THE NUMBER OF SOLUTIONS OF DECOMPOSABLE FORM EQUATIONS.

Jan-Hendrik Evertse

University of Leiden, Department of Mathematics and Computer Science
P.O. Box 9512, 2300 RA Leiden, The Netherlands, email evertse@wi.leidenuniv.nl

§1. Introduction.

Let $F(X,Y) = a_0X^r + a_1X^{r-1}Y + ... + a_rY^r$ be a binary form in $\mathbb{Z}[X,Y]$ and let $p_1, ..., p_t$ be distinct prime numbers. F is a product of linear forms $\prod_{i=1}^r(\alpha_i X + \beta_i Y)$ with algebraic α_i, β_i. From results of Thue [23] and Mahler [11] it follows that if among the linear forms $\alpha_i X + \beta_i Y$ there are three pairwise non-proportional ones, then the equation

$$(1.1) \qquad |F(x,y)| = p_1^{z_1} ... p_t^{z_t} \quad \text{in} \quad x, y, z_1, ..., z_t \in \mathbb{Z} \text{ with } \gcd(x,y) = 1$$

has only finitely many solutions. The Diophantine approximation techniques of Thue and Mahler and improvements by Siegel, Dyson, Roth and Bombieri made it possible to derive good explicit upper bounds for the number of solutions of (1.1). The best such upper bound to date, due to Bombieri [1] is $2 \times (12r)^{12(t+1)}$ (Bombieri assumed that F is irreducible and $r \geq 6$ which was not essential in his proof). For $t = 0$, i.e. $|F(x,y)| = 1$ in $x, y \in \mathbb{Z}$, Bombieri and Schmidt [2] derived the upper bound constant$\times r$ which is best possible in terms of r.

In this paper we consider generalisations of (1.1) where instead of a binary form we take a decomposable form in $n \geq 3$ variables, that is, a homogeneous polynomial $F(\mathbf{X})$ with integer coefficients which is expressible as a product of linear forms with algebraic coefficients, i.e. $F(\mathbf{X}) = \prod_{i=1}^r(\alpha_{i1}X_1 + ... + \alpha_{in}X_n)$. More precisely, we consider *decomposable form equations*

$$(1.2) \qquad |F(\mathbf{x})| = p_1^{z_1} ... p_t^{z_t} \quad \text{in } \mathbf{x} \in \mathbb{Z}^n, z_1, ..., z_t \in \mathbb{Z} \text{ with } \mathbf{x} \text{ primitive,}$$

where $\mathbf{x} = (x_1, ..., x_n) \in \mathbb{Z}^n$ is called primitive if gcd $(x_1, ..., x_n) = 1$. Schmidt initiated the study of decomposable form equations and after that several qualitative and quantitative finiteness results on (1.2) have been derived analogous to those mentioned above for eq. (1.1). Below we give an overview. For instance, from results of Schmidt, Schlickewei and Győry it follows that for 'non-degenerate' decomposable forms F, there exists an explicit upper bound, depending on $r = \deg F$, n and t only, for the number of solutions of (1.2), with a doubly exponential dependence on n. In this paper we improve this bound to $2 \times (2^{33}r^2)^{n^3(t+1)}$. Our result as well as the previous ones all go back to Schmidt's Subspace theorem.

The probably best known type of a decomposable form equation (apart from (1.1)) is a *norm form equation*, that is, an equation (1.2) in which F is a norm form, i.e.

$$(1.3) \qquad F(\mathbf{X}) = c \cdot N_{M/\mathbb{Q}}(\alpha_1 X_1 + ... + \alpha_n X_n) = c \prod_{i=1}^{[M:\mathbb{Q}]} (\alpha_1^{(i)} X_1 + ... + \alpha_n^{(i)} X_n),$$

where $M = \mathbb{Q}(\alpha_1, ..., \alpha_n)$ is an algebraic number field, c is a non-zero integer, and $x \mapsto x^{(i)} (i = 1, ..., [M : \mathbb{Q}])$ are the isomorphic embeddings of M into $\mathbb{C}$. We consider more generally decomposable form equations rather than norm form equations since several problems can be reduced to decomposable form equations which are not norm form equations; for instance the S-unit equation

$$(1.4) \qquad x_0 + ... + x_n = 0 \text{ in } \mathbf{x} = (x_0, ..., x_n) \in \mathbb{Z}^{n+1}$$
$$\text{with } \mathbf{x} \text{ primitive, } |x_1 ... x_n| \text{ composed of } p_1, ..., p_t$$

can be reduced to eq. (1.2) with $F(\mathbf{X}) = X_1 \cdots X_n (X_1 + ... + X_n)$.

We remark that for every solution $(\mathbf{x}, z_1, ..., z_t)$ of (1.2) we have $F(\mathbf{x}) \in R^*$, where R is the ring $\mathbb{Z}[(p_1 ... p_t)^{-1}]$ and R^* is the unit group of R. We consider the more general decomposable form equation over number fields,

$$(1.5) \qquad F(\mathbf{x}) \in \mathcal{O}_S^* \text{ in } \mathbf{x} \in \mathcal{O}_S^n,$$

where $\mathcal{O}_S$ is the ring of S-integers for some finite set of places S on some algebraic number field K, $\mathcal{O}_S^*$ is the unit group of $\mathcal{O}_S$, i.e. the group of S-units, and where F is a decomposable form in n variables with coefficients from $\mathcal{O}_S$. We recall that $\mathcal{O}_S = \mathcal{O}_K[(\wp_1 ... \wp_t)^{-1}]$, where $\mathcal{O}_K$ is the ring of integers of K and $\wp_1, ..., \wp_t$ are the prime ideals of $\mathcal{O}_K$, i.e. finite places, belonging to S. Further, $\mathcal{O}_S^* = \{x \in K : (x) = \wp_1^{z_1} ... \wp_t^{z_t} \text{ for } z_1, ..., z_t \in \mathbb{Z}\}$.
Obviously, if $\mathbf{x}$ is a solution of (1.5), then so is $\epsilon \mathbf{x}$ for every $\epsilon \in \mathcal{O}_S^*$. Therefore, we will give upper bounds for the maximal number of *$\mathcal{O}_S^*$-cosets* $\{\epsilon \mathbf{x} : \epsilon \in \mathcal{O}_S^*\}$ contained in the set of solutions of (1.5).

Below we give an overview of previous results on decomposable form equations, norm form equations, and S-unit equations, and then state our improvements. These improvements are consequences of our main result, stated in §2.

Overview of previous results.

In 1972, Schmidt [19, 20] was the first to derive non-trivial finiteness results for norm form equations. He formulated a non-degeneracy condition for norm forms $F \in \mathbb{Z}[X_1, ..., X_n]$ and showed that the equation

$$F(\mathbf{x}) = b \text{ in } \mathbf{x} \in \mathbb{Z}^n$$

has only finitely many solutions for every non-zero integer b if and only if F is non-degenerate. Schmidt derived this from his own higher dimensional generalisation of Roth's theorem, the Subspace Theorem [20].

Schmidt's result on norm form equations was generalised successively by Schlickewei in 1977 [14] to p-adic norm form equations (1.2) and by Laurent in 1984 [10] to norm form equations (1.5) over number fields; Laurent derived this from a result of his on linear tori conjectured by Lang. In 1988, Győry and the author [8] proved an extension of Laurent's result to arbitrary decomposable form equations (1.5) over number fields. Later, Gaál, Győry and the author [7] proved the following 'semi-quantitative' refinement of this: given a finite extension L of K, there exists a uniform bound C, depending only on n, S and L, such that for every 'non-degenerate' decomposable form $F \in \mathcal{O}_S[X_1, ..., X_n]$ which can be factored into linear forms over L, the number of $\mathcal{O}_S^*$-cosets of solutions of (1.5) is at most C; however, their method of proof did not enable an explicit computation of C. We mention that all these results follow from Schmidt's Subspace theorem and its p-adic generalisation by Schlickewei [13].

In 1989, Schmidt [21] made another breakthrough by proving a quantitative version of his Subspace theorem from [20] and then deriving an explicit upper bound for the number of solutions of norm form equations of the type $|F(\mathbf{x})| = 1$ in $\mathbf{x} \in \mathbb{Z}^n$, where F is a norm form as in (1.3) [22]. We state his result in detail. We can rewrite the equation $|F(\mathbf{x})| = 1$ as

$$(1.6) \qquad |cN_{M/\mathbb{Q}}(\xi)| = 1 \quad \text{in } \xi \in \mathcal{M},$$

where $\mathcal{M}$ is the $\mathbb{Z}$-module generated by $\alpha_1, ..., \alpha_n$. We assume that $cN_{M/\mathbb{Q}}(\alpha_1 X_1 + ... + \alpha_n X_n)$ has its coefficients in $\mathbb{Z}$. Let $V = \mathbb{Q}\mathcal{M}$ denote the $\mathbb{Q}$-vector space generated by $\mathcal{M}$, i.e. by $\alpha_1, ..., \alpha_n$. For each subfield J of M, define the subspace of V,

$$(1.7) \qquad V^J = \{\xi \in V : \ \lambda\xi \in V \text{ for every } \lambda \in J\}.$$

Thus, V^J is the largest linear subspace of V closed under scalar multiplication by elements from J. V is called *non-degenerate* if

$$(1.8) \qquad V^J = (0) \text{ for each subfield } J \text{ of } M \text{ which is not equal to } \mathbb{Q} \text{ or an imaginary quadratic field.}$$

Let $r = [M : \mathbb{Q}]$ and $n = \dim V$. Among other things, Schmidt showed that if V is non-degenerate then the number of solutions of (1.6) is at most

$$(1.9) \qquad \min\left(r^{2^{30n} r^2}, r^{(2n)^{n \cdot 2^{n+4}}} \right).$$

Soon afterwards, Schlickewei proved a p-adic generalisation of Schmidt's quantitative Subspace theorem and extended this to number fields [15]. As an application, he obtained an explicit upper bound for the number of solutions of S-unit equations over number fields [16]. Győry [9] used this to derive an explicit upper bound for the number of $\mathcal{O}_S^*$-cosets of solutions of eq. (1.5) for arbitrary decomposable forms F. We state the results of Schlickewei and Győry below.

Let K be an algebraic number field, and S a finite set of places on K, containing all infinite places. Thus, S has cardinality $r_1 + r_2 + t$ where r_1 is the number of embeddings of K into $\mathbb{R}$, r_2 is the number of complex conjugate pairs of embeddings of K into $\mathbb{C}$ and t is the number of prime ideals in S. From the Subspace theorem it follows (cf. [4,12]) that the so-called S-unit equation

$$a_1 u_1 + ... + a_n u_n = 1 \quad \text{in } u_1, ..., u_n \in \mathcal{O}_S^*, \tag{1.10}$$

where $a_1, ..., a_n$ are non-zero coefficients from K, has only finitely many solutions with non-vanishing subsums,

$$\sum_{i \in I} a_i u_i \neq 0 \quad \text{for each non-empty } I \subseteq \{1, ..., n\}. \tag{1.11}$$

In [16], Schlickewei proved that the number of such solutions is at most

$$(4sD)^{2^{36nD} s^6}, \tag{1.12}$$

where s is the cardinality of S and D is the degree of the normal closure of $K/\mathbb{Q}$. Later [17] he improved this to $2^{2^{26n} s}$.

In [9], Győry generalised finiteness results of Schmidt and Schlickewei on 'families of solutions' of (possibly degenerate) norm form equations to decomposable form equations over number fields and obtained explicit upper bounds for the number of families. As a consequence, he obtained the following: suppose that $F \in \mathcal{O}_S[X_1, ..., X_n]$ is a decomposable form such that the number of $\mathcal{O}_S^*$-cosets of solutions of (1.5) is finite (for instance, if F satisfies the non-degeneracy condition of [8]); then this number of cosets is at most

$$\{5sG\}^{2^{37nG} s^6} \tag{1.13}$$

where G is the degree of the normal closure of the field generated by K and the coefficients of the linear forms dividing F.

Győry derived his bound by reducing eq. (1.5) to a system of S-unit equations in some large extension of K (following the arguments in [10,8,7]) and using Schlickewei's bound (1.12) for the latter. If $[K : \mathbb{Q}] = d$ and if F is a norm form of degree r then $dr \leq G \leq (dr)!$; this implies that Győry 's bound is at least doubly exponential in r. Győry's approach might give something better by using an improvement of (1.12), but the best one can get in this way is a bound depending exponentially on r. We mention that Schmidt obtained his bound (1.9) with a polynomial dependence on r by reducing norm form equation (1.6) directly to his quantitative Subspace theorem.

New Results.

In contrast to the results mentioned above, our estimates are not only for the number of solutions of 'non-degenerate' decomposable form equations but also for the number of 'non-degenerate' solutions of possibly degenerate decomposable form equations. Let K be an algebraic number field and S a finite set of places on K of cardinality s, containing all infinite places. Let

$$F(\mathbf{X}) = l_1(\mathbf{X}) \cdots l_r(\mathbf{X}) \in \mathcal{O}_S[X_1, ..., X_n]$$

be a decomposable form of degree r, where $l_1, ..., l_r$ are linear forms with coefficients in some extension of K such that

$$\{\mathbf{x} \in K^n : l_1(\mathbf{x}) = 0, ..., l_r(\mathbf{x}) = 0\} = \{\mathbf{0}\}. \tag{1.14}$$

In §2 we shall define what it means for $\mathbf{x} \in \mathcal{O}_S^n$ to be (F, S)-*non-degenerate* or (F, S)-degenerate. A more restrictive condition independent of S is that for every proper, non-empty subset I of $\{1, ..., r\}$ there are algebraic numbers $c_1, ..., c_r$ such that

$$\begin{aligned} &c_1 l_1 + ... + c_r l_r \text{ is identically zero,} \\ &\sum_{i \in I} c_i l_i(\mathbf{x}) \neq 0 \ . \end{aligned} \tag{1.15}$$

(cf. §7, Remark 4). For instance, if F is a binary form with at least three pairwise non-proportional linear factors then every $\mathbf{x} \in \mathcal{O}_S^2$ with $F(\mathbf{x}) \neq 0$ is (F, S)-non-degenerate. We shall prove:

Theorem 1. *The set of (F, S)-non-degenerate solutions of*

$$F(\mathbf{x}) \in \mathcal{O}_S^* \quad \textit{in } \mathbf{x} \in \mathcal{O}_S^n \tag{1.5}$$

is the union of at most

$$\left(2^{33} r^2\right)^{n^3 s}$$

$\mathcal{O}_S^*$-cosets $\{\epsilon \mathbf{x} : \epsilon \in \mathcal{O}_S^*\}$.

We obtain the upper bound $2 \times (2^{33} r^2)^{n^3(t+1)}$ for the number of (F, S)-non-degenerate solutions of (1.2) by taking $K = \mathbb{Q}$ and $S = \{\infty, p_1, ..., p_t\}$ (where

∞ is the infinite place of $\mathbb{Q}$) and observing that each $\mathcal{O}_S^*$-coset contains precisely two primitive $\mathbf{x}$.

It will turn out (cf. §2, Remark 2) that from a given (F, S)-degenerate solution it is possible to construct infinitely many $\mathcal{O}_S^*$-cosets of such solutions. Together with Theorem 1 this yields:

Corollary. *Suppose that the number of $\mathcal{O}_S^*$-cosets of solutions of (1.5) is finite. Then this number is at most $(2^{33}r^2)^{n^3 s}$.*

We are going to state a general result on norm form equations. let K, S be as before, and let M be a finite extension of K. Further, let $\mathcal{M}$ be a finitely generated $\mathcal{O}_S$-module contained in M. Choose $c \in K^*$ such that for some set of generators $\alpha_1, ..., \alpha_m$ for $\mathcal{M}$, the form

$$F(\mathbf{X}) = cN_{M/K}(\alpha_1 X_1 + ... + \alpha_m X_m)$$

has its coefficients in $\mathcal{O}_S$. It is easy to see that this holds for any set of generators for $\mathcal{M}$ if it holds for one set of generators. We consider the norm form equation

$$(1.16) \qquad cN_{M/K}(\xi) \in \mathcal{O}_S^* \quad \text{in } \xi \in \mathcal{M}.$$

Let $V = K\mathcal{M} = \{a\xi : a \in K, \xi \in \mathcal{M}\}$ denote the K-vector space generated by $\mathcal{M}$. We denote the integral closure of $\mathcal{O}_S$ in some finite extension J of K by $\mathcal{O}_{J,S}$ and the unit group of this ring by $\mathcal{O}_{J,S}^*$. Similarly to (1.7) we define for each subfield J of M containing K,

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

Definition. $\xi \in V$ is called S-non-degenerate if

$$(1.17) \qquad \xi \notin V^J \text{ for every subfield } J \text{ of } M \text{ with } J \supseteq K \text{ for which } \mathcal{O}_{J,S}^*/\mathcal{O}_S^* \text{ is infinite.}$$

It is easy to show that $\mathcal{O}_{J,S}^*/\mathcal{O}_S^*$ is finite in the following two situations only: (i) $J = K$; (ii) K is totally real, J is a totally complex quadratic extension of K and none of the prime ideals in S splits into two different prime ideals in J.
We may partition the set of solutions of (1.16) into $\mathcal{O}_S^*$-cosets $\{\epsilon\xi : \epsilon \in \mathcal{O}_S^*\}$. Clearly, if one element in an $\mathcal{O}_S^*$-coset is S-non-degenerate, then so is every other element.

Theorem 2. *Suppose that $[M : K] = r$, that $\dim_K V = n$, and that S has cardinality s. Then eq. (1.16) has at most $\left(2^{33}r^2\right)^{n^3 s}$ $\mathcal{O}_S^*$-cosets of S-non-degenerate solutions.*

We call the space V S-non-degenerate if every non-zero $\xi \in V$ is S-non-degenerate. For $K = \mathbb{Q}$, $S = \{\infty\}$, this is precisely definition (1.8). Note that in this case $\mathcal{O}_S^*$-cosets consist of $\pm\xi$. It follows that Schmidt's bound (1.9) for the number of solutions of eq. (1.6): $|cN_{M/\mathbb{Q}}(\xi)| = 1$ in $\xi \in \mathcal{M}$ can be improved to

$$2 \times \left(2^{33} r^2\right)^{n^3}.$$

Finally, we mention an improvement of Schlickewei's upper bound (1.12) for the number of solutions of S-unit equations. Let K, S be as before.

Theorem 3. *Let $a_1, ..., a_n \in K^*$. Suppose that S has cardinality s. Then the equation*

$$a_1 u_1 + ... + a_n u_n = 1 \quad \text{in } u_1, ..., u_n \in \mathcal{O}_S^* \text{ with}$$
$$\sum_{i \in I} a_i u_i \neq 0 \quad \text{for each non-empty } I \subseteq \{1, ..., n\}$$

has at most $\left(2^{35} n^2\right)^{n^3 s}$ solutions.

Although norm form equations and S-unit equations may be considered as special types of decomposable form equations, there are problems with deriving Theorems 2 and 3 from Theorem 1, caused by the fact that in general $\mathcal{O}_S$ is not a principal ideal domain. Therefore, we will derive Theorems 1,2,3 from Theorem 4 in §2, which is a result on "Galois-symmetric S-unit-vectors."
Let K, S be as before and let $F(\mathbf{X}) \in \mathcal{O}_S[X_1, ..., X_n]$ be a decomposable form of degree r. Then $F(\mathbf{X}) = l_1(\mathbf{X}) \cdots l_r(\mathbf{X})$ where $l_1, ..., l_r$ are linear forms with coefficients in some normal extension L of K. For $\mathbf{x} \in K^n$, put $u_i := l_i(\mathbf{x})$ for $i = 1, ..., r$ and $\mathbf{u} = (u_1, ..., u_r)$. By multiplying them with constants if necessary, we may assume that the linear factors $l_1, ..., l_r$ of F are permuted by applying any automorphism from $\mathrm{Gal}(L/K)$ to their coefficients. Thus, every $\sigma \in \mathrm{Gal}(L/K)$ permutes $u_1, ..., u_r$, in other words, for every $\sigma \in \mathrm{Gal}(L/K)$ there is a permutation $\sigma(1), ..., \sigma(r)$ of $1, ..., r$ such that $\sigma(u_i) = u_{\sigma(i)}$ for $i = 1, ..., r$. Such a vector is said to be Galois-symmetric. There is a finite set of places S' on L such that for every solution $\mathbf{x}$ of (1.5), $u_1, ..., u_r$ are S'-units. Thus, every solution $\mathbf{x}$ of (1.5) corresponds to a Galois-symmetric S'-unit vector $\mathbf{u}$.

The main tool in the proof of Theorem 4 is our improved quantitative Subspace theorem from [6]. We use several ideas from Schmidt's paper [22] but our arguments differ from that of [22] in that we do not apply the Diophantine approximation techniques to the solutions $\mathbf{x}$ of (1.5) but to the corresponding Galois-symmetric vectors $\mathbf{u}$; for instance, we use the reformulation of the quantitative Subspace theorem in terms of $\mathbf{u}$ which is stated in §4. In this way we can avoid generalising the reduction theory for norm form equations in [22] to the p-adic case.

Acknowledgements. I would like to thank Professor Schlickewei for his remarks on a much earlier draft, some of which have been used in the present paper. Further, I would like to thank Professor Győry for our discussion on the subject of the paper and Professor Tijdeman for his remarks.

§2. The main result.

The unit group of a ring or algebra R (always assumed to have a unit element) is denoted by R^*. The r-fold direct sum or product of a ring, group, etc. G is denoted by G^r. Let $\overline{\mathbb{Q}}$ denote the algebraic closure of $\mathbb{Q}$. For tuples $\lambda = (\lambda_1, ..., \lambda_r), \mu = (\mu_1, ..., \mu_r) \in \overline{\mathbb{Q}}^r$ we define coordinatewise addition $\lambda + \mu = (\lambda_1 + \mu_1, ..., \lambda_r + \mu_r)$, scalar multiplication $a\lambda = (a\lambda_1, ..., a\lambda_r)$ (for $a \in \overline{\mathbb{Q}}$), and multiplication $\lambda\mu = (\lambda_1\mu_1, ..., \lambda_r\mu_r)$. The Galois group of a Galois extension F'/F is denoted by $\mathrm{Gal}(F'/F)$. For any algebraic number field it will be assumed that it is contained in $\overline{\mathbb{Q}}$.
In what follows, K is an algebraic number field and S a finite set of places on K, containing all infinite places. We denote by $\overline{\mathcal{O}}_S$ the integral closure in $\overline{\mathbb{Q}}$ of the ring of S-integers $\mathcal{O}_S$.

