ON THE QUANTITATIVE SUBSPACE THEOREM

JAN-HENDRIK EVERTSE

Abstract. In this survey we give an overview of recent improvements upon the Quantitative Subspace Theorem, obtained jointly with R. Ferretti, which follow from work in [9]. Further, we give a new gap principle with which we can estimate the number of subspaces containing the "small solutions" of the systems of inequalities being considered. As an introduction, we start with a quantitative version of Roth's Theorem.

1. A quantitative Roth's Theorem

Recall that the (absolute) height of an algebraic number ξ of degree d is given by

$$H(\xi) := \Big(a \cdot \prod_{i=1}^{d} \max(1, |\xi^{(i)}|)\Big)^{1/d},$$

where $\xi^{(1)}, \dots, \xi^{(d)}$ are the conjugates of ξ in $\mathbb{C}$ and where a is the positive integer such that the polynomial $a \cdot \prod_{i=1}^{d}(X - \xi^{(i)})$ has rational integral coefficients with gcd 1. In particular, if $\xi \in \mathbb{Q}$, then $H(\xi) = \max(|x|, |y|)$, where x, y are coprime integers such that $\xi = x/y$.

Roth's celebrated theorem from 1955 (see [17]) states that if ξ is any real algebraic number and δ any real with $\delta > 0$, then the inequality

$$(1.1) \qquad |\xi - \alpha| \leqslant H(\alpha)^{-2-\delta} \quad \text{in } \alpha \in \mathbb{Q}$$

has only finitely many solutions. Already in 1955, Davenport and Roth [3] computed an upper bound for the number of solutions of (1.1), and their bound was subsequently improved by Mignotte [16], Bombieri and van der Poorten [1], and the author [7]. We formulate a slight improvement of the

Date: June 18, 2010.

2000 Mathematics Subject Classification: 11J68, 11J25.

Keywords and Phrases: Diophantine approximation, Subspace Theorem.

latter result which follows from the Appendix of [2]. We mention that this improvement is obtained by simply going through the existing methods; its proof did not involve anything new. We distinguish between *large* and *small* solutions α of (1.1), where a rational number α is called large if

$$H(\alpha) \geqslant \max\bigl(H(\xi), 2\bigr) \tag{1.2}$$

and small otherwise.

Theorem 1.1. *Let ξ be a real algebraic number of degree d and $0 < \delta \leqslant 1$. Then the number of large solutions of* (1.1) *is at most*

$$2^{25}\delta^{-3}\log(2d)\log\bigl(\delta^{-1}\log(2d)\bigr)$$

and the number of small solutions at most

$$10\delta^{-1}\log\log\max\bigl(H(\xi), 4\bigr).$$

The proof of this result can be divided into two parts: a so-called *interval result* and a *gap principle.* The interval result may be stated as follows.

Proposition 1.2. *Let*

$$m := 1 + [25600\delta^{-2}\log(2d)], \quad \omega := 162m^2\delta^{-1},$$
$$C := \exp\left(3m\binom{d}{2}\delta^{-1}\bigl(240m^2\delta^{-1}\bigr)^m \cdot \log\bigl(36H(\xi)\bigr)\right).$$

Then there are reals $Q_1, \dots, Q_{m-1}$ with

$$C \leqslant Q_1 < Q_2 < \cdots < Q_{m-1}$$

such that if $\alpha \in \mathbb{Q}$ is a solution of (1.1) *with $H(\alpha) \geqslant C$, then*

$$H(\alpha) \in \bigcup_{i=1}^{m-1} \bigl[Q_i, Q_i^{\omega}\bigr).$$

The proof is by means of the usual "Roth machinery." Assume Theorem 1.2 is false. Then (1.1) has solutions $\alpha_1, \dots, \alpha_m$ such that $H(\alpha_1) \geqslant C$ and $H(\alpha_i) \geqslant H(\alpha_{i-1})^{\omega}$ for $i = 1, \dots, m$. One constructs a polynomial $F(X_1, \dots, X_m)$ which has integer coefficients of small absolute value, and which is of degree d_i in the variable X_i for $i = 1, \dots, m$, such that $H(\alpha_1)^{d_1} \approx \cdots \approx H(\alpha_m)^{d_m}$, and such that F has large "index" (some sort of weighted multiplicity) at the point $(\alpha_1, \dots, \alpha_m)$. Then one applies Roth's Lemma (a non-vanishing result for polynomials) to conclude that F cannot have large index at $(\alpha_1, \dots, \alpha_m)$. In fact, we use a refinement of Roth's original Lemma from 1955 (see [5]) which was proved by means of the techniques going into the proof of Faltings' Product Theorem [14].

The second ingredient is the following very basic gap principle.

Proposition 1.3. *Let $Q \geqslant 2$. Then* (1.1) *has at most one solution α such that $Q \leqslant H(\alpha) < Q^{1+\delta/2}$ and $\alpha > \xi$, and also at most one solution α such that $Q \leqslant H(\alpha) < Q^{1+\delta/2}$ and $\alpha < \xi$.*

Proof. Suppose for instance that (1.1) has two solutions α_1, α_2 which are both larger than ξ, and $Q \leqslant H(\alpha_1) \leqslant H(\alpha_2) < Q^{1+\delta/2}$ for $i = 1, 2$. Then

$$\begin{aligned} Q^{-2(1+\delta/2)} &< \big(H(\alpha_1)H(\alpha_2)\big)^{-1} \leqslant |\alpha_1 - \alpha_2| \\ &\leqslant \max_i |\xi - \alpha_i| \leqslant H(\alpha_1)^{-2-\delta} \leqslant Q^{-2-\delta} \end{aligned}$$

which is obviously impossible. □

An immediate consequence of this gap principle is that for any $Q \geqslant 2$, $E > 1$, inequality (1.1) has at most $1 + 2\log E/\log(1+\delta/2)$ solutions $\alpha \in \mathbb{Q}$ with $Q \leqslant H(\alpha) < Q^E$. Using this fact in combination with Proposition 1.2, the deduction of Theorem 1.1 is straightforward.

Also in more advanced situations, the general pattern to obtain explicit upper bounds for the number of solutions of certain Diophantine equations or inequalities, is first to prove that the number of solutions is finite by means of an involved Diophantine approximation method, and second to estimate from above the number of solutions using a more or less elementary gap principle. However, there are also many situations where we do have at our disposal a method to prove finiteness for the number of solutions but where we do not have a gap principle. So in these situations we know that there are only finitely many solutions, but we are not able to estimate their number.

2. The quantitative Subspace Theorem

We generalize the results from Section 1 to higher dimensions.

Let $n \geqslant 2$ be an integer. We denote by $\|\cdot\|$ the maximum norm on $\mathbb{R}^n$. Let

$$L_i = \alpha_{i1}X_1 + \cdots + \alpha_{in}X_n \quad (i = 1, \ldots, n)$$

be linearly forms with algebraic coefficients $\alpha_{ij} \in \mathbb{C}$ which are linearly independent, that is, their coefficient determinant $\det(L_1, \ldots, L_n) = \det(\alpha_{ij})$ is non-zero. Further, let $\delta > 0$ and consider the inequality

$$|L_1(\mathbf{x}) \cdots L_n(\mathbf{x})| \leqslant \|\mathbf{x}\|^{-\delta} \quad \text{in } \mathbf{x} \in \mathbb{Z}^n. \tag{2.1}$$

W. Schmidt's celebrated Subspace Theorem from 1972 (see [21]) states that the set of solutions of (2.1) lies in a union of finitely many proper linear subspaces of $\mathbb{Q}^n$. In 1989, Schmidt proved [23] a quantitative result, which in a slightly modified form reads as follows.

Suppose that the algebraic numbers α_{ij} have height at most H and degree at most D and that $0 < \delta \leqslant 1$. Then the solutions of

$$|L_1(\mathbf{x}) \cdots L_n(\mathbf{x})| \leqslant |\det(L_1, \ldots, L_n)| \cdot \|\mathbf{x}\|^{-\delta} \quad \textit{in } \mathbf{x} \in \mathbb{Z}^n$$

with $\|\mathbf{x}\| \geqslant \max(2H, n^{2n/\delta})$ lie in a union of at most $2^{2^{27n}\delta^{-2}}$ proper linear subspaces of $\mathbb{Q}^n$.

This quantitative result has been improved and generalized in various directions, mainly due to work of Schlickewei and the author.

We now discuss versions of the Subspace Theorem which involve non-archimedean absolute values and which take their unknowns from algebraic number fields. All our algebraic number fields considered below are contained in a given algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$.

