should be made here. Given a basis  $\mathfrak{B} = \{\underline{b}_1, \dots, \underline{b}_t\}$  of  $T$ , we may write  $\underline{x} \in T$  as  $\underline{x} = y_1 \underline{b}_1 + \dots + y_t \underline{b}_t$ , and the restriction of  $L_i^v$  to  $T$  becomes a linear form in  $\underline{y} = (y_1, \dots, y_t)$ . We may set  $|\underline{x}|_{v\mathfrak{B}} = |\underline{y}|_v = \max(|y_1|_v, \dots, |y_t|_v)$ , and we may define a height  $H_{k\mathfrak{B}}(\underline{x})$  accordingly. However, it is easily seen that  $|\underline{x}|_v \gg \ll |\underline{x}|_{v\mathfrak{B}}$ , in fact  $|\underline{x}|_v = |\underline{x}|_{v\mathfrak{B}}$  for all but finitely many  $v \in M(k)$ , so that  $H_k(\underline{x}) \gg \ll H_{k\mathfrak{B}}(\underline{x})$ . Since (in view of  $\dim T > 1$ ) a dense set on  $T$  contains infinitely many nonproportional elements  $\underline{x}$ , it now follows that in the definition of  $A(T)$  and  $r(T)$ , it does not matter whether we use  $|\underline{x}|_v$ ,  $H_k(\underline{x})$  or  $|\underline{x}|_{v\mathfrak{B}}$ ,  $H_{k\mathfrak{B}}(\underline{x})$  for any given basis  $\mathfrak{B}$  of  $T$ .

We will now derive Theorem 1 from Theorem 2. Let  $T$  be a component of  $\bar{X}'(\rho)$ . By Lemma 2, the solutions of (1.6) with  $\underline{x} \in T$  are dense in  $T$ . This implies that  $r(T) \geq \rho$ . Conversely, when

$$(3.1) \quad \dim T > 1 \quad \text{and} \quad r(T) > \rho,$$

then  $T \subseteq \bar{X}'(\rho)$ . By the Corollary,  $r(T) \in \mathfrak{R}$ . Therefore when  $\rho \notin \mathfrak{R}$ , then  $\bar{X}'(\rho)$  is the union of the subspaces  $T$  with (3.1). It is constant when  $\rho$  ranges through an interval disjoint from  $\mathfrak{R}$ .

We still have to derive the bound (1.8) on the number of components. Suppose  $\rho \notin \mathfrak{R}$ . Every  $\underline{x}$  has  $\lambda_i^v(\underline{x}) \ll 1$  and hence  $f(\underline{x}) \ll 1$ , so that  $X(\rho) = k^n \setminus \mathcal{L}$  when  $\rho < 0$ . We may therefore suppose that  $\rho > 0$ .

We had seen that  $\bar{X}'(\rho)$  is the union of the subspaces  $T$  with (3.1). Therefore the components of  $\bar{X}'(\rho)$  are the subspaces with (3.1) which are maximal, i.e., they are not properly contained in another subspace with (3.1).

Let  $T$  be such a component. Then  $r(T) > \rho$ , and there is a vertex  $\underline{\alpha}(T)$  of  $A(T)$  with  $\sum_v \sum_i \alpha_i^v = r(T)$ . The number of possibilities for  $\underline{\alpha}(T)$  is finite: We have  $0 \leq \alpha_i^v(T) \leq n$  by (2.4). Given a common denominator  $q$ , the number of possibilities for  $\underline{\alpha}(T)$  is therefore  $\leq (nq + 1)^a$ , and summation over  $q$  in  $1 \leq q \leq (n\sqrt{s})^a$  gives

$$(3.2) \quad \leq (n(ns)^a + 1)^a (ns)^{2a} \leq (ns)^{2a^2}$$

possibilities. (We may suppose  $a > 1$ , since  $A(T)$  for  $a = 1$  turns out to be an interval, hence has two vertices.) The bound in (1.8) is then a consequence of the following

**Lemma 3.** *If  $T_1, T_2$  are components with  $\underline{\alpha}(T_1) = \underline{\alpha}(T_2)$ , then  $T_1 = T_2$ .*

*Proof.* We begin with a general remark. Let  $M_\infty(k), M_0(k)$  be the set of archimedean and nonarchimedean absolute values, respectively. A point  $\underline{y}$  is called an *integer point* if  $|\underline{y}|_v \leq 1$  for every  $v \in M_0(k)$ . If we have unique factorization in  $k$ , then every point  $\underline{x} \in k^n \setminus \underline{0}$  is proportional to a point  $\underline{y}$  with  $|\underline{y}|_v = 1$  for every  $v \in M_0(k)$ . In general, the class number is finite, and  $\underline{x}$  is proportional to an integer point  $\underline{y}$  with

$$\prod_{v \in M_0(k)} |\underline{y}|_v^{d_v} > c_1,$$

where  $c_1 = c_1(k) > 0$ . Then

$$H_k(\underline{y}) \leq \prod_{v \in M_\infty(k)} |\underline{y}|_v^{d_v} < c_1^{-1} H_k(\underline{y}).$$

By multiplication with a suitable unit we get an integer point  $\underline{y} \neq \underline{0}$  having

$$c_2^{-1} H_k(\underline{y})^{1/d} < |\underline{y}|_v < c_2 H_k(\underline{y})^{1/d} \quad (v \in M_\infty(k))$$

where  $d = [k : \mathbb{Q}]$ . Such a point will be called *primitive*; every  $\underline{x} \in k^n \setminus \underline{0}$  is proportional to a primitive point.

Let  $T_1$  be a component of  $\bar{X}'(\rho)$  of dimension  $t_1$ , and let  $\underline{\alpha} = \underline{\alpha}(T_1)$ . We see from the constancy principle of Theorem 2 that when  $0 < \varepsilon < 1$  and  $H > H_0$ , there are  $t_1$  linearly independent points  $\underline{x} \in T_1$ ,  $\underline{x} \notin \mathcal{L}$  with

$$|\underline{\alpha}(\underline{x}) - \underline{\alpha}| < \varepsilon \quad \text{and} \quad H < H_k(\underline{x}) < H^{1+\varepsilon}.$$

In fact there are primitive points with this property. Let  $\underline{x}$  be such a point. Then

$$(3.3) \quad c_3^{-1} H^{\Delta_v/d} < |\underline{x}|_v < c_3 H^{\Delta_v/d+\varepsilon} \quad (v \in S),$$

where we set  $\Delta_v = 1$  when  $v \in M_\infty(k)$ , and  $\Delta_v = 0$  otherwise. Moreover, when  $H$  is large,

$$(3.4) \quad |L_i^v(\underline{x})|_v = |\underline{x}|_v \lambda_i^v(\underline{x}) = |\underline{x}|_v H_k(\underline{x})^{-\alpha_i^v(\underline{x})/d_v} < |\underline{x}|_v H_k(\underline{x})^{-(\alpha_i^v/d_v)+\varepsilon} < H^{-(\alpha_i^v/d_v)+(\Delta_v/d)+3\varepsilon}.$$

Let  $T_1, T_2$  be components of  $\bar{X}'(\rho)$  with  $\underline{\alpha}(T_1) = \underline{\alpha}(T_2) = \underline{\alpha}$ . For  $H > H_0$ , there will be  $t_1 = \dim T_1$  independent integer points in  $T_1 \setminus \mathcal{L}$  with (3.3), (3.4), and similarly there will be  $t_2 = \dim T_2$  such independent integer points in  $T_2$ . Thus if  $\dim(T_1 + T_2) = t$ , there will be  $t$  linearly independent integer points  $\underline{x}_1, \dots, \underline{x}_t$  in  $T = T_1 + T_2$  with (3.3), (3.4). Consider points

$$(3.5) \quad \underline{z} = u_1 \underline{x}_1 + \dots + u_t \underline{x}_t$$

where  $u_1, \dots, u_t$  lie in  $\mathbb{Z}$  with  $1 \leq u_j \leq H^\varepsilon$  ( $j = 1, \dots, t$ ). These points have

$$(3.6) \quad |\underline{z}|_v < c_4 H^{(\Delta_v/d)+2\varepsilon} \quad (v \in S),$$

$$(3.7) \quad |L_i^v(\underline{z})|_v < c_4 H^{-(\alpha_i^v/d_v)+(\Delta_v/d)+4\varepsilon} \quad (v \in S, 1 \leq i \leq a(v)).$$

We claim that a positive proportion of the points (3.5) has

$$(3.8) \quad |\underline{z}|_v > c_5 H^{\Delta_v/d} \quad (v \in S).$$

Suppose  $u_2, \dots, u_t$  are given. Since  $\underline{x}_1$  satisfies (3.3), we see that there is at most one value  $u_1 \in \mathbb{Z}$  for which  $|\underline{z}|_v < \frac{1}{2} c_3^{-1} H^{1/d}$  when  $v \in M_\infty(k)$ . When  $v \in M_0(k)$  extends the  $p$ -adic absolute value, and  $m$  is natural, there is at most one residue class of  $u_1$  modulo  $p^m$  where  $|\underline{z}|_v < p^{-m} c_3^{-1}$ . This proves the claim.

Let  $Z$  be the set of points  $\underline{z}$  of type (3.5) with (3.8). For  $\underline{z} \in Z$ ,

$$\lambda_i^v(\underline{z}) < c_6 H^{-(\alpha_i^v/d_v)+4\varepsilon},$$

therefore

$$(3.9) \quad f(\underline{z}) < c_7 H^{-r+4d\varepsilon},$$

where  $r = r(T_1) = r(T_2) = \sum_v \sum_i \alpha_i^v$ . Recall that  $r > \rho > 0$ . Now (3.6) yields  $H_k(\underline{z}) < c_8 H^{1+2d\varepsilon}$ , so that (3.9) gives

$$f(\underline{z}) < H^{-\rho(1+3d\varepsilon)} < H_k(\underline{z})^{-\rho}$$

if  $\varepsilon$  is sufficiently small and  $H$  sufficiently large. The set  $Z$  contains a positive proportion of the points (3.5), and therefore will not be contained in any given finite collection of proper subspaces of  $T$  if  $H$  is sufficiently large. Therefore points  $\underline{z}$  with  $f(\underline{z}) < H_k(\underline{z})^{-\rho}$  are dense in  $T$ , so that  $r(T) \geq \rho$ , in fact  $r(T) > \rho$  since  $\rho \notin \mathfrak{R}$ . Since  $T_1, T_2$  were maximal with (3.1), we have  $T_1 = T_2 = T$ .

*Remark.* It might not be hard to prove  $A(T_1 + T_2) \supset A(T_1) \cap A(T_2)$ , which would imply the lemma.

#### 4. THE MODIFIED APPROXIMATION POLYHEDRON

With the same hypotheses and the same notation as in §2, set

$$\beta_i^v(\underline{x}) = \min(\alpha_1^v(\underline{x}), \dots, \alpha_i^v(\underline{x}))$$

for  $\underline{x} \in k^n \setminus \mathcal{L}$ . Let  $\underline{\beta}(\underline{x}) \in \mathbb{R}^a$  be the point with components  $\beta_i^v(\underline{x})$ . A point  $\underline{\beta} \in \mathbb{R}^a$  will be called a *modified approximation point* if for every open neighborhood  $O$  of  $\underline{\beta}$ , the points  $\underline{x} \in k^n \setminus \mathcal{L}$  with  $\underline{\beta}(\underline{x}) \in O$  are dense. The *modified approximation set*  $B$  is the set of modified approximation points. A point  $\underline{\beta} \in B$  has

$$(4.1) \quad \beta_1^v \geq \dots \geq \beta_{a(v)}^v \geq 0 \quad (v \in S),$$

where the last inequality follows like  $\alpha_i^v \geq 0$  in (2.4). Also, just as in (2.4), we get

$$(4.2) \quad \sum_v \beta_1^v \leq n$$

for  $\underline{\beta} \in B$ . Just as in Lemma 1 we see that  $B$  is compact. Further if  $C \subset \mathbb{R}^a$  is a closed set disjoint from  $B$ , the points  $\underline{x}$  with  $\underline{\beta}(\underline{x}) \in C$  are not dense.

We will initially describe the modified approximation set under special hypotheses. In order to avoid confusion later on, we will introduce a different notation to deal with this special case. For  $k \subseteq K$ ,  $v \in M(k)$  and a given extension of  $|\cdot|_v$  from  $k$  to  $K$ , let  $k^v$  be the completion of  $k$  in  $K$  with respect to  $|\cdot|_v$ . Thus  $k^v$  consists of elements of  $K$  which are limits of elements of  $k$  under  $|\cdot|_v$ . Then  $k \subseteq k^v \subseteq K$ , and  $k^v$  is a field with  $[K : k^v] = \delta_v$ , the local degree of  $K$  over  $k$ .