Let Σ be a $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$-action on $\{1, ..., r\}$, i.e. a homomorphism from $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$ to the permutation group of $\{1, ..., r\}$; thus, Σ attaches to every $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$ a permutation $(\sigma(1), ..., \sigma(r))$ of $(1, ..., r)$. To Σ we associate the K-algebra

$$\Lambda_\Sigma := \left\{ \lambda = (\lambda_1, ..., \lambda_r) \in \overline{\mathbb{Q}}^r : \begin{array}{ll} \sigma(\lambda_i) = \lambda_{\sigma(i)} & \text{for } i = 1, ..., r \\ & \text{and } \sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K) \end{array} \right\} \tag{2.1}$$

endowed with coordinatewise addition, multiplication and scalar multiplication by elements of K. (Verify that Λ_Σ is closed under these operations). Note that Λ_Σ has unit element $\mathbf{1} := (1, ..., 1)$. Further, Λ_Σ has unit group $\Lambda_\Sigma^* = \{\lambda \in \Lambda_\Sigma : \lambda_1 \cdots \lambda_r \neq 0\}$. *) The *diagonal homomorphism* $\delta : a \mapsto (a, ..., a) = a \cdot \mathbf{1}$ maps K injectively into Λ_Σ. For instance, if Σ is the trivial $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$-action on $\{1, ..., r\}$ then Λ_Σ is the K-algebra K^r with coordinatewise operations.

Let $\mathcal{P}$ be a (Σ-)*symmetric partition* of $\{1, ..., , r\}$, that is, a collection of sets $\mathcal{P} = \{P_1, ..., P_t\}$ such that:

$P_1, ..., P_t$ are non-empty and pairwise disjoint,
$P_1 \cup ... \cup P_t = \{1, ..., r\}$,
$\sigma(P_i) := \{\sigma(k) : k \in P_i\}$ belongs to $\mathcal{P}$ for $i = 1, ..., r$ and for $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$.

*) It is easy to show that Λ_Σ is isomorphic to a direct K-algebra sum $K_1 \oplus ... \oplus K_t$ of finite field extensions of K with $[K_1{:}K] + ... + [K_t{:}K] = r$, cf. [5], Lemma 6.

A *pair* $i \overset{\mathcal{P}}{\sim} j$ is a pair $i, j \in \{1, ..., r\}$ such that i, j belong to the same set of $\mathcal{P}$. To $\mathcal{P}$ we associate the following K-subalgebra of Λ_Σ:

$$\Lambda_{\mathcal{P}} = \Lambda_{\Sigma,\mathcal{P}} = \{\lambda = (\lambda_1, ..., \lambda_r) \in \Lambda_\Sigma : \; \lambda_i = \lambda_j \text{ for each pair } i \overset{\mathcal{P}}{\sim} j\}. \quad ^{\dagger)}$$

For instance, if $\mathcal{P} = \{\{1\}, ..., \{r\}\}$ then $\Lambda_{\mathcal{P}} = \Lambda_\Sigma$ while if $\mathcal{P} = \{\{1, ..., r\}\}$ then $\Lambda_{\mathcal{P}} = \delta(K)$. Further, we define the $\mathcal{O}_S$-algebra

$$\mathcal{O}_{\mathcal{P},S} := \Lambda_{\mathcal{P}} \cap (\overline{\mathcal{O}}_S)^r = \{\lambda = (\lambda_1, ..., \lambda_r) \in \Lambda_{\mathcal{P}} : \lambda_i \in \overline{\mathcal{O}}_S \text{ for } i = 1, ..., r\}.$$

Note that $\mathcal{O}_{\mathcal{P},S}$ has unit group

$$\mathcal{O}^*_{\mathcal{P},S} = \{\lambda \in \Lambda^*_{\mathcal{P}} : \lambda_i \in \overline{\mathcal{O}}^*_S \;\text{ for } i = 1, ..., r\}$$

and that $\delta(\mathcal{O}^*_S) = \{(a, ..., a) : a \in \mathcal{O}^*_S\}$ is a subgroup of $\mathcal{O}^*_{\mathcal{P},S}$.

Now let W be an n-dimensional K-linear subspace of Λ_Σ, where $n \geq 2$. For each symmetric partition $\mathcal{P}$ of Λ_Σ, define the K-linear subspace of W,

$$W_{\mathcal{P}} = W_{\Sigma,\mathcal{P}} := \{\mathbf{u} \in W : \lambda\mathbf{u} \in W \;\text{ for every }\; \lambda \in \Lambda_{\mathcal{P}}\}.$$

For every $\lambda, \mu \in \Lambda_{\mathcal{P}}$, $\mathbf{u} \in W_{\mathcal{P}}$ we have $\lambda(\mu\mathbf{u}) = (\lambda\mu)\mathbf{u} \in W$; hence $\mu\mathbf{u} \in W_{\mathcal{P}}$. Therefore, $W_{\mathcal{P}}$ is closed under multiplication by elements of $\Lambda_{\mathcal{P}}$. It is in fact the largest subspace of W with this property. The spaces $W_{\mathcal{P}}$ appeared also in Győry's paper [9].

Definition. $\mathbf{u} \in W$ is called *S-non-degenerate* if

(2.2) $\quad \mathbf{u} \notin W_{\mathcal{P}}$ for each symmetric partition $\mathcal{P}$ of $\{1, ..., r\}$ for which $\mathcal{O}^*_{\mathcal{P},S}/\delta(\mathcal{O}^*_S)$ is infinite,

and *S-degenerate* otherwise.

At the end of this section (cf. Remark 3), we have listed the symmetric partitions $\mathcal{P}$ for which $\mathcal{O}^*_{\mathcal{P},S}/\delta(\mathcal{O}^*_S)$ is finite.

For our applications to decomposable form equations, norm form equations and S-unit equations we need a result on the set of vectors $\mathbf{u} = (u_1, ..., u_r) \in W$ with $u_1, ..., u_r \in \overline{\mathcal{O}}^*_S$. Since for this we did not have to change our arguments, we proved a slightly more general "projective" result about elements $\mathbf{u}$ of W for which the quotients u_i/u_j belong to $\overline{\mathcal{O}}^*_S$. A K^*-coset is a set $\{a\mathbf{u} : \; a \in K^*\}$ with some fixed $\mathbf{u} \in \Lambda_\Sigma$.

$\dagger)$ In fact, in this way we obtain all K-subalgebras of Λ_Σ containing $\mathbf{1}$.

Theorem 4. *Let K be an algebraic number field, S a finite set of places of K of cardinality s containing all infinite places, Σ a $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$- action on $\{1,...,r\}$ and W an n-dimensional K-linear subspace of*

$$\Lambda_\Sigma = \{\lambda \in \overline{\mathbb{Q}}^r : \sigma(\lambda_i) = \lambda_{\sigma(i)} \ \text{ for } i = 1,...,r,\ \sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)\}$$

where $r \geq 2, n \geq 2$. Then the set of $\mathbf{u} \in W$ for which

$$u_1 \cdots u_r \neq 0, \quad u_i/u_j \in \overline{\mathcal{O}}_S^* \text{ for } i,j = 1,...,r, \tag{2.3}$$

$$\mathbf{u} \text{ is } S\text{-non-degenerate} \tag{2.4}$$

is the union of at most

$$\left(2^{33}r^2\right)^{n^3 s} \tag{2.5}$$

K^*-cosets.

Clearly, if $\mathbf{u}$ satisfies (2.3), (2.4), then so does $a\mathbf{u}$ for every $a \in K^*$; therefore, it makes sense to count the number of K^*-cosets of elements $\mathbf{u} \in W$ with (2.3), (2.4). We remark that a K^*-coset of elements of Λ_Σ satisfying (2.3) need not contain an $\mathbf{u} = (u_1, ..., u_r)$ with $u_1, ..., u_r \in \overline{\mathcal{O}}_S^*$.

Remark 1. From an S-degenerate element $\mathbf{u}$ of W with (2.3) one can construct infinitely many K^*-cosets of such elements. Namely, let $\mathbf{u} \in W_\mathcal{P}$ for some symmetric partition $\mathcal{P}$ of $\{1,...,r\}$ for which $\mathcal{O}_{\mathcal{P},S}^*/\delta(\mathcal{O}_S^*)$ is infinite. Every element of the set $H := \{\zeta\mathbf{u} : \zeta \in \mathcal{O}_{\mathcal{P},S}^*\}$ belongs to $W_\mathcal{P}$ and satisfies (2.3). Moreover, since $\mathcal{O}_{\mathcal{P},S}^* \cap \delta(K^*) = \delta(\mathcal{O}_S^*)$, the set H is not contained in the union of finitely many K^*-cosets.

We now define the notion of (F,S)-degeneracy for decomposable forms F. Let as before S be a finite set of places on K of cardinality s, containing all infinite places. Further, let $F(\mathbf{X}) = l_1(\mathbf{X}) \cdots l_r(\mathbf{X}) \in \mathcal{O}_S[X_1, ..., X_n]$ be a decomposable form of degree r in $n \geq 2$ variables, where $l_1, ..., l_r$ are homogeneous linear forms in n variables with coefficients from $\overline{\mathbb{Q}}$ such that

$$\{\mathbf{x} \in K^n :\ l_1(\mathbf{x}) = 0, ..., l_r(\mathbf{x}) = 0\} = \{\mathbf{0}\}. \tag{1.14}$$

Define the K-linear map and the K-vector space

$$\varphi : K^n \to \overline{\mathbb{Q}}^r : \quad \varphi(\mathbf{x}) = (l_1(\mathbf{x}), ..., l_r(\mathbf{x})) \ \text{ and } W = \varphi(K^n),$$

respectively. φ is injective because of (1.14). By applying $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$ to the coefficients of $l_1, ..., l_r$ we obtain the same linear forms, but multiplied with certain constants and in permuted order. In other words, there is a $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$-action Σ on $\{1,...,r\}$ such that

$$\sigma(l_i) = c_{\sigma i} l_{\sigma(i)} \quad \text{for } i = 1,...,r,\ \sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K), \tag{2.6}$$

where $\sigma(l_i)$ is the linear form obtained by applying σ to the coefficients of l_i and where $c_{\sigma i}$ is some constant. We define again the space $W_{\mathcal{P}}$ by $\{\mathbf{u} \in W : \lambda\mathbf{u} \in W$ for every $\lambda \in \Lambda_{\mathcal{P}}\}$.

Definition. $\mathbf{x} \in K^n$ is called *(F,S)-non-degenerate* if $\varphi(\mathbf{x})$ is S-non-degenerate, i.e. if $\varphi(\mathbf{x}) \notin W_{\mathcal{P}}$ for every symmetric partition $\mathcal{P}$ of $\{1,...,r\}$ for which $\mathcal{O}^*_{\mathcal{P},S}/\delta(\mathcal{O}^*_S)$ is infinite. Otherwise, $\mathbf{x}$ is called (F,S)-degenerate.

Clearly, if $\mathbf{x}$ is (F,S)-(non-) degenerate, then so is every element in the $\mathcal{O}^*_S$-coset $\{\epsilon\mathbf{x} : \ \epsilon \in \mathcal{O}^*_S\}$.

We claim that the set of (F,S)-non-degenerate elements of K^n does not depend on the choice of the factorisation $l_1 \cdots l_r$ of F into linear forms and moreover, does not change when F is replaced by cF for some $c \in K^*$. Namely, let $l'_1 \cdots l'_r$ be a factorisation of cF into linear forms. Then there is a tuple of non-zero algebraic numbers $\mathbf{c} = (c_1,...,c_r)$ such that $l'_1,...,l'_r$ is a permutation of $c_1l_1,...,c_rl_r$. Put $\varphi'(\mathbf{x}) = (l'_1(\mathbf{x}),...,l'_r(\mathbf{x}))$, $W' := \varphi'(K^n)$. Then $\varphi' = \tau \circ t \circ \varphi$, where t denotes coordinatewise multiplication with $\mathbf{c}$ and τ is some permutation of coordinates. t maps the S-non-degenerate elements of W bijectively to those of $t(W)$ since $t(W)_{\mathcal{P}} = t(W_{\mathcal{P}})$ for each symmetric partition $\mathcal{P}$ of $\{1,...,r\}$. Further, it is easy to verify that τ maps Λ_Σ to $\Lambda_{\Sigma'}$ where Σ' is some other $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$-action of $\{1,...,r\}$ and that for each Σ-symmetric partition $\mathcal{P}$ of $\{1,...,r\}$, τ maps $t(W)_{\mathcal{P}}$ to $W'_{\mathcal{P}'}$ where $\mathcal{P}'$ is some Σ'-symmetric partition of $\{1,...,r\}$. Hence τ maps the S-non-degenerate elements of $t(W)$ bijectively to those of W'. This proves our claim.

Below, we shall derive Theorems 1,2 and 3 from Theorem 4. We need the following lemma.

Lemma 1. *Let $G(\mathbf{X}) = G_1(\mathbf{X})...G_r(\mathbf{X})$ be a form in $\mathcal{O}_S[X_1,...,X_n]$, where $G_1,...,G_r$ are homogeneous polynomials with coefficients from $\overline{\mathbb{Q}}$. Then for every $\mathbf{x},\mathbf{y} \in \mathcal{O}^n_S$ with $G(\mathbf{x}) \in \mathcal{O}^*_S$, $G(\mathbf{y}) \in \mathcal{O}^*_S$ we have*

$$G_i(\mathbf{x})/G_i(\mathbf{y}) \in \overline{\mathcal{O}}^*_S \ \text{for } i = 1,...,r.$$

Proof. Let L be a finite extension of K containing the coefficients of $G_1,...,G_r$. Let R denote the integral closure of $\mathcal{O}_S$ in L; then R^* is a subgroup of $\overline{\mathcal{O}}^*_S$. For a polynomial P with coefficients from L, denote by (P) the fractional ideal with respect to R generated by the coefficients of P. Using Gauss' lemma for Dedekind domains we obtain $(G_1)...(G_r) = (G) \subseteq R$. Further, we have $G_i(\mathbf{x}) \in (G_i)$ for $i = 1, ..., r$ and $G_1(\mathbf{x})...G_r(\mathbf{x}) = G(\mathbf{x}) \in R^*$. It follows that for $i = 1, ..., r$ we have $(G_i(\mathbf{x})) = (G_i)$. Clearly the same holds for $\mathbf{y}$. It follows that for $i = 1, ..., r$, $G_i(\mathbf{x})$ and $G_i(\mathbf{y})$ generate the same ideal i.e. their quotient belongs to R^*. $\square$

Proof of Theorem 1 (on decomposable form equations).
Recall that we are considering eq. (1.5) $F(\mathbf{x}) \in \mathcal{O}_S^*$ in $\mathbf{x} \in \mathcal{O}_S^n$. We assume that (1.5) has a solution, $\mathbf{y}$, say. Replacing l_i by $l_i(\mathbf{y})^{-1}l_i$ for $i = 1, ..., r$ and F by $F(\mathbf{y})^{-1}F$ does not affect the set of (F, S)-non-degenerate solutions of (1.5). Therefore, we may assume that $l_i(\mathbf{y}) = 1$ for $i = 1, ..., r$, and shall do so in the sequel. Thus, the constants $c_{\sigma i}$ in (2.6) are equal to 1, i.e. $\sigma(l_i) = l_{\sigma(i)}$ for $i = 1, ..., r$, $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$. It follows that $W = \varphi(K^n)$ is a K-linear subspace of Λ_Σ. Further, $\varphi(\mathbf{x}) = (l_1(\mathbf{x})/l_1(\mathbf{y}), ..., l_r(\mathbf{x})/l_r(\mathbf{y}))$. From the injectivity of φ it follows that $\dim W = n$ and that φ maps different $\mathcal{O}_S^*$-cosets in $\mathcal{O}_S^n$ into different K^*-cosets in W. If $\mathbf{x}$ is any (F, S)-non-degenerate solution of (1.5), then by definition, $\varphi(\mathbf{x})$ is an S-non-degenerate element of W; further, by Lemma 1 with $G_i = l_i$, the coordinates of $\varphi(\mathbf{x})$ belong to $\overline{\mathcal{O}}_S^*$, whence $\varphi(\mathbf{x})$ satisfies (2.3). Now Theorem 1 follows at once by applying Theorem 4 to W. $\square$

Remark 2. We now show that from an (F, S)-degenerate solution of (1.5) it is possible to construct infinitely many $\mathcal{O}_S^*$-cosets of such solutions. Let $\mathbf{x}$ be an (F, S)-degenerate solution of (1.5) with $\varphi(\mathbf{x}) \in W_\mathcal{P}$, where $\mathcal{P}$ is a symmetric partition of $\{1, ..., r\}$ for which $\mathcal{O}_{\mathcal{P},S}^*/\delta(\mathcal{O}_S^*)$ is infinite. For every $\lambda \in \Lambda_\mathcal{P}$, put $\mathbf{x}_\lambda := \varphi^{-1}(\lambda.\varphi(\mathbf{x}))$; note that $\mathbf{x}_\lambda \in \varphi^{-1}(W_\mathcal{P})$. $\mathcal{O}_{\mathcal{P},S}$ is a finitely generated $\mathcal{O}_S$-module; let $\{\lambda_1, ..., \lambda_t\}$ be a set of generators. There is a non-zero $d \in \mathcal{O}_S$ such that $d\mathbf{x}_{\lambda_i} \in \mathcal{O}_S^n$ for $i = 1, ..., t$. Then $d\mathbf{x}_\lambda \in \mathcal{O}_S^n$ for every $\lambda \in \mathcal{O}_{\mathcal{P},S}$. There is a positive integer m such that $\eta^m - \mathbf{1} \in d\mathcal{O}_{\mathcal{P},S}$ for every $\eta \in \mathcal{O}_{\mathcal{P},S}^*$ since $\mathcal{O}_{\mathcal{P},S}/d\mathcal{O}_{\mathcal{P},S}$ is finite.

Now let

$$G := \{\epsilon\eta^m : \epsilon \in \mathcal{O}_S^*, \eta \in \mathcal{O}_{\mathcal{P},S}^*\}.$$

For $\zeta = \epsilon\eta^m \in G$ we have $\zeta = \epsilon\mathbf{1} + d\lambda$ for some $\lambda \in \mathcal{O}_{\mathcal{P},S}$. Hence $\mathbf{x}_\zeta = \varphi^{-1}(\{\epsilon\mathbf{1} + d\lambda\}\varphi(\mathbf{x})) = \epsilon\mathbf{x} + d\mathbf{x}_\lambda \in \mathcal{O}_S^n$. Moreover, if $\zeta = (\zeta_1, ..., \zeta_r)$, then $\zeta_1 \cdots \zeta_r \in \overline{\mathcal{O}}_S^* \cap K^* = \mathcal{O}_S^*$. Therefore, $F(\mathbf{x}_\zeta) = \zeta_1...\zeta_r F(\mathbf{x}) \in \mathcal{O}_S^*$, i.e., $\mathbf{x}_\zeta$ satisfies (1.5). Since $\mathcal{O}_{\mathcal{P},S}^*$ is finitely generated, G has finite index in $\mathcal{O}_{\mathcal{P},S}^*$. Therefore, $G/\delta(\mathcal{O}_S^*)$ is infinite. Hence the set of vectors $\zeta\varphi(\mathbf{x})$, and so the set of vectors $\mathbf{x}_\zeta$ $(\zeta \in G)$, is not contained in the union of finitely many $\mathcal{O}_S^*$-cosets. $\square$

Proof of Theorem 2 (on norm form equations).
Let S be as before. Further, let M be a finite extension of K of degree r and let $\mathcal{M}$ be a finitely generated $\mathcal{O}_S$-submodule of M, such that the K-vector space $V := K\mathcal{M} = \{a\xi : a \in K, \xi \in \mathcal{M}\}$ has dimension n. Choose a set of generators

$\{\alpha_1, ..., \alpha_m\}$ for $\mathcal{M}$. We assume that

$$(2.7) \qquad 1 \in \mathcal{M}, \quad N_{M/K}(\alpha_1 X_1 + ... + \alpha_m X_m) \in \mathcal{O}_S[X_1, ..., X_m]$$

and we consider the equation

$$(2.8) \qquad N_{M/K}(\xi) \in \mathcal{O}_S^* \ \text{ in } \xi \in \mathcal{M}.$$

We show that (2.8) has at most $C := (2^{33}r^2)^{n^3 s}$ $\mathcal{O}_S^*$-cosets of solutions. In Theorem 2 we considered the more general equation (1.16) $cN_{M/K}(\xi) \in \mathcal{O}_S^*$ in $\mathbf{x} \in \mathcal{O}_S^n$ and we assumed that $cN_{M/K}(\alpha_1 X_1 + ... + \alpha_m X_m)$ has its coefficients in $\mathcal{O}_S$, for some $c \in K^*$. However, taking a solution ξ_0 of (1.16), we have for any solution ξ_1 of (1.16) that $\xi_1' := \xi_1/\xi_0$ is a solution of $N_{M/K}(\xi') \in \mathcal{O}_S^*$ in $\xi' \in \mathcal{M}'$ where $\mathcal{M}' = \xi_0^{-1}\mathcal{M}$; moreover, $\mathcal{M}'$ satisfies (2.7). So it suffices to consider eq. (2.8).

Let $\xi \mapsto \xi^{(1)}, ..., \xi \mapsto \xi^{(r)}$ denote the K-isomorphic embeddings of M into $\overline{\mathbb{Q}}$, where $\xi^{(1)} = \xi$. The mapping

$$\psi : \ M \hookrightarrow \overline{\mathbb{Q}}^r : \ \psi(\xi) = (\xi^{(1)}, ..., \xi^{(r)})$$

is a K-algebra isomorphism from M to Λ_Σ, where Σ is the $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$-action on $\{1, ..., r\}$ defined by $\sigma(\xi^{(i)}) = \xi^{(\sigma(i))}$ for $i = 1, ..., r$. Put $W := \psi(V)$.