Let $M_{\mathbb{Q}} := \{\infty\} \cup \{\text{primes}\}$ denote the set of places of $\mathbb{Q}$. We write $|\cdot|_\infty$ for the ordinary absolute value on $\mathbb{Q}$ and $|\cdot|_p$ (p prime number) for the p-adic absolute value, normalized such that $|p|_p = p^{-1}$. Further, we denote by $\mathbb{Q}_p$ the completion of $\mathbb{Q}$ at p; in particular, $\mathbb{Q}_\infty = \mathbb{R}$.

Let K be an algebraic number field and denote by M_K the set of places of K. To every place $v \in M_K$, we associate an absolute value $|\cdot|_v$ which is such that if v lies above $p \in M_{\mathbb{Q}}$, then the restriction of $|\cdot|_v$ to $\mathbb{Q}$ is $|\cdot|_p^{[K_v:\mathbb{Q}_p]/[K:\mathbb{Q}]}$, where K_v is the completion of K at v. The absolute value $|\cdot|_v$ can be continued uniquely to the algebraic closure $\overline{K_v}$ of K_v. The place v is called finite if $v \nmid \infty$, infinite if $v | \infty$, real if $K_v = \mathbb{R}$ and complex if $K_v = \mathbb{C}$. The absolute values thus chosen satisfy the product formula $\prod_{v \in M_K} |x|_v = 1$ for $x \in K^*$.

We define the height (not the standard definition) of $\mathbf{x} = (x_1, \dots, x_n) \in K^n$ by

$$H(\mathbf{x}) := \prod_{v \in M_K} \max(1, |x_1|_v, \dots, |x_n|_v).$$

Let S be a finite subset of M_K, containing all infinite places. Denote by $O_S = \{x \in K : |x|_v \leqslant 1 \text{ for } v \in M_K \setminus S\}$ the ring of S-integers. For $v \in S$, let

$$L_i^{(v)} = \alpha_{i1}^{(v)} X_1 + \dots + \alpha_{in}^{(v)} X_n \quad (i = 1, \dots, n)$$

be linearly independent linear forms with coefficients $\alpha_{ij}^{(v)} \in \overline{K_v}$ that are algebraic over K.

In 1977, Schlickewei [18] proved that the set of solutions of the inequality

$$\prod_{v \in S} |L_1^{(v)}(\mathbf{x}) \cdots L_n^{(v)}(\mathbf{x})|_v \leqslant H(\mathbf{x})^{-\delta} \quad \text{in } \mathbf{x} \in O_S^n \tag{2.2}$$

is contained in a union of finitely many proper linear subspaces of K^n.

By an elementary combinatorial argument (see for instance [12, Section 21]), one can show that every solution $\mathbf{x}$ of (2.2) satisfies one of a finite number of systems of inequalities

$$|L_i^{(v)}(\mathbf{x})|_v \leqslant C_v H(\mathbf{x})^{c_{iv}} \quad (v \in S,\ i = 1, \dots, n) \quad \text{in } \mathbf{x} \in O_S^n \tag{2.3}$$

where $C_v > 0$ for $v \in S$ and $\sum_{v \in S} \sum_{i=1}^n c_{iv} < 0$. Thus, an equivalent version of Schlickewei's extension of the Subspace Theorem is the following result which we state for reference purposes:

Theorem A. *Suppose $C_v > 0$ for $v \in S$ and $\sum_{v \in S} \sum_{i=1}^n c_{iv} < 0$. Then the solutions of* (2.3) *lie in finitely many proper linear subspaces of K^n.*

Put

$$s(v) := 1/[K : \mathbb{Q}] \ \text{ if } v \text{ is real}, \ s(v) := 2/[K : \mathbb{Q}] \ \text{ if } v \text{ is complex},$$
$$s(v) := 0 \ \text{ if } v \text{ is finite}.$$

The following technical conditions on the linear forms $L_i^{(v)}$, the constants C_v and the exponents c_{iv} will be kept throughout:

$$(2.4)\qquad \begin{cases} H(\alpha_{ij}^{(v)}) \leqslant H,\ [K(\alpha_{ij}^{(v)}) : K] \leqslant D \text{ for } v \in S,\ i,j = 1,\dots,n; \\ \# \bigcup_{v\in S} \{L_1^{(v)},\dots,L_n^{(v)}\} \leqslant R; \\ 0 < \prod_{v\in S} C_v \leqslant \prod_{v\in S} |\det(L_1^{(v)},\dots,L_n^{(v)})|_v^{1/n}; \\ \sum_{v\in S}\sum_{i=1}^{n} c_{iv} \leq -\delta \text{ with } 0 < \delta \leqslant 1; \\ \max(c_{1v},\dots,c_{nv}) = s(v) \text{ for } v \in S. \end{cases}$$

The following result is an easy consequence of a general result of Schlickewei and the author [12, Theorem 2.1]:

Theorem B. *Assume* (2.4)*. Then the set of solutions* $\mathbf{x} \in O_S^n$ *of* (2.3) *with*

$$H(\mathbf{x}) \geqslant \max(2H, n^{2n/\delta})$$

is contained in a union of at most

$$4^{(n+9)^2}\delta^{-n-4}\log(2RD)\log\log(2RD)$$

proper linear subspaces of K^n*.*

In fact, Schlickewei and the author proved a more general "absolute" version where the unknowns may be algebraic numbers not necessarily belonging to a fixed number field.

For applications it is important that the upper bound for the number of subspaces is independent of the field K. The quantity R may be replaced by ns, where s is the cardinality of S. But in many cases, R can be taken independently of s. For instance in applications to linear equations with unknowns from a finitely generated multiplicative group and to linear recurrence sequences (see [20], [13], [26]) one has to apply the above Theorem with $L_i^{(v)} \in \{X_1,\dots,X_n, X_1+\cdots+X_n\}$ for $v \in S$, $i = 1,\dots,n$, and in that case, one may take $R = n+1$.

Theorem B was the outcome of a development resulting from Schmidt's quantitative version of the Subspace Theorem mentioned above and subsequent improvements and generalizations by Schlickewei and the author [19], [20], [7], [12].

The proof of Theorem B is basically a quantification of Schmidt's method of proof of his Subspace Theorem from 1972 (see [21], [22]). It consists of geometry of numbers, a construction of an auxiliary polynomial, and an application of Roth's Lemma. In 1994, Faltings and Wüstholz [15] gave a totally new proof of the Subspace Theorem. In their proof they did not use geometry of numbers, and instead of Roth's Lemma they applied the much more powerful Faltings' Product Theorem. Another important ingredient of the proof of Faltings and Wüstholz is a stability theory for multi-filtered vector spaces. The method of Faltings and Wüstholz also allows to compute an upper bound for the number of subspaces containing the solutions of (2.3), but this is much larger than the one from Theorem B. In fact, in the proof of Faltings and Wüstholz one has to construct global line bundle sections on products of algebraic varieties of very large degrees (as opposed to Schmidt's proof where one encounters only linear varieties) and this leads to poor estimates for the number of subspaces.

However, the upper bound from Theorem B can be improved further if one combines ideas from Schmidt's method of proof with ideas from Faltings and Wüstholz. Essentially, one may follow Schmidt's method of proof, but replace Schmidt's construction of an auxiliary polynomial by that of Faltings and Wüstholz, see Section 6 for more details.

In this way, Ferretti and the author [9] obtained the following. A solution $\mathbf{x}$ of (2.3) is called *large* if

$$H(\mathbf{x}) \geqslant \max(H, n^{2n/\delta})$$

and *small* otherwise.

Theorem 2.1. *Assume* (2.4). *Then the set of large solutions of* (2.3) *lies in a union of at most*

$$10^9 2^{2n} n^{14} \delta^{-3} \log(3\delta^{-1}RD) \cdot \log(\delta^{-1} \log 3RD)$$

proper linear subspaces of K^n*.*

So compared with Theorem B, the dependence on n has been brought down from c^{n^2} to c^n, while the dependence on δ has been improved from δ^{-n-4} to $\delta^{-3}(\log\delta^{-1})^2$. With this improvement, the dependence on δ is almost as good as that in the quantitative Roth's Theorem from the previous section. One might still hope for a further improvement in terms of n, for instance to something polynomial in n, but probably this would require a new method of proof for the Subspace Theorem.

For the small solutions we have the following elementary result which is proved in Section 4 of the present paper. Here, in contrast to the large solutions, we do get a dependence on the field K.

