THE SUBSPACE THEOREM

JAN-HENDRIK EVERTSE

April 2011

Literature:

W.M. Schmidt, Diophantine approximation, Lecture Notes in Mathematics 785, Springer Verlag 1980, Chap. V,VI,VII

1. Introduction

The Subspace Theorem is a higher dimensional generalization of Roth's Theorem on the approximation of algebraic numbers by rational numbers. We explain the Subspace Theorem, give some applications to simultaneous Diophantine approximation, and then an application to higher dimensional generalizations of Thue equations, the so-called *norm form equations.*

Let α be an irrational, real number. Given a rational number x/y with $x, y \in \mathbb{Z}$, $\gcd(x, y) = 1$ and $y > 0$, we define the quality of x/y to be the real M such that $|\alpha - x/y| = y^{-M}$ if $|\alpha - x/y| < 1$; if $|\alpha - x/y| \geqslant 1$ we set $M = 0$. Dirichlet's Theorem implies that every irrational real number has infinitely many rational approximations of quality $\leqslant 2$:

Theorem 1.1. (Dirichlet, 1842). *Let $\alpha \in \mathbb{R}$, $\alpha \notin \mathbb{Q}$. Then there are infinitely many pairs of integers x, y such that*

$$\left|\alpha - \frac{x}{y}\right| \leqslant y^{-2}, \quad y > 0. \tag{1.1}$$

The famous theorem of Roth implies that if α is an irrational algebraic number, then for every $\delta > 0$, it has only finitely many rational approximations of quality $\geqslant 2 + \delta$.

Theorem 1.2. (Roth, 1955) *Let $\alpha \in \mathbb{C}$ be an irrational algebraic number and let $\delta > 0$. Then there are only finitely many pairs of integers (x, y) such* *that*

$$\left|\alpha - \frac{x}{y}\right| \leqslant y^{-2-\delta}, \quad y > 0. \tag{1.2}$$

Roth's Theorem is easy to prove if $\alpha \in \mathbb{C}\backslash\mathbb{R}$, or if α is a real quadratic number. For real algebraic numbers α of degree $\geqslant 3$, the proof of Roth's Theorem is very difficult.

In the formulation of the Subspace Theorem, we need some notions from linear algebra, which we recall below. Let n be an integer $\geqslant 1$ and $r \leqslant n$. We say that linear forms $L_1 = \sum_{j=1}^n \alpha_{1j}X_j, \dots, L_r = \sum_{j=1}^n \alpha_{nj}X_j$ with coefficients in $\mathbb{C}$ are linearly dependent if there are $c_1, \dots, c_r \in \mathbb{C}$, not all 0, such that $c_1L_1 + \cdots + c_rL_r \equiv 0$. Otherwise, $L_1, \dots, L_r$ are called linearly independent. If $r = n$, then $L_1, \dots, L_n$ are linearly independent if and only if their coefficient determinant $\det(L_1, \dots, L_n) = \det(\alpha_{ij})_{1 \leqslant i,j \leqslant n} \neq 0$.

A linear subspace T of $\mathbb{Q}^n$ of dimension r can be described as

$$T = \left\{ \sum_{i=1}^r z_i\mathbf{a}_i \,:\, z_1, \dots, z_r \in \mathbb{Q} \right\},$$

where $\mathbf{a}_1, \dots, \mathbf{a}_r$ are linearly independent vectors from $\mathbb{Q}^n$. Alternatively, T can be described as

$$T = \{\mathbf{x} \in \mathbb{Q}^n \,:\, L_1(\mathbf{x}) = 0, \dots, L_{n-r}(\mathbf{x}) = 0\}$$

where $L_1, \dots, L_{n-r}$ are linearly independent linear forms in $X_1, \dots, X_n$ with coefficients in $\mathbb{Q}$.

The norm of $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{Z}^n$ is given by

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

Theorem 1.3. (Subspace Theorem, W.M. Schmidt, 1972). *Let $n \geqslant 2$, let*

$$L_i(\mathbf{X}) = \alpha_{i1}X_1 + \cdots + \alpha_{in}X_n \ (i = 1, \dots, n)$$

be n **linearly independent** *linear forms with* **algebraic** *coefficients in $\mathbb{C}$ and let $\delta > 0$. Then the set of solutions of the inequality*

$$|L_1(\mathbf{x}) \cdots L_n(\mathbf{x})| \leqslant \|\mathbf{x}\|^{-\delta} \ \text{in } \mathbf{x} \in \mathbb{Z}^n \tag{1.3}$$

is contained in a union $T_1 \cup \cdots \cup T_t$ of finitely many proper linear subspaces of $\mathbb{Q}^n$.

Remark. The proof of the Subspace Theorem is *ineffective*, i.e., it does not enable to determine the subspaces. There is however a quantitative version of the Subspace Theorem which gives an explicit upper bound for the number of subspaces. This is an important tool for estimating the number of solutions of various types of Diophantine equations.

We show that the Subspace Theorem implies Roth's Theorem.