We claim that if ξ is an S-non-degenerate element of V then $\psi(\xi)$ is an S-non-degenerate element of W. Namely, suppose that $\xi \in V$ is such that $\psi(\xi)$ is an S-degenerate element of W. Then $\psi(\xi) \in W_{\mathcal{P}}$ for some symmetric partition $\mathcal{P}$ of $\{1, ..., r\}$ for which $\mathcal{O}_{\mathcal{P},S}^*/\delta(\mathcal{O}_S^*)$ is infinite. Let $J := \psi^{-1}(\Lambda_{\mathcal{P}})$. Then J is a K-subalgebra of M containing 1, hence a subfield of M containing K. Denote by $\mathcal{O}_{J,S}$ the integral closure of $\mathcal{O}_S$ in J. Then for $\varepsilon \in J$ we have $\varepsilon \in \mathcal{O}_{J,S}^*$ if and only if $\psi(\varepsilon) = (\varepsilon^{(1)}, ..., \varepsilon^{(r)}) \in (\overline{\mathcal{O}}_S^*)^r$. Hence $\psi(\mathcal{O}_{J,S}^*) = \Lambda_{\mathcal{P}} \cap (\overline{\mathcal{O}}_S^*)^r = \mathcal{O}_{\mathcal{P},S}^*$. Further, $\psi(\mathcal{O}_S^*) = \delta(\mathcal{O}_S^*)$, since both maps make an r-fold copy of $\xi \in K$. Hence $\mathcal{O}_{J,S}^*/\mathcal{O}_S^*$ is infinite. Moreover, since $\psi(\xi) \in W_{\mathcal{P}}$ we have $\psi(\xi)\psi(J) = \psi(\xi)\Lambda_{\mathcal{P}} \subseteq W = \psi(V)$, i.e. $\xi J \subseteq V$, which implies $\xi \in V^J$. It follows that ξ is an S-degenerate element of V. This proves our claim.

Let $\xi \in \mathcal{M}$ be a solution of (2.8). Choose a vector $\mathbf{x} = (x_1, ..., x_m) \in \mathcal{O}_S^m$ with $\xi = \sum_i x_i \alpha_i$; by (2.7) there is a vector $\mathbf{y} = (y_1, ..., y_m) \in \mathcal{O}_S^m$ with $1 = \sum_i y_i \alpha_i$. Define the linear forms $l_i(\mathbf{X}) = \alpha_1^{(i)} X_1 + ... + \alpha_m^{(i)} X_m$ for $i = 1, ..., r$. Then by (2.7), their product F has its coefficients in $\mathcal{O}_S$. Hence by Lemma 1, $\xi^{(i)} = l_i(\mathbf{x})/l_i(\mathbf{y}) \in \overline{\mathcal{O}}_S^*$ for $i = 1, ..., r$. Therefore, $\psi(\xi)$ satisfies (2.3). Now by applying Theorem 4 to $W = \psi(V)$, using our claim from above and observing that ψ maps different $\mathcal{O}_S^*$-cosets into different K^*-cosets we infer that (2.8) has at most C $\mathcal{O}_S^*$-cosets of solutions. This implies Theorem 2. $\square$

Proof of Theorem 3 (on S-unit equations).
Let S be as before, and $a_1, ..., a_n \in K^*$. Recall that we have to estimate the number of solutions of

$$(1.10) \qquad a_1u_1 + ... + a_nu_n = 1 \quad \text{in } u_1, ..., u_n \in \mathcal{O}_S^*$$

with

$$(1.11) \qquad \sum_{i\in I} a_iu_i \neq 0 \quad \text{for each non-empty subset } I \text{ of } \{1, ..., n\}.$$

Let Σ be the trivial action on $\{1, ..., n+1\}$ so that $\Lambda_\Sigma = K^{n+1}$ endowed with coordinatewise operations. Put $a_{n+1} := -1$. Define the K-linear subspace of Λ_Σ:

$$W = \{\mathbf{u} = (u_1, ..., u_{n+1}) \in \Lambda_\Sigma : \ a_1u_1 + ... + a_nu_n + a_{n+1}u_{n+1} = 0\}.$$

Let $(u_1, ..., u_n)$ be a solution of (1.10) with (1.11) and put $\mathbf{u} := (u_1, ..., u_n, 1)$. Thus, $\mathbf{u} \in W$. We show that $\mathbf{u}$ is S-non-degenerate. Assume the contrary. Then $\mathbf{u} \in W_\mathcal{P}$ for some partition $\mathcal{P} = \{P_1, ..., P_t\}$ of $\{1, ..., n+1\}$ with $t \geq 2$. We have $\Lambda_\mathcal{P} = \{\lambda = (\lambda_1, ..., \lambda_{n+1}) \in K^{n+1} : \ \lambda_i = \lambda_j$ for each pair $i \stackrel{\mathcal{P}}{\sim} j\}$; there are no conjugacy relations between the λ_i since Σ is trivial. We assume without loss of generality that $n+1 \notin P_1$. Choose $\lambda \in \Lambda_\mathcal{P}$ with $\lambda_i = 1$ for $i \in P_1$, $\lambda_i = 0$ for $i \in P_2 \cup ... \cup P_t$. We have $\lambda\mathbf{u} \in W$ which implies that $\sum_{i\in P_1} a_iu_i = 0$. But this contradicts (1.11).
By Theorem 4 with $r = n+1$, there are at most $(2^{33}(n+1)^2)^{n^3s}$ S-non-degenerate vectors $\mathbf{u} = (u_1, ..., u_n, 1) \in W$ with $u_1, ..., u_n \in \mathcal{O}_S^*$. This implies Theorem 3. $\square$

Remark 3. Let as before K be an algebraic number field, S a finite set of places on K containing all infinite places, and Σ a $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$-action on $\{1, ..., r\}$. Lemma 8 of [5] (equivalence (ii)$\Longleftrightarrow$(iii) with $\mathcal{O}_{\mathcal{P},S}^*$ replacing $G(F)$) gives a description of the symmetric partitions $\mathcal{P}$ of $\{1, ..., r\}$ for which $\mathcal{O}_{\mathcal{P},S}^*/\delta(\mathcal{O}_S^*)$ is finite. For the sake of completeness we recall this result.
Let $\mathcal{P} = \{P_1, ..., P_t\}$ be a symmetric partition of $\{1, ..., r\}$. Define the fields $K_1, ..., K_t$ by

$$\mathrm{Gal}(\overline{\mathbb{Q}}/K_j) = \{\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K) : \sigma(P_j) = P_j\} \quad \text{for } j = 1, ..., t.$$

Divide $\{P_1, ..., P_t\}$ into orbits such that P_i, P_j belong to the same orbit if and only if $P_j = \sigma(P_i)$ for some $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$. In that case, $\sigma(K_i) = K_j$. Let u be the number of orbits. Then $\mathcal{O}_{\mathcal{P},S}^*/\delta(\mathcal{O}_S^*)$ is finite if and only if one of the conditions (2.9.a,b,c) below is satisfied:

(2.9.a) $u = 1, t = 1$, i.e. $\mathcal{P} = \{\{1, ..., r\}\}$;

(2.9.b) $u = 1, t = 2$, K is totally real and K_1 is a totally complex quadratic extension of K such that none of the prime ideals in S splits into two prime ideals in K_1; further, $K_2 = \sigma(K_1) = K_1$ for some $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$;

(2.9.c) $u \geq 2$, each field among $K, K_1, ..., K_t$ is either $\mathbb{Q}$ or an imaginary quadratic field, and $S = \{\infty\}$, where ∞ is the only infinite place of K.

§3. Absolute values and heights.

Let K be an algebraic number field and denote by M_K the set of places of K. M_K consists of the embeddings $\sigma : K \hookrightarrow \mathbb{R}$ which are called *real infinite places*; the pairs of complex conjugate embeddings $\sigma, \overline{\sigma} : K \hookrightarrow \mathbb{C}$ which are called *complex infinite places*; and the prime ideals of $\mathcal{O}_K$ which are also called *finite places*. For every $v \in M_K$ we define an absolute value $|.|_v$ as follows:

$|.|_v := |\sigma(.)|^{1/[K:\mathbb{Q}]}$ if v is a real infinite place $\sigma : K \hookrightarrow \mathbb{R}$;
$|.|_v := |\sigma(.)|^{2/[K:\mathbb{Q}]} = |\overline{\sigma}(.)|^{2/[K:\mathbb{Q}]}$ if v is a complex infinite place $\{\sigma, \overline{\sigma} : K \hookrightarrow \mathbb{C}\}$;
$|.|_v := (N\wp)^{-\mathrm{ord}_\wp(.)/[K:\mathbb{Q}]}$ if v is a finite place, i.e. prime ideal $\wp$ of $\mathcal{O}_K$;

here $N\wp$ is the norm of $\wp$, i.e. the cardinality of $\mathcal{O}_K/\wp$, and $\mathrm{ord}_\wp(x)$ is the exponent of $\wp$ in the prime ideal decomposition of (x). For every $v \in M_K$ we choose a continuation of $|.|_v$ to $\overline{\mathbb{Q}}$, denoted also by $|.|_v$, and fix this in the sequel. Note that if S is a finite subset of M_K containing all (real and complex) infinite places, then

$$\mathcal{O}_S = \{x \in K : |x|_v \leq 1 \text{ for } v \notin S\}, \quad \mathcal{O}_S^* = \{x \in K : |x|_v = 1 \text{ for } v \notin S\};$$

here we write $v \notin S$ for $v \in M_K \backslash S$. The absolute values defined above satisfy the *Product formula*

$$\prod_v |a|_v = 1 \quad \text{for } a \in K^*$$

(product over M_K) and the *Extension formulas* for each finite extension L of K,

$$\prod_{w|v} |a|_w = |N_{L/K}(a)|_v^{1/[L:K]} \quad \text{for } a \in L, v \in M_K,$$

$$\prod_{w|v} |a|_w = |a|_v \quad \text{for } a \in K, v \in M_K$$

where the product is taken over all $w \in M_L$ lying above v (i.e. over all w such that the restriction of $|.|_w$ to K is a power of $|.|_v$).

For a vector $\mathbf{x} = (x_1, ..., x_r) \in \overline{\mathbb{Q}}^r$, put

$$|\mathbf{x}|_v = |x_1, ..., x_r|_v := \max(|x_1|_v, ..., |x_r|_v) \quad \text{for} \quad v \in M_K.$$

Define the *height* of $\mathbf{x} = (x_1, ..., x_r) \in \overline{\mathbb{Q}}^r$ by

$$H(\mathbf{x}) = H(x_1, ..., x_r) := \prod_{w \in M_L} |\mathbf{x}|_w,$$

where L is any number field containing $x_1, ..., x_r$ and $|\mathbf{x}|_w$ is defined similarly as $|\mathbf{x}|_v$. By the Extension formula, this does not depend on L. By the Product formula, we have

$$H(\lambda \mathbf{x}) = H(\mathbf{x}) \quad \text{for } \mathbf{x} \in \overline{\mathbb{Q}}^r, \lambda \in \overline{\mathbb{Q}}^*.$$

Note that $H(\mathbf{x}) \geq 1$ if $\mathbf{x} \neq \mathbf{0}$.

For $v \in M_K$, put

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

For the numbers $r_1,\ r_2$ of real infinite places, complex infinite places, respectively we have $r_1 + 2r_2 = [K:\mathbb{Q}]$. Hence

$$\sum_v s(v) = \sum_{v|\infty} s(v) = 1.$$

We define the *scalar product* of $\mathbf{x} = (x_1, ..., x_r),\ \mathbf{y} = (y_1, ..., y_r) \in \overline{\mathbb{Q}}^r$ as usual by

$$(\mathbf{x}, \mathbf{y}) = x_1 y_1 + ... + x_r y_r.$$

We shall frequently use the following straightforward inequalities which are valid for every $v \in M_K$:

(3.1) $\quad |n_1 a_1 + ... + n_r a_r|_v \leq (|n_1| + ... + |n_r|)^{s(v)} \max(|a_1|_v, ..., |a_r|_v)$ for $n_1, ..., n_r \in \mathbb{Z}, a_1, ..., a_r \in \overline{\mathbb{Q}}$,

(3.2) $\quad |(\mathbf{x}, \mathbf{y})|_v \leq r^{s(v)} |\mathbf{x}|_v |\mathbf{y}|_v$ for $\mathbf{x}, \mathbf{y} \in \overline{\mathbb{Q}}^r$,

(3.3) $\quad |\det(\mathbf{x}_1, ..., \mathbf{x}_r)|_v \leq (r!)^{s(v)} |\mathbf{x}_1|_v \cdots |\mathbf{x}_r|_v$ for $\mathbf{x}_1, ..., \mathbf{x}_r \in \overline{\mathbb{Q}}^r$.

We can generalise (3.3) to exterior products. Let $n \in \{1, ..., r\}$ and $\mathbf{x}_1, ..., \mathbf{x}_n \in \overline{\mathbb{Q}}^r$. For $I = \{i_1, ..., i_n\}$ with $1 \leq i_1 < ... < i_n \leq r$, define the $n \times n$-determinant

$$(\mathbf{x}_1 \wedge ... \wedge \mathbf{x}_n)_I := \begin{vmatrix} x_{1,i_1} & \dots & x_{1,i_n} \\ \vdots & & \vdots \\ x_{n,i_1} & \dots & x_{n,i_n} \end{vmatrix}$$

where $\mathbf{x}_i = (x_{i1}, ..., x_{ir})$. Letting $I_1, ..., I_{\binom{r}{n}}$ be the subsets of $\{1, ..., r\}$ of cardinality n in some fixed order, we define

$$\mathbf{x}_1 \wedge ... \wedge \mathbf{x}_n := ((\mathbf{x}_1 \wedge ... \wedge \mathbf{x}_n)_{I_1}, ..., (\mathbf{x}_1 \wedge ... \wedge \mathbf{x}_n)_{I_{\binom{r}{n}}}) \in \overline{\mathbb{Q}}^{\binom{r}{n}}.$$

Note that $\mathbf{x}_1 \wedge ... \wedge \mathbf{x}_n \neq \mathbf{0}$ if and only if $\{\mathbf{x}_1, ..., \mathbf{x}_n\}$ is ($\overline{\mathbb{Q}}$-) linearly independent. Clearly (3.3) can be generalised to

(3.4) $\quad |\mathbf{x}_1 \wedge ... \wedge \mathbf{x}_n|_v \leq (n!)^{s(v)} |\mathbf{x}_1|_v ... |\mathbf{x}_n|_v$ for $v \in M_K$.

Schmidt [18] introduced the following height for a $\overline{\mathbb{Q}}$-linear subspace Y of $\overline{\mathbb{Q}}^r$: $H((0)) = H(\overline{\mathbb{Q}}^r) = 1$ and, if Y has basis $\{\mathbf{a}_1, ..., \mathbf{a}_n\}$, say, then

$$H(Y) := H(\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n).$$

This is independent of the choice of $\{\mathbf{a}_1, ..., \mathbf{a}_n\}$. For if $\{\mathbf{b}_1, ..., \mathbf{b}_n\}$ is any other basis of Y, with $\mathbf{b}_i = \sum_{j=1}^n \alpha_{ij}\mathbf{a}_j$ for $i = 1, ..., n$ where $\Delta := \det(\alpha_{ij}) \neq 0$, then $\mathbf{b}_1 \wedge ... \wedge \mathbf{b}_n = \Delta \cdot \mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n$ and this implies that $H(\mathbf{b}_1 \wedge ... \wedge \mathbf{b}_n) = H(\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n)$.

The *orthogonal complement* of Y in $\overline{\mathbb{Q}}^r$ is defined by

$$Y^\perp = \{\mathbf{c} \in \overline{\mathbb{Q}}^r : \ (\mathbf{c}, \mathbf{u}) = 0 \text{ for all } \mathbf{u} \in Y\}.$$

By [18], p. 433 we have

$$H(Y^\perp) = H(Y). \tag{3.5}$$

Express $\overline{\mathbb{Q}}^r$ as a direct sum $\overline{\mathbb{Q}}^{r_1} \oplus \overline{\mathbb{Q}}^{r_2}$ where $r = r_1 + r_2$. Suppose that $Y \subseteq \overline{\mathbb{Q}}^r$ is a direct sum of $\overline{\mathbb{Q}}$-linear subspaces $Y_1,\ Y_2$ of $\overline{\mathbb{Q}}^{r_1},\ \overline{\mathbb{Q}}^{r_2}$, respectively, i.e.

$$Y = Y_1 \oplus Y_2 = \Big\{(\mathbf{u}_1, \mathbf{u}_2) : \mathbf{u}_1 \in Y_1, \mathbf{u}_2 \in Y_2\Big\}.$$

Then

$$H(Y) = H(Y_1) \cdot H(Y_2). \tag{3.6}$$

Namely, choose bases $\{\mathbf{b}_1, ..., \mathbf{b}_{n_1}\}$, $\{\mathbf{c}_{n_1+1}, ..., \mathbf{c}_n\}$ of $Y_1,\ Y_2$, respectively; then $\{\mathbf{a}_1, ..., \mathbf{a}_n\}$ with $\mathbf{a}_i = (\mathbf{b}_i, \mathbf{0})$ for $i = 1, ..., n_1, \mathbf{a}_i = (\mathbf{0}, \mathbf{c}_i)$ for $i = n_1 + 1, ..., n$ is a basis of Y. Thus, if $(\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n)_I \neq 0$ then $I = I_1 \cup I_2$, where $I_1 \subseteq \{1, ..., r_1\}$ has cardinality n_1 and $I_2 \subseteq \{r_1 + 1, ..., r\}$ has cardinality $n - n_1$. In that case it is easy to verify that

$$(\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n)_I = (\mathbf{b}_1 \wedge ... \wedge \mathbf{b}_{n_1})_{I_1} \cdot (\mathbf{c}_{n_1+1} \wedge ... \wedge \mathbf{c}_n)_{I_2}.$$

Now (3.6) follows from the identity

$$H(x_1y_1, x_1y_2, ..., x_my_{n-1}, x_my_n) = H(x_1, ..., x_m) \cdot H(y_1, ..., y_n) \quad \text{for } x_1, ..., x_m, y_1, ..., y_n \in \overline{\mathbb{Q}}.$$

Let Σ be a $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$-action on $\{1, ..., r\}$ $(r \geq 1)$ and let Λ_Σ be the corresponding K-algebra defined by (2.1). We have $\Lambda_\Sigma \subseteq L^r$, where L is the field defined by

$$\mathrm{Gal}(\overline{\mathbb{Q}}/L) = \{\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K) : \sigma(i) = i \text{ for } i = 1, ..., r\}.$$

L is a finite, normal extension of K. If $w_1, ..., w_g$ are the places on L lying above $v \in M_K$, then there are $\sigma_1, ..., \sigma_g \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$ such that $|.|_{w_i} = |\sigma_i(.)|_v^{1/g}$ for $i = 1, ..., g$ (recall that $|\cdot|_v$ has been extended from K to $\overline{\mathbb{Q}}$). This implies that for $\lambda = (\lambda_1, ..., \lambda_r) \in \Lambda_\Sigma$,

$$\begin{aligned} |\lambda|_{w_i} = |\lambda_1, ..., \lambda_r|_{w_i} &= |\sigma_i(\lambda_1), ..., \sigma_i(\lambda_r)|_v^{1/g} \\ &= |\lambda_{\sigma_i(1)}, ..., \lambda_{\sigma_i(r)}|_v^{1/g} = |\lambda|_v^{1/g}, \end{aligned}$$

whence

$$|\lambda|_v = \prod_{\substack{w|v \\ w \in M_L}} |\lambda|_w.$$

It follows that $|\lambda|_v$ is independent of the choice of the continuation of $|.|_v$ to $\overline{\mathbb{Q}}$. Further,

$$H(\lambda) = \prod_{v \in M_K} \prod_{\substack{w|v \\ w \in M_L}} |\lambda|_w = \prod_{v \in M_K} |\lambda|_v \quad \text{for } \lambda \in \Lambda_\Sigma \tag{3.7}$$

(so $H(\lambda)$ can be defined by taking the product over $v \in M_K$ although the coordinates of λ are not all in K).

In order to define a height for K-linear subspaces of Λ_Σ, we need the following lemma.

Lemma 2. *Let W be a K-linear subspace of Λ_Σ and let $\{\mathbf{a}_1, ..., \mathbf{a}_n\}$ be a basis of W. Then $\{\mathbf{a}_1, ..., \mathbf{a}_n\}$ is $\overline{\mathbb{Q}}$-linearly independent.*

Proof. Assume the contrary. Without loss of generality we assume that for some $i < n$, $\{\mathbf{a}_1, ..., \mathbf{a}_i\}$ is a maximal $\overline{\mathbb{Q}}$-linearly independent subset of $\{\mathbf{a}_1, ..., \mathbf{a}_n\}$. Then $\mathbf{a}_{i+1} = \alpha_1\mathbf{a}_1 + ... + \alpha_i\mathbf{a}_i$ for certain, uniquely determined $\alpha_1, ..., \alpha_i \in \overline{\mathbb{Q}}$. Since every $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$ permutes the coordinates of $\mathbf{a}_j$ ($j = 1, ..., n$) we have $\mathbf{a}_{i+1} = \sigma(\alpha_1)\mathbf{a}_1 + ... + \sigma(\alpha_i)\mathbf{a}_i$ for $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$. So by the unicity of α_j, $\sigma(\alpha_j) = \alpha_j$ for every $\sigma \in \mathrm{Gal}(\mathbb{Q}/K)$, i.e. $\alpha_j \in K$ for $j = 1, ..., i$. Hence $\{\mathbf{a}_1, ..., \mathbf{a}_{i+1}\}$ is K-linearly dependent. But this contradicts that $\{\mathbf{a}_1, ..., \mathbf{a}_n\}$ is a basis of W. □

Now we define the height of a K-linear subspace W of Λ_Σ by

$$H(W) := H(\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n), \tag{3.8}$$

where $\{\mathbf{a}_1, ..., \mathbf{a}_n\}$ is any basis of W. Under the action Σ, every $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$ maps a subset I of $\{1, ..., r\}$ of cardinality n to another such subset but it does not necessarily preserve the increasing order. It follows that every σ permutes, up to signs, the coordinates of $\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n$. Since signs do not affect the absolute values, we can repeat the above arguments leading to (3.7) and conclude that for $v \in M_K$ the quantity $|\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n|_v$ does not depend on the choice of the continuation of v to $\overline{\mathbb{Q}}$, and that the height can be obtained by taking the product over $v \in M_K$,

$$H(W) = \prod_{v \in M_K} |\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n|_v. \tag{3.9}$$

§4. The quantitative Subspace theorem.