Theorem 2.2. *Assume* (2.4). *Let* $d := [K : \mathbb{Q}]$. *Then the set of small solutions of* (2.3) *lies in a union of at most*

$$\delta^{-1}\Big((10^3 n)^{nd} + 4n\log\log 4H\Big)$$

proper linear subspaces of K^n.

In the case $K = \mathbb{Q}$ *this bound can be replaced by*

$$\delta^{-1}\Big(10^{3n} + 4n\log\log 4H\Big).$$

It is an open problem whether the bounds in Theorem 2.2 can be replaced by something depending only polynomially on n and/or d. Recent work by Schmidt [27] on Roth's Theorem over number fields suggests that a polynomial dependence on d should be possible.

3. A refinement of the Subspace Theorem and an interval result

We keep the notation and assumptions from the previous section. So K, S, $L_i^{(v)}$ ($v \in S$, $i = 1, \ldots, n$), δ, have the same meaning as before, and they satisfy (2.4). The following refinement of the Subspace Theorem follows from work of Faltings and Wüstholz [15] and Vojta [28] but there is a heavy overlap with ideas of Schmidt [25].

Theorem C. *There is a proper linear subspace* U_0 *of* K^n, *such that* (2.3) *has only finitely many solutions outside* U_0.

This space U_0 *can be determined effectively. Moreover, it can be chosen* *from a finite collection, which depends only on the linear forms* $L_i^{(v)}$ *(*$v \in S$*,* $i = 1, \ldots, n$*) and is independent of the constants* C_v *and the exponents* c_{iv}*.*

The first part giving the mere existence of U_0 is Theorem 9.1 of [15]. The second part follows from [28].

We first give a description of the space U_0 occurring in Theorem 9.1 of [15], where we have translated Faltings' and Wüstholz' terminology into ours. Let $v \in M_K$. Two linear forms $L = \sum_{i=1}^n \alpha_i X_i$ and $M = \sum_{i=1}^n \beta_i X_i$ with coefficients in $\overline{K_v}$ are said to be conjugate over K_v if there is an automorphism σ of $\overline{K_v}$ over K_v such that $\sigma(\alpha_i) = \beta_i$ for $i = 1, \ldots, n$. Given $v \in M_K$ and a system of linear forms $L_1, \ldots, L_r$ with coefficients in $\overline{K_v}$, this system is called v-symmetric if with any linear form in the system, also all its conjugates over K_v belong to this system.

Given a linear subspace U of K^n and linear forms $L_1, \ldots, L_r$ with coefficients generating a field extension F of K, we say that $L_1, \ldots, L_r$ are linearly independent on U if there is no non-trivial linear combination of $L_1, \ldots, L_r$ with coefficients in F that vanishes identically on U.

For each $v \in S$, we obtain a v-symmetric system $L_1^{(v)}, \ldots, L_{n_v}^{(v)}$, consisting of the linear forms $L_1^{(v)}, \ldots, L_n^{(v)}$ from (2.3) and their conjugates over K_v. Using $|L(\mathbf{x})|_v = |M(\mathbf{x})|_v$ for any $\mathbf{x} \in K^n$ and any linear forms L, M with coefficients in $\overline{K_v}$ which are conjugate over K_v, we see that (2.3) is equivalent to the system of inequalities

$$|L_i^{(v)}(\mathbf{x})|_v \leqslant C_v H(\mathbf{x})^{c_{iv}} \ (v \in S,\ i = 1, \ldots, n_v) \ \text{ in } \mathbf{x} \in O_S^n. \tag{3.1}$$

Now for any linear subspace U of K^n and any $v \in S$, define $\nu_v(U) = 0$ if $U = (0)$ and

$$\nu_v(U) := \min c_{i_1,v} + \cdots + c_{i_u,v}$$

if $U \neq (0)$, where $u = \dim U$, and the minimum is taken over all subsets $\{i_1, \ldots, i_u\}$ of $\{1, \ldots, n_v\}$ of cardinality u such that $L_{i_1}^{(v)}, \ldots, L_{i_u}^{(v)}$ are linearly independent on U. Further, define

$$\nu(U) := \sum_{v \in S} \nu_v(U),$$

and, if $U \neq K^n$,

$$\mu(U) := \frac{\nu(K^n) - \nu(U)}{n - \dim U}.$$

Let μ_0 be the mimimum of the quantities $\nu(U)$, taken over all proper linear subspaces U of K^n.

Now one can show that there is a unique proper linear subspace U_0 of K^n, which is the one from Theorem C, such that

$$(3.2) \qquad \begin{cases} \mu(U_0) = \mu_0; \\ U_0 \subseteq U \text{ for every linear subspace } U \text{ of } K^n \text{ with } \mu(U) = \mu_0. \end{cases}$$

It is important to remark, that Theorem C can be deduced from the apparently weaker Theorem A. The argument is roughly as follows. First assume that $U_0 = (\mathbf{0})$. (In this case, following the terminology of Faltings and Wüstholz, system (2.3) is called *semistable*.) This assumption implies that if U is any linear subspace of K^n of dimension at least 2, then Theorem A is applicable to the restriction of (2.3) to U, and thus, the solutions of (2.3) in U lie in a finite union of proper linear subspaces of U. Now by induction, it follows easily that (2.3) has only finitely many solutions.

If $U_0 \neq (\mathbf{0})$, one may derive from (2.3) a semistable system of inequalities, with solutions from the quotient vector space K^n/U_0. We infer that the solutions of (2.3) outside U_0 lie in finitely many cosets modulo U_0. Then one completes the proof by showing that each coset contains only finitely many solutions.

The space U_0 can be determined effectively in principle using a combinatorial algorithm based on ideas of Vojta [28]. In fact, let $M_1, \dots, M_t$ be the conjugates in $\overline{\mathbb{Q}}[X_1, \dots, X_n]$ of the linear forms $L_i^{(v)}$ ($v \in S$, $i = 1, \dots, n$). Let F be the extension of K generated by the coefficients of $M_1, \dots, M_t$. Define the F-vector spaces $H_i := \{\mathbf{x} \in L^n : M_i(\mathbf{x}) = 0\}$ ($i = 1, \dots, t$). From ideas of Vojta [28] it follows that $U_0 \otimes_K F$ can be obtained by an algorithm taking as input the spaces $H_1, \dots, H_t$ and applying repeatedly the operations $+$ (sum of two vector spaces) and $\cap$ (intersection) to two previously obtained spaces. The number of steps of this algorithm is bounded above effectively in terms of t only.

Alternatively, from an auxiliary result in [9] it follows that U_0 has a basis, consisting of vectors of which the coordinates have heights at most $\big(\sqrt{n}H\big)^{4^n}$, where H is given by (2.4).

The special case that $L_i^{(v)} \in \{X_1, \ldots, X_n, X_1 + \cdots + X_n\}$ for $v \in S$, $i = 1, \ldots, n$ is of particular importance for applications. It is shown in [9] that in this case we have

$$U_0 = \left\{ \mathbf{x} = (x_1, \ldots, x_n) \in K^n : \sum_{j \in I_i} x_j = 0 \text{ for } i = 1, \ldots, t \right\}$$

where $I_1, \ldots, I_t$ are certain pairwise disjoint subsets of $\{1, \ldots, n\}$.

The solutions of (2.3) outside U_0 can not be determined effectively. Moreover, it is also beyond reach to estimate the number of solutions outside U_0. But Ferretti and the author [9] proved the following more precise version of Theorem C which may be considered as an analogue of the interval result Proposition 1.2.

Theorem 3.1. *Assume* (2.4). *Put*

$$m := \big[10^8 2^{2n} n^{14} \delta^{-2} \log(3\delta^{-1} RD)\big], \quad \omega := 3n\delta^{-1} \log 3RD.$$

Then there are reals $Q_1, \ldots, Q_m$ *with*

$$\max(2H, n^{2n/\delta}) \leqslant Q_1 < Q_2 < \cdots < Q_m$$

such that for every solution $\mathbf{x} \in O_S^n$ *of* (2.3) *outside* U_0 *we have*

$$H(\mathbf{x}) < \max(2H, n^{2n/\delta}) \text{ or } H(\mathbf{x}) \in \bigcup_{i=1}^{m} \big[Q_i, Q_i^{\omega}\big).$$

In [9] we proved a more general absolute result where the unknowns are taken from $\overline{\mathbb{Q}}$ instead of K.

4. Gap principles

In this section we state and prove two gap principles. Further, we deduce Theorems 2.1 and 2.2. Theorem 2.1 is a consequence of Theorem 3.1 and our first gap principle, while Theorem 2.2 follows from our second gap principle.