Suppose now that for each  $v \in S$  we are given linearly independent linear forms  $M_1^v, \dots, M_{b(v)}^v$  in  $n$  variables with coefficients in  $k^v$ . In analogy to (2.1) we put  $\mu_j^v(\underline{x}) = |M_j^v(\underline{x})|_v / |\underline{x}|_v$ . The analogues of  $\alpha_j^v(\underline{x})$ ,  $\underline{\alpha}(\underline{x})$ ,  $\beta_j^v(\underline{x})$ ,  $\underline{\beta}(\underline{x})$  will be denoted by  $\hat{\alpha}_j^v(\underline{x})$ ,  $\hat{\underline{\alpha}}(\underline{x})$ ,  $\hat{\beta}_j^v(\underline{x})$ ,  $\hat{\underline{\beta}}(\underline{x})$ . Here and below,  $v \in S$  and  $1 \leq j \leq b(v)$ . The points  $\hat{\underline{\alpha}}(\underline{x})$ ,  $\hat{\underline{\beta}}(\underline{x})$  lie in  $\mathbb{R}^b$ , where  $b = \sum_v b(v)$ . They are defined for  $\underline{x} \in k^n \setminus \mathfrak{M}$ , where  $\mathfrak{M}$  consists of points  $\underline{x}$  with  $H_k(\underline{x}) < e$  or with  $M_j^v(\underline{x}) = 0$  for some  $v, j$ . Define  $\hat{A}, \hat{B}$  in analogy to  $A, B$ .

Linear forms in  $n$  variables with coefficients in  $k^v$  form a vector space. Let  $\mathfrak{T}_j^v$  ( $v \in S$ ,  $1 \leq j \leq b(v)$ ) be the space of linear forms spanned by  $M_1^v, \dots, M_j^v$ , so that  $\mathfrak{T}_1^v \subset \mathfrak{T}_2^v \subset \dots \subset \mathfrak{T}_{b(v)}^v$ , and  $\dim \mathfrak{T}_j^v = j$ .

Let  $T$  be a subspace of  $k^n$  of codimension  $c > 0$ . Let  $\mathfrak{T}$  be the space of linear forms vanishing on  $T$ ; then  $\dim \mathfrak{T} = c$ . Put  $g(v) = \dim(\mathfrak{T} \cap \mathfrak{T}_{b(v)}^v)$ , so that  $g(v) \leq \min(c, b(v))$ . When  $g(v) > 0$ , let  $j_1(v)$  be least with  $\dim(\mathfrak{T} \cap \mathfrak{T}_{j_1(v)}^v) = 1$ , then  $j_2(v)$  least with  $\dim(\mathfrak{T} \cap \mathfrak{T}_{j_2(v)}^v) = 2$ , etc. We obtain numbers  $1 \leq j_1(v) < j_2(v) < \dots < j_{g(v)}(v) \leq b(v)$ . Let  $\Omega^v = \Omega^v(T)$  be the set  $\{j_1(v), \dots, j_{g(v)}(v)\}$ . When  $g(v) = 0$ , set  $\Omega^v = \emptyset$ . In every case,  $\Omega^v(T)$  is a subset of  $\{1, \dots, b(v)\}$ .

Let  $\Phi^v = \Phi^v(T)$  be the complement of  $\Omega^v$  in  $\{1, \dots, b(v)\}$ . If  $M_j^{vT}$  denotes the restriction of  $M_j^v$  to  $T$ , then for given  $v$  the forms  $M_j^{vT}$  with  $j \in \Phi^v$  are linearly independent, and each  $M_j^{vT}$  is a linear combination of these  $|\Phi^v|$  forms.

**Theorem 3.** *The modified approximation set  $\hat{B}$  of the forms  $M_j^v$  is a convex polyhedron. It is defined by*

$$(4.3) \quad \hat{\beta}_1^v \geq \dots \geq \hat{\beta}_{b(v)}^v \geq 0,$$

$$(4.4) \quad \hat{\beta}_n^v = 0 \quad \text{for } v \text{ with } b(v) = n,$$

and

$$(4.5) \quad \sum_v \sum_{j \in \Omega^v(T)} \hat{\beta}_j^v \leq \text{codim } T$$

for every subspace  $T$  of  $k^n$ . The constancy principle holds: Given an open set  $\hat{O} \subset \mathbb{R}^b$  intersecting  $\hat{B}$ , given  $\varepsilon > 0$  and  $H > H_1$ , there are independent points  $\underline{x}_1, \dots, \underline{x}_n \in k^n$  with  $\hat{\underline{\beta}}(\underline{x}_l) \in \hat{O}$  and  $H < H_k(\underline{x}_l) < H^{1+\varepsilon}$  ( $l = 1, \dots, n$ ).

*Remarks.* Although there are infinitely many subspaces  $T \subset k^n$ , they lead to only finitely many different conditions (4.5) for  $\hat{\underline{\beta}}$ . When  $T = k^n$ , each

$\Omega^v(T) = \emptyset$ , and the empty sum in (4.5) is to be interpreted as 0. Therefore (4.5) with  $T = k^n$  imposes no condition on  $\underline{\hat{\beta}}$ . When  $T = \underline{0}$ , then  $\Omega^v(T) = \{1, \dots, b(v)\}$ , and (4.5) becomes

$$(4.6) \quad \sum_v \sum_j \hat{\beta}_j^v \leq n.$$

This is stronger than (4.2), but recall that we have extra hypotheses, namely that  $M_1^v, \dots, M_{b(v)}^v$  are linearly independent and have coefficients in  $k^v$ . It is now clear that  $\hat{B}$  is a convex polyhedron in  $\mathbb{R}^b$ .

Here and in §§5, 6, 7, 8 we will derive consequences of Theorem 3. The proof of Theorem 3 is postponed to the end, to §§9, 10, 11, 12.

Let  $L_1^v, \dots, L_{a(v)}^v$  be the linear forms we had at the beginning. Let  $\{b_l^v\}$ ,  $1 \leq l \leq \delta_v$ , be a basis of  $K$  over  $k^v$ . It is well known that for  $\xi_1, \dots, \xi_{\delta_v}$  in  $k^v$ ,

$$(4.7) \quad |\xi_1 b_1 + \dots + \xi_{\delta_v} b_{\delta_v}|_v \gg \ll \max(|\xi_1|_v, \dots, |\xi_{\delta_v}|_v).$$

We may write  $L_i^v(\underline{x})$  uniquely as

$$L_i^v(\underline{x}) = \sum_{l=1}^{\delta_v} M_{il}^v(\underline{x}) b_l^v \quad (v \in S, 1 \leq i \leq a(v))$$

where the  $M_{il}^v$  are linear forms with coefficients in  $k^v$ . Let  $\mathfrak{V}_i^v$  where  $1 \leq i \leq a(v)$  be the vector space of linear forms spanned by the forms  $M_{jl}^v$  with  $1 \leq j \leq i$ ,  $1 \leq l \leq \delta_v$ . Let  $M_1^v, \dots, M_{w(v,1)}^v$  be a basis of  $\mathfrak{V}_1^v$ . Extend it to a basis  $M_1^v, \dots, M_{w(v,1)}^v, \dots, M_{w(v,2)}^v$  of  $\mathfrak{V}_2^v$ , where  $w(v, 2) = w(v, 1)$  if  $\mathfrak{V}_2^v = \mathfrak{V}_1^v$ . Continuing in this manner, pick

$$M_1^v, \dots, M_{b(v)}^v \quad \text{where } b(v) = w(v, a(v))$$

such that  $M_1^v, \dots, M_{w(v,i)}^v$  is a basis of  $\mathfrak{V}_i^v$  ( $1 \leq i \leq a(v)$ ). We have

$$1 \leq w(v, 1) \leq \dots \leq w(v, a(v)) = b(v) \leq n.$$

As above, let  $\hat{B} \in \mathbb{R}^b$  be the modified approximation polyhedron of the forms  $M_j^v$  ( $v \in S$ ,  $1 \leq j \leq b(v)$ ). We will exhibit a relation between  $B$  and  $\hat{B}$ .

Let  $p$  be the linear map  $\mathbb{R}^b \rightarrow \mathbb{R}^a$  with  $p(\underline{\hat{\beta}}) = \underline{\beta}$  where  $\beta_i^v = \hat{\beta}_{w(v,i)}^v$  ( $1 \leq i \leq a(v)$ ). Intuitively I like to think of  $p$  as of a projection, but observe that when  $w(v, i) = w(v, i+1) = \dots = w(v, l-1)$ , the coordinate  $\hat{\beta}_{w(v,i)}^v$  gets "blown up" into the  $l-i$  coordinates  $\beta_i^v = \dots = \beta_{l-1}^v = \hat{\beta}_{w(v,i)}^v$ .

**Lemma 4.**  $B = p(\hat{B})$ . In particular,  $B$  is a convex polyhedron. The constancy principle holds: Given an open set  $O \subset \mathbb{R}^a$  intersecting  $B$ , given  $\varepsilon > 0$  and  $H > H_2$ , there are independent points  $\underline{x}_1, \dots, \underline{x}_n \in k^n \setminus \mathcal{L}$  with  $\underline{\beta}(\underline{x}_l) \in O$  and  $H < H_k(\underline{x}_l) < H^{1+\varepsilon}$  ( $l = 1, \dots, n$ ).

*Proof.* In view of (4.7) we have for  $\underline{x} \in k^n$ ,

$$(4.8) \quad \max(|L_1^v(\underline{x})|_v, \dots, |L_i^v(\underline{x})|_v) \gg \ll \max(|M_1^v(\underline{x})|_v, \dots, |M_{w(v,i)}^v(\underline{x})|_v).$$

Therefore for  $\underline{x} \in k^n \setminus \mathfrak{M}$  with large height  $H_k(\underline{x})$ ,

$$(4.9) \quad |\beta_i^v(\underline{x}) - \widehat{\beta}_{w(v,i)}^v(\underline{x})| < \varepsilon.$$

Note that  $\underline{x} \in k^n \setminus \mathfrak{M}$  has each  $M_j^v(\underline{x}) \neq 0$ , therefore each  $L_i^v(\underline{x}) \neq 0$ . But (4.9) is the same as

$$(4.10) \quad |\underline{\beta}(\underline{x}) - p(\widehat{\underline{\beta}}(\underline{x}))| < \varepsilon.$$

Suppose  $\widehat{\underline{\beta}} \in \widehat{B}$ , and set  $\underline{\beta} = p(\widehat{\underline{\beta}})$ . For any  $\varepsilon > 0$ , points  $\underline{x}$  with  $|\widehat{\underline{\beta}}(\underline{x}) - \widehat{\underline{\beta}}| < \varepsilon$  are dense. But then  $|p(\widehat{\underline{\beta}}(\underline{x})) - \underline{\beta}| < \varepsilon$  and  $|\underline{\beta}(\underline{x}) - \underline{\beta}| < 2\varepsilon$ . We may conclude that  $\underline{\beta} \in B$  and  $p(\widehat{B}) \subseteq B$ . Conversely, suppose  $\underline{\beta} \in B$ . We have to show that there is a  $\widehat{\underline{\beta}} \in \widehat{B}$  with  $p(\widehat{\underline{\beta}}) = \underline{\beta}$ . We first claim that

$$(4.11) \quad \beta_i^v = \beta_l^v \quad \text{for } i, l \text{ with } w(v, i) = w(v, l).$$

For if, say,  $i < l$ , the forms  $L_{i+1}^v, \dots, L_l^v$  lie in  $\mathfrak{V}_i^v$ , so that

$$(4.12) \quad \begin{aligned} \max(|L_1^v(\underline{x})|_v, \dots, |L_l^v(\underline{x})|_v) &\ll \max(|M_1^v(\underline{x})|_v, \dots, |M_{w(v,i)}^v(\underline{x})|_v) \\ &\ll \max(|L_1^v(\underline{x})|_v, \dots, |L_i^v(\underline{x})|_v), \end{aligned}$$

and  $\beta_l^v(\underline{x}) \geq \beta_i^v(\underline{x}) - \varepsilon$  for  $\underline{x} \in k^n \setminus \mathfrak{M}$  with large enough height  $H_k(\underline{x})$ . Therefore a point  $\underline{\beta} \in B$  has  $\beta_l^v \geq \beta_i^v$ , and since clearly  $\beta_l^v \leq \beta_i^v$  by the definition of  $\underline{\beta}(\underline{x})$ , we have (4.11). Define  $\underline{\beta}^*(\underline{x}) \in \mathbb{R}^b$  by

$$\beta_j^{*v}(\underline{x}) = \begin{cases} \beta_i^v & \text{if } j = w(v, i) \text{ for some } i, \\ \widehat{\beta}_j^v(\underline{x}) & \text{if } j \text{ is not of this type.} \end{cases}$$

The definition of  $\underline{\beta}^*(\underline{x})$  makes sense in view of (4.11). We have  $p(\underline{\beta}^*(\underline{x})) = \underline{\beta}$  and  $|\underline{\beta}^*(\underline{x})| \leq c_1$  for a certain constant  $c_1$ . In other words,  $\underline{\beta}^*(\underline{x})$  lies in the compact set  $C \subset \mathbb{R}^b$  consisting of  $\widehat{\underline{\beta}}$  with  $p(\widehat{\underline{\beta}}) = \underline{\beta}$  and  $|\widehat{\underline{\beta}}| \leq c_1$ .