Our main tool is the quantitative Subspace theorem from [6] which we recall below. This is an improvement of the quantitative Subspace theorem of Schmidt [21] and its p-adic generalisation by Schlickewei [15].

For a linear form $l(\mathbf{X}) := \alpha_1 X_1 + ... + \alpha_n X_n$ where $\mathbf{a} := (\alpha_1, ..., \alpha_n) \in \overline{\mathbb{Q}}^n$ is non-zero and for a number field K put

$$H(l) := H(\mathbf{a}), \quad K(l) := K(\alpha_1/\alpha_i, ..., \alpha_n/\alpha_i)$$

where i is any index from $\{1, ..., n\}$ with $\alpha_i \neq 0$. For any field K, any finite-dimensional K-vector space V, and any subset $\mathcal{S}$ of V, the *linear scattering* of $\mathcal{S}$ in V is defined as the smallest integer h for which there are proper K-linear subspaces $W_1, ..., W_h$ of V with $\mathcal{S} \subset W_1 \cup ... \cup W_h$; if such an integer h does not exist, then the linear scattering of $\mathcal{S}$ in V is defined to be ∞. For instance, $\mathcal{S}$ has linear scattering ≥ 2 in V if and only if $\mathcal{S}$ contains a basis of V.

Now let K be an algebraic number field, S a finite set of places on K containing all infinite places, n an integer ≥ 2, δ a real with $0 < \delta < 1$ and for $v \in S$, $\{l_{1v}, ..., l_{nv}\}$ a linearly independent set of linear forms in n variables with algebraic coefficients such that

$$H(l_{iv}) \leq H, \quad [K(l_{iv}) : K] \leq D \quad \text{for } v \in S, \ i = 1, ..., n, \tag{4.1}$$

where $H \geq 1, D \geq 1$. By $\det(l_1, ..., l_n)$ we denote the coefficient determinant of n linear forms $l_1, ..., l_n$ in n variables.

Lemma 3. *The set of* $\mathbf{x} \in K^n$ *with*

$$\begin{cases} 0 < \displaystyle\prod_{v \in S} \prod_{i=1}^{n} \frac{|l_{iv}(\mathbf{x})|_v}{|\mathbf{x}|_v} \leq n^{-\frac{1}{2}(n+\delta)} \Big\{ \prod_{v \in S} |\det(l_{1v}, ..., l_{nv})|_v \Big\} H(\mathbf{x})^{-n-\delta}, \\ H(\mathbf{x}) \geq n^{1/2} H \end{cases} \tag{4.2}$$

has linear scattering in K^n *at most*

$$\left\{2^{60n^2}\delta^{-7n}\right\}^s \log 4D \cdot \log\log 4D \ .$$

Proof. For $\mathbf{x} = (x_1, ..., x_n) \in K^n, v \in M_K$, put $||\mathbf{x}||_v := |\mathbf{x}|_v$ if v is finite and $||\mathbf{x}||_v := (|x_1|^{2/s(v)} + ... + |x_n|_v^{2/s(v)})^{s(v)/2}$ if v is infinite, and define the Euclidean height $H_2(\mathbf{x}) := \prod_{v\in M_K} ||\mathbf{x}||_v$. This height depends only on $\mathbf{x}$ and not on K; hence the Euclidean height can be extended uniquely to a height $H_2 : \overline{\mathbb{Q}}^n \to \mathbb{R}$. It is easy to see that $|\mathbf{x}|_v \leq ||\mathbf{x}||_v \leq n^{s(v)/2}|\mathbf{x}|_v$ for $\mathbf{x} \in \overline{\mathbb{Q}}^n, v \in M_K$; hence $H(\mathbf{x}) \leq H_2(\mathbf{x}) \leq n^{1/2}H(\mathbf{x})$ for $\mathbf{x} \in \overline{\mathbb{Q}}^n$. Let $H_2 := \max_{i,v} H_2(l_{iv})$, where $H_2(l_{iv})$ is the Euclidean height of the coefficient vector of l_{iv}. Clearly, every $\mathbf{x} \in K^n$ with (4.2) satisfies

$$(4.3) \qquad \begin{cases} 0 < \displaystyle\prod_{v\in S}\prod_{i=1}^{n} \frac{|l_{iv}(\mathbf{x})|_v}{||\mathbf{x}||_v} \leq \prod_{v\in S} |\det(l_{1v}, ..., l_{nv})|_v \cdot H_2(\mathbf{x})^{-n-\delta}, \\ H_2(\mathbf{x}) \geq H_2, \end{cases}$$

and by the Theorem of [6], the set of $\mathbf{x} \in K^n$ with (4.3) has linear scattering at most $\{2^{60n^2}\delta^{-7n}\}^s \log 4D \cdot \log\log 4D$. □

Let K be a number field and S a finite set of places as above, and let Σ be a $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$-action on $\{1, ..., r\}$ where $r \geq 2$. Further, let W be an n-dimensional K-linear subspace of Λ_Σ, with $n \geq 2$. We need an analogue of Lemma 3 for W. As before, we denote vectors of W by $\mathbf{u} = (u_1, ..., u_r)$. A non-empty subset I of $\{1, ..., r\}$ is called *independent* if $\{u_i : i \in I\}$ is $\overline{\mathbb{Q}}$-linearly independent on W, that is, for $c_i \in \overline{\mathbb{Q}}$ $(i \in I)$ we have

$$(4.4) \qquad \sum_{i\in I} c_i u_i = 0 \text{ for all } \mathbf{u} \in W \Longrightarrow c_i = 0 \text{ for } i \in I.$$

Let $\{\mathbf{a}_1 = (a_{11}, ..., a_{1r}), ..., \mathbf{a}_n = (a_{n1}, ..., a_{nr})\}$ be a basis of W. Then for a subset I of $\{1, ..., r\}$ of cardinality n we have that I is independent$\Longleftrightarrow \sum_{i\in I} c_i u_i = 0$ for $\mathbf{u} = \mathbf{a}_1, ..., \mathbf{a}_n$ implies that $c_i = 0$ for $i \in I \Longleftrightarrow (\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n)_I =$ $\det((a_{ij})_{i=1,...,n,\ j\in I}) \neq 0$. Now Lemma 2 implies that $\{1, ..., r\}$ has independent subsets of cardinality n.

In what follows, by I_v $(v \in S)$ we always denote independent subsets of $\{1, ..., r\}$ of cardinality n, and by $\mathbf{I} = (I_v : v \in S)$ a collection of such subsets. For such a collection $\mathbf{I}$ we define

$$\Delta(\mathbf{I}, W) := \prod_{v\in S} |(\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n)_{I_v}|_v \cdot \prod_{v\notin S} |\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n|_v;$$

this quantity plays the role of $\prod_{v\in S} |\det(l_{1v}, ..., l_{nv})|_v$. $\Delta(\mathbf{I}, W)$ does not depend on the choice of the basis $\{\mathbf{a}_1, ..., \mathbf{a}_n\}$ of W. Namely, let $\{\mathbf{b}_1, ..., \mathbf{b}_n\}$ be another basis of W, where $\mathbf{b}_i = \sum_{j=1}^n \alpha_{ij}\mathbf{a}_j$ for $i = 1, ..., n$, and put $\Delta := \det(\alpha_{ij})$. Then for each subset I of $\{1, ..., r\}$ of cardinality n we have $(\mathbf{b}_1 \wedge ... \wedge \mathbf{b}_n)_I = \Delta \cdot (\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n)_I$; together with the Product formula this implies that

$$\prod_{v\in S} |(\mathbf{b}_1 \wedge ... \wedge \mathbf{b}_n)_{I_v}|_v \cdot \prod_{v\notin S} |\mathbf{b}_1 \wedge ... \wedge \mathbf{b}_n|_v = \prod_{v\in M_K} |\Delta|_v \cdot \Delta(\mathbf{I}, W) = \Delta(\mathbf{I}, W) \ .$$

Lemma 4. *Let n, r be integers ≥ 2, δ a real with $0 < \delta < 1$, K an algebraic number field, S a finite set of places on K of cardinality s containing all infinite places, Σ a $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$-action on $\{1, ..., r\}$, W an n-dimensional K-linear subspace of Λ_Σ, and $\mathbf{I} = (I_v : v \in S)$ a collection of independent subsets of $\{1, ..., r\}$ of cardinality n. Then the set of $\mathbf{u} = (u_1, ..., u_r) \in W$ with*

$$\prod_{v\in S}\prod_{i\in I_v} \frac{|u_i|_v}{|\mathbf{u}|_v} \leq \Delta(\mathbf{I}, W) \cdot H(\mathbf{u})^{-n-\delta} \tag{4.5}$$

has linear scattering in W at most

$$\left\{2^{61n^2} r^{2n} \delta^{-7n}\right\}^s.$$

In the proof of Lemma 4 we use Lemma 3 and a gap principle. This gap principle is developed in §5. Lemma 4 will be proved in §6.

§5. A gap principle.

Let again K be an algebraic number field with $[K : \mathbb{Q}] = d$, S a finite set of places on K of cardinality s containing all infinite places, Σ a $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$- action on $\{1, ..., r\}$ where $r \geq 2$, and W an n-dimensional K-linear subspace of Λ_Σ, where $n \geq 2$. Further, $\mathbf{I} = (I_v : v \in S)$ is a collection of independent subsets of $\{1, ..., r\}$ of cardinality n. We shall estimate the linear scattering in W of sets of $\mathbf{u} \in W$ with

$$\prod_{v\in S}\prod_{i\in I_v} \frac{|u_i|_v}{|\mathbf{u}|_v} \leq \Delta(\mathbf{I}, W) H(\mathbf{u})^{-n} \cdot \frac{Q}{P}, \qquad H(\mathbf{u}) < B, \tag{5.1}$$

where $P \geq 1, Q \geq 1, B \geq 2P$. First we prove some auxiliary results. By $|\mathcal{S}|$ we denote the cardinality of a set $\mathcal{S}$.

Lemma 5. *Let θ be a real with $0 < \theta \leq \frac{1}{2}$ and q an integer ≥ 1. Then there exists a set Γ with the following properties:*
(i). $|\Gamma| \leq \{e(2 + \theta^{-1})\}^q$, where $e = 2.7182...$;

(ii). Γ *consists of* q*-tuples* $\gamma = (\gamma_1, ..., \gamma_q)$ *with* $\gamma_i \geq 0$ *for* $i = 1, ..., q$*;*
(iii). for every set of reals $F_1, ..., F_q, M$ *with*

$$0 < F_i \leq 1, \quad F_1 \cdots F_q \geq M, \quad M < 1 \tag{5.2}$$

there is a tuple $\gamma \in \Gamma$ *with*

$$M^{\gamma_i + (\theta/q)} < F_i \leq M^{\gamma_i} \quad \textit{for} \quad i = 1, ..., q. \tag{5.3}$$

Proof. This is Lemma 9 (ii) of [6]. □

Lemma 6. *Let* F *be a real* ≥ 1 *and let* $\mathcal{S}$ *be a subset of* W *of linear scattering in* W *at least*

$$\max\left(6F^{5d}, 4 \times 7^{d+2s}\right).$$

Then there are $\mathbf{u}_1, ..., \mathbf{u}_n \in \mathcal{S}$ *with*

$$0 < \prod_{v \notin S} \frac{|\mathbf{u}_1 \wedge ... \wedge \mathbf{u}_n|_v}{|\mathbf{u}_1|_v ... |\mathbf{u}_n|_v} \leq F^{-1}. \tag{5.4}$$

Proof. This is a variation on Lemma 10 of [6]. Let $\wp$ be a prime ideal outside S of which the norm $N\wp$ is minimal. We assume that $\mathcal{S}$ satisfies the condition of Lemma 6 and moreover that $\mathbf{0} \notin \mathcal{S}$ and that

$$(N\wp)^{-1/d} < |\mathbf{u}|_\wp \leq 1 \quad \text{for } \mathbf{u} \in \mathcal{S}. \tag{5.5}$$

This is no restriction. Namely, for non-zero $\mathbf{u} \in W$ we have $|\mathbf{u}|_\wp = (N\wp)^{-\alpha/d}$ with $\alpha \in \mathbb{Q}$. Write $\alpha = m + \beta$ with $m \in \mathbb{Z}$ and $0 \leq \beta < 1$, choose $\lambda_\mathbf{u} \in K^*$ with $|\lambda_\mathbf{u}|_\wp = (N\wp)^{-m/d}$, and put $\mathbf{u}' := \lambda_\mathbf{u}^{-1}\mathbf{u}$. Then $\mathbf{u}' \in W$ and $(N\wp)^{-1/d} < |\mathbf{u}'|_\wp \leq 1$. Since the left-hand side of (5.4) does not change when $\mathbf{u}_1, ..., \mathbf{u}_n$ are multiplied with scalars from K^*, it suffices to prove Lemma 6 for $\mathcal{S}' := \{\mathbf{u}' : \mathbf{u} \in \mathcal{S}\}$ instead of $\mathcal{S}$.
We assume that Lemma 6 is false, i.e. for all linearly independent $\mathbf{u}_1, ..., \mathbf{u}_n \in \mathcal{S}$ we have

$$\prod_{v \in S} \frac{|\mathbf{u}_1 \wedge ... \wedge \mathbf{u}_n|_v}{|\mathbf{u}_1|_v ... |\mathbf{u}_n|_v} > F^{-1}. \tag{5.6}$$

Fix a linearly independent subset $\{\mathbf{u}_1, ..., \mathbf{u}_n\}$ of $\mathcal{S}$ (this exists since $\mathcal{S}$ has linear scattering ≥ 2). For $\mathbf{y}_1, ..., \mathbf{y}_m \in W$, let $[\mathbf{y}_1, ..., \mathbf{y}_m]$ be the K-linear subspace of W generated by $\mathbf{y}_1, ..., \mathbf{y}_m$. In the remainder of the proof we consider only those $\mathbf{x} \in \mathcal{S}$ with

$$\mathbf{x} \notin [\mathbf{u}_1, ..., \mathbf{u}_{n-1}], \quad \mathbf{x} \notin [\mathbf{u}_1, ..., \mathbf{u}_{n-2}, \mathbf{u}_n]; \tag{5.7}$$

thus, we exclude at most two proper linear subspaces of W. We write $\mathbf{x} = x_1\mathbf{u}_1 + ... + x_n\mathbf{u}_n$ with $x_1, ..., x_n \in K$. For $i = 1, 2, ...$ $\{C_{vi} : v \notin S\}$ will denote a set of numbers of the form $\{|\mathbf{x}_1 \wedge ... \wedge \mathbf{x}_n|_v / |\mathbf{x}_1|_v ... |\mathbf{x}_n|_v : v \notin S\}$ where $\mathbf{x}_1, ..., \mathbf{x}_n$ are any linearly independent elements from $\mathcal{S}$ which may be different for each i; by (5.6) and (3.4) such numbers satisfy

$$0 < C_{vi} \le 1 \quad \text{for } v \notin S, \quad \prod_{v \notin S} C_{vi} > F^{-1}. \tag{5.8}$$

Fix $\mathbf{x} \in \mathcal{S}$ with (5.7). Then $\{\mathbf{u}_1, ..., \mathbf{u}_{n-1}, \mathbf{x}\}$, $\{\mathbf{u}_1, ..., \mathbf{u}_{n-2}, \mathbf{u}_n, \mathbf{x}\}$ are linearly independent. Hence for $v \in M_K \backslash S$ we have

$$\begin{cases} |x_{n-1}|_v = \dfrac{|\mathbf{u}_1 \wedge ... \wedge \mathbf{u}_{n-2} \wedge \mathbf{u}_n \wedge \mathbf{x}|_v}{|\mathbf{u}_1 \wedge ... \wedge \mathbf{u}_n|_v} = \dfrac{C_{v1}}{C_{v2}} \dfrac{|\mathbf{x}|_v}{|\mathbf{u}_{n-1}|_v}, \\ |x_n|_v = \dfrac{|\mathbf{u}_1 \wedge ... \wedge \mathbf{u}_{n-1} \wedge \mathbf{x}|_v}{|\mathbf{u}_1 \wedge ... \wedge \mathbf{u}_n|_v} = \dfrac{C_{v3}}{C_{v2}} \dfrac{|\mathbf{x}|_v}{|\mathbf{u}_n|_v}. \end{cases} \tag{5.9}$$

Let $\mathbf{y}$ be any vector in $\mathcal{S}$ satisfying

$$\mathbf{y} \notin [\mathbf{u}_1, ..., \mathbf{u}_{n-1}],\ \mathbf{y} \notin [\mathbf{u}_1, ..., \mathbf{u}_{n-2}, \mathbf{u}_n],\ \mathbf{y} \notin [\mathbf{u}_1, ..., \mathbf{u}_{n-2}, \mathbf{x}] \tag{5.10}$$

and write $\mathbf{y} = y_1\mathbf{u}_1 + ... + y_n\mathbf{u}_n$ with $y_1, ..., y_n \in K$. Similarly to (5.9) we have

$$|y_{n-1}|_v = \frac{C_{v4}}{C_{v2}} \frac{|\mathbf{y}|_v}{|\mathbf{u}_{n-1}|_v}, \quad |y_n|_v = \frac{C_{v5}}{C_{v2}} \frac{|\mathbf{y}|_v}{|\mathbf{u}_n|_v} \quad \text{for } v \notin S. \tag{5.11}$$

Further, since $\{\mathbf{u}_1, ..., \mathbf{u}_{n-2}, \mathbf{x}, \mathbf{y}\}$ is linearly independent we have

$$\begin{aligned} |x_{n-1}y_n - x_n y_{n-1}|_v &= \frac{|\mathbf{u}_1 \wedge ... \wedge \mathbf{u}_{n-2} \wedge \mathbf{x} \wedge \mathbf{y}|_v}{|\mathbf{u}_1 \wedge ... \wedge \mathbf{u}_n|_v} \\ &= \frac{C_{v6}}{C_{v2}} \frac{|\mathbf{x}|_v |\mathbf{y}|_v}{|\mathbf{u}_{n-1}|_v |\mathbf{u}_n|_v} \quad \text{for } v \notin S. \end{aligned} \tag{5.12}$$

Now (5.9), (5.11), (5.12), (5.8) imply

$$\begin{aligned} &\frac{|x_{n-1}y_n|_v}{\max(|x_{n-1}y_n|_v, |x_n y_{n-1}|_v, |x_{n-1}y_n - x_n y_{n-1}|_v)} \\ &\qquad = \frac{C_{v1}C_{v5}}{\max(C_{v1}C_{v5}, C_{v3}C_{v4}, C_{v2}C_{v6})} \ge C_{v1}C_{v5}\ . \end{aligned}$$

Hence, again by (5.8),

$$\prod_{v \notin S} \frac{|x_{n-1}y_n|_v}{\max(|x_{n-1}y_n|_v, |x_n y_{n-1}|_v, |x_{n-1}y_n - x_n y_{n-1}|_v)} > F^{-2}. \tag{5.13a}$$

Similarly, we have

$$\text{(5.13b)} \qquad \prod_{v \notin S} \frac{|x_n y_{n-1}|_v}{\max(|x_{n-1}y_n|_v, |x_n y_{n-1}|_v, |x_{n-1}y_n - x_n y_{n-1}|_v)} > F^{-2},$$

$$\text{(5.13c)} \qquad \prod_{v \notin S} \frac{|x_{n-1}y_n - x_n y_{n-1}|_v}{\max(|x_{n-1}y_n|_v, |x_n y_{n-1}|_v, |x_{n-1}y_n - x_n y_{n-1}|_v)} > F^{-2}.$$

Define the integer m by

$$\text{(5.14)} \qquad (N\wp)^{(m-1)/d} < F^2 \leq (N\wp)^{m/d}.$$

From (5.6) and (3.4) it follows that $F > 1$. Hence $m \geq 1$. We distinguish two cases.

First case: $m = 1$.
From the definition of the absolute values it follows that the left-hand side of each of the inequalities (5.13 a, b, c) is of the form $(N\mathfrak{a})^{-1/d}$, where $\mathfrak{a}$ is an ideal of $\mathcal{O}_K$ composed of prime ideals outside S. This ideal is integral since for every $v \notin S$, the corresponding factors in the left-hand sides of (5.13 a, b, c) are ≤ 1. We have $\mathfrak{a} = (1)$ for otherwise, since $\wp \notin S$ was chosen to have minimal norm and since $m = 1$, we would have had $(N\mathfrak{a})^{-1/d} \leq (N\wp)^{-1/d} \leq F^{-2}$. Therefore, the left-hand sides of (5.13 a, b, c) are equal to 1. But then, the factors in the products on the left-hand sides of (5.13 a,b,c) are also equal to 1. It follows that if $\mathbf{x}, \mathbf{y} \in \mathcal{S}$ satisfy (5.7), (5.10), respectively, then

$$|x_{n-1}y_n|_v = |x_n y_{n-1}|_v = |x_n y_{n-1} - x_{n-1}y_n|_v \quad \text{for } v \notin S.$$

Hence $u := x_{n-1}y_n / x_n y_{n-1}$ satisfies $|u|_v = |1 - u|_v = 1$ for $v \notin S$, i.e. $u, 1 - u \in \mathcal{O}_S^*$ (note that $u \in K$). By Theorem 1 of [3], there are at most $3 \times 7^{d+2s}$ possibilities for u. Fix $\mathbf{x} \in \mathcal{S}$ with (5.7). Then there is a set U of cardinality $\leq 3 \times 7^{d+2s}$ such that every $\mathbf{y} \in \mathcal{S}$ with (5.10) satisfies $y_{n-1}/y_n \in U$. In other words, $\mathcal{S}$ is contained in the union of the proper linear subspaces $[\mathbf{u}_1, ..., \mathbf{u}_{n-1}], [\mathbf{u}_1, ..., \mathbf{u}_{n-2}, \mathbf{u}_n], [\mathbf{u}_1, ..., \mathbf{u}_{n-2}, \mathbf{x}], [\mathbf{u}_1, ..., \mathbf{u}_{n-2}, \alpha\mathbf{u}_{n-1} + \mathbf{u}_n]$ $(\alpha \in U)$ of W. Hence $\mathcal{S}$ has linear scattering in W at most $3 + 3 \times 7^{d+2s} < 4 \times 7^{d+2s}$. This contradicts our assumption on $\mathcal{S}$ in the statement of Lemma 6; therefore, our assumption that Lemma 6 is false was incorrect.

