Compositio Mathematica **131:** 319–340, 2002.

Some New Applications of the Subspace Theorem

PIETRO CORVAJA[1] and UMBERTO ZANNIER[2]
Dipartimento di Mat. e Inf., via delle Scienze, 206, 33100 Udine, Italy.
e-mail: corvaja@dimi.uniud.it
[2]*IUAV – DCA, S. Croce, 191, 30135 Venice, Italy. e-mail: zannier@iuav.unive.it*

(Received: 28 July 2000; accepted in final form: 21 December 2000)

Abstract. We present some applications of the Subspace Theorem to the investigation of the arithmetic of the values of Laurent series $f(z)$ at S-unit points. For instance we prove that if $f(q^n)$ is an algebraic integer for infinitely many n, then $h(f(q^n))$ must grow faster than n. By similar principles, we also prove diophantine results about power sums and transcendency results for lacunary series; these include as very special cases classical theorems of Mahler. Our arguments often appear to be independent of previous techniques in the context.

Mathematics Subject Classifications (2000). 11J25; 11J81

Key words. Subspace Theorem. Mahler's series. Exponential diophantine equations

Introduction

In the recent papers [CZ1] and [CZ2] we exploited a simple application of the Subspace Theorem to exponential diophantine equations, which seemed new. This originates in Lemma 2 of [CZ1] which, roughly speaking, states that a power-sum cannot be too near to an integer unless it is itself an integer.

In the present paper we expand this principle to more general situations and we give different types of applications. This development is represented by Theorem 4 below, which we view as a main tool for all the other results of this paper. It concerns the approximation of an S-integer by a sum of S-units.

As a first application, we shall prove a theorem (Theorem 1) about values at S-units of an infinite Laurent series with algebraic coefficients. Roughly speaking we shall prove that if infinitely many such values are S-integers in a given number field, then their height must grow rapidly.

From this result we shall deduce (Corollary 1) information on the values of such series at q^n, for an algebraic number q: if the values are S-integers in a given number field for infinitely many n, then their height must grow faster then n. As a further corollary (Corollary 2), they cannot all be rational integers.

As another application of Theorem 4, we prove (Theorem 2) a strengthening of a previous diophantine result in [CZ1] for power sums with a 'dominant root'. This involves Puiseux expansions.

https://doi.org/10.1023/A:1015594913393 Published online by Cambridge University Press

Another type of application of the same principle concerns algebraic values of lacunary series. The present Theorem 3 provides a transcendency criterion for series including those of the form $\sum_{i=1}^{\infty} \alpha^{m_i}$ where α is an algebraic number and the sequence of integers m_i grows exponentially. We recover in particular the classical results of Mahler on the Fredholm series. We remark that these latter results may be vastly generalized by expanding the techniques introduced by Mahler (see [M], [Lo-vdP], [N], [Ma], [T]), which allow even algebraic independence results. On the other hand, such techniques always rely on certain functional equations satisfied by the relevant series, while our method only exploits lacunarity of the series.

We recall that the Subspace Theorem had been used also by Nishioka [N] in the context of transcendency of Mahler's series. However her application follows a completely different pattern and leads to different results.

All the above theorems hold by interpreting the values of the occurring series with respect to any valuation of the involved number field.

NOTATION AND STATEMENTS

We let K be a number field and S be a finite set of absolute values of K containing the Archimedean ones. For every place υ of K we note by $|\cdot|_\upsilon$ a continuation of it to $\bar{\mathbf{Q}}$ and normalize it 'with respect to K': according to this normalization, for $x \in K^\times$ the absolute logarithmic Weil height reads $h(x) = \sum_\upsilon \log^+ |x|_\upsilon$ and the product formula $\prod_\upsilon |x|_\upsilon = 1$ holds. We note that these conditions determine uniquely our normalizations. We also note that even in the Archimedean case the triangle-inequality holds with these normalizations. In fact, the present absolute value is obtained from the usual one by raising to a power between 0 and 1.

We fix an absolute value v of K and denote by $\mathbf{C}_v$ a completion of an algebraic closure of K_v. Our notion of convergence, unless otherwise specified, refers to $\mathbf{C}_v$.

We also define the S-height of a nonzero element $x \in K^\times$ to be $h_S(x) = \sum_{\upsilon \notin S} \log^+ |x|_\upsilon$. For S-integers this height vanishes, so it gives a mesure of 'how far' x is from being an S-integer.

For a vector $\mathbf{z} = (z_0, z_1, \dots, z_h) \in K^{h+1} \setminus \{0\}$, $(h \geqslant 1)$, we define $h(\mathbf{z})$ as the usual projective logarithmic height. Also, we denote by $\hat{h}(\mathbf{z})$ (resp. $\hat{h}_S(\mathbf{z})$), the sum of the $h(z_i)$ (resp. $h_S(z_i)$), $0 \leqslant i \leqslant h$. Moreover, we put, for an absolute value υ, $||\mathbf{z}||_\upsilon := \max\{|z_0|_v, \dots, |z_h|_v\}$.

Throughout, the capital 'H' will denote the exponential height.

By *Laurent series* (resp. *Laurent polynomial*) we mean, as usual, a series of the form $\sum_{n \geqslant -d} a_n z^n$ (resp. finite sum of the form $\sum_{-d \leqslant n \leqslant k} a_n z^n$).

THEOREM 1. *Let $f(z) = \sum_{i \geqslant -d} a_i z^i$ be a Laurent series with algebraic coefficients in $\mathbf{C}_v$, converging for $0 < |z|_v < 1$. Let S be a finite set of absolute values of K containing the Archimedean ones. Let z_n $(n = 1, 2, \dots)$ be an infinite sequence of pairwise distinct elements of $K^\times$ such that $f(z)$ is defined and belongs to K. Suppose that:*

https://doi.org/10.1023/A:1015594913393 Published online by Cambridge University Press

(1) $\dfrac{h_S(z_n) + h_S(z_n^{-1})}{h(z_n)} \to 0$ *as* $n \to \infty$*;*

(2) $\dfrac{h(z_n)}{-\log |z_n|_v}$ *is a bounded sequence;*

(3) $\lim_{n\to\infty} \dfrac{h_S(f(z_n))}{h(z_n)} = 0$*;*

(4) $\liminf_{n\to\infty} \dfrac{h(f(z_n))}{h(z_n)} < \infty$.

Then f is a Laurent polynomial.

Remark. Condition (1) states that the algebraic numbers z_n 'tend to be' S-units: in fact, the vanishing of both $h_S(z)$ and $h_S(z^{-1})$ characterizes S-units. Analogously, condition (3) requires that $f(z_n)$ tend to be S-integers for $n \to \infty$; it automatically holds when the $f(z_n)$ are S-integers.

Note that, though we require that the coefficients a_i are algebraic, we allow the possibility that they generate a field of infinite degree over $\mathbf{Q}$. This appears to be new in this context.

COROLLARY 1. *Let $f(z)$, S be as in Theorem 1, f not a Laurent polynomial. Let $q \in \mathbf{C}_v^\times$ be algebraic with $|q|_v < 1$. If $f(q^n)$ is an S-integer in K for all n in an infinite sequence $\mathcal{A} \subset \mathbf{N}$, then*

$$\lim_{n\in\mathcal{A}} \frac{h(f(q^n))}{n} = \infty.$$

If we just assume that $f(q^n)$ is an algebraic number in K, the conclusion does not follow in general, as is shown by simple examples like $f(z) = \sum z^{m}$.

As a particular case, we shall easily obtain the following corollary:

COROLLARY 2. *Let f be a Laurent series with complex algebraic coefficients, $q \in \mathbf{C}$ be an algebraic number with $0 < |q| < 1$. If f is not a Laurent polynomial, then the set of positive integers n such that $f(q^n)$ is a rational integer is finite.*

Remark. Let us consider the Fredholm series $f(z) = \sum_i z^{2^i}$ satisfying the functional equation $f(z^2) = f(z) - z$. Iterating this relation we see that if $f(q) \in K$ for a $q \in K^\times$, then $f(q^{2^n})$ is an S-integer in K for all n; moreover, by the same iteration, it may be also easily seen that $h(f(q^{2^n}))/2^n$ is bounded. This contradicts Corollary 1; therefore $f(q)$ is transcendental and we recover Mahler's theorem. In Theorem 3 we shall prove a much more general result, holding for series which do not necessarily satisfy simple functional equations.

In the context of algebraic functions we have the following result about diophantine equations with power sums. Given $c_i, \alpha_i \in K^\times$, $i = 1, \ldots, h$, we consider

https://doi.org/10.1023/A:1015594913393 Published online by Cambridge University Press

the power sum

$$z_n = c_1\alpha_1^n + \cdots + c_h\alpha_h^n. \tag{0.1}$$

THEOREM 2. *Let* z_n *be given by* (0.1). *Assume that for some absolute value* v, *we have* $1 \neq |\alpha_1|_v > \max(|\alpha_2|_v, \ldots, |\alpha_h|_v)$. *Let* $g \in K[Z, X]$ *be monic in* X *and suppose that for an infinite sequence of* $n \in \mathbf{N}$, *the equation* $g(z_n, X) = 0$ *has a solution* $X = x_n \in K$. *Then there exist* $d_j, \beta_j \in \bar{K}^\times$, $j = 1, \ldots, k$, *and an arithmetic progression* $\mathcal{P}$ *such that we have*

$$g\left(\sum_{i=1}^h c_i\alpha_i^n, \sum_{j=1}^k d_j\beta_j^n\right) = 0, \qquad \textit{for } n \in \mathcal{P}.$$

Remark. For simplicity we have introduced the condition that g is monic. It is a well-known trick how to get rid of this (with a corresponding modification of the conclusion); namely, we may replace $g(Z, X)$ with the polynomial $a(Z)^{d-1}g(Z, X/a(Z))$, where $a(Z)$ is the leading coefficient of g, with respect to X.