Points  $\underline{x}$  with  $|\underline{\beta}(\underline{x}) - \underline{\beta}| < \varepsilon$  are dense, since  $\underline{\beta} \in B$ . For such points,

$$|\widehat{\beta}_{w(v,i)}^v(\underline{x}) - \beta_{w(v,i)}^{*v}(\underline{x})| = |\widehat{\beta}_{w(v,i)}^v(\underline{x}) - \beta_i^v| < 2\varepsilon$$

by (4.9). Therefore  $|\widehat{\underline{\beta}}(\underline{x}) - \underline{\beta}^*(\underline{x})| < 2\varepsilon$ , and  $\widehat{\underline{\beta}}(\underline{x})$  lies in the compact set  $C(2\varepsilon)$  consisting of  $\widehat{\underline{\beta}}$  which are within  $2\varepsilon$  of  $C$ , i.e., having  $|\widehat{\underline{\beta}} - \widehat{\gamma}| \leq 2\varepsilon$  for some  $\widehat{\gamma} \in C$ . By Lemma 1, which holds, mutatis mutandis, for  $\widehat{\underline{\beta}}(\underline{x})$  in place of  $\underline{\alpha}(\underline{x})$ , we may conclude that  $C(2\varepsilon) \cap \widehat{B} \neq \emptyset$ . Since this is true for every  $\varepsilon > 0$ , in fact  $C \cap \widehat{B} \neq \emptyset$ , and  $\underline{\beta} = p(\widehat{\underline{\beta}})$  for some  $\widehat{\underline{\beta}} \in C \cap \widehat{B}$ .

This proves  $B = p(\widehat{B})$ . The last assertion of the lemma follows easily from the last assertion of Theorem 3. For let  $\underline{\beta} \in B$ , so that  $\underline{\beta} = p(\widehat{\underline{\beta}})$  with  $\widehat{\underline{\beta}} \in \widehat{B}$ . For  $H > H_3$ , there are  $n$  independent points  $\underline{x} \in k^n \setminus \mathfrak{M}$  with  $H < H_k(\underline{x}) < H^{1+\varepsilon}$  and  $|\widehat{\underline{\beta}}(\underline{x}) - \widehat{\underline{\beta}}| < \varepsilon$ . Then by (4.10),

$$|\underline{\beta}(\underline{x}) - \underline{\beta}| \leq |\underline{\beta}(\underline{x}) - p(\widehat{\underline{\beta}}(\underline{x}))| + |p(\widehat{\underline{\beta}}(\underline{x})) - p(\widehat{\underline{\beta}})| < 2\varepsilon.$$

Since  $\varepsilon > 0$  is arbitrary, everything follows.

**Lemma 5.**  $B$  consists of  $\underline{\underline{\beta}} \in \mathbb{R}^a$  with (4.11), with

$$(4.13) \quad n \geq \beta_1^v \geq \cdots \geq \beta_{a(v)}^v \geq 0,$$

$$(4.14) \quad \beta_{a(v)}^v = 0 \quad \text{for } v \text{ with } b(v) = n,$$

and with

$$(4.15) \quad \sum_v \sum_i c_i^v(T) \beta_i^v \leq \text{codim } T$$

for every subspace  $T$  of  $k^n$ , where (with  $w(v, 0) = 0$ )

$$c_i^v(T) = \sum_{\substack{j \in \Omega^v(T) \\ w(v, i-1) < j \leq w(v, i)}} 1.$$

*Proof.* We begin with the necessity. The necessity of (4.11) has already been shown. (4.13) follows from (4.1) and (4.2). Now  $\underline{\underline{\beta}} = p(\widehat{\underline{\underline{\beta}}})$  with  $\widehat{\underline{\underline{\beta}}} \in \widehat{B}$ ; and when  $b(v) = n$ , then  $\beta_{a(v)}^v = \widehat{\beta}_{w(v, a(v))}^v = \widehat{\beta}_{b(v)}^v = \widehat{\beta}_n^v = 0$ , so that (4.14) holds. Also,

$$c_i^v(T) \beta_i^v = c_i^b(T) \widehat{\beta}_{w(v, i)}^v \leq \sum_{\substack{j \in \Omega^v(T) \\ w(v, i-1) < j \leq w(v, i)}} \widehat{\beta}_j^v,$$

so that (4.15) follows from (4.5).

Conversely, when  $\underline{\underline{\beta}}$  satisfies (4.11), (4.13), (4.14), (4.15), let  $\widehat{\underline{\underline{\beta}}}$  be given by

$$(4.16) \quad \widehat{\beta}_j^v = \beta_i^v \text{ for each } j \text{ with } w(v, i-1) < j \leq w(v, i).$$

Then  $\widehat{\underline{\underline{\beta}}}$  satisfies (4.3), (4.4), (4.5), so that  $\widehat{\underline{\underline{\beta}}} \in \widehat{B}$ , and  $\underline{\underline{\beta}} = p(\widehat{\underline{\underline{\beta}}}) \in p(\widehat{B}) = B$ .

## 5. DEDUCTION OF THEOREM 2 FROM THEOREM 3

Let  $A$  be the approximation set of the forms  $L_i^v$ , and  $A'$  the intersection of  $A$  with the cone

$$(5.1) \quad \alpha_1^v \geq \cdots \geq \alpha_{a(v)}^v \geq 0 \quad (v \in S).$$

We will show that  $A'$  is a convex polyhedron in  $\mathbb{R}^a$  contained in the set (2.4), whose vertices satisfy (2.5). The positive quadrant of  $\mathbb{R}^a$  is the union of  $\prod_v (a(v)!)^v$  obvious cones, and by symmetry the intersection of  $A$  with each cone is contained in (2.4), and is a convex polyhedron with vertices having (2.5). Therefore it will suffice to deal with  $A'$ .

For  $i, l$  in  $1 \leq i, l \leq a(v)$ , write  $i \overset{v}{\approx} l$  if  $i, l$  lie in some subinterval  $g \leq x \leq h$  such that  $L_h^v$  lies in  $\mathfrak{V}_g^v$ . Write  $i \overset{v}{\sim} l$  if there is a chain  $i = l_0, l_1, \dots, l_t = l$  with  $l_{z-1} \overset{v}{\approx} l_z$  for  $z = 1, \dots, t$ . Then  $\overset{v}{\sim}$  is an equivalence relation whose equivalence classes consist of integers in certain subintervals of  $1 \leq i \leq a(v)$ . Clearly  $w(v, i) = w(v, l)$  implies  $i \overset{v}{\sim} l$ .

**Lemma 6.**  $A'$  is given by

$$(5.2) \quad n \geq \alpha_1^v \geq \cdots \geq \alpha_{a(v)}^v \geq 0,$$

$$(5.3) \quad \alpha_{a(v)}^v = 0 \quad \text{for } v \text{ with } b(v) = n,$$

$$(5.4) \quad \sum_v \sum_i c_i^v(T) \alpha_i^v \leq \text{codim } T$$

for every subspace  $T \subseteq k^n$ , and by

$$(5.5) \quad \alpha_i^v = \alpha_l^v \quad \text{if } i \sim^v l.$$

The constancy principle holds: Given an open set  $O \subset \mathbb{R}^a$  which meets  $A'$ , given  $\varepsilon > 0$  and  $H > H_4$ , there are  $n$  independent points  $\underline{x}$  with  $\underline{\alpha}(\underline{x}) \in O$  and  $H < H_k(\underline{x}) < H^{1+\varepsilon}$ .

Note that the only difference with  $B$  (as given by (4.11), (4.13), (4.14), (4.15)) is that the condition (4.11) is replaced by the stronger condition (5.5). Put differently,  $A'$  is the intersection of  $B$  with the subspace determined by (5.5).

*Proof.* We begin with the necessity. When  $\underline{\alpha} \in A'$ , points  $\underline{x}$  with

$$(5.6) \quad |\underline{\alpha}(\underline{x}) - \underline{\alpha}| < \varepsilon$$

are dense. By definition,  $\beta_i^v(\underline{x}) \leq \alpha_i^v(\underline{x})$ . On the other hand, (5.1) and (5.6) yield  $\alpha_i^v(\underline{x}) < \alpha_l^v(\underline{x}) + 2\varepsilon$  when  $l \leq i$ , so that  $\beta_i^v(\underline{x}) > \alpha_i^v(\underline{x}) - 2\varepsilon$ , and  $|\underline{\beta}(\underline{x}) - \underline{\alpha}| < 3\varepsilon$ . Therefore, since such points  $\underline{x}$  are dense,  $\underline{\alpha} \in B$ . Thus  $A' \subseteq B$ . Moreover, if  $l \geq i$  and  $L_l^v$  lies in  $\mathfrak{V}_i^v$ , we have (similar to (4.8)),

$$|L_l^v(\underline{x})|_v \ll \max(|L_1^v(\underline{x})|_v, \dots, |L_i^v(\underline{x})|_v),$$

so that  $\alpha_l^v(\underline{x}) \geq \min(\alpha_1^v(\underline{x}), \dots, \alpha_i^v(\underline{x})) - \varepsilon > \alpha_i^v - 2\varepsilon$  when  $\underline{x} \in k^n \setminus \mathfrak{M}$  has large height. Therefore  $\alpha_l^v > \alpha_i^v - 3\varepsilon$  for each  $\varepsilon > 0$ , so that  $\alpha_i^v = \alpha_l^v$ . Thus (5.5) holds when  $i \approx^v l$ , therefore more generally when  $i \sim^v l$ .

We now come to the sufficiency. Suppose  $\underline{\alpha}$  satisfies (5.2), (5.3), (5.4), (5.5). We have  $\underline{\alpha} = p(\underline{\hat{\beta}})$  with  $\underline{\hat{\beta}} \in \widehat{B}$ . We may choose  $\underline{\hat{\beta}}$  with (4.16), i.e., with

$$(5.7) \quad \hat{\beta}_j^v = \alpha_i^v \quad \text{for each } j \text{ in } w(v, i-1) < j \leq w(v, i).$$

It is easily seen that there is a  $\underline{\hat{\gamma}} \in \widehat{B}$  with  $|\underline{\hat{\gamma}} - \underline{\hat{\beta}}| < \varepsilon$  and  $\hat{\gamma}_j^v < \hat{\gamma}_g^v$  for every  $v, g, j$  with  $1 \leq g < j \leq b(v)$ . Choose  $\eta$  in  $0 < \eta < \varepsilon$  so small that

$$(5.8) \quad \hat{\gamma}_j^v < \hat{\gamma}_g^v - \eta \quad \text{when } g < j.$$

Given  $L_l^v$ , let  $i$  be least such that  $L_l^v \in \mathfrak{V}_i^v$ ; then

$$(5.9) \quad L_l^v = \sum_{j=1}^{w(v, i)} c_j M_j^v,$$

and the largest subscript  $j_0$  with  $c_{j_0} \neq 0$  has  $w(v, i-1) < j_0 \leq w(v, i)$ . Consider points  $\underline{x}$  with

$$(5.10) \quad |\underline{\hat{\beta}}(\underline{x}) - \underline{\hat{\gamma}}| < \eta/3.$$

For such points,  $g < j$  yields  $\hat{\beta}_g^v(\underline{x}) > \hat{\beta}_j^v(\underline{x}) + \eta/3$  (because of (5.8)), so that  $|M_j^v(\underline{x})|_v$  is of larger order of magnitude than  $|M_g^v(\underline{x})|_v$ , and by (5.9),  $|L_l^v(\underline{x})|_v$  is of the order of magnitude of  $|M_{j_0}^v(\underline{x})|_v$ , so that

$$|\log |L_l^v(\underline{x})|_v - \log |M_{j_0}^v(\underline{x})|_v| < c_2.$$

Thus when  $H_k(\underline{x})$  is large,  $|\alpha_l^v(\underline{x}) - \hat{\alpha}_{j_0}^v(\underline{x})| < \varepsilon$ . On the other hand  $\hat{\beta}_1^v(\underline{x}) > \hat{\beta}_2^v(\underline{x}) > \dots$  yields  $\hat{\alpha}(\underline{x}) = \hat{\beta}(\underline{x})$ . We obtain

$$\begin{aligned} |\alpha_l^v(\underline{x}) - \hat{\beta}_{j_0}^v| &< |\alpha_l^v(\underline{x}) - \hat{\beta}_{j_0}^v(\underline{x})| + |\hat{\beta}_{j_0}^v(\underline{x}) - \hat{\gamma}_{j_0}^v| + |\hat{\gamma}_{j_0}^v - \hat{\beta}_{j_0}^v| \\ &< \varepsilon + (\eta/3) + \varepsilon < 3\varepsilon. \end{aligned}$$

Since  $w(v, i-1) < j_0 \leq w(v, i)$ , (5.7) gives  $\hat{\beta}_{j_0}^v = \alpha_i^v = \alpha_l^v$ , since  $l \approx i$ . Therefore

$$(5.11) \quad |\underline{\alpha}(\underline{x}) - \underline{\alpha}| < 3\varepsilon.$$

Points with (5.10), therefore with (5.11), are dense, so that indeed  $\underline{\alpha} \in A'$ . Moreover, given  $H > H_4$ , there are  $n$  independent points  $\underline{x}$  with  $H < H_k(\underline{x}) < H^{1+\varepsilon}$  and with (5.10), hence (5.11). This proves the constancy principle.