Second case: $m \geq 2$.
We consider again $\mathbf{x} = x_1\mathbf{u}_1 + ... + x_n\mathbf{u}_n \in \mathcal{S}$ with (5.7). By (5.9) with $v = \wp$ we have $|x_{n-1}/x_n|_\wp = C_{\wp 1}|\mathbf{u}_n|_\wp / C_{\wp 3}|\mathbf{u}_{n-1}|_\wp$. Together with (5.5), (5.8), (5.14) this implies

$$(N\wp)^{-(1+\frac{1}{2}m)/d} \leq F^{-1}(N\wp)^{-1/d} < \left|\frac{x_{n-1}}{x_n}\right|_\wp < F(N\wp)^{1/d} \leq (N\wp)^{(1+\frac{1}{2}m)/d}.$$

We have $|x_{n-1}/x_n|_\wp = (N\wp)^{k_\mathbf{x}/d}$ for some $k_\mathbf{x} \in \mathbb{Z}$. Clearly,

$$-1 - \frac{1}{2}m < k_\mathbf{x} < 1 + \frac{1}{2}m.$$

Therefore, for $k_\mathbf{x}$ there are at most $m+2$ possibilities.

Fix $\pi \in \wp$ such that $|\pi|_\wp = (N\wp)^{-1/d}$. Then

$$x_{n-1} = x_n \pi^{k_\mathbf{x}} \alpha_\mathbf{x}, \quad \text{with } \alpha_\mathbf{x} \in K,\ |\alpha_\mathbf{x}|_\wp = 1.$$

We call $\mathbf{x}, \mathbf{y} \in \mathcal{S}$ *congruent* if they both satisfy (5.7) and if $k_\mathbf{x} = k_\mathbf{y}$ and $\alpha_\mathbf{x} \equiv \alpha_\mathbf{y} \pmod{\wp^m}$, i.e. $|\alpha_\mathbf{x} - \alpha_\mathbf{y}|_\wp \le (N\wp)^{-m/d}$. We claim that each congruence class has linear scattering at most 1. Namely, fix $\mathbf{x}$ from a given congruence class, and take $\mathbf{y}$ from the same class. Suppose that $\mathbf{y} \notin [\mathbf{u}_1, ..., \mathbf{u}_{n-2}, \mathbf{x}]$. Then

$$\begin{aligned}
&\frac{|x_{n-1}y_n - x_n y_{n-1}|_\wp}{\max(|x_{n-1}y_n|_\wp, |x_n y_{n-1}|_\wp, |x_{n-1}y_n - x_n y_{n-1}|_\wp)} \\
&\qquad = \frac{|\pi|_\wp^{k_\mathbf{x}} |x_n y_n|_\wp |\alpha_\mathbf{x} - \alpha_\mathbf{y}|_\wp}{|\pi|_\wp^{k_\mathbf{x}} |x_n y_n|_\wp \max(|\alpha_\mathbf{x}|_\wp, |\alpha_\mathbf{y}|_\wp, |\alpha_\mathbf{x} - \alpha_\mathbf{y}|_\wp)} \\
&\qquad = \frac{|\alpha_\mathbf{x} - \alpha_\mathbf{y}|_\wp}{\max(1, |\alpha_\mathbf{x} - \alpha_\mathbf{y}|_\wp)} \le (N\wp)^{-m/d} \le F^{-2}.
\end{aligned}$$

Since the other factors in the left-hand side of (5.13c) are ≤ 1, this contradicts (5.13c). It follows that every $\mathbf{y} \in \mathcal{S}$ with (5.7) which is congruent to $\mathbf{x}$ belongs to $[\mathbf{u}_1, ..., \mathbf{u}_{n-2}, \mathbf{x}]$. This proves our claim.
By $m \ge 2$ and (5.14) we have $(N\wp)^m < F^{4d}$. Taking into consideration the number of possibilities for $k_\mathbf{x}$, it follows that the number of congruence classes is

$$< (m+2)(N\wp)^m < \left(\frac{4d \log F}{\log N\wp} + 2\right) F^{4d}.$$

Hence $\mathcal{S}$ has linear scattering

$$< 2 + \left(\frac{4d \log F}{\log 2} + 2\right) F^{4d} < 6F^{5d};$$

here we must add 2 because of (5.7). So again, the assumption that Lemma 6 is false leads to a contradiction with the condition on $\mathcal{S}$. We conclude that Lemma 6 holds true in both cases $m = 1$ and $m \ge 2$. □

Below we state our gap principle:

Lemma 7. *Let P, Q, B be reals with $P \ge 1, Q \ge 1, B \ge 2P$. Then the set of $\mathbf{u} \in W$ with*

$$\prod_{v \in S} \prod_{i \in I_v} \frac{|u_i|_v}{|\mathbf{u}|_v} \le \Delta(\mathbf{I}, W) H(\mathbf{u})^{-n} . \frac{Q}{P}, \quad H(\mathbf{u}) < B \tag{5.1}$$

has linear scattering in W at most $Q^{5d}\left(270n^{10}r\cdot\dfrac{\log B}{\log 2P}\right)^{ns+1}$.

Proof. Let $\mathcal{F}$ be the set of $\mathbf{u}\in W$ satisfying (5.1) and also $u_1\cdots u_r\neq 0$. Thus, from W we exclude solutions of (5.1) belonging to at most r proper linear subspaces of W. For $\mathbf{u}\in\mathcal{F}$ we have $u_1\cdots u_r\in K^*$, so by the Product formula and (3.7),

$$(5.15)\quad \prod_{v\in S}\prod_{i\in I_v}\frac{|u_i|_v}{|\mathbf{u}|_v}\geq\prod_{v\in S}\frac{|u_1\ldots u_r|_v}{|\mathbf{u}|_v^r}=\left(\prod_{v\in S}|\mathbf{u}|_v\right)^{-r}\cdot\left(\prod_{v\notin S}|u_1\ldots u_r|_v\right)^{-1}$$
$$\geq\left(\prod_{v\in M_K}|\mathbf{u}|_v\right)^{-r}=H(\mathbf{u})^{-r}\geq B^{-r}.$$

Put

$$\kappa:=\frac{\log B}{\log 2P},\quad \theta:=\frac{1}{2r\kappa},\quad N:=[(2n-2)\kappa]+1.$$

By (5.15) and Lemma 5 with $M=B^{-r}$, there is a set Γ of ns-tuples of non-negative reals $\gamma=(\gamma_{iv}:v\in S,i\in I_v)$ of cardinality

$$|\Gamma|\leq\{e(2+\theta^{-1})\}^{ns}$$

such that for every $\mathbf{u}\in\mathcal{F}$ there is a tuple $\gamma\in\Gamma$ with

$$(5.16)\qquad (B^{-r})^{\gamma_{iv}+(\theta/ns)}<\frac{|u_i|_v}{|\mathbf{u}|_v}\leq(B^{-r})^{\gamma_{iv}}\quad\text{for } v\in S,\ i\in I_v\ .$$

Further, for every $\mathbf{u}\in\mathcal{F}$ there is a $j\in\{1,...,N\}$ such that

$$(5.17)\qquad B^{(j-1)/N}\leq H(\mathbf{u})<B^{j/N}\ .$$

Let $\mathcal{F}(\gamma,j)$ be the set of $\mathbf{u}\in\mathcal{F}$ satisfying (5.16), (5.17). Since $\kappa=\log B/\log 2P\geq 1$, the number of sets $\mathcal{F}(\gamma,j)$ is at most

$$|\Gamma|\cdot N\leq\{e(2+2r\kappa)\}^{ns}\{(2n-2)\kappa+1\}\leq 2n\kappa\cdot(3er\kappa)^{ns}.$$

We shall show that each set $\mathcal{F}(\gamma,j)$ has linear scattering in W

$$(5.18)\qquad <\max(6F^{5d},4\times 7^{d+2s})\quad\text{with } F:=2n!Q.$$

Since $d\leq 2s$ we have $\max(6F^{5d},4\times 7^{d+2s})\leq 6(2n!)^{10s}Q^{5d}$. Assuming (5.18), it follows that $\mathcal{F}$ has linear scattering in W at most

$$2n\kappa(3er\kappa)^{ns}\cdot 6(2n!)^{10s}Q^{5d}\leq Q^{5d}\cdot 12n\kappa(n!)^{10s}(96er\kappa)^{ns}$$

$$\leq Q^{5d}\left(270n^{10}r\frac{\log B}{\log 2P}\right)^{ns+1} - r;$$

taking into consideration the solutions with $u_1 \cdots u_r = 0$, this implies Lemma 7.

So it remains to prove that the linear scattering of $\mathcal{F}(\gamma, j)$ is smaller than the bound in (5.18). Suppose the contrary. Then by Lemma 6, there are linearly independent $\mathbf{u}_1, ..., \mathbf{u}_n \in \mathcal{F}(\gamma, j)$ with

$$\prod_{v \in M_K \setminus S} \frac{|\mathbf{u}_1 \wedge ... \wedge \mathbf{u}_n|_v}{|\mathbf{u}_1|_v ... |\mathbf{u}_n|_v} \leq (2n!)^{-1} Q^{-1}. \tag{5.19}$$

Take $v \in S$. Let $\mathbf{u}_i = (u_{i1}, ..., u_{ir})$ for $i = 1, ..., n$. Suppose for convenience that $I_v = \{1, ..., n\}$. Then by (5.16),

$$\begin{aligned}
\frac{|(\mathbf{u}_1 \wedge ... \wedge \mathbf{u}_n)_{I_v}|_v}{|\mathbf{u}_1|_v ... |\mathbf{u}_n|_v} &\leq (n!)^{s(v)} \cdot \max_{\tau} \frac{|u_{\tau(1),1}|_v}{|\mathbf{u}_{\tau(1)}|_v} \cdots \frac{|u_{\tau(n),n}|_v}{|\mathbf{u}_{\tau(n)}|_v} \\
&\leq (n!)^{s(v)} \big(B^{-r}\big)^{\sum_{i \in I_v} \gamma_{iv}}
\end{aligned}$$

where the maximum is taken over all permutations τ of $(1, ..., n)$. By taking the product over $v \in S$ we obtain, using the first inequality of (5.16) with $\mathbf{u}_1$ replacing $\mathbf{u}$ and $\kappa = \log B / \log 2P$, $\theta = 1/2r\kappa$,

$$\begin{aligned}
\prod_{v \in S} \frac{|(\mathbf{u}_1 \wedge ... \wedge \mathbf{u}_n)_{I_v}|_v}{|\mathbf{u}_1|_v ... |\mathbf{u}_n|_v} &\leq n! \big(B^{-r}\big)^{\sum_{i,v} \gamma_{iv}} \\
&< n! \left(\prod_{v \in S} \prod_{i \in I_v} \frac{|u_{1i}|_v}{|\mathbf{u}_1|_v}\right) B^{r\theta} \\
&\leq n! \Delta(\mathbf{I}, W) H(\mathbf{u}_1)^{-n} Q P^{-1} B^{1/2\kappa} \\
&= 2n! Q \Delta(\mathbf{I}, W) H(\mathbf{u}_1)^{-n} B^{-1/2\kappa}.
\end{aligned} \tag{5.20}$$

Since $\{\mathbf{u}_1, ..., \mathbf{u}_n\}$ is a basis of W we have, in view of (3.7),

$$\begin{aligned}
\Delta(\mathbf{I}, W) &= \prod_{v \in S} |(\mathbf{u}_1 \wedge ... \wedge \mathbf{u}_n)_{I_v}|_v \cdot \prod_{v \notin S} |\mathbf{u}_1 \wedge ... \wedge \mathbf{u}_n|_v \\
&= \prod_{v \in S} \frac{|(\mathbf{u}_1 \wedge ... \wedge \mathbf{u}_n)_{I_v}|_v}{|\mathbf{u}_1|_v \cdots |\mathbf{u}_n|_v} \cdot \prod_{v \notin S} \frac{|\mathbf{u}_1 \wedge ... \wedge \mathbf{u}_n|_v}{|\mathbf{u}_1|_v \cdots |\mathbf{u}_n|_v} \cdot H(\mathbf{u}_1) \cdots H(\mathbf{u}_n) \ .
\end{aligned}$$

Together with (5.20), (5.19), (5.17) and $N = [(2n-2)\kappa + 1]$ this implies that

$$\begin{aligned}
\Delta(\mathbf{I}, W) &< 2n! Q \Delta(\mathbf{I}, W) H(\mathbf{u}_1)^{-n} B^{-1/2\kappa} \cdot (2n!)^{-1} Q^{-1} \cdot H(\mathbf{u}_1) H(\mathbf{u}_2) ... H(\mathbf{u}_n) \\
&= \Delta(\mathbf{I}, W) B^{-1/2\kappa} H(\mathbf{u}_1)^{1-n} H(\mathbf{u}_2) ... H(\mathbf{u}_n) \\
&< \Delta(\mathbf{I}, W) B^{-1/2\kappa} \cdot B^{(1-n)(j-1)/N} B^{(n-1)j/N} \\
&= \Delta(\mathbf{I}, W) \cdot B^{-1/2\kappa + (n-1)/N} \\
&\leq \Delta(\mathbf{I}, W),
\end{aligned}$$

which is impossible. Thus, our assumption that the linear scattering of $\mathcal{F}(\gamma, j)$ is larger than the bound in (5.18) leads to a contradiction. This completes the proof of Lemma 7. $\square$

§6. Proof of Lemma 4.

Let again K be an algebraic number field, S a finite set of places on K of cardinality s containing all infinite places, Σ a Gal$(\overline{\mathbb{Q}}/K)$-action on $\{1, ..., r\}$ where $r \geq 2$, W an n-dimensional K-linear subspace of Λ_Σ where $n \geq 2$ and $\mathbf{I} = (I_v : v \in S)$ a collection of independent subsets of $\{1, ..., r\}$ of cardinality n. Further, let $0 < \delta < 1$. We consider the solutions of

$$\text{(4.5)} \qquad \prod_{v \in S} \prod_{i \in I_v} \frac{|u_i|_v}{|\mathbf{u}|_v} \leq \Delta(\mathbf{I}, W) \cdot H(\mathbf{u})^{-n-\delta} \quad \text{in } \mathbf{u} \in W \ .$$

We shall distinguish between "large" and "small" solutions. The small solutions are treated by the gap principle. We shall deal with the large solutions $\mathbf{u}$ of (4.5) by choosing a suitable basis $\{\mathbf{a}_1, ..., \mathbf{a}_n\}$ for W and then showing that the corresponding vectors $\mathbf{x} = (x_1, ..., x_n) \in K^n$ defined by $\mathbf{u} = x_1\mathbf{a}_1 + ... + x_n\mathbf{a}_n$ satisfy an inequality of type (4.2) to which Lemma 3 is applicable.

We first choose the basis. Put

$$W_0 := \{\mathbf{u} = (u_1, ..., u_r) \in W : u_1 \cdots u_r \neq 0\}.$$

Choose a vector $\mathbf{a}_0 = (a_{01}, ..., a_{0r})$ from W_0 with minimal height and define the linear form

$$l(\mathbf{X}) = \sum_{j=1}^{r} a_{0j}^{-1} X_j.$$

Let

$$T_1 := \{\mathbf{u} \in W : l(\mathbf{u}) = 0\}.$$

T_1 is a proper K-linear subspace of W since $l(\mathbf{a}_0) = r \neq 0$. Choose a basis $\{\mathbf{a}_1, ..., \mathbf{a}_n\}$ of W such that $\mathbf{a}_i \in W_0$ and $\mathbf{a}_i \notin T_1$ for $i = 1, ..., r$ and such that subject to these conditions, the product

$$H(\mathbf{a}_1)...H(\mathbf{a}_n) \text{ is minimal.}$$

Clearly, we may assume that $H(\mathbf{a}_1) \leq ... \leq H(\mathbf{a}_n)$. Thus, putting $M := H(\mathbf{a}_n)$, we have

$$\text{(6.1)} \qquad H(\mathbf{a}_0) \leq H(\mathbf{a}_1) \leq ... \leq H(\mathbf{a}_n) = M.$$

We must normalise $\mathbf{a}_1, ..., \mathbf{a}_n$. The properties of $\mathbf{a}_1, ..., \mathbf{a}_n$ just mentioned are not affected when for $i = 1, ..., n$, the vector $\mathbf{a}_i$ is replaced by $\lambda_i \mathbf{a}_i$ for any non-zero $\lambda_i \in K$. Take $\lambda_i := (a_{01} \cdots a_{0r})^{-1} l(\mathbf{a}_i)^{-1}$. Each λ_i belongs to K^* (as can be verified by checking that $\sigma(\lambda_i) = \lambda_i$ for $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$) and $l(\lambda_i \mathbf{a}_i) = (a_{01}...a_{0r})^{-1}$ for $i = 1, ..., r$. Therefore, we may assume that

$$(6.2) \qquad l(\mathbf{a}_i) = (a_{01} \cdots a_{0r})^{-1} \quad \text{for } i = 1, ..., r$$

and shall do so in the sequel. Because of (6.2) we have some control over the quantities $|\mathbf{a}_i|_v (v \in M_K, i = 1, ..., n)$. Namely, let $\mathbf{a}_i = (a_{i1}, ..., a_{ir})$; then by (6.2) we have $1 = (a_{01} \cdots a_{0r})(\sum_{j=1}^{r} a_{ij}/a_{0j})$ and together with (3.1) this implies that

$$(6.3) \qquad 1 \le r^{s(v)} |\mathbf{a}_0|_v^{r-1} |\mathbf{a}_i|_v \quad \text{for } v \in M_K, \ i = 1, ..., n.$$

We need (6.3) to estimate $|\mathbf{u}|_v$ and $|\mathbf{x}|_v$ in terms of each other where $\mathbf{x} = (x_1, ..., x_n) \in K^n$ is defined by $x_1 \mathbf{a}_1 + ... + x_n \mathbf{a}_n$.

Let T_2 be the K-linear subspace of W generated by $\mathbf{a}_1, ..., \mathbf{a}_{n-1}$. For every $\mathbf{u} \in W_0 \backslash (T_1 \cup T_2)$, $\{\mathbf{a}_1, ..., \mathbf{a}_{n-1}, \mathbf{u}\}$ is a basis of W. So since $\mathbf{a}_1, ..., \mathbf{a}_n$ were chosen such that $H(\mathbf{a}_1) \cdots H(\mathbf{a}_n)$ is minimal we have $H(\mathbf{a}_1)...H(\mathbf{a}_{n-1})H(\mathbf{u}) \ge H(\mathbf{a}_1)...H(\mathbf{a}_n)$. Hence

$$(6.4) \qquad H(\mathbf{u}) \ge M \quad \text{for } \mathbf{u} \in W_0 \backslash (T_1 \cup T_2).$$

Because of (6.4), the set of solutions of (4.5) can be divided into three sets:

$$A = \{\mathbf{u} \in W_0 \backslash (T_1 \cup T_2) : \ \mathbf{u} \text{ satisfies } (4.5), \ H(\mathbf{u}) \ge (2M)^{400r^3/\delta}\},$$
$$B = \{\mathbf{u} \in W_0 \backslash (T_1 \cup T_2) : \ \mathbf{u} \text{ satisfies } (4.5), \ M \le H(\mathbf{u}) < (2M)^{400r^3/\delta}\},$$
$$C = \{\mathbf{u} \in T_1 \cup T_2 \cup \{u_1 = 0\} \cup ... \cup \{u_r = 0\} : \ \mathbf{u} \text{ satisfies } (4.5)\}.$$

Clearly, C has linear scattering $\le r + 2$ in W. It remains to estimate the linear scatterings of A and B.

