On the norm form inequality $|F(\mathrm{x})| \leq h$.

Jan-Hendrik Evertse [1)]

To Professor Kalman Győry on his 60-th birthday.

Abstract. Let $F \in \mathbb{Z}[X_1, \ldots, X_n]$ be a non-degenerate norm form of degree r. In his paper [17] from 1990, Schmidt conjectured that for the number $Z_F(h)$ of solutions of the inequality $|F(\mathbf{x})| \leq h$ in $\mathbf{x} \in \mathbb{Z}^n$ one has $Z_F(h) \leq c(n,r)h^{n/r}$, with $c(n,r)$ depending on n and r only. In this paper, we show that $Z_F(h) \leq (16r)^{\frac{1}{3}(n+7)^3} h^{(n+\sum_{m=2}^{n-1} \frac{1}{m})/r} (1+\log h)^{\frac{1}{2}n(n-1)}$.

1991 Mathematics Subject Classification: 11D57

1. Introduction.

We start with recalling some results about inequalities as in the title in two variables, i.e., Thue inequalities

$$|F(x,y)| \leq h \quad \text{in } x, y \in \mathbb{Z} \tag{1.1}$$

where $F(X,Y) = a_r X^r + a_{r-1} X^{r-1} Y + \cdots + a_0 Y^r \in \mathbb{Z}[X,Y]$ is a binary form which is irreducible over $\mathbb{Q}$. Assume that F has degree $r \geq 3$. In 1933, Mahler [10] showed that for the number $Z_F(h)$ of solutions of (1.1) one has

$$Z_F(h) = C_F \cdot h^{2/r} + O(h^{1/(r-1)}) \ \text{ as } h \to \infty \qquad \text{with } C_F = \iint\limits_{|F(x,y)| \leq 1} dxdy$$

where the constant implied by the O-symbol depends on F. Note that the main term $C_F \cdot h^{2/r}$ is just the area of the region $\{(x,y) \in \mathbb{R}^2 : |F(x,y)| \leq h\}$.

---

1) The author has done part of the research for this paper while visiting the Institute for Advanced Study in Princeton during the fall of 1997. The author is very grateful to the IAS for its hospitality.

In 1987, Bombieri and Schmidt [3] showed that the Thue equation $|F(x,y)| = 1$ has only $\ll r$ solutions in $x, y \in \mathbb{Z}$ (where here and below constants implied by $\ll$ are absolute) and that the dependence on r is best possible. Also in 1987, Schmidt [15] proved more generally that $Z_F(h) \ll rh^{2/r}(1 + \frac{1}{r}\log h)$ for every $h \geq 1$ and he conjectured that the $\log h$-factor is unnecessary. Thunder [18], [19] showed that $Z_F(h) \ll rh^{2/r}$ if $\log\log m > r^9$ and $Z_F(h) \ll (r^{10}/\log r)h^{2/r}$ otherwise.

In 1993, Bean [1] showed that for every binary form $F \in \mathbb{Z}[X,Y]$ of degree $r \geq 3$ one has $C_F < 16$. In certain special cases, Mueller and Schmidt [11] and Thunder [18], [19] obtained explicit estimates $|Z_F(h) - C_F h^{2/r}| \leq c(r)h^{d(r)}$ with $c(r)$ and $d(r)$ depending only on r and $d(r) < 2/r$. Recently, Thunder [20] showed that for a binary cubic form $F \in \mathbb{Z}[X,Y]$ of discriminant $D(F)$ which is irreducible over $\mathbb{Q}$, one has $|Z_F(h) - C_F h^{2/3}| < 9 + 2008h^{1/2}|D(F)|^{-1/12} + 3156h^{1/3}$.

Now let F be a norm form of degree r in $n \geq 3$ variables, that is,

$$F = cN_{K/\mathbb{Q}}(\alpha_1 X_1 + \cdots + \alpha_n X_n) = c\prod_{i=1}^{r}(\alpha_1^{(i)}X_1 + \cdots + \alpha_n^{(i)}X_n) , \tag{1.2}$$

where $K = \mathbb{Q}(\alpha_1, \ldots, \alpha_n)$ is a number field of degree r, $\alpha \mapsto \alpha^{(i)}$ $(i = 1, \ldots, r)$ denote the isomorphic embeddings of K into $\mathbb{C}$, and c is a non-zero rational number such that F has its coefficients in $\mathbb{Z}$. To F we associate the $\mathbb{Q}$-vector space

$$V := \{\alpha_1 x_1 + \cdots + \alpha_n x_n : \; x_1, \ldots, x_n \in \mathbb{Q}\}. \tag{1.3}$$

For each subfield J of K we define the linear subspace of V,

$$V^J := \{\xi \in V : \xi\lambda \in V \;\text{ for every } \lambda \in J\} . \tag{1.4}$$

It is easy to see that $\xi\lambda \in V^J$ for $\xi \in V^J$, $\lambda \in J$, so V^J is the largest subspace of V closed under multiplication by elements from J. The norm form F is said to be *non-degenerate* if $\alpha_1, \ldots, \alpha_n$ are linearly independent over $\mathbb{Q}$ and if $V^J = (0)$ for each subfield J of K which is not equal to $\mathbb{Q}$ or to an imaginary quadratic field. It is easy to show that this notion of non-degeneracy does not depend on the choice of $c, \alpha_1, \ldots, \alpha_n$ in (1.2).

Denote by $Z_F(h)$ the number of solutions of the norm form inequality

$$|F(\mathbf{x})| \leq h \;\text{ in } \mathbf{x} \in \mathbb{Z}^n , \tag{1.5}$$

where $h > 0$. Schmidt's famous result on norm form equations from 1971 ([14], Satz 2, p.5) can be rephrased as follows:

$$Z_F(h) \text{ is finite for every } h > 0 \Longleftrightarrow F \text{ is non-degenerate.}$$

In view of Mahler's result one expects that for arbitrary non-degenerate norm forms F there is an asymptotic formula

$$Z_F(h) = C_F \cdot h^{n/r} + O(h^{d(n,r)}) \quad \text{as } h \to \infty \tag{1.6}$$

where C_F is the volume of the region $\{\mathbf{x} \in \mathbb{R}^n : |F(\mathbf{x})| \leq 1\}$ and where $d(n,r) < n/r$. By a result of Bean and Thunder [2] we have $C_F \leq n^{cn}$ for some absolute constant c. Note that the main term is precisely the volume of the region $\{\mathbf{x} \in \mathbb{R}^n : |F(\mathbf{x})| \leq h\}$.

As yet, only for norm forms from a restricted class such an asymptotic formula has been derived. In 1969, Ramachandra [12] proved that for norm forms F of the special shape $F = cN_{K/\mathbb{Q}}(X_1 + \alpha X_2 + \alpha^2 X_3 + \cdots + \alpha^{n-1} X_n)$, where $K = \mathbb{Q}(\alpha)$ is a number field of degree r and $r \geq 8n^6$, one has an asymptotic formula (1.6) with $(n-1)/(r-n+2) < d(n,r) < n/r$. This was generalised recently by de Jong [9], who showed that there is an asymptotic formula (1.6) for norm forms F as in (1.2) satisfying the following three conditions: a) each n-tuple among the linear factors of F is linearly independent; b) the Galois group of the normal closure of K over $\mathbb{Q}$ acts $n-1$ times transitively on the set of conjugates $\{\alpha^{(1)}, \ldots, \alpha^{(r)}\}$ of $\alpha \in K$; c) $r \geq 2n^{5/3}$. In the results of Ramachandra and de Jong, the constant in the error term depends on F and is ineffective. Thunder [21] obtained a formula (1.6) for norm forms F in $n = 3$ variables satisfying de Jong's conditions a) and b) and no further restriction on r with an effective error term depending on F. For arbitrary norm forms F in $n \geq 4$ variables, Thunder [21] could show only that the set of solutions of (1.5) can be divided into two sets, S_1 and S_2, say, where for the cardinality of S_1 we have an effective asymptotic formula like the right-hand side of (1.6) and where the set S_2 lies in the union of not more than $c(F)(1 + \log h)^{n-1}$ proper linear subspaces of $\mathbb{Q}^n$.

In this paper we do not consider the problem to derive an asymptotic formula such as (1.6) but instead to derive an explicit upper bound for $Z_F(h)$. In 1989, Schmidt [17] showed that for arbitrary non-degenerate norm forms F of degree r in n variables, the number $Z_F(1)$ of solutions of $|F(\mathbf{x})| = 1$ in $\mathbf{x} \in \mathbb{Z}^n$ is at most $\min\left(r^{2^{30n}r^2}, r^{(2n)^{n^{n+4}}}\right)$. This was improved by the author [6] to $\left(2^{33}r^2\right)^{n^3}$. Also in his paper [17], Schmidt conjectured that in general one has $Z_F(h) \le c(n,r)h^{n/r}$ where $c(n,r)$ depends only on n and r.

What we can prove is much less. Our main result is as follows:

Theorem 1. *Let $F \in \mathbb{Z}[X_1,\dots,X_n]$ be a non-degenerate norm form of degree r in $n \ge 2$ variables and let $h \ge 1$. Then for the number of solutions $Z_F(h)$ of $|F(\mathbf{x})| \le h$ in $\mathbf{x} \in \mathbb{Z}^n$ one has*

$$Z_F(h) \le (16r)^{\frac{1}{3}(n+7)^3} \cdot h^{(n+\sum_{m=2}^{n-1}\frac{1}{m})/r} \cdot (1+\log h)^{\frac{1}{2}n(n-1)} . \tag{1.7}$$

Except for Ramachandra's, all results on norm form equations mentioned above use Schmidt's Subspace theorem in a qualitative or quantitiative form; in particular, the results giving explicit upper bounds for $Z_F(1)$ use Schmidt's quantitative Subspace Theorem from 1989 [16] or improvements of the latter. In our proof of Theorem 1 we use a recent quantitative version of the Subspace Theorem due to Schlickewei and the author [8]. In fact, using this we first compute an upper bound for the number of proper linear subspaces of $\mathbb{Q}^n$ containing the set of solutions of (1.5) (cf. Theorem 2 below) and then obtain Theorem 1 by induction on n.

We introduce some notation used in the statement of Theorem 2. For a linear form $L = \alpha_1X_1 + \cdots + \alpha_nX_n$ with complex coefficients, we write $\overline{L} := \overline{\alpha}_1X_1 + \cdots + \overline{\alpha}_nX_n$ where $\overline{\alpha}$ denotes the complex conjugate of $\alpha \in \mathbb{C}$. Let F be the norm form given by (1.2). We assume henceforth that the isomorphic embeddings of K into $\mathbb{C}$ are so ordered that $\alpha \mapsto \alpha^{(i)}$ $(i = 1,\dots,r_1)$ map K into $\mathbb{R}$ and that $\alpha^{(i+r_2)} = \overline{\alpha^{(i)}}$ for $i = r_1+1,\dots,r_1+r_2$, where $r_1 + 2r_2 = r$. There are linear forms $L_1,\dots,L_r$ in n variables such that

$$\left.\begin{array}{l} F = \pm L_1\cdots L_r, \\ L_1,\dots,L_{r_1} \text{ have real coefficients,} \\ L_{i+r_2} = \overline{L}_i \text{ for } i = r_1+1,\dots,r_1+r_2 \end{array}\right\} \tag{1.8}$$

(for instance, one may take $L_i = \sqrt[r]{|c|}\cdot(\alpha_1^{(i)}X_1 + \cdots + \alpha_n^{(i)}X_n)$ for $i = 1,\dots,r$). Linear forms $L_1,\dots,L_r$ are not uniquely determined by (1.8). For any set of linear forms $L_1,\dots,L_r$ with (1.8) we define the quantity

$$\Delta(L_1,\dots,L_r) := \max_{\{i_1,\dots,i_n\}\subset\{1,\dots,r\}} |\det(L_{i_1},\dots,L_{i_n})| , \tag{1.9}$$

where $\det(L_{i_1},\ldots,L_{i_n})$ is the coefficient determinant of $L_{i_1},\ldots,L_{i_n}$, and where the maximum is taken over all subsets $\{i_1,\ldots,i_n\}$ of $\{1,\ldots,r\}$ of cardinality n. We define the *invariant height* of F by

$$H^*(F) := \inf \Delta(L_1,\ldots,L_r)\ , \tag{1.10}$$

where the infimum is taken over all r-tuples of linear forms $L_1,\ldots,L_r$ with (1.8). By Lemma 1 in Section 2 of the present paper, we have $H^*(F) \geq 1$.

As usual, we write e for $2.7182\cdots$. Let again F be the norm form given by (1.2) and V the vector space given by (1.3). Theorem 2 below holds for norm forms F satisfying instead of non-degeneracy the weaker condition

$$\left.\begin{array}{l} \alpha_1,\ldots,\alpha_n \text{ are linearly independent over } \mathbb{Q}, \\ V^J \subsetneqq V \text{ for each subfield } J \text{ of } K \\ \qquad\qquad \text{not equal to } \mathbb{Q} \text{ or an imaginary quadratic number field.} \end{array}\right\} \tag{1.11}$$

Theorem 2. *Let $F \in \mathbb{Z}[X_1,\ldots,X_n]$ be a norm form of degree r in $n \geq 2$ variables with (1.11) and let P be any real ≥ 1. Then the set of solutions of (1.5) is contained in the union of not more than*

$$(16r)^{(n+6)^2} \cdot \max\left(1, \left(\frac{h^{n/r} \cdot P}{H^*(F)}\right)^{\frac{1}{n-1}}\right) \cdot \left(1 + \frac{\log(eh \cdot H^*(F))}{\log eP}\right)^{n-1}$$

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

In the proof of Theorem 2 we make as usual a distinction between "small" and "large" solutions. We estimate the number of subspaces containing the small solutions by means of a gap principle which is derived in Section 5. In the proof of this gap principle we partly use arguments from Schmidt [16]; the main new idea is probably Lemma 5 in Section 5. We deal with the large solutions by reducing eq. (1.5) to a number of inequalities of the type occurring in the Subspace Theorem (where we more or less follow [6]) and then applying the quantitative result from [8].

