Acta Math., 182 (1999), 243–282

The zero multiplicity of linear recurrence sequences

by

WOLFGANG M. SCHMIDT

University of Colorado at Boulder
Boulder, Colorado, U.S.A.

1. Introduction

A *linear recurrence sequence of order t* is a sequence $\{u_n\}_{n\in\mathbb{Z}}$ of complex numbers satisfying a relation

$$u_n = c_1 u_{n-1} + \ldots + c_t u_{n-t} \quad (n \in \mathbb{Z}) \tag{1.1}$$

with $t>0$ and fixed coefficients $c_1, \ldots, c_t$, but no relation with fewer than t summands, i.e., no relation $u_n = c'_1 u_{n-1} + \ldots + c'_{t-1} u_{n-t+1}$. This implies in particular that the sequence is not the zero sequence, and that $c_t \neq 0$. The *companion polynomial* of the relation (1.1) is

$$\mathcal{P}(z) = z^t - c_1 z^{t-1} - \ldots - c_t.$$

Write

$$\mathcal{P}(z) = \prod_{i=1}^{k} (z-\alpha_i)^{a_i} \tag{1.2}$$

with distinct roots $\alpha_1, \ldots, \alpha_k$. The sequence is said to be *nondegenerate* if no quotient α_i/α_j $(1 \leqslant i < j \leqslant k)$ is a root of 1. The *zero multiplicity* of the sequence is the number of $n \in \mathbb{Z}$ with $u_n = 0$. For an introduction to linear recurrences and exponential equations, see [10].

A classical theorem of Skolem–Mahler–Lech [4] says that a nondegenerate linear recurrence sequence has finite zero multiplicity. Schlickewei [6] and van der Poorten and Schlickewei [5] derived upper bounds for the zero multiplicity when the members of the sequence lie in a number field K. These bounds depended on the order t, the degree of K, as well as on the number of distinct prime ideal factors in the decomposition of the fractional ideals (α_i) in K. More recently, Schlickewei [7] gave bounds which depend

The author was partially supported by NSF Grant DMS-9710314.

only on t and the degree of K. The linear recurrence sequence is called *simple* if all the roots of the companion polynomial are simple. Evertse, Schlickewei and Schmidt [3] showed that a simple, nondegenerate linear recurrence sequence of complex numbers has zero multiplicity bounded in terms of t only. The purpose of the present paper is to show that this holds for any nondegenerate sequence.

THEOREM. *Suppose that $\{u_n\}_{n\in\mathbb{Z}}$ is a nondegenerate linear recurrence sequence whose companion polynomial has k distinct roots of multiplicity $\leqslant a$. Then its zero multiplicity is under some bound $c(k,a)$. We may take*

$$c(k,a)=\exp((7k^a)^{8k^a}). \tag{1.3}$$

Our value for $c(k,a)$ is admittedly rather large; but it is preferable to give some value at all, rather than to say that "$c(k,a)$ is effectively computable". No special significance attaches to the numbers 7 and 8 in (1.3), which could easily be reduced. In the case of a simple linear recurrence, $a=1$, and our bound (1.3) is of the same general shape as the one given in [3].

COROLLARY. *The zero multiplicity of a nondegenerate recurrence sequence of order t is less than*

$$c(t)=\exp\exp\exp(3t\log t). \tag{1.4}$$

Proof. This is certainly true when $t=1$ or 2. When $t\geqslant 3$ we note that $k\leqslant t$, $a\leqslant t$, so that the zero multiplicity is

$$\begin{aligned}&\leqslant c(t,t)=\exp((7t^t)^{8t^t})=\exp\exp(8t^t(t\log t+\log 7))\\&<\exp\exp(t^{3t})=\exp\exp\exp(3t\log t).\end{aligned} \qquad \square$$

At the cost of some extra complication, the $\log t$ in (1.4) could be replaced by an absolute constant.

It is well known that a recurrence with the companion polynomial (1.2) is of the form

$$u_n=P_1(n)\alpha_1^n+\ldots+P_k(n)\alpha_k^n$$

where P_i is a polynomial of degree $\leqslant a_i-1$. The zero multiplicity therefore is the number of solutions $x\in\mathbb{Z}$ of the polynomial–exponential equation

$$P_1(x)\alpha_1^x+\ldots+P_k(x)\alpha_k^x=0. \tag{1.5}$$

Given a nonzero k-tuple $\mathbf{P}=(P_1,\ldots,P_k)$ of polynomials with

$$\deg P_i=t_i \quad (i=1,\ldots,k),$$

set

$$a = 1+\max_i t_i, \tag{1.6}$$

$$t = t(\mathbf{P}) = \sum_{i=1}^{k} (t_i+1). \tag{1.7}$$

Our Theorem and its Corollary can now be formulated as follows. *Suppose that $\alpha_1, ..., \alpha_k$ are in $\mathbb{C}^\times$, with no quotient α_i/α_j $(i\neq j)$ a root of unity. Then the number of solutions $x\in\mathbb{Z}$ of (1.5) does not exceed $c(k,a)$ or $c(t)$.*

A first, intuitive response to an equation (1.5) probably is that if all quotients α_i/α_j $(i\neq j)$ are "large" or "small", the summands in (1.5) will have different magnitudes when x is outside a limited range, so that there will be few zeros. As is basically known, and as we will explain again in §2, the Theorem can be reduced to the special case when $\alpha_1, ..., \alpha_k$ and the coefficients of the polynomials $P_1, ..., P_k$ are algebraic. The intuition can then be put into the more precise form that there should be few solutions if the (absolute logarithmic) heights $h(\alpha_i/\alpha_j)$ $(1\leqslant i,j\leqslant k;\ i\neq j)$ are not too small. As will be shown in §4, this intuition is correct. Note that $h(\alpha_i/\alpha_j)>0$ precisely when α_i/α_j is not a root of 1. A major difficulty now comes from the fact that when α_i/α_j is of large degree, the height, though positive, may be quite small.

The idea to overcome this difficulty is as follows. Write

$$P_i(x) = \sum_{j=1}^{a} a_{ij} x^{j-1} \quad (i=1, ..., k),$$

and set

$$N_j(X_1, ..., X_k) = \sum_{i=1}^{k} a_{ij} X_i \quad (j=1, ..., a).$$

The equation (1.5) may be rewritten as

$$\sum_{j=1}^{a} N_j(\alpha_1^x, ..., \alpha_k^x) x^{j-1} = 0. \tag{1.8}$$

Suppose that $\alpha_1, ..., \alpha_k$ and the coefficients a_{ij} lie in a number field K of degree D, and let $\xi\mapsto\xi^{(\sigma)}$ $(\sigma=1, ..., D)$ signify the embeddings $K\hookrightarrow\mathbb{C}$. Then, in an obvious notation, (1.8) gives rise to

$$\sum_{j=1}^{a} N_j^{(\sigma)}(\alpha_1^{(\sigma)x}, ..., \alpha_k^{(\sigma)x}) x^{j-1} = 0 \quad (\sigma=1, ..., D).$$

For each σ, this is a linear equation in $1, x, ..., x^{a-1}$. Hence given embeddings $\sigma_1, ..., \sigma_a$, we obtain a system of linear equations whose determinant must vanish, i.e.,

$$|N_j^{(\sigma_i)}(\alpha_1^{(\sigma_i)x}, ..., \alpha_k^{(\sigma_i)x})|_{1\leqslant i,j\leqslant a} = 0. \tag{1.9}$$

This equation is of purely exponential type, i.e., the coefficient of each exponential is a constant, and hence can be dealt with by methods developed elsewhere, e.g., in [3]. A difficulty in dealing with (1.9) is that the determinant is likely to have many exponentials

$$\big(\alpha_{i_1}^{(\sigma_1)} ... \alpha_{i_a}^{(\sigma_a)}\big)^x$$

with nonzero coefficients. A possible advantage for us is that when D is large, there will be many a-tuples $\sigma_1, ..., \sigma_a$, hence many equations (1.9) at our disposal.

A needed auxiliary result which may be of independent interest will be treated in an appendix.

Let us finally introduce the notation

$$\alpha \approx \beta$$

to mean that α, β are in $\mathbb{C}^\times$ and that α/β is a root of 1.

2. Specialization(1)

Let $\overline{\mathbb{Q}}$ be the algebraic closure of $\mathbb{Q}$ in $\mathbb{C}$. Let X, Y be algebraic varieties in $\mathbb{C}^k$ defined over $\mathbb{Q}$. It is well known that when $X\backslash Y$ is not empty, i.e., if there is a point $\boldsymbol{\alpha}\in\mathbb{C}^k$ lying in $X\backslash Y$, then there is in fact a point $\boldsymbol{\beta}\in\overline{\mathbb{Q}}^k$ lying in $X\backslash Y$. Moreover, when X is irreducible and of degree Δ, there is such a point $\boldsymbol{\beta}$ with degree $d(\boldsymbol{\beta}) := [\mathbb{Q}(\boldsymbol{\beta}) : \mathbb{Q}] \leqslant \Delta$.

When V is an algebraic variety defined over $\mathbb{Q}$ and $V\backslash Y$ is not empty, write $\delta(V\backslash Y)$ for the minimum degree of the points $\boldsymbol{\beta}\in\overline{\mathbb{Q}}^k$ in $V\backslash Y$. Write $\delta(V\backslash Y) = \infty$ when $V\backslash Y$ is empty.

LEMMA 1. *Let $X, Y, V_1, V_2, ...$ be algebraic varieties defined over $\mathbb{Q}$, and set $\mathcal{V} = \bigcup_{n=1}^{\infty} V_n$. Suppose that $\delta(V_n\backslash Y) \to \infty$ as $n \to \infty$, and that $X\backslash(Y\cup\mathcal{V})$ is not empty. Then there is a point $\boldsymbol{\beta}\in\overline{\mathbb{Q}}^k$ with*

$$\boldsymbol{\beta} \in X\backslash(Y\cup\mathcal{V}). \tag{2.1}$$

Proof. There is an irreducible component X' of X such that $X'\backslash(Y\cup\mathcal{V})$ is not empty. Let Δ be the degree of X', and $\mathcal{V}_\Delta$ the union of the varieties V_n with $\delta(V_n\backslash Y) \leqslant \Delta$.

(1) Some results of this and the next two sections first appeared in an unpublished manuscript of Schlickewei, Schmidt and Waldschmidt. *Added in proof.* This work has now been published: Zeros of linear recurrences. *Manuscripta Math.*, 98 (1998), 225–241.

Whereas $\mathcal{V}$ is not necessarily a variety, $\mathcal{V}_\Delta$ certainly is, and hence so is $Y \cup \mathcal{V}_\Delta$. Since $X' \backslash (Y \cup \mathcal{V}_\Delta)$ is not empty, there is by what we said above a point $\boldsymbol{\beta} \in X' \backslash (Y \cup \mathcal{V}_\Delta)$ with degree $d(\boldsymbol{\beta}) \leqslant \Delta$. This point cannot lie in a set $V_n \backslash Y$ with $\delta(V_n \backslash Y) > \Delta$, and hence cannot lie in $Y \cup \mathcal{V}$. □

When $\alpha_1, ..., \alpha_k$ in $\mathbb{C}^\times$ are given, and when $x \in \mathbb{Z}$, the equation (1.5) is linear in the $t = (t_1+1) + ... + (t_k+1)$ coefficients of the polynomials $P_1, ..., P_k$ of respective degrees $\leqslant t_1, ..., t_k$. Hence when $\mathcal{Z}$ is a subset of $\mathbb{Z}$, the totality of equations (1.5) with $x \in \mathcal{Z}$ defines a linear space in these coefficients. This linear space is $\neq \{\mathbf{0}\}$ precisely when $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_k)$ lies in a certain algebraic variety $X = X(\mathcal{Z}, t_1, ..., t_k)$. Thus when $\alpha_1, ..., \alpha_k$ are nonzero and if (1.5) holds for $x \in \mathcal{Z}$, then $\boldsymbol{\alpha} \in X \backslash Y$ where Y is given by $\alpha_1 ... \alpha_k = 0$.

Let $\Phi_m(x)$ be the mth cyclotomic polynomial, and $\Phi_m(x,y) = y^{\phi(m)} \Phi_m(x/y)$ its homogeneous version. For $1 \leqslant i < j \leqslant k$, let V_{ijm} be the variety in $\mathbb{C}^k$ defined by $\Phi_m(\alpha_i, \alpha_j) = 0$. Then $\delta(V_{ijm} \backslash Y) = \phi(m)$. Now if, in addition to the condition on $\boldsymbol{\alpha}$ given above, we have $\alpha_i \not\approx \alpha_j$ for $i \neq j$, then $\boldsymbol{\alpha} \notin \mathcal{V} = \bigcup_i \bigcup_j \bigcup_m V_{ijm}$, so that $\boldsymbol{\alpha} \in X \backslash (Y \cup \mathcal{V})$. By Lemma 1, there is a $\boldsymbol{\beta} \in \overline{\mathbb{Q}}^k$ with (2.1). This $\boldsymbol{\beta}$ has nonzero components with $\beta_i \not\approx \beta_j$ for $i \neq j$, and there are polynomials $\widetilde{P}_1, ..., \widetilde{P}_k$ of respective degrees $\leqslant t_1, ..., t_k$, not all zero, so that

$$\widetilde{P}_1(x)\beta_1^x + ... + \widetilde{P}_k(x)\beta_k^x = 0$$

for $x \in \mathcal{Z}$.

It is therefore clear that in proving our Theorem, we may suppose from now on that $\alpha_1, ..., \alpha_k$ are algebraic. They will lie in some number field K. The equation (1.5) with $x \in \mathcal{Z}$ is linear, with coefficients in K, in the coefficients of $P_1, ..., P_k$, and if these equations have a nontrivial solution, they have a nontrivial solution with components in K.

In summary: *We may suppose that $\alpha_1, ..., \alpha_k$ and the coefficients of $P_1, ..., P_k$ lie in a number field K.*

3. A survey of some known results

We will quote a few facts which will be used in our proof of the Theorem.

LEMMA 2. *Let $\alpha_1, ..., \alpha_q, a_1, ..., a_q$ be in $\mathbb{C}^\times$, and consider the exponential equation*

$$a_1\alpha_1^x + ... + a_q\alpha_q^x = 0. \tag{3.1}$$

When $\alpha_i \not\approx \alpha_j$ for $i \neq j$ in $1 \leqslant i, j \leqslant q$, the number of solutions $x \in \mathbb{Z}$ is less than

$$A(q) = \exp((6q)^{4q}).$$

Proof. This follows immediately from Theorem 1.2 in [3].(2) □

A solution x of (3.1) is called *nondegenerate* if no subsum vanishes.

LEMMA 3. *Again let* $\alpha_1, ..., \alpha_q, a_1, ..., a_q$ *be in* $\mathbb{C}^\times$, *but this time suppose* $\alpha_1 \approx ... \approx \alpha_q$. *There are*

$$B(q) = q^{3q^2}$$

vectors $\mathbf{c}^{(w)} = (c_1^{(w)}, ..., c_q^{(w)})$ $(w=1, ..., B(q))$ *such that for any nondegenerate solution of* (3.1), *the vector* $(\alpha_1^x, ..., \alpha_q^x)$ *is proportional to some vector* $\mathbf{c}^{(w)}$.

Proof. We may suppose that $q>1$. Setting $n=q-1$, $b_i=-a_i/a_q$, $\zeta_i=(\alpha_i/\alpha_q)^x$ $(i=1, ..., n)$, we obtain

$$b_1\zeta_1 + ... + b_n\zeta_n = 1 \tag{3.2}$$

where $\zeta_1, ..., \zeta_n$ are roots of 1. By a recent result of Evertse [2] which improves on earlier work of Schlickewei [8], the equation (3.2) has at most $B(n+1)=B(q)$ solutions in roots of unity where no subsum of $b_1\zeta_1+...+b_n\zeta_n$ vanishes. Given such a solution $\zeta_1, ..., \zeta_n$, the vector $(\alpha_1^x, ..., \alpha_q^x)$ is proportional to $(\zeta_1, ..., \zeta_n, 1)$. □

A solution $\mathbf{x}=(x_1, ..., x_q)$ of an equation

$$a_1x_1 + ... + a_qx_q = 0 \tag{3.3}$$

is called *nondegenerate* if no subsum vanishes.

LEMMA 4. *Let* Γ *be a finitely generated subgroup of* $(\mathbb{C}^\times)^q = \mathbb{C}^\times \times ... \times \mathbb{C}^\times$ *of rank* r, *and let* $a_1, ..., a_q$ *be in* $\mathbb{C}^\times$. *Then up to a factor of proportionality,* (3.3) *has at most*

$$C(q,r) = \exp((r+1)(6q)^{4q}) \tag{3.4}$$

nondegenerate solutions $\mathbf{x} \in \Gamma$.

Proof. This is just a homogeneous version of a theorem in [3]. Again set $n=q-1$, $b_i=-a_i/a_q$, and write $y_i=x_i/x_q$ $(i=1, ..., n)$. Then (3.3) becomes

$$b_1y_1 + ... + b_ny_n = 1, \tag{3.5}$$

and $(y_1, ..., y_n)$ lies in a group Γ' of rank $\leqslant r$. By Theorem 1.1 of [3], (3.5) has at most(3)

$$\exp((r+1)(6n)^{4n}) < C(q,r)$$

(2) *Added in proof.* The estimate in the final version of [3] is slightly better.

(3) *Added in proof.* Again the estimate in the final version of [3] is better, but effects no essential improvement of our main results.

solutions $(y_1,...,y_n)\in\Gamma'$ where no subsum of $b_1y_1+...+b_ny_n$ vanishes. Since $x_i=y_ix_q$, the lemma follows. $\square$

Let $h(x_1:...:x_q)$ denote the absolute logarithmic height of a point $\mathbf{x}=(\mathbf{x}_1:...:\mathbf{x}_q)$ in projective space $\mathbb{P}_{q-1}(\overline{\mathbb{Q}})$. Let $h_{\text{in}}(x_1,...,x_n)$ be the inhomogeneous height of a point $\mathbf{x}\in\overline{\mathbb{Q}}^n$, so that $h_{\text{in}}(x_1,...,x_n)=h(x_1:...:x_n:1)$. Given a number $\alpha\in\overline{\mathbb{Q}}$, there should hopefully be no confusion writing $h(\alpha)=h_{\text{in}}(\alpha)=h(\alpha:1)$.

