Ann. Scient. Éc. Norm. Sup.,
4e série, t. 38, 2005, p. 76 à 92.

S-UNIT POINTS ON ANALYTIC HYPERSURFACES

BY PIETRO CORVAJA AND UMBERTO ZANNIER

ABSTRACT. – In analogy with algebraic equations with S-units, we shall deal with S-unit points in an *analytic* hypersurface, or more generally with values of analytic functions at S-unit points.

After proving a general theorem, we shall give diophantine applications to certain problems of integral points on subvarieties of $\mathbf{A}^1 \times \mathbf{G}_m^n$. Also, we shall prove an analogue of a theorem of Masser, important in Mahler's method for transcendence.

In the course of the proofs we shall also develop a theory for those algebraic subgroups of $\mathbf{G}_m^n$ whose Zariski closure in $\mathbf{A}^n$ contains the origin. Among others, we shall prove a structure theorem for the family of such subgroups contained in a given analytic hypersurface, obtaining conclusions similar to the case of algebraic varieties.

RÉSUMÉ. – En analogie avec les équations algébriques en S-unités, on considère ici des points S-unités sur une hypersurface *analytique*. On dérive d'un énoncé général des applications à des problèmes sur les points entiers des sous-variétés de $\mathbf{A}^1 \times \mathbf{G}_m^n$. En plus, on déduit un analogue d'un lemme de zéros de Masser, qui intervient dans la méthode de Mahler en transcendance.

Lors de la démonstration de ces résultats, on développe une théorie des sous-groupes algébriques de $\mathbf{G}_m^n$ dont la clôture de Zariski dans $\mathbf{A}^n$ contient l'origine. En particulier, on démontre un théorème de structure pour la famille de tels sous-groupes contenus dans une hypersurface analytique, en obtenant une conclusion analogue à celle du cas des variétés algébriques.

1. Introduction

Algebraic equations with S-units are now rather classical: we may recall for instance the S-unit equation $X_1 + \cdots + X_n = 1$ dealt with by Evertse and van der Poorten–Schlickewei, leading to Laurent's solution [5] of a conjecture by Lang; in particular this implies that the S-unit points (over a number field k) in a subvariety V of $\mathbf{G}_m^n$ (which are the S-integral points for V) must lie in a finite number of algebraic cosets entirely contained in V.

Here we shall deal with S-unit points in an *analytic* hypersurface, or more generally with values at S-unit points of analytic functions. (See Theorem 1 below.)

Motivations for this study come from several sources.

First, there is the case of power series in several variables, representing algebraic functions, evaluated at S-unit points; the problem is now to establish whether the values are rationals or S-integers in a number field. This case for instance occurs in the problem of studying the S-integral points on subvarieties of $\mathbf{A}^1 \times \mathbf{G}_m^n$; here no general analogue of Laurent's theorem is known, even in simply described cases, like the equation $y^2 = 1 + 2^m + 3^n$. We shall approach some of these questions. (See corollary below.)

Then, we may recall, e.g., the "Vanishing Theorem" of D. Masser [6, Theorem], important in the applications of Mahler's method in transcendence. That result concerns precisely analytic functions vanishing on certain special sequences of S-unit points. Here, with completely different methods, we have conclusions in the same direction. In a way they are less general than Masser's (for instance we require algebraic coefficients for the series in question), but on the other hand we can deal with sequences of S-unit points which have no special structure, but satisfy only certain growth conditions. (See Theorem 3 below.)

As in the conclusions of Laurent's theorems, the S-unit points under consideration are found to lie in a finite union of *algebraic* cosets in $\mathbf{G}_m^n$, entirely contained in the hypersurfaces in question. (In a different language, this represents also conclusion (ii) in Masser's Theorem in [6].)

Here, the algebraic cosets important for us are those containing the origin in their Zariski closure in $\mathbf{A}^n$ (see the definitions below). We have a structure theorem for these cosets (see Theorem 2), which represents the analogue of what is known in the algebraic case (see [8]); however we have found the analysis rather more delicate in the present case of analytic varieties.

A study of similar problems in the one-variable case had been started in [3]; the conclusions were much simpler, due among others to the fact that the proper cosets of $\mathbf{G}_m$ are just points. The higher dimensional situation presents several new features.

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 v of K we denote by $|\cdot|_v$ 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 $h(x) = \sum_v \log^+ |x|_v$ [1] and the product formula $\prod_v |x|_v = 1$ 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 ν of K and denote by $\mathbf{C}_\nu$ a completion of an algebraic closure of K_ν. Our notion of convergence, unless otherwise specified, refers to $\mathbf{C}_\nu$.

We also define the S-height of a non zero element $x \in K^\times$ to be

$$h_S(x) = \sum_{v \notin S} \log^+ |x|_v.$$

For S-integers this height vanishes, so it gives a measure of "how far" x is from being an S-integer.

For a vector $\mathbf{z} = (z_0, z_1, \ldots, z_h) \in K^{h+1} \setminus \{0\}$ ($h \geqslant 1$), we define $h(\mathbf{z})$ as the usual projective logarithmic height. We put $h(0) := h((0:1)) = 0$. 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 v, $||\mathbf{z}||_v := \max\{|z_0|_v, \ldots, |z_h|_v\}$.

Throughout, we put $H(\cdot) = \exp(h(\cdot))$ and $H_S(\cdot) = \exp(h_S(\cdot))$.

We shall consider power series $\sum_{\mathbf{i}} a_{\mathbf{i}} \mathbf{X}^{\mathbf{i}} \in \mathbf{C}_\nu [\![X_1, \ldots, X_n]\!]$ where, as usual, for a multiindex $\mathbf{i} = (i_1, \ldots, i_n)$, we put $\mathbf{X}^{\mathbf{i}} := X_1^{i_1} \cdots X_n^{i_n}$, $\|\mathbf{i}\| := \max\{i_1, \ldots, i_n\}$. We shall normally assume that the relevant series are convergent in a neighborhood of the origin in $\mathbf{C}_\nu^n$.

Usually we shall consider points $\mathbf{x}$ with nonzero coordinates in some field K, and view them as points of $\mathbf{G}_m^n(K)$, so operations are intended to be made coordinatewise.

We shall use the familiar "little o" and "big O" notations.

[1] As usual, $\log^+ t = \max(0, \log t)$ for a $t > 0$.

Our main result is the following theorem, where by "coset" $\mathbf{u}H$ we refer to a coset for some connected algebraic subgroup H of $\mathbf{G}_m^n$. In Section 2 we shall briefly recall standard notation and basic facts about the corresponding theory.

THEOREM 1. – *Let $f(\mathbf{X}) = \sum_{\mathbf{i}} a_{\mathbf{i}} \mathbf{X}^{\mathbf{i}}$ be a power series with algebraic coefficients in $\mathbf{C}_\nu$ converging in a neighborhood of the origin in $\mathbf{C}_\nu^n$. Let S be a finite set of absolute values of K containing the archimedean ones. Let $\mathbf{x}_h = (x_{h1}, \ldots, x_{hn})$ $(h = 1, 2, \ldots)$ be a sequence in $K^{\times n}$, tending to zero in K_ν^n and such that $f(\mathbf{x}_h)$ is defined and belongs to K. Suppose that*:

(1) *For $i = 1, \ldots, n$, we have $h_S(x_{hi}) + h_S(x_{hi}^{-1}) = \mathrm{o}(h(x_{hi}))$ as $h \to \infty$.*

(2) $\hat{h}(\mathbf{x}_h) = \mathrm{O}(-\log(\max_i |x_{hi}|_\nu))$.