This leaves us to deal with the bound (2.5). The vertices of  $A'$  are given as intersections of certain linear manifolds obtained by having equality sign in  $a$  of the equalities and inequalities (5.2), (5.3), (5.4), (5.5). Each coefficient vector (coefficients of the  $\alpha_i^v$ 's) lies in  $\mathbb{Z}^a$  and has Euclidean length  $\leq n\sqrt{s}$ ; for notice that in particular

$$\sum_v \sum_i c_i^v(T)^2 \leq \sum_v b(v)^2 \leq sn^2.$$

The determinant of the coefficient vectors is  $\leq (n\sqrt{s})^a$  by Hadamard's Theorem, so that the solutions of the system of equations are rational with denominator satisfying (2.5).

## 6. LOW DIMENSIONAL CASES

We will deduce Theorems 1' and 2'. We begin with a study of the approximation polyhedron when

$$(6.1) \quad n = 2.$$

First, what can we say under the hypotheses of Theorem 3? We have  $b(v) \leq 2$ . For given  $v \in S$ , either  $b(v) = 1$ , or  $b(v) = 2$  and  $\hat{\beta}_2^v = 0$  by (4.4). Therefore only the coordinates  $\hat{\beta}_1^v$  are of interest. We have  $\hat{\beta}_1^v \geq 0$ , and

$$(6.2) \quad \sum_v \hat{\beta}_1^v \leq 2$$

by (4.6). We yet have to interpret (4.5) for 1-dimensional spaces  $T \subset k^2$ . For each  $v \in S$ , the subspace  $M_1^v = 0$  will either not be defined over  $k$ , or will be among a finite set of distinct 1-dimensional subspaces  $T_1, \dots, T_m$  defined over  $k$ . Write  $S = S_0 \cup S_1 \cup \dots \cup S_m$ , where  $v \in S_0$  if  $M_1^v = 0$  is not defined

over  $k$ , and  $v \in S_l$ ,  $1 \leq l \leq m$ , if the space  $M_1^v = 0$  is  $T_l$ . Now if  $T$  is any 1-dimensional subspace of  $k^2$ ,

$$\Omega^v(T) = \begin{cases} \{2\} & \text{when } v \in S_0, \\ \{2\} & \text{when } v \in S_l, \ 1 \leq l \leq m, \text{ but } T \neq T_l, \\ \{1\} & \text{when } v \in S_l, \ 1 \leq l \leq m, \text{ with } T = T_l. \end{cases}$$

The condition (4.5) for spaces  $T \subset k^2$  of dimension 1 becomes

$$(6.3) \quad \sum_{v \in S_l} \hat{\beta}_1^v \leq 1 \quad (l = 1, \dots, m).$$

Secondly, we turn to the situation of Lemma 4. Set  $t(v) = 0$  when  $w(v, 1) = 2$ ; otherwise let  $t(v)$  be the largest number with  $w(v, 1) = \dots = w(v, t(v)) = 1$ . Thus if  $p(\hat{\underline{\beta}}) = \underline{\beta}$  with  $\hat{\underline{\beta}} \in \hat{B}$ , then

$$\beta_i^v = \hat{\beta}_1^v \text{ for } 1 \leq i \leq t(v), \text{ and } \beta_i^v = 0 \text{ for } t(v) < i \leq a(v).$$

Put differently, the vector  $\underline{\beta}^v := (\beta_1^v, \dots, \beta_{a(v)}^v)$  is of the type

$$(6.4) \quad \underline{\beta}^v = (\underset{\leftarrow t(v) \rightarrow}{\hat{\beta}_1^v}, \dots, \underset{\leftarrow t(v) \rightarrow}{\hat{\beta}_1^v}, 0, \dots, 0).$$

Thus  $B$  consists of  $\underline{\beta}$  with (6.4), (6.2), (6.3).

Thirdly, consider  $A'$ . Let  $y(v)$  be largest such that  $L_{y(v)}$  lies in  $\mathfrak{W}_1^v$ . Then by (5.5),  $\alpha_i^v = \alpha_l^v$  if  $1 \leq i, l \leq y(v)$ . When  $y(v) > t(v)$ , then  $\underline{\alpha} \in A'$  has  $\underline{\alpha}^v := (\alpha_1^v, \dots, \alpha_{a(v)}^v) = (0, \dots, 0)$ . Setting

$$z(v) = \begin{cases} t(v) & \text{if } y(v) = t(v), \\ 0 & \text{if } y(v) > t(v), \end{cases}$$

we have

$$(6.5) \quad \underline{\alpha}^v = (\underset{\leftarrow z(v) \rightarrow}{\alpha_1^v}, \dots, \underset{\leftarrow z(v) \rightarrow}{\alpha_1^v}, 0, \dots, 0).$$

We further have (6.2), (6.3), i.e.,

$$(6.6) \quad \sum_v \alpha_1^v \leq 2,$$

$$(6.7) \quad \sum_{v \in S_l} \alpha_1^v \leq 1 \quad (l = 1, \dots, m).$$

Let  $A''$  be the set of  $\{\alpha_1^v\}_{v \in S}$  in  $\mathbb{R}^s$  where  $s = \text{card } S$ , having  $\alpha_1^v \geq 0$ , with  $\alpha_1^v = 0$  when  $z(v) = 0$ , (6.6), (6.7). Clearly  $A''$  is a convex polyhedron. An *extreme* vertex is defined as a vertex  $\{\alpha_1^v\}$  such that there is no other vertex  $\{\beta_1^v\}$  having  $\beta_1^v \geq \alpha_1^v$  for each  $v \in S$ . It is easily seen that the extremal vertices of  $A''$  are

- (i) For  $u \in S_0$  with  $z(u) > 0$ , the vertex  $\varphi''(u)$  having  $\alpha_1^u = 2$ , but  $\alpha_1^v = 0$  for  $v \neq u$ .
- (ii) For  $u, q$  not in the same class  $S_l$ , with  $z(u), z(q) > 0$ , the vertex  $\varphi''(u, q)$  having  $\alpha_1^u = \alpha_1^q = 1$ , but  $\alpha_1^v = 0$  for  $v \neq u, q$ .
- (iii) When  $S_0 = \emptyset$  and  $m = 1$ , for each  $u$  with  $z(u) > 0$  the vertex  $\psi''(u)$  having  $\alpha_1^u = 1$ , but  $\alpha_1^v = 0$  for  $v \neq u$ .

There is a 1-1 correspondence between extreme vertices of  $A''$  and  $A'$ . For example, the vertex  $\underline{\underline{\varphi}}''(u)$  of  $A''$  corresponds to the vertex  $\underline{\underline{\varphi}}'(u)$  of  $A'$  having

$$\underline{\underline{\alpha}}^u = (2, 2, \dots, 2, 0, \dots, 0)$$

$\leftarrow z(u) \rightarrow$

and having  $\underline{\underline{\alpha}}^v = \underline{0}$  for  $v \neq u$ . Similarly,  $\underline{\underline{\varphi}}''(u, q)$  corresponds to  $\underline{\underline{\varphi}}'(u, q)$  having

$$\underline{\underline{\alpha}}^u = (1, 1, \dots, 1, 0, \dots, 0), \quad \underline{\underline{\alpha}}^q = (1, \dots, 1, 0, \dots, 0),$$

$\leftarrow z(u) \rightarrow \quad \leftarrow z(q) \rightarrow$

and having  $\underline{\underline{\alpha}}^v = \underline{0}$  for  $v \neq u, q$ .

Therefore the extreme vertices of  $A'$  are integer points. In a similar way one could determine all the vertices of  $A'$ . All of them, and in fact all the vertices of  $A$  are integer points. This establishes Theorem 2'.

We now turn to Theorem 1'. Suppose

$$(6.8) \quad n = 3,$$

and for  $v \in S$  let  $L_1^v, L_2^v, L_3^v$  be linearly independent. We have to investigate  $\bar{X}'(3 + \varepsilon)$  when  $\varepsilon > 0$ , i.e.,  $\bar{X}'(\rho)$  when  $\rho > 3$ . Now  $X(\rho)$  is not dense by the Subspace Theorem; therefore the components of  $\bar{X}'(\rho)$  will be 2-dimensional subspaces  $V \subseteq k^3$ . These spaces will have  $r(V) \geq \rho$ , therefore  $r(V) \geq 4$ , since  $r(V) \in \mathbb{Z}$  by the Corollary to Theorem 2'.

Let  $L_i^{vV}$  be the restriction of  $L_i^v$  to  $V$ , and  $A(V)$  the approximation polyhedron for these restrictions. Note that by  $f(\underline{x}) > 0$  in (1.4), no  $L_i^{vV} = 0$ . There will be a vertex  $\underline{\underline{\alpha}}$  of  $A(V)$  with

$$(6.9) \quad \sum_v \sum_{i=1}^3 \alpha_i^v = r(V) \geq 4.$$

In fact there will be an extreme vertex with this property. Since  $L_1^v, L_2^v, L_3^v$  are linearly independent, their restrictions  $L_1^{vV}, L_2^{vV}, L_3^{vV}$  cannot all be proportional, so that  $z(v) \leq 2$ . The vertex  $\underline{\underline{\varphi}}'(u)$  therefore will have (6.9) only if  $z(u) = 2$ , so that  $\underline{\underline{\alpha}}^u = (2, 2, 0)$ . When dealing with  $A(V)$  rather than  $A'(V)$  we have to allow permutations of the variables, so that we obtain possible extreme vertices  $\underline{\underline{\varphi}}\left(\begin{smallmatrix} u \\ i \end{smallmatrix}\right)$  ( $i = 1, 2, 3$ ) with

$$\underline{\underline{\alpha}}^u = (2, 2, 0) \text{ or } (2, 0, 2) \text{ or } (0, 2, 2) \text{ and } \underline{\underline{\alpha}}^v = \underline{0} \text{ when } v \neq u.$$

The other possible extreme vertices with (6.9) are  $\underline{\underline{\varphi}}\left(\begin{smallmatrix} u & q \\ i & h \end{smallmatrix}\right)$  ( $1 \leq i, h \leq 3$ ), having  $\underline{\underline{\alpha}}^u, \underline{\underline{\alpha}}^q$  among  $(1, 1, 0), (1, 0, 1), (0, 1, 1)$ , and  $\underline{\underline{\alpha}}^v = \underline{0}$  when  $v \neq u, q$ . Altogether, there are  $\leq 3s + 9\binom{s}{2}$  possible vertices. To each component  $V$  we associate such a vertex. The argument in §3 shows that we have at most  $3s + 9\binom{s}{2}$  components.

To obtain Theorem 1' we have to go a little further. If the forms  $L_i^v$  have no zeros in  $k^3 \setminus \underline{0}$ , then their restrictions to  $V$  do not have nontrivial zeros defined over  $k$  either, so that  $S_0^V = S$  (where  $S_0^V$  is  $S_0$  defined in terms of the restrictions of our forms to  $V$ ). Then vertices  $\underline{\underline{\varphi}}\left(\begin{smallmatrix} u & q \\ i & h \end{smallmatrix}\right)$  cannot occur, and the number of components is  $\leq 3s$ .

