The Absolute Subspace Theorem and linear equations with unknowns from a multiplicative group

Jan-Hendrik Evertse and Hans Peter Schlickewei

To Professor Andrzej Schinzel on his 60-th birthday

0. Introduction

We deal with the equation

$$a_1x_1 + \ldots + a_nx_n = 1 \quad \text{in } x_1, \ldots, x_n \in G \tag{0.1}$$

where the coefficients $a_1, \ldots, a_n$ are non-zero elements of a given algebraic number field K and where G is a finitely generated subgroup of the multiplicative group K^*. Independently, Evertse [8] and van der Poorten and Schlickewei [24] showed that (0.1) has only finitely many solutions with non-vanishing subsums, i.e.

$$\sum_{i \in I} a_ix_i \neq 0 \quad \text{for each non-empty subset } I \text{ of } \{1, \ldots, n\}. \tag{0.2}$$

They both gave essentially the same proof, based on the Subspace Theorem (more precisely, Schlickewei's generalisation to p-adic absolute values and number fields [30] of the Subspace Theorem proved by Schmidt in 1972 [41]).

In 1984, Evertse [7] showed that if G is the group of S-units in K and $a, b \in K^*$, then the equation $ax + by = 1$ has at most 3×7^{4s} solutions in $x, y \in G$, where s is the cardinality of S. The significant feature of this bound is its uniformity. It does not depend upon the coefficients a and b and it involves only the cardinality of the set S but not the particular primes belonging to S. Schmidt's pioneering work from 1989 [42] in which he obtained a quantitative version of his Subspace Theorem from 1972 giving an explicit upper bound for the number of subspaces involved, opened the possibility to determine explicit upper bounds for the number of solutions of Diophantine equations from several classes, including eq. (0.1) in $n \geq 3$ unknowns. In fact, many of the generalisations and improvements of Schmidt's result obtained later were motivated by the desire to derive good explicit uniform upper bounds for

1991 Mathematics Subject Classification: 11J68, 11D61.

the number of solutions of (0.1). Schlickewei [32] obtained a quantitative version of the p-adic Subspace Theorem over number fields from [30] and was the first to derive an explicit uniform upper bound for the number of solutions of (0.1) for arbitrary n [31]. His results were improved later by Evertse [11], [9]. In another direction, Schlickewei [34] derived a quantitative version of the so-called Parametric Subspace Theorem and deduced from this an explicit upper bound for the number of solutions of (0.1) depending only on the rank r of G (that is the rank of G modulo its torsion subgroup), the number of unknowns n and the degree d of K [36]. An important open problem was, to remove the dependence on d, that is, to derive an upper bound depending only on r and n. The dependence on d was caused inter alia by a dependence of the bound in Schlickewei's quantitative Parametric Subspace Theorem on the discriminant of K; so another important open problem was to remove the discriminant from this result of Schlickewei.

In this survey paper, we present among others an improvement of Schlickewei's quantitative Parametric Subspace Theorem which is indeed independent of the discriminant. In fact, we present an "absolute" generalisation in which the unknowns are taken from the algebraic closure $\overline{\mathbf{Q}}$ of $\mathbf{Q}$ instead of from a number field K. A complete proof will be published in [14]. The main new ingredient is what may be viewed as an absolute Minkowski's theorem proved by Roy and Thunder ([27] Thm. 6.3; [28] Thm. 2).

As a consequence we proved, together with W.M. Schmidt, the following result:

Suppose that G has rank r; then the number of solutions of eq. (0.1) *with property* (0.2) *is at most* $c(n)^{r+2}$ *with* $c(n) = \exp\{(6n)^{4n}\}$.

The proof will be published in [15].

In Section 1 we introduce some notation. In Section 2 we give an overview of the history and explain the interrelationship between eq. (0.1) and the Subspace Theorem and in Section 3 we present our new results.

1. Notation

We introduce absolute values, norms and heights.

Let $M(\mathbf{Q}) = \{\infty\} \cup \{\text{prime numbers}\}$ be the set of places of $\mathbf{Q}$, $|\cdot|_\infty = |\cdot|$ the ordinary absolute value on $\mathbf{Q}$ and for every prime number p, $|\cdot|_p$ the p-adic absolute value on $\mathbf{Q}$ with $|p|_p = 1/p$. Now let K be an algebraic number field and denote by

$M(K)$ the set of all places of K,

$M^\infty(K)$ the set of infinite (archimedean) places of K,

$M^{\text{fin}}(K)$ the set of finite (non-archimedean) places of K.

We denote by $\mathbf{Q}_p$ the completion of $\mathbf{Q}$ at p and by K_v the completion of K at v. For every $v \in M(K)$, choose the absolute value $|\cdot|_v$ such that if v lies above $p \in M(\mathbf{Q})$, then $|\cdot|_v$ is a continuation of $|\cdot|_p$, i.e. $|x|_v = |x|_p$ for $x \in \mathbf{Q}$. We mostly deal with the normalised absolute value $\|\cdot\|_v$ on K given by

$$\|\cdot\|_v = |\cdot|_v^{d(v)} \quad \text{where } d(v) = \frac{[K_v : \mathbf{Q}_p]}{[K : \mathbf{Q}]}. \tag{1.1}$$

These normalised absolute values satisfy the product formula

$$\prod_{v \in M(K)} \|x\|_v = 1 \quad \text{for } x \in K^*.$$

Such normalised absolute values are introduced in precisely the same way for every finite extension of K. Thus we obtain for every finite extension F of K, every $v \in M(K)$ and every place w of F lying above v the relation

$$\|x\|_w = \|x\|_v^{d(w|v)} \quad \text{for } x \in K, \text{ where } d(w|v) = \frac{[F_w : K_v]}{[F : K]} \tag{1.2}$$

(here F_w denotes the completion of F at w). Recall that

$$\sum_{w|v} d(w|v) = 1, \tag{1.3}$$

where '$w|v$' means that the sum is taken over all places $w \in M(F)$ lying above v.

We fix an algebraic closure $\overline{\mathbf{Q}}$ of $\mathbf{Q}$ and assume that every number field K is contained in $\overline{\mathbf{Q}}$. For every $v \in M(K)$ we choose and then fix henceforth a continuation of $\|\cdot\|_v$ to $\overline{\mathbf{Q}}$ (by continuing $\|\cdot\|_v$ to the algebraic closure $\overline{K_v}$ of K_v and choosing an isomorphic embedding of $\overline{\mathbf{Q}}$ into $\overline{K_v}$) and denote this also by $\|\cdot\|_v$. Thus, for every number field K and every $v \in M(K)$ we have an absolute value $\|\cdot\|_v$ on $\overline{\mathbf{Q}}$.

We introduce v-adic norms and heights for points $\mathbf{x} = (x_1, \ldots, x_n) \in \overline{\mathbf{Q}}^n$. Given $\mathbf{x}$, let K be a number field with $\mathbf{x} \in K^n$. For $v \in M(K)$ put

$$\|\mathbf{x}\|_v := \max(\|x_1\|_v, \ldots, \|x_n\|_v).$$

Then the height of $\mathbf{x}$ is defined by

$$H(\mathbf{x}) := \prod_{v \in M(K)} \|\mathbf{x}\|_v\,.$$

By (1.2), (1.3) this does not depend on the choice of K. Occasionally, we need another height H_2 which is defined by taking Euclidean norms at the infinite places. That is, for $\mathbf{x} = (x_1, \ldots, x_n) \in \overline{\mathbf{Q}}^n$ we define

$$H_2(\mathbf{x}) = \prod_{v \in M(K)} \|\mathbf{x}\|_{v,2}\,,$$

where K is any number field with $\mathbf{x} \in K^n$ and where

$$\|\mathbf{x}\|_{v,2} = \Big(\big(\sum_{i=1}^{n} |x_i|_v^2\big)^{1/2}\Big)^{d(v)} \quad \text{for } v \in M^\infty(K),$$
$$\|\mathbf{x}\|_{v,2} = \|\mathbf{x}\|_v \quad \text{for } v \in M^{\text{fin}}(K).$$

Thus,

$$H(\mathbf{x}) \le H_2(\mathbf{x}) \le n^{1/2} H(\mathbf{x}) \quad \text{for } \mathbf{x} \in \overline{\mathbf{Q}}^n .$$

For a linear form $L = a_1 X_1 + \ldots + a_n X_n$ with coefficient vector $\mathbf{a} = (a_1, \ldots, a_n)$ in $\overline{\mathbf{Q}}^n$, a number field K and $v \in M(K)$ we define

$$\|L\|_v = \|\mathbf{a}\|_v, \quad H(L) = H(\mathbf{a}), \quad H_2(L) = H_2(\mathbf{a}).$$

As usual, for a number field K and a finite set of places S on K containing the infinite places, we define

$\mathcal{O}_S = \{x \in K : \|x\|_v \le 1 \quad \text{for } v \notin S\}$: the ring of S-integers,

$\mathcal{O}_S^* = \{x \in K : \|x\|_v = 1 \quad \text{for } v \notin S\}$: the multiplicative group of S-units.