(3) $h_S(f(\mathbf{x}_h)) = \mathrm{o}(h(\mathbf{x}_h))$.

(4) $h(f(\mathbf{x}_h)) = \mathrm{O}(h(\mathbf{x}_h))$.

Then there exist a finite number of cosets $\mathbf{u}_1 H_1, \ldots, \mathbf{u}_r H_r \subset \mathbf{G}_m^n$ such that $\{\mathbf{x}_h\}_{h\in\mathbf{N}} \subset \bigcup_{i=1}^r \mathbf{u}_i H_i$ and such that, for $i = 1, \ldots, r$, the restriction of $f(\mathbf{X})$ to $\mathbf{u}_i H_i$ coincides with a polynomial in $K[\mathbf{X}]$.

Here, as in the sequel, the "restriction" may be interpreted both formally and in the sense of functions near the origin.

We pause on the meaning of the cumbersome conditions in the theorem. Condition (1) states that the coordinates of $\mathbf{x}_h$ "tend" to be S-units. (In fact, the vanishing of both $h_S(\mathbf{x})$ and $h_S(\mathbf{x}^{-1})$ characterizes S-units.) In most of our applications (e.g. Theorem 3 below) we shall deal with S-units, so (1) will be automatic.

Condition (2) is crucial and states that the ν-adic contribution to the height of $\mathbf{x}_h$ is not negligible for $h \to \infty$, for *each* coordinate x_{hi}.

In analogy with condition (1), conditions (3) and (4) mean that the values $f(\mathbf{x}_h)$ tend to be S-integers with not too large height.

Theorem 1, involved as it is, yields nevertheless some diophantine conclusions, as for instance the following corollary on sums of S-units which are perfect dth powers.

COROLLARY. – *Let $d \in \mathbf{N}$, $\delta > 0$. Let Σ be a set of points $\mathbf{x} = (x_1, \ldots, x_n) \in (\mathcal{O}_S^*)^n$ such that*:

(i) $|x_1|_\nu \geqslant (\max_{j\geqslant 2} |x_j|_\nu) H(\mathbf{x})^\delta$.

(ii) *There exists $y = y_{\mathbf{x}} \in K$ with $x_1 + \cdots + x_n = y^d$.*

Then Σ is contained in a finite union of algebraic translates $\mathbf{u}H \subset \mathbf{G}_m^n$, $\mathbf{u} \in (\mathcal{O}_S^*)^n$, $H \subset \mathbf{G}_m^n$ an algebraic subgroup, such that, for a $P = P_{\mathbf{u}H} \in K[X_1^{\pm 1}, X_2, \ldots, X_n]$ and a $c = c_{\mathbf{u}H} \in K$, we have $X_1 + \cdots + X_n = cX_1 P(X_1, \ldots, X_n)^d$, as functions in $K[\mathbf{u}H]$.

Condition (i) on the "dominant term", which amounts to condition (2) in Theorem 1, is probably not needed, but it seems a very difficult problem to remove it. Note that the conclusion is rather restrictive on the relevant translates $\mathbf{u}H$, and admits a partial converse.

See also [9] for this corollary and for the following general conjecture on integral points on subvarieties of $\mathbf{A}^1 \times \mathbf{G}_m^n$: *let V be an irreducible subvariety of $\mathbf{A}^1 \times \mathbf{G}_m^n$ with a Zariski-dense set of S-integral points, such that the projection $\pi : V \to \mathbf{G}_m^n$ is finite. Then $\pi(V)$ is an algebraic translate $\mathbf{u}H$ and there exist an integer $d \geqslant 1$ and a morphism $\tau : H \to V$ such that $\mathbf{u}[d] = \pi \circ \tau$.* (Here $[d]$ denotes the dth power map.) The conjecture is true for $\dim V = 1$ as may be shown, e.g. by Siegel's Theorem on integral points (see [9]); Theorem 1 quickly yields another solution of this case: one reduces to the integral solutions for an equation $F(x, y) = 0$, where x is an S-unit. Taking for f in Theorem 1 an expansion of y as a Puiseux series gives the conclusion. We leave the details to the interested reader.

The corollary allows to prove easily the finiteness of solutions to diophantine equations such as $y^2 = 1 + 2^m + 6^n$ (one takes either the infinite absolute value or the 2-adic one as the case may be); see also [4], especially the corollary therein, for this and more general equations. On the contrary, the apparently similar equation $y^2 = 1 + 2^m + 3^n$ is not known to have only finitely many integer solutions; our method does not apply since the required growth conditions may be not verified; what can be deduced from the present corollary is that, for any possible infinite sequence (m, n, y) of integral solutions, the ratio $m \log 2 / n \log 3$ converges to 1.

A situation of Theorem 1 which has been implicitly investigated in [6] is when $f(\mathbf{x}_h) = 0$ for a sequence of S-unit points $\mathbf{x}_h$. Note that in this case conditions (1), (3) and (4) are automatically verified. We shall emphasize this relevant special case in Theorem 3 below. We remark that Masser's Theorem is quite relevant in Mahler's method for transcendence. In this respect we also remark that Mahler's method applied to our function f would yield, with the assumptions above, that f coincides with an algebraic function on the sequence $\{\mathbf{x}_h\}$. In this situation, both an application of Masser's Theorem or of Theorem 1 (as the case may be) give further sharper conclusions.

Especially in the treatment of the case $f(\mathbf{x}_h) = 0$, a prominent role is played by the class of tori H such that $f(\mathbf{X})$ vanishes on some coset $\mathbf{u}H$. (It turns out that this vanishing may be interpreted both formally or as a function near the origin.) In other words, we are interested in the cosets of subtori of $\mathbf{G}_m^n$ which near the origin are contained in a certain analytic hypersurface. In the case when "analytic" is replaced by "algebraic", a structure theorem is known: given an algebraic variety $Z \subset \mathbf{G}_m^n$, there are only finitely many tori H such that some coset $\mathbf{u}H$ is contained in Z and is maximal for this property (see, e.g. [1]; see also [2] for another argument, valid also in the case of abelian varieties).

Such a statement is not generally true in the analytic case (see Section 2 for a simple counterexample). However, the tori which appear in our context have the further property of containing a sequence converging to the origin; hence we restrict our considerations to the class of tori *passing through zero*, in the sense of Definition 1 below. For this class of tori, the mentioned finiteness theorems valid in the case of algebraic varieties generalize to an analytic hypersurface V (around the origin), defined by $f(\mathbf{x}) = 0$.

In fact, we shall prove a structure theorem saying that in the set of maximal cosets $\mathbf{u}H$ contained in V, only finitely many tori H appear, and for fixed H the set of cosets is parametrized by an algebraic variety. This will be relevant for the last conclusion of Theorem 1 as well as for the special case when we assume $f(\mathbf{x}_h) = 0$ (e.g. Theorem 3). In this last case in particular, this will imply the existence of an algebraic variety $Z \subset V$, depending only on $f(\mathbf{X})$ but not on the sequence of points, containing almost all the points in our sequence. Moreover, the points will be contained in a finite number of cosets, each entirely contained in Z. (While Z depends only on f, the cosets may depend on the sequence.)

We shall state precisely all of this in the following two theorems (as well as in Section 2).

THEOREM 2. – *Given a power series $F \in \mathbf{C}[\![\mathbf{X}]\!]$, the set of tori $H \subset \mathbf{G}_m^n$ such that* (i) *the Zariski closure of H in $\mathbf{A}^n$ contains 0,* (ii) *for some $\mathbf{u} = \mathbf{u}_H$, F vanishes on $\mathbf{u}H$ and* (iii) *no coset larger than $\mathbf{u}H$ verifies* (ii)*, is a finite set.*

Also, for a given torus H satisfying (i)*, the set of $\mathbf{u} \in \mathbf{G}_m^n$ such that $F = 0$ on $\mathbf{u}H$ is an algebraic set.*

In fact, the last conclusion, which proves the existence of the algebraic variety Z alluded to above, may be further specified (see Theorem 6 below).