Similar results appear in our paper [CZ1], where however we considered in detail only the case when the α_i are natural numbers. Unlike that paper, the present treatment is completely independent of Siegel's theorem on integral points on curves. Although the condition on the 'dominant root' α_1 is crucial, the method allows much flexibility concerning the assumptions on the z_n. For instance, by a direct application of Theorem 1 to Puiseux expansions, one can treat the case when the z_n are S-units, again independently of Siegel's theorem or any tool from algebraic geometry. (Siegel's theorem had been used by Dèbes [D] in connection with similar equations.)

A particular case concerns the equation $g(m) = z_n$, where g is a polynomial and z_n is as in the theorem. Such an equation is not covered by quite general results by M. Laurent [L1], [L2], dealing with relations $u_m = v_n$, where u, v are recurrence sequences.

We finally remark that results of this type lead to explicit versions of Hilbert Irreducibility Theorem and to the construction of simple Universal Hilbert Sets (see, e.g., [D], [CZ1], [DZ]). For instance it can be proved that $\{z_n + n\}$ is a Universal Hilbert Set if $z_n \in \mathbf{Z}$ is as in the statement of Theorem 2.

Our next results deal with the transcendency of values of lacunary series at algebraic points.

THEOREM 3. *Let* $m_1 < m_2 < \cdots$ *be positive integers and* $a_1, a_2, \ldots \in K^\times$ *be such that* $\sum_{i=1}^n h(a_i) = \mathrm{o}(m_n)$. *Let* $\alpha \in K^\times$, $|\alpha|_v < 1$ *and consider the number* $\gamma = \sum_{i=1}^\infty a_i\alpha^{m_i}$, *which is well defined as a limit in the* v-*topology. Let* N *be a positive integer and let* $L > h(\alpha)/(\log|\alpha|_v^{-1})$ *be a real number. Consider the sequence* $\mathcal{N}$ *of integers* n *such that* $m_{n+N} > L \cdot m_n$. *Then, either* γ *is transcendental, or for all but finitely many* $n \in \mathcal{N}$ *there exists a set of integers* $\mathcal{A}_n$ *with* $\{1, \ldots, n\} \subset \mathcal{A}_n \subset \{1, \ldots, n+N\}$ *such that* $\gamma = \sum_{i\in\mathcal{A}_n} a_i\alpha^{m_i}$.

This statement is rather involved, but leads quite easily to several simpler corollaries. We now give a few examples:

COROLLARY 3. *Let m_i be an increasing sequence of integers satisfying*

$$\sup_N \limsup_n \frac{m_{n+N}}{m_n} = \infty.$$

Let a_i be as in Theorem 3 and assume in addition that the a_i are positive reals. Then the real function defined in $(0, 1)$ by the series $\sum_{i=1}^{\infty} a_i x^{m_i}$ takes transcendental values at all algebraic points in $(0, 1)$.

In the p-adic case we have the following analogue:

COROLLARY 3′. *Let m_i be as in Corollary 3 and consider an ultrametric absolute value v. If the a_i are as in Theorem 3 and satisfy $|a_i|_v = 1$ for all i, then the v-adic function defined in the unit disk by the series $\sum_{i=1}^{\infty} a_i x^{m_i}$ assumes therein transcendental values for algebraic $x \neq 0$.*

COROLLARY 4. *Let $\alpha \in K^{\times}$, $|\alpha|_v < 1$ and let $m_i \in \mathbf{N}$ be an increasing sequence of integers. Assume that for some positive integers h, N,*

(i) $m_{n+h} - m_n \to \infty$,

(ii) $\limsup_n \dfrac{m_{n+N}}{m_n} > \dfrac{h(\alpha)}{\log |\alpha|_v^{-1}}$.

Then either $\sum_i \alpha^{m_i}$ is transcendental, or there exist pairwise disjoint finite sets $\mathcal{B}_1, \mathcal{B}_2, \ldots$ of natural numbers with the following properties:

(1) *the union $\bigcup \mathcal{B}_i$ has finite complement in* $\mathbf{N}$.

(2) $\sum_{i \in \mathcal{B}_n} \alpha^{m_i} = 0$, *for $n = 1, 2, \ldots$.*

(3) *For each n, the set $\mathcal{B}_n$ is contained in some interval of length $\leqslant h$.*

In particular, we see that if we may take $h \leqslant 2$ in (i), then $\sum_i \alpha^{m_i}$ is transcendental. It can be shown that this is the case, e.g. when $\{m_i\}_{i \in \mathbf{N}}$ are the values of a recurrence sequence such that for no positive integers d, r, the sequence $\{m_{di+r}\}_{i \in \mathbf{N}}$ is polynomial.

Note that in general the transcendency conclusion may be false. In fact, take for α a root in the unit disk of a polynomial $p(x) = x^{d_1} + \cdots + x^{d_h}$, with $0 \leqslant d_1 < \cdots < d_h$. Take now any lacunary series $q(x) = \sum_{i=0}^{\infty} x^{e_i}$, where $e_{i+1} - e_i \to \infty$. Consider finally an expansion of the product $p(x)q(x)$. Evaluating at α, we immediately obtain the required counterexample. (This construction however does not work in the non-Archimedean context, as shown by Corollary 3′.)

Another case when the transcendency conclusion holds unconditionally is obtained by requiring that the sequence m_i is strongly lacunary. In fact, we have this further result.

COROLLARY 5. *Let $m_i \in \mathbf{N}$ be an increasing sequence of integers such that* $\liminf_n (m_{n+1})/m_n > 1$. *Let α, a_i be as in Theorem 3. Then $\sum_i a_i \alpha^{m_i}$ is transcendental.*

Remark. These results appear to be new. They plainly include Mahler's theorems on lacunary series $\sum_n \alpha^{d^n}$, as well as some results recently announced by Tanaka (for instance, he obtains the conclusion of Corollary 5 under the assumptions $a_i = 1$ and $m_{i+1} \geqslant 2m_i$ for all i). On the contrary, it seems that Mahler's techniques do not apply in such generality. Our method could yield more general results; for the sake of simplicty, we limited ourselves to the present statements.

1. An Auxiliary Result

The main technical point in all the proofs is the following consequence of the Subspace Theorem:

THEOREM 4. *Let K be a number field, S a finite set of absolute values of K containing the Archimedean ones, v be an absolute value from S, ε be a positive real number, $N \geqslant 0$ an integer. Finally, let $c_0, \ldots, c_N \in \bar{K}^*$. For $\delta > (N+2)\varepsilon$, there are only finitely many $(N+1)$-tuples $\mathbf{w} = (w_0, \ldots, w_N) \in (K^*)^{N+1}$ such that the inequalities*

(i) $h_S(w_i) + h_S(w_i^{-1}) < \varepsilon h(w_i)$ *for* $i = 1, \ldots, N$

(ii) $|c_0 w_0 + c_1 w_1 + \cdots + c_N w_N|_v < \left(H(w_0) H_S(w_0)^{N+1}\right)^{-1} \hat{H}(\mathbf{w})^{-\delta}$
hold and no subsum of the $c_i w_i$ involving $c_0 w_0$ vanishes.

Introducing the coefficients c_i is important for the application to Theorems 1,2 above. However, for other applications (e.g. to Theorem 3), we can take $c_i = 1$ for all i. In fact, the proof in the case of general c_i is exactly the same as in the case $c_i = 1$.

The condition about the subsum is somewhat typical of the theory of S-unit equations and inequalities: we quote for instance the celebrated S-unit Theorem by Evertse [E] and van der Poorten and Schlickewei [vdP-S], valid for vanishing sums of S-units (see also [Sc, Thm. 2A]).

The new feature in our statement is represented by the role of the term w_0. In our applications it is crucial that w_0 is an (almost) S-integer, i.e. $H_S(w_0)$ is small compared to $H(w_0)$; this fact allows (ii) to be verified in the cases of interest for us. The above quoted results are usually proved under the stronger assumption that w_0 is an (almost) S-unit. For our applications it is very important not to have w_0 restricted to (almost) S-units. (As we have remarked above, particular cases of this principle were already applied in [CZ1] and [CZ2].)