When $\mathbf{x}=(x_1,...,x_q)$, $\mathbf{y}=(y_1,...,y_q)$, set

$$\mathbf{x}*\mathbf{y}=(x_1y_1,...,x_qy_q). \tag{3.6}$$

LEMMA 5. *Let $q>1$ and Γ be a finitely generated subgroup of $(\overline{\mathbb{Q}}^\times)^q$ of rank r. Then the solutions of*

$$z_1+...+z_q=0, \tag{3.7}$$

with $\mathbf{z}=(z_1,...,z_q)=\mathbf{x}*\mathbf{y}$ where $\mathbf{x}\in\Gamma$, $\mathbf{y}\in(\mathbb{Q}^\times)^q$ and

$$h(\mathbf{y})\leqslant\frac{1}{4q^2}h(\mathbf{x}),$$

are contained in the union of not more than $C(q,r)$ proper subspaces of the ($(q-1)$-dimensional) space defined by (3.7).

Proof. Set $n=q-1$. The lemma is an immediate consequence of the following inhomogeneous version.

LEMMA 5′. *Let Γ be a finitely generated subgroup of $(\overline{\mathbb{Q}}^\times)^n$ of rank r. Then the solutions of*

$$z_1+...+z_n=1, \tag{3.8}$$

with $\mathbf{z}=(z_1,...,z_n)=\mathbf{x}*\mathbf{y}$ where $\mathbf{x}\in\Gamma$, $\mathbf{y}\in\mathbb{Q}^n$ and

$$h_{\text{in}}(\mathbf{y})\leqslant\frac{1}{4n^2}h_{\text{in}}(\mathbf{x}), \tag{3.9}$$

are contained in the union of not more than $C(n,r)$ proper subspaces of $\overline{\mathbb{Q}}^n$.

This is a variation on Proposition A of [9]. In that proposition, the bound on the number of subspaces depended on the degree of the number field generated by Γ. But in contrast to our estimate $C(n,r)$, that bound was not doubly exponential.

Proof of Lemma 5′. In the proof of Proposition A we distinguished three kinds of solutions.

(i) Solutions where some $y_i=0$, i.e., some $z_i=0$. These clearly lie in n subspaces.

(ii) Solutions where each $y_i \neq 0$, and where $h_{\text{in}}(\mathbf{x}) > 2n \log n$. These were called *large* solutions in [9], and it was shown in (10.4) of that paper that they lie in the union of fewer than

$$2^{30n^2}(21n^2)^r$$

proper subspaces.

(iii) Solutions where each $y_i \neq 0$ and where $h_{\text{in}}(\mathbf{x}) \leqslant 2n \log n$. These were called *small* solutions in [9]. Here we argue as follows. We have $h_{\text{in}}(\mathbf{y}) \leqslant (2n \log n)/(4n^2) < \log 2$ by (3.9). Then each component has $h_{\text{in}}(y_i) < \log 2$, and since $y_i \in \mathbb{Q}^\times$, we have $y_i = \pm 1$. The equation (3.8) now becomes

$$\pm x_1 \pm x_2 \pm \ldots \pm x_n = 1. \tag{3.10}$$

The group Γ' generated by Γ and the points $(\pm 1, \ldots, \pm 1)$ again is finitely generated, and of rank r. By Proposition 2.1 of [3], the solution of (3.10) with $(\pm x_1, \ldots, \pm x_n) \in \Gamma'$ lies in the union of not more than

$$\exp((4n)^{3n} \cdot 2(r+1))$$

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

Combining our estimates we obtain

$$n + 2^{30n^2}(21n^2)^r + \exp((4n)^{3n} \cdot 2(r+1)) < C(n, r). \qquad \square$$

LEMMA 6. *Let β, b in $\overline{\mathbb{Q}}^\times$ be given. Then there is a $u \in \mathbb{Z}$ such that*

$$h(b\beta^{x-u}) \geqslant \tfrac{1}{4} h(\beta)|x|$$

for $x \in \mathbb{Z}$.

This is the case $r = n = 1$ of Lemma 15.1 in [9]. For the convenience of the reader, we will present the proof of our special case.

Proof. We may suppose that $h(\beta) > 0$. Let $K = \mathbb{Q}(b, \beta)$ and M be the set of places of K. With $v \in M$ we associate the absolute value $|\cdot|_v$ on K which extends the standard or a p-adic absolute value on $\mathbb{Q}$, as well as the renormalized absolute value $\|\cdot\|_v = (|\cdot|_v)^{d_v/D}$, where $D = \deg K$ and d_v is the local degree belonging to v. Then when $\alpha \in K^\times$,

$$h(\alpha) = \sum_{v \in M} \max(0, \log \|\alpha\|_v) = \frac{1}{2} \sum_{v \in M} |\log \|\alpha\|_v|$$

by the product formula. Hence

$$h(b\beta^x) = \frac{1}{2} \sum_{v \in M} |\log \|b\|_v + x \log \|\beta\|_v|.$$

Defining

$$\psi(\xi,\zeta)=\frac{1}{2}\sum_{v\in M}|\xi\log\|\beta\|_v+\zeta\log\|b\|_v|$$

for $(\xi,\zeta)\in\mathbb{R}^2$, we have

$$\psi(x,1)=h(b\beta^x),\quad \psi(\xi,0)=|\xi|h(\beta). \tag{3.11}$$

The function ψ has $\psi(\xi+\xi',\zeta+\zeta')\leqslant\psi(\xi,\zeta)+\psi(\xi',\zeta')$, as well as $\psi(\lambda\xi,\lambda\zeta)=|\lambda|\psi(\xi,\zeta)$ for $\lambda\in\mathbb{R}$. The set $\Psi\subset\mathbb{R}^2$ consisting of points (ξ,ζ) with $\psi(\xi,\zeta)\leqslant 1$ is convex, symmetric about $\mathbf{0}$, closed, and it contains $\mathbf{0}$ in its interior. But it may be unbounded.

When Ψ is unbounded, there is some $(\xi_0,\zeta_0)\neq(0,0)$ with $\psi(\xi_0,\zeta_0)=0$. Since $\psi(1,0)=h(\beta)>0$, we have $\zeta_0\neq 0$. By homogeneity, there is some ξ_1 with $\psi(\xi_1,1)=0$. On the other hand, when Ψ is bounded, hence compact, pick (ξ_0,ζ_0) in Ψ with maximal possible ζ_0. Writing ξ_0 as $\xi_0=\zeta_0\xi_1$ we obtain $\zeta_0(\xi_1,1)\in\Psi$, hence $\zeta_0\psi(\xi_1,1)\leqslant 1$.

Let (ξ,ζ) be given. When Ψ is unbounded, $\psi(\zeta\xi_1,\zeta)=|\zeta|\psi(\xi_1,1)=0\leqslant\psi(\xi,\zeta)$. When Ψ is bounded, we have $\psi(\zeta\xi_1,\zeta)=|\zeta|\psi(\xi_1,1)\leqslant|\zeta|/\zeta_0\leqslant\psi(\xi,\zeta)$, with the last inequality due to homogeneity and the maximality of ζ_0. Taking the difference of (ξ,ζ) and $(\zeta\xi_1,\zeta)$, we obtain $\psi(\xi-\zeta\xi_1,0)\leqslant 2\psi(\xi,\zeta)$, and hence

$$|\xi-\zeta\xi_1|h(\beta)\leqslant 2\psi(\xi,\zeta)$$

by (3.11). Setting $\zeta=1$ and replacing ξ by $x\in\mathbb{Z}$, we have

$$h(b\beta^x)=\psi(x,1)\geqslant\tfrac{1}{2}|x-\xi_1|h(\beta).$$

We pick $u\in\mathbb{Z}$ such that $\xi_1=-u+\mu$ with $|\mu|\leqslant\frac{1}{2}$. Then

$$h(b\beta^{x-u})\geqslant\tfrac{1}{2}|x-u-\xi_1|h(\beta)=\tfrac{1}{2}|x-\mu|h(\beta)\geqslant\tfrac{1}{4}h(\beta)|x|. \qquad \square$$

4. Consequences of having some height $h(\alpha_i/\alpha_j)$ not too small

Define the degree of the zero polynomial to be -1. Given a k-tuple $\mathbf{P}=(P_1,\ldots,P_k)$ of polynomials where $\deg P_i=t_i$ $(i=1,\ldots,k)$, define $t(\mathbf{P})$ by (1.7), and set

$$t^*(\mathbf{P})=1+\max_i t_i.$$

Note that a zero polynomial does not contribute to $t(\mathbf{P})$.

LEMMA 7. *Consider the equation* (1.5), *i.e.,*

$$P_1(x)\alpha_1^x+\ldots+P_k(x)\alpha_k^x=0, \tag{4.1}$$

where $(\alpha_1,\ldots,\alpha_k)\in(\overline{\mathbb{Q}}^\times)^k$ *and where each* P_i *is nonzero and has coefficients in* $\overline{\mathbb{Q}}$. *Suppose that* $t(\mathbf{P})\geqslant 3$ *and that*

$$\max_{i,j} h(\alpha_i:\alpha_j)\geqslant \hbar, \tag{4.2}$$

where $0<\hbar\leqslant 1$. *Set* $t=t(\mathbf{P})$, $t^*=t^*(\mathbf{P})$,

$$E=16t^2\cdot t^*/\hbar, \quad F=\exp((6t)^{5t})+5E\log E.$$

Then there are k-tuples

$$\mathbf{P}^{(w)}=(P_1^{(w)},\ldots,P_k^{(w)})\neq(0,\ldots,0) \quad (1\leqslant w<F)$$

of polynomials with

$$\begin{aligned}&\deg P_i^{(w)}\leqslant t_i \quad (1\leqslant w<F,\ 1\leqslant i<k),\\&\deg P_k^{(w)}<t_k \quad (1\leqslant w<F),\end{aligned}$$

such that every solution $x\in\mathbb{Z}$ *of* (4.1) *satisfies*

$$P_1^{(w)}(x)\alpha_1^x+\ldots+P_k^{(w)}(x)\alpha_k^x=0 \tag{4.3}$$

for some w *in* $1\leqslant w<F$.

Proof. Suppose $u\in\mathbb{Z}$, and set $y=x+u$. Then (4.1) may be rewritten as

$$P_1(y-u)\alpha_1^{-u}\alpha_1^y+\ldots+P_k(y-u)\alpha_k^{-u}\alpha_k^y=0,$$

which is the same as

$$Q_1(y)\alpha_1^y+\ldots+Q_k(y)\alpha_k^y=0 \tag{4.4}$$

with

$$Q_i(y)=P_i(y-u)\alpha_i^{-u} \quad (i=1,\ldots,k).$$

Suppose our assertion is true for (4.4), with polynomial k-tuples $\mathbf{Q}^{(w)}=(Q_1^{(w)},\ldots,Q_k^{(w)})$ $(1\leqslant w<F)$. Thus every solution $y\in\mathbb{Z}$ of (4.4) satisfies

$$Q_1^{(w)}(y)\alpha_1^y+\ldots+Q_k^{(w)}(y)\alpha_k^y=0 \tag{4.5}$$

for some w. But then $x=y-u$ satisfies (4.3) with $P_i^{(w)}(x)=Q_i^{(w)}(x+u)\alpha_i^u$ $(i=1, ..., k)$. We therefore may make a change of variables $x\mapsto y=x+u$.

We may suppose that $h(\alpha_1:\alpha_2)\geqslant\hbar$. Write

$$P_i(x)=a_{i0}+a_{i1}x+...+a_{i,t_i}x^{t_i}.$$

Pick u according to Lemma 6 such that

$$h(a_{1,t_1}\alpha_1^{y-u}:a_{2,t_2}\alpha_2^{y-u})=h\left(\frac{a_{1,t_1}}{a_{2,t_2}}\left(\frac{\alpha_1}{\alpha_2}\right)^{y-u}\right)\geqslant\frac{1}{4}h\left(\frac{\alpha_1}{\alpha_2}\right)|y|\geqslant\frac{1}{4}\hbar|y|.$$

Setting

$$Q_i(y)=P_i(y-u)\alpha_i^{-u}=b_{i0}+b_{i1}y+...+b_{i,t_i}y^{t_i},$$

we have $b_{1,t_1}=a_{1,t_1}\alpha_1^{-u}$, $b_{2,t_2}=a_{2,t_2}\alpha_2^{-u}$, so that

$$h(b_{1,t_1}\alpha_1^y:b_{2,t_2}\alpha_2^y)\geqslant\tfrac{1}{4}\hbar|y| \tag{4.6}$$

for $y\in\mathbb{Z}$.

The equation (4.4) is of the form

$$(b_{10}+b_{11}y+...+b_{1,t_1}y^{t_1})\alpha_1^y+...+(b_{k0}+b_{k1}y+...+(b_{k,t_k}y^{t_k})\alpha_k^y=0.$$

Some coefficients may be zero; omitting the zero coefficients we rewrite this as

$$(b'_{10}y^{v_{10}}+...+b_{1,t_1}y^{t_1})\alpha_1^y+...+(b'_{k0}y^{v_{k0}}+...+b_{k,t_k}y^{t_k})\alpha_k^y=0.$$

Let q be the total number of (nonzero) coefficients here, and consider the following vectors in q-dimensional space:

$$\begin{aligned}\mathbf{X}&=(b'_{10}\alpha_1^y, ..., b_{1,t_1}\alpha_1^y, ..., b'_{k0}\alpha_k^y, ..., b_{k,t_k}\alpha_k^y),\\ \mathbf{Y}&=(\ y^{v_{10}}\ , ..., \quad y^{t_1}\quad , ..., \ y^{v_{k0}}\ , ..., \quad y^{t_k}\quad).\end{aligned}$$

Our equation becomes

$$Z_1+...+Z_q=0 \tag{4.7}$$

with $\mathbf{Z}=\mathbf{X}*\mathbf{Y}=(X_1Y_1, ..., X_qY_q)$. Here $\mathbf{X}$ lies in the group Γ of rank $r\leqslant 2$ generated by the points $(b'_{10}, ..., b_{1,t_1}, ..., b'_{k0}, ..., b_{k,t_k})$ and $(\alpha_1, ..., \alpha_1, ..., \alpha_k, ..., \alpha_k)$. Further

$$h(\mathbf{X})\geqslant h(b_{1,t_1}\alpha_1^y:b_{2,t_2}\alpha_2^y)\geqslant\tfrac{1}{4}\hbar|y| \tag{4.8}$$

by (4.6). On the other hand, $\mathbf{Y}\in\mathbb{Q}^q$, in fact $\mathbf{Y}\in(\overline{\mathbb{Q}}^\times)^q$ when $y\neq 0$, and $h(\mathbf{Y})\leqslant t^*\log|y|$ since each $t_i\leqslant t^*$. Therefore when

$$|y|\geqslant 2E\log E, \tag{4.9}$$

so that $|y|\geqslant(32q^2t^*/\hbar)\log(16q^2t^*/\hbar)$ in view of $q\leqslant t$, then

$$|y|>\frac{16q^2t^*}{\hbar}\log|y|,$$

and

$$h(\mathbf{Y})\leqslant t^*\log|y|<\frac{\hbar}{16q^2}|y|=\frac{1}{4q^2}\cdot\frac{\hbar}{4}|y|\leqslant\frac{1}{4q^2}h(\mathbf{X})$$

by (4.8). Invoking Lemma 5, we see that for such y the vector $\mathbf{Z}$ is contained in the union of at most

$$C(q,2)<\exp((6q)^{5q})\leqslant\exp((6t)^{5t}) \tag{4.10}$$

proper subspaces of the space (4.7). Consider such a subspace $c_1Z_1+\ldots+c_qZ_q=0$ (where $(c_1,\ldots,c_q)$ is not proportional to $(1,\ldots,1)$). Taking a linear combination of this and (4.7) we obtain a nontrivial relation $c_1'Z_1+\ldots+c_{q-1}'Z_{q-1}=0$. But this means exactly that y satisfies a nontrivial equation

$$\widetilde{Q}_1(y)\alpha_1^y+\ldots+\widetilde{Q}_k(y)\alpha_k^y=0, \tag{4.11}$$

where $\deg\widetilde{Q}_i\leqslant t_i$ $(i=1,\ldots,k-1)$, $\deg\widetilde{Q}_k<t_k$.

There are not more than $5E\log E$ values of y where (4.9) is violated. For fixed y, and since $t\geqslant 3$, there will certainly be polynomials $\widetilde{Q}_1,\ldots,\widetilde{Q}_k$, not all zero, with (4.11) and the same restriction on their degrees. Altogether we get fewer than F polynomial k-tuples $\widetilde{\mathbf{Q}}=(\widetilde{Q}_1,\ldots,\widetilde{Q}_k)$, where F is the sum of the right-hand side of (4.10), and of $5E\log E$. $\square$

Lemma 7 gives us a possible opening to prove our Theorem. Note that each $\mathbf{P}^{(w)}$ has $t(\mathbf{P}^{(w)})<t(\mathbf{P})$, so that we can start induction on $t=t(\mathbf{P})$, *provided* (4.2) holds with some $\hbar=\hbar(t)>0$ independent of the degrees of $\alpha_1,\ldots,\alpha_k$. But in general such a condition (4.2) will be hard to satisfy.