THEOREM 3. – *Let $f(\mathbf{X}) = \sum_{\mathbf{i}} a_{\mathbf{i}} \mathbf{X}^{\mathbf{i}}$ be a power series with algebraic coefficients in $\mathbf{C}_\nu$ converging in a neighborhood of the origin in $\mathbf{C}_\nu^n$. Let S be a finite set of absolute values of K containing the archimedean ones. Let $\mathbf{x}_h = (x_{h1}, \ldots, x_{hn})$ $(h = 1, 2, \ldots)$ be a sequence of S-units in ${K^\times}^n$, tending to zero in K_ν^n and such that $f(\mathbf{x}_h) = 0$. Suppose also that $\hat{h}(\mathbf{x}_h) = \mathrm{O}(-\log(\max_i |x_{hi}|_\nu))$.*

Then there exist a finite number of cosets $\mathbf{u}_1H_1,\dots,\mathbf{u}_rH_r\subset\mathbf{G}_m^n$ *on which* f *vanishes and such that* $\{\mathbf{x}_h\}_{h\in\mathbf{N}}\subset\bigcup_{i=1}^r\mathbf{u}_iH_i$.

Moreover, the tori $H_1,\dots,H_r$ *depend only on* f*, not on the sequence* $\{\mathbf{x}_h\}$.

Note that we do not have the last conclusion in Theorem 1; namely, in that case the tori might depend *a priori* on the sequence $\{\mathbf{x}_h\}$.

Remark. – As in Theorem 1 of [3], the Taylor coefficients of $f(\mathbf{X})$, though algebraic, are allowed to generate an infinite degree extension of $\mathbf{Q}$.

2. Subgroups of $\mathbf{G}_m^n$ passing through 0 and corresponding restriction of power series

In this section K will denote any field of characteristic zero.

We briefly recall the structure of algebraic subgroups of $\mathbf{G}_m^n$ (see [8]).

Let H be such a subgroup. Then H is defined by binomial equations $X_1^{a_1}\cdots X_n^{a_n}=1$, where $(a_1,\dots,a_n)$ varies through a lattice (possibly of rank $<n$) $\Lambda=\Lambda_H\subset\mathbf{Z}^n$. In this way we obtain a 1–1 correspondence between algebraic subgroups of $\mathbf{G}_m^n$ and sub-lattices of $\mathbf{Z}^n$. Also, we have $\dim H+\operatorname{rank}\Lambda=n$.

The group H is connected (i.e. is an irreducible variety) precisely if Λ is *primitive*, i.e. $\mathbf{Z}^n/\Lambda$ is torsion-free. When H is connected, we also say it is a *torus*. In any case, H is a direct product of a finite group by its connected component of the identity, which is a torus. By *coset* we shall mean a translate of a torus by a point.

When H is a torus we also have an algebraic group isomorphism $H\cong\mathbf{G}_m^{\dim H}$ given by a parametrisation

$$X_i=\mathbf{t}^{\mathbf{u}_i},\quad i=1,\dots,n,\quad \mathbf{t}=(t_1,\dots,t_k),$$

where $k=\dim H$, the vectors $\mathbf{u}_i$ lie in $\mathbf{Z}^k$ and the column vectors of the matrix whose row vectors are the $\mathbf{u}_i$ are a basis for the orthogonal complement of Λ_H in $\mathbf{Z}^n$.

The class of algebraic subgroups of $\mathbf{G}_m^n$ which will appear in our setting are those containing the origin in their topological closure (for the valuation ν) in $\mathbf{A}^n\supset\mathbf{G}_m^n$. The main reason lies in the assumptions of Theorem 1, involving a sequence converging to the origin. We shall analyse this class of groups, which appears to have properties rather relevant for our applications.

DEFINITION 1. – We say that an algebraic subgroup $H\subset\mathbf{G}_m^n$ *passes through* 0 if the Zariski closure of H in $\mathbf{A}^n$ contains $(0,\dots,0)$.

We note that a subgroup H passes through 0 if and only if the same property is true for at least one among its translates, in which case it is true for all translates.

PROPOSITION 1. – *The following conditions are equivalent for a torus* H.

(i) H *passes through* 0.

(ii) *The lattice* Λ_H *does not contain any nonzero vector with all its coordinates* $\geqslant 0$.

(iii) *There exists a parametrisation* $\varphi:\mathbf{G}_m^k\to H$ *as above*

$$X_i=\mathbf{t}^{\mathbf{u}_i},\quad i=1,\dots,n,\quad \mathbf{t}=(t_1,\dots,t_k),$$

where the vectors $\mathbf{u}_i\in\mathbf{Z}^k$ *have all their coordinates* >0.

(iv) *There exists a point* $(x_1,\dots,x_n)\in H\cap \mathbf{C}_\nu^n$ *with* $|x_i|_\nu<1$ *for* $i=1,\dots,n$.
(v) *There exists a sequence in* $H\cap \mathbf{C}_\nu^n$ *converging to* $(0,\dots,0)$ *in the* ν*-adic topology.*

Proof. – To prove that (i) implies (ii), suppose that Λ_H contains a nonzero vector $(a_1,\dots,a_n)$ in the first orthant, i.e. with $a_i\geqslant 0$ for $i=1,\dots,n$. By definition, the function $X_1^{a_1}\cdots X_n^{a_n}-1$ vanishes on H. On the other hand, this function is a polynomial and therefore it must vanish on the Zariski closure of H on $\mathbf{A}^n$, proving that this closure does not contain the origin.

To prove that (ii) implies (iii) requires more work. We start by proving the following

CLAIM. – *Let* $V\subset\mathbf{R}^n$ *be a vector space not containing any nonzero vector in the first orthant. Then its orthogonal complement in* $\mathbf{R}^n$ *contains a basis in the interior of the first orthant.*

To prove the claim, we first consider a weak form of it, namely we show that the orthogonal complement V^* to V contains a nonzero vector in the first orthant. Let $\mathbf{v}_1,\dots,\mathbf{v}_r\in\mathbf{R}^n$ be a basis for V. We view the $\mathbf{v}_i$ as row vectors of a matrix M and define $\mathbf{w}_1,\dots,\mathbf{w}_n\in\mathbf{R}^r$ to be the column vectors of M.

Consider the convex span W of the $\mathbf{w}_i$, namely

$$W:=\Big\{x_1\mathbf{w}_1+\cdots+x_n\mathbf{w}_n\colon\ x_i\geqslant 0,\ \sum x_i=1\Big\}.$$

Naturally, W is a compact subset of $\mathbf{R}^r$ and we contend that it contains the origin. Assume the contrary, and take a vector $\mathbf{z}\in W$ of minimal norm. Then $\mathbf{z}\neq 0$ and, denoting by $(\cdot,\cdot)$ the euclidean scalar product in $\mathbf{R}^r$, we have

$$(\mathbf{z},\mathbf{w})\geqslant 0,\quad \text{for all } \mathbf{w}\in W.$$

In fact, if $(\mathbf{z},\mathbf{w})=-s<0$ for a $\mathbf{w}\in W$, consider $\mathbf{z}(t):=t\mathbf{w}+(1-t)\mathbf{z}$ for a real positive $t<1$. We have $\mathbf{z}(t)\in W$ and $|\mathbf{z}(t)|^2=t^2|\mathbf{w}|^2+(1-t)^2|\mathbf{z}|^2-2t(1-t)s$. Denoting by $f(t)$ the right side, we see that $f(0)=|\mathbf{z}|^2$ and $f'(0)=-2|\mathbf{z}|^2-2s<0$. Therefore $f(t)<|\mathbf{z}|^2$ for small positive t, contradicting the minimality of $|\mathbf{z}|$ for $\mathbf{z}\in W$. (This argument is fairly standard.)