We first estimate the linear scattering of A in W. Take $\mathbf{u} \in A$. Then $\mathbf{x} = (x_1, ..., x_n) \in K^n$ defined by $\mathbf{u} = x_1 \mathbf{a}_1 + ... + x_n \mathbf{a}_n$ is non-zero. By (6.3) and $r \ge n$ we have

$$(6.5) \qquad \begin{aligned} |\mathbf{u}|_v &= n^{s(v)} |\mathbf{x}|_v \max(|\mathbf{a}_1|_v, ..., |\mathbf{a}_n|_v) \\ &\le n^{s(v)} r^{(n-1)s(v)} |\mathbf{a}_0|_v^{(r-1)(n-1)} |\mathbf{x}|_v |\mathbf{a}_1|_v ... |\mathbf{a}_n|_v \\ &\le r^{ns(v)} |\mathbf{a}_0|_v^{(r-1)(n-1)} |\mathbf{a}_1|_v ... |\mathbf{a}_n|_v . |\mathbf{x}|_v \quad \text{for } v \in M_K. \end{aligned}$$

By taking the product over $v \in M_K$ we obtain, using (6.1),

$$(6.6) \qquad \begin{aligned} H(\mathbf{u}) &\le r^n H(\mathbf{a}_0)^{(r-1)(n-1)} H(\mathbf{a}_1)...H(\mathbf{a}_n) H(\mathbf{x}) \\ &\le r^n M^{(r-1)(n-1)+n} H(\mathbf{x}) \\ &\le r^n M^{rn} H(\mathbf{x}). \end{aligned}$$

For $v \in M_K$, choose an independent set $I \subset \{1, ..., r\}$ of cardinality n such that $|(\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n)_I|_v = |\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n|_v$. Suppose for convenience that $I = \{1, ..., n\}$. Then

$$\Delta := (\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n)_I = \begin{vmatrix} a_{11} & \dots & a_{n1} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{nn} \end{vmatrix} . \tag{6.7}$$

We have $u_i = \sum_{j=1}^{v} a_{ji} x_j$ for $i = 1, ..., n$. Hence by Cramer's rule,

$$x_i = \sum_{j=1}^{n} (\Delta_{ij}/\Delta) u_j, \tag{6.8}$$

where Δ_{ij} is $\pm$ the determinant obtained by deleting the i-th row and j-th column from Δ. By (3.3) and (6.3) we have for $v \in M_K$,

$$|\Delta_{ij}|_v \le \{(n-1)!\}^{s(v)} \prod_{\substack{i=1 \\ i \ne j}}^{n} |\mathbf{a}_i|_v \le \{(n-1)!\}^{s(v)} r^{s(v)} |\mathbf{a}_0|_v^{r-1} |\mathbf{a}_1|_v ... |\mathbf{a}_n|_v.$$

By inserting this into (6.8) and using (6.7) we obtain

$$|\mathbf{x}|_v \le (n!r)^{s(v)} |\mathbf{a}_0|_v^{r-1} \frac{|\mathbf{a}_1|_v ... |\mathbf{a}_n|_v}{|\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n|_v} . |\mathbf{u}|_v \quad \text{for} \quad v \in M_K;$$

so, by taking the product over v, using $n \le r$ and (3.9),

$$H(\mathbf{x}) \le r^{n+1} \frac{M^{r+n-1}}{H(W)} H(\mathbf{u}) \le r^{n+1} M^{r+n-1} H(\mathbf{u}). \tag{6.9}$$

For $j = 1, ..., r$ define the linear form l_j by taking the j-th coordinates of $\mathbf{a}_1, ..., \mathbf{a}_n$, i.e.

$$l_j(\mathbf{X}) = a_{1j} X_1 + ... + a_{nj} X_n.$$

For $j = 1, ..., r$ define the field K_j by $\mathrm{Gal}(\overline{\mathbb{Q}}/K_j) = \{\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K) : \sigma(j) = j\}$ (where the action by $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$ on $\{1, ..., r\}$ is by means of Σ). The cosets of $\mathrm{Gal}(\overline{\mathbb{Q}}/K_j)$ in $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$ are $\{\sigma : \sigma(j) = k\}$ for certain $k \in \{1, ..., r\}$. Hence $[K_j : K] \le r$. For $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K_j)$, $j = 1, ..., r$, $i = 1, ..., n$ we have $\sigma(a_{ij}) = a_{i,\sigma(j)} = a_{ij}$ hence l_j has its coefficients in K_j. It follows that

$$[K(l_j) : K] \le r \quad \text{for} \quad j = 1, ..., r. \tag{6.10}$$

Let $L = K_1 ... K_r$. For $w \in M_L$ we have, by (6.3),

$$\begin{aligned} |l_j|_w = |a_{1j}, ..., a_{nj}|_w &\le \max(|\mathbf{a}_1|_w, ..., |\mathbf{a}_n|_w) \\ &\le r^{(n-1)s(w)} |\mathbf{a}_0|_w^{(r-1)(n-1)} |\mathbf{a}_1|_w ... |\mathbf{a}_n|_w. \end{aligned}$$

So by taking the product over $w \in M_L$, on using (6.1),

$$H(l_j) \le r^{n-1} M^{(r-1)(n-1)+n} \le r^{n-1} M^{rn} \quad \text{for } j = 1, ..., r.$$

Together with $H(\mathbf{u}) \ge (2M)^{400r^3/\delta}$ and (6.6) this implies that

$$\begin{aligned} H(\mathbf{x}) &\ge r^{-n} M^{-rn} H(\mathbf{u}) \ge r^{-n} M^{-rn} (2M)^{400r^3/\delta} \\ &\ge n^{1/2} r^{n-1} M^{rn} \\ &\ge n^{1/2} \cdot \max_{j=1,\dots,r} H(l_j). \end{aligned} \tag{6.11}$$

We shall show that

$$\prod_{v \in S} \prod_{i \in I_v} \frac{|l_i(\mathbf{x})|_v}{|\mathbf{x}|_v} \le n^{-1/2} \prod_{v \in S} |\det(l_i : i \in I_v)|_v \cdot H(\mathbf{x})^{-n-99\delta/100}. \tag{6.12}$$

First observe that by (3.4),

$$\begin{aligned} \prod_{v \in S} |\det(l_i : i \in I_v)|_v &= \prod_{v \in S} |(\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n)_{I_v}|_v \\ &= \Delta(\mathbf{I}, W) \Big(\prod_{v \notin S} |\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n|_v \Big)^{-1} \\ &\ge \Delta(\mathbf{I}, W) \Big(\prod_{v \notin S} |\mathbf{a}_1|_v ... |\mathbf{a}_n|_v \Big)^{-1}. \end{aligned} \tag{6.13}$$

Further, by (6.5) and (6.3),

$$\begin{aligned} \prod_{v \in S} |\mathbf{u}|_v &\le \prod_{v \in S} \{ r^{ns(v)} |\mathbf{a}_0|_v^{(r-1)(n-1)} |\mathbf{a}_1|_v ... |\mathbf{a}_n|_v \} \cdot \prod_{v \in S} |\mathbf{x}|_v \\ &= r \prod_{v \in S} \Big\{ \prod_{i=1}^{n} r^{s(v)} |\mathbf{a}_0|_v^{r-1} |\mathbf{a}_i|_v \Big\}^{\frac{n-1}{n}} \Big(\prod_{v \in S} |\mathbf{a}_1|_v ... |\mathbf{a}_n|_v \Big)^{1/n} \prod_{v \in S} |\mathbf{x}|_v \\ &\le r \prod_{v \in M_K} \Big\{ \prod_{i=1}^{n} r^{s(v)} |\mathbf{a}_0|_v^{r-1} |\mathbf{a}_i|_v \Big\}^{\frac{n-1}{n}} \Big(\prod_{v \in S} |\mathbf{a}_1|_v ... |\mathbf{a}_n|_v \Big)^{1/n} \prod_{v \in S} |\mathbf{x}|_v \\ &= r^n H(\mathbf{a}_0)^{(r-1)(n-1)} H(\mathbf{a}_1) ... H(\mathbf{a}_n) \cdot \Big(\prod_{v \notin S} |\mathbf{a}_1|_v ... |\mathbf{a}_n|_v \Big)^{-1/n} \cdot \prod_{v \in S} |\mathbf{x}|_v \\ &\le r^n M^{rn} \Big(\prod_{v \notin S} |\mathbf{a}_1|_v ... |\mathbf{a}_n|_v \Big)^{-1/n} \cdot \prod_{v \in S} |\mathbf{x}|_v. \end{aligned} \tag{6.14}$$

We have $l_i(\mathbf{x}) = u_i$ for $i = 1, ..., r$. Now (6.14), (6.13), (6.9) imply that

$$\prod_{v\in S}\prod_{i\in I_v}\frac{|l_i(\mathbf{x})|_v}{|\mathbf{x}|_v}\cdot n^{\frac{1}{2}(n+\delta)}\left(\prod_{v\in S}|\det(l_i : i\in I_v)|_v\right)^{-1}\cdot H(\mathbf{x})^{n+99\delta/100}$$

$$\leq \left(\prod_{v\in S}\prod_{i\in I_v}\frac{|u_i|_v}{|\mathbf{u}|_v}\right)\left\{r^n M^{rn}\right\}^n\left(\prod_{v\notin S}|\mathbf{a}_1|_v ... |\mathbf{a}_n|_v\right)^{-1}\cdot$$

$$\cdot\, n^{\frac{1}{2}(n+\delta)}\Delta(\mathbf{I}, W)^{-1}\left(\prod_{v\notin S}|\mathbf{a}_1|_v ... |\mathbf{a}_n|_v\right)\cdot$$

$$\cdot\left(r^{n+1}M^{r+n-1}\right)^{n+99\delta/100} H(\mathbf{u})^{n+99\delta/100}$$

$$= \left(\prod_{v\in S}\prod_{i\in I_v}\frac{|u_i|_v}{|\mathbf{u}|_v}\right)\Delta(\mathbf{I}, W)^{-1}H(\mathbf{u})^{n+\delta}\cdot$$

$$\cdot\, n^{\frac{1}{2}(n+\delta)}r^{n^2+(n+1)(n+99\delta/100)}M^{rn^2+(r+n-1)(n+99\delta/100)}H(\mathbf{u})^{-\delta/100}.$$

By (4.5) and by $H(\mathbf{u}) \geq (2M)^{400r^3/\delta}$, $0 < \delta < 1, r \geq n \geq 2$ this is ≤ 1. This proves (6.12).

By Lemma 3 and (6.10), the set of $\mathbf{x} \in K^n$ with (6.11), (6.12) has linear scattering in K^n at most

$$N_A := \left\{2^{60n^2}\left(\frac{100}{99}\right)^{7n}\delta^{-7n}\right\}^s \log 4r \cdot \log\log 4r\ .$$

Since $\mathbf{u} \mapsto \mathbf{x}$ is a bijective linear mapping from W to K^n, it follows that the linear scattering of A in W is at most N_A.

We now estimate the linear scattering of B in W. Put

$$\theta := r^{(ns-1)/(ns+1)}.$$

Let k be the smallest integer such that

$$\theta^k \geq 400r^3/\delta\ .$$

As $r \geq 2$ and $(ns+1)/(ns-1) \leq 3$, we have

$$k \leq 1 + \frac{ns+1}{ns-1}\left(3 + \frac{\log(400/\delta)}{\log r}\right) \leq 10 + 5\log(400/\delta). \tag{6.15}$$

For each $\mathbf{u} \in B$, there is a $j \in \{1, ..., k\}$ such that

$$\frac{1}{2}(2M)^{\theta^{j-1}} \leq H(\mathbf{u}) < (2M)^{\theta^j}. \tag{6.16}$$

From (4.5) and the lower bound in (6.16) it follows that

$$\prod_{v\in S}\prod_{i\in I_v}\frac{|u_i|_v}{|\mathbf{u}|_v}\leq \Delta(\mathbf{I},W)H(\mathbf{u})^{-n}.\Big\{\frac{1}{2}(2M)^{\theta^{j-1}}\Big\}^{-\delta}.$$

By applying Lemma 7 with $P=\{\frac{1}{2}(2M)^{\theta^{j-1}}\}^{\delta}, Q=1, B=(2M)^{\theta^j}$, on observing that $\log B/\log 2P\leq \theta/\delta$, we obtain that the set of $\mathbf{u}\in B$ with (6.16) has linear scattering in W at most

$$\big(270n^{10}r\theta/\delta\big)^{ns+1}=\Big(270n^{10}/\delta\Big)^{ns+1}r^{2ns}\ .$$

Together with (6.15) this implies that B has linear scattering in W at most

$$\begin{aligned}N_B&:=\Big\{10+5\log(400/\delta)\Big\}(270n^{10}/\delta)^{ns+1}r^{2ns}\\&\leq\Big(\frac{270n^{10}r}{\delta}\Big)^{2ns}.\end{aligned}$$

Recalling that C has linear scattering in W at most $N_C:=r+2$, we conclude that the set of $\mathbf{u}\in W$ with (4.5) has total linear scattering in W at most

$$\begin{aligned}&N_A+N_B+N_C\\&\quad=\Big(2^{60n^2}\Big(\frac{100}{99}\Big)^{7n}\delta^{-7n}\Big)^{s}\log 4r\cdot\log\log 4r+\Big(\frac{270n^{10}r}{\delta}\Big)^{2ns}+r+2\\&\quad<\Big(2^{61n^2}r^{2n}\delta^{-7n}\Big)^{s}.\end{aligned}$$

This completes the proof of Lemma 4. □

§7. Some linear algebra.

We will have to derive Theorem 4 from Lemma 4 and for this we need some linear algebra. Let K be an algebraic number field, Σ a $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$-action on $\{1,...,r\}$, and W an n-dimensional K-linear subspace of Λ_Σ. In what follows, by $\mathbf{u}$, $\mathbf{c}$, λ we denote vectors $(u_1,...,u_r)$, $(c_1,...,c_r)$, $(\lambda_1,...,\lambda_r)$. Note that we use both the coordinatewise product $\mathbf{cu}=(c_1u_1,...,c_ru_r)$ and the scalar product $(\mathbf{c},\mathbf{u})=c_1u_1+...+c_ru_r$. We define the orthogonal complement of W by the $\overline{\mathbb{Q}}$-vector space

$$W^\perp=\{\mathbf{c}\in\overline{\mathbb{Q}}^r:\ (\mathbf{c},\mathbf{u})=0\text{ for every }\mathbf{u}\in W\}.$$

We have to be a little bit careful since $W^\perp$ is a $\overline{\mathbb{Q}}$-vector space, whereas W is a K-vector space.

Lemma 8. $W = \{\mathbf{u} \in \Lambda_\Sigma : \ (\mathbf{c}, \mathbf{u}) = 0$ *for every* $\mathbf{c} \in W^\perp\}$.

Proof. Let $W' = \{\mathbf{u} \in \Lambda_\Sigma : \ (\mathbf{c}, \mathbf{u}) = 0$ for every $\mathbf{c} \in W^\perp\}$. Then W' is a K-linear subspace of Λ_Σ with $W' \supseteq W$ and $W'^\perp = W^\perp$. Let $\{\mathbf{a}_1, ..., \mathbf{a}_n\}$ be any basis of W. Then $W^\perp = \{\mathbf{c} \in \overline{\mathbb{Q}}^r : (\mathbf{a}_i, \mathbf{c}) = 0$ for $i = 1, ..., n\}$. By Lemma 2, $\{\mathbf{a}_1, ..., \mathbf{a}_n\}$ is $\overline{\mathbb{Q}}$-linearly independent. Hence $\dim_{\overline{\mathbb{Q}}} W^\perp = r - n$. Similarly, $\dim_{\overline{\mathbb{Q}}} W'^\perp = r - \dim_K W'$. Hence $\dim_K W' = n$. It follows that $W = W'$. □

For $\mathbf{c} = (c_1, ..., c_r) \in \overline{\mathbb{Q}}^r$, $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$ define the vector

$$\mathbf{c}^{(\sigma)} := (\sigma(c_{\sigma^{-1}(1)}), ..., \sigma(c_{\sigma^{-1}(r)})). \tag{7.1}$$

Obviously, $\mathbf{c} \in \Lambda_\Sigma$ if and only if $\mathbf{c}^{(\sigma)} = \mathbf{c}$ for every $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$. We use this to prove:

Lemma 9. *Let* $\mathcal{P} = \{P_1, ..., P_t\}$ *be a symmetric partition of* $\{1, ..., r\}$. *Then* $\dim_K \Lambda_{\mathcal{P}} = t$.

Proof. Define the $\overline{\mathbb{Q}}$-vector space $Y = \{\mathbf{u} \in \overline{\mathbb{Q}}^r : \ u_i = u_j$ for each pair $i \overset{\mathcal{P}}{\sim} j\}$ (recall that $i \overset{\mathcal{P}}{\sim} j$ if and only if i, j belong to the same set of $\mathcal{P}$). Then $\Lambda_{\mathcal{P}} = Y \cap \Lambda_\Sigma$. We claim that every $\mathbf{u} \in Y$ is a linear combination of vectors from $\Lambda_{\mathcal{P}}$. Namely, take $\mathbf{u} \in Y$. Since $\mathcal{P}$ is a symmetric partition we have also that $\mathbf{u}^{(\sigma)} \in Y$ for every $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$. Let G be the group of $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$ such that $\sigma(u_i) = u_i$ and $\sigma(i) = i$ for $i = 1, ..., r$. Then G has finite index in $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$; let $\sigma_1, ..., \sigma_g$ be a full system of left coset-representatives. Let L be the extension of K with $\mathrm{Gal}(\overline{\mathbb{Q}}/L) = G$ and choose a K-basis $\omega_1, ..., \omega_g$ of L. Then the vectors $\mathbf{b}_i := \sum_{j=1}^g \sigma_j(\omega_i)\mathbf{u}^{(\sigma_j)}$ $(i = 1, ..., g)$ belong to Y; further, by the definition of G and L, $\mathbf{b}_i$ is independent of the choice of the coset representatives. Since left multiplication with $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$ permutes the left cosets of G in $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$, we have $\mathbf{b}_i^{(\sigma)} = \sum \sigma\sigma_j(\omega_i)\mathbf{u}^{(\sigma\sigma_j)} = \mathbf{b}_i$ for $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$, i.e. $\mathbf{b}_i \in Y \cap \Lambda_\Sigma = \Lambda_{\mathcal{P}}$ for $i = 1, ..., g$. Now since the matrix $(\sigma_j(\omega_i))$ is invertible, we have that $\mathbf{u}$ is a linear combination of $\mathbf{b}_1, ..., \mathbf{b}_g$. This proves our claim.

Clearly, $\dim_{\overline{\mathbb{Q}}} Y = t$. From what we proved above, it follows that Y is generated by a K-basis of $\Lambda_{\mathcal{P}}$. By Lemma 2, this K-basis is $\overline{\mathbb{Q}}$-linearly independent, hence consists of t vectors. This implies Lemma 9. □

We can now give an alternative description for the space $W_{\mathcal{P}}$:

Lemma 10. *Let $\mathcal{P} = \{P_1, ..., P_t\}$ be a symmetric partition of $\{1, ..., r\}$. Then*

$$W_{\mathcal{P}} = \{\mathbf{u} \in W : \sum_{i \in P_j} c_i u_i = 0 \text{ for } j = 1, ..., t \text{ and for every } \mathbf{c} \in W^{\perp}\}.$$

Proof. For $\lambda = (\lambda_1, ..., \lambda_r) \in \Lambda_{\mathcal{P}}$ define $\lambda' = (\lambda'_1, ..., \lambda'_t)$ where $\lambda'_i = \lambda_j$ for each $j \in P_i$. Further, for $\mathbf{u} \in W$, $\mathbf{c} \in W^{\perp}$, $j = 1, ..., t$ we put $(\mathbf{c}, \mathbf{u})_j := \sum_{i \in P_j} c_i u_i$. Thus, for $\mathbf{u} \in W$, $\mathbf{c} \in W^{\perp}$, $\lambda \in \Lambda_{\mathcal{P}}$ we have

$$(\mathbf{c}, \lambda\mathbf{u}) = \sum_{j=1}^{t} \lambda'_j (\mathbf{c}, \mathbf{u})_j. \tag{7.2}$$

Recalling that $W_{\mathcal{P}} = \{\mathbf{u} \in W : \mathbf{u}\Lambda_{\mathcal{P}} \subseteqq W\}$, we infer by Lemma 8 that $\mathbf{u} \in W_{\mathcal{P}}$ if and only if $(\lambda\mathbf{u}, \mathbf{c}) = 0$ for every $\lambda \in \Lambda_{\mathcal{P}}$, $\mathbf{c} \in W^{\perp}$. Further, by Lemma 9 and Lemma 2, among the vectors $\lambda \in \Lambda_{\mathcal{P}}$, hence also among the vectors λ', there are t $\overline{\mathbb{Q}}$-linearly independent ones. Together with (7.2) this implies that $(\lambda\mathbf{u}, \mathbf{c}) = 0$ for every $\lambda \in \Lambda_{\mathcal{P}}$ if and only if $(\mathbf{c}, \mathbf{u})_j = 0$ for $j = 1, ..., t$. This proves Lemma 10. $\square$

Remark 4. (cf. §1 for notation.) We show that solutions of (1.5) satisfying (1.15) are (F, S)-non-degenerate. Let $l_1, ..., l_r$ be the linear factors of F, define the linear mapping $\varphi(\mathbf{x}) = (l_1(\mathbf{x}), ..., l_r(\mathbf{x}))$ and put $W = \varphi(K^n)$. Suppose that $\mathbf{x} \in \mathcal{O}_S^n$ satisfies (1.15). Put $\mathbf{u} = \varphi(\mathbf{x})$. Then for each proper, non-empty subset I of $\{1, ..., r\}$ there is a $\mathbf{c} \in W^{\perp}$ such that $\sum_{i \in I} c_i u_i \neq 0$. Lemma 10 implies that $\mathbf{u} \notin W_{\mathcal{P}}$ for each symmetric partition $\mathcal{P} \neq \{\{1, ..., r\}\}$; hence $\mathbf{u}$ is S-non-degenerate. Now by definition, $\mathbf{x}$ is (F, S)-non-degenerate.

We define the *support* of $\mathbf{c} = (c_1, ..., c_r) \in W^{\perp}$ by

$$\mathrm{supp}(\mathbf{c}) := \{i \in \{1, ..., r\} : c_i \neq 0\}.$$

We say that $\mathbf{c}$ has minimal support in $W^{\perp}$ if $\mathbf{c} \in W^{\perp}$ is non-zero and if there is no non-zero $\mathbf{c}' \in W^{\perp}$ with $\mathrm{supp}(\mathbf{c}') \subsetneqq \mathrm{supp}(\mathbf{c})$.
We associate a *hypergraph* $\mathcal{H}(W)$ to W, by taking as set of vertices $\{1, ..., r\}$ and as edges those subsets I of $\{1, ..., r\}$ for which there exists a vector $\mathbf{c}$ of minimal support in $W^{\perp}$ with $\mathrm{supp}(\mathbf{c}) = I$.

We say that two vertices $i, j \in \{1, ..., r\}$ are *connected* if there are $i_1, i_2, ..., i_t \in \{1, ..., r\}$, such that each of the pairs $\{i, i_1\}, \{i_1, i_2\}, ..., \{i_t, j\}$ is a subset of an edge of $\mathcal{H}(W)$. Connectedness is clearly an equivalence relation on $\{1, ..., r\}$ and the equivalence classes are called the *connected components* of $\mathcal{H}(W)$.
The connected components of $\mathcal{H}(W)$ form a symmetric partition. Namely, let $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$. For every $\mathbf{c} \in W^{\perp}$, $\mathbf{u} \in W$ we have $(\mathbf{u}, \mathbf{c}^{(\sigma)}) = \sigma((\mathbf{u}, \mathbf{c})) = 0$, where $\mathbf{c}^{(\sigma)}$ is defined by (7.1); hence $\mathbf{c}^{(\sigma)} \in W^{\perp}$. It follows that if I is an edge of

$\mathcal{H}(W)$ then so is $\sigma(I)$. Consequently, σ maps each connected component of $\mathcal{H}(W)$ to a connected component.