We keep the notation introduced before. Further, we put

$$\Delta_v := |\det(L_1^{(v)},\dots,L_n^{(v)})|_v \quad \text{for } v\in S.$$

We state our first gap principle. This result is well-known but we have included a proof for convenience of the reader.

Proposition 4.1. *Assume* (2.4). *Let* $Q \geqslant n^{2n/\delta}$. *Then the set of solutions* $\mathbf{x}\in O_S^n$ *of* (2.3) *with*

$$Q \leqslant \|\mathbf{x}\| < Q^{1+\delta/2n}$$

is contained in a single proper linear subspace of K^n.

Proof. Let $\mathcal{T}$ denote the set of solutions $\mathbf{x}\in O_S^n$ to (2.3) with

$$Q \leqslant H(\mathbf{x}) < Q^{1+\delta/2n} \text{ for } i=1,\dots,n.$$

Notice that for $\mathbf{x}\in\mathcal{T}$ we have, by the last condition of (2.4),

$$(4.1)\quad |L_i^{(v)}(\mathbf{x})|_v \leqslant C_v H(\mathbf{x})^{s(v)+(c_{iv}-s(v))} \leqslant C_v Q^{c_{iv}+s(v)\delta/2n} \text{ for } i=1,\dots,n.$$

Take $\mathbf{x}_1,\dots,\mathbf{x}_n\in\mathcal{T}$. First let v be an infinite place of K. Then $|\cdot|_v$ can be extended to $\overline{K_v}=\mathbb{C}$ and for this extension we have $|\cdot|_v=|\cdot|^{s(v)}$. Now by Hadamard's inequality,

$$(4.2)\qquad \begin{aligned} |\det(\mathbf{x}_1,\dots,\mathbf{x}_n)|_v &= \Delta_v^{-1}\cdot |\det\big(L_i^{(v)}(\mathbf{x}_j)\big)_{i,j}|_v \\ &\leqslant (n^{n/2})^{s(v)}\Delta_v^{-1}\prod_{i=1}^n \max_{j=1}^n |L_i^{(v)}(\mathbf{x}_j)|_v \\ &\leqslant (n^{n/2})^{s(v)}\Delta_v^{-1}C_v^n Q^{(\sum_{i=1}^n c_{iv})+s(v)\delta/2}. \end{aligned}$$

For finite $v\in S$ we have by a similar argument, but now using $s(v)=0$ and the ultrametric inequality instead of Hadamard's inequality,

$$(4.3)\qquad |\det(\mathbf{x}_1,\dots,\mathbf{x}_n)|_v \leqslant \Delta_v^{-1}C_v^n Q^{\sum_{i=1}^n c_{iv}},$$

while for the places v outside S we have, trivially,

$$(4.4)\qquad |\det(\mathbf{x}_1,\dots,\mathbf{x}_n)|_v \leqslant 1.$$

Now taking the product over $v \in M_K$ and using (2.4), $\sum_{v|\infty} s(v) = 1$, (4.2)–(4.4) and our assumption $Q > n^{2n/\delta}$ we obtain

$$\begin{aligned}\prod_{v \in M_K} |\det(\mathbf{x}_1, \ldots, \mathbf{x}_n)|_v &\leqslant n^{n/2} \prod_{v \in S} (\Delta_v^{-1} C_v^n) \cdot Q^{(\delta/2) + \sum_{v \in S} \sum_{i=1}^n c_{iv}} \\ &\leqslant n^{n/2} Q^{-\delta/2} < 1,\end{aligned}$$

and so, $\det(\mathbf{x}_1, \ldots, \mathbf{x}_n) = 0$ by the product formula. Hence $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are linearly dependent. This holds for arbitrary $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathcal{T}$. Therefore, $\mathcal{T}$ is contained in a single proper linear subspace of K^n. □

Proof of Theorem 2.1. According to Theorem 3.1, for the large solutions $\mathbf{x}$ of (2.3) outside U_0 we have $H(\mathbf{x}) \in \mathcal{U} := \bigcup_{i=1}^m [Q_i, Q_i^\omega)$. We have to cover $\mathcal{U}$ by intervals of the shape $[Q, Q^{1+\delta/2n})$ and then apply Proposition 4.1. It is not difficult to show that $\mathcal{U}$ is contained in a union of not more than

$$m\left(1 + \left[\frac{\log \omega}{\log(1 + \delta/2n)}\right]\right)$$

intervals of the shape $[Q, Q^{1+\delta/2n})$. By Proposition 4.1, this quantity, with one added to it to take care of the space U_0, is then an upper bound for the number of subspaces containing the large solutions of (2.3). This is bounded above by the quantity in Theorem 2.1. □

We now deduce a gap principle to deal with the small solutions of (2.3) which is more intricate than the one deduced above.

Proposition 4.2. *Let $d := [K : \mathbb{Q}]$ and $Q \geqslant 1$. Then the set of solutions $\mathbf{x} \in O_S^n$ of (2.3) with*

$$Q \leqslant H(\mathbf{x}) < 2Q^{1+\delta/2n}$$

is contained in a union of at most

$$(90n)^{nd}$$

proper linear subspaces of K^n.

If $K = \mathbb{Q}$ this upper bound can be replaced by

$$200^n.$$

In the proof we need a number of lemmas. For $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{C}^n$, define $\|\mathbf{y}\| := \max(|y_1|, \ldots, |y_n|)$.

Lemma 4.3. *Let $M \geqslant 1$. We can partition $\mathbb{C}^n$ into at most $(20n)^n M^2$ subsets, such that for any $\mathbf{y}_1, \dots, \mathbf{y}_n \in \mathbb{C}^n$ belonging to the same subset,*

$$|\det(\mathbf{y}_1, \dots, \mathbf{y}_n)| \leqslant M^{-1} \|\mathbf{y}_1\| \cdots \|\mathbf{y}_n\|. \tag{4.5}$$

Proof. We can express any non-zero $\mathbf{y} \in \mathbb{C}^n$ uniquely as $\lambda \cdot \mathbf{z}$, where λ is a complex number with $|\lambda| = \|\mathbf{y}\|$, and where $\mathbf{z} = (z_1, \dots, z_n) \in \mathbb{C}^n$ with $\|\mathbf{z}\| = 1$ and with $|z_j| < 1$ for $j < i$ and $z_i = 1$ for some $i \in \{1, \dots, n\}$. For $j = 1, \dots, n$, $j \neq i$ we write $z_j = u_j + \sqrt{-1} v_j$ with $u_j, v_j \in \mathbb{R}$. Further, we express $\mathbf{0}$ as $0 \cdot \mathbf{z}$ with $\mathbf{z} = (1, 0, \dots, 0)$, and put $u_j = 0, v_j = 0$ for $j = 2, \dots, n$. Thus, with every $\mathbf{y} \in \mathbb{C}^n$ we associate a unique index $i \in \{1, \dots, n\}$ and a unique vector $\mathbf{w} = (u_j, v_j : j \neq i) \in [-1, 1]^{2n-2}$.

Let $K := \big(M \cdot n^{n/2}\big)^{1/(n-1)}$. We divide the $(2n-2)$-dimensional cube $[-1, 1]^{2n-2}$ into at most $([2\sqrt{2}K]+1)^{2n-2}$ subcubes of size at most $(\sqrt{2} \cdot K)^{-1}$. Then we divide $\mathbb{C}^n$ into at most $n([2\sqrt{2} \cdot K] + 1)^{2n-2}$ classes such that two vectors $\mathbf{y}$ belong to the same class if the indices i associated with them are equal, and the vectors $\mathbf{w}$ associated with them belong to the same subcube. Notice that the number of classes is bounded above by

$$n\Big(2\sqrt{2} \cdot \big(M \cdot n^{n/2}\big)^{1/(n-1)} + 1\Big)^{2n-2} \leqslant (20n)^n M^2.$$