If $\mathbf{z}=(z_1,\dots,z_r)$ we then have that $z_1\mathbf{v}_1+\cdots+z_r\mathbf{v}_r$ is a nonzero vector lying in the first orthant (namely, all of its coordinates are nonnegative): in fact, its coordinates are just the scalar products $(\mathbf{z},\mathbf{w}_i)$. Thus we have a contradiction, proving that W contains in fact the origin. Take an equation $y_1\mathbf{w}_1+\cdots+y_n\mathbf{w}_n=0$ with nonnegative y_i having sum equal to 1. The vector $\mathbf{y}=(y_1,\dots,y_n)$ is nonzero, lies in the first orthant, and is orthogonal to V, proving the weak form of the claim.

To prove the full claim, take a vector $\mathbf{v}^*\in V^*$ in the first orthant, having a maximum number of nonzero coordinates. We show that in fact all of its coordinates are positive. Assume the contrary and renumber coordinates to assume $\mathbf{v}^*=(x_1,\dots,x_h,0,\dots,0)$, where $x_i>0$ for $i=1,\dots,h$. Define $U\subset\mathbf{R}^{n-h}$ to be the image of V^* under the projection on the last $n-h$ coordinates. Suppose that U has a nonzero vector $\mathbf{u}$ in the first orthant of $\mathbf{R}^{n-h}$ and let $\mathbf{u}'\in V^*$ be a vector whose projection is $\mathbf{u}$. Then we see that, for small positive t, the vector $\mathbf{v}^*+t\mathbf{u}'\in V^*$ has at least $h+1$ positive coordinates and lies in the first orthant, a contradiction with the maximality of h. Therefore U has no nonzero vectors in the first orthant. By the weak form of the claim, proved above, applied this time to U, there exists a nonzero vector $\mathbf{w}\in\mathbf{R}^{n-h}$, in the first orthant, which is orthogonal to U. Consider now the vector $\mathbf{w}'\in\mathbf{R}^n$ whose first h coordinates are 0 and whose projection on the last $n-h$ coordinates is $\mathbf{w}$. Then $\mathbf{w}'$ is nonzero, lies in the first orthant and is orthogonal to V^*, namely lies in V, against the assumption.

This contradiction therefore proves that $\mathbf{v}^*$ has strictly positive coordinates. In particular, a whole neighborhood of $\mathbf{v}^*$ in $\mathbf{R}^n$ lies in the interior of the first orthant. Intersecting this neighborhood with V^*, we get a nonempty open subset of V^* contained in the interior of the first orthant. Since every nonempty open subset of a real vector space contains a basis of it, the claim finally follows.

We can now apply the claim to the real vector space V spanned by Λ_H. Observe that V cannot contain any nonzero vector in the first orthant: in fact, assume the contrary, and let $\mathbf{v}$ be such a vector. Renumbering coordinates, we may assume that $\mathbf{v}$ has the first h coordinates equal to 0 and has the remaining coordinates strictly positive. The subspace of V consisting of vectors with vanishing first h coordinates is defined over $\mathbf{Q}$, whence its rational points are dense in it. Since $\mathbf{v}$ lies in this space, there exists a rational nonzero vector in V contained in the first orthant. But then, by multiplying this vector by a suitable positive integer, we would obtain a nonzero vector in Λ_H, contained in the first orthant, against the present assumption.

In virtue of the claim, we thus obtain a basis for the orthogonal V^*, contained in the interior of the first orthant. Since V^* is also defined over $\mathbf{Q}$, we obtain an integral primitive vector $\mathbf{w}_1$ in V^* having strictly positive coordinates. Complete now, as is certainly possible, $\mathbf{w}_1$ to a basis $\mathbf{w}_1, \ldots, \mathbf{w}_k$ for the lattice $V^* \cap \mathbf{Z}^n$. Replacing if necessary $\mathbf{w}_i$ by $\mathbf{w}_i + N\mathbf{w}_1$ for $i = 2, \ldots, k$, where N is a sufficiently large positive integer, we may assume that all the $\mathbf{w}_i$ lie in the first orthant.

We now define the matrix $\mathbf{u}_1, \ldots, \mathbf{u}_n$ (where $\mathbf{u}_i \in \mathbf{Z}^k$) as the transpose of the matrix formed with $\mathbf{w}_1, \ldots, \mathbf{w}_k$ and consider the map

$$\varphi : \mathbf{G}_m^k \to \mathbf{G}_m^n, \quad \mathbf{t} \mapsto (\mathbf{t}^{\mathbf{u}_1}, \ldots, \mathbf{t}^{\mathbf{u}_n}).$$

This map is plainly regular and is an algebraic group homomorphism. Its image is contained in H, since the $\mathbf{w}_i$ are orthogonal to Λ_H. Note also that

$$k =: \dim V^* = n - \dim V = n - \operatorname{rank} \Lambda_H = \dim H.$$

Since the $\mathbf{w}_i$ are linearly independent and since H is a torus, the map is dominant, whence, being a group homomorphism, is surjective; moreover, the fact that the vectors $\mathbf{w}_i$ are a basis for the lattice orthogonal to Λ_H (and not merely for V^*), easily implies that it is injective. This finally proves that (ii) implies (iii).

Now, suppose given a parametrisation φ satisfying the stated conditions and choose $\mathbf{t} = (t_1, \ldots, t_k) \in \mathbf{G}_m^k$ with $|t_i|_\nu < 1$. Plainly the point $(x_1, \ldots, x_n) := \varphi(\mathbf{t})$ has the required property, proving that (iii) implies (iv).

Next, starting from a point $(x_1, \ldots, x_n)$ as in (iv), the sequence of its positive powers converges to the origin. This proves the implication (iv) $\Longrightarrow$ (v).

Finally, that (v) implies (i) is clear, since a Zariski closed subset of $\mathbf{A}^n(\mathbf{C}_\nu)$ is closed also for the ν-adic topology. $\square$

Remark. – In the arithmetical results of the present paper, we shall often need to work with the restriction of a power series $F \in K[\![\mathbf{X}]\!]$ to suitable algebraic subgroups or cosets passing through 0. The parametrisation occurring in (iv) of the preceding lemma, combined with the next simple lemma, allows us to view such a restriction as a new power series in $K[\![\mathbf{t}]\!]$: we thus reduce to work on a smaller power of $\mathbf{G}_m$. A parametrisation φ as in the proposition sends a torus in $\mathbf{G}_m^k$ passing through 0 in a subtorus of H with the same property. Note however that there may exist proper subtori $H' \subset H$ such that H' passes through 0 but $\varphi^{-1}(H')$ does not! (Consider, e.g. the example $H := \{x = z\} \subset \mathbf{G}_m^3$, $H' := \{x = y = z\}$ and $\varphi(t, u) = (tu, tu^2, tu)$.) This phenomenon slightly complicates the proof of Theorems 4, 5 below.