Subspace Theorem $\Longrightarrow$ Roth's Theorem. Let (x, y) (with $y > 0$) be a pair of integers satisfying $|\alpha - x/y| \leqslant y^{-2-\delta}$. Multiplying with y^2 gives

$$|y(x - \alpha y)| \leqslant y^{-\delta}.$$

Since the linear forms Y and $X - \alpha Y$ are linearly independent, this is an inequality to which the Subspace Theorem is applicable, except that on the right hand side we have y instead of $\|\mathbf{x}\| = \max(|x|, |y|)$. However, we certainly have $|x/y| \leqslant |\alpha| + y^{-2} \leqslant |\alpha| + 1$. So $|x| \leqslant (|\alpha| + 1)y =: Cy$ and then also, $\|\mathbf{x}\| \leqslant Cy$. With this observation, our inequality becomes

$$|y(x - \alpha y)| \leqslant C^{\delta}\|\mathbf{x}\|^{-\delta},$$

and now we can apply the Subspace Theorem. We infer that the pairs (x, y) lie in only finitely many proper, i.e., one-dimensional, linear subspaces of $\mathbb{Q}^2$.

Pick one of these subspaces, say T. We show that T contains only finitely many solutions of (1.2). We have $T = \{\lambda(x_0, y_0) : \lambda \in \mathbb{Q}\}$ where for (x_0, y_0) we may take a pair of integers with $\gcd(x_0, y_0) = 1$. If $(x, y) \in T \cap \mathbb{Z}^2$ and satisfies (1.2), then $(x, y) = \lambda(x_0, y_0)$ with $\lambda \in \mathbb{Z}$, and

$$0 < \left|\alpha - \frac{x_0}{y_0}\right| = \left|\alpha - \frac{x}{y}\right| \leqslant |\lambda|^{-2} y_0^{-2},$$

implying that $|\lambda|$ is bounded in terms of T. So indeed T contains only finitely many solutions of (1.2). Roth's theorem easily follows. □

The Subspace Theorem states that the set of solutions of (1.3) is contained in a finite union of proper linear subspaces of $\mathbb{Q}^n$, but one may wonder whether (1.3) doesn't have only finitely many solutions. For instance, it may be that there is a non-zero $\mathbf{x}_0 \in \mathbb{Z}^n$ with $L_1(\mathbf{x}_0) = 0$. Then for every $d \in \mathbb{Z}$, the point $d\mathbf{x}_0$ is a solution to (1.3), and this gives infinitely many solutions to (1.3). To avoid such a construction, let us consider

$$0 < |L_1(\mathbf{x}) \cdots L_n(\mathbf{x})| \leqslant \|\mathbf{x}\|^{-\delta} \text{ in } \mathbf{x} \in \mathbb{Z}^n. \tag{1.4}$$

For instance, let $n = 2$. Then the solutions of (1.4) lie in finitely one-dimensional subspacss of $\mathbb{Q}^2$, and similarly as in the proof of Roth's Theorem, each of these subspaces contains only finitely many solutions. So altogether, if $n = 2$ then (1.4) has only finitely many solutions. We now give an example showing that (1.4) may have infinitely many solutions if $n \geqslant 3$.

Example. Let $0 < \delta < 1$ and consider the inequality

$$0 < |(x_1+\sqrt{2}x_2+\sqrt{3}x_3)(x_1-\sqrt{2}x_2+\sqrt{3}x_3)(x_1-\sqrt{2}x_2-\sqrt{3}x_3)| \leqslant \|\mathbf{x}\|^{-\delta} \tag{1.5}$$

to be solved in $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{Z}^3$. Notice that the three linear forms on the left-hand side are linearly independent.

Consider the triples of integers $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{Z}^3$ with $x_3 = 0, x_1x_2 \neq 0$. For these points, $\|\mathbf{x}\| = \max(|x_1|, |x_2|, 0)$. By Dirichlet's Theorem, the inequality

$$\left|\sqrt{2} - \frac{x_1}{x_2}\right| \leqslant |x_2|^{-2}$$

has infinitely many solutions $(x_1, x_2) \in \mathbb{Z}^2$ with $x_2 \neq 0$. For these solutions, $\|\mathbf{x}\|$ has the same order of magnitude as $|x_2|$. Indeed,

$$|x_1/x_2| \leqslant |x_2|^{-2} + \sqrt{2} \leqslant 1 + \sqrt{2},$$

and so, $\|\mathbf{x}\| = \max(|x_1|, |x_2|) \leqslant (1 + \sqrt{2})|x_2|$.

So for the points under consideration,

$$\begin{aligned}0 &< |(x_1 + \sqrt{2}x_2 + \sqrt{3}x_3)(x_1 - \sqrt{2}x_2 + \sqrt{3}x_3)(x_1 - \sqrt{2}x_2 - \sqrt{3}x_3)| \\ &= |(x_1 + \sqrt{2}x_2)(x_1 - \sqrt{2}x_2)^2| \\ &\leqslant (1 + \sqrt{2})\|\mathbf{x}\| \cdot (x_2^{-1})^2 \leqslant (1 + \sqrt{2})^3\|\mathbf{x}\|^{-1} \\ &\leqslant \|\mathbf{x}\|^{-\delta},\end{aligned}$$

provided $\|\mathbf{x}\|$ is sufficiently large. It follows that (1.5) has infinitely many solutions $\mathbf{x}$ lying in the subspace $x_3 = 0$. In a similar way, it can be shown that (1.5) has infinitely many solutions in the subspaces $x_1 = 0$, $x_2 = 0$, respectively (exercise).

By the Subspace Theorem, the solutions of (1.5) with $x_1x_2x_3 \neq 0$ lie in finitely many proper linear subspaces of $\mathbb{Q}^3$. With a more precise argument one shows that (1.5) has only finitely many solutions with $x_1x_2x_3 \neq 0$ (exercise).

We give a generalization of the Subspace Theorem which may be useful for certain applications.

We say that a system of $r \geqslant n$ linear forms $L_1, \dots, L_r$ in the variables $X_1, \dots, X_n$ is *in general position* if each n-tuple of linear forms among $L_1, \dots, L_r$ is linearly independent.