We state another consequence of Theorem 2. For a homogeneous polynomial $Q \in \mathbb{C}[X_1,\ldots,X_n]$ and a non-singular complex $n{\times}n$-matrix B we define the homogeneous polynomial

$$Q^B(\mathbf{X}) := Q(\mathbf{X}B)\ ,$$

where $\mathbf{X} = (X_1, \dots, X_n)$ is the row vector consisting of the n variables. Now let $F \in \mathbb{Z}[X_1, \dots, X_n]$ be a norm form of degree r and $L_1, \dots, L_r$ linear forms with (1.8). Further, let B be a non-singular $n \times n$-matrix with entries in $\mathbb{Z}$. From definition (1.9) it follows at once that

$$\Delta(L_1^B, \dots, L_r^B) = |\det B| \cdot \Delta(L_1, \dots, L_r) \ .$$

Now clearly, if $L_1, \dots, L_r$ run through all factorisations of F with (1.8), then $L_1^B, \dots,$ L_r^B run through all factorisations of F^B with (1.8). Hence the invariant height defined by (1.10) satisfies

$$H^*(F^B) = |\det B| \cdot H^*(F) \ . \tag{1.12}$$

Two norm forms $F, G \in \mathbb{Z}[X_1, \dots, X_n]$ are said to be *equivalent* if $G = F^B$ for some matrix $B \in GL_n(\mathbb{Z})$, i.e., with $\det B = \pm 1$. Thus, a special case of (1.12) is that

$$H^*(G) = H^*(F) \ \text{ for equivalent norm forms } F, G. \tag{1.13}$$

For a norm form $F \in \mathbb{Z}[X_1, \dots, X_n]$, let $||F||$ denote the maximum of the absolute values of its coefficients. In [17], Schmidt developed a reduction theory for norm forms, which implies that for every norm form $F \in \mathbb{Z}[X_1, \dots, X_n]$ of degree r there is a matrix $B \in GL_n(\mathbb{Z})$ such that $||F^B||$ is bounded from above in terms of r, n and $H^*(F)$. By combining this with Theorem 2, we show in an explicit form the following: there is a finite union of equivalence classes depending on h, such that for all norm forms F outside this union, the set of solutions of (1.5) is contained in the union of at most a quantity independent of h proper linear subspaces of $\mathbb{Q}^n$.

Theorem 3. *Let $h \geq 1$ and let $F \in \mathbb{Z}[X_1, \dots, X_n]$ be a norm form of degree r in $n \geq 2$ variables with (1.11) and with*

$$\min_{B \in GL_n(\mathbb{Z})} ||F^B|| \geq (32n)^{nr/2} h^{2n} \ . \tag{1.14}$$

Then the set of solutions of (1.5) is contained in the union of not more than

$$(16r)^{(n+7)^2}$$

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

The case $n = 2$ of Theorem 3 was considered earlier by Győry and the author in [7]. Note that each one-dimensional subspace of $\mathbb{Q}^2$ contains precisely two primitive points, i.e., points with coordinates in $\mathbb{Z}$ whose gcd is equal to 1. Combining the method of Bombieri and Schmidt [3] with linear forms in logarithms estimates, Győry and the author showed that the Thue inequality $|F(x,y)| \leq h$ has at most $12r$ primitive solutions $(x,y) \in \mathbb{Z}^2$ provided that $F \in \mathbb{Z}[X,Y]$ is an irreducible binary form of degree $r \geq 400$ with $\min_{B \in GL_2(\mathbb{Z})} ||F^B|| \geq \exp(c_1(r) h^{10(r-1)^2})$ for some effectively computable function $c_1(r)$ of r. Theorem 3 gives the much worse upper bound $2 \times (16r)^{81}$ for the number of primitive solutions of $|F(x,y)| \leq h$ but subject to the much weaker constraints that F have degree $r \geq 3$ and $\min_{B \in GL_2(\mathbb{Z})} ||F^B|| \geq 2^{6r} h^4$. The polynomial dependence on h of this last condition is because in the proof of Theorem 3 no linear forms in logarithms estimates were used. Under a similar condition on F and for all $r \geq 3$, Győry obtained the upper bound $28r$ (personal communication).

One may wonder whether there is a sharpening of Theorem 3 which gives for all norm forms F in $n \geq 3$ variables lying outside some union of finitely many equivalence classes, an upper bound independent of h for the number of primitive solutions of equation (1.5) instead of just for the number of subspaces. It was already explained in [5] that such a sharpening does not exist. To construct a counterexample, one takes a number field K of degree r and fixes $\mathbb{Q}$-linearly independent $\alpha_1, \ldots, \alpha_{n-1} \in K$ and $c \in \mathbb{Q}^*$ such that the norm form $cN_{K/\mathbb{Q}}(\alpha_1 X_1 + \cdots + \alpha_{n-1} X_{n-1})$ has its coefficients in $\mathbb{Z}$. Now if α_n runs through all algebraic integers of K, then $F := cN_{K/\mathbb{Q}}(\alpha_1 X_1 + \cdots + \alpha_{n-1} X_{n-1} + \alpha_n X_n)$ runs through infinitely many pairwise inequivalent norm forms in $\mathbb{Z}[X_1, \ldots, X_n]$. Clearly, one has $|F(\mathbf{x})| \leq h$ for every algebraic integer $\alpha_n \in K$ and each primitive vector $\mathbf{x} = (x_1, \ldots, x_{n-1}, 0)$ with $x_i \in \mathbb{Z}$, $|x_i| \ll h^{1/r}$ for $i = 1, \ldots, n-1$, the constant implied by $\ll$ depending only on $K, c, \alpha_1, \ldots, \alpha_{n-1}$. Thus, there are infinitely many pairwise inequivalent norm forms $F \in \mathbb{Z}[X_1, \ldots, X_n]$ of degree r such that for every such F and for every $h \gg 1$, the inequality $|F(\mathbf{x})| \leq h$ has $\gg h^{(n-1)/r}$ primitive solutions $\mathbf{x} \in \mathbb{Z}^n$ lying in the subspace $x_n = 0$.

2. Proof of Theorem 1.

In this section, we deduce Theorem 1 from Theorem 2. Let $F \in \mathbb{Z}[X_1, \ldots, X_n]$ be the norm form of degree r given by (1.2) and let K, $\alpha_1, \ldots, \alpha_n$ and c be as in (1.2). Assume that F is non-degenerate. The cardinality of a set $\mathcal{I}$ is denoted by $|\mathcal{I}|$. We need the following lemma:

Lemma 1. $H^*(F) \geq 1$.

Proof. Choose linear forms $L_1, \dots, L_r$ with (1.8). Let $\mathcal{I}$ denote the collection of ordered n-tuples $(i_1, \dots, i_n)$ from $\{1, \dots, r\}$ for which $\det(L_{i_1}, \dots, L_{i_n}) \neq 0$. According to Schmidt ([17], p. 203), the semi-discriminant

$$D(F) := \prod_{(i_1,\dots,i_n)\in\mathcal{I}} |\det(L_{i_1}, \dots, L_{i_n})|$$

is a positive integer. This implies that

$$\Delta(L_1, \dots, L_r) = \max_{(i_1,\dots,i_n)\in\mathcal{I}} |\det(L_{i_1}, \dots, L_{i_n})| \geq D(F)^{1/|\mathcal{I}|} \geq 1 \ .$$

By taking the infimum over all $L_1, \dots, L_r$ with (1.8) we obtain Lemma 1. □

Assume that F is non-degenerate. Denote by $Z_F^*(h)$ the number of primitive solutions of (1.5), i.e., with $\gcd(x_1, \dots, x_n) = 1$.

Lemma 2. $Z_F^*(h) \leq \frac{1}{3} \cdot (16r)^{\frac{1}{3}(n+7)^3} \cdot h^{(n+\sum_{m=2}^{n-1}\frac{1}{m})/r} \cdot (1 + \log h)^{\frac{1}{2}n(n-1)}$.

Proof. Denote by $A(n, r, h)$ the right-hand side of the inequality in Lemma 2. We proceed by induction on n. First, let $n = 2$. Since F is non-degenerate, condition (1.11) is satisfied. On applying Theorem 2 with $P = H^*(F)$ (which is allowed by Lemma 1), we infer that the set of solutions of (1.5) is contained in the union of not more than

$$(16r)^{64} h^{2/r} \Big(1 + \frac{\log(eh \cdot H^*(F))}{\log eH^*(F)}\Big) \leq (16r)^{64} h^{2/r}(2 + \log h) \leq \frac{1}{2} A(2, r, h)$$

proper one-dimensional linear subspaces of $\mathbb{Q}^2$. Using that each one-dimensional subspace contains at most two primitive solutions, we get $Z_F^*(h) \leq A(2, r, h)$.

Now let $n \geq 3$. Again, (1.11) holds since F is non-degenerate, and again from Theorem 2 with $P = H^*(F)$ we infer that the set of solutions of (1.5) is contained in the union of not more than

$$\begin{aligned}(16r)^{(n+6)^2} h^{\frac{n}{(n-1)r}} \Big(1 + \frac{\log(eh \cdot H^*(F))}{\log eH^*(F)}\Big)^{n-1} \\ \leq (16r)^{(n+6)^2} 2^{n-1} \cdot h^{\frac{n}{(n-1)r}} (1 + \log h)^{n-1} \quad =: B(n, r, h)\end{aligned}$$

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

There is no loss of generality to assume that these subspaces have dimension $n-1$. Let T be one of these subspaces and consider the solutions of (1.5) lying in T. Fix a basis $\{\mathbf{a}_i = (a_{i1}, \dots, a_{in}) :\ i = 1, \dots, n-1\}$ of the $\mathbb{Z}$-module $T \cap \mathbb{Z}^n$ and define the norm form $G \in \mathbb{Z}[Y_1, \dots, Y_{n-1}]$ in $n-1$ variables by

$$G := F(Y_1\mathbf{a}_1 + \cdots + Y_{n-1}\mathbf{a}_{n-1}) \ .$$

Clearly, there is a one-to-one correspondence between the primitive solutions of (1.5) lying in T and the primitive solutions of

$$|G(\mathbf{y})| \le h \quad \text{ in } \mathbf{y} \in \mathbb{Z}^{n-1} \ . \tag{2.1}$$

Note that since $F = cN_{K/\mathbb{Q}}(\alpha_1 X_1 + \cdots + \alpha_n X_n)$ we have

$$G = cN_{K/\mathbb{Q}}(\beta_1 Y_1 + \cdots + \beta_{n-1} Y_{n-1}) \ \text{ with } \beta_i = \sum_{j=1}^{n} a_{ij}\alpha_j \ \text{ for } i = 1, \dots, n-1.$$

The vector space associated to G is $W := \{\beta_1 y_1 + \cdots + \beta_{n-1} y_{n-1} :\ y_1, \dots, y_{n-1} \in \mathbb{Q}\}$. As usual, for each subfield J of K we define $W^J := \{\xi \in W : \lambda\xi \in W \text{ for every } \lambda \in J\}$. We verify that G is non-degenerate. First, the numbers $\beta_1, \dots, \beta_{n-1}$ are linearly independent over $\mathbb{Q}$ since $\alpha_1, \dots, \alpha_n$ are $\mathbb{Q}$-linearly independent and since the vectors $\mathbf{a}_1, \dots, \mathbf{a}_{n-1}$ are $\mathbb{Q}$-linearly independent. Second, since $W \subset V$ and F is non-degenerate we have that $W^J \subset V^J = (0)$ if J is not equal to $\mathbb{Q}$ or to an imaginary quadratic field.

We infer from the induction hypothesis that the number of primitive solutions of (2.1), and hence the number of primitive solutions of (1.5) lying in T, is at most $A(n-1, r, h)$. Since we have at most $B(n, r, h)$ possibilities for T, we conclude that the total number of primitive solutions of (1.5) is at most

$$\begin{aligned}
&A(n-1, r, h) \cdot B(n, r, h) \\
&= \frac{1}{3} \cdot 2^{n-1} (16r)^{\frac{1}{3}(n+6)^3 + (n+6)^2} \cdot h^{(n-1+(\sum_{m=2}^{n-2} \frac{1}{m}) + \frac{n}{n-1})/r} \cdot (1 + \log h)^{\frac{1}{2}(n-1)(n-2)+n-1} \\
&\le \frac{1}{3} \cdot (16r)^{\frac{1}{3}(n+7)^3} \cdot h^{(n + \sum_{m=2}^{n-1} \frac{1}{m})/r} \cdot (1 + \log h)^{\frac{1}{2}n(n-1)} = A(n, r, h) \ . \qquad \square
\end{aligned}$$

Proof of Theorem 1. We have to prove that $Z_F(h) \le \psi(h) h^{n/r}$, where

$$\psi(h) = (16r)^{\frac{1}{3}(n+7)^3} h^{(\sum_{m=2}^{n-1} \frac{1}{m})/r} (1 + \log h)^{\frac{1}{2}n(n-1)}.$$

For $c = 0, \ldots, h$, denote by $a(c)$ the number of primitive solutions $\mathbf{x} \in \mathbb{Z}^n$ of $|F(\mathbf{x})| = c$ and by $b(c)$ the number of all solutions $\mathbf{x} \in \mathbb{Z}^n$ of $|F(\mathbf{x})| = c$. Thus,

$$
\begin{aligned}
&a(0) = 0, \quad b(0) = 1, \quad b(c) = \sum_{d:\ d^r | c, d>0} a(c/d^r) \text{ for } c > 0, \\
&Z_F(h) = \sum_{c=0}^{h} b(c), \quad Z_F^*(h) = \sum_{c=0}^{h} a(c), \quad \text{for } h \geq 0.
\end{aligned}
$$

This implies, on interchanging the summation and then using $a(c) = Z_F^*(c) - Z_F^*(c-1)$ for $c \geq 1$,

$$
\begin{aligned}
Z_F(h) &= 1 + \sum_{c=1}^{h} \sum_{d:\ d^r | c, d>0} a(c/d^r) \\
&\leq 1 + \sum_{c=1}^{h} a(c) \sqrt[r]{h/c} \\
&= 1 + Z_F^*(h) + \sum_{c=1}^{h-1} \left(\sqrt[r]{h/c} - \sqrt[r]{h/(c+1)} \right) \cdot Z_F^*(c) \ .
\end{aligned}
$$