2. History

We start with recalling Schmidt's quantitative Subspace Theorem from 1989. Let $L_i = \alpha_{i1} X_1 + \ldots + \alpha_{in} X_n$ $(i = 1, \ldots, n)$ be n linearly independent linear forms with coefficients in $\overline{\mathbf{Q}}$ such that

$$H_2(L_i) \le H_2 \quad \text{for } i = 1, \ldots, n, \qquad [\mathbf{Q}(\{\alpha_{ij} : 1 \le i, j \le n\}) : \mathbf{Q}] \le D. \tag{2.1}$$

Consider the inequality

$$|L_1(\mathbf{x}) \cdots L_n(\mathbf{x})| \le |\det(L_1, \ldots, L_n)| \cdot H_2(\mathbf{x})^{-\delta} \quad \text{in } \mathbf{x} \in \mathbf{Z}^n, \tag{2.2}$$

where $\det(L_1, \ldots, L_n) = \det((\alpha_{ij})_{1 \le i,j \le n})$ and where $0 < \delta \le 1$.

Theorem A (Schmidt [42]). *The set of solutions of* (2.2) *with*

$$H_2(\mathbf{x}) \ge \max\left((n!)^{8/\delta}, H_2\right) \tag{2.3}$$

is contained in some finite union $T_1 \cup \ldots \cup T_a$ of proper linear subspaces of $\mathbf{Q}^n$ with

$$a \le (2D)^{2^{26n}\delta^{-2}}.$$

Schlickewei [32] proved a generalisation of Theorem A over number fields allowing an arbitrary finite set of absolute values. One of the main ingredients in the proofs of Schmidt and Schlickewei was Roth's lemma, a non-vanishing result for polynomials proved by Roth in 1955 [26]. In [10], Evertse derived a sharpening of Roth's lemma[1)] and by means of this, in [11] he considerably improved upon

1) In his paper [16] (cf. Section 3), Faltings proved in a non-explicit form his Product theorem which is a far-reaching generalisation of Roth's lemma. Part of the arguments in Faltings' proof were made explicit by van der Put [25]. Evertse went further on this and worked out a completely explicit version of Faltings' Product theorem [10]. A similar explicit version of the Product theorem was obtained independently by Ferretti [17]. Evertse obtained his sharpening of Roth's lemma by slightly refining the techniques used in the proof of the Product theorem.

Theorem A and Schlickewei's generalisation. We recall only Evertse's quantitative Subspace Theorem. For a field F and a linear form $L = a_1X_1 + \ldots + a_nX_n$ with coefficients in some extension of F, we define the field $F(L) = F(a_1/a_i, , \ldots, a_n/a_i)$ for any index i with $a_i \neq 0$. Let K be an algebraic number field and S a finite set of places on K of cardinality s, containing all infinite places. For $v \in S$, let $L_1^{(v)}, \ldots, L_n^{(v)}$ be linearly independent linear forms in $X_1, \ldots, X_n$ with coefficients in $\overline{\mathbf{Q}}$ such that

$$H_2(L_i^{(v)}) \leq H_2, \quad [K(L_i^{(v)}) : K] \leq D \quad \text{for } v \in S,\ i = 1, \ldots, n. \tag{2.4}$$

Consider the analogue of (2.2) for number fields,

$$\prod_{v \in S} \prod_{i=1}^{n} \frac{\|L_i^{(v)}(\mathbf{x})\|_v}{\|\mathbf{x}\|_{v,2}} \leq \Big(\prod_{v \in S} \|\det(L_1^{(v)}, \ldots, L_n^{(v)})\|_v \Big) \cdot H_2(\mathbf{x})^{-n-\delta} \quad \text{in } \mathbf{x} \in K^n \tag{2.5}$$

with $0 < \delta \leq 1$. Then one has [11]:

Theorem B. *The set of solutions of* (2.5) *with*

$$H_2(\mathbf{x}) \geq H_2 \tag{2.6}$$

is contained in some finite union $T_1 \cup \ldots \cup T_a$ of proper linear subspaces of K^n with

$$a \leq \big(2^{60n^2}\delta^{-7n}\big)^s \cdot \log 4D \log\log 4D. \tag{2.7}$$

Let us now turn to applications of the Subspace Theorem. Let K be an algebraic number field and let G be a finitely generated subgroup of the multiplicative group K^*. Consider the equation

$$\left.\begin{array}{l} a_1x_1 + \ldots + a_nx_n = 1 \quad \text{in } x_1, \ldots, x_n \in G \\ \qquad \text{with } \sum_{i \in I} a_ix_i \neq 0 \text{ for each non-empty subset } I \text{ of } \{1, \ldots, n\}, \end{array}\right\} \tag{2.8}$$

where $a_1, \ldots, a_n \in K^*$. As mentioned in the Introduction, the (qualitative) Subspace Theorem implies that (2.8) has only finitely many solutions.

By using his quantitative version of the Subspace Theorem from [32], Schlickewei [31] derived an explicit upper bound for the number of solutions of (2.8) in the case when G is the group of S-units:

Theorem C. *Let K be an algebraic number field of degree d, let S be a set of places of K of finite cardinality s containing all infinite places, and let $G = \mathcal{O}_S^*$. Then* (2.8) *has at most*

$$(4sd!)^{2^{36nd!}s^6}$$

solutions.

Later, he improved this to $2^{2^{27n}s}$ [36]. Using his Theorem B, Evertse [9] further improved Schlickewei's bound to

$$(2^{35}n^2)^{n^3s}. \tag{2.9}$$

Note that $\mathcal{O}_S^*$ (that is, $\mathcal{O}_S^*$ modulo its torsion subgroup) has rank $s-1$ and that $s \geq d/2$ (since S contains all infinite places). Hence Theorem C and its improvements give for $G = \mathcal{O}_S^*$ an upper bound for the number of solutions of (2.8) depending only on n and the rank of $\mathcal{O}_S^*$. The ultimate goal was, to obtain for arbitrary finitely generated multiplicative groups G an upper bound for the number of solutions of (2.8) depending only on n and the rank of G. Theorem C and its improvements imply for finitely generated subgroups G of K^* only an upper bound depending on s, where s is the cardinality of the smallest set of places S such that S contains all infinite places and $G \subset \mathcal{O}_S^*$. The number s can be much larger than the rank of G, for instance if G is a cyclic group with a generator α with $\|\alpha\|_v \neq 1$ for precisely s places v.

If one applies to (2.8) a quantitative version of the Subspace Theorem such as Theorem B with an upper bound for the number of subspaces depending on s, one necessarily obtains an upper bound for the number of solutions of (2.8) depending on s and one cannot exploit the fact that G has rank much smaller than s. Schlickewei considered a different approach, by reducing (2.8) to the *Parametric Subspace Theorem.* The latter can be stated as follows. Let K, S be as above. For $v \in S$, let $L_1^{(v)}, \dots, L_n^{(v)}$ be linearly independent linear forms in $X_1, \dots, X_n$ with coefficients in $\overline{\mathbf{Q}}$. For a *fixed* tuple of reals $\mathbf{c} = (c_{iv} : v \in S,\ i = 1, \dots, n)$ and a *varying* parameter $Q \geq 1$, define the "parallelepiped"

$$\Pi(Q, \mathbf{c}) = \{\mathbf{x} \in \mathcal{O}_S^n : \|L_i^{(v)}(\mathbf{x})\|_v \leq Q^{c_{iv}} \quad \text{for } v \in S,\ i = 1, \dots, n\}. \tag{2.10}$$

Parametric Subspace Theorem (Qualitative version). *Let* $\mathbf{c}$ *be a fixed tuple with*

$$\sum_{v \in S} \sum_{i=1}^{n} c_{iv} =: -\delta < 0.$$

Then there are finitely many proper linear subspaces $T_1, \dots, T_b$ *of* K^n *such that for every* $Q \geq 1$ *we have*

$$\Pi(Q, \mathbf{c}) \subset T_1 \cup \ldots \cup T_b\,.$$

The Parametric Subspace Theorem is in fact equivalent to the (qualitative) Subspace Theorem. We sketch how the Parametric Subspace Theorem implies the Subspace Theorem, i.e., that the set of solutions of

$$\prod_{v \in S} \prod_{i=1}^{n} \frac{\|L_i^{(v)}(\mathbf{x})\|_v}{\|\mathbf{x}\|_v} \leq H(\mathbf{x})^{-n-\delta} \quad \text{in } \mathbf{x} \in K^n \tag{2.11}$$

is contained in the union of finitely many proper linear subspaces of K^n. Let $\mathbf{x}$ be a solution of (2.11) with $L_i^{(v)}(\mathbf{x}) \neq 0$ for $v \in S$, $i = 1, \dots, n$. Note that (2.11) is homogeneous in $\mathbf{x}$. By replacing $\mathbf{x}$ by a scalar multiple if necessary, we may assume that $\mathbf{x} \in \mathcal{O}_S^n$ and $\prod_{v \in S} \|\mathbf{x}\|_v \ll H(\mathbf{x})$, where the constant implied by $\ll$ depends on K. Then (2.11) implies $\prod_{v \in S} \prod_{i=1}^{n} \|L_i^{(v)}(\mathbf{x})\|_v \ll H(\mathbf{x})^{-\delta}$. Define Q

and the tuple $\mathbf{c}$ by

$$Q = H(\mathbf{x}), \qquad \|L_i^{(v)}(\mathbf{x})\|_v = Q^{c_{iv}} \quad \text{for } v \in S,\ i = 1, \dots, n.$$