Theorem 1.4. *Let*

$$L_i(\mathbf{X}) = \alpha_{i1}X_1 + \cdots + \alpha_{in}X_n \ (i = 1, \dots, r,\ r \geqslant n)$$

be r linear forms with algebraic coefficients in $\mathbb{C}$ in general position and let $\delta > 0$. Then the set of solutions of the inequality

$$|L_1(\mathbf{x}) \cdots L_r(\mathbf{x})| \leqslant \|\mathbf{x}\|^{r-n-\delta} \ \text{in } \mathbf{x} \in \mathbb{Z}^n \tag{1.6}$$

is contained in a union $T_1 \cup \cdots \cup T_t$ of finitely many proper linear subspaces of $\mathbb{Q}^n$.

This can be deduced by combining the Subspace Theorem with the following lemma.

Lemma 1.5. *Let $M_1, \dots, M_n$ be linearly independent linear forms in $X_1, \dots, X_n$ with complex coefficients. Then there is a constant $C > 0$ such that*

$$\|\mathbf{x}\| \leqslant C \max\big(|M_1(\mathbf{x})|, \dots, |M_n(\mathbf{x})|\big) \ \text{for all } \mathbf{x} \in \mathbb{C}^n.$$

Proof. Since the linear forms $M_1, \dots, M_n$ are linearly independent, they span the complex vector space of all linear forms in $X_1, \dots, X_n$ with complex coefficients. So we can express $X_1, \dots, X_n$ as linear combinations of $M_1, \dots, M_n$, i.e.,

$$X_i = \sum_{j=1}^{n} \beta_{ij} M_j \text{ with } \beta_{ij} \in \mathbb{C} \ (i = 1, \dots, n).$$

Take $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{C}^n$ and put $M := \max_{1 \leqslant i \leqslant n} |M_i(\mathbf{x})|$. Then

$$\max_{1 \leqslant i \leqslant n} |x_i| \leqslant \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} |\beta_{ij}| \cdot |M_j(\mathbf{x})| \leqslant C \cdot M \ \text{ with } C := \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} |\beta_{ij}|.$$

□

Proof of Theorem 1.4. We partition the solutions $\mathbf{x}$ of (1.6) into a finite number of subsets according to the ordering of the numbers $|L_1(\mathbf{x})|, \dots, |L_r(\mathbf{x})|$, and show that each of these subsets lies in at most finitely many proper linear subspaces of $\mathbb{Q}^n$. Consider the solutions $\mathbf{x} \in \mathbb{Z}^n$ from one of these subsets, say for which

$$|L_1(\mathbf{x})| \leqslant \cdots \leqslant |L_r(\mathbf{x})|.$$

By Lemma 1.5, for $i = n+1, \dots, r$, there is a constant C_i such that for all solutions $\mathbf{x}$ under consideration,

$$\|\mathbf{x}\| \leqslant C_i |L_i(\mathbf{x})|,$$

since $L_1, \dots, L_{n-1}, L_i$ are linearly independent, and $|L_i(\mathbf{x})|$ has the largest value. Substituting this into (1.6) gives

$$\begin{aligned} |L_1(\mathbf{x}) \cdots L_n(\mathbf{x})| &\leqslant C\|\mathbf{x}\|^{r-n-\delta} \prod_{i=n+1}^{r} |L_i(\mathbf{x})|^{-1} \\ &\leqslant C \cdot (C_{n+1} \cdots C_r) \|\mathbf{x}\|^{-\delta}. \end{aligned}$$

So the solutions $\mathbf{x}$ under consideration lie in at most finitely many proper linear subspaces of $\mathbb{Q}^n$. □

We deduce a result on simultaneous approximation.

Theorem 1.6. *Let* $\alpha_1, \dots, \alpha_n$ *be algebraic numbers in* $\mathbb{C}$ *and let* $C > 0, \delta > 0$. *Then the inequality*

$$0 < |\alpha_1 x_1 + \cdots + \alpha_n x_n| \leqslant C\|\mathbf{x}\|^{1-n-\delta} \text{ in } \mathbf{x} = (x_1, \dots, x_n) \in \mathbb{Z}^n \tag{1.7}$$

has only finitely many solutions.

Remark. For $n = 2$ this result is equivalent to Roth's Theorem (exercise).

Proof. We proceed by induction on n. For $n = 1$ the assertion is obvious. (Here we use our assumption $\alpha_1 x_1 \neq 0$). Let $n > 1$ and suppose Theorem 1.6 is true for linear forms in fewer than n variables.

We apply the Subspace Theorem. We may assume that at least one of the coefficients $\alpha_1, \dots, \alpha_n$ is non-zero, otherwise there are no solutions. Suppose that $\alpha_1 \neq 0$. Then (1.7) implies

$$|(\alpha_1 x_1 + \cdots + \alpha_n x_n) x_2 \cdots x_n| \leqslant C\|\mathbf{x}\|^{-\delta}$$

and by the Subspace Theorem, the solutions of the latter lie in a union of finitely many proper linear subspaces $T_1, \dots, T_t$ of $\mathbb{Q}^n$. We consider only solutions with $\alpha_1 x_1 + \cdots + \alpha_n x_n \neq 0$. Therefore, without loss of generality we may assume that $\alpha_1 x_1 + \cdots + \alpha_n x_n$ is not identically 0 on any of the spaces $T_1, \dots, T_t$.