LEMMA 1. – *Let $H \subset \mathbf{G}_m^n$ be a torus passing through 0. Then, for every given monomial in $K[\![\mathbf{X}]\!]$, there exist only finitely many monomials having the same restriction on H.*

Proof. – By Proposition 1, there exists $P := (x_1, \dots, x_n) \in H$ with $\|P\|_\nu < 1$. Observe that $|\mathbf{X}^\mathbf{i}(P)|_\nu \leqslant \|P\|_\nu^{\deg \mathbf{X}^\mathbf{i}} = \|P\|_\nu^{i_1+\cdots+i_n}$. Therefore the values $\mathbf{X}^\mathbf{i}(P)$ tend ν-adically to 0, whence at most a finite number of them can assume the same value on P and a fortiori on the whole H. □

We observe that, given a parametrisation $\varphi : \mathbf{G}_m^k \to H$ as in Proposition 1, (iii), two monomials in $\mathbf{X}$ have the same restriction to H if and only if their composition with φ produces the same monomial in $\mathbf{t}$.

This same positivity, together with Lemma 1 implies that, for a power series $F \in K[\![\mathbf{X}]\!]$, the composition $F \circ \varphi$ is again a power series, in $K[\![\mathbf{t}]\!]$. This power series represents the restriction of F to H, through the given isomorphism $H \cong \mathbf{G}_m^k$. The restriction of F to a coset $\mathbf{u}H$ will be simply the restriction to H of the power series $F(\mathbf{u}\mathbf{X})$.

Observe that in this way, if H passes through 0, we can always define the restriction to $\mathbf{u}H$ of a formal power series, regardless of convergence in any topology. In other words, we do not need that the series represents a function in order to define its restriction to $\mathbf{u}H$.

Actually, one can define the restriction even without a parametrization as above, simply by grouping together monomials which have the same restriction to H. We shall make this precise in a moment.

DEFINITION 2. – For an algebraic subgroup $H \subset \mathbf{G}_m^n$, we say that two monomials are equivalent with respect to H if their restrictions to H are equal functions, namely if their difference belongs to the ideal defining H as an algebraic variety.

We can now define the concept of *normal form* of a formal power series with respect to H.

DEFINITION 3. – Let $F = \sum a_\mathbf{i} \mathbf{X}^\mathbf{i} \in K[\![X_1, \dots, X_n]\!]$ be a formal power series. We say that F is in normal form with respect to H if the monomials which appear in F with nonzero coefficient are inequivalent with respect to H.

The above lemma implies that given any power series F, we may *normalize* it, by grouping together the monomials which are equivalent with respect to H. Namely, by the lemma, each equivalence class of monomials is a finite set, hence the sum $a_C := \sum_{\mathbf{X}^\mathbf{i} \in C} a_\mathbf{i}$ is well defined for every class C. Hence the power series $G := \sum_C a_C \mathbf{X}_C$, where $\mathbf{X}_C$ runs through a set of representatives for the equivalence classes, is in normal form with respect to H and has the same restriction to H as F.

Now, first, we stress that the set of representatives is quite arbitrary. Second, we remark that this notion is referred to formal power series; however it is easily seen that in the case when F converges absolutely near the origin, any of its normal forms coincides with F as a function in the intersection of H with the relevant convergence region.

These notions automatically extend to the case of a coset $\mathbf{u}H$ in place of a torus H: it suffices to argue with H in place of $\mathbf{u}H$ and with the series $F(\mathbf{u}\mathbf{X})$ in place of $F(\mathbf{X})$. Note that the relevant equivalence classes are the same as before.

Let $F \in K[\![\mathbf{X}]\!]$ be a formal power series over the field K. We shall consider the family of cosets $\mathbf{u}H$ (for a torus H passing through 0), such that the restriction of F to $\mathbf{u}H$ vanishes.

We have the following simple criterion for this vanishing.

PROPOSITION 2. – *Let $f(\mathbf{X}) = \sum_\mathbf{i} a_\mathbf{i} \mathbf{X}^\mathbf{i} \in K[\![\mathbf{X}]\!]$ and let H be a torus passing through 0. Suppose that the restriction of $f(\mathbf{X})$ to $\mathbf{u}H$ coincides with a polynomial (resp. vanishes). Then, for all but finitely many (resp. for all) classes C of equivalence of monomials, with respect to H,* *we have*

$$\sum_{\mathbf{X}^{\mathbf{i}}\in C} a_{\mathbf{i}}\mathbf{u}^{\mathbf{i}} = 0,$$

the sum being finite for each C.

The proof is immediate using the parametrisation φ and the preceding lemma. (A proof may be also obtained directly by invoking the concept of normal form, just introduced.)

As a matter of fact, the same arguments prove that, when the power series is convergent ν-adically, the same criterion holds if we interpret the restriction of F in the usual sense, i.e. considering it as a function. We state this fact separately.

PROPOSITION $2'$. – *Let* $f(\mathbf{X}) = \sum_{\mathbf{i}} a_{\mathbf{i}}\mathbf{X}^{\mathbf{i}} \in \mathbf{C}_\nu[\![\mathbf{X}]\!]$ *be convergent in a neighborhood* U *of the origin in* $\mathbf{C}_\nu^n$*, and let* H *be a torus passing through* 0*. Suppose that* $f(\mathbf{X})$ *coincides with a polynomial* (*resp. vanishes*) *on* $U \cap \mathbf{u}H$*. Then, for all but finitely many* (*resp. for all*) *classes* C *of equivalence of monomials, with respect to* H*, we have*

$$\sum_{\mathbf{X}^{\mathbf{i}}\in C} a_{\mathbf{i}}\mathbf{u}^{\mathbf{i}} = 0,$$

the sum being finite for each C.

We shall now prove a structure theorem for the set of cosets on which a given power series vanishes.

DEFINITION 4. – We shall say that a coset $\mathbf{u}H$ is *maximal for* F if F vanishes on $\mathbf{u}H$ but does not vanish on any strictly larger coset $\mathbf{u}H'$, where H' is a torus containing H properly.

For dimensional reasons, every coset on which F vanishes is contained in a maximal one for F.

In the case of algebraic varieties (e.g. if the power series is a polynomial), some finiteness results have been proved for the set of tori H occurring in a maximal coset; see for instance [1] (and also [2] for another proof, valid also in the context of abelian varieties). In the present situation, it is a fundamental assumption that the relevant tori pass through 0. Without this proviso, examples like $F(\mathbf{X}) := \prod_{\mathbf{i}}(1 - \mathbf{X}^{\mathbf{i}})$, where $\mathbf{i}$ varies on $\mathbf{N}^n$ show that we may have vanishing on an infinite family of maximal tori.

THEOREM 4. – *Given a power series* $F \in K[\![\mathbf{X}]\!]$*, the set of tori* H *passing through* 0 *and such that, for some* $\mathbf{u} = \mathbf{u}_H$*, the coset* $\mathbf{u}H$ *is maximal for* F*, is a finite set.*

It will be convenient to prove a more general fact. To state it, let Z be an affine irreducible variety over the field K and let $\Gamma \subset \mathbf{G}_m^n$ be a subtorus. We prove:

THEOREM 5. – *Given a power series* $G(\mathbf{z}, \mathbf{X}) \in K[Z][\![\mathbf{X}]\!]$*, the set of tori* $H \subset \Gamma$*,* H *passing through* 0*, with the property that there exists* $\mathbf{u} \in Z$ *such that* H *is a maximal subtorus of* Γ *on which* $G(\mathbf{u}, \mathbf{X})$ *vanishes, is a finite set.*

To deduce Theorem 4 from this statement, just take $Z = \Gamma = \mathbf{G}_m^n$ and $G(\mathbf{z}, \mathbf{X}) := F(\mathbf{z}\mathbf{X})$ (componentwise multiplication).