Then clearly,

$$\mathbf{x} \in \Pi(Q, \mathbf{c}), \qquad \sum_{v \in S} \sum_{i=1}^{n} c_{iv} \ll -\delta < 0.$$

But the tuple $\mathbf{c}$ varies with $\mathbf{x}$ whereas the Parametric Subspace Theorem requires $\mathbf{c}$ to be fixed. However, one can show that if $\mathbf{x}$ runs through the set of solutions of (2.11), then $\mathbf{c}$ runs through a bounded subset of $\mathbf{R}^{ns}$, where s is the cardinality of S. By covering this bounded set with small cubes, one infers that there is a finite set $\mathcal{C}$ in $\mathbf{R}^{ns}$ such that every $\mathbf{c}$ in the bounded set is very close to some $\mathbf{c}' \in \mathcal{C}$. More precisely one can show that there is a finite set $\mathcal{C}$ of cardinality $\leq c(n, \delta)^{ns}$ (with $c(n, \delta)$ a function of n and δ only) such that for every solution $\mathbf{x}$ of (2.11) there is a $\mathbf{c}' = (c'_{iv} : v \in S,\ i = 1, \dots, n) \in \mathcal{C}$ very close to $\mathbf{c}$ with

$$\mathbf{x} \in \Pi(Q, \mathbf{c}'), \qquad \sum_{v \in S} \sum_{i=1}^{n} c'_{iv} \leq -\frac{\delta}{2}\,.$$

Now by applying the Parametric Subspace Theorem to $\Pi(Q, \mathbf{c}')$ for every $\mathbf{c}' \in \mathcal{C}$ we infer that there is a union of finitely many proper linear subspaces of K^n containing the set of solutions of (2.11). $\square$

This argument implies that if we had a quantitative version of the Parametric Subspace Theorem with a uniform upper bound for the number b of subspaces, then by multiplying this with $c(n, \delta)^{ns}$ we would obtain, similarly as in Theorem B, an upper bound depending on s for the number of subspaces in the Subspace Theorem. We will see later that in contrast, the number of subspaces in the Parametric Subspace Theorem can be estimated from above independently of s.

We consider again eq. (2.8) where G is a subgroup of K^* of rank r. We sketch Schlickewei's argument to reduce (2.8) to the Parametric Subspace Theorem. If we want to derive an upper bound for the number of solutions of (2.8) depending only on n and r we may as well assume that all coefficients $a_1, \dots, a_n$ of (2.8) are equal to 1, since if we add $a_1, \dots, a_n$ as new generators to G, then the rank of G increases by at most n. That is, we may restrict ourselves to the equation

$$x_1 + \dots + x_n = 1 \quad \text{in } x_1, \dots, x_n \in G. \tag{2.12}$$

We choose a number field K and a finite set of places S on K, containing all infinite places, such that $G \subset \mathcal{O}_S^*$. Let $\mathbf{x} = (x_1, \dots, x_n)$ be a solution of (2.12), put $x_0 := 1$ and choose

$$Q = H(\mathbf{x}') \quad \text{where } \mathbf{x}' = (x_0, x_1, \dots, x_n). \tag{2.13}$$

Define the tuple of reals $\mathbf{e} = (e_{iv} : v \in S, i = 0, \dots, n)$ by

$$\|x_i\|_v = Q^{e_{iv}} \quad \text{for } v \in S,\ i = 0, \dots, n. \tag{2.14}$$

For $v\in S$ choose $i(v)$ from $\{0,\dots,n\}$ such that

$$e_{i(v),v}=\max(e_{0v},\dots,e_{nv}).$$

Now choose linear forms $L_i^{(v)}$ ($v\in S,\ i=1,\dots,n$) and a tuple of reals $\mathbf{c}=(c_{iv}:v\in S,\ i=1,\dots,n)$ such that for $v\in S$ we have

$$\begin{aligned}\{L_1^{(v)},\dots,L_n^{(v)}\}&=\{X_0,X_1,\dots,X_n\}\backslash\{X_{i(v)}\}\\ &\qquad\text{where } X_0:=X_1+\ldots+X_n\,,\end{aligned}\tag{2.15}$$

$$c_{iv}=e_{jv}\quad\text{if } L_i^{(v)}=X_j.\tag{2.16}$$

Then clearly, $\mathbf{x}\in\Pi(Q,\mathbf{c})$, where $\Pi(Q,\mathbf{c})$ is given by (2.10). Further, by the product formula and $x_0,\dots,x_n\in\mathcal{O}_S^*$ we have

$$\prod_{v\in S}\|x_i\|_v=1,\quad\text{whence }\sum_{v\in S}e_{iv}=0\quad\text{for } i=0,\dots,n$$

and by $\|\mathbf{x}'\|_v=Q^{e_{i(v),v}}$ for $v\in S$, $\|\mathbf{x}'\|_v=1$ for $v\notin S$ and (2.13) we have

$$\sum_{v\in S}e_{i(v),v}=1.$$

Hence

$$\sum_{v\in S}\sum_{i=1}^{n}c_{iv}=-1<0.$$

Now the tuple $\mathbf{e}$ defined by (2.14) varies with $\mathbf{x}$ and therefore so do the tuple $(i(v):v\in S)$ and the tuple $\mathbf{c}$. If $\mathbf{x}$ runs through the solutions of (2.12), then the vector $\mathbf{x}'=(1,x_1,\dots,x_n)$ runs through a group of rank nr. Using this, one can prove that the tuple $\mathbf{e}$ defined by (2.14) runs through an nr-dimensional linear subspace of $\mathbf{R}^{(n+1)s}$, where s is the cardinality of S. Schlickewei showed in [34] that if $\mathbf{x}$ runs through the solutions of (2.12), then $\mathbf{e}$ runs through a bounded subset of this linear subspace and moreover, that every element from this bounded subset can be closely approximated by an element from a finite set $\mathcal{D}$ of cardinality at most a function $c_1(n,r)$ of n and r. More precisely, Schlickewei proved that for the tuple $\mathbf{e}$ defined by (2.14), one can choose a tuple $\mathbf{e}'$ from $\mathcal{D}$ close to $\mathbf{e}$ such that if $(i(v):v\in S)$ and $\mathbf{c}$ are defined in precisely the same way as above but with $\mathbf{e}'$ replacing $\mathbf{e}$, then $\mathbf{x}\in\Pi(Q,\mathbf{c})$, where Q is slightly larger than $H(\mathbf{x}')$ and

$$\sum_{v\in S}\sum_{i=1}^{n}c_{iv}\le-\tfrac{99}{100},\tag{2.17}$$

say. Stated otherwise, we have $\mathbf{x}\in\Pi(Q,\mathbf{c})$, where the tuple of linear forms and reals $(\{L_i^{(v)}\};\ \mathbf{c})$ satisfies (2.15) and (2.17) and belongs to a collection of cardinality $c_1(n,r)$ independent of $\mathbf{x}$. $\square$

Let us speculate and let us suppose that (for linear forms $L_i^{(v)}$ defined by (2.15)) the number of subspaces b in the Parametric Subspace Theorem is bounded above by a quantity depending only on $\delta=-\left(\sum_{v\in S}\sum_{i=1}^{n}c_{iv}\right)$ and n, say. Then by substituting $\delta = 99/100$ and multiplying this quantity with $c_1(n,r)$ we would obtain an upper bound $c_2(n,r)$ for the number of proper linear subspaces of K^n containing the set of solutions of (2.12). By an argument using induction on n we would then deduce an upper bound for the number of solutions of (2.12) or of (2.8) depending on n and r only.

We recall Schlickewei's quantitative Parametric Subspace Theorem from [34]. Let K be a number field, D_K the discriminant of K and $d = [K : \mathbf{Q}]$. Further, let S be a finite set of places of K, containing all infinite places and for $v \in S$, let $L_1^{(v)}, \dots, L_n^{(v)}$ be linearly independent linear forms with

$$\{L_1^{(v)}, \dots, L_n^{(v)}\} \subset \{X_1, \dots, X_n, X_1 + \dots + X_n\} \tag{2.18}$$

(which is sufficient for applications to (2.8) in view of (2.15)). Lastly, let $\mathbf{c} = (c_{iv} : v \in S,\ i = 1, \dots, n)$ be a tuple of reals with

$$\sum_{v\in S}\sum_{i=1}^{n} c_{iv} \le -\delta \quad \text{with } 0 < \delta \le 1, \qquad \sum_{v\in S}\sum_{i=1}^{n} |c_{iv}| \le 1 \tag{2.19}$$

(the second inequality is some normalisation assumption). The following result is a slight reformulation of Lemma 6.1 of [34].