For this reason, moreover, our arguments follow a different pattern with respect to the proof of the S-unit Theorem, as given, e.g., in [Sc, Thm. 2A]. In both contexts, an application of the Subspace Theorem leads to a linear relation with fixed coefficients among the w_i's, and this allows to reduce the number of variables by a linear substitution. In the classical case, this procedure eventually leads to some

https://doi.org/10.1023/A:1015594913393 Published online by Cambridge University Press

constant ratio w_i/w_j ($i \neq j$), and the conclusion readily follows. In our case on the contrary, due to the lack of symmetry of (ii), the stronger conclusion about w_i/w_j is not generally true (consider, e.g., the example $w_0 = 2^n + 3^n$, $w_1 = -2^n$, $w_2 = -3^n$, $w_3 = 2^{-n^2}$). Therefore, at this point we have used a supplementary combinatorial argument from linear algebra (Lemma 1 below) to obtain a vanishing sum.

An alternative procedure, more similar to the classical one, would consist in applying a result by Evertse [E] to linear relations with fixed coefficients among $w_1, \ldots, w_N$.

As noted above, if we are satisfied with any linear relation with fixed coefficients (not necessarily 0 or 1), a straightforward application of the Subspace Theorem is sufficient.

We immediately state and prove the mentioned linear algebra result, which will be used towards the end of the proof of Theorem 4.

We introduce just a bit of notation. We let k be an infinite field. We say that $T: k^{n+1} \to k^{n+1}$ is a *truncation operator* if there exists a set $s(T) \subset \{0, 1, \ldots, n\}$ such that

$$T((x_0, \ldots, x_n)) = (y_0, \ldots, y_n), \quad \text{where} \quad y_j = \begin{cases} x_j & \text{if } j \in s(T), \\ 0 & \text{if } j \notin s(T). \end{cases}$$

We say that $s(T)$ is the *support* of T. Note that $s(T)$ determines T.

We let $\mathbf{u} := (1, \ldots, 1) \in k^{n+1}$ be the vector with all components equal to 1.

LEMMA 1. *Let V be a nonzero vector subspace of k^{n+1}. Suppose that for every $\mathbf{x} = (x_0, \ldots, x_n) \in V$, and for every index $i \in \{1, \ldots, n\}$, either $x_i = 0$, or $x_i = x_0$ or $T(\mathbf{x}) - x_i T(\mathbf{u}) \in V$ for some truncation operator $T = T_{i,\mathbf{x}}$ with $0 \in s(T)$. Then V contains a vector $\mathbf{v} = (v_0, \ldots, v_n)$, where $v_0 = 1$ and $v_i \in \{0, 1\}$ for all $i = 0, 1, \ldots, n$.*

Proof. We note at once that it suffices to find any nonzero element $\mathbf{v} \in V$ such that $v_i \in \{0, 1\}$ for all i. In fact, suppose this is the case, but $v_0 = 0$. Let $v_i = 1$ and choose $T = T_{i,\mathbf{v}}$ as in the assumption. Then the vector $T(\mathbf{u}) - T(\mathbf{v})$ satisfies the conclusion.

We now argue by induction on n, the result being true for $n = 0$. Let now $n > 0$.

Suppose that $x_j = 0$ for a certain $j \in \{1, \ldots, n\}$ and all $\mathbf{x} \in V$ and let P_j be the projection outside the jth coordinate. Then P_j induces an isomorphism of V with a subspace of k^n (whose coordinates we number with $0, \ldots, j-1, j+1, \ldots, n$). On the other hand, if T is a truncation operator on k^{n+1}, we have $P_j \circ T = T' \circ P_j$, where T' is the truncation operator on k^n with support $s(T) \setminus \{j\}$. It follows at once that the assumptions of the Proposition are verified for $P_j(V)$ in place of V. The induction assumption then implies the existence of a nonzero $\mathbf{v} \in V$ such that $P_j(\mathbf{v})$ has all its coordinates in $\{0, 1\}$. However this must then hold also for $\mathbf{v}$ itself, proving the desired conclusion. The same argument works if $x_j = x_0$ for a certain nonzero j and all $\mathbf{x} \in V$.

Therefore we may assume that neither $x_0 = 0$ nor $x_0 = x_j$ holds identically on V.

Fix for the moment a nonzero index i and let T be a truncation operator with $0 \in s(T)$. Put $\Lambda_T(\mathbf{x}) = T(\mathbf{x}) - x_i T(\mathbf{u})$, so Λ_T is a linear operator on k^{n+1}. Our assump-

https://doi.org/10.1023/A:1015594913393 Published online by Cambridge University Press

tion reads

$$V \subset \left(\bigcup_T \Lambda_T^{-1}(V) \right) \cup V' \cup V'',$$

where V' (resp. V'') is the subspace of V defined by $x_i = 0$ (resp. $x_i = x_0$). Recall that we are assuming that both V', V'' are proper subspaces of V. Therefore, since k is infinite, V must be contained in some space $\Lambda_T^{-1}(V)$.

Namely, we may replace our assumption by the (apparently) stronger one asserting that *for all nonzero i there exists a truncation operator T (depending only on i) with $0 \in s(T)$, such that $\Lambda_T(V) \subset V$.*

We now assume that the conclusion of the Lemma is not true and proceed to derive a contradiction.

We consider all decompositions $V = V_1 \oplus V_2$, where V_1, V_2 are subspaces of V satisfying the following: *for some renumbering of* $\{1, 2, \ldots, n\}$ *and for a certain index* $i \in \{0, 1, \ldots, n\}$ *we have:*

(a) $x_j = x_0$ *for* $0 \leqslant j \leqslant i$ *and all* $\mathbf{x} \in V_1$;
(b) $x_j = 0$ *for all* $j \geqslant i$ *and all* $\mathbf{x} \in V_2$.

Taking $i = 0$, $V_1 = V$, $V_2 = \{0\}$, we see that such a decomposition exists. We then choose some decomposition corresponding to a maximal possible i. Suppose first $i = n$. Since we are assuming that no coordinate vanishes identically on V, we see that $V_2 \neq V$. Therefore V_1 contains a nonzero vector, which has by definition all of its coordinates equal. Hence $\mathbf{u}$ lies in V and in particular the conclusion of the lemma holds in this case. Therefore we may assume $i < n$.

Let $\pi: V = V_1 \oplus V_2 \to V_1$ be the projection on V_1. Pick $\ell > i$ and let T be a truncation operator as above corresponding to ℓ, namely with $0 \in s(T)$ and such that $T(\mathbf{x}) - x_\ell T(\mathbf{u}) \in V$ for all $\mathbf{x} \in V$. We put $\Lambda(\mathbf{x}) := T(\mathbf{x}) - x_\ell T(\mathbf{u})$ and, for $\mathbf{x} \in V_1$,

$$L(\mathbf{x}) = (\pi \circ \Lambda|_{V_1})(\mathbf{x}) = \pi(T(\mathbf{x}) - x_\ell T(\mathbf{u})),$$

so L is an endomorphism of V_1. (Note that we do not define L outside V_1.)

Observe that $y_\ell = 0$ for all $\mathbf{y} = (y_0, \ldots, y_n) \in \operatorname{Im} L$. In fact, the ℓth coordinate of $\Lambda(\mathbf{x})$ is plainly 0, for all $\mathbf{x} \in V$. Also, since $\ell > i$, the projection on V_1 does not alter the ℓth coordinate, since by property (b) we have $x_\ell = 0$ for all $\mathbf{x} \in V_2$.

Suppose $\mathbf{x} \in \ker L$. This means $\mathbf{y} := \Lambda(\mathbf{x}) \in V_2$. For all j, we have that y_j is either $x_j - x_\ell$ or 0 according as j does belong to $s(T)$ or not. Since $\mathbf{x} \in V_1$, we then see by (a) above that y_j is either $x_0 - x_\ell$ or 0 for $j \leqslant i$. Also, since $\mathbf{y} \in V_2$, y_j is 0 for all $j \geqslant i$, by (b) above. Therefore the coordinates of $\mathbf{y}$ assume at most one nonzero value. If $\mathbf{y} \neq 0$, a nonzero multiple of $\mathbf{y}$ leads to the conclusion. So we may assume that $\mathbf{y} = 0$ for all $\mathbf{x} \in \ker L$, which amounts to the inclusion $\ker L \subset \ker \Lambda$. We contend that

$$\ker L \cap \operatorname{Im} L = \{0\}. \tag{1.1}$$

In fact, let $\mathbf{x} \in \ker L \cap \operatorname{Im} L$. In particular, we may write $\mathbf{x} = (\pi \circ \Lambda)(\mathbf{y})$ for some $\mathbf{y} \in V_1$, so

$$\mathbf{x} = T(\mathbf{y}) - y_\ell T(\mathbf{u}) + \mathbf{v}, \quad \text{where } \mathbf{y} \in V_1, \mathbf{v} \in V_2. \tag{1.2}$$

This implies $x_\ell = 0$, so $\Lambda(\mathbf{x}) = T(\mathbf{x})$. Since $\mathbf{x} \in \ker L \subset \ker \Lambda$ we have $T(\mathbf{x}) = 0$, which implies $x_0 = 0$, because $0 \in s(T)$. Since $\mathbf{x} \in V_1$, this fact in turn leads to $x_j = 0$ for $0 \leqslant j \leqslant i$.