5. A proposition which implies the Theorem

An n-tuple of linear forms $M_1,\ldots,M_n$ in a variable vector $\mathbf{X}$ will be called linearly independent over $\mathbb{Q}$ if there is no $\mathbf{y}=(y_1,\ldots,y_n)\in\mathbb{Q}^n\setminus\{\mathbf{0}\}$ such that we have identically

$$y_1M_1(\mathbf{X})+\ldots+y_nM_n(\mathbf{X})=0.$$

PROPOSITION. *Suppose that linear forms* $M_1, ..., M_n$ *in* $\mathbf{X}=(X_1, ..., X_k)$ *have algebraic coefficients and are linearly independent over* $\mathbb{Q}$. *Let* $\alpha_1, ..., \alpha_k$ *be algebraic and have* $\alpha_i \not\approx \alpha_j$ *when* $i \neq j$. *Consider numbers* $x \in \mathbb{Z}$ *for which*

$$M_1(\alpha_1^x, ..., \alpha_k^x), \quad ..., \quad M_n(\alpha_1^x, ..., \alpha_k^x) \tag{5.1}$$

are linearly dependent over $\mathbb{Q}$. *These number fall into at most*

$$H(k,n) = \exp((7k^n)^{6k^n})$$

classes with the following property. For each class C, *there is a natural number* m *such that*

(a) *solutions* x, x' *in* C *have* $x \equiv x' \pmod{m}$,

(b) *there are* i, j *with* $h(\alpha_i^m : \alpha_j^m) \geqslant \hbar$, *where*

$$\hbar = \hbar(k,n) = e^{-10k^{2n}}. \tag{5.2}$$

We will now deduce the Theorem. We are concerned with (1.5), where we write P_i in the form

$$P_i(x) = \sum_{j=1}^{a} a_{ij} x^{j-1} \quad (i = 1, ..., k)$$

with $a = 1 + \max_i \deg P_i$. Define linear forms

$$N_j(\mathbf{X}) = \sum_{i=1}^{k} a_{ij} X_i \quad (j = 1, ..., a)$$

in $\mathbf{X}=(X_1, ..., X_k)$. Then as already noted in the Introduction, (1.5) may be written as

$$\sum_{j=1}^{a} N_j(\alpha_1^x, ..., \alpha_k^x) x^{j-1} = 0. \tag{5.3}$$

Here $N_1, ..., N_a$ are not necessarily independent over $\mathbb{Q}$. Let n be the maximum number of independent ones among them. There are linear forms $M_1, ..., M_n$, linearly independent over $\mathbb{Q}$, such that

$$N_j(\mathbf{X}) = \sum_{r=1}^{n} c_{jr} M_r(\mathbf{X}) \quad (j = 1, ..., a)$$

with rational coefficients c_{jr}. Then (5.3) becomes

$$\sum_{r=1}^{n} \left(\sum_{j=1}^{a} c_{jr} x^{j-1} \right) M_r(\alpha_1^x, ..., \alpha_k^x) = 0. \tag{5.4}$$

There are less than a numbers x where

$$\sum_{j=1}^{a} c_{jr} x^{j-1} = 0 \quad (r=1, \dots, n).$$

For the other solutions of (5.4), the n numbers in (5.1) are linearly dependent over $\mathbb{Q}$. By the Proposition, these numbers fall into at most $H(k,n)$ classes. Let us look at solutions in a fixed class.

When x_0 is a solution in the class, every solution in the class is of the type $x = x_0 + mz$ with $z \in \mathbb{Z}$. In terms of z, the original equation (1.5) becomes

$$\widehat{P}_1(z)\widehat{\alpha}_1^z + \dots + \widehat{P}_k(z)\widehat{\alpha}_k^z = 0, \tag{5.5}$$

where $\widehat{P}_i(z) = \alpha_i^{x_0} P_i(x_0 + mz)$, $\widehat{\alpha}_i = \alpha_i^m$ $(i=1, \dots, k)$. But now for some i, j,

$$h(\widehat{\alpha}_i : \widehat{\alpha}_j) \geqslant \hbar(k,n).$$

We will prove that when $t(\mathbf{P}) = t$, the equation (1.5) has at most

$$Z(t, k^a) = \exp(t(7k^a)^{7k^a}) \tag{5.6}$$

solutions x. We clearly may suppose that $k \geqslant 2$, $t \geqslant 3$. We will prove our assertion by induction on t in $3 \leqslant t \leqslant k^a$. We apply Lemma 7 to (5.5). Since $t^*(\mathbf{P}) \leqslant t$, $n \leqslant a$, we have

$$\begin{aligned} E &\leqslant 16t^3/\hbar(k,n) \leqslant 16k^{3a} e^{10k^{2a}} < e^{13k^{2a}}, \\ 5E \log E &< 65k^{2a} \cdot e^{13k^{2a}} < e^{18k^{2a}}, \\ F &\leqslant \exp((6t)^{5t}) + \exp(18k^{2a}) \leqslant \exp((6k^a)^{5k^a}) + \exp(18k^{2a}) < \exp((7k^a)^{5k^a}). \end{aligned}$$

By Lemma 7, each solution of (5.5) satisfies an equation with a polynomial vector $\mathbf{P}^{(w)} = (P_1^{(w)}, \dots, P_k^{(w)}) \neq (0, \dots, 0)$ with $1 \leqslant w < F$ having $t(\mathbf{P}^{(w)}) < t$. By induction on t, each such equation has at most $Z(t-1, k^a)$ solutions. We therefore obtain

$$\begin{aligned} &< a + H(k,n) F \cdot Z(t-1, k^a) \\ &\leqslant a + \exp((7k^n)^{6k^n} + (7k^a)^{5k^a}) \cdot \exp((t-1)(7k^a)^{7k^a}) \\ &< \exp(t(7k^a)^{7k^a}) = Z(t, k^a) \end{aligned}$$

solutions, establishing (5.6).

Since $t \leqslant k^a$, the number of solutions of (1.5) certainly is

$$< \exp((7k^a)^{8k^a}). \qquad \square$$

6. Splitting of exponential equations

Let nonzero $a_1, ..., a_q$, $\alpha_1, ..., \alpha_q$ be given. We consider the function

$$f(x) = a_1\alpha_1^x + ... + a_q\alpha_q^x. \tag{6.1}$$

We group together summands $a_i\alpha_i^x$ and $a_j\alpha_j^x$ where $\alpha_i \approx \alpha_j$. After relabeling, we may write (uniquely up to ordering)

$$f(x) = f_1(x) + ... + f_g(x) \tag{6.2}$$

where

$$f_i(x) = a_{i1}\alpha_{i1}^x + ... + a_{i,q_i}\alpha_{i,q_i}^x \quad (i = 1, ..., g)$$

with $q_1 + ... + q_g = q$ and

$$\alpha_{ij} \approx \alpha_{ik} \quad \text{when } 1 \leqslant i \leqslant g,\ 1 \leqslant j, k \leqslant q_i,$$
$$\alpha_{ij} \not\approx \alpha_{i'k} \quad \text{when } 1 \leqslant i \neq i' \leqslant g,\ 1 \leqslant j \leqslant q_i,\ 1 \leqslant k \leqslant q_{i'}.$$

LEMMA 8. *All but at most*

$$G(q) = \exp((7q)^{4q}) \tag{6.3}$$

solutions $x \in \mathbb{Z}$ *of* $f(x) = 0$ *have*

$$f_1(x) = ... = f_g(x) = 0. \tag{6.4}$$

We will say that the equation $f(x)=0$ *splits* into the g equations (6.4).

Proof. The lemma is nontrivial only when $g \geqslant 2$; and then $q = q(f) \geqslant 2$. We proceed by induction on q. When $q=2$ and $g=2$, we have in fact $f(x) = a\alpha_{11}^x + b\alpha_{21}^x$ with $ab \neq 0$ and $\alpha_{11} \not\approx \alpha_{21}$. There can be at most one $x \in \mathbb{Z}$ with $f(x)=0$.

We now turn to the step $q-1 \to q$ where $q \geqslant 3$. Observe that $(\alpha_1^x, ..., \alpha_q^x)$ lies in a group Γ of rank $r \leqslant 1$. By Lemma 4, there are at most $C(q,1) = \exp(2 \cdot (6q)^{4q})$ vectors $\mathbf{c}^{(l)} = (c_1^{(l)}, ..., c_q^{(l)})$, $1 \leqslant l \leqslant C(q,2)$, such that for every nondegenerate solution $x \in \mathbb{Z}$ of $f(x)=0$ we have $(\alpha_1^x, ..., \alpha_q^x)$ proportional to some $\mathbf{c}^{(l)}$. Thus the quotients $(\alpha_i/\alpha_j)^x$ depend only on l. But since $g \geqslant 2$, some α_i/α_j is not a root of 1, so that for given l, there can be at most one solution $x \in \mathbb{Z}$.

When x is a degenerate solution of $f(x)=0$, there is a nontrivial partition of $\{1, ..., q\}$ into subsets $\{i_1, ..., i_n\}$, $\{j_1, ..., j_m\}$ (with $n+m=q$) such that

$$a_{i_1}\alpha_{i_1}^x + ... + a_{i_n}\alpha_{i_n}^x = 0, \quad a_{j_1}\alpha_{j_1}^x + ... + a_{j_m}\alpha_{j_m}^x = 0.$$

There are $<2^{q-1}$ such partitions. But each partition yields nonzero f^*, f^{**} with

$$f^*(x) = f^{**}(x) = 0, \tag{6.5}$$

and with $f^*+f^{**}=f$, as well as $q(f^*), q(f^{**})<q=q(f)$ (where $q(f)$ is the number of nonzero summands of a function f). Write

$$\begin{aligned} f^*(x) &= f_1^*(x)+\ldots+f_g^*(x), \\ f^{**}(x) &= f_1^{**}(x)+\ldots+f_g^{**}(x), \end{aligned}$$

where f_i^*, f_i^{**} are linear combinations of $\alpha_{i1}^x, \ldots, \alpha_{i.q_i}^x$. By induction, all but at most $2G(q-1)$ solutions of (6.5) have

$$\begin{aligned} f_i^*(x) &= 0 \quad (1 \leqslant i \leqslant g), \\ f_i^{**}(x) &= 0 \quad (1 \leqslant i \leqslant g), \end{aligned}$$

hence (6.4). The number of exceptions to (6.4) therefore is

$$\begin{aligned} &< \exp(2(6q)^{4q}) + 2^q G(q-1) \\ &< \exp(2(6q)^{4q}) + 2^q \exp((7q)^{4q-4}) \\ &< \exp((7q)^{4q}) = G(q). \end{aligned} \qquad \square$$

A summand $a_i\alpha_i^x$ in (6.1) will be called a *singleton* if $\alpha_i \not\approx \alpha_j$ for every $j \neq i$, $1 \leqslant j \leqslant q$. Then one of the g summands in (6.2) equals just $a_i\alpha_i^x$, and hence has no zero. We therefore obtain the following

COROLLARY. *Suppose that f as given by* (6.1) *contains a singleton. Then $f(x)=0$ has at most $G(q)$ zeros $x \in \mathbb{Z}$.*

The α_{ij} $(1 \leqslant j \leqslant q_i)$ occurring in f_i are all $\approx$ to each other. However, given a solution x of $f_i(x)=0$, there may be a subsum of f_i which vanishes. We will refer to such a possible phenomenon as a *subsplitting*. It causes considerable complications in our proof of the Theorem; in particular, it necessitates the Appendix.

A solution x of $f_i(x)=0$ where no subsplitting occurs is called a *nondegenerate* solution. To ease notation, let us suppose that f itself as given by (6.1) has $\alpha_1 \approx \ldots \approx \alpha_q$. By Lemma 3, there are vectors $\mathbf{c}^{(w)}$ $(1 \leqslant w \leqslant B(q))$ such that for a nondegenerate solution, $(\alpha_1^x, \ldots, \alpha_q^x)$ is proportional to some $\mathbf{c}^{(w)}$.

7. Algebraic numbers having many conjugates which are $\approx$ to each other

Throughout, $\alpha, \beta, \gamma, \delta$ will be in $\overline{\mathbb{Q}}^{\times}$.

LEMMA 9. (i) $\approx$ *is an equivalence relation on* $\overline{\mathbb{Q}}^{\times}$.

(ii) *If* $\alpha\approx\beta$, $\gamma\approx\delta$, *then* $\alpha\gamma\approx\beta\delta$.

(iii) *If* $\alpha^l\approx\beta^l$ *for some* $l\in\mathbb{Z}\backslash\{0\}$, *then* $\alpha\approx\beta$.

(iv) *If* $\alpha\approx\beta$ *and* σ *is an embedding of* $\mathbb{Q}(\alpha,\beta)$ *into* $\overline{\mathbb{Q}}$, *then* $\sigma(\alpha)\approx\sigma(\beta)$.

Note that (i) has already been tacitly used above.

Proof. Let $T\subset\overline{\mathbb{Q}}^{\times}$ be the torsion subgroup, i.e., the group of roots of 1. Then $\alpha\approx\beta$ precisely when α,β have the same image in the factor group $\overline{\mathbb{Q}}^{\times}/T$. This implies (i), (ii). When $\xi^l\in T$ for some $l\in\mathbb{Z}\backslash\{0\}$, then $\xi\in T$; and this implies (iii). Finally, if $\xi\in T\cap\mathbb{Q}(\alpha,\beta)$ and σ is an embedding of $\mathbb{Q}(\alpha,\beta)$ into $\overline{\mathbb{Q}}$, then $\sigma(\xi)\in T$; and this yields (iv). $\square$

LEMMA 10. *Let* β *be of degree* d, *and* $S=\{\beta^{[1]},...,\beta^{[d]}\}$ *the set of its conjugates. Partition* S *as*

$$S=S_1\cup...\cup S_m$$

into equivalence classes under $\approx$. *Then*(4) $d=mn$ *with some* $n\in\mathbb{Z}$, *and*

$$|S_1|=...=|S_m|=n.$$

Proof. Let G be the Galois group of $K=\mathbb{Q}(\beta^{[1]},...,\beta^{[d]})$. When $\sigma\in G$, let $\sigma(S_i)$ be the set of elements $\sigma(\beta^{[a]})$ where $\beta^{[a]}$ runs through S_i. By (iv) of the preceding lemma, G permutes the sets $S_1,...,S_m$, i.e., G acts on the m-element set $\Sigma=\{S_1,...,S_m\}$. Since G acts transitively on S, it acts transitively on Σ. Given S_i, S_j and $\sigma\in G$ with $\sigma(S_i)=S_j$, we have $|S_i|=|\sigma(S_i)|=|S_j|$. Therefore $S_1,...,S_m$ have some common cardinality n, and $d=mn$. $\square$

Lehmer's conjecture says that if $\beta\not\approx 1$ is of degree d, then $h(\beta)\geqslant c_1/d$ with an absolute constant $c_1>0$. The best that is known in this direction is Dobrowolski's [1] estimate $h(\beta)\geqslant(c_2/d)(\log^+\log^+ d/\log^+ d)^3$, with the notation $\log^+\xi=\max(1,\log\xi)$. According to Voutier [11], we may take $c_2=\frac{1}{4}$. We will use the slightly weaker version

$$h(\beta)\geqslant\frac{1}{4d(\log^+ d)^3}. \tag{7.1}$$

The following lemma can sometimes be used in place of Lehmer's conjecture.

(4) The number n here and in $n(\beta), n_K(\beta)$ below should not be confused with the number n in the Proposition.

LEMMA 11. *Let β be as in Lemma* 10, *and suppose $\beta \not\approx 1$. Then*

$$h(\beta) \geqslant \frac{1}{4d(\log^+ m)^3}. \tag{7.2}$$

Proof. In the notation of Lemma 10, we may suppose that $\beta \in S_1$. Let γ_i $(i=1, ..., m)$ be the product of the elements of S_i, i.e.,

$$\gamma_i = \prod_{\beta^{[a]} \in S_i} \beta^{[a]}.$$

Then G permutes $\gamma_1, ..., \gamma_m$, so that every conjugate of γ_1 is among $\gamma_1, ..., \gamma_m$. We may infer that γ_1 is of degree $\leqslant m$. Moreover, $\gamma_1 \not\approx 1$, for otherwise $\beta^n \approx \gamma_1 \approx 1$, and hence $\beta \approx 1$, against the hypothesis. Therefore $h(\gamma_1) \geqslant 1/(4m(\log^+ m)^3)$. But

$$h(\gamma_1) \leqslant \sum_{\beta^{[a]} \in S_1} h(\beta^{[a]}) = |S_1| h(\beta) = nh(\beta).$$

We may conclude that

$$h(\beta) \geqslant \frac{h(\gamma_1)}{n} \geqslant \frac{1}{4d(\log^+ m)^3}. \qquad \square$$

Henceforth we will use the notation $n(\beta)=n$ where n is as in Lemma 10. Suppose that $\mathbb{Q}(\beta) \subset K$ where K is of degree D. Let $\xi \mapsto \xi^{(\sigma)}$ $(\sigma=1, ..., D)$ signify the embeddings $K \hookrightarrow \mathbb{C}$. Then each $\beta^{[a]}$ $(1 \leqslant a \leqslant d)$ occurs D/d times among $\beta^{(1)}, ..., \beta^{(D)}$. Therefore among $\beta^{(1)}, ..., \beta^{(D)}$, there are

$$n_K(\beta) := \frac{D}{d} n(\beta)$$

elements which are $\approx$ to each other. Note that $D = mn_K(\beta)$. We immediately get the following

COROLLARY. $h(\beta) \geqslant 1/(4d(\log^+(D/n_K(\beta)))^3)$.

Again let β be as in Lemma 10, and suppose $S_1 = \{\beta^{[1]}, ..., \beta^{[n]}\}$. So $\beta^{[1]}, ..., \beta^{[n]}$ have a common absolute value b, and we may write

$$\beta^{[i]} = b \cdot e^{2\pi i \varrho_i} \quad (i=1, ..., n) \tag{7.3}$$

with $0 \leqslant \varrho_i < 1$. The differences $\varrho_i - \varrho_j$ are rational, since $\beta^{[i]}/\beta^{[j]} \approx 1$.

More generally, let $R = \{\varrho_1, ..., \varrho_n\}$ be a system of reals such that each difference $\varrho_i - \varrho_j \in \mathbb{Q}$, but $\varrho_i - \varrho_j \notin \mathbb{Z}$ when $i \neq j$. Let r_{ij} be the denominator of $\varrho_i - \varrho_j$, i.e., the least natural number such that $r_{ij}(\varrho_i - \varrho_j) \in \mathbb{Z}$. Given $x \in \mathbb{N}$, let $u_i(x)$ be the number of j in $1 \leqslant j \leqslant n$ with $r_{ij} | x$. The system R will be called *homogeneous* if $u_1(x) = ... = u_n(x)$ for $x \in \mathbb{N}$.

LEMMA 12. *Let* $\{\beta^{[1]}, \ldots, \beta^{[n]}\}$ *be as above, and* $\varrho_1, \ldots, \varrho_n$ *defined by* (7.3). *Then* $R=\{\varrho_1, \ldots, \varrho_n\}$ *is homogeneous.*

Proof. Write $v_i(x)$ for the number of j in $1 \leqslant j \leqslant n$ with $r_{ij}=x$. Since $u_i(x)=\sum_{y|x} v_i(y)$, it will suffice to check that $v_1(x)=\ldots=v_n(x)$. Since $\beta^{[i]}/\beta^{[j]}=e^{2\pi i r'_{ij}/r_{ij}}$ with $\gcd(r_{ij}, r'_{ij})=1$, we have $r_{ij}=x$ precisely when $\beta^{[i]}/\beta^{[j]}$ is a primitive xth root of 1.