Theorem D. *There are proper linear subspaces $T_1, \dots, T_b$ of K^n, with*

$$b \le 2^{2^{22n}\delta^{-2}}$$

such that for every Q satisfying

$$Q \ge \max(n^{2n/\delta}, |D_K|^{1/2d}) \quad \textit{and a technical condition} \tag{2.20}$$

we have $\Pi(Q, \mathbf{c}) \subset T_1 \cup \dots \cup T_b$.

The technical condition in (2.20) is too complicated to be stated here but quite harmless. The lower bound for Q is a much more serious problem. When applying Theorem D to (2.8), Q roughly speaking corresponds to the height of a solution of (2.8) and δ is a constant (cf. (2.13), (2.17)). So Theorem D can be applied only to the "large" solutions of (2.8). Schlickewei managed to determine an explicit upper bound for the number of "small" solutions of (2.8), depending on n, r and the degree d of $[K : \mathbf{Q}]$. Thus, he obtained the following result [36]:

Theorem E. *Let K be an algebraic number field of degree d and G a subgroup of rank r of the multiplicative group K^*. Then the number of solutions of* (2.8) *is at most*

$$2^{2^{26n}} \cdot 2^{16n^4 r + 4n^2 r^2} \cdot d^{6n^2(r+1)}.$$

We recall that for $G = \mathcal{O}_S^*$ we have $d \le 2(\operatorname{rank} \mathcal{O}_S^* + 1)$.

By applying Evertse's sharpening of Roth's lemma [10], Schlickewei [38] improved the upper bound for b in Theorem D to $8^{(n+5)^2}\delta^{-n-4}$, but with the same condition (2.20) imposed on Q. Schlickewei and Schmidt [39] then improved the bound in Theorem E to $(2d)^{41n^3r}r^{n^2r}$.

Schmidt (Theorem 5 of [44]; cf. Theorem I in Section 3 of the present paper) obtained an upper bound for the number of "very small" solutions of (2.8). In combination with a hypothetical version of Theorem D where (2.20) is replaced by a lower bound for Q depending only on n and δ, this would have given an upper bound for the total number of solutions of (2.8) depending only on n and r. Thus, to obtain such an upper bound, the term $|D_K|^{1/2d}$ in (2.20) is the only remaining obstacle.

We explain why the term $|D_K|^{1/2d}$ in (2.20) introduces a dependence on d in the upper bound for the number of solutions of (2.8). To this end, we use the following "Gap principle," ([34], Lemma 8.6) which states that if Q runs through a small interval, then the points in $\Pi(Q, \mathbf{c})$ run through a small number of subspaces.

Gap principle. *Let $L_i^{(v)}$ $(v \in S,\ i = 1, \ldots, n)$ be linear forms satisfying* (2.18) *and $\mathbf{c}$ a tuple satisfying* (2.19). *For every $Q_0 \geq n^{2n/\delta}$, $E > 1$, there are proper linear subspaces $T_1, \ldots, T_k$ of K^n with*

$$k \leq 1 + \frac{4n}{\delta} \log E,$$

such that for every $Q \in [Q_0, Q_0^E)$ we have $\Pi(Q, \mathbf{c}) \subset T_1 \cup \ldots \cup T_k$.

Assume (2.18), (2.19). Let $\mathbf{x} = (x_1, \ldots, x_n) \in \Pi(Q, \mathbf{c})$. Suppose that for some j with $x_j \neq 0$, the quotients x_i/x_j $(1 \leq i \leq n)$ generate K and are not all equal to roots of unity. Under this hypothesis, Silverman [46] (Theorem 2) showed that

$$H(\mathbf{x}) \geq d^{-1/d} |D_K|^{1/2d(d-1)}.$$

Now from

$$\|\mathbf{x}\|_v \ll \max_{1 \leq i \leq n} \|L_i^{(v)}\|_v \ll Q^{\max(c_{1v}, \ldots, c_{nv})} \ll Q^{\sum_{i=1}^n |c_{iv}|} \quad \text{for } v \in S,$$
$$\|\mathbf{x}\|_v \leq 1 \quad \text{for } v \notin S \text{ since } \mathbf{x} \in \mathcal{O}_S^n\ ,$$

and (2.19) it follows that $H(\mathbf{x}) \ll Q$. (Here and below constants implied by $\ll$ depend only on n and δ.) Hence

$$Q \gg |D_K|^{1/2d(d-1)}.$$

By the Gap principle, the union of the sets $\Pi(Q, \mathbf{c})$, with $|D_K|^{1/2d(d-1)} \ll Q < |D_K|^{1/2d}$, is contained in the union of $\ll \log d$ proper linear subspaces of K^n. Therefore, in order to incorporate the solutions $\mathbf{x}$ with $Q \leq |D_K|^{1/2d}$ we have to add a quantity $\ll \log d$ to the upper bound for the number of subspaces in Theorem D. Thus, the final result on the number of solutions of (2.8) involves the parameter d. □

We now review some results about the equation in two unknowns

$$ax + by = 1 \quad \text{in } x, y \in G, \tag{2.21}$$

where as above G is a finitely generated subgroup of K^* of rank r and where $a, b \in K^*$. Schlickewei [34] proved for the case $n = 2$ a version of Theorem D independent of the discriminant. Further, he derived [37] an estimate for the number of "small" solutions of (2.21) (which preceded Schmidt's result mentioned above). By combining these results he obtained in the same paper [37] an upper bound for the total number of solutions of (2.21) depending only on the rank r. Schlickewei considered the general case that G is a *finite type* subgroup of the multiplicative group of complex numbers $\mathbf{C}^*$. Here, G is called a finite type group if it has a free subgroup G_0 of finite rank such that G/G_0 is a torsion group; the rank of G is then defined as the rank of G_0. By a simple argument, Schlickewei reduced the general case to the special case that G is a finitely generated multiplicative group in some number field. His result is as follows:

Theorem F. *Let G be a finite type subgroup of $\mathbf{C}^*$ of rank r and $a, b \in \mathbf{C}^*$. Then* (2.21) *has at most $2^{2^{26}+9r^2}$ solutions.*

Later, Schlickewei and Schmidt [39] improved this to $2^{14r+63}r^{2r}$. By using hypergeometric functions instead of Theorem D in dimension 2, Beukers and Schlickewei [1] obtained the bound $2^{16(r+1)}$. This last result is comparable to Evertse's upper bound 3×7^{4s} for the case $G = \mathcal{O}_S^*$ where S has cardinality s [7].

By a very different method, Bombieri, Mueller and Poe [3] showed that if G has rank r and is contained in a number field of degree d, then (2.21) has at most $d^{9r}2^{125r^2}$ solutions. They obtained their result by extending an idea of Poe [23] to a general "cluster principle" entailing that the solutions of (2.21) can be divided into clusters of solutions lying close together, and by combining this with an effective upper bound for the heights of the solutions of (2.21) obtained by means of lower bounds for linear forms in logarithms.

Silverman [47] showed that for any algebraic number field K of degree d and any given element α of the unit group $\mathcal{O}_K^*$ of the ring of integers of K, the equation $\alpha^m + \varepsilon = 1$ has at most $d^{1+o(1)}$ solutions in $m \in \mathbf{Z}$, $\varepsilon \in \mathcal{O}_K^*$, in other words, there are at most $d^{1+o(1)}$ integers m such that α^m is an exceptional unit. This does not follow from any of the results mentioned above.

Further information about equations (2.8) and (2.21) and their applications can be found in the survey papers [18] and [13].

3. New results

We present an improvement of Theorem D which is independent of the discriminant. In his proof of Theorem D, Schlickewei used a generalisation to number fields of Minkowski's theorem on successive minima, proved by McFeat [22] and later independently by Bombieri and Vaaler [4]. We used instead an "absolute Minkowski's theorem" of Roy and Thunder ([27], Thm. 6.3; [28], Thm. 2). First, we recall the result of McFeat and Bombieri and Vaaler, as well as that of Roy and Thunder.

Let K be a number field of degree d and of discriminant D_K. For every $v \in M(K)$ we extend $\|\cdot\|_v$ to the completion K_v. Let S be a finite set of places on K, containing all infinite places. Thus,

$$S = M^\infty(K) \cup S^{\text{fin}},$$

where S^{fin} consists of the finite places in S. For $v \in S$, let $L_1^{(v)}, \ldots, L_n^{(v)}$ be linearly independent linear forms in $X_1, \ldots, X_n$ with coefficients in K_v. Let $Q \geq 1$ be a real, and $\mathbf{c} = (c_{iv} : v \in S,\ i = 1, \ldots, n)$ a tuple of reals and define as before,

$$\Pi(Q, \mathbf{c}) = \{\mathbf{x} \in \mathcal{O}_S^n : \|L_i^{(v)}(\mathbf{x})\|_v \leq Q^{c_{iv}} \quad \text{for } v \in S,\ i = 1, \ldots, n\}.$$

For reals $\lambda > 0$, define

$$\lambda\Pi(Q, \mathbf{c}) = \left\{ \mathbf{x} \in \mathcal{O}_S^n : \begin{array}{l} \|L_i^{(v)}(\mathbf{x})\|_v \leq \lambda^{d(v)} Q^{c_{iv}} \quad \text{for } v \in M^\infty(K),\ i = 1, \ldots, n, \\ \|L_i^{(v)}(\mathbf{x})\|_v \leq Q^{c_{iv}} \quad \text{for } v \in S^{\text{fin}},\ i = 1, \ldots, n \end{array} \right\},$$

where $d(v) = 1/d$ if $K_v = \mathbf{R}$ and $d(v) = 2/d$ if $K_v = \mathbf{C}$. For $i = 1, \ldots, n$, the i-th successive minimum $\lambda_i = \lambda_i(Q, \mathbf{c})$ of $\Pi(Q, \mathbf{c})$ is the infimum of all $\lambda > 0$ such that $\lambda\Pi(Q, \mathbf{c})$ contains i linearly independent vectors. Obviously, $\lambda_1 \leq \ldots \leq \lambda_n$. Put

$$\Delta = \prod_{v \in S} \|\det(L_1^{(v)}, \ldots, L_n^{(v)})\|_v, \qquad \delta = -\Big(\sum_{v \in S} \sum_{i=1}^n c_{iv}\Big). \tag{3.1}$$

The following result, which was used by Schlickewei in his proof of Theorem D, is a consequence of [22], Thm. 5, p. 15 and Thm. 6, p. 23 and of [4], Thm. 3, p. 18 and Thm. 6, p. 23.