In general, suppose  $\underline{\varphi}\left(\begin{smallmatrix} u \\ 1 \end{smallmatrix}\right)$  is a vertex of some  $A(V)$ . Then  $u \in S_0^V$ . Also,  $z^V(u) = 2$ , so that  $L_1^{uV}, L_2^{uV}$  are proportional. This implies that the 1-dimensional space  $R$  where  $L_1^u = L_2^u = 0$  is contained in  $V$ . Thus  $L_1^{uV}$  vanishes on  $R$ , and  $R$  is not defined over  $k$  since  $u \in S_0^V$ . Then  $V$  must be the only 2-dimensional subspace defined over  $k$  and containing  $R$ . Suppose  $\underline{\varphi}\left(\begin{smallmatrix} u & q \\ 1 & h \end{smallmatrix}\right)$  is a vertex of some  $A(V')$ . Then also  $V' \supset R$ , so that  $V = V'$ . But we can associate  $\underline{\varphi}\left(\begin{smallmatrix} u \\ 1 \end{smallmatrix}\right)$  to  $A(V)$ , hence do not need  $\underline{\varphi}\left(\begin{smallmatrix} u & q \\ 1 & h \end{smallmatrix}\right)$ . More generally, if we need  $\underline{\varphi}\left(\begin{smallmatrix} u \\ i \end{smallmatrix}\right)$ , then we do not need any  $\underline{\varphi}\left(\begin{smallmatrix} u & q \\ i & h \end{smallmatrix}\right)$ . We now can construct a set  $\mathfrak{S}$  of vertices of cardinality  $|\mathfrak{S}| \leq 9\binom{s}{2}$  such that every component  $V$  of  $\bar{X}'(\rho)$  has a vertex  $\underline{\alpha} = \underline{\alpha}(V)$  with (6.9) in  $A(V) \cap \mathfrak{S}$ . The upper bounds of Theorem 1' follow.

We still have to show that these bounds are best possible. When  $k = \mathbb{Q}$ ,  $s = 1$ , consider the forms

$$\begin{aligned} L_1 &= x + \alpha y + \beta z, \\ L_2 &= \alpha^{-1}x + y + \gamma z, \\ L_3 &= \beta^{-1}x + \gamma^{-1}y + z \end{aligned}$$

where  $\alpha, \beta, \gamma$  are real algebraic and such that none of  $L_1, L_2, L_3$  has a zero in  $\mathbb{Q}^3 \setminus \underline{0}$ , and the three forms are independent. This is certainly true for "general"  $\alpha, \beta, \gamma$ . Let  $W$  be the subspace  $z = 0$ ; then  $L_1^W, L_2^W$  are proportional but the subspace  $L_1^W = 0$  is not defined over  $\mathbb{Q}$ , since  $\alpha \notin \mathbb{Q}$ . Therefore the approximation polyhedron of  $L_1^W, L_2^W, L_3^W$  contains the point  $(2, 2, 0)$ , so that  $r(W) = 4$  and  $W$  is a component of  $\bar{X}'(3 + \varepsilon)$  for  $0 < \varepsilon < 1$ . Similar the subspaces  $U: x = 0$  and  $V: y = 0$  are components, so that we have 3 components.

When  $s > 1$ , construct maps  $\tau^v \in GL(\mathbb{Q}, 3)$  such that the  $3s$  spaces  $\tau^v(U), \tau^v(V), \tau^v(W)$  with  $v \in S$  are distinct. Set  $L_i^v(\underline{x}) = L_i((\tau^v)^{-1}\underline{x})$  where  $L_1, L_2, L_3$  are as above. It is easily seen that our  $3s$  subspaces are components of  $\bar{X}'(3 + \varepsilon)$  for  $0 < \varepsilon < 1$ . Note that our forms  $L_i^v$  have no zeros in  $\mathbb{Q}^3 \setminus \underline{0}$ .

Next, again when  $s > 1$ , consider  $3s$  rational points  $\underline{x}_i^v$  ( $v \in S, 1 \leq i \leq 3$ ) such that any 3 of them are linearly independent. Let  $L_1^v, L_2^v, L_3^v$  be nonzero linear forms which vanish respectively on the planes spanned by  $\underline{x}_2^v, \underline{x}_3^v$  and  $\underline{x}_3^v, \underline{x}_1^v$ , and  $\underline{x}_1^v, \underline{x}_2^v$ . These 3 forms are independent. Let  $V_{ij}^{uv}$  for  $u \neq v$  be the plane spanned by  $\underline{x}_i^u, \underline{x}_j^v$ . By our initial assumption on the  $\underline{x}_i^v$ , these  $9\binom{s}{2}$  planes are distinct. We will show that they have  $r(V_{ij}^{uv}) = 4$ , so that they are components of  $\bar{X}'(3 + \varepsilon)$  for  $0 < \varepsilon < 1$ . It suffices to deal with  $V_{11}^{uv}$ . Both  $L_2^u, L_3^u$  vanish on  $\underline{x}_1^u$ , so that their restrictions to  $V = V_{11}^{uv}$  are proportional, and  $z^V(u) = 2$ . Similarly  $z^V(v) = 2$ , and  $\underline{\varphi}\left(\begin{smallmatrix} u & v \\ 1 & 1 \end{smallmatrix}\right)$  with  $\underline{\alpha}^u = \underline{\alpha}^v = (1, 1, 0)$  is a vertex of  $A(V)$ , so that indeed  $r(V) = 4$ .

We note that Theorem 1' can be proved more directly; all the essential ingredients are in Vojta [5, §2].

## 7. "GENERAL SYSTEMS," AND OTHER EXAMPLES

We will always assume that the forms  $L_i^v$  have coefficients in  $k^v$ . Suppose at first that  $a(v) = n$  and  $L_1^v, \dots, L_n^v$  are linearly independent for each  $v \in S$ .

Then  $\mathfrak{W}_i^v$  ( $1 \leq i \leq n$ ) is spanned by  $L_1^v, \dots, L_i^v$ , we may set  $M_i^v = L_i^v$ , and  $\mathfrak{T}_i^v = \mathfrak{W}_i^v$ . It easily follows that  $\widehat{B} = B = A'$ .

We will call  $L_1^v, \dots, L_n^v$  *general* if for any  $i$  in  $1 \leq i \leq n$  and any  $i$ -dimensional subspace  $T$  defined over  $k$ , the intersection of  $T$  with the subspace  $L_1^v = \dots = L_i^v = 0$  is  $\{0\}$ . This notion depends on the ordering of  $L_1^v, \dots, L_n^v$ . It means that when  $N_{i+1}^v, \dots, N_n^v$  are any  $n-i$  linearly independent linear forms with coefficients in  $k$ , then  $L_1^v, \dots, L_i^v, N_{i+1}^v, \dots, N_n^v$  are linearly independent. In the notation of §4, we have for every  $k$ -rational subspace  $T$  that  $g(v) = \dim(\mathfrak{T} \cap \mathfrak{T}_n^v) = \dim \mathfrak{T} = \text{codim } T = c$ , and  $j_1(v) = n-c+1$ ,  $j_2(v) = n-c+2, \dots, j_c(v) = n$ , so that  $\Omega^v(T) = \{n-c+1, \dots, n\}$ . When  $T = \{0\}$ , then (4.5) as always becomes (4.6). When  $\text{codim } T = c < n$ , then (4.5) becomes

$$(7.1) \quad \sum_v \sum_{j=n-c+1}^n \widehat{\beta}_j^v \leq c.$$

But in view of  $\widehat{\beta}_1^v \geq \dots \geq \widehat{\beta}_n^v$ , (7.1) follows from (4.6). Therefore:

When  $L_1^v, \dots, L_n^v$  is general for each  $v \in S$ , then  $A'$  is given by

$$\alpha_1^v \geq \dots \geq \alpha_n^v = 0$$

and

$$(7.2) \quad \sum_v \sum_{j=1}^n \alpha_j^v \leq n.$$

Therefore if  $L_1^v, \dots, L_n^v$  remain general under every permutation,  $A$  consists of  $\{\alpha_j^v\}$  having

$$(7.3) \quad \min(\alpha_1^v, \dots, \alpha_n^v) = 0 \quad (v \in S)$$

and (7.2). In this case,  $r = n$ .

Next, suppose that  $a(v) = m \geq n$  and that any  $n$  among  $L_1^v, \dots, L_m^v$  are linearly independent for each  $v \in S$ . Then  $b(v) = n$  and  $p(\widehat{\underline{\beta}}) = \underline{\underline{\beta}}$  with

$$\beta_i^v = \begin{cases} \widehat{\beta}_i^v & \text{when } 1 \leq i \leq n, \\ \widehat{\beta}_n^v & \text{when } n \leq i \leq m. \end{cases}$$

Call the system *very general* if for each  $v \in S$ , any ordered set of  $n$  forms  $L_{i_1}^v, \dots, L_{i_n}^v$ , where  $i_1, \dots, i_n$  are distinct, is general. It is not hard to see that  $A$  consists of  $\underline{\alpha} = \{\alpha_i^v\}$  such that

$$(7.4) \quad \text{for each } v, \text{ at least } n-m+1 \text{ among } \alpha_1^v, \dots, \alpha_m^v \text{ are } 0,$$

$$(7.5) \quad \sum_v \sum_{j=1}^m \alpha_j^v \leq n.$$

Again we have  $r = n$ .

Return to the situation where  $a(v) = n$  for  $v \in S$ . We call our system *supergeneral*, if for every  $k$ -rational subspace  $T$  of dimension  $t \geq 2$ , the restrictions of the forms  $L_i^v$  to  $T$  form a very general system (with  $t, n$  in

place of  $n, m$ ). For example, if for given  $v$  and  $t$ ,  $L_i^v = \gamma_{i1}x_1 + \cdots + \gamma_{in}x_n$  is such that the quantities

$$\gamma_{i_1 j_1} \cdots \gamma_{i_t j_t}$$

with  $1 \leq i_1 < i_2 < \cdots < i_t \leq n$  and  $1 \leq j_1, \dots, j_t \leq n$  are linearly independent over  $k$ , then our system is supergeneral. For such a system,  $r(T) = t = \dim T$  for every subspace  $T$  as above. Thus no such subspace can be a component of  $\bar{X}(n + \varepsilon)$  when  $\varepsilon > 0$ , and  $\bar{X}'(n + \varepsilon) = \emptyset$ . Therefore:

*For a supergeneral system, the solutions to (1.4) lie in finitely many 1-dimensional subspaces of  $k^n$ . In particular, if the forms in (1.1) are supergeneral, then (1.1) has only finitely many solutions  $\underline{x} \in \mathbb{Z}^n \setminus \underline{0}$ .*

It is easily seen that the above assertion is true also when the forms  $L_i^v$  do not have coefficients in  $k^v$ .

We give a final example. Let  $k = \mathbb{Q}$  and  $s = 1$ , so that  $S$  consists of the archimedean absolute value. Suppose  $L_1, \dots, L_{n-1}, L_{n+1}$  is a very general system of forms with real coefficients, and  $L_n = L_1 + \cdots + L_{n-1}$ . With respect to the ordered set  $L_1, \dots, L_n, L_{n+1}$ , we may take  $M_1 = L_1, \dots, M_{n-1} = L_{n-1}$ ,  $M_n = L_{n+1}$ , and  $\hat{B}$  consists of  $n$ -tuples with  $\hat{\beta}_1 \geq \cdots \geq \hat{\beta}_n = 0$  and  $\sum_{i=1}^n \hat{\beta}_i^v \leq n$ . Here  $p(\hat{\beta}) = (\hat{\beta}_1, \dots, \hat{\beta}_{n-1}, \hat{\beta}_{n-1}, \hat{\beta}_n)$ , so that  $A' = B$  consists of  $\underline{\alpha} = (\alpha_1, \dots, \alpha_{n+1})$  with

$$(7.6) \quad \alpha_1 \geq \cdots \geq \alpha_{n-1} = \alpha_n \geq \alpha_{n+1} = 0 \quad \text{and} \quad \alpha_1 + \cdots + \alpha_{n-1} \leq n.$$

The situation is similar for any ordering of  $L_1, \dots, L_{n+1}$  where  $L_{n+1}$  is last. But with respect to the ordered set  $L_1, \dots, L_{n-1}, L_{n+1}, L_n$ , we may again take  $M_1 = L_1, \dots, M_{n-1} = L_{n-1}$ ,  $M_n = L_{n+1}$ , and  $\hat{B}$  is as before. This time  $p(\hat{\beta}) = (\hat{\beta}_1, \dots, \hat{\beta}_n, \hat{\beta}_n)$ , so that  $A'$  consists of  $\underline{\alpha}$  with  $\alpha_1 \geq \cdots \geq \alpha_n = \alpha_{n+1} = 0$  and  $\alpha_1 + \cdots + \alpha_{n-1} \leq n$ . The situation is similar for any ordering where  $L_{n+1}$  is not last.