Now let $\mathbf{y}_1, \dots, \mathbf{y}_n$ belong to the same class. For $k = 1, \dots, n$, write $\mathbf{y}_k = \lambda_k \mathbf{z}_k$ as above and let $\mathbf{w}_k$ be the corresponding vector from $[-1, 1]^{2n-2}$. Since $\mathbf{w}_1, \dots, \mathbf{w}_n$ belong to the same subcube we have

$$\|\mathbf{z}_k - \mathbf{z}_1\| \leqslant \sqrt{2} \cdot \|\mathbf{w}_k - \mathbf{w}_1\| \leqslant K^{-1}$$

for $k = 2, \dots, n$. Hence, using Hadamard's inequality,

$$\begin{aligned} |\det(\mathbf{z}_1, \dots, \mathbf{z}_n)| &= |\det(\mathbf{z}_1, \mathbf{z}_2 - \mathbf{z}_1, \dots, \mathbf{z}_n - \mathbf{z}_1)| \\ &\leqslant n^{n/2} (K^{-1})^{n-1} = M^{-1} \end{aligned}$$

which implies

$$\begin{aligned} |\det(\mathbf{y}_1, \dots, \mathbf{y}_n)| &= |\lambda_1 \cdots \lambda_n| \cdot |\det(\mathbf{z}_1, \dots, \mathbf{z}_n)| \\ &\leqslant M^{-1} \|\mathbf{y}_1\| \cdots \|\mathbf{y}_n\|. \end{aligned}$$

This completes our proof. □

Lemma 4.4. *Let D be a positive real, and let $\mathcal{S}$ be a subset of $\mathbb{Z}^n$ such that*

$$|\det(\mathbf{x}_1,\dots,\mathbf{x}_n)| \leqslant D \quad \text{for } \mathbf{x}_1,\dots,\mathbf{x}_n \in \mathcal{S}.$$

Then $\mathcal{S}$ is contained in a union of at most

$$100^n D^{1/(n-1)}$$

proper linear subspaces of $\mathbb{Q}^n$.

Proof. This is Lemma 5 of [8]. □

We deduce the following consequence.

Lemma 4.5. *Let D_v $(v \in M_{\mathbb{Q}})$ be positive reals such that $D_v = 1$ for all but finitely many v and put $D := \prod_{v\in M_{\mathbb{Q}}} D_v$. Let $\mathcal{T}$ be a subset of $\mathbb{Q}^n$ such that*

$$|\det(\mathbf{x}_1,\dots,\mathbf{x}_n)|_v \leqslant D_v \text{ for } v \in M_{\mathbb{Q}},\ \mathbf{x}_1,\dots,\mathbf{x}_n \in \mathcal{T}. \tag{4.6}$$

Then $\mathcal{T}$ is contained in a union of at most

$$100^n D^{1/(n-1)} \tag{4.7}$$

proper linear subspaces of $\mathbb{Q}^n$.

Proof. Without loss of generality we assume that $\mathcal{T}$ is not contained in a proper linear subspace of $\mathbb{Q}^n$. Further, without loss of generality we assume that for every finite place v of $\mathbb{Q}$,

$$D_v = \max\{|\det(\mathbf{x}_1,\dots,\mathbf{x}_n)|_v : \mathbf{x}_1,\dots,\mathbf{x}_n \in \mathcal{T}\}.$$

Indeed, if the maximum were $D_v' < D_v$, we could replace D_v by D_v' without strengthening (4.6), and replace (4.7) by a smaller upper bound.

Fix a finite place v and let $\mathbb{Z}_v := \{x \in \mathbb{Q} : |x|_v \leqslant 1\}$, i.e., $\mathbb{Z}_v$ is the localization of $\mathbb{Z}$ at v. Choose $\mathbf{y}_1,\dots,\mathbf{y}_n \in \mathcal{T}$ such that $|\det(\mathbf{y}_1,\dots,\mathbf{y}_n)|_v = D_v$, and let $\mathcal{M}_v$ denote the $\mathbb{Z}_v$-module generated by $\mathbf{y}_1,\dots,\mathbf{y}_n$. Now if $\mathbf{x} \in \mathcal{T}$, then $\mathbf{x} = \sum_{i=1}^n u_i\mathbf{y}_i$ with $u_1,\dots,u_n \in \mathbb{Q}$. We can express u_i as a quotient of two determinants, where in the denominator we have $\det(\mathbf{y}_1,\dots,\mathbf{y}_n)$, and in the numerator the determinant obtained by replacing $\mathbf{y}_i$ by $\mathbf{x}$. Using (4.6), this implies that $|u_i|_v \leqslant 1$ for $i = 1,\dots,n$. Hence $\mathcal{T}$ is contained in $\mathcal{M}_v$.

Applying this for every finite place v, we infer that $\mathcal{T}$ is contained in $\mathcal{M} := \bigcap_{v\neq\infty} \mathcal{M}_v$, where the intersection is over all finite places. The set $\mathcal{M}$ is a lattice of rank n in $\mathbb{Q}^n$ of determinant $\Delta := \big(\prod_{v\neq\infty} D_v\big)^{-1}$. Choose a basis $\mathbf{z}_1, \dots, \mathbf{z}_n$ of $\mathcal{M}$. Then $|\det(\mathbf{z}_1, \dots, \mathbf{z}_n)| = \Delta$. Define the linear map $\varphi :\ \mathbf{u} = (u_1, \dots, u_n) \mapsto \sum_{i=1}^n u_i\mathbf{z}_i$ and let $\mathcal{S} := \varphi^{-1}(\mathcal{T})$. Then $\mathcal{S} \subseteq \mathbb{Z}^n$ and for any $\mathbf{u}_1, \dots, \mathbf{u}_n \in \mathcal{S}$ we have

$$\begin{aligned} |\det(\mathbf{u}_1, \dots, \mathbf{u}_n)| &= \Delta^{-1} \cdot |\det(\varphi(\mathbf{u}_1), \dots, \varphi(\mathbf{u}_n))| \\ &\leqslant \Delta^{-1} D_\infty = \prod_{v\in M_\mathbb{Q}} D_v = D. \end{aligned}$$

Now by Lemma 4.4, the set $\mathcal{S}$, and hence also $\mathcal{T}$, is contained in a union of not more than $100^n D^{1/(n-1)}$ proper linear subspaces of $\mathbb{Q}^n$. □

We leave as an open problem to generalize the above Lemma to arbitrary algebraic number fields.

Proof of Proposition 4.2. We start with the case that K is an arbitrary number field. Let $\mathcal{T}'$ be the set of solutions $\mathbf{x} \in O_S^n$ of (2.3) with $Q \leqslant H(\mathbf{x}) < 2Q^{1+\delta/2n}$. Completely analogously to (4.1) we have for $\mathbf{x} \in \mathcal{T}'$, $v \in S$, $i = 1, \dots, n$,

$$|L_i^{(v)}(\mathbf{x})|_v \leqslant 2^{s(v)} C_v Q^{c_{iv}+s(v)\delta/2n}. \tag{4.8}$$

For $\mathbf{x} \in \mathcal{T}'$ and any infinite place v of K, define the vector

$$\varphi_v(\mathbf{x}) := \big(Q^{-c_{1v}/s(v)} L_1^{(v)}(\mathbf{x}), \dots, Q^{-c_{nv}/s(v)} L_n^{(v)}(\mathbf{x})\big).$$

Notice that for each infinite place v of K we have $\varphi_v(\mathbf{x}) \in \mathbb{C}^n$. Put $M := (9/2)^{n/2}$. By Lemma 4.3, and since K has at most d infinite places, we can partition $\mathcal{T}'$ into at most

$$(20n)^{nd} M^{2d} \leqslant (90n)^{nd}$$

classes, such that if $\mathbf{x}_1, \dots, \mathbf{x}_n$ belong to the same class, then for each infinite place v,

$$|\det\big(\varphi_v(\mathbf{x}_1), \dots, \varphi_v(\mathbf{x}_n)\big)| \leqslant M^{-1} \prod_{i=1}^n \|\varphi_v(\mathbf{x})\|.$$

We show that the set of elements of $\mathcal{T}'$ from a given class is contained in a proper linear subspace of K^n, that is, that any n elements of $\mathcal{T}'$ from the same class have determinant 0. So let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be elements of $\mathcal{T}'$ from the same class. Then by (4.8) and what we just proved, we have for every infinite place v of K, using $|\cdot|_v = |\cdot|_v^{s(v)}$ on $\overline{K_v} = \mathbb{C}$,

$$\begin{aligned} |\det(\mathbf{x}_1, \ldots, \mathbf{x}_n)|_v &= \Delta_v^{-1} Q^{\sum_{i=1}^n c_{iv}} \cdot |\det\big(\varphi_v(\mathbf{x}_1), \ldots, \varphi_v(\mathbf{x}_n)\big)|^{s(v)} \\ &\leqslant \Delta_v^{-1} Q^{\sum_{i=1}^n c_{iv}} M^{-s(v)} \prod_{i=1}^n \|\varphi_v(\mathbf{x})\|^{s(v)} \\ &\leqslant \Delta_v^{-1} Q^{\sum_{i=1}^n c_{iv}} M^{-s(v)} C_v^n 2^{ns(v)} Q^{s(v)\delta/2}, \end{aligned}$$

which, thanks to our choice of M, yields

$$|\det(\mathbf{x}_1, \ldots, \mathbf{x}_n)|_v < \Delta_v^{-1} C_v Q^{(\sum_{i=1}^n c_{iv}) + s(v)\delta/2}.$$

For the finite places $v \in S$ we have (4.3) and for the places v outside S, (4.4). By taking the product over $v \in M_K$, using (2.4), we obtain

$$\prod_{v \in M_K} |\det(\mathbf{x}_1, \ldots, \mathbf{x}_n)|_v < \prod_{v \in S} (\Delta_v^{-1} C_v^n) Q^{\sum_{v\in S}\sum_{i=1}^n c_{iv} + (\delta/2)} \leqslant Q^{-\delta/2} \leqslant 1.$$

Now the product formula implies indeed that for any $\mathbf{x}_1, \ldots, \mathbf{x}_n$ in the same class we have $\det(\mathbf{x}_1, \ldots, \mathbf{x}_n) = 0$. This proves Proposition 4.2 in the case that K is an arbitrary algebraic number field.