Proof of Theorem 5. – We start by recalling a well-known fact, whose proof we give for completeness. We shall say that $(x_1, \ldots, x_n) \geqslant (y_1, \ldots, y_n)$ if $x_i \geqslant y_i$ for $i = 1, \ldots, n$. This defines a partial order on $\mathbf{R}^n$.

CLAIM. – *Given a subset $\mathcal{A} \subset \mathbf{N}^n$, the set of minimal points in $\mathcal{A}$ for the above defined order relation, is a finite set.*

We argue by induction on n, the claim being clear for $n = 1$. Suppose by contradiction that there exists an infinite set $\mathcal{A}' \subset \mathcal{A}$ of minimal points, and let $(x_1, \ldots, x_n)$ be one of them having a minimal first coordinate. Every other point $(y_1, \ldots, y_n) \in \mathcal{A}'$ has $y_1 \geqslant x_1$, whence must verify $y_i < x_i$ for at least one index $i \in \{2, \ldots, n\}$. Going to an infinite subset, we may assume that a given coordinate, say y_n, is bounded and, going to a smaller, but still infinite, subset, we may assume that $y_n = c$ is fixed. Now we get a contradiction with the inductive assumption. □

To prove the theorem, we shall argue by induction on $m := \dim \Gamma$, the result being clear for $m = 0$. Suppose now that $m \geqslant 1$ and that the statement is true up to $m - 1$.

We now argue by a second induction, this time on $\dim Z$. It turns out to be convenient to start the induction with the case of an empty Z, i.e. $\dim Z = -1$. In this case the set of relevant tori is empty as well, so the theorem is true.

We now assume that Z is nonempty and that the statement has been proved for Γ and all varieties of dimension smaller than $\dim Z$.

Let Z, Γ, G be as in the statement and put $G(\mathbf{z}, \mathbf{X}) = \sum_{\mathbf{i}} a_{\mathbf{i}}(\mathbf{z})\mathbf{X}^{\mathbf{i}}$. We start by replacing $G(\mathbf{z}, \mathbf{X})$ by a possibly different series obtained by removing certain terms appearing in G. We let C be an equivalence class of monomials with respect to Γ (see Definition 2 above) and consider the sum $\sum_{\mathbf{X}^{\mathbf{i}} \in C} a_{\mathbf{i}}(\mathbf{z})$. If this sum is zero in $K[Z]$, we throw away the entire class from G and otherwise we simply leave unchanged all the terms corresponding to C. We do this for all classes, obtaining a series $G^*(\mathbf{z}, \mathbf{X}) = \sum_{\mathbf{i}} a^*_{\mathbf{i}}(\mathbf{z})\mathbf{X}^{\mathbf{i}}$. Note that the restriction of $G(\mathbf{z}, \mathbf{X})$ to $Z \times \Gamma$ coincides with the restriction of $G^*(\mathbf{z}, \mathbf{X})$ to $Z \times \Gamma$; in particular, for all $\mathbf{u}$, $G(\mathbf{u}, \mathbf{X})$ vanishes on a subtorus H of Γ if and only if the same holds for $G^*(\mathbf{u}, \mathbf{X})$. Hence it is sufficient to prove our contention for G^* in place of G. Also, note that by construction a sum $\sum_{\mathbf{X}^{\mathbf{i}} \in C} a^*_{\mathbf{i}}(\mathbf{z})$ vanishes if and only if all of its terms vanish.

We define $\mathcal{A}$ to be the set of vectors $\mathbf{i}$ such that the monomial $\mathbf{X}^{\mathbf{i}}$ has a nonzero coefficient $a^*_{\mathbf{i}}$ in the formal power series $G^*(\mathbf{z}, \mathbf{X})$. We may assume that G^* is nonzero, since otherwise the result is plainly true; hence we may assume that $\mathcal{A}$ is nonempty. We define $\mathcal{A}'$ to be the (finite) set of minimal vectors in $\mathcal{A}$, with respect to the above defined partial order relation.

Let $\mathbf{u} \in Z(K)$ and let H be a maximal torus for the stated property, so in particular $G^*(\mathbf{u}, \mathbf{X})$ vanishes on H (but not on larger tori contained in Γ). We distinguish two cases.

First case: *there exists a class C with respect to Γ, C corresponding to some vector in $\mathcal{A}'$, such that the sum $\sum_{\mathbf{X}^{\mathbf{i}} \in C} a^*_{\mathbf{i}}(\mathbf{z})$ vanishes at $\mathbf{u}$*. Each of these conditions, which are finite in number, defines a finite number of proper subvarieties Z_j of Z, i.e. the components of a hypersurface $\sum_{\mathbf{X}^{\mathbf{i}} \in C} a^*_{\mathbf{i}}(\mathbf{z}) = 0$ in Z. Note that $\sum_{\mathbf{X}^{\mathbf{i}} \in C} a^*_{\mathbf{i}}(\mathbf{z})$ is not identically zero: this is because of our construction of G^* and because C contains at least one monomial corresponding to some vector in $\mathcal{A}$. Restricting the coefficients of G^* to each of these components gives rise to finitely many power series $G_j \in K[Z_j][\![\mathbf{X}]\!]$, to which we may apply the inductive assumption. In fact, $\dim Z_j < \dim Z$, since Z is irreducible. Also, note that (for given $\mathbf{u} \in Z_j(K)$) H remains *a fortiori* maximal even with respect to the new series. In this case, the proof is concluded.

Second case: *no sum $\sum_{\mathbf{X}^{\mathbf{i}} \in C} a^*_{\mathbf{i}}(\mathbf{z})$, where C contains some monomial corresponding to a vector in $\mathcal{A}'$, vanishes at $\mathbf{u}$*. This case requires further work.

Since H passes through 0, by Proposition 1 there exists a vector $\mathbf{v}_H \in \mathbf{Z}^n$ with strictly positive coordinates and orthogonal to Λ_H. Let $\mathcal{A}''$ be the subset of $\mathbf{i} \in \mathcal{A}$ such that the scalar product $(\mathbf{i}, \mathbf{v}_H)$ is minimal. This subset is nonempty, since $(\mathbf{i}, \mathbf{v}_H)$ is a nonnegative integer for $\mathbf{i} \in \mathcal{A}$. We contend that $\mathcal{A}'' \subset \mathcal{A}'$: in fact, if $\mathbf{i} \in \mathcal{A}''$ and $\mathbf{j} \in \mathcal{A}$ is distinct from $\mathbf{i}$, we get $(\mathbf{j}, \mathbf{v}_H) \geqslant (\mathbf{i}, \mathbf{v}_H)$, whence $(\mathbf{j} - \mathbf{i}, \mathbf{v}_H) \geqslant 0$, proving that $\mathbf{i} - \mathbf{j}$ cannot have only nonnegative coordinates. Therefore $\mathcal{A}'' \subset \mathcal{A}'$, whence in particular, in virtue of the claim, the set $\mathcal{A}''$ is finite (and nonempty).

Fix $\mathbf{j} \in \mathcal{A}''$ and define E as the equivalence class of $\mathbf{X}^{\mathbf{j}}$ with respect to H. In view of Proposition 2, the sum $\sum_{\mathbf{X}^{\mathbf{i}} \in E} a_{\mathbf{i}}^*(\mathbf{u})$ vanishes. On the other hand, the class E is a finite union of equivalence classes with respect to Γ, since $H \subset \Gamma$. Since the subsum corresponding to the equivalence class C of $\mathbf{j}$ with respect to Γ is nonzero (in fact we are in the "second case" and $\mathbf{j} \in \mathcal{A}'' \subset \mathcal{A}'$), there must be another class $\bar{C} \subset E$, with respect to Γ, such that the corresponding subsum is also nonzero. Therefore there exists $\bar{\mathbf{j}} \in \mathcal{A}$ such that $\mathbf{j} - \bar{\mathbf{j}} \in \Lambda_H \setminus \Lambda_\Gamma$. In particular $\mathbf{j} - \bar{\mathbf{j}}$ is orthogonal to $\mathbf{v}_H$, whence $\bar{\mathbf{j}} \in \mathcal{A}'' \subset \mathcal{A}'$.