Consider the solutions of (1.7) in T_i. Choose a non-trivial linear form vanishing identically on T_i, $a_1 x_1 + \cdots + a_n x_n = 0$. Suppose for instance, that

$a_n \neq 0$. Then x_n can be expressed as a linear combination of $x_1, \dots, x_{n-1}$. By substituting this into (1.7) we obtain an inequality

$$0 < |\beta_1 x_1 + \cdots + \beta_{n-1}x_{n-1}| \leqslant C\|\mathbf{x}\|^{1-n-\delta} \leqslant C\big(\max_{1\leqslant i\leqslant n-1} |x_i|\big)^{2-n-\delta}.$$

By the induction hypothesis, the latter inequality has only finitely many solutions $(x_1, \dots, x_{n-1})$. So T_i contains only finitely many solutions $\mathbf{x}$ of (1.7). Applying this to $T_1, \dots, T_t$ we obtain that (1.7) has altogether only finitely many solutions. □

2. Galois theory

We have collected some facts from Galois theory which are needed later on.

Let K be an algebraic number field of degree r. We can express K as $\mathbb{Q}(\theta)$. Let $f(X)$ denotes the minimal polynomial of θ over $\mathbb{Q}$. Then f has precisely r zeros in $\mathbb{C}$, $\theta_1, \dots, \theta_r$, say. The field K has precisely r distinct embeddings in $\mathbb{C}$, $\sigma_1, \dots, \sigma_r$, say, which are determined by $\sigma_i(\theta) = \theta_i$ for $i = 1, \dots, r$.

Let G denote the *Galois closure* of K, that is $G = \mathbb{Q}(\theta_1, \dots, \theta_r)$. A $\mathbb{Q}$-automorphism of G is an isomorphism of G to itself, leaving the elements of $\mathbb{Q}$ unchanged. Clearly, the $\mathbb{Q}$-automorphisms of G form a group under composition, the *Galois group* of G, notation $\mathrm{Gal}(G/\mathbb{Q})$. The field G consists of all polynomials in $\theta_1, \dots, \theta_r$ with rational coefficients. Since a $\mathbb{Q}$-automorphism τ preserves addition and multiplication and leaves $\mathbb{Q}$ unchanged, it is uniquely determined by $\tau(\theta_1), \dots, \tau(\theta_r)$.

We recall some facts about the Galois group:

- The cardinality of $\mathrm{Gal}(G/\mathbb{Q})$ is equal to $[G : \mathbb{Q}]$.
- The set of $x \in G$ such that $\tau(x) = x$ for all $\tau \in \mathrm{Gal}(G/\mathbb{Q})$ is equal to $\mathbb{Q}$.
- Let $\tau \in \mathrm{Gal}(G/\mathbb{Q})$. Then $\tau(\theta_1), \dots, \tau(\theta_r)$ is a permutation of $\theta_1, \dots, \theta_r$ and τ is uniquely determined by this permutation. Thus, $\mathrm{Gal}(G/\mathbb{Q})$ may be viewed as a subgroup of the permutation group on r elements S_r.
- Let $\tau \in \mathrm{Gal}(G/\mathbb{Q})$. Then $\tau \circ \sigma_1, \dots, \tau \circ \sigma_r$ is a permutation of the embeddings $\sigma_1, \dots, \sigma_r$ of K in $\mathbb{C}$.

Example. Let $K = \mathbb{Q}(\sqrt[3]{2})$. Then $G = \mathbb{Q}(\sqrt[3]{2}, \omega\sqrt[3]{2}, \omega^{-1}\sqrt[3]{2}) = \mathbb{Q}(\sqrt[3]{2}, \omega)$, where ω is a primitive cube root of unity. The field $\mathbb{Q}(\sqrt[3]{2})$ has degree 3, and

ω has degree 2 over $\mathbb{Q}(\sqrt[3]{2})$. Hence G has degree 6. An automorphism of G is determined by its images of $\sqrt[3]{2}$ and ω. Its image of $\sqrt[3]{2}$ must be another root of X^3-2, that is, one of $\sqrt[3]{2}, \omega\sqrt[3]{2}, \omega^{-1}\sqrt[3]{2}$, and its image of ω must be another root of X^2+X+1, that is, one of ω, ω^{-1}. Thus, G has at most 6 automorphisms. So in fact, G has precisely 6 automorphisms, given by the permutations of $\sqrt[3]{2}, \omega\sqrt[3]{2}, \omega^{-1}\sqrt[3]{2}$. Hence $\mathrm{Gal}(G/\mathbb{Q}) \cong S_3$.

Example. Let $K=\mathbb{Q}(\sqrt[4]{2})$. Then $G=\mathbb{Q}(\sqrt[4]{2}, i\sqrt[4]{2}, -\sqrt[4]{2}, -i\sqrt[4]{2}))=\mathbb{Q}(\sqrt[4]{2}, i)$. The field G has degree 8, since $\sqrt[4]{2}$ has degree 4 over $\mathbb{Q}$, and i degree 2 over $\mathbb{Q}(\sqrt[4]{2})$. An automorphism of G is determined by its images of $\sqrt[4]{2}$ and i. Its image of $\sqrt[4]{2}$ is one of $\sqrt[4]{2}, i\sqrt[4]{2}, -\sqrt[4]{2}, -i\sqrt[4]{2}$ and its image of i is one of $i, -i$. So G has at most 8, and hence precisely 8, automorphisms. Thus, the Galois group $\mathrm{Gal}(G/\mathbb{Q})$ is a subgroup of order 8 of S_4. We obtain the action of $\mathrm{Gal}(G/\mathbb{Q})$ on $\{\sqrt[4]{2}, i\sqrt[4]{2}, -\sqrt[4]{2}, -i\sqrt[4]{2}\}$ by labelling the vertices of a square by $\sqrt[4]{2}, i\sqrt[4]{2}, -\sqrt[4]{2}, -i\sqrt[4]{2}$, respectively, and applying the symmetries of the square.