Given i and x set $v=v_i(x)$, and suppose that $\beta^{[i]}/\beta^{[l_k]}$ $(1 \leqslant k \leqslant v)$ is a primitive xth root of 1 for v distinct numbers $l_1, \ldots, l_v$ in $1 \leqslant l \leqslant n$.

Let G' be the subgroup of the Galois group G of $\mathbb{Q}(\beta^{[1]}, \ldots, \beta^{[d]})$ which permutes $\beta^{[1]}, \ldots, \beta^{[n]}$, i.e., which acts on $S_1=\{\beta^{[1]}, \ldots, \beta^{[n]}\}$. Since G acts transitively on S and permutes $S_1, \ldots, S_m$, the group G' acts transitively on S_1. Now let j in $1 \leqslant j \leqslant n$ be given, and pick $\sigma \in G'$ with $\sigma(\beta^{[i]})=\beta^{[j]}$. We have $\sigma(\beta^{[l_k]})=\beta^{[l'_k]}$ where $l'_1, \ldots, l'_v$ are v distinct integers in $1 \leqslant l' \leqslant n$. Further

$$\frac{\beta^{[j]}}{\beta^{[l'_k]}}=\sigma\left(\frac{\beta^{[i]}}{\beta^{[l_k]}}\right) \quad (1 \leqslant k \leqslant v)$$

are primitive xth roots of 1. Therefore $v_j(x) \geqslant v=v_i(x)$. By symmetry, $v_i(x)=v_j(x)$, and the lemma follows. □

When α, β, γ are in $\overline{\mathbb{Q}}^\times$ or more generally in $\mathbb{C}^\times$, write

$$G(\alpha : \beta : \gamma) \tag{7.4}$$

for the subgroup of $\mathbb{C}^\times$ generated by α/β and α/γ. Clearly $G(\alpha:\beta:\gamma)$ is finite precisely when $\alpha \approx \beta \approx \gamma$. With β and $S_1=\{\beta^{[1]}, \ldots, \beta^{[n]}\}$ as above, a triple of integers i, j, h in $1 \leqslant i, j, h \leqslant n$ will be said to be ε-*bad* if

$$|G(\beta^{[i]} : \beta^{[j]} : \beta^{[h]})| \leqslant \varepsilon n.$$

In the notation of (7.3), this happens precisely when

$$\operatorname{lcm}(r_{ij}, r_{ih}) \leqslant \varepsilon n.$$

Now let $l \geqslant 3$, and consider l-tuples of integers $u_1, \ldots, u_l$ in $1 \leqslant u \leqslant n$. Such an l-tuple will be called ε-*bad* if some triple u_i, u_j, u_h with distinct i, j, h in $1 \leqslant i, j, h \leqslant l$ is ε-*bad*, i.e., if it has

$$|G(\beta^{[u_i]} : \beta^{[u_j]} : \beta^{[u_h]})| \leqslant \varepsilon n.$$

Since $R=\{\varrho_1, \ldots, \varrho_n\}$ is homogeneous, and by the Corollary of the Appendix,

the number of ε-*bad* l-*tuples is* $< \varepsilon^{1/2} l^3 n^l$.

Suppose again that $\mathbb{Q}(\beta)\subset K$ and that $\xi\mapsto\xi^{(\sigma)}$ $(\sigma=1,\dots,D)$ signify the embeddings $K\hookrightarrow\mathbb{C}$. There are $n_K(\beta)=nD/d$ numbers μ in $1\leqslant\mu\leqslant D$ such that $\beta^{(\mu)}\in\{\beta^{[1]},\dots,\beta^{[n]}\}$. Let $\mathcal{M}$ be the set of these numbers. Given $l\geqslant 3$, an l-tuple $\mu_1,\dots,\mu_l$ of numbers in $\mathcal{M}$ will be called ε-**bad** if there are distinct numbers i,j,h in $1\leqslant i,j,h\leqslant l$ such that

$$|G(\beta^{(\mu_i)}:\beta^{(\mu_j)}:\beta^{(\mu_h)})|\leqslant\varepsilon n.$$

Since for each u in $1\leqslant u\leqslant n$ there are D/d numbers μ in $\mathcal{M}$ with $\beta^{(\mu)}=\beta^{[u]}$, and since $n_K(\beta)=nD/d$, we see that the number of ε-**bad** l-tuples of numbers in $\mathcal{M}$ is less than

$$\varepsilon^{1/2}l^3 n_K(\beta)^l. \tag{7.5}$$

Here $\mathcal{M}$ is typical of a subset of $\{1,\dots,D\}$ such that the numbers $\beta^{(\mu)}$ with $\mu\in\mathcal{M}$ make up an equivalence class under $\approx$. Any such set $\mathcal{M}$ has $|\mathcal{M}|=n_K(\beta)$. We have

LEMMA 13. *Let $\mathcal{M}\subset\{1,\dots,D\}$ be such that the numbers $\beta^{(\mu)}$ with $\mu\in\mathcal{M}$ make up an equivalence class under $\approx$ of the numbers $\beta^{(1)},\dots,\beta^{(D)}$. Then the number of ε-**bad** l-tuples $\mu_1,\dots,\mu_l$ with $\mu_i\in\mathcal{M}$ $(i=1,\dots,l)$ is less than* (7.5).

8. Two easy lemmas

Let K be a number field of degree D, and let $\xi\mapsto\xi^{(\sigma)}$ $(\sigma=1,\dots,D)$ signify the embeddings $K\hookrightarrow\mathbb{C}$. When $\mathbf{a}=(a_1,\dots,a_n)\in K^n$, set $\mathbf{a}^{(\sigma)}=(a_1^{(\sigma)},\dots,a_n^{(\sigma)})$ $(\sigma=1,\dots,D)$.

LEMMA 14. *Suppose that $\mathbf{a}\in K^n$. Then the vectors $\mathbf{a}^{(\sigma)}$ $(\sigma=1,\dots,D)$ span a rational subspace of K^n.*

Proof. This is well known. □

LEMMA 15. *Suppose that $\mathbf{a}\in K^n\subset\mathbb{C}^n$ but $\mathbf{a}\notin T$ where T is some subspace of $\mathbb{C}^n$. Then there are at least D/n integers σ in $1\leqslant\sigma\leqslant D$ with $\mathbf{a}^{(\sigma)}\notin T$.*

Proof. We will first suppose that $a_1,\dots,a_n$ are linearly independent over $\mathbb{Q}$. If the lemma were false for $\mathbf{a}$, there would be a set of more than $D-D/n$ vectors $\mathbf{a}^{\sigma}$ in T. Since $T\neq\mathbb{C}^n$, it will suffice to show that any set of more than $(1-1/n)D$ vectors $\mathbf{a}^{(\sigma)}$ spans $\mathbb{C}^n$.

So let $\mathcal{A}\subset\{1,\dots,D\}$ be given with $|\mathcal{A}|>(1-1/n)D$, and let $\mathcal{B}$ be the complement of $\mathcal{A}$, so that $|\mathcal{B}|<D/n$. Since $a_1,\dots,a_n$ are linearly independent over $\mathbb{Q}$, the vectors $\mathbf{a}^{(\sigma)}$ $(\sigma=1,\dots,D)$ span $\mathbb{C}^n$. We may suppose without loss of generality that $\mathbf{a}^{(1)},\dots,\mathbf{a}^{(n)}$ are linearly independent. Suppose that K is generated by α, i.e., $K=\mathbb{Q}(\alpha)$. Let G be the Galois group of its normal closure $\mathbb{Q}(\alpha^{(1)},\dots,\alpha^{(D)})$. For $g\in G$ we have $g(\alpha^{(\sigma)})=\alpha^{(\sigma_g)}$, where $1_g,\dots,D_g$ is a permutation of $1,\dots,D$. Given σ and τ, there is a $g\in G$ with

$\sigma_g=\tau$; in fact, the number of such $g\in G$ is $|G|/D$. Given σ, the number of $g\in G$ with $\sigma_g\in\mathcal{B}$ is $|G|\cdot|\mathcal{B}|/D$. The number of $g\in G$ such that at least one of $1_g,...,n_g$ is in $\mathcal{B}$ is $\leqslant|G|\cdot|\mathcal{B}|n/D<|G|$. Hence there is a $g\in G$ such that $1_g,...,n_g$ all lie in $\mathcal{A}$. Since $\mathbf{a}^{(1)},...,\mathbf{a}^{(n)}$ are independent and $g(\mathbf{a}^{(i)})=\mathbf{a}^{(i_g)}$ $(i=1,...,n)$ with $1_g,...,n_g$ in $\mathcal{A}$, the vectors $\mathbf{a}^{(\sigma)}$ with $\sigma\in\mathcal{A}$ indeed span $\mathbb{C}^n$.

This takes care of the case when $a_1,...,a_n$ are linearly independent over $\mathbb{Q}$. In general, we may suppose that $a_1,...,a_r$ are linearly independent over $\mathbb{Q}$, and that

$$a_j=\sum_{i=1}^{r}c_{ij}a_i \quad (r<j\leqslant n)$$

with rational coefficients c_{ij}. Since $\mathbf{a}\notin T$, there is a relation $\gamma_1x_1+...+\gamma_nx_n=0$ valid on T, such that $\gamma_1a_1+...+\gamma_na_n\neq0$. But then

$$\gamma_1'a_1+...+\gamma_r'a_r\neq0$$

with $\gamma_i'=\gamma_i+\sum_{j=r+1}^{n}c_{ij}\gamma_j$. Thus $\hat{\mathbf{a}}=(a_1,...,a_r)$ does not lie in the space $T'\subset\mathbb{C}^r$ defined by $\gamma_1'x_1+...+\gamma_r'x_r=0$. By the case of the lemma already shown, there are at least $D/r\geqslant D/n$ integers σ with $\hat{\mathbf{a}}^{(\sigma)}\notin T'$, so that $\gamma_1a_1^{(\sigma)}+...+\gamma_na_n^{(\sigma)}\neq0$, and hence $\mathbf{a}^{(\sigma)}\notin T$. $\square$

9. Nonvanishing of determinants

After the preliminary work of the preceding sections, we can finally commence with the proof of the Proposition. We first dispose of two simple cases.

(a) When $k=1$, $M_j(X)=b_jX$, and the linear independence condition means that $b_1,...,b_n$ are linearly independent over $\mathbb{Q}$. Then for any $\xi\neq0$, in particular for $\xi=\alpha_1^x$, the numbers $M_1(\xi)=b_1\xi,...,M_n(\xi)=b_n\xi$ are linearly independent over $\mathbb{Q}$.

(b) When $n=1$, $M_1(\mathbf{X})=a_1X_1+...+a_kX_k$ is not identically zero, and furthermore $M_1(\alpha_1^x,...,\alpha_k^x)=0$ becomes $a_1\alpha_1^x+...+a_k\alpha_k^x=0$. By Lemma 2, this equation has at most

$$A(k)\leqslant H(k,1)$$

solutions. We now put each solution into a class by itself. Hence in each class we may choose m arbitrarily large, in particular so large that some $h(\alpha_i^m:\alpha_j^m)\geqslant\hbar(k,1)$.

We may then suppose from now on that $k\geqslant2$, $n\geqslant2$. Again K will be a field containing $\alpha_1,...,\alpha_k$ and the coefficients of our linear forms. Again we set $D=\deg K$, and $\xi\mapsto\xi^{(\sigma)}$ $(\sigma=1,...,D)$ will signify the embeddings $K\hookrightarrow\mathbb{C}$. When $M_j(\mathbf{X})=a_{1j}X_1+...+a_{kj}X_k$, set $M_j^{(\sigma)}(\mathbf{X})=a_{1j}^{(\sigma)}X_1+...+a_{kj}^{(\sigma)}X_k$. We will write $\mathbf{a}_i=(a_{i1},...,a_{in})$ and $\mathbf{a}_i^{(\sigma)}=(a_{i1}^{(\sigma)},...,a_{in}^{(\sigma)})$. Now if the n numbers (5.1) are linearly dependent over $\mathbb{Q}$, we have

$$y_1M_1(\alpha_1^x,...,\alpha_k^x)+...+y_nM_n(\alpha_1^x,...,\alpha_k^x)=0$$

with $y_1, ..., y_n$ in $\mathbb{Q}$, not all zero. Then for $\sigma=1, ..., D$,

$$y_1 M_1^{(\sigma)}(\alpha_1^{(\sigma)x}, ..., \alpha_k^{(\sigma)x}) + ... + y_n M_n^{(\sigma)}(\alpha_1^{(\sigma)x}, ..., \alpha_k^{(\sigma)x}) = 0.$$

Therefore the matrix with rows

$$(M_1^{(\sigma)}(\alpha_1^{(\sigma)x}, ..., \alpha_k^{(\sigma)x}), ..., M_n^{(\sigma)}(\alpha_1^{(\sigma)x}, ..., \alpha_k^{(\sigma)x})) \quad (\sigma = 1, ..., D)$$

has rank $<n$. Let $\mathcal{D}(\sigma_1, ..., \sigma_n; x)$ be the determinant formed from the rows $\sigma_1, ..., \sigma_n$ of that matrix; then

$$\mathcal{D}(\sigma_1, ... \sigma_n; x) = 0. \tag{9.1}$$

LEMMA 16.

$$\mathcal{D}(\sigma_1, ..., \sigma_n; x) = \sum_{i_1=1}^{k} ... \sum_{i_n=1}^{k} \Delta\big(\mathbf{a}_{i_1}^{(\sigma_1)}, ..., \mathbf{a}_{i_n}^{(\sigma_n)}\big)\big(\alpha_{i_1}^{(\sigma_1)} ... \alpha_{i_n}^{(\sigma_n)}\big)^x, \tag{9.2}$$

where $\Delta(\mathbf{a}_{i_1}^{(\sigma_1)}, ..., \mathbf{a}_{i_n}^{(\sigma_n)})$ *is the determinant of the matrix with rows* $\mathbf{a}_{i_1}^{(\sigma_1)}, ..., \mathbf{a}_{i_n}^{(\sigma_n)}$. *When* $M_1, ..., M_n$ *are linearly independent over* $\mathbb{Q}$, *this determinant is* $\neq 0$ *for certain* $\sigma_1, ..., \sigma_n$ *and* $i_1, ..., i_n$.

Proof. Since $M_j^{(\sigma)}(\alpha_1^{(\sigma)x}, ..., \alpha_k^{(\sigma)x}) = a_{1j}^{(\sigma)}\alpha_1^{(\sigma)x} + ... + a_{kj}^{(\sigma)}\alpha_k^{(\sigma)x}$, we see that

$$\begin{aligned}
&\mathcal{D}(\sigma_1, ..., \sigma_n; x) \\
&= \begin{vmatrix} a_{11}^{(\sigma_1)}\alpha_1^{(\sigma_1)x} + ... + a_{k1}^{(\sigma_1)}\alpha_k^{(\sigma_1)x} & ... & a_{1n}^{(\sigma_1)}\alpha_1^{(\sigma_1)x} + ... + a_{kn}^{(\sigma_1)}\alpha_k^{(\sigma_1)x} \\ \vdots & & \vdots \\ a_{11}^{(\sigma_n)}\alpha_1^{(\sigma_n)x} + ... + a_{k1}^{(\sigma_n)}\alpha_k^{(\sigma_n)x} & ... & a_{1n}^{(\sigma_n)}\alpha_1^{(\sigma_n)x} + ... + a_{kn}^{(\sigma_n)}\alpha_k^{(\sigma_n)x} \end{vmatrix} \\
&= \sum_{\pi} \varepsilon_\pi (a_{1,\pi(1)}^{(\sigma_1)}\alpha_1^{(\sigma_1)x} + ... + a_{k,\pi(1)}\alpha_k^{(\sigma_1)x}) ... (a_{1,\pi(n)}^{(\sigma_n)}\alpha_1^{(\sigma_n)x} + ... + a_{k,\pi(n)}^{(\sigma_n)}\alpha_k^{(\sigma_n)x})
\end{aligned}$$

where π runs through the permutations of $1, ..., n$, and where ε_π is the sign of π. We obtain

$$\begin{aligned}
\mathcal{D}(\sigma_1, ..., \sigma_n; x) &= \sum_{i_1=1}^{k} ... \sum_{i_n=1}^{k} \alpha_{i_1}^{(\sigma_1)x} ... \alpha_{i_n}^{(\sigma_n)x} \sum_{\pi} \varepsilon_\pi a_{i_1,\pi(1)}^{(\sigma_1)} ... a_{i_n,\pi(n)}^{(\sigma_n)} \\
&= \sum_{i_1=1}^{k} ... \sum_{i_n=1}^{k} \Delta\big(\mathbf{a}_{i_1}^{(\sigma_1)}, ..., \mathbf{a}_{i_n}^{(\sigma_n)}\big)\big(\alpha_{i_1}^{(\sigma_1)} ... \alpha_{i_n}^{(\sigma_n)}\big)^x.
\end{aligned}$$

Given i, the vectors $\mathbf{a}_i^{(\sigma)}$ $(\sigma=1, ..., D)$ span a subspace S_i of $\mathbb{C}^n$ which is rational by Lemma 14. We claim that when $M_1, ..., M_n$ are linearly independent over $\mathbb{Q}$, then

$S_1+\ldots+S_k=\mathbb{C}^n$. For otherwise, there is a nontrivial relation $y_1X_1+\ldots+y_nX_n=0$ valid on $S_1+\ldots+S_k$, with coefficients $y_1, \ldots, y_n$ in $\mathbb{Q}$. Since $\mathbf{a}_i \in S_i$, we have $y_1a_{i1}+\ldots+y_na_{in}=0$ $(i=1, \ldots, k)$, which leads to $y_1M_1(\mathbf{X})+\ldots+y_nM_n(\mathbf{X})=0$, against our assumption. Now $\mathbb{C}^n=S_1+\ldots+S_k$ is spanned by the vectors $\mathbf{a}_i^{(\sigma)}$ $(i=1, \ldots, k; \sigma=1, \ldots, D)$, hence is spanned by certain vectors $\mathbf{a}_{i_1}^{(\sigma_1)}, \ldots, \mathbf{a}_{i_n}^{(\sigma_n)}$. But then $\Delta(\mathbf{a}_{i_1}^{(\sigma_1)}, \ldots, \mathbf{a}_{i_n}^{(\sigma_n)}) \neq 0$. □

Changing our notation, suppose that

$$\Delta(\mathbf{a}_{u_1}^{(\tau_1)}, \ldots, \mathbf{a}_{u_n}^{(\tau_n)}) \neq 0. \tag{9.3}$$

The n-tuple $u_1, \ldots, u_n$ will be fixed from now on. By relabeling embeddings, we may suppose that $\tau_1=1$. In view of (9.3), $\mathbf{a}_{u_2}^{(\tau_2)}$ does not lie in the space spanned by the vectors $\mathbf{a}_{u_1}^{(1)}, \mathbf{a}_{u_3}^{(\tau_3)}, \ldots, \mathbf{a}_{u_n}^{(\tau_n)}$. By Lemma 15, there is a subset $\mathcal{S}_2$ of $\{1, \ldots, D\}$ of cardinality $|\mathcal{S}_2| \geqslant D/n$ such that $\mathbf{a}_{u_2}^{(\sigma)}$ does not lie in this subspace when $\sigma \in \mathcal{S}_2$; thus

$$\Delta(\mathbf{a}_{u_1}^{(1)}, \mathbf{a}_{u_2}^{(\sigma)}, \mathbf{a}_{u_3}^{(\tau_3)}, \ldots, \mathbf{a}_{u_n}^{(\tau_n)}) \neq 0$$

when $\sigma \in \mathcal{S}_2$. When $n>2$, we continue as follows. Let $\sigma_2 \in \mathcal{S}_2$ be given. Then $\mathbf{a}_{u_3}^{(\tau_3)}$ does not lie in the space spanned by $\mathbf{a}_{u_1}^{(1)}, \mathbf{a}_{u_2}^{(\sigma_2)}, \mathbf{a}_{u_4}^{(\tau_4)}, \ldots, \mathbf{a}_{u_n}^{(\tau_n)}$. By Lemma 15, there is a set $\mathcal{S}_3(\sigma_2) \subset \{1, \ldots, D\}$ of cardinality $\geqslant D/n$ such that $\mathbf{a}_{u_3}^{(\sigma)}$ does not lie in this subspace when $\sigma \in \mathcal{S}_3(\sigma_2)$. Thus

$$\Delta(\mathbf{a}_{u_1}^{(1)}, \mathbf{a}_{u_2}^{(\sigma_2)}, \mathbf{a}_{u_3}^{(\sigma_3)}, \mathbf{a}_{u_4}^{(\tau_4)}, \ldots, \mathbf{a}_{u_n}^{(\tau_n)}) \neq 0$$

when $\sigma_2 \in \mathcal{S}_2$, $\sigma_3 \in \mathcal{S}_3(\sigma_2)$.