Theorem G. *Suppose that for $v \in S^{\text{fin}}$, $i = 1, \ldots, n$, the number $Q^{c_{iv}}$ belongs to the value set of $\|\cdot\|_v$. Then*

$$n^{-n/2} \Delta Q^\delta \leq \lambda_1 \cdots \lambda_n \leq |D_K|^{n/2d} \Delta Q^\delta.$$

It is important to remark that the occurrence of the term $|D_K|^{1/2d}$ in (2.20) was caused only by the factor $|D_K|^{n/2d}$ in the upper bound for $\lambda_1 \cdots \lambda_n$.

Below, we give an analogous result for the "algebraic closures" of the sets $\Pi(Q, \mathbf{c})$, which is a consequence of the result of Roy and Thunder. Let F be a finite extension of K and denote by S_F the set of places of F lying above those in S. Thus, $\mathcal{O}_{S_F}$ is the integral closure of $\mathcal{O}_S$ in F. For each place $v \in S$ and for each place $w \in S_F$ lying above v we introduce linear forms $L_i^{(w)}$ and reals c_{iw} by

$$L_i^{(w)} = L_i^{(v)}, \qquad c_{iw} = d(w|v) \cdot c_{iv} \quad (i = 1, \ldots, n), \tag{3.2}$$

where $d(w|v)$ is given by (1.2). Define

$$\Pi_F(Q, \mathbf{c}) = \{\mathbf{x} \in \mathcal{O}_{S_F}^n : \|L_i^{(w)}(\mathbf{x})\|_w \leq Q^{c_{iw}} \quad \text{for } w \in S_F,\ i = 1, \ldots, n\}.$$

By (1.2), for every pair of finite extensions F, E of K with $F \subseteq E$ we have $\Pi_E(Q, \mathbf{c}) \cap F^n = \Pi_F(Q, \mathbf{c})$. Now we define the algebraic closure of $\Pi(Q, \mathbf{c})$ by

$$\overline{\Pi}(Q, \mathbf{c}) = \bigcup_{F \supseteq K} \Pi_F(Q, \mathbf{c}),$$

where the union is taken over all finite extensions F of K. Note that $\overline{\Pi}(Q,\mathbf{c}) \subset \overline{\mathcal{O}}_S^n$, where $\overline{\mathcal{O}}_S$ is the integral closure of $\mathcal{O}_S$ in $\overline{\mathbf{Q}}$.

For $\lambda > 0$ we define

$$\lambda\overline{\Pi}(Q,\mathbf{c}) = \bigcup_{F\supseteq K} \Big(\lambda\Pi_F(Q,\mathbf{c})\Big),$$

where for every finite extension F of K the set $\lambda\Pi_F(Q,\mathbf{c})$ is given by

$$\left\{ \mathbf{x} \in \mathcal{O}_{S_F}^n : \begin{array}{l} \|L_i^{(w)}(\mathbf{x})\|_w \le \lambda^{d(w)} Q^{c_{iw}} \text{ for } w \in M^\infty(F),\ i = 1,\dots,n, \\ \|L_i^{(w)}(\mathbf{x})\|_w \le Q^{c_{iw}} \quad \text{for } w \in S_F^{\mathrm{fin}},\ i = 1,\dots,n, \end{array} \right\}$$

(with $d(w) = \frac{1}{[F:\mathbf{Q}]}$ if $F_w = \mathbf{R}$ and $d(w) = \frac{2}{[F:\mathbf{Q}]}$ if $F_w = \mathbf{C}$). The i-th successive minimum $\overline{\lambda}_i = \overline{\lambda}_i(Q,\mathbf{c})$ of $\overline{\Pi}(Q,\mathbf{c})$ is the infimum of all $\lambda > 0$ such that $\lambda\overline{\Pi}(Q,\mathbf{c})$ contains i linearly independent vectors from $\overline{\mathcal{O}}_S^n$. The next result is a consequence of Roy and Thunder [28], Thm. 2, which in turn is a slight improvement of [27], Thm. 6.3.

Theorem H. $\overline{\Pi}(Q,\mathbf{c})$ *has precisely* n *successive minima with* $0 < \overline{\lambda}_1 \le \dots \le \overline{\lambda}_n < \infty$ *and*

$$n^{-n/2}\Delta Q^\delta \le \overline{\lambda}_1 \cdots \overline{\lambda}_n \le e^{n(n-1)/4}\Delta Q^\delta.$$

In Theorem H, there is no dependence on the discriminant of some number field but the price is, that we have no information about the number field generated by the coordinates of the vectors corresponding to the successive minima. In his proof of Theorem D, Schlickewei dealt only with vectors in K^n for some given number field K. Fortunately, we were able to extend Schlickewei's arguments in such a way that we could work with arbitrary vectors from $\overline{\mathbf{Q}}^n$. This allowed us to apply Theorem H instead of Theorem G. Thus, we succeeded to prove a Parametric Subspace Theorem which does not involve anymore the discriminant. In fact, since we had to deal with vectors in $\overline{\mathbf{Q}}^n$ anyhow, we were able to prove a quantitative "absolute" Parametric Subspace Theorem dealing with algebraic closures $\overline{\Pi}(Q,\mathbf{c}) \subset \overline{\mathbf{Q}}^n$ rather than sets $\Pi(Q,\mathbf{c}) \subset K^n$. Further, we considerably relaxed conditions (2.18) and (2.19).

Let K, S be as above, and let $L_i^{(v)}$ ($v \in S,\ i = 1,\dots,n$) be linear forms with the following properties:

$$\left.\begin{array}{l} \text{for } v \in S,\ \{L_1^{(v)},\dots,L_n^{(v)}\} \text{ is a linearly independent set of linear forms} \\ \text{in } X_1,\dots,X_n \text{ with coefficients in } K, \\ H(L_i^{(v)}) \le H,\ \|L_i^{(v)}\|_v = 1 \text{ for } v \in S,\ i = 1,\dots,n, \\ \text{there are exactly } R \text{ distinct sets among } \{L_1^{(v)},\dots,L_n^{(v)}\}\ (v \in S). \end{array}\right\} \tag{3.3}$$

Further, let $\mathbf{c} = (c_{iv} : v \in S,\ i = 1,\dots,n)$ be a fixed tuple of reals with

$$\sum_{v\in S}\sum_{i=1}^n c_{iv} \le -\delta \quad \text{with } 0 < \delta \le 1, \qquad \sum_{v\in S} \max(c_{1v},\dots,c_{nv}) \le 1 \tag{3.4}$$

and put

$$\Delta = \prod_{v \in S} \|\det(L_1^{(v)}, \dots, L_n^{(v)})\|_v \,.$$

The complete proof of the following result will be published in [14]:

Theorem 1. *There are proper linear subspaces* $T_1, \dots, T_b$ *of* $\overline{\mathbf{Q}}^n$*, all defined over* K*, with*

$$b \le 4^{(n+5)^2} \delta^{-n-4} \log 4R \cdot \log\log 4R$$

such that for every Q *with*

$$Q \ge \max\big(H,\ (n^{n/2}\Delta^{-1})^{2/\delta}\big) \tag{3.5}$$

there is an $i \in \{1, \dots, b\}$ *with*

$$\overline{\Pi}(Q, \mathbf{c}) \subset T_i \,.$$

In the special case when $\{L_1^{(v)}, \dots, L_n^{(v)}\} \subset \{X_1, \dots, X_n, X_1 + \dots + X_n\}$ for $v \in S$, we have $H = 1$, $R \le n+1$ and $\Delta = 1$, whence both the upper bound for b and the lower bound for Q depend only on n and δ. So Theorem 1 gives us precisely the improvement of Theorem D we were aiming at. Let G be a finitely generated multiplicative group contained in a number field. Using, as indicated in Section 1, Theorem 1 for the "large solutions" of (2.8) and Schmidt's result ([44], Thm. 5) for the "small" solutions, we obtained together with Schmidt an upper bound for the total number of solutions of (2.8) depending on n and the rank r of G only.