We prove some properties of $\mathcal{H}(W)$. Let $\mathcal{P} = \{P_1, ..., P_t\}$ be the symmetric partition of connected components of $\mathcal{H}(W)$. For $j = 1, ..., t$, define the linear subspace of $W^\perp$,

$$Y_j := \{\mathbf{c} \in W^\perp : \operatorname{supp}(\mathbf{c}) \subseteq P_j\}. \tag{7.3}$$

Lemma 11. *(i).* $W^\perp = Y_1 \oplus ... \oplus Y_t$.
(ii). For the height of W *we have* $H(W) = H(Y_1) \cdots H(Y_t)$.
(iii). $W = W_{\mathcal{P}}$.

Proof. (i). Every non-zero $\mathbf{c} \in W^\perp$ is a sum of vectors with minimal support in $W^\perp$. Namely, if $\mathbf{c} \in W^\perp$ has minimal support then we are done; otherwise, choose a vector $\mathbf{c}'$ with minimal support in $W^\perp$ such that $\operatorname{supp}(\mathbf{c}') \subsetneqq \operatorname{supp}(\mathbf{c})$ and arrange, by taking a suitable scalar multiple, that a non-zero coordinate of $\mathbf{c}'$ is equal to the corresponding coordinate of $\mathbf{c}$; then $\operatorname{supp}(\mathbf{c} - \mathbf{c}') \subsetneqq \operatorname{supp}(\mathbf{c})$ and we may repeat the argument with $\mathbf{c} - \mathbf{c}'$.
Let $\mathbf{c} \in W^\perp$. We have $\mathbf{c} = \sum_{j=1}^{t} \sum_{i \in T_j} \mathbf{d}_{ij}$, where $\mathbf{d}_{ij}$ has minimal support in $W^\perp, \operatorname{supp}(\mathbf{d}_{ij}) \subseteq P_j$, and T_j is some set of indices. Put $\mathbf{c}_j := \sum_{i \in T_j} \mathbf{d}_{ij}$. Then $\mathbf{c} = \mathbf{c}_1 + ... + \mathbf{c}_t$, $\mathbf{c}_j \in W^\perp$, $\operatorname{supp}(\mathbf{c}_j) \subseteq P_j$, i.e. $\mathbf{c}_j \in Y_j$. $\mathbf{c}_1, ..., \mathbf{c}_t$ are uniquely determined by $\mathbf{c}$, since for each $i \in P_j$, the i-th coordinate of $\mathbf{c}_j$ is equal to the i-th coordinate of $\mathbf{c}$. This proves (i).

(ii). By (3.5), (3.6) we have $H(W) = H(W^\perp) = H(Y_1) \cdots H(Y_t)$. We mention that (3.5) has been stated for subspaces of $\overline{\mathbb{Q}}^r$ and not for subspaces of Λ_Σ. However, letting $\overline{W} = \overline{\mathbb{Q}}W$ be the $\overline{\mathbb{Q}}$-vector space generated by W, we know by Lemma 2 that any basis of W is a basis of $\overline{W}$, and by definition (3.8) that $H(W) = H(\overline{W})$; further, it is obvious that $W^\perp = \overline{W}^\perp$.

(iii). By (i) we can express $\mathbf{c} \in W^\perp$ as $\mathbf{c}_1 + ... + \mathbf{c}_t$ with $\mathbf{c}_i \in W^\perp$, $\operatorname{supp}(\mathbf{c}_i) \subseteq P_i$ for $i = 1, ..., t$. Hence for $\mathbf{u} \in W, \mathbf{c} \in W^\perp$ we have $(\mathbf{c}_i, \mathbf{u}) = 0$ for $i = 1, ..., t$; together with Lemma 10 this implies that $\mathbf{u} \in W_{\mathcal{P}}$. This shows (iii). □

We recall that a non-empty subset I of $\{1, ..., r\}$ is called independent if (4.4) holds; otherwise I is called dependent. Below we have collected some simple facts for later use. $\mathcal{P} = \{P_1, ..., P_t\}$ denotes again the symmetric partition of connected components of $\mathcal{H}(W)$.

Lemma 12. *(i). A non-empty subset* I *of* $\{1, ..., r\}$ *is dependent if and only if there is an edge* J *of* $\mathcal{H}_W$ *with* $J \subseteq I$.

(ii). Each maximal independent subset of $\{1, ..., r\}$ has cardinality n.
(iii). Each maximal independent subset of P_j ($j = 1, ..., t$) has the same cardinality, n_j, say.
(iv). Let I_j be a maximal independent subset of P_j for $j = 1, ..., t$. Then $I_1 \cup ... \cup I_t$ is a maximal independent subset of $\{1, ..., r\}$. Hence $n_1 + ... + n_t = n$.
(v). Let I be an edge of $\mathcal{H}(W)$, contained in the connected component P_j. Then there is a subset H of P_j such that for each $i \in I$, $(I\backslash\{i\}) \cup H$ is an independent set of cardinality n_j.

Proof. We first make some remarks. Let $\overline{W} = \overline{\mathbb{Q}}W$ be the $\overline{\mathbb{Q}}$-vector space generated by W. Then by Lemma 2, $\overline{W}$ has dimension n. Denote by f_i the $\overline{\mathbb{Q}}$-linear function $\mathbf{u} \mapsto u_i$ on $\overline{W}$. Thus, $f_1, ..., f_r$ generate the vector space of $\overline{\mathbb{Q}}$-linear functions on $\overline{W}$. Hence $\text{rank}_{\overline{\mathbb{Q}}}\{f_1, ..., f_r\} = \dim_{\overline{\mathbb{Q}}} \overline{W} = n$. Now (4.4) and the fact that every vector in $\overline{W}$ is a $\overline{\mathbb{Q}}$-linear combination of vectors in W imply that I is (in)dependent if and only if $\{f_i : \ i \in I\}$ is $\overline{\mathbb{Q}}$-linearly (in)dependent. Further, $\mathbf{c} = (c_1, ..., c_r) \in W^{\perp}$ if and only if $\sum_{i=1}^{n} c_i f_i = 0$.
(i). By the remarks just made we have that I is dependent$\iff$ there exists a non-zero $\mathbf{c} = (c_1, ..., c_r) \in \overline{\mathbb{Q}}^r$ with $\text{supp}(\mathbf{c}) \subseteq I$, $\sum_{i=1}^{r} c_i f_i = 0 \iff$ there is a non-zero $\mathbf{c} \in W^{\perp}$ with supp $(\mathbf{c}) \subseteq I \iff I$ contains an edge of $\mathcal{H}(W)$.
(ii). The remarks made above imply that a maximal independent subset of $\{1, ..., r\}$ has cardinality $\text{rank}_{\overline{\mathbb{Q}}}\{f_1, ..., f_r\} = \dim_{\overline{\mathbb{Q}}} W = n$.
(iii). This holds with $n_j := \text{rank}_{\overline{\mathbb{Q}}}\{f_i : \ i \in P_j\}$.
(iv). We apply (i). Put $I = I_1 \cup ... \cup I_t$. If I is dependent then there is an edge J of $\mathcal{H}(W)$ with $J \subseteq I$. But then, $J \subseteq I \cap P_j = I_j$ for some j which contradicts that I_j is independent. Further, for each $i \in \{1, ..., r\}\backslash I$ we have that $I \cup \{i\}$ is dependent; namely if $i \in P_j$, say, then $I_j \cup \{i\}$ is already dependent by the maximality of I_j.
(v). Since I is an edge we have by (i) and the remarks made above that for each $i \in I$, the set $B_i := \{f_k \ : k \in I\backslash\{i\}\}$ is linearly independent, and that f_i is linearly dependent on B_i. Choose $i_0 \in I$. By (iii) there is a subset H of P_j such that $\{f_k : \ k \in (I\backslash\{i_0\}) \cup H\}$ is linearly independent and of cardinality n_j. It follows easily that also for each $i \in I$, $\{f_k : \ k \in (I\backslash\{i\}) \cup H\}$ is linearly independent and of cardinality n_j. This implies (v). □

§8. Reduction of Theorem 4 to the Subspace theorem.

We recall that by $|\mathcal{S}|$ we denote the cardinality of a set $\mathcal{S}$. As before, K is an algebraic number field, S is a finite set of places on K containing all infinite places, Σ is a $\text{Gal}(\overline{\mathbb{Q}}/K)$-action on $\{1, ..., r\}$ and W is an n-dimensional K-linear subspace of Λ_{Σ}, where $r \geq n \geq 2$. By $\mathbf{u}$ we denote vectors $(u_1, ..., u_r)$, and for $I \subseteq \{1, ..., r\}$ we define the partial vector

$$\mathbf{u}_I = (u_i : \ i \in I).$$

Recall that Theorem 4 deals with S-non-degenerate elements $\mathbf{u} \in W$ satisfying

$$u_1 \cdots u_r \neq 0, \quad u_i/u_j \in \overline{\mathcal{O}}_S^* \text{ for } i,j = 1,...,r. \tag{2.3}$$

In this section we show that each such $\mathbf{u}$ satisfies a Diophantine inequality of the type (4.5) considered in Lemma 4.

Let $\mathcal{P} = \{P_1, ..., P_t\}$ be the symmetric partition of connected components of the hypergraph $\mathcal{H}(W)$ associated to W. We recall that $\mathcal{O}_{\mathcal{P},S}^* = \Lambda_{\mathcal{P}} \cap (\overline{\mathcal{O}}_S^*)^r$ and that $\delta(a) = (a, ..., a)$ (r times) for $a \in K$.

Lemma 13. *Suppose that $\mathcal{O}_{\mathcal{P},S}^*/\delta(\mathcal{O}_S^*)$ is finite. Then for every $\mathbf{u} \in W$ with (2.3) we have*

$$H(\mathbf{u}) \leq H(\mathbf{u}_{P_1})...H(\mathbf{u}_{P_t}).$$

Proof. Take $\mathbf{u} = (u_1, ..., u_r) \in W$ with (2.3). Define

$$v_i := \frac{u_i^r}{u_1 ... u_r} \quad \text{for } i = 1, ..., r, \tag{8.1}$$

and put

$$\lambda_k := \prod_{j \in P_i} v_j \quad \text{for } k \in P_i, \ i = 1, ..., t. \tag{8.2}$$

From (2.3) it follows that for $k \in P_i$, $i = 1, ..., t$ we have

$$\lambda_k = \prod_{j \in P_i} \prod_{l=1}^{r} \frac{u_j}{u_l} \in \overline{\mathcal{O}}_S^*.$$

Obviously, $\lambda_k = \lambda_l$ for each pair $k \overset{\mathcal{P}}{\sim} l$. Further, for $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K), k \in P_i, i = 1, ..., t$ we have

$$\sigma(\lambda_k) = \prod_{j \in P_i} \sigma(v_j) = \prod_{j \in P_i} v_{\sigma(j)} = \prod_{j \in \sigma(P_i)} v_j = \lambda_{\sigma(k)}.$$

Hence $\lambda \in \Lambda_{\mathcal{P}} \cap (\overline{\mathcal{O}}_S^*)^r = \mathcal{O}_{\mathcal{P},S}^*$. By the finiteness of $\mathcal{O}_{\mathcal{P},S}^*/\delta(\mathcal{O}_S^*)$, there is a positive integer m such that $\lambda^m \in \delta(\mathcal{O}_S^*)$, i.e.

$$\lambda_k^m = \rho \in \mathcal{O}_S^* \quad \text{for } k = 1, ..., r.$$

Together with (8.1), (8.2) this implies that

$$1 = (v_1 ... v_r)^m = \prod_{i=1}^{t} (\prod_{j \in P_i} v_j)^m = \rho^t,$$

hence ρ is a root of unity. It follows that $\prod_{j\in P_i} v_j$ is a root of unity for $i = 1, ..., t$. Let $L = K(u_1, ..., u_r)$, $\mathbf{v} = (v_1, ..., v_r)$, $\mathbf{v}_I = (v_i : i \in I)$. Then for $w \in M_L, i = 1, ..., t$ we have

$$|\mathbf{v}_{P_i}|_w = \max(|v_j|_w : j \in P_i) \geq \left(\prod_{j\in P_i} |v_j|_w\right)^{1/|P_i|} = 1,$$

hence $|\mathbf{v}|_w \leq |\mathbf{v}_{P_1}|_w ... |\mathbf{v}_{P_t}|_w$. By taking the product over $w \in M_L$ we get

$$H(\mathbf{v}) \leq H(\mathbf{v}_{P_1}) ... H(\mathbf{v}_{P_t}). \tag{8.3}$$

Now (8.1) and the Product formula imply that $H(\mathbf{v}_I) = H(\mathbf{u}_I)^r$ for each subset I of $\{1, ..., r\}$. Together with (8.3) this implies Lemma 13. $\square$

For $i = 1, ..., t$, let n_i be the cardinality of a maximal independent subset of P_i (cf. Lemma 12) and let $Y_i = \{\mathbf{u} \in W^\perp : \ \mathrm{supp}(\mathbf{u}) \subseteq P_i\}$ (cf. (7.3)). We have:

Lemma 14. *Let $\mathbf{u} \in W$ with $u_1 \cdots u_r \neq 0$. For each $i \in \{1, ..., t\}$ there is an edge J_i of $\mathcal{H}(W)$ with $J_i \subseteq P_i$ and*

$$H(\mathbf{u}_{J_i}) \geq 2^{-n_i/2}\{H(\mathbf{u}_{P_i})/H(Y_i)\}^{1/(n-1)}. \tag{8.4}$$

Proof. Fix $i \in \{1, ..., t\}$ and put $r_i := |P_i|$. If $r_i = 1$ then $\mathbf{u}_{J_i}$, $\mathbf{u}_{P_i}$ and the basis of Y_i are vectors with only one non-zero coordinate and by the Product formula these have height equal to 1; so in that case (8.4) is trivially true. Suppose that $r_i \geq 2$. Write P for P_i, Y for Y_i. Let H be a maximal independent subset of P. By Lemma 12 (iii), H has cardinality n_i. We assume that $P = \{1, ..., r_i\}$, $H = \{1, ..., n_i\}$ which is no loss of generality. By Lemma 12 (i), for each $k \in P\backslash H = \{n_i + 1, ..., r_i\}$, there is an edge J_k of $\mathcal{H}(W)$ with $J_k \subseteq H \cup \{k\}$ and $k \in J_k$. For each $k \in P\backslash H$, there is a vector $\mathbf{c}_k \in Y$ with $\mathrm{supp}(\mathbf{c}_k) = J_k$. We assume that the k-th coordinate of $\mathbf{c}_k$ is 1, which is no restriction. That is, we have

$$\begin{array}{rcllr} \mathbf{c}_{n_i+1} & = & (c_{n_i+1,1}, ..., & c_{n_i+1,n_i}, & 1, 0, ..., 0), \\ \mathbf{c}_{n_i+2} & = & (c_{n_i+2,1}, ..., & c_{n_i+2,n_i}, & 0, 1, ..., 0), \\ & & \cdots & & \\ \mathbf{c}_{r_i} & = & (c_{r_i,1}, ..., & c_{r_i,n_i}, & 0, 0, ..., 1), \end{array} \tag{8.5}$$

where we have omitted the zero coordinates for the indices in the connected components $\neq P$. Here, $c_{ki} \neq 0$ only if $i \in J_k$. The set $\{\mathbf{c}_k : k \in P\backslash H\}$ is a basis of Y. Namely, let $\mathbf{c} = (\alpha_1, ..., \alpha_r) \in Y$. Then $\mathrm{supp}(\mathbf{c}) \subseteq P$. Further, $\mathbf{c} - \sum_{k\in P\backslash H} \alpha_k \mathbf{c}_k$ has its support in H. But H does not contain an edge of $\mathcal{H}(W)$, hence $\mathbf{c} = \sum_{k\in P\backslash H} \alpha_k \mathbf{c}_k$.

For every subset D of H with $D \neq \emptyset$, $H\backslash D \neq \emptyset$, there is a $k \in P\backslash H$ with $J_k \cap D \neq \emptyset, J_k \cap (H\backslash D) \neq \emptyset$. Namely suppose this is not true for some D. Then for each $k \in P\backslash H$ we have either $J_k \subseteq D \cup \{k\}$ or $J_k \subseteq (H\backslash D) \cup \{k\}$. Hence every vector $\mathbf{c}_k$ $(k \in P\backslash H)$ defined above has its support either in $D' = D \cup \{k \in P\backslash H : J_k \subseteq D \cup \{k\}\}$ or in $P\backslash D'$. But then, as $\{\mathbf{c}_k : k \in P\backslash H\}$ is a basis of Y, every $\mathbf{c} \in Y$ can be expressed as $\mathbf{c}' + \mathbf{c}''$ with $\mathbf{c}', \mathbf{c}'' \in Y$ and $\mathrm{supp}(\mathbf{c}') \subseteq D', \mathrm{supp}(\mathbf{c}'') \subseteq P\backslash D'$. Hence $\mathcal{H}(W)$ has no edge containing elements of both D' and $P\backslash D'$ and this contradicts that P is a connected component of $\mathcal{H}(W)$.

Hence there is a $k_1 \in P\backslash H$ with $|J_{k_1} \cap H| \geq 2$. If $J_{k_1} \not\supseteq H$, then there is a $k_2 \in P\backslash H$ with $J_{k_2} \cap J_{k_1} \cap H \neq \emptyset$, $(J_{k_1} \cup J_{k_2}) \cap H \supsetneqq J_{k_1} \cup H$. Continuing in this way, we get a sequence of elements $k_1, ..., k_a$ of $P\backslash H$, with $a \leq n_i - 1$, such that

$$\text{(8.6)} \qquad \begin{cases} |J_{k_1} \cap H| \geq 2, \\ J_{k_j} \cap (J_{k_1} \cup ... \cup J_{k_{j-1}}) \cap H \neq \emptyset \quad \text{for } j = 2, ..., a, \\ (J_{k_1} \cup ... \cup J_{k_j}) \cap H \supsetneqq (J_{k_1} \cup ... \cup J_{k_{j-1}}) \cap H \quad \text{for } j = 2, ..., a, \\ J_{k_1} \cup ... \cup J_{k_a} \supseteq H. \end{cases}$$

Now let $\mathbf{u} \in W$ with $u_1 ... u_r \neq 0$. For $A, B \subseteq \{1, ..., r\}$ with $A \cap B \neq \emptyset$ we have

$$\text{(8.7)} \qquad H(\mathbf{u}_{A \cup B}) \leq H(\mathbf{u}_A) H(\mathbf{u}_B).$$

Namely, let $\mathbf{v} = \lambda^{-1}\mathbf{u}$ with $\lambda = u_i$ for some $i \in A \cap B$ and let $L := K(u_1, ..., u_r)$. Both $\mathbf{v}_A, \mathbf{v}_B$ have a coordinate 1. Hence $|\mathbf{v}_{A \cup B}|_w \leq |\mathbf{v}_A|_w |\mathbf{v}_B|_w$ for $w \in M_L$. By taking the product over $w \in M_L$ we get $H(\mathbf{v}_{A \cup B}) \leq H(\mathbf{v}_A) H(\mathbf{v}_B)$ which is the same as (8.7).
By (8.6) and (8.7) we have

$$H(\mathbf{u}_H) \leq H(\mathbf{u}_{J_{k_1}}) ... H(\mathbf{u}_{J_{k_a}}) \leq \left\{ \max_i H(\mathbf{u}_{J_{k_i}}) \right\}^{n_i - 1}.$$

Hence $\mathcal{H}(W)$ has an edge $J \subseteq P$ with

$$\text{(8.8)} \qquad H(\mathbf{u}_J) \geq H(\mathbf{u}_H)^{1/(n_i - 1)}$$

where we agree that the right-hand side is 1 if $n_i = 1$ (in this case, $|H| = 1$ and hence $H(\mathbf{u}_H) = 1$). We have to relate $H(\mathbf{u}_H)$ to $H(\mathbf{u}_P)$.

It is obvious that the exterior product of the basis of Y given in (8.5), $\mathbf{c}_{n_i+1} \wedge ... \wedge \mathbf{c}_{r_i}$ has among its coordinates, up to sign, all numbers 1, c_{kj} $(k \in P\backslash H,\ j \in H)$. Let L be a finite extension of K containing $u_1, ..., u_r$ and all c_{kj}. Then

$$\text{(8.9)} \qquad 1 \leq |\mathbf{c}_k|_w \leq |\mathbf{c}_{n_i+1} \wedge ... \wedge \mathbf{c}_{r_i}|_w \quad \text{for } k \in P\backslash H,\ w \in M_L .$$

We have that the scalar product $(\mathbf{u}, \mathbf{c}_k) = 0$, whence

$$u_k = -\sum_{j \in H} c_{kj} u_j \quad \text{for } k \in P \backslash H.$$

Together with (3.2) and (8.9) this implies that

$$\begin{aligned} |\mathbf{u}_P|_w = \max_{k \in P} |u_k|_w &\leq n_i^{s(w)} \max_{k \in P} |\mathbf{c}_k|_w \cdot |\mathbf{u}_H|_w \\ &\leq n_i^{s(w)} |\mathbf{c}_{n_i+1} \wedge ... \wedge \mathbf{c}_{r_i}|_w |\mathbf{u}_H|_w \quad \text{for } w \in M_L \ . \end{aligned}$$

By taking the product over $w \in M_L$ we obtain

$$H(\mathbf{u}_P) \leq n_i H(\mathbf{c}_{n_i+1} \wedge ... \wedge \mathbf{c}_{r_i}) H(\mathbf{u}_H) = n_i H(Y) H(\mathbf{u}_H).$$

If $H(\mathbf{u}_P) > H(Y)$ then this implies, together with (8.8),

$$\begin{aligned} H(\mathbf{u}_J) &\geq \max\bigl\{1,\ n_i^{-1} H(Y)^{-1} H(\mathbf{u}_P)\bigr\}^{1/(n_i-1)} \\ &\geq n_i^{-1/(n_i-1)} \bigl\{H(\mathbf{u}_P)/H(Y)\bigr\}^{1/(n_i-1)} \geq 2^{-n_i/2} \Bigl\{ H(\mathbf{u}_P)/H(Y) \Bigr\}^{1/(n-1)}. \end{aligned}$$

(8.4) is clearly true also if $H(\mathbf{u}_P) \leq H(Y)$. This proves Lemma 14. □

The next lemma is the main result of this section:

Lemma 15. *Suppose that $\mathcal{O}^*_{\mathcal{P},S}/\delta(\mathcal{O}^*_S)$ is finite. Then for every $\mathbf{u} \in W$ satisfying (2.3) there is a collection $\mathbf{I} = (I_v : v \in S)$ of independent subsets of $\{1, ..., r\}$ of cardinality n such that*

$$(8.10) \qquad \prod_{v \in S} \prod_{i \in I_v} \frac{|u_i|_v}{|\mathbf{u}|_v} \leq 2^{n/2} H(W)^{\binom{r}{n}} \cdot \Delta(\mathbf{I}, W) H(\mathbf{u})^{-n-\{1/(n-1)\}}.$$

Proof. Take $i \in \{1, ..., t\}$. Let J_i be the edge contained in P_i from Lemma 14. By Lemma 12 (v), there is a subset H' of $P_i \backslash J_i$ such that for every $j \in J_i$, $H' \cup (J_i \backslash \{j\})$ is a maximal independent subset of P_i of cardinality n_i.