Continuing in this way, we inductively construct sets $\mathcal{S}_2, \mathcal{S}_3(\sigma_2), \ldots, \mathcal{S}_n(\sigma_2, \ldots, \sigma_{n-1})$ of cardinality at least D/n, such that $\mathcal{S}_j(\sigma_2, \ldots, \sigma_{j-1})$ is defined when

$$\sigma_2 \in \mathcal{S}_2, \quad \sigma_3 \in \mathcal{S}_3(\sigma_2), \quad \ldots, \quad \sigma_{j-1} \in \mathcal{S}_{j-1}(\sigma_2, \ldots, \sigma_{j-2}), \tag{9.4}$$

and such that

$$\Delta(\mathbf{a}_{u_1}^{(1)}, \mathbf{a}_{u_2}^{(\sigma_2)}, \ldots, \mathbf{a}_{u_n}^{(\sigma_n)}) \neq 0 \tag{9.5}$$

when

$$\sigma_2 \in \mathcal{S}_2, \quad \sigma_3 \in \mathcal{S}_3(\sigma_2), \quad \ldots, \quad \sigma_n \in \mathcal{S}_n(\sigma_2, \ldots, \sigma_{n-1}). \tag{9.6}$$

10. Selection of exponential equations

It will be convenient to set

$$\Delta\begin{pmatrix} \sigma_1, \ldots, \sigma_n \\ i_1, \ldots, i_n \end{pmatrix} = \Delta(\mathbf{a}_{i_1}^{(\sigma_1)}, \ldots, \mathbf{a}_{i_n}^{(\sigma_n)}),$$

$$\mathcal{A}\begin{pmatrix} \sigma_1, \ldots, \sigma_n \\ i_1, \ldots, i_n \end{pmatrix} = \alpha_{i_1}^{(\sigma_1)} \ldots \alpha_{i_n}^{(\sigma_n)},$$

and when $\boldsymbol{\sigma}=(\sigma_1, ..., \sigma_n)$,

$$f_{\boldsymbol{\sigma}}(x)=\sum_{i_1=1}^{k} \cdots \sum_{i_n=1}^{k} \Delta\begin{pmatrix}\sigma_1, ..., \sigma_n \\ i_1, ..., i_n\end{pmatrix}\left(\mathcal{A}\begin{pmatrix}\sigma_1, ..., \sigma_n \\ i_1, ..., i_n\end{pmatrix}\right)^{x}. \tag{10.1}$$

Then (9.1) becomes

$$f_{\boldsymbol{\sigma}}(x)=0. \tag{10.2}$$

Here $f_{\boldsymbol{\sigma}}$ is of the type f considered in §6. According to Lemma 8, the equation (10.2) will split, with up to $G(q)$ exceptions. The number $q=q(\boldsymbol{\sigma})$ of nonzero summands in (10.1) has $q\leqslant k^n$, so that splitting occurs with at most $G(k^n)$ exceptions.(5) In principle, we can do this for any $\boldsymbol{\sigma}$ with $1\leqslant\sigma_i\leqslant D$ $(i=1, ..., n)$, which should give us a lot of information. However, if we carried out this splitting for every n-tuple $\boldsymbol{\sigma}$, the number of exceptions would depend on a factor involving the degree D, which we have to avoid. We therefore have to select a small set of n-tuples $\boldsymbol{\sigma}$ for which we will study (10.2).

Let $\mathcal{S}$ be the set of n-tuples $\boldsymbol{\sigma}=(\sigma_1, \sigma_2, ..., \sigma_n)$ with $\sigma_1=1$, and with $\sigma_2, ..., \sigma_n$ satisfying (9.6). When $\boldsymbol{\sigma}\in\mathcal{S}$, the coefficient

$$\Delta\begin{pmatrix}\sigma_1, ..., \sigma_n \\ u_1, ..., u_n\end{pmatrix}$$

in (10.1) is nonzero, so that not all coefficients of $f_{\boldsymbol{\sigma}}$ vanish. We will restrict ourselves to $\boldsymbol{\sigma}\in\mathcal{S}$; but the set $\mathcal{S}$ is still too large and will have to be pared down.

As in (6.2), we may write $f_{\boldsymbol{\sigma}}=f_{\boldsymbol{\sigma}1}+...+f_{\boldsymbol{\sigma},g(\boldsymbol{\sigma})}$. Here we may suppose that $f_{\boldsymbol{\sigma}1}$ has the summand

$$\Delta\begin{pmatrix}\sigma_1, ..., \sigma_n \\ u_1, ..., u_n\end{pmatrix}\left(\mathcal{A}\begin{pmatrix}\sigma_1, ..., \sigma_n \\ u_1, ..., u_n\end{pmatrix}\right)^{x}. \tag{10.3}$$

Let $\mathcal{I}(\boldsymbol{\sigma})$ be the set of n-tuples $(i_1, ..., i_n)$ with

$$\Delta\begin{pmatrix}\sigma_1, ..., \sigma_n \\ i_1, ..., i_n\end{pmatrix}\neq 0, \tag{10.4}$$

$$\mathcal{A}\begin{pmatrix}\sigma_1, ..., \sigma_n \\ i_1, ..., i_n\end{pmatrix}\approx\mathcal{A}\begin{pmatrix}\sigma_1, ..., \sigma_n \\ u_1, ..., u_n\end{pmatrix}. \tag{10.5}$$

Clearly $(u_1, ..., u_n)\in\mathcal{I}(\boldsymbol{\sigma})$, and

$$f_{\boldsymbol{\sigma}1}(x)=\sum_{(i_1, ..., i_n)\in\mathcal{I}(\sigma)} \Delta\begin{pmatrix}\sigma_1, ..., \sigma_n \\ i_1, ..., i_n\end{pmatrix}\left(\mathcal{A}\begin{pmatrix}\sigma_1, ..., \sigma_n \\ i_1, ..., i_n\end{pmatrix}\right)^{x}. \tag{10.6}$$

(5) When the linear forms $M_1, ..., M_n$ come from a polynomial vector $\mathbf{P}$ with $t(\mathbf{P})=t$, an estimate $q\leqslant c^t$ with an absolute constant c may be shown to hold, enabling one to replace $\log t$ in (1.4) by a constant.

We will first deal with the case where $|\mathcal{I}(\boldsymbol{\sigma})|=1$ for some $\boldsymbol{\sigma}\in\mathcal{S}$. Then $f_{\boldsymbol{\sigma}1}$ equals (10.3), so that $f_{\boldsymbol{\sigma}}$ contains a singleton. It suffices in this case to restrict ourselves to (10.2) with this particular $\boldsymbol{\sigma}$. By the corollary to Lemma 8, (10.2) has at most

$$G(q(\boldsymbol{\sigma}))\leqslant G(k^n)\leqslant H(k,n)$$

solutions x. Here we put each solution in a class by itself. We can choose m so large that some $h(\alpha_i^m:\alpha_j^m)\geqslant\hbar(k,n)$.

We may then suppose from now on that $|\mathcal{I}(\boldsymbol{\sigma})|>1$ for each $\boldsymbol{\sigma}\in\mathcal{S}$. The number of n-tuples $(i_1,...,i_n)$ is k^n, and the number of sets of such n-tuples is 2^{k^n}. Therefore the number of possibilities for $\mathcal{I}(\boldsymbol{\sigma})$ is $<2^{k^n}$. Given $\sigma_1,\sigma_2,...,\sigma_{n-1}$ with $\sigma_1=1$ and $\sigma_2,...,\sigma_{n-1}$ satisfying (9.6), there is a set $\mathcal{I}(\sigma_1,\sigma_2,...,\sigma_{n-1})$ such that $\mathcal{I}(\sigma_1,...,\sigma_{n-1},\sigma_n)=\mathcal{I}(\sigma_1,...,\sigma_{n-1})$ when σ_n lies in a subset $\mathcal{S}_n'(\sigma_2,...,\sigma_{n-1})$ of $\mathcal{S}_n(\sigma_2,...,\sigma_{n-1})$ of cardinality

$$|\mathcal{S}_n'(\sigma_2,...,\sigma_{n-1})|>2^{-k^n}|\mathcal{S}_n(\sigma_2,...,\sigma_n)|\geqslant Dn^{-1}\cdot 2^{-k^n}.$$

Given $\sigma_1,\sigma_2,...,\sigma_{n-2}$ with $\sigma_1=1$ and $\sigma_2,...,\sigma_{n-2}$ satisfying (9.6), then there is a set $\mathcal{I}(\sigma_1,\sigma_2,...,\sigma_{n-2})$ such that $\mathcal{I}(\sigma_1,...,\sigma_{n-2},\sigma_{n-1})=\mathcal{I}(\sigma_1,...,\sigma_{n-2})$ when σ_{n-1} lies in a subset $\mathcal{S}_{n-1}'(\sigma_2,...,\sigma_{n-2})$ of $\mathcal{S}_{n-1}(\sigma_2,...,\sigma_{n-2})$ of cardinality $>D/(n\cdot 2^{k^n})$. After carrying out $n-1$ such steps, we obtain a set $\mathcal{I}$ of n-tuples $(i_1,...,i_n)$, as well as sets

$$\mathcal{S}_2',\quad \mathcal{S}_3'(\sigma_2),\quad ...,\quad \mathcal{S}_n'(\sigma_2,...,\sigma_{n-1}),\tag{10.7}$$

where $\mathcal{S}_j'(\sigma_2,...,\sigma_{j-1})$ is defined for

$$\sigma_2\in\mathcal{S}_2',\quad \sigma_3\in\mathcal{S}_3'(\sigma_2),\quad ...,\quad \sigma_{j-1}\in\mathcal{S}_{j-1}'(\sigma_2,...,\sigma_{j-2}).$$

Each of the sets (10.7) has cardinality

$$>\frac{D}{n\cdot 2^{k^n}}.\tag{10.8}$$

Further, when $\mathcal{S}'$ consists of $\boldsymbol{\sigma}$ with $\sigma_1=1$ and

$$\sigma_2\in\mathcal{S}_2',\quad \sigma_3\in\mathcal{S}_3'(\sigma_2),\quad ...,\quad \sigma_n\in\mathcal{S}_n'(\sigma_2,...,\sigma_{n-1}),$$

then

$$\mathcal{I}(\boldsymbol{\sigma})=\mathcal{I}\quad\text{when } \boldsymbol{\sigma}\in\mathcal{S}'.\tag{10.9}$$

For $2\leqslant j\leqslant n$, let $\mathcal{T}_j$ be the set of numbers $i_j\neq u_j$ in $1\leqslant i_j\leqslant k$ such that

$$(i_1,...,i_{j-1},i_j,u_{j+1},...,u_n)\in\mathcal{I}\tag{10.10}$$

for certain $i_1,...,i_{j-1}$. (When $j=n$, (10.10) becomes $(i_1,...,i_{n-1},i_n)\in\mathcal{I}$.)

LEMMA 17. *Suppose* $i_j \in \mathcal{T}_j$. *Then*

$$h\left(\frac{\alpha_{i_j}}{\alpha_{u_j}}\right) > \frac{1}{k^{6n}\deg(\alpha_{i_j}/\alpha_{u_j})}.$$

Proof. In view of (10.5), which holds for any n-tuple $(i_1, ..., i_n) \in \mathcal{I}$,

$$\mathcal{A}\begin{pmatrix} \sigma_1, ..., \sigma_j, \sigma_{j+1}, ..., \sigma_n \\ i_1, ..., i_j, u_{j+1}, ..., u_n \end{pmatrix} \approx \mathcal{A}\begin{pmatrix} \sigma_1, ..., \sigma_n \\ u_1, ..., u_n \end{pmatrix}$$

for $\boldsymbol{\sigma} \in \mathcal{S}'$. Thus

$$\alpha_{i_1}^{(\sigma_1)} ... \alpha_{i_j}^{(\sigma_j)} \alpha_{u_{j+1}}^{(\sigma_{j+1})} ... \alpha_{u_n}^{(\sigma_n)} \approx \alpha_{u_1}^{(\sigma_1)} ... \alpha_{u_j}^{(\sigma_j)} \alpha_{u_{j+1}}^{(\sigma_{j+1})} ... \alpha_{u_n}^{(\sigma_n)},$$

which is

$$\left(\frac{\alpha_{i_j}}{\alpha_{u_j}}\right)^{(\sigma_j)} \approx \left(\frac{\alpha_{u_1}}{\alpha_{i_1}}\right)^{(\sigma_1)} ... \left(\frac{\alpha_{u_{j-1}}}{\alpha_{i_{j-1}}}\right)^{(\sigma_{j-1})}. \tag{10.11}$$

This holds when $\sigma_1 = 1$, $\sigma_2 \in \mathcal{S}'_2$, ..., $\sigma_j \in \mathcal{S}'_j(\sigma_2, ..., \sigma_{j-1})$. Let such $\sigma_2, ..., \sigma_{j-1}$ be fixed, and let σ_j range through $\mathcal{S}'_j(\sigma_2, ..., \sigma_{j-1})$. Then the right-hand side of (10.11) is fixed, so that the number of $(\alpha_{i_j}/\alpha_{u_j})^{(\sigma)}$ $(\sigma = 1, ..., D)$ which are $\approx$ to each other is at least $|\mathcal{S}'_j(\sigma_2, ..., \sigma_{j-1})| > D/(n \cdot 2^{k^n})$. In other words, in the notation of §7,

$$n_K\left(\frac{\alpha_{i_j}}{\alpha_{u_j}}\right) > \frac{D}{n \cdot 2^{k^n}}. \tag{10.12}$$

Since $i_j \neq u_j$, and hence $\alpha_{i_j}/\alpha_{u_j} \not\approx 1$, the corollary to Lemma 11 yields

$$h\left(\frac{\alpha_{i_j}}{\alpha_{u_j}}\right) > \frac{1}{4(\log(n \cdot 2^{k^n}))^3 \deg(\alpha_{i_j}/\alpha_{u_j})} > \frac{1}{k^{6n}\deg(\alpha_{i_j}/\alpha_{u_j})},$$

on recalling that $k \geqslant 2$, $n \geqslant 2$. □

For $2 \leqslant j \leqslant n$, let $\mathcal{T}_j^*$ be the set of numbers $\alpha_{i_j}/\alpha_{u_j}$ with $i_j \in \mathcal{T}_j$. Say $\mathcal{T}_j^* = \{\beta_1, ..., \beta_r\}$. Clearly $r < k$; possibly $r = 0$, and $\mathcal{T}_j^*$ is empty. We had seen in (10.12) and Lemma 17 that

$$n_K(\beta_s) > \frac{D}{n \cdot 2^{k^n}} \quad (s = 1, ..., r) \tag{10.13}$$

and that

$$h(\beta_s) > \frac{1}{k^{6n}\deg \beta_s} \quad (s = 1, ..., r). \tag{10.14}$$

Set

$$l = 3k^n. \tag{10.15}$$

Recall the definition of the group $G(\alpha{:}\beta{:}\gamma)$ in §7.

LEMMA 18. *Suppose*

$$D > e^{4k^{2n}}. \tag{10.16}$$

Let $2 \leqslant j \leqslant n$, *and let* $\sigma_1, ..., \sigma_{j-1}$ *with* $\sigma_1 = 1$, $\sigma_2 \in \mathcal{S}'_2, ..., \sigma_{j-1} \in \mathcal{S}'_{j-1}(\sigma_2, ..., \sigma_{j-2})$ *be given. Then there is a subset* $\mathcal{S}''_j = \mathcal{S}''_j(\sigma_1, ..., \sigma_{j-1})$ *of* $\mathcal{S}'_j(\sigma_1, ..., \sigma_{j-1})$ *of cardinality*

$$|\mathcal{S}''_j(\sigma_1, ..., \sigma_{j-1})| = l,$$

such that

$$|G(\beta_s^{(\phi)} : \beta_s^{(\psi)} : \beta_s^{(\omega)})| > e^{-9k^{2n}} \deg \beta_s \quad (s = 1, ..., r) \tag{10.17}$$

for any triple of distinct numbers ϕ, ψ, ω *in* $\mathcal{S}''_j(\sigma_2, ..., \sigma_{j-1})$.