Van der Poorten and Schlickewei [24] showed that eq. (2.8) has only finitely many solutions for every finitely generated subgroup G of $\mathbf{C}^*$, by means of some specialisation argument, reducing to the case that G is contained in a number field. Together with some Kummer theory worked out by Laurent [19] this implies that (2.8) has only finitely many solutions if G is a finite type subgroup of $\mathbf{C}^*$. The specialisation argument can be considerably simplified and the Kummer theory can be avoided if one already knows that in the case when G is contained in some number field the number of solutions is bounded above by a function of n and r. Thus, together with Schmidt, we obtained the following result for arbitrary finite type subgroups of $\mathbf{C}^*$ [15]:

Theorem 2. *Let* G *be a finite type subgroup of* $\mathbf{C}^*$ *of rank* r *and* $a_1, \dots, a_n \in \mathbf{C}^*$*. Then the number of solutions of the equation)*

$$\left.\begin{array}{l} a_1x_1 + \dots + a_nx_n = 1 \quad \textit{in } x_1, \dots, x_n \in G \\ \qquad \textit{with } \sum_{i \in I} a_ix_i \ne 0 \textit{ for each non-empty subset } I \textit{ of } \{1, \dots, n\} \end{array}\right\}$$

is at most

$$c(n)^{r+2} \quad \textit{with } c(n) = \exp\big((6n)^{4n}\big).$$

From Theorem 1 we derived (in the case that G is contained in a number field) an upper bound for the number of "large" solutions of (2.8) depending exponentially on n^3. Unfortunately, Schmidt's estimate for the number of "small" solutions gave a contribution to the upper bound of Theorem 2 depending doubly exponentially on n. At the end of this section we discuss Schmidt's result in more detail.

In certain special cases, results much better than Theorem 2 are known. If G is the group of S-units $\mathcal{O}_S^*$ in some number field, Evertse's bound (2.9) is much sharper. Now suppose that G is the group of roots of unity in $\mathbf{C}^*$ (i.e. G has rank 0). Let $n_1 = n+1$. ¿From results of Mann [21] and Conway and Jones [6] it follows that if the coefficients $a_1, \ldots, a_n$ are rational numbers, then (2.8) has at most $\exp(2n_1^{3/2} \log n_1)$ solutions. Schlickewei [33] derived the upper bound $\exp(4n_1!)$ if $a_1, \ldots, a_n$ are arbitrary complex numbers. Recently, Evertse [12] improved this to $\exp(3n_1^2 \log n_1)$.

We mention that to prove Theorem 2 already a "non-absolute" Parametric Subspace Theorem (i.e., dealing with sets $\Pi(Q, \mathbf{c})$ and not with their algebraic closures) would have sufficed, as long as the result would not have involved the discriminant. The absolute generalisation as stated in Theorem 1 dealing with algebraic closures $\overline{\Pi}(Q, \mathbf{c})$ was not necessary but we obtained this as a by-product of some independent interest.

We present some corollaries of Theorems 1 and 2. Our first corollary is a consequence of Theorem 2 for recurrence sequences. Let $U = \{u_m\}_{m \in \mathbf{Z}}$ be a sequence of complex numbers satisfying a recurrence relation of order q,

$$u_m = c_1 u_{m-1} + \ldots + c_q u_{m-q}$$

with $c_1, \ldots, c_q \in \mathbf{C}$, $c_q \neq 0$. As is well-known, we have

$$u_m = \sum_{i=1}^{n} g_i(m) \alpha_i^m \quad \text{for } m \in \mathbf{Z},$$

where $\alpha_1, \ldots, \alpha_n$ are distinct, non-zero complex numbers and $g_1, \ldots, g_n \in \mathbf{C}[X]$ polynomials with

$$\prod_{i=1}^{n} (X - \alpha_i)^{\deg g_i + 1} = X^q - c_1 X^{q-1} - \ldots - c_q \, .$$

Denote by $N_U(a)$ the number of integers m with

$$u_m = a.$$

The sequence U is called *non-degenerate* if neither $\alpha_1, \ldots, \alpha_n$, nor any of the quotients α_i/α_j $(1 \leq i < j \leq n)$ is a root of unity. From the Theorem of Skolem-Mahler-Lech (cf. [20]) it follows that then $N_U(a)$ is finite for every $a \in \mathbf{C}$. Using his Theorem D, Schlickewei [34] showed that if U is non-degenerate, and $\alpha_1, \ldots, \alpha_n$ and the coefficients of $g_1, \ldots, g_n$ generate an algebraic number field K of degree d, then for every $a \in K$ we have

$$N_U(a) \leq d^{6q^2} 2^{2^{28q!}}.$$

If we assume that $g_1, \dots, g_n$ are all constant, we obtain the following improvement, by applying Theorem 2 to the group G generated by $\alpha_1, \dots, \alpha_n$ which has at most rank n:

Corollary 3. *Let U be a recurrence sequence satisfying*

$$u_m = g_1\alpha_1^m + \ldots + g_n\alpha_n^m \quad \textit{for } m \in \mathbf{Z},$$

where $\alpha_1, \dots, \alpha_n$ are non-zero complex numbers such that neither $\alpha_1, \dots, \alpha_n$, nor any of the quotients α_i/α_j $(1 \le i < j \le n)$ is a root of unity and where $g_1, \dots, g_n$ are non-zero complex numbers. Then for every $a \in \mathbf{C}$ we have

$$N_U(a) \le \exp\big((n+2) \times (6n)^{4n}\big).$$

We mention that for $n = 2$, Schlickewei [35] had previously established an absolute bound for $N_U(a)$. His bound had been substantially improved by Beukers and Schlickewei [1] who showed $N_U(a) \le 61$. Very recently, Schmidt [45] obtained the remarkable result that for arbitrary non-degenerate complex recurrence sequences U of order q (i.e., with arbitrary polynomials $g_1, \dots, g_n$) one has $N_U(a) \le C(q)$, where $C(q)$ depends only (and in fact triply exponentially) on q. His proof uses Corollary 3 stated above.

We now present some consequences of Theorem 1 for Diophantine inequalities whose proofs will be published in [14]. In what follows, K, S are as above, $L_i^{(v)}$ $(v \in S,\ i = 1, \dots, n)$ are linear forms satisfying (3.3) and $\mathbf{c}$ is a tuple of reals with (3.4). Further, we put $\Delta = \prod_{v \in S} \|\det(L_1^{(v)}, \dots, L_n^{(v)}\|_v$. For every finite extension F of K, the linear forms $L_i^{(w)}$ and the reals c_{iw} are defined by (3.2).

Consider for every finite extension F of K the system of inequalities

$$\|L_i^{(w)}(\mathbf{x})\|_w \le H(\mathbf{x})^{c_{iw}} \quad (w \in S_F,\ i = 1, \dots, n) \quad \text{in } \mathbf{x} \in \mathcal{O}_{S_F}^n. \tag{3.6}$$

Note that every $\mathbf{x}$ satisfying (3.6) for some finite extension F of K belongs to $\overline{\Pi}(Q, \mathbf{c})$ with $Q = H(\mathbf{x})$. Therefore, Theorem 1 implies at once:

Corollary 4. *There are proper linear subspaces $T_1, \dots, T_b$ of $\overline{\mathbf{Q}}^n$, all defined over K, with*

$$b \le 4^{(n+5)^2} \delta^{-n-4} \log 4R \cdot \log\log 4R$$

such that for every finite extension F of K, the set of solutions of (3.6) *with*

$$H(\mathbf{x}) \ge \max\big(H,\ (n^{n/2}\Delta^{-1})^{2/\delta}\big)$$

is contained in $T_1 \cup \ldots \cup T_b$.

Now consider for every finite extension F of K the inequality

$$\prod_{w \in S_F} \prod_{i=1}^{n} \frac{\|L_i^{(w)}(\mathbf{x})\|_w}{\|\mathbf{x}\|_w} \le H(\mathbf{x})^{-n-\delta} \quad \text{in } \mathbf{x} \in F^n.$$

The usual procedure is to split up the product at the left-hand side in separate factors so to obtain a system of inequalities of type (3.6). But the number of ways to split up the product may depend on F and this results in an upper bound for the number of subspaces depending on F. Instead, we consider for each finite extension F of K the inequality

$$\prod_{v\in S}\prod_{i=1}^{n}\Big(\max_{w|v}\frac{\|L_i^{(v)}(\mathbf{x})\|_w}{\|\mathbf{x}\|_w}\Big)^{1/d(w|v)}\leq H(\mathbf{x})^{-n-\delta}\quad\text{in } \mathbf{x}\in F^n \tag{3.7}$$

where the maximum is taken over all places $w\in M(F)$ lying above v. The exponents $1/d(w|v)$ are needed to normalise the absolute values with respect to K. Note that for $F=K$ we get (2.11). By a combinatorial argument going back to Mahler which we do not work out, we can show that every solution $\mathbf{x}\in F^n$ of (3.7) has a scalar multiple which satisfies (3.6) for some possibly other number field F' and some tuple $\mathbf{c}$ having (3.4) with $\delta/2$ instead of δ. Here $\mathbf{c}$ belongs to a finite set independent of $\mathbf{x}$ of cardinality depending on n, δ and s, where s denotes the cardinality of S. This introduces a dependence on s in the upper bound for the number of subspaces. Using $R\leq s$ we get rid of the parameter R. Thus, we obtain the following "quantitative Absolute Subspace Theorem:"

Corollary 5. *There are proper linear subspaces $T_1,\dots,T_a$ of $\overline{\mathbf{Q}}^n$, all defined over K, with*

$$a\leq 8^{(n+6)^2}(50/\delta)^{ns+n+3},$$

such that for every finite extension F of K, the set of solutions of (3.7) *with*

$$H(\mathbf{x})\geq\max\big(H,\ (n^{n/2}\Delta^{-1})^{4/\delta}\big)$$

is contained in $T_1\cup\ldots\cup T_a$.