Put $\mathbf{v} = (v_0, \dots, v_n)$. Taking into account that $\mathbf{y} \in V_1$ and $\mathbf{v} \in V_2$ we get

$$v_j = \begin{cases} x_j - y_j + y_\ell = -y_0 + y_\ell & \text{if } j \in s(T) \text{ and } 0 \leqslant j \leqslant i \\ 0 & \text{if } j > i \text{ or } j \notin s(T). \end{cases}$$

In particular the coordinates of $\mathbf{v} \in V$ have at most one nonzero value. If $\mathbf{v}$ would be nonzero, our conclusion would be true, so we can assume $\mathbf{v} = 0$. Plugging this into (1.2) we see that

$$\mathbf{x} = T(\mathbf{y}) - y_\ell T(\mathbf{u}), \quad T(\mathbf{x}) = 0.$$

The first equation shows that $\mathbf{x} \in \operatorname{Im} T$, so $\mathbf{x} \in \operatorname{Im} T \cap \ker T$, which is plainly 0. This concludes the verification of (1.1).

It is now easy to conclude the proof of Lemma 1. By (1.1) we may write $V_1 = \ker L \oplus \operatorname{Im} L$, whence $V = (\ker L) \oplus (\operatorname{Im} L \oplus V_2)$. Observe that $\mathbf{x} \in \ker L \subset \ker \Lambda$ implies that $x_0 = x_j$ for $0 \leqslant j \leqslant i$ (since $\mathbf{x} \in V_1$) and $x_0 = x_\ell$ (since $\mathbf{x} \in \ker \Lambda$, i.e. $T(\mathbf{x}) = x_\ell T(\mathbf{u})$, and $0 \in s(T)$). The equation $T(\mathbf{x}) = x_\ell T(\mathbf{u})$ also implies that $x_j = x_\ell = x_0$ for all $j \in s(T)$.

In conclusion, $x_0 = x_j$ for $j \in R := \{0, 1, \dots i\} \cup \{\ell\} \cup s(T)$ and all $\mathbf{x} \in \ker L$.

Suppose now that $\mathbf{x} \in \operatorname{Im} L \oplus V_2$. Then $\mathbf{x} \in \operatorname{Im} \Lambda \oplus V_2$, so we may write $\mathbf{x} = T(\mathbf{y}) - y_\ell T(\mathbf{u}) + \mathbf{v}$ for some $\mathbf{y} \in V$, $\mathbf{v} \in V_2$. This shows that $x_j = 0$ for $j \notin R$ and for $j = \ell$.

Renumbering the indices $\{i+1, \dots, n\}$ by means of a permutation σ, we may assume that $R = \{0, 1, \dots, h\}$, where $\sigma(\ell) = h$. We have thus found a new renumbering of $\{1, 2, \dots, n\}$ and a decomposition for V of the required shape, where i however has been replaced by $h > i$. This contradicts the maximality of i and concludes the proof of the lemma. □

Proof of Theorem 4. We argue by induction on N. The case $N = 0$ is easy: in fact, the assumption (ii) contradicts Liouville's bound for large $H(w_0)$. Therefore, the finiteness of the sequence of such w_0 follows.

We now suppose that the theorem holds up to $N - 1$ and assume by contradiction that there exist infinitely many $(N+1)$-tuples $\mathbf{w}$ satisfying (i) and (ii), and admitting no vanishing subsums of the terms $c_i w_i$ involving $c_0 w_0$.

Let us define, for every $v \in S$, $N + 1$ independent linear forms in $\mathbf{X} := (X_0, \dots, X_N)$ as follows: put

$$L_{v,0}(\mathbf{X}) = c_0 X_0 + \cdots + c_N X_N$$

and for $v \in S$, $0 \leqslant i \leqslant n$, $(\upsilon, i) \neq (v, 0)$ put $L_{\upsilon,i}(\mathbf{X}) = X_i$. Let as before $\mathbf{w} = (w_0, \dots, w_N) \in K^{N+1}$ and consider the double product

$$\prod_{\upsilon \in S} \prod_{i=0}^{N} \frac{|L_{\upsilon,i}(\mathbf{w})|_\upsilon}{||\mathbf{w}||_\upsilon},$$

where $||\mathbf{w}||_\upsilon = \max_{0 \leqslant j \leqslant N} |w_j|_\upsilon$. By putting

$$\sigma = c_0 w_0 + \cdots + c_N w_N = L_{v,0}(\mathbf{w}),$$

we can rewrite the double product as

$$\prod_{\upsilon \in S} \prod_{i=0}^{N} \frac{|L_{\upsilon,i}(\mathbf{w})|_\upsilon}{||\mathbf{w}||_\upsilon} = |\sigma|_v \cdot \left(\prod_{\upsilon \in S \setminus \{v\}} |w_0|_\upsilon \right) \left(\prod_{v \in S} \prod_{i=1}^{N} |w_i|_v \right) \left(\prod_{\upsilon \in S} ||\mathbf{w}||_\upsilon \right)^{-(N+1)}.$$

By applying the product formula to $w_1 \cdots w_N$ we can replace the term $\left(\prod_{\upsilon \in S} \prod_{i=1}^{N} |w_i|_\upsilon\right)$ in the above equation by $\left(\prod_{\upsilon \notin S} \prod_{i=1}^{N} |w_i|_\upsilon\right)^{-1}$ which is bounded above by $H_S(w_1^{-1}) \cdots H_S(w_N^{-1})$. Then we obtain the upper bound

$$\prod_{\upsilon \in S} \prod_{i=0}^{N} \frac{|L_{\upsilon,i}(\mathbf{w})|_\upsilon}{||\mathbf{w}||_\upsilon} \leqslant |\sigma|_v \left(\prod_{\upsilon \in S \setminus \{v\}} |w_0|_\upsilon \right) \prod_{i=1}^{N} H_S(w_i^{-1}) \left(\prod_{\upsilon \in S} ||\mathbf{w}||_\upsilon \right)^{-(N+1)}.$$

Since

$$\prod_{\upsilon \in S} ||\mathbf{w}||_\upsilon = H(\mathbf{w}) \cdot \left(\prod_{\upsilon \notin S} ||\mathbf{w}||_\upsilon \right)^{-1} \geqslant H(\mathbf{w}) \cdot \prod_{i=0}^{N} H_S(w_i)^{-1}$$

and since $H_S(w_i) \cdot H_S(w_i^{-1}) < H(w_i)^\varepsilon$ for every $i \geqslant 1$, by assumption, we obtain after some calculations the upper bound

$$\prod_{\upsilon \in S} \prod_{i=0}^{n} \frac{|L_{\upsilon,i}(\mathbf{w})|_\upsilon}{||\mathbf{w}||_\upsilon} \leqslant |\sigma|_v \left(\prod_{\upsilon \in S \setminus \{v\}} |w_0|_\upsilon \right) \prod_{i=1}^{N} H(w_i)^{\varepsilon(N+1)} H(\mathbf{w})^{-N-1} H_S(w_0)^{N+1}.$$

Also, we trivially have $\prod_{\upsilon \in S \setminus \{v\}} |w_0|_\upsilon \leqslant H(w_0)$, whence

$$\prod_{\upsilon \in S} \prod_{i=0}^{n} \frac{|L_{\upsilon,i}(\mathbf{w})|_\upsilon}{||\mathbf{w}||_\upsilon} \leqslant |\sigma|_v H(w_0) \prod_{i=1}^{N} H(w_i)^{\varepsilon(N+1)} H(\mathbf{w})^{-N-1} H_S(w_0)^{N+1}. \tag{1.3}$$

Observe that $L_{v,0}(\mathbf{w}) = \sigma$ is nonzero, because of our assumption about vanishing subsums. Therefore, by the Subspace Theorem, as formulated in [Sc], the lower bound

$$\prod_{\upsilon \in S} \prod_{i=0}^{N} \frac{|L_{\upsilon,i}(\mathbf{w})|_\upsilon}{||\mathbf{w}||_\upsilon} > H(\mathbf{w})^{-N-1-\varepsilon} \tag{1.4}$$

holds outside the union of a finite set of hyperplanes of K^{N+1}. Since for the height of the projective point $(w_0 : \ldots : w_N)$ we have the bound $H(w_0 : \ldots : w_N) = H(\mathbf{w}) \leqslant$

$H(w_0)\cdots H(w_N)$, we get from (1.3), (1.4) that outside the mentioned exceptional set,

$$1 \leqslant |\sigma|_v \cdot H(w_0)^{1+\varepsilon} \cdot H_S(w_0)^{N+1} \prod_{i=1}^{N} H(w_i)^{(N+2)\varepsilon}. \tag{1.5}$$

Since $\delta > (N+2)\varepsilon$, the bound (1.5) contradicts assumption (ii) of Theorem 4.

Therefore, almost all our $(N+1)$-tuples lie in the said exceptional set and for our purpose we deal from now on with infinitely many of them lying in a fixed hyperplane of equation

$$A_0 c_0 X_0 + \ldots + A_N c_N X_N = 0, \quad A_i \in \bar{K}.$$

Let V be a maximal vector subspace of $\bar{K}^{N+1}$ with the following property: there exists an infinite sequence $\mathcal{R}$ of the above $(N+1)$-tuples such that, for every $(x_0, \ldots, x_N) \in V$ and for every $\mathbf{w} = (w_0, \ldots, w_N) \in \mathcal{R}$, we have $x_0 c_0 w_0 + \cdots + x_N c_N w_N = 0$.