By Lemma 2 we have $Z_F^*(c) \leq \frac{1}{3}\psi(h) c^{n/r}$ for $c \leq h$. Hence

$$
\begin{aligned}
Z_F(h) &\leq 1 + \frac{1}{3}\psi(h) \left(h^{n/r} + \sum_{c=1}^{h-1} \left(\sqrt[r]{h/c} - \sqrt[r]{h/(c+1)} \right) \cdot c^{n/r} \right) \\
&= 1 + \frac{1}{3}\psi(h) \sum_{c=1}^{h} \sqrt[r]{h/c} \cdot \left(c^{n/r} - (c-1)^{n/r} \right) \\
&\leq 1 + \frac{1}{3}\psi(h) \int_0^h \sqrt[r]{h/x} \cdot \frac{n}{r} x^{\frac{n}{r}-1} dx = 1 + \frac{1}{3}\psi(h) \frac{n}{n-1} h^{n/r} \\
&\leq \psi(h) h^{n/r} \ .
\end{aligned}
$$

This proves Theorem 1. $\square$

3. Proof of Theorem 3.

Let $F \in \mathbb{Z}[X_1, \dots, X_n]$ be a norm form of degree r satisfying (1.2), (1.11). Define the quantity

$$H_2^*(F) := \inf \Big(\sum_{\{i_1,\dots,i_n\}\subset\{1,\dots,r\}} |\det(L_{i_1}, \dots, L_{i_n})|^2 \Big)^{1/2} ,$$

where the sum is taken over all subsets of $\{1, \dots, r\}$ of cardinality n and the infimum over all r-tuples of linear forms $L_1, \dots, L_r$ with (1.8). By Schmidt's reduction theory for norm forms (cf. [17], Lemma 4) we have

$$\min_{B\in GL_n(\mathbb{Z})} ||F^B|| \le \big(2^n n^{3/2} V(n)^{-1}\big)^r \cdot H_2^*(F)^r ,$$

where $V(n)$ is the volume of the n-dimensional Euclidean ball with radius 1. Together with $V(n) \ge (n!)^{1/2}$ and $H_2^*(F) \le \binom{r}{n}^{1/2} H^*(F)$, this implies

$$\min_{B\in GL_n(\mathbb{Z})} ||F^B|| \le (32n)^{nr/2} \cdot H^*(F)^r .$$

Now assume that F satisfies (1.14). Then it follows that

$$H^*(F) \ge h^{2n/r} .$$

By applying Theorem 2 with $P = H^*(F)^{1/2}$, we infer that the set of solutions of (1.5) is contained in the union of at most

$$(16r)^{(n+6)^2} \cdot \Big(1 + \frac{\log(eh \cdot H^*(F))}{\log eH^*(F)^{1/2}}\Big)^n \le (16r)^{(n+6)^2} \cdot \Big(1 + \frac{\log eH^*(F)^{1+(r/2n)}}{\log eH^*(F)^{1/2}}\Big)^{n-1}$$
$$\le (16r)^{(n+7)^2}$$

proper linear subspaces of $\mathbb{Q}^n$. This proves Theorem 3. □

4. Choice of the linear factors.

By $\overline{\mathbb{Q}}$ we denote the algebraic closure of $\mathbb{Q}$ in $\mathbb{C}$. We agree that algebraic number fields occurring in this paper are contained in $\overline{\mathbb{Q}}$. Vectors from $\mathbb{Q}^n$, $\mathbb{R}^n$, etc., will always be row vectors.

Let $F \in \mathbb{Z}[X_1, \ldots, X_n]$ be a norm form of degree r satisfying (1.2) for some number field K, some $\alpha_1, \ldots, \alpha_n \in K$, and some non-zero $c \in \mathbb{Q}$. As before, we order the isomorphic embeddings of K into $\mathbb{C}$ in such a way that $\alpha \mapsto \alpha^{(i)}$ map K into $\mathbb{R}$ for $i = 1, \ldots, r_1$ and $\alpha^{(i+r_2)} = \overline{\alpha^{(i)}}$ for $i = r_1+1, \ldots, r_1+r_2$, where $r_1+2r_2 = r = [K:\mathbb{Q}]$. In this section we choose appropriate linear factors $L_1, \ldots, L_r$ of F satisfying (1.8). The quantities $\Delta(L_1, \ldots, L_r)$ and $H^*(F)$ are defined by (1.9), (1.10), respectively.

Lemma 3. *There are linear forms $L_1, \ldots, L_r$ which satisfy (1.8) and which have the following additional properties:*

$$L_i = \sqrt[k]{\beta^{(i)}} \cdot (\alpha_1^{(i)} X_1 + \cdots + \alpha_n^{(i)} X_n) \quad (i = 1, \ldots, r) \tag{4.1}$$

for some $\beta \in K$, $k \in \mathbb{Z}_{>0}$ and choices for the roots $\sqrt[k]{\beta^{(1)}}, \ldots, \sqrt[k]{\beta^{(r)}}$;

$$L_1, \ldots, L_r \text{ have algebraic integer coefficients;} \tag{4.2}$$

$$H^*(F) \leq \Delta(L_1, \ldots, L_r) \leq 2H^*(F)\ . \tag{4.3}$$

Proof. We will frequently use that if $L_1, \ldots, L_r$ satisfy (1.8) then $c_1 L_1, \ldots, c_r L_r$ satisfy (1.8) if and only if $c_i \in \mathbb{R}$ for $i = 1, \ldots, r_1$, $c_{i+r_2} = \overline{c_i}$ for $i = r_1+1, \ldots, r_1+r_2$ and $c_1 \cdots c_r = \pm 1$.

Suppose that K has class number h. Let $\mathfrak{a}$ denote the fractional ideal in K generated by $\alpha_1, \ldots, \alpha_r$. Then $\mathfrak{a}^h$ is a principal ideal with generator $\gamma \in K$, say. Choose roots $\delta_i := \sqrt[2h]{(\gamma^{(i)})^2}$ such that $\delta_i \in \mathbb{R}_{>0}$ for $i = 1, \ldots, r_1$ and such that $\delta_{i+r_2} = \overline{\delta_i}$ for $i = r_1+1, \ldots, r_1+r_2$ (note that for $i = 1, \ldots, r_1$, $(\gamma^{(i)})^2$ is positive so that it has a positive real $2h$-th root).

Define the linear forms

$$\tilde{L}_i := \sqrt[r]{|c| \cdot \delta_1 \cdots \delta_r} \cdot \delta_i^{-1} (\alpha_1^{(i)} X_1 + \cdots + \alpha_n^{(i)} X_n) \quad \text{for } i = 1, \ldots, r, \tag{4.4}$$

where the r-th root is a positive real. From (1.2) it follows easily that $\tilde{L}_1, \ldots, \tilde{L}_r$ satisfy (1.8). We claim that $\tilde{L}_i$ has algebraic integer coefficients for $i = 1, \ldots, r$. Let M be a finite extension of K containing the numbers $\alpha_i^{(j)}$, δ_j $(i = 1, \ldots, n,\ j = 1, \ldots, r)$ and $\sqrt[r]{|c| \cdot \delta_1 \cdots \delta_r}$. For $\beta_1, \ldots, \beta_m \in M$, denote by $[\beta_1, \ldots, \beta_m]$ the fractional ideal in M generated by $\beta_1, \ldots, \beta_m$. For a polynomial $Q \in M[X_1, \ldots, X_n]$, denote by $[Q]$ the fractional ideal in M generated by the coefficients in Q. By the choice of the δ_i we have $[\delta_i]^{2h} = [\gamma^{(i)}]^2 = [\alpha_1^{(i)}, \ldots, \alpha_n^{(i)}]^{2h}$, hence $[\delta_i] = [\alpha_1^{(i)}, \ldots, \alpha_n^{(i)}]$. Therefore, $[\tilde{L}_i] = [\sqrt[r]{|c| \cdot \delta_1 \cdots \delta_r}\,]$. But according to Gauss' lemma for Dedekind domains we have $[F] = [\tilde{L}_1] \cdots [\tilde{L}_r] = [\,|c| \cdot \delta_1 \cdots \delta_r]$. By assumption, F has its coefficients in $\mathbb{Z}$, so $|c| \cdot \delta_1 \cdots \delta_r$ is an algebraic integer. This proves our claim.

Let $\theta > 0$. From the definition of $H^*(F)$ it follows at once that there are complex numbers $c_1, \ldots, c_r$ with $c_1, \ldots, c_{r_1} \in \mathbb{R}$, $c_{i+r_2} = \overline{c_i}$ for $i = r_1 + 1, \ldots, r_1 + r_2$ and $c_1 \cdots c_r = \pm 1$ such that

$$\Delta(c_1 \tilde{L}_1, \ldots, c_r \tilde{L}_r) \le (1+\theta) H^*(F)\ . \tag{4.5}$$

We approximate $c_1, \ldots, c_r$ by algebraic units. Let U_K denote the unit group of the ring of integers of K. According to Dirichlet's unit theorem, the set $\{(\log|\varepsilon^{(1)}|, \ldots, \log|\varepsilon^{(r)}|) : \varepsilon \in U_K\}$ is a lattice which spans the linear subspace $H \subset \mathbb{R}^r$ given by the equations $x_1 + \cdots + x_r = 0$, $x_{i+r_2} = x_i$ for $i = r_1 + 1, \ldots, r_1 + r_2$. This implies that there is a constant $C_K > 0$ such that for every positive integer m there is an $\varepsilon \in U_K$ with

$$\big|\log|\varepsilon^{(i)}| - m\log|c_i|\,\big| \le C_K \quad \text{for } i = 1, \ldots, r.$$

Choose m so large that $C_K < m\log(1+\theta)$. Choose roots $\eta_i := \sqrt[2m]{(\varepsilon^{(i)})^2}$ such that $\eta_i \in \mathbb{R}$ for $i = 1, \ldots, r_1$ and $\eta_{i+r_2} = \overline{\eta_i}$ for $i = r_1 + 1, \ldots, r_1 + r_2$. Thus, $(\log|\eta_1|, \ldots, \log|\eta_r|) \in H$ and

$$\big|\log|\eta_i| - \log|c_i|\big| \le \log(1+\theta) \quad \text{for } i = 1, \ldots, r\ . \tag{4.6}$$

Define the linear forms

$$L_i := \eta_i \cdot \tilde{L}_i = \eta_i \sqrt[r]{|c|\delta_1 \cdots \delta_r} \cdot \delta_i^{-1}(\alpha_1^{(i)} X_1 + \cdots + \alpha_n^{(i)} X_n) \quad \text{for } i = 1, \ldots, r.$$

Note that with $k = 2hmr$ we have

$$\begin{aligned}(\eta_i \sqrt[r]{|c|\delta_1 \cdots \delta_r} \cdot \delta_i^{-1})^k &= (\varepsilon^{(i)})^{2hr} |c|^{2hm} (\gamma^{(1)} \cdots \gamma^{(r)})^{2m} (\gamma^{(i)})^{-2mr} = \beta^{(i)} \\ &\qquad \text{with } \beta := \varepsilon^{2hr} |c|^{2hm} N_{K/\mathbb{Q}}(\gamma)^{2m} \gamma^{-2mr} \in K\ .\end{aligned}$$

Hence $L_1, \ldots, L_r$ satisfy (4.1) for some β, k. It is easy to check that $L_1, \ldots, L_r$ satisfy (1.8) and (4.2). To verify (4.3), we observe that by (4.6), (4.5) we have, on choosing $i_1, \ldots, i_n \in \{1, \ldots, r\}$ with $\Delta(L_1, \ldots, L_r) = |\det(L_{i_1}, \ldots, L_{i_n})|$,

$$\begin{aligned}\Delta(L_1, \ldots, L_r) &= \frac{|\eta_{i_1} \cdots \eta_{i_n}|}{|c_{i_1} \cdots c_{i_n}|} \cdot |\det(c_{i_1} \tilde{L}_{i_1}, \ldots, c_{i_n} \tilde{L}_{i_n})| \\ &\le (1+\theta)^n \cdot |\det(c_{i_1} \tilde{L}_{i_1}, \ldots, c_{i_n} \tilde{L}_{i_n})| \le (1+\theta)^{n+1} H^*(F) \le 2H^*(F)\end{aligned}$$

for sufficiently small θ. Lastly, since $L_1, \ldots, L_r$ satisfy (1.8) we have $\Delta(L_1, \ldots, L_r) \ge H^*(F)$. This proves Lemma 3. □

We recall that for a homogeneous polynomial $Q \in \mathbb{C}[X_1,\dots,X_n]$ and a non-singular complex $n\times n$-matrix B we define $Q^B(\mathbf{X}) := Q(\mathbf{X}B)$. Further, we denote by $||Q||$ the maximum of the absolute values of the coefficients of Q.