The point  $\alpha_1 = \cdots = \alpha_n = n/(n-1)$ ,  $\alpha_{n+1} = 0$ , is a vertex of (7.6). Here  $\alpha_1 + \cdots + \alpha_{n+1} = n^2/(n-1)$ . It is now easily seen that  $r = n^2/(n-1)$ .

## 8. THE DEPENDENCY ON $S$ AND ON $k$

This section may be omitted at first reading. We assumed that the coefficients of our linear forms were in a number field  $K$  and that  $|\cdot|_v$  for  $v \in M(k)$  was extended in some way to  $K$ . We may replace  $K$  by an algebraic number field containing it, in fact by the field  $\mathbb{Q}^a$  of algebraic numbers. Therefore the field  $K$  is irrelevant. For the rest of this section,  $|\cdot|_{\text{vex}}$  for  $v \in M(k)$  will be a fixed extension of  $|\cdot|_v$  to  $\mathbb{Q}^a$ . Now  $f(\underline{x})$  is defined in terms of  $|\cdot|_{\text{vex}}$ .

We next wish to partially eliminate the dependency on  $S$  and  $k$ . Let  $E$  be a finite Galois extension of  $\mathbb{Q}$  containing  $k$  and the coefficients of the  $L_i^v$ . The extensions of  $v \in M(k)$  to absolute values  $w \in M(E)$  are given by

$$(8.1) \quad |\xi|_w = |\sigma(\xi)|_{\text{vex}},$$

where  $\sigma$  runs through  $\text{Gal}(E/k)$ . Let  $S_E$  consists of  $w \in M(E)$  which are extensions of some  $v \in S$  (in symbols:  $w | v$ ). When  $w \in S_E$  with  $w | v$ , set  $a(w) = a(v)$ .

Let  $E_0$  be the smallest Galois extension of  $\mathbb{Q}$  as above. For  $w \in S_{E_0}$ , pick some  $\sigma \in \text{Gal}(E_0/k)$  such that (8.1) holds, and set  $L_i^w = \sigma^{-1}(L_i^v)$ , i.e.,  $\sigma^{-1}$  is

applied to the coefficients. There is some arbitrary choice here. We get forms  $L_i^w$  ( $w \in S_{E_0}$ ,  $1 \leq i \leq a(w)$ ).

When  $E$  is an arbitrary Galois extension as above and  $u \in S_E$ , there is a unique  $w \in S_{E_0}$  with  $u \mid w$ . Set  $L_i^u = L_i^w$ . We get forms  $L_i^u$  ( $u \in S_E$ ,  $1 \leq i \leq a(u)$ ). For  $\underline{x} \in E^n \setminus \underline{0}$  put  $\lambda_i^u(\underline{x}) = |L_i^u(\underline{x})|_u / |\underline{x}|_u$  ( $u \in S_E$ ,  $1 \leq i \leq a(u)$ ) and

$$f_E(\underline{x}) = \prod_{u \in S_E} \prod_{i=1}^{a(u)} (\lambda_i^u(\underline{x}))^{d_u},$$

where  $d_u$  is the local degree of  $E$  at  $u$ .

For  $\underline{x} \in k^n \setminus \underline{0}$  and  $w \in S_{E_0}$  where  $w \mid v$ ,  $v \in S$ , with (8.1),

$$|\underline{x}|_w = |\underline{x}|_v \quad \text{and} \quad |L_i^w(\underline{x})|_w = |\sigma(L_i^w(\underline{x}))|_{\text{vex}} = |L_i^v(\underline{x})|_{\text{vex}}.$$

More generally, for  $u \in S_E$  and  $u \mid v$ ,  $v \in S$ , we have  $|\underline{x}|_w = |\underline{x}|_v$  and  $|L_i^u(\underline{x})|_u = |L_i^v(\underline{x})|_{\text{vex}}$ , therefore  $\lambda_i^u(\underline{x}) = \lambda_i^{\text{vex}}(\underline{x})$ . Since  $\sum_{u \mid v} d_u = [E : k]d_v$ ,

$$f_E(\underline{x}) = f(\underline{x})^{[E : k]}.$$

Also, if  $\underline{x} \in E^n \setminus \underline{0}$  where  $\mathbb{Q} \subset k \subset E \subset F$  with  $E, F$  Galois extensions as above, then

$$f_F(\underline{x}) = f_E(\underline{x})^{[F : E]}.$$

Therefore it is permissible to define

$$f_{\text{abs}}(\underline{x}) = (f_E(\underline{x}))^{1/[E : \mathbb{Q}]}$$

for  $\underline{x} \in (\mathbb{Q}^a)^n \setminus \underline{0}$ , where  $E$  is any Galois extension as above, which contains the coordinates of  $\underline{x}$ . The relation

$$f_{\text{abs}}(\underline{x}) < H(\underline{x})^{-\rho},$$

where  $H$  is the absolute height, is now well defined. For  $\underline{x} \in k^n \setminus \underline{0}$ , this relation is the same as (1.6).

When constructing the modified approximation polyhedron  $B$  in §4, it mattered how the spaces  $\mathfrak{W}_i^v$  were constructed in terms of  $k^v$ , and therefore  $v$  really mattered. Now consider the same construction  $B_E$  of the forms  $L_i^w$  ( $w \in S_E$ ,  $1 \leq i \leq a(v)$ ), and for points  $\underline{x} \in E^n \setminus \underline{0}$ . Here the  $w$ -adic completion of  $E$  in  $E$  is  $E$  itself, and therefore the nature of the absolute values  $w$  does not matter for  $B_E$  or for  $A_E$ .

Suppose now that  $L_1^v, \dots, L_n^v$  are linearly independent forms with coefficients in  $k$ . Let  $X_E(n + \varepsilon)$  consist of  $\underline{x} \in E^n$  with

$$(8.2) \quad f_{\text{abs}}(\underline{x}) < H(\underline{x})^{-n-\varepsilon}.$$

Then  $\bar{X}'_k(n + \varepsilon) \subseteq \bar{X}'_E(n + \varepsilon) \subseteq Y$ , where  $Y$  is a finite union of proper subspaces of  $E^n$ , which does not depend on the absolute values  $v \in S$ .

Conceivably, the number of components of  $\bar{X}'_E(n + \varepsilon)$  might increase to infinity as  $E$  increases. But we will see that such a calamity cannot actually happen.

Let us return for a moment to the proof of Theorem 1. Write  $v \simeq u$  if  $a(v) = a(u)$  and the tuples  $L_1^v, \dots, L_{a(v)}^v$  and  $L_1^u, \dots, L_{a(u)}^u$  are the same. In the characterization of  $A'$  in §5, the components  $\alpha_i^v$  occur only in blocks  $\sum_v \alpha_i^v$ ,

where the sum is over  $v$  in a given equivalence class. Pick some representative  $v_0$  in each class. If there is a vertex  $\underline{\alpha} \in A'$  with  $\sum_v \sum_i \alpha_i^v = r$ , then there is also a vertex which is *special* in the sense that  $\alpha_i^v = 0$  when  $v$  is not the representative  $v_0$  in its class. We are then reduced to points in a polyhedron of dimension  $\sum_{v_0} a(v_0) = a_0$ , say, where the sum is over representatives  $v_0$  of the classes. In analogy to (3.2), we get  $\leq (nt)^{2a_0^2}$  vertices, where  $t$  is the number of equivalence classes. Thus the number of components of  $\bar{X}'(\rho)$  is  $\leq (nt)^{2a_0^2} \leq (nt)^{2n^2t^2} \leq 2^{2n^3t^3}$ .

The forms  $L_1^w, \dots, L_n^w$  with  $w \in S_E$  are obtained from the forms  $L_1^v, \dots, L_n^v$  ( $v \in S$ ) by conjugation, so that there are at most  $s[K : \mathbb{Q}]$  such  $n$ -tuples, where  $K$  contains the coefficients of the  $L_i^v$ . But in  $A'_E$  the ordering of each  $n$ -tuple matters, and we may get  $\leq n!s[K : \mathbb{Q}]$  distinct ordered tuples. We may apply what we said above with  $t = n!s[K : \mathbb{Q}]$ . Therefore the number of components of  $\bar{X}'_E(N + \varepsilon)$  is

$$(8.3) \quad \leq 2^{2(n \cdot n!s[K : \mathbb{Q}])^3}.$$

The following is easily seen. Suppose to each number field  $E$  we associate a closed set  $C_E \subseteq (\mathbb{Q}^a)^n$  with at most  $c$  components (so that it is the union of at most  $c$  subspaces of  $(\mathbb{Q}^a)^n$ ), and having  $C_E \subset C_F$  when  $E \subset F$ . Then there is a field  $F_0$  such that  $C_E \subseteq C_{F_0}$  for every field  $E$ . We obtain

**Theorem 4.** *There is a closed proper subset  $Y \subseteq (\mathbb{Q}^a)^n$  with at most (8.3) components such that for any  $\varepsilon > 0$  and any field  $E$  containing  $k$  and the coefficients of the forms  $L_i^v$ , and where  $E/\mathbb{Q}$  is Galois,*

$$\bar{X}'_E(n + \varepsilon) \subseteq Y.$$

Thus for any such field  $E$ , the solutions of (8.2) lie in  $Y$ , with the exception of points in a finite number of 1-dimensional subspaces which may depend on  $E$  and on  $\varepsilon$ .

## 9. SOME LEMMAS

**Lemma 7.** *Let  $v \in M(k)$  and  $k_v$  the completion of  $k$  with respect to  $v$ . Then the algebraic closure  $k_v^a$  of  $k$  in  $k_v$  has  $[k_v^a : k] = \infty$ .*

*Proof.* This is well known.

Write  $[k : \mathbb{Q}] = d$ , and  $d = r_1 + 2r_2$ , if there are  $r_1$  real embeddings  $x \mapsto x^{(i)}$  ( $1 \leq i \leq r_1$ ) of  $k$  into  $\mathbb{C}$ , and  $r_2$  pairs of complex conjugate embeddings  $x \mapsto x^{(i)}$ ,  $x \mapsto x^{(i+r_2)}$  for  $r_1 < i \leq r_1 + r_2$ . Put

$$\underline{\underline{\psi}}(x) = (x^{(1)}, \dots, x^{(r_1)}, \operatorname{Re} x^{(r_1+1)}, \operatorname{Im} x^{(r_1+1)}, \dots, \operatorname{Re} x^{(r_1+r_2)}, \operatorname{Im} x^{(r_1+r_2)}),$$

where  $\operatorname{Re}$ ,  $\operatorname{Im}$  denote the real and imaginary part, respectively. As  $x$  runs through the ring  $\mathfrak{O}_k$  of integers of  $k$ , then  $\underline{\underline{\psi}}(x)$  runs through a certain lattice of  $\mathbb{R}^d$ . Given  $\underline{x} = (x_1, \dots, x_n) \in k^n$ , set  $\underline{\underline{\psi}}(\underline{x}) = (\underline{\underline{\psi}}(x_1), \dots, \underline{\underline{\psi}}(x_n)) \in \mathbb{R}^{nd}$ . As  $\underline{x}$  ranges through  $\mathfrak{O}_k^n$ , then  $\underline{\underline{\psi}}(\underline{x})$  runs through a lattice in  $\mathbb{R}^{nd}$ .

Given a set  $S \subseteq M(k)$  as described in the Introduction, let  $S_0, S_\infty$  respectively consist of the nonarchimedean and the archimedean absolute values in  $S$ . (This supersedes the notation  $S_0$  of §6.) Then  $S_\infty$  contains all the archimedean absolute values.

**Lemma 8.** Suppose for each  $v \in S_0$  we are given linearly independent linear forms  $M_1^v, \dots, M_n^v$  in  $n$  variables with coefficients in  $k_v$ . Suppose we are given real numbers  $C_j^v$  with  $0 < C_j^v \leq 1$  ( $v \in S_0$ ,  $1 \leq j \leq n$ ). Then as  $\underline{x}$  runs through the points of  $\mathfrak{D}_k^n$  having

$$(9.1) \quad |M_j^v(\underline{x})|_v \leq C_j^v \quad (v \in S_0, 1 \leq j \leq n),$$

the points  $\underline{\psi}(\underline{x})$  will run through a lattice  $\Lambda \subset \mathbb{R}^{nd}$  of determinant

$$\det \Lambda \gg \ll \prod_{v \in S_0} \prod_{j=1}^n (C_j^v)^{-d_v}.$$

Here the constants in  $\gg \ll$  may depend on everything except the  $C_j^v$ . The routine proof will not be carried out here.