Now let $K = \mathbb{Q}$. Let $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathcal{T}'$. First let $v = \infty$ be the infinite place of $\mathbb{Q}$. Notice that $s(\infty) = 1$. Then using (4.8) we obtain in a similar manner as (4.2),

$$|\det(\mathbf{x}_1, \ldots, \mathbf{x}_n)|_\infty \leqslant n^{n/2} \Delta_\infty^{-1} C_\infty^n \cdot 2Q^{(\sum_{i=1}^n c_{i\infty}) + \delta/2}.$$

For the finite places in S and for the places outside S we have (4.3), (4.4). Now using Lemma 4.5, (2.4), we infer that $\mathcal{T}'$ is contained in a union of at most

$$\begin{aligned} 100^n \Big(2n^{n/2} \prod_{v \in S} (\Delta_v^{-1} C_v^n) \cdot Q^{\sum_{v\in S}\sum_{i=1}^n c_{iv} + (\delta/2)}\Big)^{1/(n-1)} \\ \leqslant 100^n \big(2n^{n/2}\big)^{1/(n-1)} < 200^n \end{aligned}$$

proper linear subspaces of $\mathbb{Q}^n$. This completes our proof. □

Proof of Theorem 2.2. Let K be an arbitrary algebraic number field of degree d. We divide the solutions into consideration into those with $H(\mathbf{x}) \in I_1$ and those with $H(\mathbf{x}) \in I_2$, where

$$I_1 = \big[n^{2n/\delta}, \max(2H, n^{2n/\delta})\big), \quad I_2 = \big[1, n^{2n/\delta}\big).$$

We have $I_1 \subseteq \bigcup_{h=0}^{A-1} \big[Q_h, Q_h^{1+\delta/2n}\big)$, where

$$Q_h = (n^{2n/\delta})^{(1+\delta/2n)^h} \quad (h = 0, 1, 2, \dots),$$

$$A = 1 + \left[\frac{\log\Big(\log\max(2H, n^{2n/\delta})/\log n^{2n/\delta}\Big)}{\log(1+\delta/2n)}\right] \leqslant 4n\delta^{-1}\log\log 4H.$$

So by Proposition 4.1, the solutions $\mathbf{x} \in O_S^n$ of (2.3) with $H(\mathbf{x}) \in I_1$ lie in a union of at most A proper linear subspaces of K^n.

Next, we have $I_2 \subseteq \bigcup_{h=0}^{B-1} \big[Q_h, 2Q_h^{1+\delta/2n}\big)$, where

$$Q_h = 2^{\gamma_h} \text{ with } \gamma_h = \frac{2n}{\delta}\Big((1 + (\delta/2n))^h - 1\Big) \quad (h = 0, 1, 2, \dots),$$

$$B = 1 + \left[\frac{\log(1 + \log n/\log 2)}{\log(1+\delta/2n)}\right] \leqslant 4n\delta^{-1}\log(3\log n).$$

So by Proposition 4.2, the solutions $\mathbf{x} \in O_S^n$ of (2.3) with $H(\mathbf{x}) \in I_2$ lie in a union of at most $(90n)^{nd}B$ proper linear subspaces of K^n.

We conclude that the number of subspaces containing the solutions $\mathbf{x} \in O_S^n$ of (2.3) with $H(\mathbf{x}) \leqslant \max(2H, n^{2n/\delta})$ is bounded above by

$$A + (90n)^{nd}B \leqslant \delta^{-1}\big((10^3 n)^{nd} + 4n\log\log 4H\big).$$

In the case $K = \mathbb{Q}$ we have a similar computation, replacing $(90n)^{nd}$ by 200^n. □

5. On the number of solutions outside the exceptional subspace U_0

It seems to be a very difficult open problem to give an upper bound for the number of solutions of (2.3) lying outside the exceptional subspace U_0 from Theorem C. To obtain such a bound we would have to combine the interval result Theorem 3.1 with some strengthening of the gap principle Proposition 4.1 giving an upper bound for the number of solutions $\mathbf{x}$ with

$Q \leqslant H(\mathbf{x}) < Q^{1+\delta/2n}$ instead of the number of subspaces containing these solutions. But this seems to be totally out of reach. However, such a strong gap principle may exist in certain applications where one considers solutions $\mathbf{x}$ with additional constraints, and then it may be possible to estimate from above the number of such restricted solutions.

In 1990, Schmidt [24] gave an example of a system of inequalities (2.3) which is known to have finitely many solutions, but which is such that from any explicit upper bound for the number of solutions of this system one can derive a very strong *effective* finiteness result for some related system of Diophantine inequalities.

We give another such example, which is a modification of a result from Hirata-Kohno and the author [10]. We consider the inequality

$$|x_1 + x_2\xi + x_3\xi^2| \leqslant H(\mathbf{x})^{-2-\delta} \quad \text{in } \mathbf{x} = (x_1, x_2, x_3) \in \mathbb{Z}^3 \text{ with } \gcd(x_1, x_2, x_3) = 1, \tag{5.1}$$

where ξ is a real algebraic number of degree $\geqslant 3$ and where $\delta > 0$. By augmenting this single inequality with the two trivial inequalities

$$|x_2| \leqslant H(\mathbf{x}), \quad |x_3| \leqslant H(\mathbf{x})$$

we obtain a system of type (2.3). Since ξ has degree at least 3, the linear form $X_1 + X_2\xi + X_3\xi^2$ does not vanish identically on any non-zero non-linear subspace U of $\mathbb{Q}^3$. Consequently, if U is a linear subspace of $\mathbb{Q}^3$ of dimension $k > 0$ we have $\nu(U) = -2 - \delta + k - 1$. Hence

$$\mu(U) = \frac{\nu(\mathbb{Q}^3) - \nu(U)}{3 - \dim U} = 1$$

if $U \neq (0)$ and $\mu((0)) = -\delta/3 < 1$. So according to the description of U_0 in Section 3, we have $U_0 = (\mathbf{0})$ and by Theorem C, (5.1) has only finitely many solutions. (This can also be deduced directly from Theorem A).

We prove the following Proposition.

Proposition 5.1. *Let N be an upper bound for the number of solutions of* (5.1). *Then for every $\alpha \in \mathbb{Q}$ we have*

$$|\xi - \alpha| \geqslant 2^{-2-\delta}(1 + |\xi|)^{-1} N^{-3-\delta} \cdot H(\alpha)^{-3-\delta}. \tag{5.2}$$

One of the most wanted achievements in Diophantine approximation would be to prove an effective version of Roth's Theorem, i.e., an inequality of the shape

$$|\xi-\alpha|\geqslant C(\xi,\delta)H(\alpha)^{-2-\delta} \text{ for } \alpha\in\mathbb{Q}$$

with some effectively computable constant $C(\xi,\delta)>0$. Our Proposition implies that from an explicit upper bound for the number of solutions of (5.1) one would be able to deduce an effective inequality with instead of an exponent $2+\delta$ an exponent $3+\delta$. Save some special cases, such a result is much stronger than any of the effective results on the approximation of algebraic numbers by rationals that have been obtained so far.

Proof. Let α be a rational number. We can express α as $\alpha=r/s$, where r,s are rational integers with $s>0$, $\gcd(r,s)=1$. Thus, $H(\alpha)=\max(|r|,|s|)$. Let u be an integer with

$$|u|\leqslant\Big(2^{2+\delta}(1+|\xi|)\cdot|\xi-\alpha|\cdot H(\alpha)^{3+\delta}\Big)^{-1/(3+\delta)}. \tag{5.3}$$

We assume that the right-hand side is at least 1; otherwise (5.2) follows at once.