Lemma 4. *Let $L_1,\dots,L_r$ be linear forms with (1.8), (4.1), (4.2), (4.3). Then there is a matrix $B \in GL_n(\mathbb{Z})$ such that*

$$||L_i^B|| \le (2n)^{n+1}H^*(F) \quad \text{for } i=1,\dots,r. \tag{4.7}$$

Proof. We modify an argument of Schmidt from [17]. We will apply Minkowski's theorem on successive minima to the symmetric convex body

$$\mathcal{C} := \{\mathbf{x} \in \mathbb{R}^n : \ |L_i(\mathbf{x})| \le 1 \ \text{ for } i=1,\dots,r\}.$$

We need a lower bound for the volume of $\mathcal{C}$. Recall (1.8). Let $M_1,\dots,M_r$ be the linear forms with real coefficients given by

$$\left.\begin{aligned} &M_i := L_i \quad (i=1,\dots,r_1),\\ &M_i := \tfrac{1}{2}(L_i+\overline{L_i}) = \tfrac{1}{2}(L_i+L_{i+r_2}) \quad (i=r_1+1,\dots,r_1+r_2),\\ &M_{i+r_2} := \tfrac{1}{2\sqrt{-1}}(L_i-\overline{L_i}) = \tfrac{1}{2\sqrt{-1}}(L_i-L_{i+r_2}) \quad (i=r_1+1,\dots,r_1+r_2). \end{aligned}\right\} \tag{4.8}$$

Let $\{j_1,\dots,j_n\}$ be a subset of $\{1,\dots,r\}$ for which $|\det(M_{j_1},\dots,M_{j_n})|$ is maximal. Suppose that $1 \le j_1 < \cdots j_s \le r_1 < j_{s+1} < \cdots < j_n$. By (4.8) we have

$$\det(M_{j_1},\dots,M_{j_n}) = \sum_I \varepsilon_I \det(L_{j_1},\dots,L_{j_s},L_{i_{s+1}},\dots,L_{i_n})$$

where the sum is taken over tuples $I=(i_{s+1},\dots,i_n)$ with precisely two possibilities for each index i_j and where $|\varepsilon_I| = 2^{s-n}$ for each tuple I. Together with (4.3) this implies

$$|\det(M_{j_1},\dots,M_{j_n})| \le \Delta(L_1,\dots,L_r) \le 2H^*(F)\ . \tag{4.9}$$

From (4.8) it follows that $\mathrm{rank}\{M_1,\dots,M_r\} = n$. Hence $|\det(M_{j_1},\dots,M_{j_n})| \ne 0$. So there are $c_{ik} \in \mathbb{C}$ with

$$L_i = \sum_{k=1}^n c_{ik}M_{j_k} \quad \text{for } i=1,\dots,r. \tag{4.10}$$

We estimate $|c_{ik}|$ from above. First suppose $i \le r_1$. Then $L_i = M_i$ by (4.8), so

$$|c_{ik}| = \frac{|\det(M_{j_1}, \ldots, M_i, \ldots, M_{j_n})|}{|\det(M_{j_1}, \ldots, M_{j_k}, \ldots, M_{j_n})|} \le 1 \ .$$

Now suppose $r_1 + 1 \le i \le r_1 + r_2$. Then by (4.8) we have $L_i = M_i + \sqrt{-1} M_{i+r_2}$, so

$$\begin{aligned} |c_{ik}| &= \frac{|\det(M_{j_1}, \ldots, L_i, \ldots, M_{j_n})|}{|\det(M_{j_1}, \ldots, M_{j_k}, \ldots, M_{j_n})|} \\ &\le \frac{|\det(M_{j_1}, \ldots, M_i, \ldots, M_{j_n})|}{|\det(M_{j_1}, \ldots, M_{j_k}, \ldots, M_{j_n})|} + \frac{|\det(M_{j_1}, \ldots, M_{i+r_2}, \ldots, M_{j_n})|}{|\det(M_{j_1}, \ldots, M_{j_k}, \ldots, M_{j_n})|} \le 2 \ . \end{aligned}$$

We have a similar estimate for $|c_{ik}|$ for $r_1 + r_2 + 1 \le i \le r$. Hence $|c_{ik}| \le 2$ for $i = 1, \ldots, r$, $k = 1, \ldots, n$. Together with (4.10) this implies

$$\mathcal{C} \supseteq \mathcal{D} := \{\mathbf{x} \in \mathbb{R}^n : \ |M_{j_k}(\mathbf{x})| \le (2n)^{-1} \ \text{ for } k = 1, \ldots, n\} \ .$$

So by (4.9),

$$\begin{aligned} \mathrm{vol}(\mathcal{C}) \ge \mathrm{vol}(\mathcal{D}) &= 2^n (2n)^{-n} |\det(M_{j_1}, \ldots, M_{j_n})|^{-1} \\ &\ge \frac{1}{2} n^{-n} H^*(F)^{-1} \ . \end{aligned} \tag{4.11}$$

Denote by $\lambda_1, \ldots, \lambda_n$ the successive minima of $\mathcal{C}$ with respect to $\mathbb{Z}^n$. Thus, there are linearly independent vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n \in \mathbb{Z}^n$ with $\mathbf{b}_j \in \lambda_j \mathcal{C}$, i.e., with $|L_i(\mathbf{b}_j)| \le \lambda_j$ for $i = 1, \ldots, r$, $j = 1, \ldots, n$. By Minkowski's theorem and (4.11) we have

$$\lambda_1 \cdots \lambda_n \le 2^n \mathrm{vol}(\mathcal{C})^{-1} \le 2^{n+1} n^n H^*(F) \ .$$

Further, by (1.8) we have $1 \le |F(\mathbf{b}_1)| = \prod_{i=1}^n |L_i(\mathbf{b}_1)| \le \lambda_1^n$. Hence

$$\lambda_n \le 2^{n+1} n^n H^*(F) \ . \tag{4.12}$$

By a result of Mahler (cf. Cassels [4], Lemma 8, p. 135), $\mathbb{Z}^n$ has a basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ with $\mathbf{b}_j \in j\lambda_j \mathcal{C}$ for $j = 1, \ldots, n$. Together with (4.12) this implies

$$|L_i(\mathbf{b}_j)| \le n\lambda_n \le (2n)^{n+1} H^*(F) \ \text{ for } i = 1, \ldots, r, \ j = 1, \ldots, n.$$

Now Lemma 4 holds for the matrix B with rows $\mathbf{b}_j$ $(j = 1, \ldots, n)$. □

Let $L_1, \ldots, L_r$ be linear forms with (1.8), (4.1)–(4.3) and let B be the matrix from Lemma 4. We now write F for F^B, L_i for L_i^B and replace everywhere the old forms F, L_i by the new ones just chosen. This affects neither the minimal number of subspaces of $\mathbb{Q}^n$ containing the set of solutions of (1.5) nor the invariant height $H^*(F)$. Further, the conditions (1.8) and (4.1)–(4.3) remain valid (but with different $\alpha_1, \dots, \alpha_n$ in (4.1)). Lastly, condition (4.7) holds but with B being replaced by the identity matrix.

So it suffices to prove Theorem 2 for these newly chosen forms $F, L_1, \dots, L_r$ and we will proceed further with these forms. This means that in the remainder of this paper, F is a norm form in $\mathbb{Z}[X_1, \dots, X_n]$ of degree r of the shape (1.2) satisfying (1.11), K is the number field and $\alpha_1, \dots, \alpha_n$ are the elements of K from (1.2), and $L_1, \dots, L_r$ are linear forms with the following properties:

$$F = \pm L_1 \cdots L_r; \tag{4.13}$$

$$L_1, \dots, L_{r_1} \text{ have real coefficients;} \tag{4.14}$$

$$L_{i+r_2} = \overline{L}_i \quad \text{for } i = r_1+1, \dots, r_1+r_2; \tag{4.15}$$

$$L_i = \sqrt[k]{\beta^{(i)}} \cdot (\alpha_1^{(i)} X_1 + \cdots + \alpha_n^{(i)} X_n) \quad (i = 1, \dots, r) \tag{4.16}$$

for some $\beta \in K$, $k \in \mathbb{Z}_{>0}$ and choices for the roots $\sqrt[k]{\beta^{(1)}}, \dots, \sqrt[k]{\beta^{(r)}}$;

$$H^*(F) \le \Delta(L_1, \dots, L_r) \le 2H^*(F)\ ; \tag{4.17}$$

$$||L_i|| \le (2n)^{n+1} H^*(F) \quad \text{for } i = 1, \dots, r; \tag{4.18}$$

$$L_1, \dots, L_r \text{ have algebraic integer coefficients.} \tag{4.19}$$

We fix once and for all a finite, normal extension $N \subset \mathbb{C}$ of $\mathbb{Q}$ such that N contains K, the images of the isomorphic embeddings $\alpha \mapsto \alpha^{(i)}$ $(i = 1, \dots, r)$ of K into $\mathbb{C}$, the coefficients of $L_1, \dots, L_r$ and the k-th roots of unity, where k is the integer from (4.16). Let

$$d := [N : \mathbb{Q}] \tag{4.20}$$

and denote by $\mathrm{Gal}(N/\mathbb{Q})$ the Galois group of $N/\mathbb{Q}$. Clearly, for each $\sigma \in \mathrm{Gal}(N/\mathbb{Q})$ there is a permutation $\sigma^*(1), \dots, \sigma^*(r)$ of $1, \dots, r$ such that

$$\sigma(\alpha^{(i)}) = \alpha^{(\sigma^*(i))} \quad \text{for } \alpha \in K,\ i = 1, \dots, r. \tag{4.21}$$

For each pair $i, j \in \{1, \dots, r\}$, we have

$$\sigma^*(i) = j \text{ for precisely } d/r \text{ elements } \sigma \in \mathrm{Gal}(N/\mathbb{Q}), \tag{4.22}$$

since the $\mathbb{Q}$-isomorphism $\alpha^{(i)} \mapsto \alpha^{(j)}$ $(\alpha \in K)$ can be extended in exactly d/r ways to an automorphism of N. For a linear form $L = \alpha_1 X_1 + \cdots + \alpha_n X_n$ with coefficients in N and for $\sigma \in \mathrm{Gal}(N/\mathbb{Q})$ define $\sigma(L) := \sigma(\alpha_1) X_1 + \cdots + \sigma(\alpha_n) X_n$. From (4.21) and (4.16) it follows that there are k-th roots of unity $\rho_{\sigma,i}$ such that

$$\sigma(L_i) = \rho_{\sigma,i} L_{\sigma^*(i)} \quad \text{for } i = 1, \dots, r, \quad \sigma \in \mathrm{Gal}(N/\mathbb{Q}). \tag{4.23}$$

Denote by ι the restriction to N of the complex conjugation. Note that $\iota \in \mathrm{Gal}(N/\mathbb{Q})$. Recall that the conjugates of $\alpha \in K$ were so ordered that $\alpha^{(i)} \in \mathbb{R}$ for $i = 1, \dots, r_1$ and $\alpha^{(i+r_2)} = \overline{\alpha^{(i)}}$ for $i = r_1 + 1, \dots, r_1 + r_2$. By (4.21) we have that $\iota^*(i) = i$ for $i = 1, \dots, r_1$ and that ι^* interchanges i and $i + r_2$ for $i = r_1 + 1, \dots, r_1 + r_2$. Together with (4.14) (i.e., $\overline{L_i} = L_i$ for $i = 1, \dots, r_1$) and (4.15) this implies

$$L_{\iota^*(i)} = \overline{L_i} \quad \text{for } i = 1, \dots, r. \tag{4.24}$$

5. The small solutions.

In this section, we develop a gap principle to deal with the small solutions of (1.5). We need a preparatory lemma.

Lemma 5. *Let $D \geq 1$ and let $\mathcal{S}$ be a subset of $\mathbb{Z}^n$ with the property that*

$$|\det(\mathbf{x}_1, \dots, \mathbf{x}_n)| \leq D \quad \textit{for each } n\textit{-tuple } \mathbf{x}_1, \dots, \mathbf{x}_n \in \mathcal{S}. \tag{5.1}$$

Then $\mathcal{S}$ is contained in the union of not more than

$$100^n \cdot D^{\frac{1}{n-1}}$$

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

Proof. We assume without loss of generality that $\mathcal{S}$ is not contained in a single proper linear subspace of $\mathbb{Q}^n$, i.e., that $\mathcal{S}$ contains n linearly independent vectors, $\mathbf{x}_1, \dots, \mathbf{x}_n$, say. Then every $\mathbf{x} \in \mathcal{S}$ can be expressed as $\mathbf{x} = \sum_{i=1}^n z_i \mathbf{x}_i$ for certain $z_i \in \mathbb{Q}$. By (5.1) we have for such an $\mathbf{x}$,

$$|z_i| = \frac{|\det(\mathbf{x}_1, \dots, \mathbf{x}, \dots, \mathbf{x}_n)|}{|\det(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_n)|} \leq D \quad \text{for } i = 1, \dots, n.$$

This implies that $\mathcal{S}$ is finite.

Let $\mathcal{S} = \{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$. Denote by $\mathcal{C}$ the smallest convex body which contains $\mathcal{S}$ and which is symmetric about $\mathbf{0}$, i.e.,

$$\mathcal{C} = \{\sum_{i=1}^{m} z_i\mathbf{x}_i : z_i \in \mathbb{R}, \sum_{i=1}^{m} |z_i| \leq 1\}.$$

Let $\lambda_1, \ldots, \lambda_n$ denote the successive minima of $\mathcal{C}$ with respect to $\mathbb{Z}^n$. The body $\mathcal{C}$ contains n linearly independent points from $\mathbb{Z}^n$ since $\mathcal{S}$ does. Therefore,

$$0 < \lambda_1 \leq \cdots \leq \lambda_n \leq 1. \tag{5.2}$$

Further, we have

$$\lambda_1 \cdots \lambda_n \geq D^{-1}. \tag{5.3}$$

To show this, take linearly independent vectors $\mathbf{y}_1, \ldots, \mathbf{y}_n \in \mathbb{Z}^n$ with $\mathbf{y}_i \in \lambda_i\mathcal{C}$ for $i = 1, \ldots, n$. Then $\lambda_i^{-1}\mathbf{y}_i \in \mathcal{C}$, i.e., $\lambda_i^{-1}\mathbf{y}_i = \sum_{j=1}^{m} z_{ij}\mathbf{x}_j$ for certain $z_{ij} \in \mathbb{R}$ with $\sum_{j=1}^{m} |z_{ij}| \leq 1$. In view of (5.1) this implies

$$\begin{aligned}(\lambda_1 \cdots \lambda_n)^{-1} &\leq |\det(\lambda_1^{-1}\mathbf{y}_1, \ldots, \lambda_n^{-1}\mathbf{y}_n)| \\ &\leq \sum_{j_1=1}^{m} \cdots \sum_{j_n=1}^{m} |z_{1,j_1}| \cdots |z_{n,j_n}| \cdot |\det(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_n})| \leq D\,,\end{aligned}$$

which is (5.3).