3. Norm form equations

Recall that if $F(X,Y)$ is an irreducible binary form in $\mathbb{Q}[X,Y]$ of degree r with coefficient of X^r equal to 1, say, then

$$F(X,Y)=\prod_{i=1}^{r}(X-\alpha_i Y)$$

where $\alpha_1,\dots,\alpha_r$ are the conjugates of an algebraic number α. If $K=\mathbb{Q}(\alpha)$, and $\sigma_1,\dots,\sigma_r$ are the embeddings of K in $\mathbb{C}$, with $\sigma_i(\alpha)=\alpha_i$, then

$$F(X,Y)=\prod_{i=1}^{r}(X-\sigma_i(\alpha)Y)=N_{K/\mathbb{Q}}(X-\alpha Y).$$

That is, F is a *norm form* in two variables. The Thue equation

$$N_{K/\mathbb{Q}}(x-\alpha y)=c \quad \text{in } x,y\in\mathbb{Z}$$

has only finitely many solutions if $[K:\mathbb{Q}]\geqslant 3$ (by Thue's Theorem) or if K is an imaginary quadratic field (then the solutions represent points with integer coordinates on an ellipsis). It may have infinitely many solutions if K is real quadratic. For instance if $K=\mathbb{Q}(\sqrt{d})$ with d a positive, non-square integer, then the Pell equation $x^2-dy^2=N_{K/\mathbb{Q}}(x-\sqrt{d}y)=1$ has infinitely many solutions.

We consider a generalization of the Thue equation, with norm forms of an arbitrary number of variables. Let $K = \mathbb{Q}(\theta)$ be an algebraic number field of degree r, $\theta_1, \ldots, \theta_r \in \mathbb{C}$ the roots of the minimal polynomial of θ, and $\sigma_1, \ldots, \sigma_r$ the embeddings of K in $\mathbb{C}$, determined by $\sigma_i(\theta) = \theta_i$ for $i = 1, \ldots, r$. Further, let $G := \mathbb{Q}(\theta_1, \ldots, \theta_r)$.

Now suppose that $r \geqslant n \geqslant 2$ and let $\alpha_1, \ldots, \alpha_n$ be elements of K which are linearly independent over $\mathbb{Q}$, that is, the only solution in $x_1, \ldots, x_n \in \mathbb{Q}$ of $x_1\alpha_1 + \cdots + x_n\alpha_n = 0$ is $x_1 = \cdots = x_n = 0$. Define the polynomial

$$F(X_1, \ldots, X_n) := N_{K/\mathbb{Q}}(\alpha_1 X_1 + \cdots + \alpha_n X_n) := \prod_{i=1}^{r} (\sigma_i(\alpha_1) X_1 + \cdots + \sigma_i(\alpha_n) X_n).$$

Notice that if we apply any τ from the Galois group $\mathrm{Gal}(G/\mathbb{Q})$, then it permutes the linear factors of F, hence it leaves the coefficients of F unchanged. So F has its coefficients in $\mathbb{Q}$.

We deal with the so-called *norm form equation*

$$N_{K/\mathbb{Q}}(\alpha_1 X_1 + \cdots + \alpha_n X_n) = c \quad \text{in } \mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}^n \tag{3.1}$$

If $n = 2$, the left-hand side is a binary form and (3.1) becomes a Thue equation.

In 1972, Schmidt gave a necessary and sufficient condition such that (3.1) has only finitely many solutions. His proof was based on the Subspace Theorem. Here, we prove a special case of his result.

Theorem 3.1. *Suppose that $n < r$, and let $\alpha_1, \ldots, \alpha_n$ be elements of K which are linearly independent over $\mathbb{Q}$. Assume that* $\mathrm{Gal}(G/\mathbb{Q}) \cong S_r$. *Then* (3.1) *has only finitely many solutions.*

Lemma 3.2. *Let $L_i := \sigma_i(\alpha_1) X_1 + \cdots + \sigma_i(\alpha_n) X_n$ for $i = 1, \ldots, n$. Then the linear forms $L_1, \ldots, L_r$ are in general position.*

Proof. Since $\alpha_1, \ldots, \alpha_n$ are linearly independent over $\mathbb{Q}$, we can add $r - n$ new elements to make a basis of K over $\mathbb{Q}$, say $\alpha_1, \ldots, \alpha_r$. Then the matrix

$$\begin{pmatrix} \sigma_1(\alpha_1) & \cdots & \sigma_1(\alpha_r) \\ \vdots & & \vdots \\ \sigma_r(\alpha_1) & \cdots & \sigma_r(\alpha_r) \end{pmatrix}$$

is non-singular; in fact the square of its determinant is the discriminant of a basis of K over $\mathbb{Q}$, hence non-zero. This means that the matrix consisting of the first n columns,

$$\begin{pmatrix} \sigma_1(\alpha_1) & \cdots & \sigma_1(\alpha_n) \\ \vdots & & \vdots \\ \sigma_r(\alpha_1) & \cdots & \sigma_r(\alpha_n) \end{pmatrix}$$

must have column rank n. Then the row rank of this matrix is also n, which implies that this matrix has n linearly independent rows. Assume without loss of generality that the first n rows are linearly independent. This means precisely that the linear forms $L_1, \dots, L_n$ are linearly independent, i.e., $\det(L_1, \dots, L_n) \neq 0$. We have to show that also $L_{i_1}, \dots, L_{i_n}$ are linearly independent, for any n distinct indices $i_1, \dots, i_n$ from $1, \dots, r$.