**Lemma 9.** Suppose that for each  $v \in S_\infty$  we are given linearly independent linear forms  $M_1^v, \dots, M_n^v$  in  $n$  variables with coefficients in  $k_v$ . Suppose we are given real numbers  $C_j^v$  with  $0 < C_j^v \leq 1$  ( $v \in S_\infty$ ,  $1 \leq j \leq n$ ). Then  $\underline{x} \in k^n$  will satisfy

$$(9.2) \quad |M_j^v(\underline{x})|_v \leq C_j^v \quad (v \in S_\infty, 1 \leq j \leq n)$$

if and only if  $\underline{\psi}(\underline{x})$  lies in a certain convex set  $\mathfrak{K} \subset \mathbb{R}^{nd}$  of volume

$$\text{vol } \mathfrak{K} \gg \ll \prod_{v \in S_\infty} \prod_{j=1}^n (C_j^v)^{d_v}.$$

Again the proof will be omitted.

## 10. THE NECESSITY OF THE CONDITIONS IN THEOREM 3

We now begin with the proof of Theorem 3. From now on, for convenience, we will write  $\beta_i^v$  for  $\hat{\beta}_i^v$  and  $B$  for  $\hat{B}$ . Our aim here is to show that  $\underline{\beta} \in B$  necessarily satisfies (4.3), (4.4), (4.5).

By definition,  $\beta_j^v(\underline{x}) \geq \beta_{j+1}^v(\underline{x})$ , so that in fact  $\beta_1^v \geq \dots \geq \beta_{b(v)}^v$  for  $\underline{\beta} \in B$ . Further  $\mu_j^v(\underline{x}) \ll 1$ , therefore  $\alpha_j^v(\underline{x}) \geq -c/\log H_k(\underline{x})$  for  $\underline{x} \in k^n \setminus \mathfrak{M}$  (recall that  $\mathfrak{M}$  consists of  $\underline{x}$  having some  $\mathfrak{M}_j^v(\underline{x}) = 0$  or having  $\bar{H}_k(\underline{x}) < e$ ), which gives  $\beta_j^v(\underline{x}) \geq -c/\log H_k(\underline{x})$  and  $\beta_j^v \geq 0$ . Thus (4.3) is established. When  $b(v) = n$ ,  $\max(|M_1^v(\underline{x})|_v, \dots, |M_n^v(\underline{x})|_v) \gg |\underline{x}|_v$ , so that

$$\max(\mu_1^v(\underline{x}), \dots, \mu_n^v(\underline{x})) \gg 1$$

and  $\beta_n^v(\underline{x}) \leq c'/\log H_k(\underline{x})$  for  $\underline{x} \in k^n \setminus \mathfrak{M}$ , and therefore  $\beta_n^v = 0$ . This establishes (4.4).

It remains for us to deal with (4.5) for a given subspace  $T$ . Since, as noticed above, (4.5) is an empty condition when  $T = k^n$ , we may suppose that  $c = \text{codim } T > 0$ . Recall that  $\mathfrak{T}$  was the space of linear forms vanishing on  $T$ ; let  $X_1, \dots, X_c$  be a basis of  $\mathfrak{T}$ . Write  $\underline{X}(\underline{x}) = (X_1(\underline{x}), \dots, X_c(\underline{x}))$ , so that  $\underline{x} \mapsto \underline{X}(\underline{x})$  is a linear map  $k^n \rightarrow k^c$ . Pick linear forms  $N_1^v, \dots, N_{g(v)}^v$  such that  $N_1^v$  is a basis of  $\mathfrak{T} \cap \mathfrak{T}_{j_1(v)}^v$ , that  $N_1^v, N_2^v$  is a basis of  $\mathfrak{T} \cap \mathfrak{T}_{j_2(v)}^v$ , etc. Then  $N_i^v$  is a linear combination of  $M_1^v, \dots, M_{j_1(v)}^v$ . Further  $N_1^v, \dots, N_{g(v)}^v$  are linearly

independent linear combinations of  $X_1, \dots, X_c$ , so that  $N_i^v(\underline{x}) = N_i'^v(\underline{X}(\underline{x}))$ , where  $N_i'^v$  is a linear form  $k^c \rightarrow k$ , and the forms  $N_1'^v, \dots, N_{g(v)}'^v$  are linearly independent. Writing  $\underline{X} = \underline{X}(\underline{x})$  for short,

$$\begin{aligned} |N_i'^v(\underline{X})|_v &= |N_i^v(\underline{x})|_v \ll \max(|M_1^v(\underline{x})|_v, \dots, |M_{j_i(v)}^v(\underline{x})|_v) \\ &= |\underline{x}|_v \max(\mu_1^v(\underline{x}), \dots, \mu_{j_i(v)}^v(\underline{x})) \\ &= |\underline{x}|_v H_k(\underline{x})^{-\beta(i, v, \underline{x})/d_v}, \end{aligned}$$

where  $\beta(i, v, \underline{x}) = \beta_{j_i(v)}^v(\underline{x})$ .

Points  $\underline{x}$  with  $\underline{X}(\underline{x}) = \underline{0}$  are not dense. When  $\underline{X} = \underline{X}(\underline{x}) \neq \underline{0}$ , then  $|\underline{X}|_v \ll |\underline{x}|_v$ , and in fact  $|\underline{X}|_v \leq |\underline{x}|_v$  for all but finitely many  $v$ , and  $H_k(\underline{X}) \ll H_k(\underline{x})$ . Setting  $\nu_i^v(\underline{X}) = |N_i^v(\underline{X})|_v/|\underline{X}|_v$  we obtain

$$\begin{aligned} \prod_v \prod_{i=1}^{g(v)} (\nu_i^v(\underline{X}))^{d_v} &\ll \prod_v \prod_{i=1}^{g(v)} \left( \left( \frac{|\underline{x}|_v}{|\underline{X}|_v} \right)^{d_v} H_k(\underline{x})^{-\beta(i, v, \underline{x})} \right) \\ &\ll \prod_v \left( \frac{|\underline{x}|_v}{|\underline{X}|_v} \right)^{cd_v} H_k(\underline{x})^{-\beta(\underline{x})}, \end{aligned}$$

because of  $g(v) \leq c$ , where

$$(10.1) \quad \beta(\underline{x}) = \sum_v \sum_{i=1}^{g(v)} \beta(i, v, \underline{x}) = \sum_v \sum_{i=1}^{g(v)} \beta_{j_i(v)}^v(\underline{x}) = \sum_v \sum_{j \in \Omega^v(T)} \beta_j^v(\underline{x}).$$

Therefore

$$\prod_v \prod_{i=1}^{g(v)} (\nu_i^v(\underline{X}))^{d_v} \ll (H_k(\underline{x})/H_k(\underline{X}))^c H_k(\underline{x})^{-\beta(\underline{x})}.$$

By the Subspace Theorem, applied to the linear forms  $N_i'^v$  in  $c$  variables, the left-hand side here is  $> H_k(\underline{X})^{-c-\varepsilon}$  except for a set of  $\underline{X} \in k^c$  which is not dense. Therefore, except for a set of  $\underline{x} \in k^c$  which is not dense, the left-hand side with  $\underline{X} = \underline{X}(\underline{x})$  is  $> H_k(\underline{X})^{-c-\varepsilon} \gg H_k(\underline{X})^{-c} H_k(\underline{x})^{-\varepsilon}$ , and therefore  $\beta(\underline{x}) < c + 2\varepsilon$ . We may conclude that  $\underline{\beta} \in B$  has (4.5).

## 11. SOME REDUCTIONS

To finish the proof of Theorem 3, in particular to prove the sufficiency of the conditions (4.3), (4.4), (4.5), we begin with some reductions. Let  $\underline{\beta} \in \mathbb{R}^b$  lie in the polyhedron  $P$  given by (4.3), (4.4), (4.5). Any open neighborhood of  $\underline{\beta}$  will contain some cube

$$(11.1) \quad |\underline{\xi} - \underline{\beta}| < \varepsilon.$$

It will suffice to show that points  $\underline{x}$  with  $\underline{\beta}(\underline{x})$  in this cube are dense, and that for  $H > H_5(\varepsilon)$  there are  $n$  independent points  $\underline{x}$  with  $H < H_k(\underline{x}) < H^{1+\varepsilon}$  and with  $\underline{\beta}(\underline{x})$  in this cube.

The point  $\underline{\beta}$  in particular satisfies (4.6).

*First Reduction.* We may suppose that (4.6) holds with equality. To justify this claim, pick  $w \in M(k)$ ,  $w \notin S$ , and set  $S^* = S \cup \{w\}$ . Set  $b(w) = 1$ ,

and let  $M_1^w$  be a linear form with coefficients which are algebraic and linearly independent over  $k$ , and which lie in  $k_v$ . Such a choice is possible by Lemma 7. By enlarging  $K$  if necessary, the coefficients will lie in  $k^v \subset K$ . Let  $\underline{\underline{\beta}}^* = \{\beta_j^v\}$  be the same as  $\underline{\underline{\beta}}$ , but with an extra component  $\beta_1^w$  chosen such that

$$(11.2) \quad \sum_{v \in S^*} \sum_{j=1}^{b(v)} \beta_j^v = n.$$

Then  $\beta_1^w \geq 0$  by (4.6).

**Lemma 10.**  $\underline{\underline{\beta}}^*$  lies in the polyhedron  $P^*$  in  $\mathbb{R}^{b+1}$  given by (4.3), (4.4), (4.5), but defined in terms of  $S^*$  and the forms  $M_j^v$  with  $v \in S^*$ ,  $1 \leq j \leq b(v)$ .

*Proof.* All we have to check is (4.5). When  $T = \{0\}$ , we have  $\Omega^v(T) = \{1, \dots, b(v)\}$  for each  $v \in S^*$ , and (4.6) with the sum over  $v \in S^*$  holds with equality by (11.2). When  $T \neq \{0\}$ , we will see below that

$$(11.3) \quad \Omega^w(T) = \emptyset.$$

Therefore the sum in (4.5) is the same whether extended over  $v \in S$  or  $\in S^*$ , so that  $\underline{\underline{\beta}} \in P$  implies  $\underline{\underline{\beta}}^* \in P^*$ .

We yet have to establish (11.3). If it were false, then  $\Omega^w(T) = \{1\}$  (for it is a subset of  $\{1, \dots, b(w)\} = \{1\}$ ). Then  $M_1^w$  would be a linear combination of  $c < n$  linear forms  $F_1, \dots, F_c$  defining the  $k$ -rational space  $T$ . There is a point  $\underline{x} \in k^n \setminus \underline{0}$  with  $F_i(\underline{x}) = 0$  ( $i = 1, \dots, c$ ), therefore  $M_1^w(\underline{x}) = 0$ , contradicting the linear independence of the coefficients of  $M_1^w$  over  $k$ . This proves the lemma.

Define  $\underline{\underline{\beta}}^*(\underline{x})$  like  $\underline{\underline{\beta}}(\underline{x})$ , but with an extra component  $\beta_1^w(\underline{x})$ . Now if  $|\underline{\underline{\beta}}^*(\underline{x}) - \underline{\underline{\beta}}^*| < \varepsilon$ , then *a fortiori*  $|\underline{\underline{\beta}}(\underline{x}) - \underline{\underline{\beta}}| < \varepsilon$ . The First Reduction follows.

*Second Reduction.* We may suppose that  $b(v) = n$  for each  $v \in S$ . For if  $b(v) < n$ , augment  $M_1^v, \dots, M_{b(v)}^v$  to get  $n$  linearly independent forms  $M_1^v, \dots, M_n^v$  with coefficients in  $k^v$ . Let  $\underline{\underline{\beta}}^\circ = \{\beta_j^{\circ v}\}$  with  $\beta_j^{\circ v} = \beta_j^v$  when  $1 \leq j \leq b(v)$ , and  $\beta_j^{\circ v} = 0$  when  $b(v) < j \leq n$ . All we need is

**Lemma 11.**  $\underline{\underline{\beta}}^\circ$  lies in the polyhedron  $P^\circ \subset \mathbb{R}^{sn}$  given by (4.3), (4.4), (4.5), but defined in terms of the linear forms  $M_j^v$  with  $v \in S$ ,  $1 \leq j \leq n$ .

*Proof.* Since (4.3), (4.4) are obvious, it remains for us to check (4.5). We have to show that (4.5) remains valid if the sum, rather than over  $\Omega^v(T)$ , is extended over  $\Omega^{\circ v}(T)$ , where  $\Omega^{\circ v}(T)$  is defined in the obvious way. Observe that  $\Omega^{\circ v}(T)$  is obtained from  $\Omega^v(T)$  by possibly adding some integers  $j > b(v)$ . But  $\beta_j^v = 0$  for such integers  $j$ .