We may and will assume that $(A_0, \ldots, A_N) \in V$, so V in particular is nonzero. We proceed to verify the assumptions of Lemma 1 for V.

Let $\mathbf{x} = (x_0, \ldots, x_N) \in V$ be a nonzero vector and pick any index $i \in \{1, \ldots, N\}$. Suppose that x_i is different from both x_0 and 0.

Let $\mathbf{w} \in \mathcal{R}$. From the relation $x_0 c_0 w_0 + \cdots + x_N c_N w_N = 0$ we express $c_i w_i$ as a linear combination of $c_j w_j$, $j \neq i$. Substituting, we find

$$\sigma = \sigma(\mathbf{w}) = \sum_{j \in J} \left(1 - \frac{x_j}{x_i}\right) c_j w_j, \tag{1.6}$$

where J is the subset of $\{0, \ldots, N\}$ made up of those j for which $x_j \neq x_i$. Note that $0 \in J$ but $i \notin J$. Also, the terms in the right side of (1.6) are nonzero.

We want to apply the inductive assumption for Theorem 4 (with suitable new data) to (1.6).

Given $\mathbf{w} \in \mathcal{R}$, we let $\mathbf{w}'$ to be the vector in K^{*J} whose components are $w'_j = w_j$, for $j \in J$. Also, we let $c'_j = (1 - (x_j/x_i))c_j$, for $j \in J$.

We show that the assumptions (i), (ii) in the statement of Theorem 4 are verified, with c'_j in place of c_j, for the $\mathbf{w}'$ corresponding to the $\mathbf{w} \in \mathcal{R}$. In fact, (i) is trivial and (ii) follows from (1.6) and the inequality $\hat{H}(\mathbf{w}') \leqslant \hat{H}(\mathbf{w})$.

We observe now that also $\mathbf{w}'$ runs through an infinite sequence. In fact, if $\mathbf{w}'$ has finitely many possibilities for $\mathbf{w} \in \mathcal{R}$, then $\sigma(\mathbf{w})$ has finitely many possibilities as well, by (1.6). Therefore $\sigma(\mathbf{w})$ must vanish for almost all $\mathbf{w}$, since $|\sigma(\mathbf{w})|_v \to 0$, in view of assumption (ii). But this contradicts our initial assumption about vanishing subsums.

Since $\mathbf{w}'$ takes infinitely many values, as we have proved, we may conclude from the inductive assumption for Theorem 4, that for all $\mathbf{w}$ in an infinite subsequence of $\mathcal{R}$, some subsum (containing the term with $j = 0$) of the right side of (1.6) vanishes.

By taking a further infinite subsequence, we may in fact assume that the same subsum occurs for all elements of the subsequence.

This means that there exists a $J' \subset J$, with $0 \in J'$, such that for $\mathbf{w}'$ in the mentioned subsequence, we have $\sum_{j \in J'}(1 - (x_j/x_i))c_j w_j = 0$. We rewrite this as $\sum_{j \in J'}(x_i - x_j)c_j w_j = 0$. By the maximality of the space V (relative to all infinite sequences), we have that V contains the vector whose j-th coordinate is 0 if $j \notin J'$ and $x_i - x_j$ otherwise. Letting T be the truncation operator with support J' (see the definitions just before the statement of Lemma 1), we then have that $T(\mathbf{x}) - x_i T(\mathbf{u}) \in V$.

Therefore, the assumptions of Lemma 1 are true for V; hence, V contains a vector $(v_0, \dots, v_N)$ such that $v_0 = 1$ and $v_i \in \{0, 1\}$ for all i.

By the present definition of V, there exists an infinite sequence of $\mathbf{w} \in \mathcal{R}$ such that for all of them we have $v_0 c_0 w_0 + \dots + v_N c_N w_N = 0$. This relation however represents a vanishing subsum for all the $(N+1)$-tuples in an infinite subsequence of $\mathcal{R}$, contrary to our assumptions. □

2. Proof of Theorem 1 and Corollaries 1,2

The idea for the proof of Theorem 1 is as follows. By using a suitable partial sum of the Laurent series defining $f(z)$ we approximate the values $f(z_n)$ by a finite sum. The assumptions for Theorem 1 will then allow an application of Theorem 4, which will immediately give the conclusion.

We now go on with the details. Since the series for $f(z)$ converges in the unit disk, we have for $i > i_0$ an estimate $|a_i|_v \leqslant 2^i$. We write, for a given $M > d$,

$$f(z) = \sum_{-d \leqslant i < M} a_i z^i + r_M(z).$$

Therefore, for $M > i_0$ and $|z|_v < 1/4$, we have an inequality

$$|r_M(z)|_v \leqslant 2(2|z|_v)^M. \tag{2.1}$$

Since the z_n are distinct, we have in particular that $h(z_n) \to \infty$. Observe that assumption (2) implies that $|z_n|_v \to 0$ and in fact the stronger inequality

$$h(z_n) \leqslant C_1 \log |z_n|_v^{-1}, \tag{2.2}$$

for some positive constant C_1. By assumption (4) there exists a positive constant C_2 such that the inequality

$$h(f(z_n)) \leqslant C_2 h(z_n), \tag{2.3}$$

holds for all n in an infinite sequence. By considering only this sequence we may assume that (2.3) holds for all $n \in \mathbf{N}$.

We are going to apply Theorem 4.

We first choose an $M > i_0 + 2C_1C_2$ and define N as the number of nonzero terms among the a_i, for $-d \leqslant i < M$.

We put $c_0 = 1$ and we let $c_1, \dots, c_N$ to be the nonzero terms a_i, $-d \leqslant i < M$, in some order.

For a given n, we define $\mathbf{w} = \mathbf{w}_n$ by putting $w_0 = -f(z_n)$ and, for $i = 1, \dots, N$, $w_i = z_n^j$ if $c_i = a_j$.

Finally, we choose a positive ε and put $\delta = (N+3)\varepsilon$.

We can assume that $v \in S$. We proceed to verify assumptions (i), (ii) for Theorem 4, for large n, provided ε is small enough.

The present assumption (1) implies (i) of Theorem 4, for all large n.

To verify (ii), observe that, with the above conventions,

$$c_0 w_0 + \cdots + c_N w_N = -f(z_n) + \sum_{-d \leqslant i \leqslant M-1} a_i z_n^i = -r_M(z_n).$$

By the inequalities (2.1) and (2.2) we thus have

$$|c_0 w_0 + \cdots + c_N w_N|_v \leqslant 2(2|z_n|_v)^M \leqslant 2^{M+1} H(z_n)^{-\frac{M}{C_1}}. \tag{2.4}$$

On the other hand we have, by (2.3),

$$H(w_0) \leqslant H(z_n)^{C_2}, \tag{2.5}$$

and

$$\hat{H}(\mathbf{w}) = H(w_0) \prod_{i=-d}^{M-1} H(z_n^i) \leqslant H(z_n)^{C_2+d^2+M^2}. \tag{2.6}$$

Further, in view of assumption (3) of Theorem 1, we have for all large n,

$$H_S(w_0) \leqslant H(z_n)^{\varepsilon}. \tag{2.7}$$

Combining (2.5), (2.6), (2.7) we find

$$H(w_0) H_S^{N+1}(w_0) \hat{H}^{\delta}(\mathbf{w}) \leqslant H(z_n)^{C_2+(N+1)\varepsilon+(N+3)\varepsilon(C_2+d^2+M^2)}. \tag{2.8}$$

We now assume that ε is so small to ensure that

$$C_2 + (N+1)\varepsilon + (N+3)\varepsilon(C_2 + d^2 + M^2) \leqslant \frac{M}{2C_1}.$$

This will be possible since $M > 2C_1C_2$ and N is fixed. By (2.4) and (2.8) we then get

$$\begin{aligned} |c_0 w_0 + \cdots + c_N w_N|_v &\leqslant 2^{M+1} H(z_n)^{-\frac{M}{C_1}} \\ &\leqslant 2^{M+1} H(z_n)^{-\frac{M}{2C_1}} \Big(H(w_0) H_S^{N+1}(w_0) \hat{H}^{\delta}(\mathbf{w}) \Big)^{-1}. \end{aligned}$$

For large n we have $2^{M+1} H(z_n)^{-\frac{M}{2C_1}} < 1$ and (ii) for Theorem 4 follows.