Since $\mathrm{Gal}(G/\mathbb{Q}) \cong S_r$, there is τ in this Galois group, such that $\tau(\theta_1) = \theta_{i_1}, \dots, \tau(\theta_n) = \theta_{i_n}$. This implies $\tau \circ \sigma_1 = \sigma_{i_1}, \dots, \tau \circ \sigma_n = \sigma_{i_n}$. As a consequence τ maps the coefficients of L_j to those of L_{i_j} for $j = 1, \dots, n$. It follows that

$$\det(L_{i_1}, \dots, L_{i_n}) = \tau(\det(L_1, \dots, L_n)) \neq 0,$$

hence $L_{i_1}, \dots, L_{i_n}$ are linearly independent. This proves our lemma. □

Proof of Theorem 3.1. We proceed by induction on n. First let $n = 1$. Then equation (3.1) becomes

$$N_{K/\mathbb{Q}}(\alpha_1 x_1) = N_{K/\mathbb{Q}}(\alpha) x_1^r = c,$$

and this clearly has only finitely many solutions.

Next, let $n \geqslant 2$, and assume the theorem is true for norm form equations in fewer than n unknowns. Since the linear forms $L_1, \dots, L_r$ are in general position, by Theorem 1.4, for any $C > 0, \delta > 0$ the set of solutions of

$$|F(\mathbf{x})| = |L_1(\mathbf{x}) \cdots L_r(\mathbf{x})| \leqslant C \|\mathbf{x}\|^{r-n-\delta}$$

lies in a union of finitely many proper linear subspaces of $\mathbb{Q}^n$. It follows that the solutions of (3.1) lie in only finitely many proper linear subspaces of $\mathbb{Q}^n$.

We show that (3.1) has only finitely many solutions in each of these subspaces. Let T be one of these subspaces. For solutions in T, one of the coordinates, can be expressed as a linear combination of the others. Say that we have $x_n = a_1 x_1 + \cdots + a_{n-1} x_{n-1}$ identically on T. By substituting this in (3.1) we get a norm form equation in $n-1$ variables

$$N_{K/\mathbb{Q}}(\beta_1 x_1 + \cdots + \beta_{n-1} x_{n-1}) = c,$$

where $\beta_i = \alpha_i + a_i\alpha_n$ for $i = 1, \dots, n-1$. Now $\beta_1, \dots, \beta_{n-1}$ are linearly independent over $\mathbb{Q}$. Hence by the induction hypothesis, this last equation has only finitely many solutions $(x_1, \dots, x_{n-1}) \in \mathbb{Z}^{n-1}$. This implies that the original equation (3.1) has only finitely many solutions $(x_1, \dots, x_n) \in T$. This completes our proof. □

We give examples of norm form equations with infinitely many solutions. We use the following fact:

Lemma 3.3. *Let K be an algebraic number field and α an element of the ring of integers $\mathcal{O}_K$ of K. Then*

$$\alpha \text{ is a unit of } \mathcal{O}_K \Longleftrightarrow N_{K/\mathbb{Q}}(\alpha) = \pm 1.$$

Proof. Well-known. □

It is more convenient to rewrite (3.1) as

$$N_{K/\mathbb{Q}}(\xi) = c \quad \text{in } \xi \in \mathcal{M}, \tag{3.2}$$

where

$$\mathcal{M} := \{\alpha_1 x_1 + \cdots + \alpha_n x_n : x_1, \dots, x_n \in \mathbb{Z}\}.$$

Notice that $\mathcal{M}$ is a free $\mathbb{Z}$-module in K of rank n, i.e., its elements can be expressed uniquely as $\mathbb{Z}$-linear combinations of a basis of n elements.

Take an algebraic number field K such that the unit group $\mathcal{O}_K^*$ of the ring of integers of K is infinite. This means precisely that K is not $\mathbb{Q}$ or an imaginary quadratic field. Take $\mathcal{M} = \mathcal{O}_K$. Recall that $\mathcal{O}_K$ is a free $\mathbb{Z}$-module of rank equal to $[K : \mathbb{Q}]$. Now clearly, if $\varepsilon \in \mathcal{O}_K^*$, then $\xi = \varepsilon^2$ is a solution to

$$N_{K/\mathbb{Q}}(\xi) = 1 \quad \text{in } \xi \in \mathcal{O}_K,$$

and so this last norm form equation has infinitely many solutions.

More generally, (3.2) has infinitely many solutions if $\mathcal{M}$ contains $\mu\mathcal{O}_L = \{\mu\xi : \xi \in \mathcal{O}_L\}$, where $\mu \in K^*$, and L is a subfield of K which is not equal to $\mathbb{Q}$ or to an imaginary quadratic field. Now Schmidt's result on norm form equations is as follows.

Theorem 3.4. (W.M. Schmidt, 1972) *Let K be an algebraic number field, $\alpha_1, \dots, \alpha_n$ elements of K which are linearly independent over $\mathbb{Q}$, and $\mathcal{M} := \{\sum_{i=1}^n \alpha_i x_i : x_i \in \mathbb{Z}\}$. Then the following two assertions are equivalent:*

(i) there do not exist $\mu \in K^*$ and a subfield L of K not equal to $\mathbb{Q}$ or to an

imaginary quadratic field such that $\mu\mathcal{O}_L \subseteq \mathcal{M}$;

(ii) for every $c \in \mathbb{Q}^*$, the equation

$$N_{K/\mathbb{Q}}(\xi) = c \quad \textit{in } \xi \in \mathcal{M} \tag{3.2}$$

has only finitely many solutions.