Proof. For brevity, write $\mathcal{S}'_j = \mathcal{S}'_j(\sigma_2, ..., \sigma_{j-1})$. When $r=0$, the condition (10.17) is vacuous. Since $\mathcal{S}'_j$ has cardinality $> D/(n \cdot 2^{k^n}) > 3k^n = l$ by (10.8), (10.16), it certainly contains a subset $\mathcal{S}''_j$ of cardinality l.

Now suppose that $r > 0$. Set

$$\varepsilon = e^{-8k^{2n}}. \tag{10.18}$$

Note that

$$2\varepsilon^{1/2} k l^3 (n \cdot 2^{k^n})^l < \varepsilon^{1/2} \cdot 54 k^{3n+1} e^{3k^{2n}} = 54 k^{3n+1} e^{-k^{2n}} < 1 \tag{10.19}$$

since $k \geqslant 2$, $n \geqslant 2$, and that

$$2l^2 (n \cdot 2^{k^n})^l < 18 k^{2n} e^{3k^{2n}} < D \tag{10.20}$$

by (10.16).

Let $\beta_s \in \mathcal{T}_j^*$ be given. We had seen in (10.11) that the numbers $\beta_s^{(\sigma)}$ with $\sigma \in \mathcal{S}'_j(\sigma_2, ..., \sigma_{j-1})$ were all $\approx$ to each other. So let $\mathcal{M}$ be the set of *all* the σ in $1 \leqslant \sigma \leqslant D$ for which $\beta_s^{(\sigma)}$ is $\approx$ to these numbers. By Lemma 13, the number of ε-**bad** l-tuples $\mu_1, ..., \mu_l$ in $\mathcal{M}$ is less than

$$\varepsilon^{1/2} l^3 n_K(\beta_s)^l \leqslant \varepsilon^{1/2} l^3 D^l. \tag{10.21}$$

In particular, the number of ε-**bad** l-tuples $\mu_1, ..., \mu_l$ with each μ_i in $\mathcal{S}'_j$ is less than (10.21). So far, $\beta_s \in \mathcal{T}_j^*$ was fixed. The number of l-tuples $\mu_1, ..., \mu_l$ in $\mathcal{S}'_j$ which are ε-**bad** for some β_s, $1 \leqslant s \leqslant r$, is

$$\leqslant r\varepsilon^{1/2} l^3 D^l < \varepsilon^{1/2} k l^3 D^l < \frac{1}{2} \left(\frac{D}{n \cdot 2^{k^n}} \right)^l \tag{10.22}$$

by (10.19). The number of l-tuples of numbers $\mu_1, ..., \mu_l$ in $\mathcal{S}'_j$ of which at least two numbers are equal is

$$\leqslant \binom{l}{2} D^{l-1} < l^2 D^{l-1} < \frac{1}{2} \left(\frac{D}{n \cdot 2^{k^n}} \right)^l \tag{10.23}$$

by (10.20). On the other hand, the number of all l-tuples $\mu_1, ..., \mu_l$ in $\mathcal{S}'_j$ is

$$|\mathcal{S}'_j(\sigma_2, ..., \sigma_{j-1})|^l \geqslant \left(\frac{D}{n\cdot 2^{k^n}}\right)^l.$$

Comparing this with (10.22), (10.23), we see that there is an l-tuple $\mu_1, ..., \mu_l$ of distinct numbers of $\mathcal{S}'_j$ which is not ε-**bad** for any of $\beta_1, ..., \beta_r$. By definition, this means that for any three distinct numbers i, j, h in $1\leqslant i, j, h\leqslant l$, we have

$$\begin{aligned} |G(\beta_s^{(\mu_i)}:\beta_s^{(\mu_j)}:\beta_s^{(\mu_h)})| &> \varepsilon n(\beta_s) = \varepsilon(\deg \beta_s) D^{-1} n_K(\beta_s) \\ &> \frac{\varepsilon \deg \beta_s}{n\cdot 2^{k^n}} > (\deg \beta_s) e^{-9k^{2n}} \end{aligned}$$

by (10.13), (10.18).

We now set $\mathcal{S}''_j(\sigma_2, ..., \sigma_{j-1}) = \{\mu_1, ..., \mu_l\}$. Then indeed for any three distinct numbers ϕ, ψ, ω in $\mathcal{S}''_j(...)$ we have (10.17). □

The condition (10.16) on D can always be achieved by enlarging the field K, if necessary. We will assume from now on that (10.16) holds.

Remark. Without (10.16) we might not produce an l-tuple $\mu_1, ..., \mu_l$ of *distinct* integers. This really would not make much difference. Note that if we enlarge K, there may be several embeddings $\sigma: K \hookrightarrow \mathbb{C}$ whose restrictions to the field generated by $\alpha_1, ..., \alpha_k$ and the coefficients of $M_1, ..., M_n$ are equal.

We now define $\mathcal{S}''$ to be the set of n-tuples $\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_n)$ with $\sigma_1 = 1$, $\sigma_2 \in \mathcal{S}''_2$, $\sigma_3 \in \mathcal{S}''_3(\sigma_2), ..., \sigma_n \in \mathcal{S}''_n(\sigma_1, ..., \sigma_{n-1})$. We will deal with the equations (10.2) where $\boldsymbol{\sigma} \in \mathcal{S}''$. The number of these equations is

$$|\mathcal{S}''| = l^{n-1} < 3^n \cdot k^{n^2}, \tag{10.24}$$

hence is bounded independently of D.

11. Conclusion

As noted above, each equation (10.2) splits, with at most $G(q) \leqslant G(k^n)$ exceptions. If we carry this out for each $\boldsymbol{\sigma} \in \mathcal{S}''$, we get

$$\leqslant |\mathcal{S}''| G(k^n) < 3^n k^{n^2} \exp((7k^n)^{4k^n}) < \exp((7k^n)^{5k^n}) \tag{11.1}$$

exceptions. We put each exceptional x into a class by itself. As we have noted before, we then can make m so large that $h(\alpha_i^m : \alpha_j^m) \geqslant \hbar(k, n)$.

For nonexceptional x, each equation (10.2) with $\boldsymbol{\sigma}\in\mathcal{S}''$ splits, so that x satisfies

$$f_{\boldsymbol{\sigma}1}(\mathbf{x})=0 \quad (\boldsymbol{\sigma}\in\mathcal{S}''). \tag{11.2}$$

We write this out in detail:

$$\sum_{(i_1,\dots,i_n)\in\mathcal{I}} \Delta\begin{pmatrix}\sigma_1,\dots,\sigma_n\\ i_1,\dots,i_n\end{pmatrix}\left(\mathcal{A}\begin{pmatrix}\sigma_1,\dots,\sigma_n\\ i_1,\dots,i_n\end{pmatrix}\right)^x=0. \tag{11.3}$$

Here in each summand we have (10.4), (10.5). One of the summands has $(i_1,\dots,i_n)=(u_1,\dots,u_n)$. A natural impulse would be to apply Lemma 3 to (11.3). But not so fast: x might be a degenerate solution of (11.3), i.e., the unpleasant phenomenon of subsplitting might occur.

Given $\boldsymbol{\sigma}\in\mathcal{S}''$ and given a solution x of (11.3), there will be a subset $\mathcal{I}(\boldsymbol{\sigma},x)\subset\mathcal{I}$ containing $(u_1,\dots,u_n)$ such that

$$\sum_{(i_1,\dots i_n)\in\mathcal{I}(\boldsymbol{\sigma},x)} \Delta\begin{pmatrix}\sigma_1,\dots,\sigma_n\\ i_1,\dots,i_n\end{pmatrix}\left(\mathcal{A}\begin{pmatrix}\sigma_1,\dots,\sigma_n\\ i_1,\dots,i_n\end{pmatrix}\right)^x=0, \tag{11.4}$$

but that splits no further, i.e., that no subsum of (11.4) vanishes. Since the coefficient

$$\Delta\begin{pmatrix}\sigma_1,\dots,\sigma_n\\ u_1,\dots,u_n\end{pmatrix}\neq 0$$

by (9.5), we have necessarily $|\mathcal{I}(\boldsymbol{\sigma},x)|>1$.

There are fewer than k^n tuples $\mathbf{i}=(i_1,\dots,i_n)\neq(u_1,\dots,u_n)$. Hence given $\sigma_1,\dots,\sigma_{n-1}$, there will be an n-tuple

$$\mathbf{i}=\mathbf{i}(\sigma_1,\dots\sigma_{n-1},x)\neq(u_1,\dots,u_n)$$

such that $\mathbf{i}\in\mathcal{I}(\boldsymbol{\sigma},x)$ for at least $l/k^n=3$ of the numbers $\sigma_n\in\mathcal{S}_n''(\sigma_2,\dots,\sigma_{n-1})$. Let $\mathcal{S}_n^*(\sigma_2,\dots,\sigma_{n-1},x)$ consist of 3 such numbers σ_n. Next, given $\sigma_1,\dots,\sigma_{n-2}$, there will be an n-tuple

$$\mathbf{i}(\sigma_1,\dots,\sigma_{n-2},x)$$

such that $\mathbf{i}(\sigma_1,\dots,\sigma_{n-2},\sigma_{n-1},x)=\mathbf{i}(\sigma_1,\dots,\sigma_{n-2},x)$ for at least 3 of the numbers σ_{n-1}. And so forth. We obtain n-tuples

$$\mathbf{i}(x),\quad \mathbf{i}(\sigma_2,x),\quad \dots,\quad \mathbf{i}(\sigma_2,\dots,\sigma_{n-1},x)$$

and 3-element sets

$$\mathcal{S}_2^*(x),\quad \mathcal{S}_3^*(\sigma_2,x),\quad \dots,\quad \mathcal{S}_n^*(\sigma_2,\dots,\sigma_{n-1},x)$$

with the following property. Let $\mathcal{S}^*(x)$ consist of $\boldsymbol{\sigma}=(\sigma_1,\ldots,\sigma_n)$ with

$$\sigma_1=1,\quad \sigma_2\in\mathcal{S}_2^*(x),\quad \sigma_3\in\mathcal{S}_3^*(\sigma_2,x),\quad \ldots,\quad \sigma_n\in\mathcal{S}_n^*(\sigma_2,\ldots,\sigma_{n-1},x).$$

Then

$$\mathbf{i}(x)\in\mathcal{I}(\boldsymbol{\sigma},x) \tag{11.5}$$

when $\boldsymbol{\sigma}\in\mathcal{S}^*(x)$.

Now let Σ be a system of 3-element sets $\mathcal{S}_2^*$, $\mathcal{S}_3^*(\sigma_2)$, ..., $\mathcal{S}_n^*(\sigma_1,\ldots,\sigma_{n-1})$, where $\mathcal{S}_j^*(\sigma_2,\ldots,\sigma_{j-1})$ is defined when $\sigma_2\in\mathcal{S}_2^*$, $\sigma_3\in\mathcal{S}_3^*(\sigma_2)$, ..., $\sigma_{j-1}\in\mathcal{S}_{j-1}^*(\sigma_2,\ldots,\sigma_{j-2})$, and where $\mathcal{S}_j^*(\sigma_2,\ldots,\sigma_{j-1})\subset\mathcal{S}_j''(\sigma_2,\ldots,\sigma_{j-1})$. The number of possible choices for $\mathcal{S}_2^*$ is $\leqslant l^3$. The number of choices for $\mathcal{S}_3^*(\sigma_2)$ is also $\leqslant l^3$, but carrying this out for each $\sigma_2\in\mathcal{S}_2^*$, we get $\leqslant l^{3\cdot 3}$ choices. The number of choices for all the sets $\mathcal{S}_4^*(\sigma_2,\sigma_3)$ with $\sigma_2\in\mathcal{S}_2^*$, $\sigma_3\in\mathcal{S}_3^*(\sigma_2)$, is $\leqslant l^{3\cdot3\cdot3}$, etc. Thus the number of possibilities for a system Σ is

$$\leqslant l^3\cdot l^{3\cdot 3}\cdot\ldots\cdot l^{3^{n-1}}<l^{3^n}.$$

When $\mathbf{i}$ is an n-tuple and Σ a system as above, let $C(\mathbf{i},\Sigma)$ be the class of solutions x with $\mathbf{i}(x)=\mathbf{i}$ and

$$\mathcal{S}_2^*(x)=\mathcal{S}_2^*,\quad \mathcal{S}_3^*(\sigma_2,x)=\mathcal{S}_3^*(\sigma_2),\quad \ldots,\quad \mathcal{S}_n^*(\sigma_2,\ldots,\sigma_{n-1},x)=\mathcal{S}_n^*(\sigma_2,\ldots,\sigma_{n-1})$$

whenever

$$\sigma_2\in\mathcal{S}_2^*,\quad \sigma_3\in\mathcal{S}_3^*(\sigma_2),\quad \ldots,\quad \sigma_n\in\mathcal{S}_n^*(\sigma_2,\ldots,\sigma_{n-1}). \tag{11.6}$$

The number of classes is less than

$$k^n\cdot l^{3^n}=k^n(3k^n)^{3^n}. \tag{11.7}$$

We will now study solutions in a given class $C(\mathbf{i},\Sigma)$. Let $j=j(\mathbf{i})$ be the number such that

$$\mathbf{i}=(i_1,\ldots,i_j,u_{j+1},\ldots,u_n)$$

and $i_j\neq u_j$. Possibly $i_n\neq u_n$, so that $j=n$. But we cannot have $j=1$, for then (10.5) would give

$$\mathcal{A}\begin{pmatrix}\sigma_1,\sigma_2,\ldots,\sigma_n\\ i_1,u_2,\ldots,u_n\end{pmatrix}\approx\mathcal{A}\begin{pmatrix}\sigma_1,\sigma_2,\ldots,\sigma_n\\ u_1,u_2,\ldots,u_n\end{pmatrix},$$

and hence $\alpha_{i_1}^{(\sigma_1)}\approx\alpha_{u_1}^{(\sigma_1)}$, which cannot happen when $i_1\neq u_1$. Therefore

$$2\leqslant j\leqslant n. \tag{11.8}$$

For $x\in C(\mathbf{i},\Sigma)$, and $\boldsymbol{\sigma}$ with $\sigma_1=1$ and (11.6), the equation (11.4) becomes

$$\Delta\binom{\sigma_1,\dots,\sigma_n}{u_1,\dots,u_n}\left(\mathcal{A}\binom{\sigma_1,\dots,\sigma_n}{u_1,\dots,u_n}\right)^x+\Delta\binom{\sigma_1,\dots,\sigma_n}{i_1,\dots,i_n}\left(\mathcal{A}\binom{\sigma_1,\dots,\sigma_n}{i_1,\dots,i_n}\right)^x \tag{11.9}$$
$$(+\text{possible further terms})=0.$$

We will now restrict $\boldsymbol{\sigma}$ with (11.6) even further. We fix $\sigma_1=1$, $\sigma_2,\dots,\sigma_{j-1}$ arbitrarily such that (11.6) holds, in so far as it applies to them. We let σ_j vary in $\mathcal{S}_j^*(\sigma_2,\dots,\sigma_{j-1})$, so that σ_j assumes three values ϕ,ψ,ω. Given a choice of σ_j, we again fix $\sigma_{j+1},\dots,\sigma_n$ such that (11.6) holds. Thus we now have three n-tuples $\boldsymbol{\sigma}$, which we will denote by $\boldsymbol{\sigma}_\phi,\boldsymbol{\sigma}_\psi,\boldsymbol{\sigma}_\omega$. We will study (11.4), which is the same as (11.9), for these three choices of $\boldsymbol{\sigma}$.

The number of possibilities for each of $\mathcal{I}(\boldsymbol{\sigma}_\phi,x)$, $\mathcal{I}(\boldsymbol{\sigma}_\psi,x)$, $\mathcal{I}(\boldsymbol{\sigma}_\omega,x)$ is $<2^{k^n}$. We subdivide the class $C(\mathbf{i},\Sigma)$ into

$$2^{3k^n} \tag{11.10}$$

subclasses $C(\mathbf{i},\Sigma,\mathcal{I}_\phi,\mathcal{I}_\psi,\mathcal{I}_\omega)$ such that $\mathcal{I}(\boldsymbol{\sigma}_\phi,x)=\mathcal{I}_\phi$, $\mathcal{I}(\boldsymbol{\sigma}_\psi,x)=\mathcal{I}_\psi$, $\mathcal{I}(\boldsymbol{\sigma}_\omega,x)=\mathcal{I}_\omega$ in the class. Let q_ϕ,q_ψ,q_ω be the number of nonzero summands in (11.9) with $\boldsymbol{\sigma}=\boldsymbol{\sigma}_\phi,\boldsymbol{\sigma}_\psi,\boldsymbol{\sigma}_\omega$, respectively. Each of these numbers is $\leqslant k^n$.