It should be noted that the exceptional set of solutions with $H(\mathbf{x})$ less than $\max\big(H,\ (n^{n/2}\Delta^{-1})^{2/\delta}\big)$ need not be contained even in the union of finitely many proper linear subspaces of $\overline{\mathbf{Q}}^n$. For instance, if $L_i^{(v)}=X_i$ for $v\in S$, $i=1,\dots,n$, then for every vector $\mathbf{x}$ consisting of roots of unity there is a finite extension F of K such that $\mathbf{x}$ satisfies (3.7).

The last consequence of Theorem 1 we mention is an absolute analogue of Schmidt's Theorem A. Denote by $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ the Galois group of $\overline{\mathbf{Q}}/\mathbf{Q}$. For $\sigma\in\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, $\mathbf{x}=(x_1,\dots,x_n)\in\overline{\mathbf{Q}}^n$, define $\sigma(\mathbf{x})=(\sigma(x_1),\dots,\sigma(x_n))$. Let $L_i=\alpha_{i1}X_1+\ldots+\alpha_{in}X_n$ $(i=1,\dots,n)$ be linearly independent linear forms with coefficients in $\overline{\mathbf{Q}}$ such that

$$H(L_i)\leq H,\qquad [\mathbf{Q}(L_i):\mathbf{Q}]\leq D,\qquad |L_i|:=\max_{1\leq j\leq n}|\alpha_{ij}|=1\qquad\text{for } i=1,\dots,n.$$

Consider the inequality

$$\prod_{i=1}^{n}\Big(\max_{\sigma\in\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})}\frac{|L_i(\sigma(\mathbf{x}))|}{|\sigma(\mathbf{x})|}\Big)\leq H(\mathbf{x})^{-n-\delta}\quad\text{in } \mathbf{x}\in\overline{\mathbf{Q}}^n \tag{3.8}$$

with $0<\delta\leq 1$, where $|\sigma(\mathbf{x})|$ denotes the maximum norm of $\sigma(\mathbf{x})$.

Corollary 6. *The set of solutions of* (3.8) *with*

$$H(\mathbf{x}) \geq \max\big(H,\ (n^{n/2} \cdot |\det(L_1, \dots, L_n)|^{-1})^{4/\delta}\big)$$

is contained in the union of finitely many proper linear subspaces $T_1, \dots, T_a$ *of* $\overline{\mathbf{Q}}^n$, *all defined over* $\mathbf{Q}$, *with*

$$a \leq 16^{(n+6)^2} \cdot \delta^{-2n-3} \log 4D \cdot \log\log 4D.$$

We return to Schmidt's result on the number of small solutions of (2.8). For $x \in \overline{\mathbf{Q}}^*$ define the logarithmic (absolute Weil-) height by

$$h(x) = \sum_{v \in M(K)} \log \max(1, \|x\|_v),$$

where K is any number field containing x. For $\mathbf{x} = (x_1, \dots, x_n) \in (\overline{\mathbf{Q}}^*)^n$ define the logarithmic norm

$$h_s(\mathbf{x}) = \sum_{i=1}^{n} h(x_i).$$

For $\mathbf{x}, \mathbf{y} \in (\overline{\mathbf{Q}}^*)^n$ let $\mathbf{x} * \mathbf{y}$ denote the coordinatewise product of $\mathbf{x}, \mathbf{y}$ and $\mathbf{x}^m$ the coordinatewise m-th power of $\mathbf{x}$ for $m \in \mathbf{Z}$. Then h_s satisfies the norm axioms $h_s(\mathbf{x}) \geq 0$ and $h_s(\mathbf{x}) = 0$ if and only if $\mathbf{x}$ is torsion, i.e. consists of roots of unity; $h_s(\mathbf{x}^m) = |m| h_s(\mathbf{x})$ for $m \in \mathbf{Z}$; $h_s(\mathbf{x} * \mathbf{y}) \leq h_s(\mathbf{x}) + h_s(\mathbf{y})$. The following is a special case of [44], Thm. 5:

Theorem I (Schmidt [44]). *Let G be a finite type subgroup of $\overline{\mathbf{Q}}^*$ of rank r and $C \geq 0$. Put $q = \exp((4n)^{2n})$. Then* (2.8) *has at most*

$$q(qC)^r$$

solutions $\mathbf{x} = (x_1, \dots, x_n) \in G^n$ with $h_s(\mathbf{x}) \leq C$.

Results such as Theorem I heavily rely on good explicit lower bounds for the logarithmic norms of algebraic points lying on algebraic varieties. The research on such lower bounds was started by Zhang [49], who by means of Arakelov theory proved a general result about the logarithmic norms of algebraic points on curves, a special case of which is as follows: there is an absolute constant $C > 0$ such that every algebraic point $\mathbf{x} = (x, y)$ for which $x + y = 1$ and x, y are not both equal to 0 or a root of unity has $h_s(\mathbf{x}) \geq C$. After that, by an elementary method, Zagier [48] showed that every such point $\mathbf{x} = (x, y)$ satisfies $h_s(\mathbf{x}) \geq \frac{1}{2} \log\{\frac{1}{2}(1+\sqrt{5})\}$. Zagier's result was further extended by Schlickewei and Wirsing [40]. Schlickewei derived from their result the estimate for the number of "small" solutions of $ax + by = 1$ in $x, y \in G$ that he needed in the proof of Theorem F. The results of Zhang, Zagier, and Schlickewei and Wirsing were further improved and generalised by Beukers and Zagier [2], Schmidt [43], again Zhang [50], Bombieri and Zannier [5] and again Schmidt [44]. Theorem 3 of the last paper gives in the most general situation an explicit lower bound for the logarithmic norm of an algebraic point on an algebraic variety. The following result is the special case of this needed in the proof of Theorem I. For a positive integer n and a positive real h, let $A(n,h)$ denote the smallest integer A such that $\{1,\dots,A\}$ contains an arithmetic progression of length n all of whose terms are composed of primes $> h$.

Theorem J (Schmidt [44]). *Let* $\mathbf{x}=(x_1,\dots,x_n)\in(\overline{\mathbf{Q}}^*)^n$ *be such that*

$$x_1+\ldots+x_n=1,\qquad \sum_{i\in I}x_i\neq 0\quad \textit{for each subset } I \textit{ of } \{1,\dots,n\},$$

$x_1,\dots,x_n$ *are not all roots of unity.*

Then $h_s(\mathbf{x})\geq 1/A(n,5n)$.

Schmidt [44] proved that

$$A(n,h)<n\cdot e^{1.017h},\quad A(n,h)<(c_1h)^{c_2n}, \tag{3.9}$$

respectively. Using a result of Schinzel [29] (Lemma 1) one gets the second estimate in the following explicit form $A(n,h)<(2h)^{20n}$.

We mention that Beukers and Zagier [2] (Cor. 2.1) obtained the much better lower bound $h_s(\mathbf{x})\geq\frac{1}{2}\log\{\frac{1}{2}(1+\sqrt{5})\}$ but only subject to the restriction $x_1^{-1}+\ldots+x_n^{-1}\neq 1$ which makes their result not applicable for our purposes.