To $\mathcal{C}$ we associate the vector norm on $\mathbb{R}^n$ given by

$$||\mathbf{x}|| := \min\{\lambda : \ \mathbf{x} \in \lambda\mathcal{C}\}.$$

According to a result of Schlickewei [13] (p. 176, Prop. 4.2), the lattice $\mathbb{Z}^n$ has a basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ such that

$$||\mathbf{x}|| \geq 4^{-n} \max(|z_1| \cdot ||\mathbf{e}_1||, \ldots, |z_n| \cdot ||\mathbf{e}_n||) \quad \text{for } \mathbf{x} \in \mathbb{Z}^n, \tag{5.4}$$

where $z_1, \ldots, z_n$ are the integers determined by $\mathbf{x} = \sum_{i=1}^{n} z_i\mathbf{e}_i$. Assuming $||\mathbf{e}_1|| \leq \cdots \leq ||\mathbf{e}_n||$ as we may, we have

$$||\mathbf{e}_i|| \geq \lambda_i \quad \text{for } i = 1, \ldots, n. \tag{5.5}$$

Since $\mathcal{S} \subseteq \mathcal{C}$ we have $||\mathbf{x}|| \leq 1$ for $\mathbf{x} \in \mathcal{S}$. So by (5.4), (5.5) we have for all $\mathbf{x} \in \mathcal{S}$,

$$|z_i| \leq 4^n\lambda_i^{-1} \quad \text{for } i = 1, \ldots, n. \tag{5.6}$$

Now since the mapping $\mathbf{x} \mapsto (z_1, \ldots, z_n)$ is a linear isomorphism from $\mathbb{Z}^n$ to itself, it suffices to prove that the set of vectors $\mathbf{z} = (z_1, \ldots, z_n) \in \mathbb{Z}^n$ with (5.6) is contained in the union of not more than $100^n \cdot D^{\frac{1}{n-1}}$ proper linear subspaces of $\mathbb{Q}^n$.

We construct a collection of $(n-1)$-dimensional linear subspaces of $\mathbb{Q}^n$ whose union contains the set of vectors with (5.6), or rather a set of linear forms with integer coefficients such that each vector $\mathbf{z} \in \mathbb{Z}^n$ with (5.6) is a zero of at least one of these forms. We first determine an index s such that λ_t is not too small for $t > s$. Let $s \in \{0, \ldots, n-2\}$ be the index for which $(\lambda_{s+1} \cdots \lambda_n)^{\frac{1}{n-s-1}}$ is maximal. From (5.3) it follows that

$$(\lambda_{s+1} \cdots \lambda_n)^{\frac{1}{n-s-1}} \ge (\lambda_1 \cdots \lambda_n)^{\frac{1}{n-1}} \ge D^{-\frac{1}{n-1}} . \tag{5.7}$$

Moreover, we have

$$\lambda_t \ge (\lambda_{s+1} \cdots \lambda_n)^{\frac{1}{n-s-1}} \quad \text{for } t = s+1, \ldots, n. \tag{5.8}$$

Indeed, for $s = n-2$ this follows at once from (5.2). Suppose $s < n-2$. Then from the definition of s it follows that

$$(\lambda_{s+1} \cdots \lambda_n)^{\frac{1}{n-s-1}} \ge (\lambda_{s+2} \cdots \lambda_n)^{\frac{1}{n-s-2}} .$$

This implies $\lambda_{s+1}^{\frac{1}{n-s-2}} \ge (\lambda_{s+1} \cdots \lambda_n)^{\frac{1}{n-s-2} - \frac{1}{n-s-1}}$, whence $\lambda_{s+1} \ge (\lambda_{s+1} \cdots \lambda_n)^{\frac{1}{n-s-1}}$, and this certainly implies (5.8).

Using an argument similar to the proof of Siegel's lemma, we show that for each vector $\mathbf{z}$ with (5.6), there is a non-zero vector $\mathbf{c} = (c_{s+1}, \ldots, c_n) \in \mathbb{Z}^{n-s}$ with

$$c_{s+1} z_{s+1} + \cdots + c_n z_n = 0, \tag{5.9}$$

$$|c_i| \le \lambda_i \cdot \Big(\frac{n \cdot 4^n + 1}{\lambda_{s+1} \cdots \lambda_n} \Big)^{\frac{1}{n-s-1}} \quad \text{for } i = s+1, \ldots, n. \tag{5.10}$$

Put

$$B := \Big(\frac{n \cdot 4^n + 1}{\lambda_{s+1} \cdots \lambda_n} \Big)^{\frac{1}{n-s-1}} .$$

Consider all vectors $\mathbf{c} = (c_{s+1}, \ldots, c_n) \in \mathbb{Z}^{n-s}$ with

$$0 \le c_i \le \lambda_i B \quad \text{for } i = s+1, \ldots, n. \tag{5.11}$$

Let $\mathbf{z} = (z_1, \ldots, z_n) \in \mathbb{Z}^n$ be a vector with (5.6), and suppose that $z_i > 0$ for exactly r indices $i \in \{s+1, \ldots, n\}$, where $r \ge 0$. Then for vectors $\mathbf{c} \in \mathbb{Z}^{n-s}$ with (5.11) we have

$$-(n-s-r)4^nB \le c_{s+1}z_{s+1}+\cdots+c_nz_n \le r\cdot 4^nB\ .$$

So the number of possible values for $c_{s+1}z_{s+1}+\cdots+c_nz_n$ is at most

$$[r\cdot 4^nB]+[(n-s-r)\cdot 4^nB]+1\le [(n-s)\cdot 4^nB]+1\ .$$

Further, the number of vectors $\mathbf{c}\in\mathbb{Z}^{n-s}$ with (5.11) is equal to

$$\prod_{i=s+1}^{n}([\lambda_iB]+1)\ .$$

By the choice of B, this number is larger than

$$\lambda_{s+1}\cdots\lambda_nB^{n-s}=\lambda_{s+1}\cdots\lambda_nB^{n-s-1}\cdot B=(n\cdot 4^n+1)B\ge [(n-s)4^nB]+1\ ,$$

noting that by (5.2) we have $B\ge 1$. Therefore, there are two different vectors $\mathbf{c}'=(c'_{s+1},\dots,c'_n)$, $\mathbf{c}''=(c''_{s+1},\dots,c''_n)\in\mathbb{Z}^{n-s}$ with $0\le c'_i,\ c''_i\le\lambda_iB$ for $i=s+1,\dots,n$ and $c'_{s+1}z_{s+1}+\cdots+c'_nz_n=c''_{s+1}z_{s+1}+\cdots+c''_nz_n$. Now clearly, the vector $\mathbf{c}:=\mathbf{c}'-\mathbf{c}''$ is non-zero and satisfies (5.9), (5.10).

For each non-zero $\mathbf{c}\in\mathbb{Z}^{n-s}$, (5.9) defines a proper linear subspace of $\mathbb{Q}^n$. By estimating from above the number of vectors $\mathbf{c}\in\mathbb{Z}^{n-s}$ with (5.10), we conclude that the set of vectors $\mathbf{z}\in\mathbb{Z}^n$ with (5.6) is contained in the union of at most

$$\begin{aligned}&\prod_{i=s+1}^{n}\Big\{2\lambda_i\Big(\frac{n\cdot 4^n+1}{\lambda_{s+1}\cdots\lambda_n}\Big)^{\frac{1}{n-s-1}}+1\Big\}\\
&\qquad\le\prod_{i=s+1}^{n}\Big\{3\lambda_i\Big(\frac{n\cdot 4^n+1}{\lambda_{s+1}\cdots\lambda_n}\Big)^{\frac{1}{n-s-1}}\Big\}\quad\text{(by (5.8))}\\
&\qquad=3^{n-s}(n\cdot 4^n+1)^{\frac{n-s}{n-s-1}}(\lambda_{s+1}\cdots\lambda_n)^{1-\frac{n-s}{n-s-1}}\le 100^n\cdot(\lambda_{s+1}\cdots\lambda_n)^{-\frac{1}{n-s-1}}\\
&\qquad\le 100^n\cdot D^{\frac{1}{n-1}}\quad\text{(by (5.7))}\end{aligned}$$

proper linear subspaces of $\mathbb{Q}^n$. This proves Lemma 5. □

The next gap principle is a generalisation of Lemma 3.1 of Schmidt [16]. For $\mathbf{x}=(x_1,\dots,x_n)\in\mathbb{C}^n$ we put $||\mathbf{x}||:=\max(|x_1|,\dots,|x_n|)$.

Lemma 6. *Let P, Q, B be reals with*

$$P \ge 1, \quad Q \ge 1, \quad B \ge 1 \tag{5.12}$$

and let $M_1, \dots, M_n$ be linearly independent linear forms in $X_1, \dots, X_n$ with complex coefficients. Then the set of $\mathbf{x} \in \mathbb{Z}^n$ with

$$|M_1(\mathbf{x}) \cdots M_n(\mathbf{x})| \le |\det (M_1, \dots, M_n)| \cdot \frac{Q}{P} \,, \tag{5.13}$$

$$||\mathbf{x}|| \le B \tag{5.14}$$

is contained in the union of not more than

$$(100n^2)^n \cdot Q^{\frac{1}{n-1}} \cdot \left(1 + \frac{\log eB}{\log eP}\right)^{n-1}$$

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

Proof. Put

$$T := (n-1)\left(1 + \left[\frac{\log eB}{\log eP}\right]\right) \,. \tag{5.15}$$

We assume that

$$||M_i\,|| = 1 \quad \text{for } i = 1, \dots, n \tag{5.16}$$

(recall that $||M_i\,||$ is the maximum of the absolute values of the coefficients of M_i). This is no loss of generality, since (5.13) does not change if $M_1, \dots, M_n$ are replaced by constant multiples. As a consequence, the solutions $\mathbf{x}$ of (5.13), (5.14) satisfy

$$|M_i(\mathbf{x})| \le nB \quad \text{for } i = 1, \dots, n.$$

This implies that for every solution $\mathbf{x} \in \mathbb{Z}^n$ of (5.13), (5.14) either there is an index $j \in \{1, \dots, n-1\}$ such that

$$|M_j(\mathbf{x})| < (nB)^{1-n} \,, \tag{5.17}$$

or there are integers $c_1, \dots, c_{n-1}$ with

$$(nB)^{c_i/T} \le |M_i(\mathbf{x})| \le (nB)^{(c_i+1)/T} \quad \text{for } i = 1, \dots, n-1, \tag{5.18}$$

$$-(n-1)T \le c_i \le T - 1 \quad \text{for } i = 1, \dots, n-1. \tag{5.19}$$

(The linear form $M_n(\mathbf{x})$ does not have to be taken into consideration).

We consider first the solutions $\mathbf{x} \in \mathbb{Z}^n$ of (5.13), (5.14) which satisfy (5.17) for some fixed $j \in \{1, \ldots, n-1\}$. Let $\mathbf{x}_1 = (x_{11}, \ldots, x_{1n}), \ldots, \mathbf{x}_n$ be any such solutions. Let $M_i = \alpha_1 X_1 + \cdots + \alpha_n X_n$ with $||M_i|| = |\alpha_t|$, say. Then $|\alpha_t| = 1$ by (5.16) and so the absolute value of the determinant $\det(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ does not change if we replace its t-th column by $M_i(\mathbf{x}_1), \ldots, M_i(\mathbf{x}_n)$. Hence

$$\begin{aligned} |\det(\mathbf{x}_1, \ldots, \mathbf{x}_n)| &= \left| \det \begin{pmatrix} x_{11} & \cdots & M_i(\mathbf{x}_1) & \cdots & x_{1n} \\ \vdots & & \vdots & & \vdots \\ x_{n1} & \cdots & M_i(\mathbf{x}_n) & \cdots & x_{nn} \end{pmatrix} \right| \\ &\leq n! \cdot \prod_{\substack{k=1 \\ k \neq t}}^{n} \max_{j=1,\ldots,n} |x_{jk}| \cdot \max_{j=1,\ldots,n} |M_i(\mathbf{x}_j)| \\ &< n! \cdot B^{n-1} (nB)^{1-n} \leq 1 \ . \end{aligned}$$

Now since $\det(\mathbf{x}_1, \ldots, \mathbf{x}_n) \in \mathbb{Z}$, this implies $\det(\mathbf{x}_1, \ldots, \mathbf{x}_n) = 0$. Hence $\mathbf{x}_1, \ldots, \mathbf{x}_n$ lie in a single subspace of $\mathbb{Q}^n$. We infer that for each $i \in \{1, \ldots, n-1\}$, the set of solutions of (5.13), (5.14) satisfying (5.17) is contained in a single proper linear subspace of $\mathbb{Q}^n$.

We now deal with the solutions $\mathbf{x}$ of (5.13), (5.14) which satisfy (5.18) for some fixed tuple $c_1, \ldots, c_{n-1}$ with (5.19). Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be any such solutions. Then

$$|\det(\mathbf{x}_1, \ldots, \mathbf{x}_n)| = |\det(M_1, \ldots, M_n)|^{-1} \cdot |\det(M_i(\mathbf{x}_j))_{1 \leq i,j \leq n}| \ . \tag{5.20}$$

The last determinant is a sum of $n!$ terms

$$\pm M_1(\mathbf{x}_{\sigma(1)}) \cdots M_n(\mathbf{x}_{\sigma(n)})$$

where σ is a permutation of $1, \ldots, n$. Consider such a term with $\sigma(n) = j$. Using that by (5.18) we have

$$|M_k(\mathbf{x}_l)| \leq (nB)^{1/T} |M_k(\mathbf{x}_j)| \quad \text{for } k = 1, \ldots, n-1, \quad l \neq j$$

we get

$$\begin{aligned} |M_1(\mathbf{x}_{\sigma(1)}) \cdots M_n(\mathbf{x}_{\sigma(n)})| &\leq (nB)^{(n-1)/T} |M_1(\mathbf{x}_j) \cdots M_n(\mathbf{x}_j)| \\ &\leq (nB)^{(n-1)/T} |\det(M_1, \ldots, M_n)| \cdot \frac{Q}{P} \qquad \text{using (5.13)} \\ &\leq e \cdot (n/e)^{(n-1)/T} |\det(M_1, \ldots, M_n)| \cdot Q \qquad \text{using (5.15).} \end{aligned}$$

By inserting this into (5.20) and using $T \geq n-1$ we obtain

$$|\det(\mathbf{x}_1, \ldots, \mathbf{x}_n)| \leq n \cdot n! \cdot Q \ .$$

Now Lemma 5 implies that the set of solutions of (5.13), (5.14) satisfying (5.18) for some fixed $c_1, \dots, c_{n-1}$ is contained in the union of at most

$$100^n \cdot (n \cdot n! \cdot Q)^{\frac{1}{n-1}}$$

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

We have $n-1$ inequalities (5.17), each giving rise to a single subspace of $\mathbb{Q}^n$. Further, in view of (5.19) we have $(nT)^{n-1}$ systems of inequalities (5.18). Together with (5.15), this implies that the set of solutions $\mathbf{x} \in \mathbb{Z}^n$ of (5.13), (5.14) is contained in the union of at most

$$\begin{aligned} n-1 \;+\; 100^n (n \cdot n!)^{\frac{1}{n-1}} Q^{\frac{1}{n-1}} \;\cdot\; (n(n-1))^{n-1} \Big(1 + \frac{\log eB}{\log eP}\Big)^{n-1} \\ \le (100n^2)^n \cdot Q^{\frac{1}{n-1}} \cdot \Big(1 + \frac{\log eB}{\log eP}\Big)^{n-1} \end{aligned}$$

proper linear subspaces of $\mathbb{Q}^n$. This proves Lemma 6. □

Now let $F \in \mathbb{Z}[X_1, \dots, X_n]$ be the norm form of degree r with (1.2), (1.11), and $L_1, \dots, L_r$ the linear forms with (4.13)–(4.19) which we have fixed in Section 4.