Define the vector $\mathbf{x}=(x_1,x_2,x_3)$ by $x_1+x_2X+x_3X^2=(u+X)(r-sX)$. Then $\mathbf{x}\in\mathbb{Z}^3$, $\gcd(x_1,x_2,x_3)=1$ and by (5.3),

$$\begin{aligned}|x_1+x_2\xi+x_3\xi^2| &= |u+\xi|\cdot|r-s\xi|\\ &\leqslant (1+|\xi|)\max(1,|u|)\max(|r|,|s|)\cdot|\xi-\alpha|\\ &\leqslant \Big(2\max(1,|u|)\max(|r|,|s|)\Big)^{-2-\delta}\leqslant H(\mathbf{x})^{-2-\delta}.\end{aligned}$$

Thus, each integer u with (5.3) gives rise to a solution of (5.1). Consequently, the number of solutions of (5.1), and hence N, is bounded from below by the right-hand side of (5.3). Now (5.2) follows by a straightforward computation. □

6. About the proofs of Theorems 2.1 and 3.1

We discuss in somewhat more detail the new ideas leading to the improved bound for the number of subspaces in Theorem 2.1 as compared with Theorem B. For simplicity, we consider only the special case $K=\mathbb{Q}$,

$S = \{\infty\}$, $O_S = \mathbb{Z}$. Notice that $H(\mathbf{x}) = \|\mathbf{x}\| = \max(|x_1|,\dots,|x_n|)$ for $\mathbf{x} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$. Thus, we consider systems of inequalities

$$(6.1)\qquad |L_i(\mathbf{x})| \leqslant C \cdot \|\mathbf{x}\|^{c_i}\ (i = 1,\dots,n) \ \text{ in } \mathbf{x} \in \mathbb{Z}^n,$$

where $L_1,\dots,L_n$ are linearly independent linear forms in $X_1,\dots,X_n$ with coefficients in $\mathbb{C}$ that are algebraic over $\mathbb{Q}$, $0 < C \leqslant |\det(L_1,\dots,L_n)|^{1/n}$, and $c_1 + \cdots + c_n \leqslant -\delta$ with $0 < \delta \leqslant 1$.

With a solution $\mathbf{x} \in \mathbb{Z}^n$ we associate a convex body $\Pi(\mathbf{x})$, consisting of those $\mathbf{y} \in \mathbb{R}^n$ such that

$$|L_i(\mathbf{y})| \leqslant C\|\mathbf{x}\|^{c_i} \text{ for } i = 1,\dots,n.$$

Denote by $\lambda_i(\mathbf{x})$ $(i = 1,\dots,n)$ the successive minima of this body. Then $\lambda_1(\mathbf{x}) \leqslant 1$, and by Minkowski's theorem, $\prod_{i=1}^n \lambda_i(\mathbf{x}) \gg \mathrm{vol}\big(\Pi(\mathbf{x})\big)^{-1} \gg \|\mathbf{x}\|^{\delta}$, where here and below, the constants implied by $\ll$, $\gg$ depend on n, $L_1,\dots,L_n$ and δ.

There is an index $k \in \{1,\dots,n-1\}$ such that $\lambda_k(\mathbf{x})/\lambda_{k+1}(\mathbf{x}) \ll \|\mathbf{x}\|^{-\delta/n}$. To apply the approximation techniques going into the Subspace Theorem, one needs that the one but last minimum $\lambda_{n-1}(\mathbf{x})$ is $\ll 1$. In general, this need not be the case. Schmidt's ingenious idea was, to construct from $\Pi(\mathbf{x})$ a new convex body $\widehat{\Pi}(\mathbf{x})$ in $\wedge^{n-k}\mathbb{R}^n \cong \mathbb{R}^N$ with $N := \binom{n}{k}$ of which the one but last minimum is indeed $\ll 1$. The body $\widehat{\Pi}(\mathbf{x})$ may be described as the set of $\widehat{\mathbf{y}} \in \mathbb{R}^N$ such that

$$(6.2)\qquad |M_i(\widehat{\mathbf{y}})| \ll \|\mathbf{x}\|^{e_i(\mathbf{x})} \text{ for } i = 1,\dots,N,$$

where $M_1,\dots,M_N$ are linearly independent linear forms in N variables with real algebraic coefficients, and $e_1(\mathbf{x}),\dots,e_N(\mathbf{x})$ are exponents, which unfortunately may depend on $\mathbf{x}$, such that $\sum_{i=1}^N e_i(\mathbf{x}) < -\delta/2n^2$, say, see [22] or [4] for more details on Schmidt's construction. As mentioned before, the one but last minimum of $\widehat{\Pi}(\mathbf{x})$ is $\ll 1$. Then by Minkowski's Theorem, the last minimum is $\gg \|\mathbf{x}\|^{\delta/2n^2}$. This implies that $\widehat{\Pi}(\mathbf{x}) \cap \mathbb{Z}^N$ spans a linear subspace $T(\mathbf{x})$ of $\mathbb{Q}^N$ of dimension $N-1$.

In their proof of Theorem B, Schlickewei and the author had to partition the set of solutions of (6.1) into classes in such a way, that for any two solutions $\mathbf{x}$, $\mathbf{x}'$ in the same class, we have $e_i(\mathbf{x}) \approx e_i(\mathbf{x}')$ for $i = 1, \dots, N$. Then they proceeded further with solutions from the same class.

The continuation of the proof of Schlickewei and the author is then as follows. Suppose there are solutions $\mathbf{x}_1, \dots, \mathbf{x}_M$ in the same class such that $\|\mathbf{x}_1\|$ is large and $\log \|\mathbf{x}_{i+1}\| / \log \|\mathbf{x}_i\|$ are large for $i = 1, \dots, M-1$, where M and "large" depend on δ, n and $L_1, \dots, L_n$. Then one constructs an auxiliary multihomogeneous polynomial $P(\mathbf{Y}_1, \dots, \mathbf{Y}_M)$ in M blocks of N variables with integer coefficients, which is of degree d_i in block $\mathbf{Y}_i$ for $i = 1, \dots, M$, where $\|\mathbf{x}_1\|^{d_1} \approx \cdots \approx \|\mathbf{x}_M\|^{d_M}$. The polynomial P is such that $|P_I(\widehat{\mathbf{y}}_1, \dots, \widehat{\mathbf{y}}_M)| < 1$ for all $\widehat{\mathbf{y}}_h \in \widehat{\Pi}(\mathbf{x}_h) \cap \mathbb{Z}^N$, $h = 1, \dots, M$, and all partial derivatives P_I of P of not too large order. Then for these I, $\widehat{\mathbf{y}}_1, \dots, \widehat{\mathbf{y}}_M$ we have that $P_I(\widehat{\mathbf{y}}_1, \dots, \widehat{\mathbf{y}}_M) = 0$. By extrapolation it then follows that all P_I vanish identically on $T(\mathbf{x}_1) \times \cdots \times T(\mathbf{x}_M)$. On the other hand, using an extension of Roth's Lemma, proved also by Schmidt, one shows that such a polynomial cannot exist.

This contradiction shows that solutions $\mathbf{x}_1, \dots, \mathbf{x}_M$ as above cannot exist. This leads to an upper bound depending on n, δ, D for the number of subspaces containing the solutions of (6.1) belonging to a given class. We have to multiply this with the number of classes to get our final bound for the number of subspaces containing the solutions from all classes together. As it turns out, the number of classes is at most $\gamma_1^{n^2} \delta^{-\gamma_2 n}$ with absolute constants γ_1, γ_2 and in terms of n, δ, this dominates the resulting bound for the number of subspaces.

In their proof of Theorem 2.1, Ferretti and the author used, instead of Schmidt's multi-homogeneous polynomial, the one constructed by Faltings and Wüstholz [15]. The latter polynomial has the great advantage, that the argument sketched above works also for solutions $\mathbf{x}_1, \dots, \mathbf{x}_M$ not necessarily belonging to the same class. Thus, a subdivision of the solutions of (6.1) into classes is not necessary, and we can save a factor $\gamma_1^{n^2} \delta^{-\gamma_2 n}$ in the final upper bound for the number of subspaces.