Theorem 4 now implies that for all large n some subsum of $c_0 w_0 + \cdots + c_N w_N$ involving w_0 vanishes. This means that $f(z)$ coincides with a fixed Laurent polyno-

https://doi.org/10.1023/A:1015594913393 Published online by Cambridge University Press

mial at the points $z = z_n$ for all n in a suitable infinite sequence. Since the z_n are distinct and converge to 0, Theorem 1 follows. □

Proof of Corollary 1. This is a simple application of Theorem 1. By enlarging K, we may assume that $q \in K$. We enlarge S so that q is an S-unit. In Theorem 1, let us put $z_n = q^n$. Assumptions (1), (2) are trivially verified, and so is (3) if $f(q^n)$ is an S-integer. If the conclusion of Corollary 1 is not true, then condition (4) for Theorem 1 would also be verified. But then f would be a Laurent polynomial, against the assumption. □

Proof of Corollary 2. We apply Corollary 1 with v equal to the absolute value on $\mathbf{C}$. We take $K = \mathbf{Q}$ and $S = \{\infty\}$. We have $|f(q^n)| = \mathrm{O}(|q|^{-nd})$, where d is the order of the pole of f at 0. Therefore, if $f(q^n)$ is a nonzero rational integer we have

$$h(f(q^n)) = \log|f(q^n)| = \mathrm{O}(n).$$

In conclusion, by Corollary 1, either f is a Laurent polynomial or the set of positive integers n with that property is finite. □

3. Proof of Theorem 2

We work only in the case $|\alpha_1|_v > 1$ and consider the Puiseux expansions at $z = \infty$ of the solutions $X = X(z)$ of $g(z, X) = 0$. The arguments in the case $|\alpha_1|_v < 1$ are completely analogous and use the expansions at $z = 0$.

For large n, any solution x_n of $g(z_n, X) = 0$ will be given by some Puiseux expansion, considered v-adically; this is because $|z_n|_v \to \infty$. We may assume that for all n in an infinite sequence $\mathcal{R}$ the same expansion occurs, so we have, for the v-adic convergence,

$$x_n = \sigma_p z_n^{\frac{p}{e}} + \sigma_{p-1} z_n^{\frac{p-1}{e}} + \cdots, \tag{3.1}$$

for some determination of the e-th root and some algebraic numbers σ_i, $i \leqslant p$. It is well known that the coefficients σ_i in fact lie in a fixed number field.

In the sequel $C_1, C_2, \ldots$ will denote positive numbers depending only on $g(Z, X)$ and on the c_i, α_i.

Since $|\alpha_1|_v > |\alpha_i|_v$ for $i = 2, \ldots, h$, we have binomial expansions

$$z_n^{\frac{j}{e}} = c_1^{\frac{j}{e}} \alpha_1^{\frac{jn}{e}} \left(1 + \sum_{i=2}^{h} \frac{c_i}{c_1}\left(\frac{\alpha_i}{\alpha_1}\right)^n\right)^{\frac{j}{e}} = c_1^{\frac{j}{e}} \alpha_1^{\frac{jn}{e}} \sum_{r=0}^{\infty} \binom{\frac{j}{e}}{r} \left(\sum_{i=2}^{h} \frac{c_i}{c_1}\left(\frac{\alpha_i}{\alpha_1}\right)^n\right)^r,$$

for some choice of the eth roots of c_1 and α_1, which we may assume to be fixed for all $n \in \mathcal{R}$.

https://doi.org/10.1023/A:1015594913393 Published online by Cambridge University Press

Combining (3.1) with this expansion it is easy to see that we may write, in the v-adic convergence,

$$x_n = \sum_{j=1}^{\infty} \tau_j \gamma_j^n, \qquad n \in \mathcal{R}, \tag{3.2}$$

where the $\tau_j \in \bar{K}$ and the γ_j are distinct and lie in the multiplicative group generated by $\alpha_1^{\frac{1}{e}}$ and $\alpha_2, \ldots, \alpha_h$. Also, it is easy to see that the γ_j tend v-adically to zero.

Suppose first that the series on the right of (3.2) is not a finite sum. Then we assume that no τ_j is zero and that the γ_j are written in decreasing order, i.e. $|\gamma_1|_v \geqslant |\gamma_2|_v \geqslant \cdots > 0$.

Now, both the binomial expansion and the series on the right of (3.1) converge absolutely for all large n. It follows that we may write

$$\sum_{j=1}^{\infty} |\tau_j|_v |\gamma_j|_v^{C_1} \leqslant C_2. \tag{3.3}$$

We are going to apply Theorem 4, after approximating x_n by a finite sum extracted from (3.2). We enlarge K at once and assume that it contains all the $\alpha_i^{\frac{1}{e}}$ and all the coefficients σ_j in the Puiseux series. In particular, we may assume that K contains all the τ_j, γ_j.

We estimate the tails of the series on the right of (3.2). We have, for $N \geqslant 0$

$$\sum_{j=N+1}^{\infty} |\tau_j|_v |\gamma_j|_v^n = \sum_{j=N+1}^{\infty} |\tau_j|_v |\gamma_j|_v^{n-C_1} |\gamma_j|_v^{C_1} \leqslant C_2 |\gamma_{N+1}|_v^{n-C_1}, \tag{3.4}$$

where we have used (3.3).

For later purposes we need an estimate of $H(x_n)$. We derive it from the equation $g(z_n, x_n) = 0$. Observe that an estimate $H(z_n) \leqslant C_3 C_4^n$ follows immediately from (0.1). On the other hand, we can estimate the height of the roots of an equation in terms of the heights of the coefficients. We finally obtain

$$H(x_n) \leqslant C_5 C_6^n. \tag{3.5}$$

We choose N so that

$$|\gamma_{N+1}|_v C_6 < 1. \tag{3.6}$$

Also, we choose a finite S so that it contains v and all Archimedean absolute values of K. Moreover we require that all the c_j, α_j, all the nonzero coefficients of $g(Z, X)$ and $\tau_1, \ldots, \tau_N$ are S-units. In particular, with this choice all the γ_j are S-units. Also, the z_n are S-integers and $g(Z, X)$ is monic in X; therefore, the x_n too are S-integers, in view of the equations $g(z_n, x_n) = 0$.

We shall apply Theorem 4 with $c_i = 1$ for $i = 0, \ldots, N$. We put $w_0 = -x_n$ and, for $i = 1, \ldots, N$, we put $w_i = \tau_i \gamma_i^n$. If $w_0 = -x_n = 0$ for infinitely many $n \in \mathbf{N}$, then

https://doi.org/10.1023/A:1015594913393 Published online by Cambridge University Press

$g(z_n, 0) = 0$ for such values, whence $g(Z, 0) = 0$ identically, since $z_n \to \infty$. In this case the conclusion of the theorem is verified by choosing $k = 0$.

Therefore we suppose from now on that $x_n \neq 0$ for all $n \in \mathcal{R}$. Observe that w_0 is an S-integer and that, due to our choice of S, the w_i are S-units for $i \geqslant 1$. In particular, assumption (i) for Theorem 4 is verified for any choice of $\varepsilon > 0$.

We proceed to verify assumption (ii) for all large $n \in \mathcal{R}$, provided we choose a small enough ε and put $\delta = (N+3)\varepsilon$.

By using (3.2) and (3.4) we find

$$|w_0 + w_1 + \cdots + w_N|_v = | - x_n + \sum_{i=1}^{N} \tau_i \gamma_i^n|_v \leqslant C_2 |\gamma_{N+1}|_v^{n-C_1}. \tag{3.7}$$

To compare this estimate with the right side of (ii) of Theorem 4, we first observe that $H_S(w_0) = 1$, since $w_0 = -x_n$ is an S-integer. Moreover, in view of (3.5) we may write

$$\hat{H}(\mathbf{w}) = H(x_n) H(\tau_1 \gamma_1^n) \cdots H(\tau_N \gamma_N^n) \leqslant C_5 C_6^n \prod_{i=1}^{N} H(\tau_i) \prod_{i=1}^{N} H(\gamma_i^n) \leqslant AB^n, \tag{3.8}$$

where we may take $A = C_5 H(\tau_1) \ldots H(\tau_N)$ and $B = C_6 H(\gamma_1) \ldots H(\gamma_N)$. We observe that A, B depend on N but not on n.

We now choose ε so that

$$|\gamma_{N+1}|_v C_6 B^\delta < 1. \tag{3.9}$$

This will be possible for small ε in view of (3.6), recalling our choice $\delta = (N+3)\varepsilon$.

In view of (3.7), (3.5) and (3.8), the verification of (ii) of Theorem 4 will follow from

$$C_2 |\gamma_{N+1}|_v^{n-C_1} < (C_5 C_6^n)^{-1} (AB^n)^{-\delta},$$

which is the same as

$$(|\gamma_{N+1}|_v C_6 B^\delta)^n < (C_2 C_5 A^\delta)^{-1} |\gamma_{N+1}|_v{}^{C_1}.$$

However, this latter inequality follows from (3.9) for large n.

Therefore, by Theorem 4 we may conclude that for all large $n \in \mathcal{R}$ some subsum of $w_0 + w_1 + \cdots + w_N$, involving w_0, vanishes. Going to an infinite subsequence $\mathcal{R}'$ of $\mathcal{R}$, we may assume that for all $n \in \mathcal{R}'$ the subsum corresponds to the same set of terms. Namely, there exists a set $I \subset \{1, \ldots, N\}$, such that $x_n = \sum_{i \in I} \tau_i \gamma_i^n$, $n \in \mathcal{R}'$. Recall that we are assuming that the right side of (3.2) has infinitely many terms. If this is not the case, however, our last conslusion is automatic.

https://doi.org/10.1023/A:1015594913393 Published online by Cambridge University Press

We change notation and write $\sum_{i\in I}\tau_i\gamma_i^n=\sum_{i=1}^k d_i\beta_i^n$; plugging into $g(z_n,x_n)=0$ we find

$$g\left(\sum_{i=1}^h c_i\alpha_i^n,\sum_{j=1}^k d_j\beta_j^n\right)=0,\quad \text{for } n\in\mathcal{R}'. \tag{3.10}$$

But now the Skolem–Lech–Mahler Theorem (see, e.g., [vdP]) implies that the left side of (3.10) vanishes for all n in a suitable arithmetic progression. This concludes the proof.