Lemma 7. *For every solution $\mathbf{x}$ of (1.5) there are $i_1, \dots, i_n \in \{1, \dots, r\}$ such that $L_{i_1}, \dots, L_{i_n}$ are linearly independent and such that*

$$|L_{i_1}(\mathbf{x}) \cdots L_{i_n}(\mathbf{x})| \le |\det(L_{i_1}, \dots, L_{i_n})| \cdot \frac{h^{n/r}}{H^*(F)}. \tag{5.21}$$

Proof. We closely follow the proof of Lemma 3 of Schmidt [17]. For $\mathbf{x} = \mathbf{0}$ we have $L_i(\mathbf{x}) = 0$ for $i = 1, \dots, r$ and (5.21) is trivial. Let $\mathbf{x}$ be a non-zero solution of (1.5). Define linear forms

$$L_i' = \frac{|F(\mathbf{x})|^{1/r}}{|L_i(\mathbf{x})|} \cdot L_i \quad (i = 1, \dots, r).$$

From (4.13)–(4.15) it follows that $L_1', \dots, L_r'$ satisfy (1.8). Pick $i_1, \dots, i_n$ such that $|\det(L_{i_1}', \dots, L_{i_n}')|$ is maximal. Then from (1.5) and the definition of $H^*(F)$ it follows that

$$|L_{i_1}(\mathbf{x}) \cdots L_{i_n}(\mathbf{x})| = \frac{|\det(L_{i_1}, \dots, L_{i_n})|}{|\det(L_{i_1}', \dots, L_{i_n}')|} \cdot |F(\mathbf{x})|^{n/r} = \frac{|\det(L_{i_1}, \dots, L_{i_n})|}{\Delta(L_1', \dots, L_r')} \cdot |F(\mathbf{x})|^{n/r}$$

$$\leq |\det(L_{i_1}, \ldots, L_{i_n})| \cdot \frac{h^{n/r}}{H^*(F)} \ . \qquad \square$$

By combining Lemmata 7 and 6 we arrive at the following result for the small solutions of norm form inequality (1.5):

Proposition 1. *Let $P \geq 1,\ B \geq 1$. Then the set of solutions $\mathbf{x}$ of (1.5) with*

$$||\mathbf{x}|| \leq B \tag{5.22}$$

is contained in the union of at most

$$(300rn)^n \cdot \max\left(1, \left(\frac{h^{n/r}P}{H^*(F)}\right)^{\frac{1}{n-1}}\right) \cdot \left(1 + \frac{\log eB}{\log eP}\right)^{n-1}$$

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

Proof. From Lemma 6 with

$$Q = \max\left(1, \frac{h^{n/r}P}{H^*(F)}\right)$$

and from Lemma 7 we infer that for each tuple $\{i_1, \ldots, i_n\}$, the set of solutions of (1.5) satisfying (5.21) and (5.22) is contained in the union of not more than

$$(100n^2)^n \cdot \max\left(1, \left(\frac{h^{n/r}P}{H^*(F)}\right)^{\frac{1}{n-1}}\right) \cdot \left(1 + \frac{\log eB}{\log eP}\right)^{n-1}$$

proper linear subspaces of $\mathbb{Q}^n$. Now Proposition 1 follows, on noting that we have at most $\binom{r}{n}$ possibilities for $\{i_1, \ldots, i_n\}$ and that $(100n^2)^n \cdot \binom{r}{n} \leq (300rn)^n$. $\square$

6. The quantitative Subspace Theorem.

We recall a special case of the quantitative Subspace Theorem from [8] and then specialise it to a situation relevant for eq. (1.5). We must first introduce the Euclidean height which has been used in the statement of the quantitative Subspace theorem of [8].

Let $\mathbf{x} = (x_1, \ldots, x_n) \in \overline{\mathbb{Q}}^n$ with $\mathbf{x} \neq \mathbf{0}$. To define the height of $\mathbf{x}$, we choose a number field K containing $x_1, \ldots, x_n$. Let $d = [K : \mathbb{Q}]$ and let $\sigma_1, \ldots, \sigma_d$ denote the isomorphic embeddings of K into $\mathbb{C}$. Further, denote by $N_{K/\mathbb{Q}}(x_1, \ldots, x_n)$ the absolute norm of the fractional ideal in K generated by $x_1, \ldots, x_n$. Then the Euclidean height of $\mathbf{x}$ is defined by

$$H_2(\mathbf{x}) := \left(\frac{\prod_{i=1}^d \left(\sum_{j=1}^n |\sigma_i(x_j)|^2 \right)^{1/2}}{N_{K/\mathbb{Q}}(x_1, \ldots, x_n)} \right)^{1/d} .$$

It is easy to see that this is independent of the choice of K. Moreover, we have $H_2(\lambda \mathbf{x}) = H_2(\mathbf{x})$ for every non-zero $\lambda \in \overline{\mathbb{Q}}$. This implies $H_2(\mathbf{x}) \geq 1$. Note that

$$H_2(\mathbf{x}) = \frac{\left(|x_1|^2 + \cdots + |x_n|^2\right)^{1/2}}{\gcd(x_1, \ldots, x_n)} \qquad \text{for } \mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}^n \backslash \{\mathbf{0}\}.$$

For a non-zero linear form $L = \alpha_1 X_1 + \cdots + \alpha_n X_n$ with $\mathbf{a} = (\alpha_1, \ldots, \alpha_n) \in \overline{\mathbb{Q}}^n$ we put $H_2(L) := H_2(\mathbf{a})$.

The following result is a special case of Corollary 2.6 of [8]. Except for the better quantitative bound, this result is of the same nature as the first quantitative version of the Subspace Theorem, obtained by Schmidt [16].

Quantitative Subspace Theorem. *Let $0 < \delta \leq 1$ and let $M_1, \ldots, M_n$ be linearly independent linear forms in $X_1, \ldots, X_n$ such that*

the coefficients of $M_1, \ldots, M_n$ generate an algebraic number field of degree D. (6.1)

Then the set of $\mathbf{x} \in \mathbb{Z}^n$ with

$$|M_1(\mathbf{x}) \cdots M_n(\mathbf{x})| \leq |\det (M_1, \ldots, M_n)| \cdot H_2(\mathbf{x})^{-\delta}, \tag{6.2}$$

$$H_2(\mathbf{x}) \geq \max \left(n^{6n/\delta}, H_2(M_1), \ldots, H_2(M_n) \right) \tag{6.3}$$

is contained in the union of at most

$$16^{(n+5)^2} \cdot \delta^{-2n-3} \log(4D) \cdot \log\log(4D) \tag{6.4}$$

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

Let F be the norm form with (1.2), (1.11) and $L_1, \ldots, L_r$ the linear forms with (4.13)–(4.19) which we have fixed throughout the paper. Let K be the number field associated to F as in (1.2), N the finite, normal extension of $\mathbb{Q}$ introduced at the end of Section 4 containing the coefficients of $L_1, \dots, L_r$ and $d = [N : \mathbb{Q}]$. We want to apply the quantitative Subspace Theorem to any set of n linearly independent forms from $L_1, \dots, L_r$. We need the following estimates:

Lemma 8. *(i). $H_2(L_i) \leq \sqrt{n}(2n)^{n+1}H^*(F)$ for $i = 1, \dots, r$.*
(ii). For each linearly independent subset $\{L_{i_1}, \dots, L_{i_n}\}$ of $\{L_1, \dots, L_r\}$ we have

$$|\det(L_{i_1}, \dots, L_{i_n})| \geq (2H^*(F))^{1-\binom{r}{n}} .$$

Proof. (i). From (4.19), (4.23), (4.18) it follows that

$$\begin{aligned} H_2(L_i) &\leq \Big(\prod_{\sigma \in \mathrm{Gal}(N/\mathbb{Q})} ||\sigma(L_i)||_2 \Big)^{1/d} = \Big(\prod_{\sigma \in \mathrm{Gal}(N/\mathbb{Q})} ||L_{\sigma^*(i)}||_2 \Big)^{1/d} \\ &\leq \sqrt{n}(2n)^{n+1}H^*(F) . \end{aligned}$$

(ii). By Schmidt's result [17], p. 203 the semi-discriminant

$$D(F) := \prod_{(j_1, \dots, j_n)} |\det(L_{j_1}, \dots, L_{j_n})|$$

is a positive integer, where the product is taken over all ordered n-tuples $(j_1, \dots, j_n)$ for which $\det(L_{j_1}, \dots, L_{j_n}) \neq 0$. Further, by (4.17) we have for each such n-tuple that $|\det(L_{j_1}, \dots, L_{j_n})| \leq 2H^*(F)$. Denote by $\mathcal{J}$ the collection of all unordered subsets $\{j_1, \dots, j_n\}$ of $\{1, \dots, r\}$ for which $\det(L_{j_1}, \dots, L_{j_n}) \neq 0$. Then $\mathcal{J}$ has cardinality $\leq \binom{r}{n}$. Hence

$$\begin{aligned} |\det(L_{i_1}, \dots, L_{i_n})| &\geq \prod_{\{j_1, \dots, j_n\} \in \mathcal{J}} |\det(L_{j_1}, \dots, L_{j_n})| \cdot (2H^*(F))^{1-\binom{r}{n}} \\ &= |D(F)|^{1/n!}(2H^*(F))^{1-\binom{r}{n}} \geq (2H^*(F))^{1-\binom{r}{n}}. \end{aligned}$$ □

Below we have stated our basic tool for dealing with the large solutions of (1.5). As before, $||\mathbf{x}||$ denotes the maximum norm of $\mathbf{x}$. After the proof of Proposition 2, we will not use anymore Euclidean heights.

Proposition 2. *Let $L_{i_1}, \ldots, L_{i_n}$ be linearly independent linear forms among $L_1, \ldots, L_r$ and let $0 < \delta \leq 1$. Then the set of primitive $\mathbf{x} \in \mathbb{Z}^n$ with*

$$|L_{i_1}(\mathbf{x}) \cdots L_{i_n}(\mathbf{x})| \leq ||\mathbf{x}||^{-\delta}, \tag{6.5}$$

$$||\mathbf{x}|| \geq (eH^*(F))^{(4r)^{n+1}/\delta} \tag{6.6}$$

is contained in the union of not more than

$$16^{(n+6)^2} \cdot \delta^{-2n-3} \log(4r) \cdot \log\log(4r) \tag{6.7}$$

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

Proof. Inequality (6.2) does not change if the linear forms $M_1, \ldots, M_n$ are replaced by constant multiples. Therefore, we may replace (6.1) by the weaker condition that $M_1, \ldots, M_n$ are constant multiples of linear forms $M_1', \ldots, M_n'$ such that the coefficients of $M_1', \ldots, M_n'$ generate an algebraic number field of degree D.

Let $\mathbf{x} \in \mathbb{Z}^n$ be a primitive solution of (6.5), (6.6). Then

$$\begin{aligned}
&|L_{i_1}(\mathbf{x}) \cdots L_{i_n}(\mathbf{x})| \\
&\quad \leq (2H^*(F))^{1-\binom{r}{n}} \cdot \left(\sqrt{n} \cdot ||\mathbf{x}||\right)^{-3\delta/4} \quad \text{by (6.5), (6.6)} \\
&\quad \leq |\det(L_{i_1}, \ldots, L_{i_n})| \cdot H_2(\mathbf{x})^{-3\delta/4} \quad \text{by } H_2(\mathbf{x}) \leq \sqrt{n} \cdot ||\mathbf{x}|| \text{ and Lemma 8 (ii).}
\end{aligned}$$

Therefore, (6.2) holds with $L_{i_1}, \ldots, L_{i_n}$ replacing $M_1, \ldots, M_n$ and $3\delta/4$ replacing δ. Further, from Lemma 8 (i) and (6.6) it follows that $\mathbf{x}$ satisfies (6.3) with $L_{i_1}, \ldots, L_{i_n}$ and $3\delta/4$ replacing $M_1, \ldots, M_n$ and δ. Finally, from the construction of $L_1, \ldots, L_r$ in Section 4 it follows that for $i = 1, \ldots, r$, L_i is a constant multiple of a linear form with coefficients in $K^{(i)}$, where $K^{(i)}$ is a conjugate of K, whence has degree r. Therefore, there are constant multiples of $L_{i_1}, \ldots, L_{i_n}$ whose coefficients generate a number field of degree at most r^n. Now by applying the quantitative Subspace Theorem with $L_{i_1}, \ldots, L_{i_n}, 3\delta/4$ and r^n replacing $M_1, \ldots, M_n$, δ and D, we obtain that the set of primitive $\mathbf{x} \in \mathbb{Z}^n$ with (6.5) and (6.6) is contained in the union of at most

$$16^{(n+5)^2} (4/3\delta)^{2n+3} \log(4r^n) \log\log(4r^n)$$

proper linear subspaces of $\mathbb{Q}^n$. This is smaller than the quantity in (6.7). □

7. Reduction to the Subspace Theorem.

We will use some results from [5] and follow the arguments of Sections 7, 8 of [6]. Like before, the norm form F satisfies (1.2) and (1.11) and the linear forms $L_1, \dots, L_r$ satisfy (4.13)–(4.19). Further, $N \subset \mathbb{C}$ is the normal extension of $\mathbb{Q}$ chosen in Section 4 and for each $\sigma \in \mathrm{Gal}(N/\mathbb{Q})$, $(\sigma^*(1), \dots, \sigma^*(r))$ is the permutation of $(1, \dots, r)$ defined by (4.21). For $\sigma \in \mathrm{Gal}(N/\mathbb{Q})$, $I \subseteq \{1, \dots, r\}$ we write $\sigma^*(I) := \{\sigma^*(i) : i \in I\}$. We denote by ι the restriction to N of the complex conjugation on $\mathbb{C}$.