The proof of the interval result Theorem 3.1 follows the same lines. First one proves Theorem 3.1 in the special case that the exceptional subspace $U_0 = (\mathbf{0})$. Assuming that Theorem 3.1 is false, one arrives at a contradiction using Schmidt's construction of $\widehat{\Pi}(\mathbf{x})$, Faltings' and Wüstholz' construction of an auxiliary polynomial, and Schmidt's extension of Roth's Lemma. Then one proves the result for arbitrary U_0 by considering a system derived from (6.1) with solutions taken from the quotient $\mathbb{Q}^n/U_0$.

We now discuss the constructions of an auxiliary polynomial by Schmidt and by Faltings and Wüstholz, respectively.

We have to construct a non-zero multihomogeneous polynomial

$$P(\mathbf{Y}_1,\dots,\mathbf{Y}_M)\in\mathbb{Z}[\mathbf{Y}_1,\dots,\mathbf{Y}_M]$$

in M blocks $\mathbf{Y}_1,\dots,\mathbf{Y}_M$ of N variables, which is homogeneous of degree d_h in the block $\mathbf{Y}_h$ for $h=1,\dots,M$. This polynomial can be expressed as

$$\sum_{\mathbf{i}} c(\mathbf{i})\prod_{h=1}^{M}\prod_{j=1}^{N} M_j(\mathbf{Y}_h)^{i_{hj}}$$

where the summation is over tuples $\mathbf{i}=(i_{hj})$ such that $\sum_{j=1}^{N} i_{hj}=d_h$ for $h=1,\dots,M$.

Schmidt's approach is to construct P with coefficients with small absolute values, such that

$$c(\mathbf{i})=0 \text{ if } \max_{1\leqslant j\leqslant N}\left|\Big(\sum_{h=1}^{M}\frac{i_{hj}}{d_h}\Big)-\frac{M}{N}\right|\geqslant\varepsilon$$

for some sufficiently small ε. The conditions $c(\mathbf{i})=0$ may be viewed as linear equations in the unknown coefficients of P. We may consider the indices i_{hj} as random variables with expectation $1/N$. Then the law of large numbers from probability theory implies that for sufficiently large M, the number of conditions $c(\mathbf{i})=0$ is smaller than the total number of coefficients of P. Now Siegel's Lemma gives a non-zero polynomial P with coefficients with small absolute values.

The approach of Faltings and Wüstholz is as follows. Let $\alpha_{hj}\in\mathbb{R}$ with $|\alpha_{hj}|\leqslant 1$ for $h=1,\dots,m$, $j=1,\dots,R$. Construct P with coefficients with small absolute values such that

$$c(\mathbf{i})=0 \text{ if } \left|\sum_{h=1}^{M}\sum_{j=1}^{N}\alpha_{hj}\Big(\frac{i_{hj}}{d_h}-\frac{1}{N}\Big)\right|\geqslant\varepsilon.$$

Again, thanks to the law of large numbers, for sufficiently large M the number of conditions $c(\mathbf{i}) = 0$ is smaller than the number of coefficients of P, and then P is obtained via an application of Siegel's Lemma.

The choice of the weights α_{hj} is completely free. In fact, if we are given solutions $\mathbf{x}_1, \dots, \mathbf{x}_M$ of (6.1) from different classes, we may choose the α_{hj} in a suitable manner depending on the exponents $e_i(\mathbf{x}_h)$ ($i = 1, \dots, N$, $h = 1, \dots, M$) from (6.2), and then show that $|P_I(\widehat{\mathbf{y}}_1, \dots, \widehat{\mathbf{y}}_M)| < 1$ for all $\mathbf{y}_h \in \widehat{\Pi}(\mathbf{x}_h) \cap \mathbb{Z}^N$, $h = 1, \dots, M$, and all partial derivatives P_I of P of not too large order. Then the proofs of Theorems 2.1 and 3.1 are completed as sketched above.

References

[1] E. Bombieri, A.J. van der Poorten, *Some quantitative results related to Roth's Theorem,* J. Austral. Math. Soc. (Ser. A) **45** (1988), 233–248.

[2] Y. Bugeaud, J.-H. Evertse, *On two notions of complexity of algebraic numbers,* Acta Arith. **133** (2008), 221–250.

[3] H. Davenport, K.F. Roth, *rational approximations to algebraic numbers,* Mathematika **2** (1955), 160–167.

[4] J.-H. Evertse, *The Subspace Theorem of W.M. Schmidt,* In: Diophantine Approximation and Abelian Varieties, Introductory lectures, ed. by B. Edixhoven, J.-H. Evertse, Lecture Notes Math. **1566**, Springer Verlag, 1993, Chap. IV, 31–50.

[5] J.-H. Evertse, *An explicit version of Faltings' Product theorem and an improvement of Roth's lemma,* Acta Arith. **73** (1995), 215–248.

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

[7] J.-H. Evertse, *The number of algebraic numbers of given degree approximating a given algebraic number,* In: Analytic Number Theory, Proc. conf. Kyoto, 1996, ed. by Y. Motohashi, Cambridge Univ. Press 1997, 53–83.

[8] J.-H. Evertse, *On the norm form inequality $|F(\mathbf{x})| \leqslant h$,* Publ. Math. Debrecen **56** (2000), 337–374.

[9] J.-H. Evertse, R.G. Ferretti, *A further improvement of the quantitative Subspace Theorem,* preprint, available from the first author.

[10] J.-H. Evertse, N. Hirata-Kohno, *Wirsing systems and resultant inequalities,* In: Number Theory for the Millennium, vol. I, ed. by M.A. Bennett, B.C. Berndt, N. Boston, H.G. diamond, A.J. Hildebrand, W. Philip, A.K. Peters, Natick, Mass. 2002, 449–461.

[11] J.-H. Evertse, H.P. Schlickewei, *The absolute subspace Theorem and linear equations with unknowns from a multiplicative group,* in: Number Theory in Progress, vol. I, ed. by K. Győry, H. Iwaniec, J. Urbanowicz, Walter de Gruyter, Berlin-New York, 1999, 121–142.

[12] J.-H. Evertse, H.P. Schlickewei, *A quantitative version of the Absolute Subspace Theorem,* J. reine angew. Math. **548** (2002), 21–127.

[13] J.-H. Evertse, H.P. Schlickewei, W.M. Schmidt, *Linear equations in variables which lie in a multiplicative group,* Annals of Math. **155** (2002), 807–836.

[14] G. Faltings, *Diophantine approximation on abelian varieties,* Annals of Math. **133** (1991), 549–576.

[15] G. Faltings, G. Wüstholz, *Diophantine approximations on projective spaces,* Invent. math. **116** (1994), 109–138.

[16] M. Mignotte, *Quelques remarques sur l'approximation rationnelle des nombres algébriques,* J. reine angew. Math. **268/269** (1974), 341–347.

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

[18] H.P. Schlickewei, *The* $\mathfrak{p}$*-adic Thue-Siegel-Roth-Schmidt theorem*, Arch. Math. (Basel) **29** (1977), 267–270.

[19] H. P. Schlickewei, *The quantitative Subspace Theorem for number fields,* Compositio Math. **82** (1992), 245–273.

[20] H.P. Schlickewei, *Multiplicities of recurrence sequences,* Acta Math. **176** (1996), 171–242.

[21] W.M. Schmidt, *Norm form equations,* Annals of Math. **96** (1972), 526–551.

[22] W.M. Schmidt, *Diophantine Approximation,* Lecture Notes Math. **785**, Springer Verlag, 1980.

[23] W.M. Schmidt, *The Subspace Theorem in diophantine approximations,* Compos. Math **69** (1989), 121–173.

[24] W.M. Schmidt, *The number of rational approximations by algebraic numbers and the number of solutions of norm form equations,* In: Number Theory and related topics, proc. Conf. Bombay, 1988, Tata Inst. Fund. Res. Stud. Math. 12, Tata Inst. Fund. Res., Bombay, 1989, 195–202.

[25] W.M. Schmidt, *Vojta's refinement of the Subspace Theorem,* Trans. Amer. Math. Soc. **340** (1993), 705–731.

[26] W.M. Schmidt, *The zero multiplicity of linear recurrence sequences,* Acta Math. **182** (1999), 243–282.

[27] W.M. Schmidt, *Counting diophantine approximations of small height,* J. Ramanujan Math. Soc. **24** (2009), 171–190.

[28] P. Vojta, *A refinement of Schmidt's Subspace Theorem,* Amer. J. Math. **111** (1989), 489–518.

J.-H. Evertse, Universiteit Leiden, Mathematisch Instituut, Postbus 9512, 2300 RA Leiden, The Netherlands

E-mail address: evertse@math.leidenuniv.nl