The implication (i)$\Longrightarrow$(ii) is deduced from the Subspace Theorem. The proof is too difficult to be included here. We prove only the other implication, that is, if (i) is false then (3.2) has infinitely many solutions. Indeed, for every $\varepsilon \in \mathcal{O}_L^*$ we have $\mu\varepsilon^2 \in \mathcal{M}$ and $N_{K/\mathbb{Q}}(\varepsilon) = \pm 1$. Thus, by letting ε run through $\mathcal{O}_L^*$, we obtain infinitely many elements $\xi = \mu\varepsilon^2 \in \mathcal{M}$ with

$$N_{K/\mathbb{Q}}(\xi) = N_{K/\mathbb{Q}}(\mu)N_{K/\mathbb{Q}}(\varepsilon)^2 = N_{K/\mathbb{Q}}(\mu).$$

□

Example. Let $K = \mathbb{Q}(\sqrt[6]{2})$, $\mathcal{M} := \{x_1\sqrt[6]{2} + x_2\sqrt{2} + x_3\sqrt[6]{2}^5 : x_1, x_2, x_3 \in \mathbb{Z}\}$. Notice that K contains the subfield $L = \mathbb{Q}(\sqrt[3]{2})$ and that

$$\mathcal{M} = \sqrt[6]{2}\{x_1 + x_2\sqrt[3]{2} + x_3\sqrt[3]{2}^2\} = \sqrt[6]{2}\mathcal{O}_L.$$

Now $\mathcal{O}_L^* = \{\pm(1-\sqrt[3]{2})^n : n \in \mathbb{Z}\}$, and $N_{K/\mathbb{Q}}(1-\sqrt[3]{2}) = 1$. Hence every $n \in \mathbb{Z}$ yields a solution $\xi := \sqrt[6]{2}(1-\sqrt[3]{2})^n \in \mathcal{M}$ of

$$N_{K/\mathbb{Q}}(\xi) = N_{K/\mathbb{Q}}(\sqrt[6]{2}) = 2.$$

4. Exercises

Exercise 1. Prove that the following three assertions are equivalent:

(a) Let α be an algebraic number from $\mathbb{C}$ with $\alpha \notin \mathbb{Q}$. Then for every $\delta > 0$ there are only finitely many pairs of integers (x, y) such that

$$\left|\alpha - \frac{x}{y}\right| \leqslant |y|^{-2-\delta}, \quad y \neq 0.$$

(b) Let α, β be two non-zero algebraic numbers from $\mathbb{C}$. Then for every $C > 0, \delta > 0$ there are only finitely many pairs of integers (x, y) such that

$$0 < |\alpha x + \beta y| \leqslant C\max(|x|, |y|)^{-1-\delta}.$$

(c) Let $\alpha_1, \beta_1, \alpha_2, \beta_2$ be non-zero algebraic numbers in $\mathbb{C}$ such that $\alpha_1\beta_2 - \alpha_2\beta_1 \neq 0$. Then for every $C > 0, \delta > 0$, there are only finitely many pairs of integers (x, y) such that

$$0 < |(\alpha_1 x + \beta_1 y)(\alpha_2 x + \beta_2 y)| \leqslant C \max(|x|, |y|)^{-\delta}.$$

Exercise 2. Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be a binary form.

(a) Prove that $F(X,Y) = \prod_{i=1}^{r}(\alpha_i X - \beta_i Y)$ where α_i, β_i are algebraic numbers from $\mathbb{C}$.

(b) Assume that F has degree $r \geqslant 3$, and that F is not divisible by $(\alpha X - \beta Y)^2$ for some $\alpha, \beta \in \mathbb{C}$. Further, let $C > 0, \delta > 0$. Prove that there are only finitely many pairs of integers (x, y) such that

$$|F(x,y)| \leqslant C \max(|x|, |y|)^{r-2-\delta}, \quad F(x,y) \neq 0.$$

(c) Let $G(X,Y)$ be a non-zero polynomial in $\mathbb{Z}[X,Y]$ of total degree at most $r-3$, i.e., $G(X,Y) = \sum_{i,j} a_{ij} X^i Y^j$, where the sum is over pairs of indices i, j with $i \geqslant 0, j \geqslant 0$ and $i + j \leqslant r - 3$. Prove that the equation

$$F(x,y) = G(x,y)$$

has at most finitely many solutions $x, y \in \mathbb{Z}$ with $F(x,y) \neq 0$.

Exercise 3. Let $0 < \delta < 1$ and consider the inequality

$$(4.1)\quad 0 < |(x_1+\sqrt{2}x_2+\sqrt{3}x_3)(x_1-\sqrt{2}x_2+\sqrt{3}x_3)(x_1-\sqrt{2}x_2-\sqrt{3}x_3)| \leqslant \|\mathbf{x}\|^{-\delta}$$

to be solved in $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{Z}^3$.

(a) Prove that the subspaces $x_1 = 0$, $x_2 = 0$ contain infinitely many solutions of (4.1).

(b) Let T be a one-dimensional linear subspace of $\mathbb{Q}^3$. Prove that T contains at most finitely many solutions of (4.1).