We define a hypergraph $\mathcal{H}$ as follows. The vertices of $\mathcal{H}$ are the indices $1, \dots, r$. Further, the edges of $\mathcal{H}$ are the sets I of cardinality ≥ 2 such that $\{L_i : i \in I\}$ is a linearly dependent set of linear forms (over N), whereas for each proper subset I' of I, the set $\{L_i : i \in I'\}$ is linearly independent. As usual, two vertices i, j of $\mathcal{H}$ are said to be connected if there is a sequence of edges $I_1, \dots, I_m$ of $\mathcal{H}$ such that $i \in I_1$, $I_j \cap I_{j+1} \neq \emptyset$ for $j = 1, \dots, m-1$ and $j \in I_m$. Denote by $C_1, \dots, C_t$ the connected components of $\mathcal{H}$. It follows at once from (4.23) that if $\{L_i : i \in I\}$ is linearly (in)dependent, then so is $\{L_i : i \in \sigma^*(I)\}$ for each $\sigma \in \mathrm{Gal}(N/\mathbb{Q})$. Hence if I is an edge of $\mathcal{H}$, then so is $\sigma^*(I)$ for each $\sigma \in \mathrm{Gal}(N/\mathbb{Q})$. This implies that for each connected component C_i of $\mathcal{H}$ and for each $\sigma \in \mathrm{Gal}(N/\mathbb{Q})$, $\sigma^*(C_i)$ is also a connected component of $\mathcal{H}$.

Lemma 9. *Either $\mathcal{H}$ is connected, or $\mathcal{H}$ has precisely two connected components, C_1 and C_2, say, and $\iota^*(C_1) = C_2$.*

Proof. We assume that the embedding $\alpha \mapsto \alpha^{(1)}$ is the identity on K and that the index $1 \in C_1$. Define the subfield J of N by

$$\mathrm{Gal}(N/J) = \{\sigma \in \mathrm{Gal}(N/\mathbb{Q}) : \sigma^*(C_1) = C_1\} .$$

By (4.21) we have

$$\mathrm{Gal}(N/J) \supseteq \{\sigma \in \mathrm{Gal}(N/\mathbb{Q}) : \sigma^*(1) = 1\} = \mathrm{Gal}(N/K);$$

so $J \subseteq K$. It is easy to see that for $\sigma \in \mathrm{Gal}(N/\mathbb{Q})$, the left coset $\sigma\mathrm{Gal}(N/J)$ is equal to $\{\tau \in \mathrm{Gal}(N/\mathbb{Q}) : \tau^*(C_1) = C_i\}$ where $\sigma^*(C_1) = C_i$. Moreover, (4.21) implies that for each index $j \in \{1, \dots, r\}$, there is a $\sigma \in \mathrm{Gal}(N/\mathbb{Q})$ with $\sigma^*(1) = j$ and so for each $i \in \{1, \dots, t\}$ there is a $\sigma \in \mathrm{Gal}(N/\mathbb{Q})$ with $\sigma^*(C_1) = C_i$. This implies that there are exactly t left cosets of $\mathrm{Gal}(N/J)$ in $\mathrm{Gal}(N/\mathbb{Q})$, and so

$$[J:\mathbb{Q}] = t\ . \tag{7.1}$$

Let V as before be the vector space defined by (1.3). We have

$$V^J = V\ . \tag{7.2}$$

This follows from some theory from Section 4, pp. 191–193 of [5]. By (4.16) of the present paper, the linear form L_i is proportional to $L_i' := \alpha_1^{(i)} X_1 + \cdots + \alpha_n^{(i)} X_n$ for $i = 1, \ldots, r$, so the hypergraph $\mathcal{H}$ does not change if in its definition, L_i is replaced by L_i'. For the space V and the field L on p. 192 of [5] we take $\mathbb{Q}^n$ and the field N of the present paper. In our situation, the quantity u defined by (4.5) of [5] is equal to 1, and the field K_1 defined by (4.3) on p. 192 of [5] is equal to J. Consider the injective map from K to N^r,

$$\psi : \xi \mapsto (\xi^{(1)}, \ldots, \xi^{(r)})\ .$$

Note that ψ maps V onto the space $\{(L_1'(\mathbf{x}), \ldots, L_r'(\mathbf{x})) : \mathbf{x} \in \mathbb{Q}^n\}$. Further, from Lemma 6, (ii) on pp. 192–193 of [5] it follows that ψ maps J onto

$$\begin{aligned}\Lambda(F) := \{\mathbf{c} = (c_1, \ldots, c_r) \in N^r :\quad & \text{for every } \mathbf{x} \in \mathbb{Q}^n \text{ there is an } \mathbf{y} \in \mathbb{Q}^n \\ & \text{with } L_i'(\mathbf{y}) = c_i L_i'(\mathbf{x}) \text{ for } i = 1, \ldots, r\}\ .\end{aligned}$$

Denoting the images of $(c_1, \ldots, c_r)$, $(L_1'(\mathbf{x}), \ldots, L_r'(\mathbf{x}))$, $(L_1'(\mathbf{y}), \ldots, L_r'(\mathbf{y}))$ under ψ^{-1} by λ, ξ, η, respectively, we obtain that J is the set of $\lambda \in K$ such that for every $\xi \in V$ there is an $\eta \in V$ with $\eta = \lambda\xi$. This implies (7.2).

By (7.2) and (1.11) we have that either $J = \mathbb{Q}$ in which case it follows from (7.1) that $\mathcal{H}$ is connected; or that J is an imaginary quadratic field, in which case (7.1) implies that $\mathcal{H}$ has precisely two connected components C_1 and C_2. Moreover, in this case we have that ι is not the identity on J, so $\iota^*(C_1) \neq C_1$, which implies $\iota^*(C_1) = C_2$. This completes the proof of Lemma 9. □

Let $\mathbf{x} \in \mathbb{Z}^n$. We write

$$u_i := L_i(\mathbf{x}) \quad (i = 1, \ldots, r), \qquad \mathbf{u} = (u_1, \ldots, u_r)\ .$$

From (4.22) it follows that for each $i \in \{1, \ldots, r\}$ we have

$$\prod_{\sigma \in \mathrm{Gal}(N/\mathbb{Q})} |u_{\sigma^*(i)}| = \prod_{\sigma \in \mathrm{Gal}(N/\mathbb{Q})} |L_{\sigma^*(i)}(\mathbf{x})| = |F(\mathbf{x})|^{d/r}\ , \tag{7.3}$$

where as before $d = [N : \mathbb{Q}]$. From (1.11) it follows that if $\mathbf{x} \neq \mathbf{0}$, then $F(\mathbf{x}) \neq 0$ and so $u_i \neq 0$ for $i = 1, \dots, r$. Further, (7.3) implies

$$\prod_{\sigma \in \mathrm{Gal}(N/\mathbb{Q})} |u_{\sigma^*(i)}| \geq 1 \quad \text{if } \mathbf{x} \neq \mathbf{0}. \tag{7.4}$$

For each subset I of $\{1, \dots, r\}$ we define a suitable height,

$$H_I(\mathbf{u}) := \Big(\prod_{\sigma \in \mathrm{Gal}(N/\mathbb{Q})} \max_{i \in I} |u_{\sigma^*(i)}| \Big)^{1/d}.$$

From (7.4) it follows at once that

$$H_I(\mathbf{u}) \geq 1 \quad \text{for each non-empty subset } I \text{ of } \{1, \dots, r\}. \tag{7.5}$$

Further, we have

$$H_{I_1 \cup I_2}(\mathbf{u}) \leq H_{I_1}(\mathbf{u}) \cdot H_{I_2}(\mathbf{u}) \quad \text{if } I_1 \cap I_2 \neq \emptyset. \tag{7.6}$$

For if $i \in I_1 \cap I_2$, then for each $\sigma \in \mathrm{Gal}(N/\mathbb{Q})$ we have

$$|u_{\sigma^*(i)}| \cdot \max_{j \in I_1 \cup I_2} |u_{\sigma^*(j)}| = \max_{j \in I_1 \cup I_2} |u_{\sigma^*(i)} \cdot u_{\sigma^*(j)}| \leq \max_{j \in I_1} |u_{\sigma^*(j)}| \cdot \max_{j \in I_2} |u_{\sigma^*(j)}|,$$

whence

$$\Big(\prod_{\sigma \in \mathrm{Gal}(N/\mathbb{Q})} |u_{\sigma^*(i)}| \Big)^{1/d} \cdot H_{I_1 \cup I_2}(\mathbf{u}) \leq H_{I_1}(\mathbf{u}) \cdot H_{I_2}(\mathbf{u}),$$

which together with (7.4) implies (7.6).

In what follows, we assume that the collection of edges of $\mathcal{H}$ is not empty. We deal with the cases that $\mathcal{H}$ is connected and that $\mathcal{H}$ has two connected components, simultaneously. Let C_1 be a connected component of $\mathcal{H}$ (so either the whole vertex set $\{1, \dots, r\}$ or one of the two components).

Lemma 10. *Let S be a maximal subset of C_1 such that $\{L_j : j \in S\}$ is linearly independent. Then for each non-zero $\mathbf{x} \in \mathbb{Z}^n$ there is an edge I of $\mathcal{H}$ contained in C_1 such that*

$$H_S(\mathbf{u}) \leq H_I(\mathbf{u})^{n-1}. \tag{7.7}$$

Proof. Fix $\mathbf{x} \in \mathbb{Z}^n$, $\mathbf{x} \neq \mathbf{0}$. We use the argument on p. 208 of [5]. We have

$$L_i = \sum_{j \in D_i} c_{ij} L_j \quad \text{for } i \in C_1,$$

where $D_i \subseteq S$ and where $c_{ij} \neq 0$ for $j \in D_i$. As has been explained on p. 208 of [5], for each subset D of S with $D \neq \emptyset$, $D \subsetneq S$, there is an i such that $D_i \cap D \neq \emptyset$, $D_i \not\subset D$. This implies in particular that there is an $i_1 \in C_1$ such that $1 \in D_{i_1}$ and D_{i_1} has cardinality ≥ 2. If D_{i_1} is not equal to the whole set S, then choose i_2 such that $D_{i_1} \cap D_{i_2} \neq \emptyset$ and $D_{i_2} \not\subset D_{i_1}$. If $D_{i_1} \cup D_{i_2} \subsetneq S$, then choose i_3 such that $D_{i_3} \cap (D_{i_1} \cup D_{i_2}) \neq \emptyset$ and $D_{i_3} \not\subset D_{i_1} \cup D_{i_2}$. Continuing like this, we obtain sets $D_{i_1}, \ldots, D_{i_s}$, such that

$$\begin{aligned}
&S = D_{i_1} \cup \cdots \cup D_{i_s}, \\
&D_{i_h} \cap (D_{i_1} \cup \cdots \cup D_{i_{h-1}}) \neq \emptyset, \quad D_{i_h} \not\subset D_{i_1} \cup \cdots \cup D_{i_{h-1}} \quad \text{for } h = 2, \ldots, s.
\end{aligned}$$

Assuming s with this property to be minimal, we have $s \leq |S| - 1 \leq n - 1$ since we started with a set D_{i_1} of cardinality ≥ 2 and each newly chosen set D_{i_h} adds at least one element to the union of the sets chosen previously. Now clearly, $I_h := \{i_h\} \cup D_{i_h}$ is an edge of $\mathcal{H}$ for $h = 1, \ldots, s$, and we have

$$S \subset I_1 \cup \cdots \cup I_s, \quad I_h \cap (I_1 \cup \cdots \cup I_{h-1}) \neq \emptyset \text{ for } h = 2, \ldots, s.$$

Together with (7.6) this implies $H_S(\mathbf{u}) \leq H_{I_1}(\mathbf{u}) \cdots H_{I_s}(\mathbf{u})$. Now this fact and (7.5) imply that there is an edge I of $\mathcal{H}$ such that

$$H_S(\mathbf{u}) \leq H_I(\mathbf{u})^s \leq H_I(\mathbf{u})^{n-1}.$$

This proves Lemma 10. □

Lemma 11. *Suppose that $\mathcal{H}$ has edges. Then for every non-zero $\mathbf{x} \in \mathbb{Z}^n$, there is an edge I of $\mathcal{H}$ such that*

$$||\mathbf{x}|| \leq n! \cdot (2n)^{n^2-1} \cdot 2^{\binom{r}{n}-1} \cdot H^*(F)^{\binom{r}{n}+n-2} \cdot H_I(\mathbf{u})^{n-1}. \tag{7.8}$$

Proof. Let S be the set from Lemma 10 and let $\mathbf{x} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$. Choose $\sigma \in \mathrm{Gal}(N/\mathbb{Q})$ such that

$$\max_{i \in S} |u_{\sigma^*(i)}|$$

is minimal. Then by Lemma 10, there is an edge I of $\mathcal{H}$ such that

$$\max_{i\in S}|u_{\sigma^*(i)}|\leq H_I(\mathbf{u})^{n-1}. \tag{7.9}$$

We show that $\max_{i\in S}|u_{\sigma^*(i)}|=\max_{i\in S'}|u_i|$ where S' is a set of cardinality n such that $\{L_j:j\in S'\}$ is linearly independent and then we estimate $||\mathbf{x}||$ from above in terms of $\max_{i\in S'}|u_i|$.