No subsum of (11.9) vanishes. Hence we may apply Lemma 3. Fix $\boldsymbol{\sigma}=\boldsymbol{\sigma}_\phi$ for the moment. Let $\mathcal{A}_{\boldsymbol{\sigma}}(x)$ be the vector in q_ϕ-space with components

$$\left(\mathcal{A}\binom{\sigma_1,\dots,\sigma_n}{i_1,\dots,i_n}\right)^x$$

where $(i_1,\dots,i_n)\in\mathcal{I}_\phi$. According to Lemma 3, there are vectors $\mathbf{c}_{\boldsymbol{\sigma}}^{(w)}$ $(1\leqslant w\leqslant B(q_\phi))$ such that $\mathcal{A}_{\boldsymbol{\sigma}}(x)$ is proportional to some $\mathbf{c}_{\boldsymbol{\sigma}}^{(w)}$ for every solution x. We subdivide $\mathbf{C}(\mathbf{i},\Sigma,\mathcal{I}_\phi,\mathcal{I}_\psi,\mathcal{I}_\omega)$ according to the $\mathbf{c}_{\boldsymbol{\sigma}}^{(w)}$ $(w=1,\dots,B(q_\phi))$ to which $\mathcal{A}_{\boldsymbol{\sigma}}(x)$ is proportional. In fact we do this for $\boldsymbol{\sigma}_\phi$ as well as for $\boldsymbol{\sigma}_\psi,\boldsymbol{\sigma}_\omega$, so that we divide into

$$\leqslant B(q_\phi)B(q_\psi)B(q_\omega)\leqslant B(k^n)^3$$

subclasses. Thus altogether, by (11.7), (11.10), the number of subclasses (which we will call "classes" from now on) is

$$\begin{aligned}&<k^n(3k^n)^{3^n}\cdot 2^{3k^n}B(k^n)^3<(3k^n)^{3^n+1}\cdot 2^{3k^n}\cdot(k^n)^{9k^{2n}}\\&<k^{12nk^{2n}}<\exp(12nk^{2n+1}).\end{aligned} \tag{11.11}$$

Considering only the two components of $\mathcal{A}_{\boldsymbol{\sigma}}(\mathbf{x})$ highlighted in (11.9), we get

$$\left(\mathcal{A}\binom{\sigma_1,\dots,\sigma_n}{i_1,\dots,i_n}\right)^x=c_\phi\left(\mathcal{A}\binom{\sigma_1,\dots,\sigma_n}{u_1,\dots,u_n}\right)^x$$

when $\boldsymbol{\sigma}=\boldsymbol{\sigma}_\phi$, where $c_\phi = c_\phi^{(w)} = c_{\boldsymbol{\sigma}}^{(w)}(i_1,\dots,i_n)/c_{\boldsymbol{\sigma}}^{(w)}(u_1,\dots,u_n)$ in a given class is fixed. By our definition of j, this gives

$$(\alpha_{i_1}^{(\sigma_1)}\dots\alpha_{i_j}^{(\sigma_j)})^x = c_\phi(\alpha_{u_1}^{(\sigma_1)}\dots\alpha_{u_j}^{(\sigma_j)})^x$$

for $\boldsymbol{\sigma}=\boldsymbol{\sigma}_\phi$, or

$$\left(\left(\frac{\alpha_{i_j}}{\alpha_{u_j}}\right)^{(\phi)}\right)^x = c_\phi\left(\left(\frac{\alpha_{u_1}}{\alpha_{i_1}}\right)^{(\sigma_1)}\dots\left(\frac{\alpha_{u_{j-1}}}{\alpha_{i_{j-1}}}\right)^{(\sigma_{j-1})}\right)^x.$$

An analogous relation holds when $\boldsymbol{\sigma}=\boldsymbol{\sigma}_\psi$ or $\boldsymbol{\sigma}=\boldsymbol{\sigma}_\omega$. Taking quotients we obtain

$$\left(\frac{(\alpha_{i_j}/\alpha_{u_j})^{(\phi)}}{(\alpha_{i_j}/\alpha_{u_j})^{(\psi)}}\right)^x = \frac{c_\phi}{c_\psi}.$$

Now $\alpha_{i_j}/\alpha_{u_j}$ is one of the numbers β_s in $\mathcal{T}_j^*$, so that we may write

$$\left(\frac{\beta_s^{(\phi)}}{\beta_s^{(\psi)}}\right)^x = \frac{c_\phi}{c_\psi}.$$

Similarly $\left(\beta_s^{(\phi)}/\beta_s^{(\omega)}\right)^x = c_\phi/c_\omega$. Hence if x, x' lie in our class, then

$$\left(\frac{\beta_s^{(\phi)}}{\beta_s^{(\psi)}}\right)^{x-x'} = \left(\frac{\beta_s^{(\phi)}}{\beta_s^{(\omega)}}\right)^{x-x'} = 1.$$

So if $|G(\beta_s^{(\phi)}:\beta_s^{(\psi)}:\beta_s^{(\omega)})| = m$, we obtain $x\equiv x' \pmod m$. On the other hand, $m > e^{-9k^{2n}}\deg\beta_s$ by (10.17), so that in view of (10.14),

$$h(\beta_s^m) = mh(\beta_s) > \frac{e^{-9k^{2n}}}{k^{6n}} > e^{-10k^{2n}} = \hbar(k,n).$$

But β_s is the quotient of two numbers α_i/α_j, so that $h(\alpha_i^m:\alpha_j^m) > \hbar(k,n)$.

So how many classes do we have? Adding (11.1) to (11.11) we obtain indeed at most

$$\exp((7k^n)^{6k^n}) = H(k,n)$$

classes. □

Appendix: Denominators of certain rational numbers

Consider a system $R=\{\varrho_1,\dots,\varrho_n\}$ of real numbers whose differences $\varrho_i-\varrho_j$ lie in $\mathbb{Q}$, but not in $\mathbb{Z}$ when $i\neq j$. We will briefly refer to such a set of reals as a *system*. Let r_{ij} be the denominator of $\varrho_i-\varrho_j$, so that $\varrho_i-\varrho_j = r'_{ij}/r_{ij}$ with $r_{ij}>0$ and $\gcd(r_{ij}, r'_{ij})=1$. In particular, $r_{ii}=1$ $(1\leqslant i\leqslant n)$. We would like most of these denominators to have order of magnitude at least n. Given $\varepsilon>0$, let $N_0(\varepsilon)$ be the number of pairs i,j in $1\leqslant i,j\leqslant n$ with

$$r_{ij}\leqslant \varepsilon n.$$

Is there a function $\delta(\varepsilon)$ (independent of n and of R) which tends to 0 as $\varepsilon\to 0$, such that

$$N_0(\varepsilon)\leqslant \delta(\varepsilon)n^2? \tag{A1}$$

The answer to this question is negative: Let $R=\{0,1/n,...,(n-1)/n\}$. In this case $N_0(\varepsilon)=nN'(\varepsilon)$ where $N'(\varepsilon)$ is the number of integers i, $1\leqslant i\leqslant n$, with $\gcd(i,n)\geqslant 1/\varepsilon$. Now $N'(\varepsilon)=n-N''(\varepsilon)$ where $N''(\varepsilon)$ counts the number of integers i, $1\leqslant i\leqslant n$, with $\gcd(i,n)<1/\varepsilon$. Clearly

$$N''(\varepsilon)\leqslant n\prod_{\substack{p\mid n\\ p\geqslant 1/\varepsilon}}(1-p^{-1}).$$

Hence let $n=n_m$ be the product of the primes p in $1/\varepsilon\leqslant p\leqslant m$. Then when $m\geqslant m_0(\varepsilon)$, we have $N''(\varepsilon)<\frac{1}{2}n$, and hence $N'(\varepsilon)>\frac{1}{2}n$, $N_0(\varepsilon)>\frac{1}{2}n^2$, which is inconsistent with (A1).

Not to give up, we write $N(\varepsilon)$ for the number of triples i,j,k in $1\leqslant i,j,k\leqslant n$ with

$$\operatorname{lcm}(r_{ij},r_{ik})\leqslant \varepsilon n. \tag{A2}$$

I conjecture that there is a function $\delta(\varepsilon)$ (independent of n and R) which tends to 0 as $\varepsilon\to 0$, such that

$$N(\varepsilon)\leqslant \delta(\varepsilon)n^3. \tag{A3}$$

I cannot prove this conjecture, unless we make an extra assumption on the system R. Given $x\in\mathbb{N}$ and $1\leqslant i\leqslant n$, write $u_i(x)$ for the number of integers j, $1\leqslant j\leqslant n$, with

$$r_{ij}\mid x.$$

We call R *homogeneous* if for every x, the number $u_i(x)$ is independent of i; say $u_i(x)=u(x)$. For example, $R=\{0,1/n,...,(n-1)/n\}$ is homogeneous. Another example is the system R_m consisting of the numbers i/m with $1\leqslant i\leqslant m$, $\gcd(i,m)=1$, so that R_m has cardinality $n=\phi(m)$: for in this case, if i/m and i'/m are in R_m, say $i'\equiv ki \pmod m$, then $r_{ij}=r_{i'j'}$ where j' is the integer in $1\leqslant j'\leqslant m$ with $j'\equiv kj \pmod m$. Occasionally we will write $u_i^R(x)$ and $u^R(x)$ to indicate that our functions come from a particular system R.

THEOREM A. *Suppose $0<\varkappa<1$. Then when R is homogeneous we have*

$$N(\varepsilon)\leqslant\zeta(2-\varkappa)\varepsilon^{\varkappa}n^3. \tag{A4}$$

Thus in this case we may take $\delta(\varepsilon)=\zeta(2-\varkappa)\varepsilon^{\varkappa}$ in (A3). It may be shown that when the denominators r_{ij} are all powers of a fixed prime p, and when R is homogeneous, then $N(\varepsilon)\leqslant\varepsilon^2n^3$. Therefore (A4) may perhaps be replaced by $N(\varepsilon)\leqslant c_0(\varkappa)\varepsilon^{\varkappa}n^3$ for $0<\varkappa<2$, and possibly even for $\varkappa=2$. To carry out the proof of Theorem A, we will need a somewhat more general theorem. Let $R=\{\varrho_1,...,\varrho_n\}$, $S=\{\sigma_1,...,\sigma_n\}$ be homogeneous systems. We will call R,S *isomorphic*, and write $R\sim S$, if $u^R(x)=u^S(x)$.

THEOREM B. *Let $R=\{\varrho_1,...,\varrho_n\}$, $S=\{\sigma_1,...,\sigma_n\}$, $T=\{\tau_1,...,\tau_n\}$ be homogeneous, and isomorphic to each other. Suppose that all the differences $\varrho_i-\sigma_j$, $\varrho_i-\tau_k$ lie in $\mathbb{Q}$. Let a_{ij} be the denominator of $\varrho_i-\sigma_j$, let b_{ik} be the denominator of $\varrho_i-\tau_k$, and $N(\varepsilon)$ the number of triples i,j,k in $1\leqslant i,j,k\leqslant n$ with*

$$\operatorname{lcm}(a_{ij},b_{ik})\leqslant\varepsilon n. \tag{A5}$$

Then (A4) holds for $\varkappa$ in $0<\varkappa<1$.

The proof of Theorem B will proceed via a series of lemmas. Let $R=\{\varrho_1,...,\varrho_n\}$ be a homogeneous system. Note that $r_{ij}\in\mathbb{N}$ is least such that $r_{ij}(\varrho_i-\varrho_j)\in\mathbb{Z}$. When $x\in\mathbb{N}$, write $\varrho_i\overset{x}{\equiv}\varrho_j$ if $r_{ij}\mid x$, i.e., if $x(\varrho_i-\varrho_j)\in\mathbb{Z}$. Clearly $\overset{x}{\equiv}$ defines an equivalence relation among the elements of R. Thus R splits into equivalence classes, where each class contains $u(x)$ elements, and the number of classes is $v(x):=n/u(x)$.

When $R=\{\varrho_1,...,\varrho_n\}$, $S=\{\sigma_1,...,\sigma_n\}$, write $R\overset{x}{\equiv}S$ if $x(\varrho_i-\sigma_j)\in\mathbb{Z}$ for $1\leqslant i,j\leqslant n$. The relation $\overset{x}{\equiv}$ for systems is symmetric and transitive but not reflexive, since not necessarily $R\overset{x}{\equiv}R$. But when $R\overset{x}{\equiv}S$, then $R\overset{x}{\equiv}S\overset{x}{\equiv}R$, hence $R\overset{x}{\equiv}R$. When $R\overset{x}{\equiv}R$, then $\varrho_i=\varrho_1+(a_i/x)$ with $a_i\in\mathbb{Z}$, but when $i\neq j$ we have $\varrho_i-\varrho_j=(a_i-a_j)/x\notin\mathbb{Z}$, so that $a_i\not\equiv a_j \pmod x$, and therefore R has cardinality $|R|\leqslant x$.

LEMMA A. *Let R be homogeneous, $x\in\mathbb{N}$, and let $R_1,...,R_v$ be the equivalence classes with respect to $\overset{x}{\equiv}$. Then $R_r\overset{x}{\equiv}R_r$ $(1\leqslant r\leqslant v)$, but $R_r\overset{x}{\not\equiv}R_s$ when $r\neq s$. The systems $R_1,...,R_v$ are homogeneous and isomorphic to each other. When $R\overset{m}{\equiv}R$ and $x\mid m$, then $v\leqslant m/x$.*

Furthermore, if S is homogeneous and isomorphic to R, with equivalence classes $S_1,...,S_v$, then $R_1\sim...\sim R_v\sim S_1\sim...\sim S_v$. Given $1\leqslant r\leqslant v$, there is at most one s with $R_r\overset{x}{\equiv}S_s$.

Proof. We have $x(\varrho_i-\varrho_j)\in\mathbb{Z}_r$ when $\varrho_i,\varrho_j\in R_r$, and therefore $R_r\overset{x}{\equiv}R_r$. But when $\varrho_i\in R_r$, $\varrho_j\in R_s$ with $r\neq s$, then $\varrho_i\overset{x}{\not\equiv}\varrho_j$, and hence $R_r\overset{x}{\not\equiv}R_s$. Now suppose that ϱ_i,ϱ_j

are in R_r. Then $\varrho_i \overset{y}{\equiv} \varrho_j$ when $r_{ij} | y$. But since $r_{ij} | x$, this holds precisely if $r_{ij} | y'$ where $y' = \gcd(y, x)$. Conversely, if $\varrho_i \in R_r$, $\varrho_j \in R$ and $r_{ij} | y'$, then $r_{ij} | y$ and $r_{ij} | x$, hence $\varrho_j \in R_r$. Therefore

$$u_i^{R_r}(y) = u_i^{R_r}(y') = u_i^R(y') = u^R(y').$$

We may conclude that R_r is homogeneous with $u^{R_r}(y) = u^R(y')$. Therefore $R_1 \sim \ldots \sim R_v$. When $R \overset{m}{\equiv} R$ and $x | m$, each $\varrho_i = \varrho_1 + a_i/m$ with $a_i \in \mathbb{Z}$. Now if $\varrho_i \in R_r$, $\varrho_j \in R_s$ with $r \neq s$, then $x(\varrho_i - \varrho_j) = x(a_i - a_j)/m \notin \mathbb{Z}$, so that $a_i \not\equiv a_j \pmod{m/x}$. This shows that the number v of classes $R_1, \ldots, R_v$ has $v \leqslant m/x$.

When S is homogeneous with $R \sim S$, each equivalence class $S_1, \ldots, S_v$ is homogeneous, and $u_i^{S_s}(y) = u^S(y') = u^R(y')$, so that indeed $S_1, \ldots, S_v$ are isomorphic to $R_1, \ldots, R_v$. When $R_r \overset{x}{\equiv} S_s$ and $R_r \overset{x}{\equiv} S_t$, then $S_s \overset{x}{\equiv} S_t$, so that $s = t$. □

Write $c(\varkappa, p) = (1 - p^{\varkappa - 2})^{-1}$,

$$c(\varkappa, m) = \prod_{p | m} c(\varkappa, p).$$

LEMMA B. *Suppose that we are in the situation of Theorem B, and that*

$$R \overset{m}{\equiv} S \overset{m}{\equiv} T. \tag{A6}$$

Then

$$N(\varepsilon) \leqslant c(\varkappa, m) \varepsilon^{\varkappa} n^3.$$

Since for any systems R, S, T as in Theorem B there is an $m \in \mathbb{N}$ with (A6), and since $c(\varkappa, m) < \zeta(2 - \varkappa)$, this lemma implies Theorem B.

Proof. When $m = 1$, we have $\varrho_i - \varrho_j \in \mathbb{Z}$ for $1 \leqslant i, j \leqslant n$, but $\varrho_i - \varrho_j \notin \mathbb{Z}$ for $i \neq j$, and therefore $n = 1$. Then (A5) cannot hold unless $\varepsilon \geqslant 1$; but then $N(\varepsilon) = 1 \leqslant \varepsilon^{\varkappa} = c(\varkappa, 1) \varepsilon^{\varkappa} \cdot 1^3$.

It will therefore suffice to prove the lemma for

$$m = p^l m_0$$

where p is a prime, $p \nmid m_0$, $l > 0$, assuming that the lemma is true for m_0.

We may apply a common translation to R, S, T. Hence we may suppose that all the elements of R, S, T lie in $m^{-1}\mathbb{Z}$. Set

$$x_q = m_0 p^{l-q} = m p^{-q} \quad (0 \leqslant q \leqslant l).$$

Let $R_1, \ldots, R_{v_1}$ be the equivalence classes of R with respect to $\overset{x_1}{\equiv}$. Thus $v_1 = v(x_1)$, and each R_r $(1 \leqslant r \leqslant v_1)$ has $u(x_1) = n/v_1$ elements. By Lemma A, we have $v_1 \leqslant m/x_1 = p$.

Given a class R_r, we split it into subclasses $R_{r,1}, \ldots, R_{r,v_2}$ with respect to $\overset{x_2}{\equiv}$. Since $R_r \overset{x_1}{\equiv} R_r$, we have $v_2 \leqslant x_1/x_2 = p$. Moreover, since $R_r \sim R_{r'}$ $(1 \leqslant r, r' \leqslant v_1)$, the number v_2 is by Lemma A independent of r in $1 \leqslant r \leqslant v_1$. Note that R splits into the classes R_{r_1, r_2} $(1 \leqslant r_1 \leqslant v_1,\ 1 \leqslant r_2 \leqslant v_2)$ with respect to $\overset{x_2}{\equiv}$, and these $v_1 v_2$ systems R_{r_1, r_2} are isomorphic to each other. Suppose now that $1 < q \leqslant l$, and that we have defined systems $R_{r_1, \ldots, r_{q-1}}$ for $1 \leqslant r_i \leqslant v_i$ $(i = 1, \ldots, q-1)$, these being the equivalence classes of R with respect to $\overset{x_{q-1}}{\equiv}$. A system $R_{r_1, \ldots, r_{q-1}}$ splits into classes $R_{r_1, \ldots, r_{q-1}, r_q}$ $(1 \leqslant r_q \leqslant v_q)$ with respect to $\overset{x_q}{\equiv}$. Here $v_q \leqslant p$, and v_q is independent of $r_1, \ldots, r_{q-1}$. The $v_1 \ldots v_q$ systems $R_{r_1, \ldots, r_q}$ are all isomorphic to each other, and they are the equivalence classes of R with respect to $\overset{x_q}{\equiv}$, so that $v(x_q) = v_1 \ldots v_q$. Each $R_{r_1, \ldots, r_q}$ contains $n/(v_1 \ldots v_q)$ elements. In this way, we eventually construct systems $R_{r_1, \ldots, r_q}$ for $0 < q \leqslant l$ and $1 \leqslant r_i \leqslant v_i$ $(i = 1, \ldots, q)$. When $q = 0$, a notation $R_{r_1, \ldots, r_q}$ will simply mean R.

In complete analogy, we construct systems $S_{s_1, \ldots, s_q}$ and $T_{t_1, \ldots, t_q}$, where again $1 \leqslant s_i \leqslant v_i$ and $1 \leqslant t_i \leqslant v_i$ $(i = 1, \ldots, q)$, with the numbers $v_1, \ldots, v_q$ the same as above by Lemma A, and since $R \sim S \sim T$. Furthermore

$$R_{r_1, \ldots, r_q} \sim S_{s_1, \ldots, s_q} \sim T_{t_1, \ldots, t_q}$$

for any $r_1, \ldots, r_q, s_1, \ldots, s_q, t_1, \ldots, t_q$ under consideration.