Observe that, since $\mathbf{j}$ and $\bar{\mathbf{j}}$ are (distinct) elements of the finite set $\mathcal{A}'$, the vector $\mathbf{y} := \mathbf{j} - \bar{\mathbf{j}}$ is nonzero and lies in a finite set independent of H. For every such $\mathbf{y}$ we consider the $(m-1)$-dimensional subtorus $\Gamma_{\mathbf{y}} \subset \Gamma$ given by the connected component of the identity in the group defined by the equation $\mathbf{X}^{\mathbf{y}} = 1$ in Γ. Note that the lattice associated to such a group is $\mathbf{Z}\mathbf{y} + \Lambda_\Gamma \subset \Lambda_H$, so H is contained in the group. Since however H is connected, it must be contained in the connected component of the identity; in other words, $\Gamma_{\mathbf{y}} \supset H$.

We may now apply the inductive assumption (with respect to m), with $\Gamma_{\mathbf{y}}$ in place of Γ. Observe that, since there are only finitely many $\mathbf{y}$ to consider, only finitely many $\Gamma_{\mathbf{y}}$ may arise in this way. Also, if $H \subset \Gamma_{\mathbf{y}}$ is maximal for Γ (relative to $\mathbf{u}$), a fortiori it is maximal for $\Gamma_{\mathbf{y}}$ (again relative to $\mathbf{u}$), concluding the proof. □

We now complete the proof of Theorem 2, by recalling:

THEOREM 6. – *Given a torus H, there exists an algebraic set $U_H \subset \mathbf{G}_m^n$ whose points are pairwise inequivalent modulo H, and such that $F(\mathbf{X})$ vanishes on a coset $\mathbf{u}H$ if and only if there exists $\mathbf{u}' \in U_H$ with $\mathbf{u}H = \mathbf{u}'H$.*

The proof is implicit in the proof of [1, Theorem 1(a)], since by Proposition 2 the set of $\mathbf{u} \in \mathbf{G}_m^n$ such that $F(\mathbf{X})$ vanishes on $\mathbf{u}H$ is an algebraic set. □

3. Proof of Theorem 1

A fundamental tool will be a consequence of the Subspace Theorem, proved as Theorem 4 in [3]. For the reader's convenience we recall the result as the following

LEMMA 2. – *Let K be a number field, S a finite set of absolute values of K containing the archimedean ones, ν 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 $\delta > (N+2)\epsilon$, there are only finitely many $(N+1)$-tuples $\mathbf{w} = (w_0, \ldots, w_N) \in (K^*)^{N+1}$ such that the inequalities*

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

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

We shall also need a result in the theory of S-unit equations, which is essentially a reformulation of a theorem of Evertse.

LEMMA 3. – *Let K, S, N be as above, and let $0 < d < 1/(N + N^2)$. Then there exists a finite set $Q \subset K^*$ such that, for all solutions $\mathbf{w} = (w_1, \ldots, w_N) \in (K^*)^N$ of the equation $w_1 + \cdots + w_N = 0$ satisfying $h_S(w_i) + h_S(w_i^{-1}) < d h(w_i)$ for $i = 1, \ldots, N$, there exist indices $1 \leqslant r < s \leqslant N$, such that $w_r / w_s \in Q$.*

Proof. – Without loss of generality we may assume that there is no proper nonempty vanishing subsum of the w_i. In fact, we may otherwise consider just the w_i appearing in a vanishing subsum, since all the assumptions plainly remain valid for this subset.

We enlarge S so that there exists a system Σ of integral ideals in K, representatives for the ideal classes, and made up of prime ideals in S. For every $\mathbf{w}$ as above, let J be the minimal integral ideal such that $Jw_1, \ldots, Jw_N$ are contained in $\mathcal{O}_S$. Let $I \in \Sigma$ be in the same class of J and let $\lambda \in K$ be a generator for $I^{-1}J$. It is known that, multiplying if necessary λ by a unit, we may assume that $h(\lambda) \leqslant h_S(w_1) + \cdots + h_S(w_N) + D$, where D depends only on the regulator of K.

We put $x_i = \lambda w_i$. The x_i are S-integers satisfying the equation $x_1 + \cdots + x_N = 0$. Also, we have

$$\sum_{v \in S} \sum_{i=1}^{N} \log |x_i|_v \leqslant Nh(\lambda) - \sum_{v \notin S} \sum_{i=1}^{N} \log |w_i|_v$$
$$\leqslant Nh(\lambda) + \sum_{i=1}^{N} h_S\big(w_i^{-1}\big) \leqslant (Nd + N^2 d)h(\mathbf{w}) + ND.$$

Put $\delta := Nd + N^2 d < 1$. We apply the theorem of Evertse appearing at p. 91 of [7]. The assumptions for that theorem are verified, since in the notation of that theorem the vector $(x_1, \ldots, x_N)$ is $(\exp(ND), \delta, S)$-admissible and since we have remarked that we may assume that no proper subsum vanishes. We conclude that the projective point

$$(x_1 : \ldots : x_N) = (w_1 : \ldots : w_N)$$

has finitely many possibilities, concluding the proof. □

Now to the proof of the theorem. We fix a positive integer B; later on we shall choose it to be large enough for our purposes.

We consider the set of monomials $\mathbf{X}^{\mathbf{i}}$ such that $\|\mathbf{i}\| \leqslant B$ and $a_{\mathbf{i}} \neq 0$. We let N be the number of such monomials and choose a numbering $\mathbf{M}_1, \ldots, \mathbf{M}_N$ of them. Further, if $\mathbf{M}_r = \mathbf{X}^{\mathbf{i}}$ we put $c_r := a_{\mathbf{i}}$ and we put $c_0 = -1$.

For any given (large) positive integer h, we put (dropping the subscript h from the w's)

$$w_0 := f(\mathbf{x}_h), \qquad w_j := \mathbf{M}_j(\mathbf{x}_h), \quad j = 1, 2, \ldots, N.$$

We go on by proving a ν-adic bound for the sum $c_0 w_0 + \cdots + c_N w_N$.

Since f converges near the origin, we have a bound $|a_{\mathbf{i}}|_\nu \leqslant C_1 \rho^{\|\mathbf{i}\|}$, where C_1, ρ are suitable positive real numbers. We also recall that $\|\mathbf{x}_h\|_\nu$ tends to 0. We assume in the sequel that h is so large that $2\rho\sqrt{\|\mathbf{x}_h\|_\nu} < 1$. Then we have

$$|c_0 w_0 + \cdots + c_N w_N|_\nu \leqslant C_1 \sum_{\|\mathbf{i}\| > B} \big(\rho \|\mathbf{x}_h\|_\nu\big)^{\|\mathbf{i}\|} \leqslant C_2 \|\mathbf{x}_h\|_\nu^{B/2}, \tag{3.1}$$

where $C_2 = C_1 \sum_{\mathbf{i}} 2^{-\|\mathbf{i}\|}$ (the sum being convergent).

In order to apply Lemma 2 (for a suitable δ), we proceed to estimate the terms $H(w_0)$, $H_S(w_0)$ and $\hat{H}(\mathbf{w})$.