Remark. It may be shown (and also follows from the proof) that the d_i, β_i and a relevant arithmetic progression may be found effectively. Examples like $z_n=2^n+1-(-1)^n$, $g(Z,X)=Z-X^2$, show that in general we cannot expect that the conclusion holds for all $n\in\mathbf{N}$. This will be possible however under suitable additional conditions on the *roots* α_i of the recurrence sequence z_n; for instance it is easy to obtain the stronger conclusion when the α_i are multiplicatively independent.

4. Proof of Theorem 3 and its Corollaries

To prove Theorem 3, we start by verifying that the series defining γ in fact converges (absolutely) in the v-adic topology. Note that the assumptions imply that $h(a_n)=o(m_n)$. In particular, for large n we have $h(a_n)<(\log|\alpha|_v^{-1}/2))m_n$. Therefore

$$|a_n|_v\leqslant H(a_n)<|\alpha|_v^{-\frac{m_n}{2}},$$

concluding the argument.

Assuming that γ is algebraic, we shall apply Theorem 4 with the following data. We first enlarge K so that $\gamma\in K$. Then we choose S such that it contains v, the Archimedean places of K and such that α is an S-unit. The integer N already appears in the statement of Theorem 3.

For an integer $n\in\mathcal{N}$ we put $w_0=-\gamma+\sum_{j=1}^n a_j\alpha^{m_j}$ and, for $i=1,\dots,N$, we put $w_i=a_{i+n}\alpha^{m_{i+n}}$.

Finally, we choose $c_i=1$ for $i=0,1,\dots,N$.

We first estimate $H(w_0)$ and for this purpose we note that for every place v of K we have

$$\begin{aligned}|w_0|_v&\leqslant \max(1,|n+1|_v)\cdot\max(|\gamma|_v,|a_1\alpha^{m_1}|_v,\dots|a_n\alpha^{m_n}|_v)\\&\leqslant \max(1,|n+1|_v)\cdot\max(|\gamma|_v,|a_1|_v,\dots|a_n|_v)\cdot\max(1,|\alpha|_v)^{m_n}.\end{aligned}\tag{4.1}$$

From this inequality we obtain, on taking the product of $\max(1,|w_0|_v)$ over all places,

$$H(w_0)\leqslant(n+1)H(\gamma)H(a_1)\dots H(a_n)H(\alpha)^{m_n}. \tag{4.2}$$

https://doi.org/10.1023/A:1015594913393 Published online by Cambridge University Press

From (4.1) we also derive, taking into account that α is an S-unit, that

$$H_S(w_0) \leqslant H(\gamma)H(a_1)\dots H(a_n). \tag{4.3}$$

For the w_i, $i = 1, 2, \dots, N$, we have

$$H(w_i) \leqslant H(a_{i+n})H(\alpha)^{m_{i+n}},$$

whence

$$\hat{H}(\mathbf{w}) \leqslant (n+1)H(\gamma)H(a_1)\dots H(a_{n+N})H(\alpha)^{m_n+\dots+m_{n+N}}. \tag{4.4}$$

Let $\varepsilon < 1/2$ be a positive real number (which will be specified later). From the assumption $h(a_r) = o(m_r)$ we deduce the inequality $|a_r|_v < |\alpha|_v^{-\varepsilon m_r}$ valid for $r > r_0(\varepsilon)$. Therefore, for $n > r_0(\varepsilon)$,

$$\begin{aligned} |w_0 + \dots + w_N|_v &\leqslant \sum_{r>n+N} |a_r|_v |\alpha|_v^{m_r} \leqslant \sum_{r>n+N} |\alpha|_v^{m_r(1-\varepsilon)} \\ &\leqslant |\alpha|_v^{m_{n+N+1}(1-\varepsilon)} \sum_{s=0}^{\infty} |\alpha|_v^{s(1-\varepsilon)} \leqslant C_1 |\alpha|_v^{m_{n+N+1}(1-\varepsilon)}, \end{aligned} \tag{4.5}$$

where we can take $C_1 = \sum_{s=0}^{\infty} |\alpha|_v^{s/2}$, since $\varepsilon < 1/2$.

We are going to apply Theorem 4 with $\delta = (N+3)\varepsilon$.

We first verify assumption (i) for every choice of ε, provided we take n to be large enough with respect to ε. In fact, for $i = 1, \dots, N$ we have $h_S(w_i) = h_S(a_i)$ (resp. $h_S(w_i^{-1}) = h_S(a_i^{-1})$). Hence, $h_S(w_i) + h_S(w_i^{-1}) \leqslant 2h(a_i)$. On the other hand $h(w_i) \geqslant m_i h(\alpha) - h(a_i)$. The conclusion follows since $h(a_i) = o(m_i)$.

To verify (ii) we shall compare the estimate for its left side, given by (4.5), with an estimate for its right side given by (4.2), (4.3) and (4.4) above. These latter inequalities give, after a short calculation,

$$\begin{aligned} &H(w_0)H_S(w_0)^{N+1}\hat{H}(\mathbf{w})^{\delta} \\ &\quad \leqslant (n+1)^{1+\delta}(H(\gamma)H(a_1)\dots H(a_n))^{N+2+\delta} \times \\ &\quad \times (H(a_{n+1})\dots H(a_{n+N}))^{\delta} H(\alpha)^{m_n + \delta\sum_{j=1}^{N} m_{n+j}}. \end{aligned}$$

We assume that n is so large that $H(a_1)\dots H(a_r) < 2^{\varepsilon m_r}$ for all $r \geqslant n$. We also assume that $\delta = (N+3)\varepsilon < 1$ and that $(n+1)^2 < 2^{\varepsilon} m_n$. Using these bounds in the last displayed inequality we get

$$H(w_0)H_S(w_0)^{N+1}\hat{H}(\mathbf{w})^{\delta} < H(\gamma)^{N+3} 2^{\varepsilon(N+4)m_n} 2^{\delta\varepsilon m_{n+N}} H(\alpha)^{m_n + \delta N m_{n+N}}.$$

Finally, for small enough ε we see that we have the bound

$$H(w_0)H_S(w_0)^{N+1}\hat{H}(\mathbf{w})^{\delta} < H(\alpha)^{m_n + C_2\varepsilon N m_{n+N}}, \tag{4.6}$$

where C_2 depends only on α and γ. We observe that this bound holds provided ε is small enough with respect to N and provided n is sufficiently large with respect to ε.

We have not yet used the fact that $n \in \mathcal{N}$. We shall exploit this in the comparison of (4.5) and (4.6). In view of these inequalities the verification of (*ii*) of Theorem 4 amounts to the following:

$$C_1 |\alpha|_v^{m_{n+N}\left(1-\varepsilon+\frac{h(\alpha)}{\log|\alpha|_v}C_2 N\varepsilon\right)-m_n\frac{h(\alpha)}{\log|\alpha|_v^{-1}}} < 1. \tag{4.7}$$

Since $n \in \mathcal{N}$, we have $m_{n+N} > Lm_n$, whence the exponent of $|\alpha|_v$ in (4.7) is at least

$$m_n\left(L\left(1-\varepsilon+\frac{h(\alpha)}{\log|\alpha|_v}C_2 N\varepsilon\right)-\frac{h(\alpha)}{\log|\alpha|_v^{-1}}\right).$$

For small ε this exceeds λm_n, for a suitable positive λ (depending on α, N, ε). Therefore (4.7) holds for all large $n \in \mathcal{N}$.

From Theorem 4 we deduce that, for all but a finite number of $n \in \mathcal{N}$, some subsum of type $w_0 + \sum_{i\in I} w_i$ vanishes, where $I = I_n$ is a subset of $\{1, \ldots, N\}$ depending on $n \in \mathcal{N}$. Putting $\mathcal{A}_n = \{1, \ldots, n\} \cup \{n + i\colon i \in I_n\}$, we get the conclusion of Theorem 3. □

Proof of Corollary 3. Let α be a real algebraic number in (0, 1). We have to show that $\gamma := \sum_{i=1}^{\infty} a_i\alpha^{m_i}$ is transcendental. (That the series is convergent is proved in Theorem 3.)

We enlarge K to contain α. The field $\mathbf{Q}(\alpha, a_1, a_2, \ldots)$ is implicitly embedded in $\mathbf{R}$. We are going to apply Theorem 3 taking v to be any extension to K of the corresponding real valuation.

Put $L := 2h(\alpha)/\log|\alpha|_v^{-1}$. By the assumptions, there exists N such that the sequence $\mathcal{N}$ of integers n satisfying $m_{n+N} > Lm_n$ is infinite.