If we have

$$R_{r_1, \ldots, r_q} \overset{x_q}{\equiv} S_{s_1, \ldots, s_q} \overset{x_q}{\equiv} T_{t_1, \ldots, t_q} \tag{A7}$$

for some $1 \leqslant q \leqslant l$ and $r_1, \ldots, r_q, s_1, \ldots, s_q, t_1, \ldots, t_q$, then

$$R_{r_1, \ldots, r_{q-1}} \overset{x_{q-1}}{\equiv} S_{s_1, \ldots, s_{q-1}} \overset{x_{q-1}}{\equiv} T_{t_1, \ldots, t_{q-1}}. \tag{A8}$$

When $q = 1$, then (A8) is to be interpreted as $R \overset{x_0}{\equiv} S \overset{x_0}{\equiv} T$, which is certainly true by (A6) and since $x_0 = m$. On the other hand, when (A8) holds, then by Lemma A the number of triples r_q, s_q, t_q with (A7) is $\leqslant v_q$ (since there are v_q choices for r_q). Write w_1 for the number of triples r_1, s_1, t_1 such that (A7) holds for $q = 1$. By what we have just said, $w_1 \leqslant v_1$. Suppose that $w_1, \ldots, w_{q-1}$ have been defined such that the number of $3(q-1)$-tuples $r_1, \ldots, r_{q-1}, s_1, \ldots, s_{q-1}, t_1, \ldots, t_{q-1}$ with (A8) equals $w_1 \ldots w_{q-1}$. Then let w_q be a number such that the number of $3q$-tuples $r_1, \ldots, r_q, s_1, \ldots, s_q, t_1, \ldots, t_q$ with (A7) equals $w_1 \ldots w_{q-1} w_q$. In particular, when $w_1 \ldots w_{q-1} = 0$, then (A8) never holds, hence (A7) never holds, and we set $w_q = 0$. In this way w_q is uniquely defined, and $0 \leqslant w_q \leqslant v_q$ for $q = 1, \ldots, l$.

For convenience we will write $\mathbf{r} = (r_1, \ldots, r_l)$, $\mathbf{s} = (s_1, \ldots, s_l)$, $\mathbf{t} = (t_1, \ldots, t_l)$. There are $(v_1 v_2 \ldots v_l)^3$ triples $\mathbf{r}, \mathbf{s}, \mathbf{t}$. For $0 \leqslant q \leqslant l$, let $\mathcal{C}_q$ be the set of triples $\mathbf{r}, \mathbf{s}, \mathbf{t}$ for which q is the largest integer in $0 \leqslant q \leqslant l$ for which (A7) holds. In particular, $\mathcal{C}_0$ consists of triples where (A7) does not hold for $q = 1$.

The number of $3q$-tuples $r_1, ..., r_q, ..., t_1, ..., t_q$ with (A7) is $w_1 \dots w_q$. Therefore $\mathcal{C}_l$ has cardinality

$$|\mathcal{C}_l| = w_1 \dots w_l. \tag{A9}$$

When $q<l$, the number of $3(q+1)$-tuples $r_1, ..., r_q, r_{q+1}, ..., t_1, ..., t_q, t_{q+1}$ with (A7) equals $w_1 \dots w_q v_{q+1}^3$. On the other hand, the number of $3(q+1)$-tuples where (A7) holds with $q+1$ in place of q is $w_1 \dots w_q w_{q+1}$. Therefore the number of $(q+1)$-tuples with (A7), but not (A7) with $q+1$ in place of q, is $w_1 \dots w_q (v_{q+1}^3 - w_{q+1})$. Given such a $3(q+1)$-tuple, the number of choices for $r_{q+2}, ..., r_l, ..., t_{q+2}, ..., t_l$ is $(v_{q+2} \dots v_l)^3$ (to be interpreted as 1 when $q=l-1$). Therefore

$$|\mathcal{C}_q| = w_1 \dots w_q (v_{q+1}^3 - w_{q+1})(v_{q+2} \dots v_l)^3 \quad (0 \leqslant q < l), \tag{A10}$$

with the right-hand side to be interpreted as $(v_1^3 - w_1)(v_2 \dots v_l)^3$ when $q=0$, and as $w_1 \dots w_{l-1}(v_l^3 - w_l)$ when $q=l-1$.

We now insert a sublemma to Lemma B. Given $\mathbf{r}, \mathbf{s}, \mathbf{t}$, write $N(\mathbf{r}, \mathbf{s}, \mathbf{t}; \varepsilon)$ for the number of triples i, j, k with $\varrho_i \in R_{\mathbf{r}}$, $\sigma_j \in S_{\mathbf{s}}$, $\tau_k \in T_{\mathbf{t}}$ having (A5).

LEMMA C. *Suppose that* $\mathbf{r}, \mathbf{s}, \mathbf{t}$ *is in the class* $\mathcal{C}_q$. *Then*

$$N(\mathbf{r}, \mathbf{s}, \mathbf{t}; \varepsilon) \leqslant c(\varkappa, m_0) \varepsilon^{\varkappa} n^3 p^{(q-l)\varkappa} (v_1 \dots v_l)^{\varkappa - 3}.$$

Proof. Numbers $\xi \in m^{-1}\mathbb{Z}$ may uniquely be written as

$$\xi = \frac{y}{m_0} + \frac{z}{p^l} = \xi' + \xi'',$$

say, where $y, z \in \mathbb{Z}$ and $0 \leqslant z < p^l$. Accordingly, when $\varrho_i \in R_{\mathbf{r}}$, write $\varrho_i = \varrho_i' + \varrho_i''$. But $m_0 = x_l$, so that $\varrho_i \overset{m_0}{\equiv} \varrho_{i^*}$ for $\varrho_i, \varrho_{i^*} \in R_{\mathbf{r}}$, and therefore ϱ_i'' is the same for every $\varrho_i \in R_{\mathbf{r}}$. Using the same argument for $S_{\mathbf{s}}, T_{\mathbf{t}}$, we have

$$\varrho_i = \varrho_i' + \varrho'', \quad \sigma_j = \sigma_j' + \sigma'', \quad \tau_k = \tau_k' + \tau''$$

for $(\varrho_i, \sigma_j, \tau_k) \in R_{\mathbf{r}} \times S_{\mathbf{s}} \times T_{\mathbf{t}}$. Since $\mathbf{r}, \mathbf{s}, \mathbf{t} \in \mathcal{C}_q$, we have $x_q(\varrho_i - \sigma_j) = m_0 p^{l-q}(\varrho_i - \sigma_j) \in \mathbb{Z}$, therefore $p^{l-q}(\varrho'' - \sigma'') \in \mathbb{Z}$, and also $p^{l-q}(\varrho'' - \tau'') \in \mathbb{Z}$. On the other hand, when $q<l$, then (A7) does not hold with $q+1$ in place of q, so that not both $x_{q+1}(\varrho_i - \sigma_j)$, $x_{q+1}(\varrho_i - \tau_k)$ lie in $\mathbb{Z}$, and hence not both $p^{l-q-1}(\varrho'' - \sigma'')$, $p^{l-q-1}(\varrho'' - \tau'')$ lie in $\mathbb{Z}$. We may infer that the respective denominators a'' and b'' of $\varrho'' - \sigma''$ and $\varrho'' - \tau''$ have

$$\operatorname{lcm}(a'', b'') = p^{l-q}. \tag{A11}$$

Let $R'_{\mathbf{r}}$ be the system consisting of the ϱ'_i where $\varrho_i \in R_{\mathbf{r}}$, and define $S'_{\mathbf{s}}, T'_{\mathbf{t}}$ similarly. Then $R'_{\mathbf{r}} \sim R_{\mathbf{r}}$, $S'_{\mathbf{s}} \sim S_{\mathbf{s}}$, $T'_{\mathbf{t}} \sim T_{\mathbf{t}}$, so that $R'_{\mathbf{r}} \sim S'_{\mathbf{s}} \sim T'_{\mathbf{t}}$. Furthermore $R'_{\mathbf{r}} \overset{m_0}{\equiv} S'_{\mathbf{s}} \overset{m_0}{\equiv} T'_{\mathbf{t}}$. When $(\varrho'_i, \sigma'_j, \tau'_k) \in R'_{\mathbf{r}} \times S'_{\mathbf{s}} \times T'_{\mathbf{t}}$, let a'_{ij}, b'_{ik} be the respective denominators of $\varrho'_i - \sigma'_j$, $\varrho'_i - \tau'_k$. Then $a_{ij} = a'_{ij} a''$, $b_{ik} = b'_{ik} b''$. By (A11) and since $p \nmid a'_{ij} b'_{ik}$,

$$\operatorname{lcm}(a_{ij}, b_{ik}) = p^{l-q} \operatorname{lcm}(a'_{ij}, b'_{ik}).$$

The condition (A5) therefore becomes

$$\operatorname{lcm}(a'_{ij}, b'_{ik}) \leqslant \varepsilon p^{q-l} n = \varepsilon p^{q-l} v_1 \dots v_l (n / v_1 \dots v_l). \tag{A12}$$

We supposed Lemma B to be true for m_0. We apply this case of the lemma to $R'_{\mathbf{r}}, S'_{\mathbf{s}}, T'_{\mathbf{t}}$ with $\varepsilon p^{q-l} v_1 \dots v_l$ in place of ε, and observe that these three systems each have cardinality $n/(v_1 \dots v_l)$. Therefore $N(\mathbf{r}, \mathbf{s}, \mathbf{t}; \varepsilon)$, which is the number of triples $(\varrho_i, \sigma_j, \tau_k) \in R_{\mathbf{r}} \times S_{\mathbf{s}} \times T_{\mathbf{t}}$ with (A5), satisfies

$$\begin{aligned} N(\mathbf{r}, \mathbf{s}, \mathbf{t}; \varepsilon) &\leqslant c(\varkappa, m_0)(\varepsilon p^{q-l} v_1 \dots v_l)^{\varkappa} (n / v_1 \dots v_l)^3 \\ &= c(\varkappa, m_0) \varepsilon^{\varkappa} n^3 p^{(q-l)\varkappa} (v_1 \dots v_l)^{\varkappa - 3}. \end{aligned}$$

This completes the proof of Lemma C. □

We now continue with the proof of Lemma B. Clearly

$$N(\varepsilon) = \sum_{\mathbf{r}} \sum_{\mathbf{s}} \sum_{\mathbf{t}} N(\mathbf{r}, \mathbf{s}, \mathbf{t}; \varepsilon),$$

so that by Lemma C,

$$N(\varepsilon) \leqslant c(\varkappa, m_0) \varepsilon^{\varkappa} n^3 (v_1 \dots v_l)^{\varkappa - 3} \sum_{q=0}^{l} |\mathcal{C}_q| p^{(q-l)\varkappa}.$$

Here w_q enters in the formulas (A9), (A10) for $|\mathcal{C}_{q-1}|, |\mathcal{C}_q|, \dots, |\mathcal{C}_l|$. When w_q increases, then $|\mathcal{C}_{q-1}|$ decreases (or remains constant), whereas $|\mathcal{C}_q|, \dots, |\mathcal{C}_l|$ increase (or remain constant), but the sum $|\mathcal{C}_{q-1}| + |\mathcal{C}_q| + \dots + |\mathcal{C}_l|$ certainly is constant. Since the coefficient $p^{(q-1-l)\varkappa}$ of $|\mathcal{C}_{q-1}|$ in the above sum is smaller than the coefficients of $|\mathcal{C}_q|, \dots, |\mathcal{C}_l|$, the sum can only increase when w_q increases. Since we had $0 \leqslant w_q \leqslant v_q$, we may replace w_q by v_q $(q = 1, \dots, l)$. In this case the sum becomes (starting with the term $q = l$)

$$\begin{aligned} &v_1 \dots v_l + (v_1 \dots v_{l-1})(v_l^3 - v_l) p^{-\varkappa} + (v_1 \dots v_{l-2})(v_{l-1}^3 - v_{l-1}) v_l^3 p^{-2\varkappa} \\ &\quad + \dots + v_1 (v_2^3 - v_2)(v_3 \dots v_l)^3 p^{-(l-1)\varkappa} + (v_1^3 - v_1)(v_2 \dots v_l)^3 p^{-l\varkappa}. \end{aligned}$$

We may infer that

$$N(\varepsilon) \leqslant c(\varkappa, m_0) \varepsilon^{\varkappa} n^3 f_1(v_1, \dots, v_l) \leqslant c(\varkappa, m_0) \varepsilon^{\varkappa} n^3 f_{\gamma}(v_1, \dots, v_l)$$

where $\gamma=c(\varkappa,p)=(1-p^{\varkappa-2})^{-1}$ and

$$f_\xi(v_1,...,v_l)$$
$$=(v_1...v_l)^{\varkappa-2}\left(\xi+\frac{v_l^2-1}{p^\varkappa}+\frac{v_{l-1}^2-1}{p^{2\varkappa}}v_l^2+\frac{v_{l-2}^2-1}{p^{3\varkappa}}(v_{l-1}v_l)^2+...+\frac{v_1^2-1}{p^{l\varkappa}}(v_2...v_l)^2\right).$$

We claim that in the domain $1\leqslant v_q\leqslant p$ $(q=1,...,l)$, we have

$$f_\gamma(v_1,...,v_l)\leqslant\gamma=c(\varkappa,p), \tag{A13}$$

and this will finish the proof of Lemma B, hence of Theorem B. Here (A13) will be shown by induction on l. Hence we will assume that $l=1$, or that $l>1$ and (A13) has been established for $l-1$. When $v_{l-1},...,v_1$ are given, $f_\gamma(v_1,...,v_l)$ is of the form

$$Av_l^\varkappa+Bv_l^{\varkappa-2}$$

with positive coefficients A,B. This function in $v_l>0$ is first decreasing, then increasing, so that its maximum in any closed interval of positive reals is taken at an end point. For $l=1$ we have $f_\gamma(1)=\gamma$, $f_\gamma(p)=1+\gamma p^{\varkappa-2}-p^{-2}<1+\gamma p^{\varkappa-2}=\gamma$, so that in $1\leqslant v_1\leqslant p$ we have indeed $f_\gamma(v_1)\leqslant\gamma$. When $l>1$ we have by induction

$$\begin{aligned}f_\gamma(v_1,...,v_{l-1},1)&=(v_1...v_{l-1})^{\varkappa-2}\left(\gamma+\frac{v_{l-1}^2-1}{p^{2\varkappa}}+...+\frac{v_1^2-1}{p^{l\varkappa}}(v_1...v_{l-1})^2\right)\\&\leqslant f_\gamma(v_1,...,v_{l-1})\leqslant\gamma,\\f_\gamma(v_1,...,v_{l-1},p)&=(v_1...v_{l-1})^{\varkappa-2}\left(1+\gamma p^{\varkappa-2}-p^{-2}+\frac{v_{l-1}^2-1}{p^\varkappa}+...+\frac{v_1^2-1}{p^{(l-1)\varkappa}}(v_1...v_{l-1})^2\right)\\&\leqslant f_\gamma(v_1,...,v_{l-1})\leqslant\gamma,\end{aligned}$$

since $1+\gamma p^{\varkappa-2}-p^{-2}<\gamma$. Our claimed estimate (A13) follows. □

With a view to application in the main part of the paper, we will now formulate a corollary to Theorem A. When R is a system as above, we will say that a triple of integers i,j,k in $1\leqslant i,j,k\leqslant n$ is *ε-bad* if (A2) holds. Note that this property is independent of the ordering of the triple. Let $l\geqslant 3$, and consider l-tuples of integers $u_1,...,u_l$ in $1\leqslant u_1,...,u_l\leqslant n$. We call such an l-tuple *ε-bad* if some triple u_i,u_j,u_k with distinct i,j,k is ε-bad.

COROLLARY. *Suppose that $R=\{\varrho_1,...,\varrho_n\}$ is homogeneous. Then for any $l\geqslant 3$, the number of ε-bad l-tuples $u_1,...,u_l$ is*

$$<\varepsilon^{1/2}l^3n^l.$$

Proof. By the case $\varkappa=\frac{1}{2}$ of Theorem A, the number of ε-bad triples is

$$\leqslant \zeta\left(\tfrac{3}{2}\right)\varepsilon^{1/2}n^3 < 3\varepsilon^{1/2}n^3.$$

Hence given a triple i, j, k with $1 \leqslant i < j < k \leqslant l$, the number of l-tuples $u_1, ..., u_l$ for which u_i, u_j, u_k is ε-bad is $< 3\varepsilon^{1/2}n^3 \cdot n^{l-3} = 3\varepsilon^{1/2}n^l$. The number of triples i, j, k in question is $\binom{l}{3}$, so that the number of ε-bad l-tuples is

$$\leqslant 3\binom{l}{3}\varepsilon^{1/2}n^l < \varepsilon^{1/2}l^3n^l. \qquad \square$$

References

[1] Dobrowolski, E., On a question of Lehmer and the number of irreducible factors of a polynomial. *Acta Arith.*, 34 (1979), 391–401.

[2] Evertse, J. H., The number of solutions of linear equations in roots of unity. To appear.

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

[4] Lech, C., A note on recurring series. *Ark. Mat.*, 2 (1953), 417–421.

[5] Poorten, A. J. van der & Schlickewei, H. P., Zeros of recurrence sequences. *Bull. Austral. Math. Soc.*, 44 (1991), 215–223.

[6] Schlickewei, H. P., Multiplicities of algebraic linear recurrences. *Acta Math.*, 170 (1993), 151–180.

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

[8] — Equations in roots of unity. *Acta Arith.*, 76 (1996), 99–108.

[9] Schlickewei, H. P. & Schmidt, W. M., The number of solutions of polynomial-exponential equations. To appear in *Compositio Math.*

[10] Shorey, T. N. & Tijdeman, R., *Exponential Diophantine Equations*. Cambridge Tracts in Math., 87. Cambridge Univ. Press, Cambridge–New York, 1986.

[11] Voutier, P., An effective lower bound for the height of algebraic numbers. *Acta Arith.*, 74 (1996), 81–95.

Wolfgang M. Schmidt
Department of Mathematics
University of Colorado at Boulder
Campus Box 395
Boulder, Colorado 80309-0395
U.S.A.
schmidt@euclid.colorado.edu

Received March 3, 1998