Take $\mathbf{u} \in W$ with (2.3). Put $L := K(u_1, ..., u_r)$ and let T be the set of places in M_L lying above those in S. Since $u_i/u_j \in \overline{\mathcal{O}}^*_S$ we have $|u_i/u_j|_w = 1$ for $w \in M_L \backslash T,\ i, j = 1, ..., r$, whence

$$|u_1|_w = ... = |u_r|_w \quad \text{for } w \in M_L \backslash T.$$

For $w \in T$, choose $i_w \in J_i$ such that $|u_{i_w}|_w = |\mathbf{u}_{J_i}|_w$ and put $I_{iw} := (J_i \backslash \{i_w\}) \cup H'$. Note that $|u_i|_w = |\mathbf{u}_{J_i}|_w = |\mathbf{u}|_w$ for $w \in M_L \backslash T$, $i = 1, ..., r$. Together with the Product formula this implies

$$\prod_{w \in T} \prod_{j \in I_{iw}} \frac{|u_j|_w}{|\mathbf{u}|_w} = \prod_{w \in T} \left(\{ \prod_{j \in J_i \cup H'} |u_j|_w \} |\mathbf{u}|_w^{-n_i} |\mathbf{u}_{J_i}|_w^{-1} \right)$$
$$= \prod_{j \in J_i \cup H'} \left(\prod_{w \in M_L} |u_j|_w \right) \prod_{w \in M_L} \left(|\mathbf{u}|_w^{-n_i} |\mathbf{u}_{J_i}|_w^{-1} \right)$$
$$= H(\mathbf{u})^{-n_i} H(\mathbf{u}_{J_i})^{-1}.$$

By inserting (8.4) we get for $i = 1, ..., t$,

$$\prod_{w \in T} \prod_{j \in I_{iw}} \frac{|u_j|_w}{|\mathbf{u}|_w} \leq 2^{n_i/2} H(Y_i)^{1/(n-1)} H(\mathbf{u})^{-n_i} H(\mathbf{u}_{P_i})^{-1/(n-1)}.$$

Put $I_w = \cup_{i=1}^t I_{iw}$ for $w \in T$. By Lemma 12 (ii),(iv), I_w is an independent subset of $\{1, ..., r\}$ of cardinality n. By taking the product over $i = 1, ..., t$ and using Lemma 12 (iii), Lemma 11 (ii) and Lemma 13 we get

$$(8.11) \quad \prod_{w \in T} \prod_{j \in I_w} \frac{|u_j|_w}{|\mathbf{u}|_w} \leq 2^{(n_1 + ... + n_t)/2} \{ H(Y_1) \cdots H(Y_t) \}^{1/(n-1)} \cdot$$
$$\cdot H(\mathbf{u})^{-(n_1 + ... + n_t)} \{ H(\mathbf{u}_{P_1}) \cdots H(\mathbf{u}_{P_t}) \}^{-1/(n-1)}$$
$$\leq 2^{n/2} H(W)^{1/(n-1)} H(\mathbf{u})^{-n - \{1/(n-1)\}}.$$

For $v \in S$, let $g(v)$ be the number of places on L lying above v. Since L is a normal extension of K, for every $w \in M_L$ lying above $v \in S$ there is a $\sigma_w \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$ such that $|\cdot|_w = |\sigma_w(\cdot)|_v^{1/g(v)}$. Hence

$$(8.12) \qquad \prod_{j \in I_w} \frac{|u_j|_w}{|\mathbf{u}|_w} = \left(\prod_{j \in I_w} \frac{|\sigma_w(u_j)|_v}{|\mathbf{u}|_v} \right)^{1/g(v)} = \left(\prod_{j \in \sigma_w^{-1}(I_w)} \frac{|u_j|_v}{|\mathbf{u}|_v} \right)^{1/g(v)}.$$

Let $I_v = \sigma_w^{-1}(I_w)$ for that place w lying above v for which the left-hand side of (8.12) is minimal and put $\mathbf{I} := (I_v : v \in S)$. Since σ_w^{-1} maps edges of $\mathcal{H}(W)$ to edges, it maps independent subsets of $\{1, ..., r\}$ to independent subsets. Therefore, $\mathbf{I}$ consists of independent sets of cardinality n. By the choices of $I_v, v \in S$ and by (8.12) we have, recalling that we have precisely $g(v)$ places w lying above v,

$$\prod_{j \in I_v} \frac{|u_j|_v}{|\mathbf{u}|_v} \leq \prod_{w | v} \prod_{j \in I_w} \frac{|u_j|_w}{|\mathbf{u}|_w} \quad \text{for } v \in S,$$

and by inserting this into (8.11) we obtain

$$\prod_{v\in S}\prod_{j\in I_v}\frac{|u_j|_v}{|\mathbf{u}|_v}\le 2^{n/2}H(W)^{1/(n-1)}H(\mathbf{u})^{-n-\{1/(n-1)\}}. \tag{8.13}$$

We complete the proof of Lemma 15 by deriving a lower bound for $\Delta(\mathbf{I},W)$. Take any basis $\{\mathbf{a}_1,...,\mathbf{a}_n\}$ of W. Let $s_1,...,s_p$ be the non-zero coordinates of $\mathbf{a}_1\wedge...\wedge\mathbf{a}_n$ and put $\mathbf{s}=(s_1,...,s_p)$. Note that $p\le\binom{r}{n}$. By the remark at the end of §3, for every $\sigma\in\mathrm{Gal}(\overline{\mathbb{Q}}/K),\sigma(s_1),...,\sigma(s_p)$ is up to signs a permutation of $s_1,...,s_p$ and by (3.9), $H(\mathbf{s})=\prod_{v\in M_K}|\mathbf{s}|_v$. Further, $(s_1\cdots s_p)^2$ is in K^* and therefore, $\prod_{v\in M_K}|s_1...s_p|_v=1$. Let $(\mathbf{a}_1\wedge...\wedge\mathbf{a}_n)_{I_v}=s_{i_v}$ for $v\in S$. Then

$$\begin{aligned}\Delta(\mathbf{I},W)&=\prod_{v\in S}|s_{i_v}|_v\cdot\prod_{v\notin S}|\mathbf{s}|_v=\left(\prod_{v\in S}\frac{|s_{i_v}|_v}{|\mathbf{s}|_v}\right)H(\mathbf{s})\\&\ge\left(\prod_{v\in S}\frac{|s_1...s_p|_v}{|\mathbf{s}|_v^p}\right)H(\mathbf{s})=\left(\prod_{v\in S}|\mathbf{s}|_v^p\cdot\prod_{v\notin S}|s_1...s_p|_v\right)^{-1}H(\mathbf{s})\\&\ge\left(\prod_{v\in M_K}|\mathbf{s}|_v\right)^{-p}H(\mathbf{s})=H(\mathbf{s})^{1-p}=H(\mathbf{a}_1\wedge...\wedge\mathbf{a}_n)^{1-p}\\&=H(W)^{1-p}\ge H(W)^{1-\binom{r}{n}}.\end{aligned}\tag{8.14}$$

By inserting this into (8.13) we arrive at

$$\begin{aligned}\prod_{j\in S}\prod_{j\in I_v}\frac{|u_j|_v}{|\mathbf{u}|_v}&\le 2^{n/2}H(W)^{\frac{1}{n-1}+\binom{r}{n}-1}\Delta(\mathbf{I},W)H(\mathbf{u})^{-n-\{1/(n-1)\}}\\&\le 2^{n/2}H(W)^{\binom{r}{n}}\Delta(\mathbf{I},W)H(\mathbf{u})^{-n-\{1/(n-1)\}}.\end{aligned}$$

This proves Lemma 15. □

§9. Proof of Theorem 4.

$K,S,s=|S|,r\ge 2,\Sigma$ and W have the same meaning as before. Let $\mathcal{P}$ be the symmetric partition of connected components of $\mathcal{H}(W)$. We recall that by $\mathbf{uc}$ we denote the coordinatewise product of the vectors $\mathbf{u}$ and $\mathbf{c}$. For $\mathbf{u}=(u_1,...,u_r)$ with $u_1...u_r\ne 0$ we write $\mathbf{u}^{-1}$ for $(u_1^{-1},...,u_r^{-1})$. Define the multiplicative group of $\mathbf{u}$ satisfying (2.3),

$$G_{S,\Sigma}=\{\mathbf{u}=(u_1,...,u_r)\in\Lambda_\Sigma^*:\ u_i/u_j\in\overline{\mathcal{O}}_S^*\text{ for }i,j=1,...,r\}.$$

Lemma 16. *Suppose that* $\dim W = n \geq 2$ *and that* $\mathcal{O}^*_{\mathcal{P},S}/\delta(\mathcal{O}^*_S)$ *is finite. Then the linear scattering of* $W \cap G_{S,\Sigma}$ *in* W *is at most*

$$(2^{66}r^4)^{n^2 s}.$$

Proof. For $\mathbf{u}_0 \in G_{S,\Sigma}$ define the K-vector space $\mathbf{u}_0^{-1}W = \{\mathbf{u}_0^{-1}\mathbf{u} : \mathbf{u} \in W\}$. Note that this is again an n-dimensional subspace of Λ_Σ. It is no restriction to assume that

$$H(\mathbf{u}^{-1}W) \geq H(W) \quad \text{for every } \mathbf{u} \in G_{S,\Sigma}. \tag{9.1}$$

Namely, suppose that W does not satisfy (9.1). Choose $\mathbf{u}_0 \in G_{S,\Sigma}$ such that

$$H(\mathbf{u}_0^{-1}W) = \min\left\{H(\mathbf{u}^{-1}W) : \mathbf{u} \in G_{S,\Sigma}\right\};$$

this minimum is assumed since the set at the right-hand side is discrete. Put $W' := \mathbf{u}_0^{-1}W$. Clearly, W' satisfies (9.1). Further, since $(\mathbf{u}_0\mathbf{c}, \mathbf{u}_0^{-1}\mathbf{u}) = (\mathbf{c}, \mathbf{u})$ we have $W'^{\perp} = \mathbf{u}_0 W^{\perp}$. Hence the hypergraphs $\mathcal{H}(W)$ and $\mathcal{H}(W')$ have the same edges, i.e. are equal. Moreover, the linear scattering of $W' \cap G_{S,\Sigma}$ in W' is the same as that of $W \cap G_{S,\Sigma}$ in W. Hence it suffices to prove Lemma 16 for W' instead of W.

Assume (9.1). We divide $W \cap G_{S,\Sigma}$ into the sets

$$\begin{aligned} D &:= \left\{\mathbf{u} \in W \cap G_{S,\Sigma} : \quad H(\mathbf{u}) \geq \{2H(W)\}^{r^n}\right\} \\ E &:= \left\{\mathbf{u} \in W \cap G_{S,\Sigma} : \quad H(\mathbf{u}) < \{2H(W)\}^{r^n}\right\} \end{aligned}$$

and estimate the linear scatterings of both sets.

Let $\mathbf{u} \in D$ and let $\mathbf{I} = (I_v : v \in S)$ be the collection of Lemma 15. Then

$$\begin{aligned} \prod_{v \in S} \prod_{i \in I_v} \frac{|u_i|_v}{|\mathbf{u}|_v} &\leq 2^{n/2} H(W)^{\binom{r}{n}} \cdot \Delta(\mathbf{I}, W) H(\mathbf{u})^{-n-\{1/(n-1)\}} \\ &\leq \Delta(\mathbf{I}, W) \cdot H(\mathbf{u})^{-n-\{1/n\}}, \end{aligned}$$

where we used that $H(\mathbf{u}) \geq \{2H(W)\}^{r^n}$. By Lemma 4, the set of $\mathbf{u} \in D$ corresponding to a *fixed* collection $\mathbf{I}$ as above has linear scattering in W at most $(2^{61n^2} r^{2n} n^{7n})^s$. For each I_v we have at most $\binom{r}{n}$ possibilities and so for $\mathbf{I}$ at most $\binom{r}{n}^s$ possibilities. It follows that the linear scattering of D in W is at most

$$N_D := \binom{r}{n}^s (2^{61n^2} r^{2n} n^{7n})^s \leq \tfrac{1}{2}(2^{66}r^4)^{n^2 s}.$$

Now let $\mathbf{u} = (u_1, ..., u_r) \in E$. We rewrite $H(\mathbf{u}^{-1}W)$. Choose any basis $\{\mathbf{a}_1, ..., \mathbf{a}_n\}$ of W. Then $(\mathbf{u}^{-1}\mathbf{a}_1 \wedge ... \wedge \mathbf{u}^{-1}\mathbf{a}_n)_I = \big(\prod_{i\in I} u_i\big)^{-1}(\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n)_I$, hence

$$H(\mathbf{u}^{-1}W) = \prod_v |\mathbf{u}^{-1}\mathbf{a}_1 \wedge ... \wedge \mathbf{u}^{-1}\mathbf{a}_n|_v = \prod_v \max_I \frac{|(\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n)_I|_v}{|\prod_{i\in I} u_i|_v},$$

where the maximum is taken over all independent subsets I of cardinality n of $\{1, ..., r\}$. For $v \in S$, let I_v be a set for which the maximum is assumed and put $\mathbf{I} := (I_v : v \in S)$. Since $u_i/u_j \in \overline{\mathcal{O}}_S^*$ we have $|u_i/u_j|_v = 1$ for $v \notin S,\ i, j = 1, ..., r$. Hence $|u_i|_v = |\mathbf{u}|_v$ for $i = 1, ..., r,\ v \notin S$. Therefore,

$$\begin{aligned} \max_I \frac{|(\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n)_I|_v}{|\prod_{i\in I} u_i|_v} &= |\mathbf{u}|_v^{-n} \max_I |(\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n)_I|_v \\ &= |\mathbf{u}|_v^{-n} |\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n|_v \quad \text{for } v \notin S. \end{aligned}$$

It follows that

$$\begin{aligned} H(\mathbf{u}^{-1}W) &= \prod_{v\in S} \frac{|(\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n)_{I_v}|_v}{\prod_{i\in I_v} |u_i|_v} \cdot \prod_{v\notin S} \frac{|\mathbf{a}_1 \wedge ... \wedge \mathbf{a}_n|_v}{|\mathbf{u}|_v^n} \\ &= \Bigg(\prod_{v\in S}\prod_{i\in I_v} \frac{|u_i|_v}{|\mathbf{u}|_v}\Bigg)^{-1} \Delta(\mathbf{I}, W) H(\mathbf{u})^{-n}. \end{aligned}$$

Together with (9.1) this implies that

$$\prod_{v\in S}\prod_{i\in I_v} \frac{|u_i|_v}{|\mathbf{u}|_v} \le \Delta(\mathbf{I}, W) H(\mathbf{u})^{-n} H(W)^{-1}. \tag{9.2}$$

Now Lemma 7 with $P = H(W),\ Q = 1,\ B = \{2H(W)\}^{r^n}$ implies that the set of $\mathbf{u} \in E$ corresponding to a fixed tuple $\mathbf{I}$ as in (9.2) has linear scattering in W at most

$$(270n^{10}r^{n+1})^{ns+1}.$$

Since for $\mathbf{I}$ we have at most $\binom{r}{n}^s$ possibilities, it follows that the linear scattering of E in W is at most

$$N_E := \binom{r}{n}^s \big(270n^{10}r^{n+1}\big)^{ns+1} \le \tfrac{1}{2}(2^{66}r^4)^{n^2 s}.$$

We conclude that $W \cap G_{S,\Sigma}$ has linear scattering in W at most

$$N_D + N_E \le (2^{66}r^4)^{n^2 s}.$$

This proves Lemma 16. □

Proof of Theorem 4.
Let $\mathcal{F}(W)$ be the set of vectors $\mathbf{u} \in W$ satisfying (2.3) and (2.4), i.e. $\mathbf{u} \in G_{S,\Sigma}$ and $\mathbf{u}$ is S-non-degenerate. By Lemma 11 (iii), we have $W = W_{\mathcal{P}}$ and so by definition (2.2), $\mathcal{F}(W)$ is empty if $\mathcal{O}^*_{\mathcal{P},S}/\delta(\mathcal{O}^*_S)$ is infinite. We have to estimate from above the number of K^*-cosets in $\mathcal{F}(W)$. We proceed by induction on $n = \dim_K W$.

First let $n = 2$. Then the number of K^*-cosets in $\mathcal{F}(W)$ is precisely the linear scattering of $\mathcal{F}(W)$ in W. This is 0 if $\mathcal{O}^*_{\mathcal{P},S}/\delta(\mathcal{O}^*_S)$ is infinite and by Lemma 16 at most

$$(2^{66}r^4)^{2^2 s} = (2^{33}r^2)^{2^3 s}$$

if $\mathcal{O}^*_{\mathcal{P},S}/\delta(\mathcal{O}^*_S)$ is finite.

Now let $n \geq 3$. Assume that for every K-linear subspace W_0 of Λ_Σ of dimension n_0 with $2 \leq n_0 < n$, the set $\mathcal{F}(W_0)$ of S-non-degenerate elements of W_0 belonging to $G_{S,\Sigma}$ is the union of at most $(2^{33}r^2)^{n_0^3 s} K^*$-cosets. We may assume that $\mathcal{O}^*_{\mathcal{P},S}/\delta(\mathcal{O}^*_S)$ is finite. Then by Lemma 16 we have

$$\mathcal{F}(W) \subseteq W_1 \cup ... \cup W_t,$$

where $W_1, ..., W_t$ are proper K-linear subspaces of W and

$$t \leq (2^{66}r^4)^{n^2 s}. \tag{9.3}$$

The notion 'S-non-degenerate element of W' depends on both the element and W but it follows easily from the definition that for any proper linear subspace W_0 of W, every S-non-degenerate element of W belonging to W_0 is an S-non-degenerate element of W_0. Therefore, $\mathcal{F}(W) \cap W_0 \subseteq \mathcal{F}(W_0)$ for $W_0 \subseteq W$. We conclude that

$$\mathcal{F}(W) \subseteq \mathcal{F}(W_1) \cup ... \cup \mathcal{F}(W_t).$$

Together with (9.3) and the induction hypothesis this implies that $\mathcal{F}(W)$ is the union of at most

$$(2^{66}r^4)^{n^2 s} \cdot (2^{33}r^2)^{(n-1)^3 s} = (2^{33}r^2)^{\{(n-1)^3+2n^2\}s} \leq (2^{33}r^2)^{n^3 s}$$

K^*-cosets. This completes the proof of Theorem 4. □

References.

[1] E. BOMBIERI, On the Thue-Mahler equation II, Acta Arith. 67 (1994), 69-96.

[2] E. BOMBIERI, W.M. SCHMIDT, On Thue's equation, Invent. Math. 88 (1987), 69-81.

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

[4] J.H. EVERTSE, On sums of S-units and linear recurrences, Compositio Math. 53 (1984), 225-244.

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

[6] J.H. EVERTSE, An improvement of the quantitative Subspace theorem, Compositio Math., to appear.

[7] J.H. EVERTSE, I. GAÁL & K. GYŐRY, On the numbers of solutions of decomposable polynomial equations, Arch. Math. 52 (1989), 337-353.

[8] J.H. EVERTSE, K. GYŐRY, Finiteness criteria for decomposable form equations, Acta Arith. 50 (1988), 357-379.

[9] K. GYŐRY, On the numbers of families of solutions of systems of decomposable form equations, Publ. Math. Debrecen 42 (1993), 65-101.

[10] M. LAURENT, Equations diophantiennes exponentielles, Invent. Math. Π78 (1984), 299-327.

[11] K. MAHLER, Zur Approximation algebraischer Zahlen, II. Über die Anzahl der Darstellungen ganzer Zahlen durch Binärformen, Math. Ann. 108 (1933), 37-55.

[12] A.J. v.d. POORTEN & H.P. SCHLICKEWEI, Additive relations in fields, J. Austral. Math. Soc (A) 51 (1991), 154-170.

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

[14] H.P. SCHLICKEWEI, On norm form equations, J. Number Th. 9 (1977), 370-380.

[15] H.P. SCHLICKEWEI, The quantitative Subspace Theorem for number fields, Comp. Math. 82 (1992), 245-274.

[16] H.P. SCHLICKEWEI, S-unit equations over number fields, Invent. Math. 102 (1990), 95-107.

[17] H.P. SCHLICKEWEI, On linear equations in elements of finitely generated groups, lecture at MSRI, Berkeley, February 1993.

[18] W.M. SCHMIDT, On heights of algebraic subspaces and diophantine approximations, Ann. Math. 85 (1967), 430-472.

[19] W.M. SCHMIDT, Linearformen mit algebraischen Koefficienten II, Math. Ann. 191 (1971), 1-20.

[20] W.M. SCHMIDT, Norm form equations, Ann. Math. 96 (1972), 526-551.

[21] W.M. SCHMIDT, The subspace theorem in Diophantine approximations, Compositio Math. 69 (1989), 121-173.

[22] W.M. SCHMIDT, The number of solutions of norm form equations, Trans. Am. Math. Soc. 317 (1990), 197-227.

[23] A. THUE, Über Annäherungswerte algebraischer Zahlen, J. reine angew. Math. 135 (1909), 284-305.