Assuming by contradiction that γ is algebraic, the conclusion of Theorem 3 implies that γ is equal to a finite partial sum of the defining series. This is however impossible since all the terms in the series are positive. □

The proof of Corollary 3′ is similar. We have only to observe that no sum $\sum_{i\in I} a_i x^{m_i}$ can vanish if $|x|_v < 1$ and I is nonempty; in fact, all the terms which appear have pairwise distinct absolute values.

Proof of Corollary 4. Let α, v be as in the statement. We start with a lemma, which is significant only in the Archimedean case.

LEMMA 2. *There is a number C_1 depending only on α with the following property. Suppose that $b_1 < b_2 < \ldots$ are positive integers, that $\sum_{i=1}^{\infty} \alpha^{b_i} = 0$ and that, for an integer r, $\sum_{i=1}^{s} \alpha^{b_i} \neq 0$ for $s = 1, 2, \ldots, r$. Then $b_{r+1} - b_1 < r!C_1^r$.*

https://doi.org/10.1023/A:1015594913393 Published online by Cambridge University Press

In the sequel $C_2, C_3, \dots$ denote positive numbers depending only on α.

First we have, for $s \geqslant 1$,

$$\left|\sum_{i=1}^{s} \alpha^{b_i}\right|_v = \left|\sum_{i=s+1}^{\infty} \alpha^{b_i}\right|_v \leqslant C_2|\alpha|_v^{b_{s+1}},$$

where we may take $C_2 = \sum_{i=0}^{\infty} |\alpha|_v^i = (1 - |\alpha|_v)^{-1}$.

For $s = 1$ this gives $|\alpha|_v^{b_1 - b_2} \leqslant C_2$, whence $b_2 - b_1 \leqslant C_3$.

For general s we get

$$\left|\sum_{i=1}^{s} \alpha^{b_i - b_1}\right|_v \leqslant C_2|\alpha|_v^{b_{s+1}-b_1}.$$

For $s \leqslant r$ the left side does not vanish. Therefore, by Liouville's inequality we have

$$-h\left(\sum_{i=1}^{s} \alpha^{b_i - b_1}\right) \leqslant -C_4(b_{s+1} - b_1) + C_5.$$

On the other hand we have $h(\sum_{i=1}^{s} \alpha^{b_i - b_1}) \leqslant sC_6(b_s - b_1)$. In conclusion

$$b_{s+1} - b_1 \leqslant sC_7(b_s - b_1) + C_8, \qquad s = 1, 2, \dots, r.$$

Iterating we have the result.

We now prove Corollary 4. Assume that $\gamma := \sum_{i=1}^{\infty} \alpha^{m_i}$ is algebraic. Let L be a real number between $h(\alpha)/\log|\alpha|_v^{-1}$ and $\limsup_n m_{n+N}/m_n$. Then the sequence $\mathcal{N}$ of Theorem 3 is infinite. We apply Theorem 3 and take $n \in \mathcal{N}$ so large that $\gamma = \sum_{i \in \mathcal{A}_n} \alpha^{m_i}$ and $m_{s+h} - m_s > h!C_1^h$ for $s \geqslant n$, where C_1 is as in Lemma 2 and h is as in the statement of the present corollary.

We denote by $\{a_1 < a_2 < \dots\}$ the sequence of m_j, for j varying over the complement of $\mathcal{A}_n$ in $\mathbf{N}$. Then $\sum_{i=1}^{\infty} \alpha^{a_i} = 0$. We apply Lemma 2 with $b_i = a_i$ for all i. Since $b_{h+1} - b_1 > h!C_1^h$ by construction, we deduce that some subsum $\alpha^{a_1} + \dots + \alpha^{a_s} = 0$, where $s \leqslant h$.

Therefore $\sum_{i=s+1}^{\infty} \alpha^{a_i} = 0$. We put $b_i = a_{i+s}$ and repeat the procedure, and so on. We plainly obtain sets $\mathcal{B}_i$ satisfying the conclusion. $\square$

Proof of Corollary 5. We may plainly suppose that $m_{i+1} > \lambda m_i$ for $i = 1, 2, \dots$, for some $\lambda > 1$. This gives in particular $m_{i+N}/m_i > \lambda^N$ for all i. Therefore, the sequence $\mathcal{N}$ of Theorem 3 is infinite if N is sufficiently large.

We fix such an N and we apply Theorem 3, assuming that $\gamma := \sum a_i\alpha^{m_i}$ is algebraic. Choose a positive $\varepsilon < \min((\lambda - 1)/(\lambda + 1), \frac{1}{2})$ and let $\mathcal{A}_n$ be as in the conclusion of Theorem 3, with $n \in \mathcal{N}$ large enough with respect to ε. In view of Liouville's inequality and the assumption $h(a_r) = o(m_r)$, we may write for large n,

$$\max(|a_r|_v, |a_r|_v^{-1}) \leqslant |\alpha|_v^{-\varepsilon m_r}, \qquad \forall r \geqslant n. \tag{4.8}$$

https://doi.org/10.1023/A:1015594913393 Published online by Cambridge University Press

The conclusion of Theorem 3 gives

$$\sum_{i \notin \mathcal{A}_n} a_i \alpha^{m_i} = 0. \tag{4.9}$$

Define b as the minimal integer not in $\mathcal{A}_n$. In particular, we have $b > n$. Equations (4.9) and (4.8) give

$$\begin{aligned} |a_b \alpha^{m_b}|_v &\leqslant \sum_{i=1}^{\infty} |a_{b+i}|_v |\alpha^{m_{b+i}}|_v \leqslant \sum_{i=1}^{\infty} |\alpha|_v^{m_{b+i}(1-\varepsilon)} \\ &\leqslant |\alpha|_v^{m_{b+1}(1-\varepsilon)} \sum_{i=1}^{\infty} |\alpha|_v^{(m_{b+i}-m_{b+1})(1-\varepsilon)} \leqslant |\alpha|_v^{m_{b+1}(1-\varepsilon)} \frac{1}{1-|\alpha|_v^{1-\varepsilon}} \\ &\leqslant C_9 |\alpha|_v^{m_{b+1}(1-\varepsilon)}, \end{aligned}$$

where we can take $C_9 = 1/(1-\sqrt{|\alpha|_v})$. Using $m_{b+1} > \lambda m_b$ we obtain $|a_b|_v \leqslant C_9 |\alpha|_v^{m_b(\lambda-1-\lambda\varepsilon)}$, and, comparing with (4.8) (recall that $b \geqslant n$) we finally get $1 \leqslant C_9 |\alpha|_v^{m_b(\lambda-1-\lambda\varepsilon-\varepsilon)}$. Since $\lambda - 1 - \lambda\varepsilon - \varepsilon > 0$ and $m_b \geqslant b > n$ this inequality cannot hold for large n. Therefore γ must be transcendental. $\square$

References

[CZ1] Corvaja, P. and Zannier, U.: Diophantine equations with power sums and universal Hilbert sets, *Indag. Math.* **9** (1998), 317–332.

[CZ2] Corvaja, P. and Zannier, U.: On the diophantine equation $f(a^m, y) = b^n$, *Acta Arith.* **94**(1) (2000), 25–40.

[D] Dèbes, P.: On the irreducibility of the polynomials $P(t^m, Y)$, *J. Number Theory* **42** (1992), 141–157.

[DZ] Dèbes, P. and Zannier, U.: Universal Hilbert subsets, *Math. Proc. Cambridge Philos. Soc.* **124** (1998), 127–134.

[E] Evertse, J.-H.: On sums of S-units and linear recurrences, *Compositio Math.* **53** (1984), 225–244.

[L1] Laurent, M.: Equations exponentielles-polynômes et suites récurrentes linéaires II, *J. Number Theory* **31** (1989), 24–53.

[L2] Laurent, M.: Equations exponentielles polynômes et suites récurrentes linéaires, *Astérisque* **147–148** (1987), 121–139, 343–344.

[Lo-vdP] Loxton, J. H. and van der Poorten, A.: Transcendence and algebraic independence by a method of Mahler, In: Baker, Masser (eds), *Transcendence Theory*, Academic Press, New York, 1977.

[M] Mahler, K.: Arithmetische Eigenschaften der Lösungen einer Klasse von Funktionalgleichungen, *Math. Ann.* **101** (1929), 342–366.

[Ma] Masser, D.: Algebraic independence properties of the Hecke–Mahler seires, *Quart. J. Math.* **50** (1999), 207–230.

[N] Nishioka, K.: Algebraic independence by Mahler's method and S-unit equations, *Compositio Math* **92** (1994), 87–110.

[vdP] van der Poorten, A. J.: Some facts that should be better known, especially about rational functions, in *Number Theory and Applications* (Banff, 1988), Kluwer Acad. Publ., Dordrecht, 1989, pp. 497–528.

[vdP-S] van der Poorten, A. J. and Schlickewei, H. P.: The growth conditions for recurrence sequences, Macquarie Univ. Math. Rep. 82–0041, North Ryde, Australia, 1982.

[Sc] Schmidt, W. M.: *Diophantine Approximations and Diophantine Equations*, Lecture Notes in Math. 1467, Springer, New York, 1991.

[T] Tanaka, T.: Algebraic independence results related to linear recurrences, *Osaka J. Math.* **36** (1999), 203–227.