We give a rough idea how Theorem J is applied to obtain Theorem I. Define the logarithmic distance of $\mathbf{x}$, $\mathbf{y}\in(\overline{\mathbf{Q}}^*)^n$ by $\delta(\mathbf{x},\mathbf{y})=h_s(\mathbf{x}*\mathbf{y}^{-1})$. Let $\mathcal{S}$ be the set of solutions $\mathbf{x}$ of (2.8) with $h_s(\mathbf{x})\leq C$. We select from $\mathcal{S}$ a maximal subset, such that any two points in this subset have logarithmic distance $\geq\varepsilon$, say, where $\varepsilon>0$ is a real that has to be chosen optimally. Using that G has rank r one shows by an elementary argument that this subset has cardinality at most

$$(1+(2C/\varepsilon))^{nr}. \tag{3.10}$$

(cf. [43], Lemma 4). So it remains to estimate from above the number of solutions lying in a "ball"

$$B(\mathbf{y},\varepsilon)=\{\mathbf{x}\in G^n:\ \delta(\mathbf{x},\mathbf{y})<\varepsilon\}$$

where $\mathbf{y}=(y_1,\dots,y_n)$ is a fixed solution of (2.8). By replacing the coefficients of (2.8) by $a_i'=a_iy_i$ for $i=1,\dots,n$, we see that it suffices to estimate from above the number of solutions of

$$a_1'x_1+\ldots+a_n'x_n=1\quad\text{in } \mathbf{x}=(x_1,\dots,x_n)\in G^n \text{ with } h_s(\mathbf{x})<\varepsilon. \tag{3.11}$$

Note that $(1,\dots,1)$ is a solution of (3.11). Take n other solutions $\mathbf{x}_i=(x_{i1},\dots,x_{in})$ of (3.11). Then we obtain the determinant equation

$$\begin{vmatrix}1 & 1 & \cdots & 1\\ 1 & x_{11} & \cdots & x_{1n}\\ \vdots & \vdots & & \vdots\\ 1 & x_{n1} & \cdots & x_{nn}\end{vmatrix}=0.$$

Put $x_{0i} = 1$, $x_{i0} = 1$ for $i = 0, \dots, n$. The determinant is an alternating sum of $(n+1)!$ terms $x_\sigma = x_{0,\sigma(0)} \cdots x_{n,\sigma(n)}$, where σ runs through the permutations of $\{0, \dots, n\}$. By taking a minimal vanishing subsum and dividing by one term we get

$$\sum_{\sigma \in I} \pm \frac{x_\sigma}{x_{\sigma_0}} = 1, \tag{3.12}$$

where I is some set of permutations of $\{0, \dots, n\}$, σ_0 is a fixed permutation, and the left-hand side has no vanishing subsums. Now one can show that if (3.11) has many solutions, there are n solutions $\mathbf{x}_1, \dots, \mathbf{x}_n$ among these for which at least one of the terms in the left-hand side of (3.12) is not a root of unity. So we can apply Theorem J to (3.12). On noting that I has cardinality smaller than $(n+1)!$ we obtain for the vector $\mathbf{X} = (x_\sigma / x_{\sigma_0} : \sigma \in I)$

$$h_s(\mathbf{X}) \geq B^{-1} \quad \text{with } B = A((n+1)!, 5(n+1)!).$$

On the other hand, by taking ε sufficiently small, one can show that $h_s(\mathbf{X}) < B^{-1}$ and this gives a contradiction. $\square$

Both estimates in (3.9) imply an upper bound for B which is doubly exponential in n and this results in an upper bound for ε^{-1} doubly exponential in n. Consequently, already the quantity (3.10) gives a contribution to the upper bound in Theorem I doubly exponential in n.

References

[1] Beukers, F., Schlickewei, H.P., The equation $x + y = 1$ in finitely generated groups. Acta Arith. 78 (1996), 189–199.

[2] Beukers, F., Zagier, D., Lower bounds for heights of points on hypersurfaces. Acta Arith. 79 (1997), 103–111.

[3] Bombieri, E., Mueller, J., Poe, M., The unit equation and the cluster principle. Acta Arith. 79 (1997), 361–389.

[4] Bombieri, E., Vaaler, J., On Siegel's lemma. Invent. Math. 73 (1983), 11–32.

[5] Bombieri, E., Zannier, U., Algebraic points on subvarieties of $\mathbf{G}_m^n$. Internat. Math. Res. Notices 7 (1995), 333–347.

[6] Conway, J.H., Jones, A.J., Trigonometric Diophantine equations (On vanishing sums of roots of unity). Acta Arith. 30 (1976), 229–240.

[7] Evertse, J.-H., On equations in S-units and the Thue-Mahler equation. Invent. Math. 75 (1984), 561–584.

[8] — On sums of S-units and linear recurrences. Compositio Math. 53 (1984), 225–244.

[9] — The number of solutions of decomposable form equations. Invent. Math. 122 (1995), 559–601.

[10] — An explicit version of Faltings' Product Theorem and an improvement of Roth's lemma. Acta Arith. 73 (1995), 215–248.

[11] Evertse, J.-H., An improvement of the quantitative Subspace Theorem. Compositio Math. 101 (1996), 225–311.

[12] — The number of solutions of linear equations in roots of unity. Preprint.

[13] Evertse, J.-H., Győry, K., Stewart, C.L., Tijdeman, R., S-unit equations and their applications. In: New advances in transcendence theory (ed. by A. Baker), 110–174. Cambridge Univ. Press, Cambridge 1988.

[14] Evertse, J.-H., Schlickewei, H.P., A quantitative version of the absolute Subspace Theorem. To appear.

[15] Evertse, J.-H., Schlickewei, H.P., Schmidt, W.M., Linear equations with variables which lie in a multiplicative group. To appear.

[16] Faltings, G., Diophantine approximation on abelian varieties. Ann. of Math. (2) 133 (1991), 549–576.

[17] Ferretti, R., An effective version of Faltings' Product Theorem. Forum Math. 8 (1996), 401–427.

[18] Győry, K., Some recent applications of S-unit equations. Astérisque 209 (1992), 17–38.

[19] Laurent, M., Equations Diophantiennes exponentielles. Invent. Math. 78 (1984), 299–327.

[20] Lech, C., A note on recurring series. Ark. Math. 2 (1953), 417–421.

[21] Mann, H.B., On linear equations between roots of unity. Mathematika 12 (1965), 107–117.

[22] McFeat, R.B., Geometry of numbers in adèle spaces. Dissertationes Math. (Rozprawy Mat.) 88 (1971).

[23] Poe, M., On distribution of solutions of S-unit equations. J. Number Theory 62 (1997), 221–241.

[24] Poorten, A.J. van der, Schlickewei, H.P., Additive relations in fields. J. Austral. Math. Soc. Ser. A 51 (1991), 154–170.

[25] Put, M. van der, The Product theorem. In: Diophantine approximation and abelian varieties (ed. by B. Edixhoven, J.-H. Evertse; Lecture Notes in Math. 1566), 77–82. Springer, Berlin 1995.

[26] Roth, K.F., Rational approximations to algebraic numbers. Mathematika 2 (1955), 1–20. Corrigendum: ibid., 168.

[27] Roy, D., Thunder, J.L., An absolute Siegel's Lemma. J. Reine Angew. Math. 476 (1996), 1–26.

[28] — A note about Siegel's Lemma on the algebraic closure of a global field. Preprint.

[29] Schinzel, A., On sums of roots of unity. Acta Arith. 11 (1966), 419–432.

[30] Schlickewei, H.P., The $\wp$-adic Thue-Siegel-Roth-Schmidt theorem. Arch. Math. (Basel) 29 (1977), 267–270.

[31] — S-unit equations over number fields. Invent. Math. 102 (1990), 95–107.

[32] — The quantitative Subspace Theorem for number fields. Compositio Math. 82 (1992), 245–274.

[33] Schlickewei, H.P., Equations in roots of unity. Acta Arith. 76 (1996), 99–108.

[34] — Multiplicities of recurrence sequences. Acta Math. 176 (1996), 171–243.

[35] — The multiplicity of binary recurrences. Invent. Math. 129 (1997), 11–36.

[36] — Linear equations over finitely generated groups. Ann. of Math. (2), to appear.

[37] — Equations $ax + by = 1$. Ann. of Math. (2), to appear.

[38] — A parametric version of the Subspace Theorem. Preprint.

[39] Schlickewei, H.P., Schmidt, W.M., Linear equations with variables which lie in a multiplicative group. Preprint.

[40] Schlickewei, H.P., Wirsing, E., Lower bounds for the heights of solutions of linear equations. Invent. Math. 129 (1997), 1–10.

[41] Schmidt, W.M., Norm form equations. Ann. of Math. (2) 96 (1972), 526–551.

[42] — The Subspace Theorem in Diophantine approximations. Compositio Math. 69 (1989), 121–173.

[43] — Heights of algebraic points lying on curves or hypersurfaces. Proc. Amer. Math. Soc. 124 (1996), 3003–3013.

[44] — Heights of points on subvarieties of $\mathbf{G}_m^n$. In: Number theory. Papers from the Séminaire de Théorie des Nombres de Paris, 1993–94 (ed. by S. David; London Math. Soc. Lecture Note Ser. 235), 157–187. Cambridge Univ. Press, Cambridge 1996.

[45] — The zero multiplicity of linear recurrence sequences. To appear.

[46] Silverman, J.H., Lower bounds for height functions. Duke Math. J. 51 (1984), 395–403.

[47] — Exceptional units and numbers of small Mahler measure. Experiment. Math. 4 (1995), 70–83.

[48] Zagier, D., Algebraic numbers close to both 0 and 1. Math. Comp. 61 (1993), 485–491.

[49] Zhang, S., Positive line bundles on arithmetic surfaces. Ann. of Math. (2) 136 (1992), 569–587.

[50] — Positive line bundles on arithmetic varieties. J. Amer. Math. Soc. 8 (1995), 187–221.

Jan-Hendrik Evertse
Mathematical Institute
University of Leiden
P.O.Box 9512, 2300 RA Leiden
The Netherlands
evertse@wi.leidenuniv.nl

Hans Peter Schlickewei
Fachbereich Mathematik
Universität Marburg
Hans-Meerwein-Straße, Lahnberge
35032 Marburg/Lahn, Germany
hps@mathematik.uni-marburg.de