(c) Let $T = \{\mathbf{x} \in \mathbb{Q}^3 : a_1x_1 + a_2x_2 + a_3x_3 = 0\}$, where $a_1, a_2, a_3 \in \mathbb{Q}$ and at least two among a_1, a_2, a_3 are non-zero (i.e., T is not equal to one of the subspaces $x_1 = 0$, $x_2 = 0$, $x_3 = 0$). Consider the solutions $\mathbf{x} = (x_1, x_2, x_3)$ of (4.1) in T and eliminate one of the variables x_1, x_2, x_3 by expressing it as a linear combination of the two others. Prove that after this elimination, the linear forms L_1, L_2, L_3 become a system of linear forms in two variables which is in general position.

Distinguish between the cases $a_3 = 0$ giving $x_2 = -(a_1/a_2)x_1$, and $a_3 \neq 0$ giving $x_3 = -(a_1/a_3)x_1 - (a_2/a_3)x_2$.

(d) Prove that (4.1) has only finitely many solutions with $x_1x_2x_3 \neq 0$.

Remark. In 1989, Vojta proved the following refinement of the Subspace Theorem. Let again $L_1, \dots, L_n$ be n linearly independent linear forms with algebraic coefficients in $\mathbb{C}$ and $\delta > 0$. Then there exist a finite collection $S_1, \dots, S_m$ of proper linear subspaces of $\mathbb{Q}^n$, which is effectively determinable and which is independent of δ, and a finite set F_δ which depends on δ, such that the set of solutions of

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

is contained in $S_1 \cup \cdots \cup S_m \cup F_\delta$.

In example (4.1) one can take $S_1 = \{x_1 = 0\}$, $S_2 = \{x_2 = 0\}$, $S_3 = \{x_3 = 0\}$.

Exercise 4. Let $L_1 = \alpha_1X_1 + \cdots + \alpha_nX_n$, $L_2 = \beta_1X_1 + \cdots + \beta_nX_n$ be two linearly independent linear forms with algebraic coefficients from $\mathbb{C}$. Let $\delta > 0$. Prove that the system of inequalities

$$0 < |L_1(\mathbf{x})| \leqslant \|\mathbf{x}\|^{1-n}, \ 0 < |L_2(\mathbf{x})| \leqslant \|\mathbf{x}\|^{1-\delta} \text{ in } \mathbf{x} \in \mathbb{Z}^n$$

has only finitely many solutions.

Exercise 5. In this exercise you are asked to prove another generalization of Roth's Theorem. Let $\alpha_1, \dots, \alpha_n$ be real algebraic numbers such that $1, \alpha_1, \dots, \alpha_n$ are linearly independent over $\mathbb{Q}$ and let $\delta > 0$. Consider the system of inequalities

$$(4.2) \qquad \left|\alpha_1 - \frac{x_1}{x_{n+1}}\right| \leqslant \|\mathbf{x}\|^{-1-\frac{1}{n}-\delta}, \dots, \left|\alpha_n - \frac{x_n}{x_{n+1}}\right| \leqslant \|\mathbf{x}\|^{-1-\frac{1}{n}-\delta}$$

to be solved simultaneously in $\mathbf{x} = (x_1, \dots, x_{n+1}) \in \mathbb{Z}^{n+1}$. Prove that (4.2) has only finitely many solutions (for $n = 1$ this gives back Roth's Theorem). To this end, work out the following steps.

(a) Prove that the set of solutions of (4.2) lies in only finitely many proper linear subspaces of $\mathbb{Q}^{n+1}$.

(b) Let T be a proper linear subspace of $\mathbb{Q}^{n+1}$. and let $b_1, \dots, b_{n+1}$ be integers, not all equal to 0, such that $b_1x_1 + \cdots + b_nx_n + b_{n+1}x_{n+1} = 0$ is identically 0 on T. Assume that T contains infinitely many solutions of (4.2). Prove that $b_1\alpha_1 + \cdots + b_n\alpha_n + b_{n+1} = 0$.

Hint. If $\mathbf{x}$ is a solution in T of (4.2), then α_i is very close to x_i/x_{n+1} for $i = 1, \dots, n$.

Exercise 6. Let $K = \mathbb{Q}(\sqrt[5]{2})$. Thus, there are precisely five embeddings $\sigma_1, \dots, \sigma_5 : K \to \mathbb{C}$, given by $\sigma_i(\sqrt[5]{2}) = \rho^{i-1}\sqrt[5]{2}$, where ρ is a primitive 5-th root of unity.

Consider the Diophantine equation

$$
\begin{aligned}
(4.3)\quad & N_{K/\mathbb{Q}}\big(x_1 + x_2\sqrt[5]{2} + x_3(\sqrt[5]{2})^2\big) \\
& \qquad = \prod_{i=1}^{5} \big(x_1 + x_2\sigma_i(\sqrt[5]{2}) + x_3\sigma_i(\sqrt[5]{2})^2\big) = 1 \text{ in } \mathbf{x} = (x_1, x_2, x_3) \in \mathbb{Z}^3 .
\end{aligned}
$$

(a) Prove that the left-hand side of (4.3) is a product of linear forms in general position (think of Vandermonde determinants).

(b) Let $\alpha, \beta \in K$ with $\beta \neq 0$ and $\alpha/\beta \notin \mathbb{Q}$. Prove that the linear forms $\sigma_i(\alpha)X + \sigma_i(\beta)Y$ $(i = 1, \dots, 5)$ are in general position.

(c) Prove that (4.3) has only finitely many solutions. It is not allowed to use Theorem 3.4 here.