Define the set S' by $S':=\sigma^*(S)$ if $\mathcal{H}$ is connected and $S':=\sigma^*(S)\cup\iota^*\sigma^*(S)$ if $\mathcal{H}$ has two connected components, where ι denotes the complex conjugation on N. First suppose that $\mathcal{H}$ is connected. Then $\{L_j:j\in\sigma^*(S)\}$ is linearly independent and it spans $\{L_1,\dots,L_r\}$. By (1.11), we have that rank $\{L_1,\dots,L_r\}=n$. Hence $S'=\sigma^*(S)$ has cardinality n. Now suppose that $\mathcal{H}$ has two connected components. Then by Lemma 9 these connected components are $\sigma^*(C_1)$ and $\iota^*\sigma^*(C_1)$. We have that $\{L_j:j\in\sigma^*(S)\}$ is linearly independent and spans $\{L_j:j\in\sigma^*(C_1)\}$ and that $\{L_j:j\in\iota^*\sigma^*(S)\}$ is linearly independent and spans $\{L_j:j\in\iota^*\sigma^*(C_1)\}$. Therefore, $\{L_j:j\in S'\}$ spans $\{L_1,\dots,L_r\}$. Suppose that $\{L_j:j\in S'\}$ is linearly dependent. Then S' contains an edge of $\mathcal{H}$. This edge is contained in one of the two connected components, so either in $S'\cap\sigma^*(C_1)=\sigma^*(S)$ or in $S'\cap\iota^*\sigma^*(C_1)=\iota^*\sigma^*(S)$. But this is impossible, since both sets $\{L_j:j\in\sigma^*(S)\}$ and $\{L_j:j\in\iota^*\sigma^*(S)\}$ are linearly independent. It follows that also in the second case, $\{L_j:j\in S'\}$ is linearly independent and S' has cardinality n.

If $\mathcal{H}$ is connected then clearly $\max_{i\in S'}|u_i|=\max_{i\in S}|u_{\sigma^*(i)}|$. If $\mathcal{H}$ has two connected components, then it follows from (4.24) and $u_j=L_j(\mathbf{x})$ for $j=1,\dots,r$ that $u_{\iota^*\sigma^*(j)}=\overline{u_{\sigma^*(j)}}$ for $j\in S$, hence also $\max_{i\in S'}|u_i|=\max_{i\in S}|u_{\sigma^*(i)}|$. By inserting this into (7.9) we get in both cases,

$$\max_{i\in S'}|u_i|\leq H_I(\mathbf{u})^{n-1}\ .$$

Therefore, (7.8) follows immediately, once we have shown that

$$||\mathbf{x}||\leq n!\cdot(2n)^{n^2-1}\cdot 2^{\binom{r}{n}-1}\cdot H^*(F)^{\binom{r}{n}+n-2}\cdot\max_{i\in S'}|u_i|\,. \tag{7.10}$$

Suppose that $S'=\{i_1,\dots,i_n\}$. Let A be the matrix, whose j-th column consists of the coefficients of L_{i_j}, for $j=1,\dots,n$. Then $\mathbf{x}=(u_{i_1},\dots,u_{i_n})A^{-1}$. The elements of A^{-1} are $\pm\Delta_{ij}/\Delta$, where Δ_{ij} is the determinant of the $(n-1)\times(n-1)$-matrix obtained by removing the j-th row and i-th column from A, and where $\Delta=\det(L_{i_1},\dots,L_{i_n})$.

Hence

$$||\mathbf{x}|| = |\max_{k=1,\dots,n} \sum_{j=1}^{n} u_{i_j} \cdot \Delta_{jk}/\Delta| \leq n \cdot \max_{j,k} |\Delta_{jk}| \cdot |\Delta|^{-1} \cdot \max_{i \in S'} |u_i| \ . \tag{7.11}$$

We have

$$|\Delta_{jk}| \leq (n-1)! \cdot \left(\max_k ||L_k||\right)^{n-1} \leq (n-1)! \cdot (2n)^{n^2-1} \cdot H^*(F)^{n-1} \quad \text{by (4.18)},$$
$$|\Delta|^{-1} \leq (2H^*(F))^{\binom{r}{n}-1} \quad \text{by Lemma 8, (ii).}$$

By inserting these inequalities into (7.11) we obtain (7.10). This proves Lemma 11.□

We finally arrive at:

Proposition 3. *Suppose that $\mathcal{H}$ has edges. Then for every solution $\mathbf{x} \in \mathbb{Z}^n$ of (1.5) with $\mathbf{x} \neq \mathbf{0}$, there are linearly independent linear forms $L_{i_1}, \dots, L_{i_n}$ among $L_1, \dots, L_r$ such that*

$$|L_{i_1}(\mathbf{x}) \cdots L_{i_n}(\mathbf{x})| \leq C \cdot ||\mathbf{x}||^{-1/(n-1)} \ , \tag{7.12}$$
$$\textit{with } C := \left(n! \cdot (2n)^{n^2-1} \cdot 2^{\binom{r}{n}-1} \cdot H^*(F)^{\binom{r}{n}+n-2}\right)^{\frac{1}{n-1}} \cdot h^{(n+1)/r} . \tag{7.13}$$

Proof. Fix a non-zero solution $\mathbf{x} \in \mathbb{Z}^n$ of (1.5). Choose linearly independent linear forms $L_{i_1}, \dots, L_{i_n}$ from $L_1, \dots, L_r$ such that the quantity

$$U := |L_{i_1}(\mathbf{x}) \cdots L_{i_n}(\mathbf{x})|$$

is minimal.

Let I be the edge from Lemma 11. Suppose that I has cardinality t. Each linear form from $\{L_j : j \in I\}$ is linearly dependent on the other forms in this set, and these other forms are linearly independent. Hence $\{L_j : j \in I\}$ has rank $t-1$. Choose a subset T of $\{1, \dots, r\}$ of cardinality $n-t+1$ such that $\{L_j : j \in T \cup I\}$ has rank n. Then for each $i \in I$, the set of linear forms $\{L_j : j \in T \cup I \backslash \{i\}\}$ is linearly independent and has cardinality n.

Pick $\sigma \in \mathrm{Gal}(N/\mathbb{Q})$. Choose $i_\sigma \in I$ such that $|u_{\sigma^*(i_\sigma)}| = \max_{i \in I} |u_{\sigma^*(i)}|$. Then the set $\{L_j : j \in \sigma^*(T \cup I \backslash \{i_\sigma\})\}$ is linearly independent and has cardinality n. So by the definition of U we have

$$U \leq \prod_{j \in \sigma^*(T \cup I \setminus \{i_\sigma\})} |u_j| = \prod_{j \in T \cup I} |u_{\sigma^*(j)}| \cdot \Big(\max_{i \in I} |u_{\sigma^*(i)}| \Big)^{-1} .$$

It follows that U is bounded above by the geometric mean of the terms at the right-hand side for all $\sigma \in \mathrm{Gal}(N/\mathbb{Q})$. By (7.3), the fact that $T \cup I$ has cardinality $n+1$ and the definition of $H_I(\mathbf{u})$, this geometric mean is equal to $|F(\mathbf{x})|^{(n+1)/r} H_I(\mathbf{u})^{-1}$. Hence

$$U \leq h^{(n+1)/r} \cdot H_I(\mathbf{u})^{-1} .$$

By inserting (7.8) we get Proposition 3. □

8. Proof of Theorem 2.

We combine Propositions 1,2 and 3. We recall that $n \geq 2$. It clearly suffices to show that the set of *primitive* solutions of (1.5) is contained in the union of not more than

$$A := (16r)^{(n+6)^2} \cdot \max\left(1, \Big(\frac{h^{n/r} P}{H^*(F)}\Big)^{\frac{1}{n-1}}\right) \cdot \Big(1 + \frac{\log(eh \cdot H^*(F))}{\log eP}\Big)^{n-1} \tag{8.1}$$

proper linear subspaces of $\mathbb{Q}^n$. We divide the primitive solutions $\mathbf{x} \in \mathbb{Z}^n$ of (1.5) into

$$\text{large solutions, i.e., with } ||\mathbf{x}|| \geq e^{-1}\Big(eh \cdot H^*(F)\Big)^{(4r)^{n+2}},$$
$$\text{small solutions, i.e., with } ||\mathbf{x}|| < e^{-1}\Big(eh \cdot H^*(F)\Big)^{(4r)^{n+2}}.$$

We first deal with the large solutions. First suppose that the hypergraph $\mathcal{H}$ defined in Section 7 has no edges. Then by Lemma 9, the hypergraph $\mathcal{H}$ has two connected components $\{1\}$ and $\{2\}$ with $2 = \iota^*(1)$. This means that $n = 2$, $r = 2$, that the linear forms L_1, L_2 are linearly independent and that $L_2 = \overline{L_1}$ in view of (4.24). Let $\mathbf{x}$ be a solution of (1.5). Then $|u_1| = |u_2| \leq h^{1/2}$ where $u_i = L_i(\mathbf{x})$. Further, we have $\mathbf{x} = (u_1, u_2)A^{-1}$, where A is the 2×2-matrix whose i-th column consists of the coefficients of L_i. Now by (4.18) and (4.17), the elements of A^{-1} have absolute values at most

$$\frac{\max(||L_1||, ||L_2||)}{|\det(L_1, L_2)|} \leq 4^3 H^*(F)/H^*(F) = 64.$$

Hence $||\mathbf{x}|| \leq 128h^{1/2}$. So (1.5) does not have large solutions.

Now assume that $\mathcal{H}$ does have edges. Let C be the quantity defined by (7.13). Then by Proposition 3, for every large solution $\mathbf{x}$ of (1.5) there are linearly independent linear forms $L_{i_1}, \ldots, L_{i_n}$ among $L_1, \ldots, L_r$ with

$$|L_{i_1}(\mathbf{x}) \cdots L_{i_n}(\mathbf{x})| \leq C \cdot ||\mathbf{x}||^{-1/(n-1)} \leq ||\mathbf{x}||^{-1/n} \,. \tag{8.2}$$

We apply Proposition 2 in Section 6 (with $\delta = 1/n$) to (8.2). Note that the large primitive solutions of (1.5) satisfy (6.6). Thus, on observing that for the set $\{i_1, \ldots, i_n\}$ we have at most $\binom{r}{n}$ possibilities, we obtain that the set of large primitive solutions of (1.5) is contained in the union not more than

$$\binom{r}{n} \cdot 16^{(n+6)^2} \cdot n^{2n+2} \log 4r \cdot \log\big(n \log 4r\big) < \frac{1}{2}A$$

proper linear subspaces of $\mathbb{Q}^n$, where A is given by (8.1).

We now deal with the small primitive solutions of (1.5) and to this end we apply Proposition 1 in Section 5. Taking $B := e^{-1}\Big(eh \cdot H^*(F)\Big)^{(4r)^{n+2}}$ and observing that

$$\Big(1 + \frac{\log eB}{\log eP}\Big)^{n-1} \leq (4r)^{(n+2)(n-1)}\Big(1 + \frac{\log(eh \cdot H^*(F))}{\log eP}\Big)^{n-1} ,$$

we infer that the set of small primitive solutions of (1.5) is contained in the union of not more than

$$\begin{aligned}(300rn)^n \cdot (4r)^{(n+2)(n-1)} \cdot \max\left(1, \Big(\frac{h^{n/r}P}{H^*(F)}\Big)^{\frac{1}{n-1}}\right) \cdot \Big(1 + \frac{\log(eh \cdot H^*(F))}{\log eP}\Big)^{n-1} \\ < \frac{1}{2}A\end{aligned}$$

proper linear subspaces of $\mathbb{Q}^n$. It follows that indeed, the set of all primitive solutions of (1.5) is contained in the union of not more than A proper linear subspaces of $\mathbb{Q}^n$. This proves Theorem 2. $\square$

References

[1] M. Bean, *Areas of Plane Regions Defined by Binary Forms,* Ph.D.thesis, Univ. of Waterloo, 1992.

[2] M. Bean and J.L. Thunder, *Isoperimetric inequalities for volumes associated with decomposable forms,* J. London Math. Soc. 54 (1996), 39–49.

[3] E. Bombieri and W.M. Schmidt, *On Thue's equation,* Invent. math. 73 (1983), 11–32

[4] J.W.S. Cassels, *An introduction to the geometry of numbers,* Grundlehren Math. Wiss., vol. 99, Springer Verlag, 1959.

[5] J.-H. Evertse, *Decomposable form equations with a small linear scattering,* J. reine angew. Math. 432 (1992), 177–217.

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

[7] J.-H. Evertse and K. Győry, *Thue inequalities with a small number of solutions,* in: *The mathematical heritage of C.F. Gauss,* G.M. Rassias, ed., World Scientific, Singapore, 1991, 204–224.

[8] J.-H. Evertse and H.P. Schlickewei, *A quantitative version of the absolute Subspace Theorem,* to appear.

[9] R.S. de Jong, *On p-adic norm form inequalities,* Master's thesis, Leiden, 1999.

[10] K. Mahler, *Zur Approximation algebraischer Zahlen III (Über die mittlere Anzahl der Darstellungen großer Zahlen durch binäre Formen),* Acta Math. 62 (1933), 91–166.

[11] J. Mueller and W.M. Schmidt, *Thue's equation and a conjecture of Siegel,* Acta Math. 160 (1988), 207–247.

[12] K. Ramachandra, *A lattice point problem for norm forms in several variables,* J. Number Theory 1 (1969), 534–555.

[13] H.P. Schlickewei, *Lower Bounds for Heights on Finitely Generated Groups,* Monatsh. Math. 123 (1997), 171–178.

[14] W.M. Schmidt, *Linearformen mit algebraischen Koeffizienten II,* Math. Ann. 191 (1971), 1–20.

[15] —, *Thue equations with few coefficients,* Trans. Amer. Math. Soc. 303 (1987), 241–255.

[16] —, *The subspace theorem in Diophantine approximations,* Compos. Math. 69 (1989), 121–173.

[17] —, *The number of solutions of norm form equations,* Trans. Amer. Math. Soc. 317 (1990), 197–227.

[18] J.L. Thunder, *The Number of Solutions to Cubic Thue Inequalities,* Acta Arith. 66 (1994), 237–243.

[19] —, *On Thue Inequalities and a Conjecture of Schmidt,* J. Number Theory 52 (1995), 319–328.

[20] —, *On cubic Thue inequalities and a result of Mahler,* Acta Arith. 83 (1998), 31–44.

[21] —, *Decomposable form inequalities. Preprint, 1999.*

Address of the author:
Universiteit Leiden
Mathematisch Instituut
Postbus 9512, 2300 RA Leiden
The Netherlands
email evertse@math.leidenuniv.nl