## 12. COMPLETION OF THE PROOF OF THEOREM 3

Let  $\underline{\underline{\beta}} \in P$ , i.e., suppose (4.3), (4.4), (4.5) hold. We may suppose that each  $b(v) = n$ , and that (4.6) holds with equality. For  $Q > 1$ , consider points  $\underline{x} \in \mathfrak{O}_k^n$  with

$$(12.1) \quad |M_j^v(\underline{x})|_v \leq Q^{-\beta_j^v/d_v} \quad (v \in S_0, 1 \leq j \leq n).$$

As  $\underline{x}$  runs through these points,  $\underline{\psi}(\underline{x})$  runs through a lattice  $\Lambda \subset \mathbb{R}^{nd}$ . Further

$$(12.2) \quad |M_j^v(\underline{x})|_v \leq Q^{-(\beta_j^v/d_v)+(1/d)} \quad (v \in S_\infty, \quad 1 \leq j \leq n)$$

holds if  $\underline{\psi}(\underline{x})$  is in a certain convex set  $\mathfrak{K} \in \mathbb{R}^{nd}$ . By Lemmas 8, 9, by (4.6), and since

$$(12.3) \quad \sum_{v \in S_\infty} d_v = d,$$

we have

$$\text{vol } \mathfrak{K} / \det \Lambda \gg Q^{-\sum_v \sum_j \beta_j^v + n} \geq 1.$$

Let  $m_1, \dots, m_{nd}$  be the successive minima of  $\mathfrak{K}$  with respect to  $\Lambda$ . By Minkowski's Theorem,

$$(12.4) \quad m_1 \cdots m_{nd} \ll 1.$$

**Lemma 12.**  $m_1 > Q^{-\varepsilon}$  when  $Q > Q_0(\varepsilon)$ .

*Proof.* When  $m_1 < Q^{-\varepsilon}$ , there is a point  $\underline{x} \in \mathfrak{D}_k^n \setminus \underline{0}$  with (12.1) and with

$$(12.5) \quad |M_j^v(\underline{x})|_v \leq Q^{-(\beta_j^v/d_v)+(1/d)-\varepsilon} \quad (v \in S_\infty, \quad 1 \leq j \leq n).$$

Such points have  $|\underline{x}|_v \leq 1$  when  $v$  is nonarchimedean,  $|\underline{x}|_v \ll Q^{(1/d)-\varepsilon}$  when  $v$  is archimedean, therefore

$$(12.6) \quad H_k(\underline{x}) \ll Q^{1-d\varepsilon} \leq Q.$$

Let  $T$  be a subspace of  $k^n$  with  $\dim T = t > 1$ . Recall the definition of  $\Omega^v(T)$ . Since now  $b(v) = n$ , we have  $|\Omega^v(T)| = g(v) = c = \text{codim } T$ . The complementary set  $\Phi^v(T)$  has  $|\Phi^v(T)| = n - c = t$ . Observe that by equality in (4.6) and by (4.5),

$$-\sum_{v \in S} \sum_{j \in \Phi^v(T)} \beta_j^v + \sum_{v \in S_\infty} \sum_{j \in \Phi^v(T)} d_v((1/d) - \varepsilon) = \sum_{v \in S} \sum_{j \in \Omega^v(T)} \beta_j^v - n + t - td\varepsilon < -\varepsilon.$$

Therefore  $\underline{x} \in \mathfrak{D}_k^n \setminus \underline{0}$ ,  $\underline{x} \notin \mathfrak{M}$  with (12.1), (12.5) has

$$(12.7) \quad \prod_{v \in S} \prod_{j \in \Phi^v(T)} (\mu_j^v(\underline{x}))^{d_v} < Q^{-\varepsilon} \prod_{v \in S} |\underline{x}|_v^{-td_v} \leq Q^{-\varepsilon} H_k(\underline{x})^{-t} \ll H_k(\underline{x})^{-t-\varepsilon}.$$

The restrictions to  $T$  of the forms  $M_j^v$  with given  $v$  and  $j \in \Phi^v(T)$  are  $t = \dim T$  linearly independent forms. By the Subspace Theorem, points  $\underline{x} \in T$  with (12.7) cannot be dense in  $T$ .

Let  $Z$  be the set of points  $\underline{x} \in \mathfrak{D}_k^n$  with (12.1), (12.5) for arbitrary  $Q$ . We have seen that  $Z \cap T$  is not dense in  $T$ . Therefore by Lemma 2,  $T$  cannot be a component of  $Z$ . Therefore the only possible components of  $Z$  are of dimension 1. When  $\dim T = t = 1$ , we still get (12.7). When  $\dim T = 1$ , then  $\Phi^v(T)$  indexes precisely the forms  $M_j^v$  which do not vanish on  $T$ . Therefore when  $\underline{x} \in T \setminus \underline{0}$ , the left-hand side of (12.7) is  $\neq 0$ , so that (12.7) cannot happen when  $Q > Q_0(\varepsilon)$ .

From (12.4) and Lemma 12,  $m_{nd} \ll Q^{nd\varepsilon-\varepsilon}$ . Since  $\varepsilon > 0$  was arbitrary we have  $m_{nd} < Q^\varepsilon$  when  $Q > Q_1(\varepsilon)$ . There are  $nd$  linearly independent points  $\underline{\psi}(\underline{x}_i)$  ( $i = 1, \dots, nd$ ) in  $\Lambda \cap (m_{nd}\mathfrak{K})$ . Among  $\underline{x}_1, \dots, \underline{x}_{nd}$  there must be

$n$  linearly independent points; say  $\underline{x}_1, \dots, \underline{x}_n$  are such points. Each of these points lies in  $\mathfrak{O}_k^n$  and satisfies

$$(12.8) \quad |M_j^v(\underline{x})|_v < Q^{-(\beta_j^v/d_v) + (\Delta_v/d) + \varepsilon} \quad (v \in S, \ 1 \leq j \leq n)$$

where as in §3,  $\Delta_v = 1$  when  $v \in M_\infty(k) = S_\infty$ , and  $\Delta_v = 0$  otherwise. These points have  $|\underline{x}|_v \leq 1$  when  $v$  is nonarchimedean,  $|\underline{x}|_v \ll Q^{(1/d) + \varepsilon}$  when  $v$  is archimedean, therefore

$$(12.9) \quad H_k(\underline{x}) \ll Q^{1+d\varepsilon}.$$

**Lemma 13.** Consider the set of points  $\underline{x} \in \mathfrak{O}_k^n$  which satisfy (12.8) for any  $Q > 1$ .

(a) The subset of points for which one of these  $n$ s inequalities holds with an extra factor  $Q^{-2dn\varepsilon}$  on the right-hand side is not dense in  $k^n$ .

(b) The subset of points with  $H_k(\underline{x}) < Q^{1-3d^2ns^2\varepsilon}$  is not dense.

*Proof.* (a) From (12.8), equality in (4.6), and (12.3),

$$\prod_{v \in S} \prod_{j=1}^n (\mu_j^v(\underline{x}))^{d_v} < Q^{dn\varepsilon} \prod_{v \in S} |\underline{x}|_v^{-nd_v} \leq Q^{dn\varepsilon} H_k(\underline{x})^{-n}.$$

If we had an extra factor  $Q^{-2dn\varepsilon}$ , this would be

$$< Q^{-dn\varepsilon} H_k(\underline{x})^{-n} \ll H_k(\underline{x})^{-n-dn\varepsilon/(1+d\varepsilon)}$$

by (12.9). In view of the Subspace Theorem, such  $\underline{x} \in k^n$  are not dense.

(b) By part (a) we may suppose that

$$|M_j^v(\underline{x})|_v \geq Q^{-(\beta_j^v/d_v) + (\Delta_v/d) + \varepsilon - 2dn\varepsilon} \quad (v \in S, \ 1 \leq j \leq n).$$

Applying this with  $j = n$  and observing that  $\beta_n^v = 0$ , we obtain  $|\underline{x}|_v \gg Q^{(\Delta_v/d) + \varepsilon - 2dn\varepsilon}$ , so that by (12.8) and equality in (4.6),

$$\mu_j^v(\underline{x}) \ll Q^{-(\beta_j^v/d_v) + 2dn\varepsilon},$$

and

$$\prod_{v \in S} \prod_{j=1}^n (\mu_j^v(\underline{x}))^{d_v} \ll Q^{-n+2d^2n^2s^2\varepsilon}$$

By the Subspace Theorem, solutions with  $H_k(\underline{x}) < Q^{1-3d^2ns^2\varepsilon}$  are not dense.

Let  $T_1, \dots, T_t$  be proper subspaces of  $k^n$  containing the exceptional points  $\underline{x}$  described in Lemma 13.

The proof of Theorem 3 is now completed as follows. For  $Q > Q_2(\varepsilon)$  pick linearly independent points  $\underline{x}_1, \dots, \underline{x}_n$  in  $\mathfrak{O}_k^n$  satisfying (12.8) with  $\varepsilon/3$  in place of  $\varepsilon$ . Consider points

$$\underline{x} = u_1 \underline{x}_1 + \dots + u_n \underline{x}_n \quad \text{with} \quad u_i \in \mathbb{Z}, \ 1 \leq u_i < Q^{\varepsilon/3} \ (i = 1, \dots, n).$$

When  $Q$  is sufficiently large, most of these points will be outside  $T_1, \dots, T_t$ , and outside any other given finite collection of proper subspaces. When  $Q$  is large, they will satisfy the right hand inequalities in

$$Q^{-(\beta_j^v/d_v) + (\Delta_v/d) - 2dn\varepsilon} < |M_j^v(\underline{x})|_v < Q^{-(\beta_j^v/d_v) + (\Delta_v/d) + \varepsilon} \quad (v \in S, \ 1 \leq j \leq n).$$

Since they lie outside  $T_1, \dots, T_t$ , they will also satisfy the left-hand inequalities, and they will have

$$(12.10) \quad Q^{1-3d^2ns^2\varepsilon} < H_k(\underline{x}) < Q^{1+2d\varepsilon};$$

here we used also (12.9). Further  $|\underline{x}|_v \gg \ll \max(|M_1^v(\underline{x})|_v, \dots, |M_n^v(\underline{x})|_v)$ , so that

$$Q^{(\Delta_v/d)-2d n s \varepsilon} \ll |\underline{x}|_v \ll Q^{(\Delta_v/d)+\varepsilon} \quad (v \in S),$$

and

$$Q^{-(\beta_j^v/d_v)-3d n s \varepsilon} \ll \mu_j^v(\underline{x}) \ll Q^{-(\beta_j^v/d_v)+3d n s \varepsilon}.$$

In conjunction with (12.10) this yields

$$|\alpha_j^v(\underline{x}) - \beta_j^v| < c_1 \varepsilon \quad (v \in S, 1 \leq j \leq n),$$

where  $c_1 = c_1(d, n, s)$ . Since  $\beta_j^v(\underline{x}) = \min(\alpha_1^v(\underline{x}), \dots, \alpha_j^v(\underline{x}))$ , and  $\beta_j^v = \min(\beta_1^v, \dots, \beta_j^v)$  by (4.3), we obtain  $|\beta_j^v(\underline{x}) - \beta_j^v| < c_1 \varepsilon$ , i.e.,

$$|\underline{\beta}(\underline{x}) - \underline{\beta}| < c_1 \varepsilon.$$

Since  $\varepsilon > 0$  was arbitrary, and since points  $\underline{x}$  with this property do not lie in any given finite collection of proper subspaces,  $\underline{\beta}$  is in fact an approximation point. As for the constancy principle, given  $H$  and small  $\varepsilon > 0$ , pick  $Q$  with  $Q^{1-3d^2ns^2\varepsilon} = H$ ; then (12.10) yields  $H < H_k(\underline{x}) < H^{1+c_2\varepsilon}$ .

## REFERENCES

1. H. P. Schlickewei, *Die p-adische Verallgemeinerung des Satzes von Thue-Siegel-Roth-Schmidt*, J. Reine Angew. Math. **288** (1976), 86–105.
2. W. M. Schmidt, *Norm form equations*, Ann. of Math. **96** (1972), 526–551.
3. ———, *Diophantine approximation*, Lecture Notes in Math., vol. 785, Springer, 1980.
4. J. H. Silverman, *The arithmetic of elliptic curves*, Graduate Texts in Math., Springer, 1985.
5. P. Vojta, *A refinement of Schmidt's Subspace Theorem*, Amer. J. Math. **111** (1989), 489–511.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF COLORADO, BOULDER, COLORADO 80309-0395  
E-mail address: schmidt@euclid.colorado.edu