We shall use assumption (2) of Theorem 1, saying that $\hat{H}(\mathbf{x}_h)$ is bounded above by positive powers of $\|\mathbf{x}_h\|_\nu^{-1}$.

We fix $\epsilon = 1/BN^2$ and put $\delta = (N+3)\epsilon$. For the rest of the proof, $C_3, C_4, \ldots$ will denote positive numbers independent of B or h.

First, by assumption (4) (and (2)), we have for large h

$$H^{-1}(w_0) \geqslant \|\mathbf{x}_h\|_\nu^{C_3}.$$

By assumption (3), we have (for large h)

$$H_S^{-1}(w_0) \geqslant \|\mathbf{x}_h\|_\nu^\epsilon.$$

Finally, by (4) (and (2)), we get, recalling that $\mathbf{w} = (w_0, w_1, \ldots, w_N)$,

$$\hat{H}^{-1}(\mathbf{w}) \geqslant \|\mathbf{x}_h\|_\nu^{C_4 BN}.$$

Combining all these estimates, we obtain

$$\bigl(H(w_0)H_S(w_0)^{N+1}\bigr)^{-1}\hat{H}(\mathbf{w})^{-\delta} \geqslant \|\mathbf{x}_h\|_\nu^{C_3+(N+1)\epsilon+C_4BN\delta}.$$

Recalling our choices of δ and ϵ, we find that the exponent of $\|\mathbf{x}_h\|_\nu$ in the last inequality is bounded above by a number C_5 independent of B or h. Therefore, in view of Eq. (3.1), the assumption (ii) for the above lemma (with the present choice of ϵ and δ) is verified for all large h, provided B has been chosen larger than $2C_5$.

Also, assumption (i) of the lemma (with the present choice of ϵ) is verified for all large h. In fact, take any monomial w in $\mathbf{x}_h$, which we write as $w = x_{h1}^{a_1} \cdots x_{hn}^{a_n}$, with $a_i \geqslant 0$. We have

$$|w|_\nu = \prod_i |x_{hi}|_\nu^{a_i} \leqslant \|\mathbf{x}_h\|_\nu^{\sum_i a_i}.$$

It follows from Liouville inequality and from the fact that $\hat{H}(\mathbf{x}_h) \leqslant \|\mathbf{x}_h\|_\nu^{-C_6}$ that

$$H(w) \geqslant |w|_\nu^{-1} \geqslant \|\mathbf{x}_h\|_\nu^{-\sum_i a_i} \geqslant \hat{H}(\mathbf{x}_h)^{C_6 \sum_i a_i}.$$

Now, observe that $h_S(w) + h_S(w^{-1}) \leqslant \sum_i a_i(h_S(x_{hi}) + h_S(x_{hi}^{-1}))$. By our assumptions the right side is $\mathrm{o}(\sum_i a_i h(x_{hi}))$. In turn, this quantity is $\mathrm{o}(h(w))$, in view of the previous lower bound for $H(w)$.

We conclude that, for all large h, a suitable subsum of the $c_j w_j$ involving $c_0 w_0$ vanishes. Since there are only finitely many subsums to consider, we may assume in the sequel that the same subsum occurs for all h.

Let $P(\mathbf{X})$ be the polynomial corresponding to such a subsum; namely,

$$P(\mathbf{X}) = \sum_{j \in J} c_j \mathbf{M}_j(\mathbf{X})$$

for some $J \subset \{1, \ldots, N\}$. Then, putting $g(\mathbf{X}) := f(\mathbf{X}) - P(\mathbf{X})$ we have that

$$g(\mathbf{x}_h) = f(\mathbf{x}_h) - P(\mathbf{x}_h) = -c_0 w_0 - \sum_{j \in J} c_j w_j = 0$$

for large h.

The theorem will then follow immediately by applying the following proposition (which may have some independent interest) to the power series $g(\mathbf{X})$.

PROPOSITION 3. – *Let $f(\mathbf{X}) = \sum_{\mathbf{i}} a_{\mathbf{i}} \mathbf{X}^{\mathbf{i}}$ be a power series with algebraic coefficients in $\mathbf{C}_\nu$ converging in a neighborhood of the origin in $\mathbf{C}_\nu^n$. Let S and $\mathbf{x}_h = (x_{h1}, \ldots, x_{hn})$ $(h = 1, 2, \ldots)$*

be a sequence satisfying (1), (2) *of Theorem* 1. *Assume that for every fixed positive B, we have*

$$\big|f(\mathbf{x}_h)\big|_\nu = \mathrm{O}\big(\|\mathbf{x}_h\|_\nu^B\big).$$

Then all but finitely many of the elements $\mathbf{x}_h$ lie in the union of finitely many cosets $\mathbf{u}_1 H_1, \ldots, \mathbf{u}_r H_r$ on which f vanishes.

Moreover the tori $H_1, \ldots, H_r$ may be chosen to depend only on f and not on the sequence $\{\mathbf{x}_h\}$.

In particular, $f(\mathbf{x}_h)$ is forced to be eventually 0.

Proof of proposition. – The idea of the proof is roughly as follows. Using Lemmas 2, 3, one can show that if f is not identically zero, then the set $\{\mathbf{x}_h\}$ may be split in a finite union of sets, each set being contained in a proper coset passing through 0. Then one can argue with the restriction of f to each coset, iterating the method whenever f does not vanish identically on the coset. At each step the dimension of the relevant coset decreases, so we eventually end up with the union of finitely many cosets, on which f vanishes, containing our original sequence. Actually, the proof may be shortened and one can just deal with the final step. This motivates the following definition:

We say that a coset $\mathbf{u}H$ is minimal with respect to a set $\{\mathbf{z}_h\}_{h\in\mathbf{N}} \subset \mathbf{u}H$ if there do not exist proper sub-cosets $\mathbf{u}_1 H_1, \ldots, \mathbf{u}_r H_r \subset \mathbf{u}H$ such that $\{\mathbf{z}_h\}_{h\in\mathbf{N}}$ is contained in $\bigcup_{i=1}^r \mathbf{u}_i H_i$.

For dimensional reasons, it is clear that the set $\{\mathbf{x}_h\}$ may be partitioned in a finite number of sets Y_j, $j = 1, \ldots, s$, contained respectively in cosets $\mathbf{v}_j H_j$ passing through 0 and minimal with respect to Y_j. (Possibly $\mathbf{G}_m^n$ is already minimal with respect to $\{\mathbf{x}_h\}$.) Plainly, if a single $\mathbf{x}_h \in \mathbf{v}_j H$, then we may choose $\mathbf{v}_j = \mathbf{x}_h$. Therefore, without loss of generality, we may assume that all the involved $\mathbf{v}_j$ are defined over K.

To prove the proposition, it thus suffices to prove that *if a coset $\mathbf{u}H$ is minimal with respect to an infinite subsequence of $\{\mathbf{x}_h\}$, then f vanishes on $\mathbf{u}H$*. We proceed to prove this claim.

To save notation, we may assume that $\mathbf{u}H$ is minimal with respect to $\{\mathbf{x}_h\}$ and, as we have just remarked, we may choose $\mathbf{u} \in \mathbf{G}_m^n(K)$.

We begin to choose a normal form for $f(\mathbf{u}\mathbf{X})$ with respect to H, in the sense of Definition 3.

Namely, if $f(\mathbf{X}) = \sum_{\mathbf{i}} a_{\mathbf{i}} \mathbf{X}^{\mathbf{i}}$, we write $g(\mathbf{X}) = \sum_C a_C(\mathbf{u}) \mathbf{X}_C$, where C runs through the classes of monomials with respect to H, the $\mathbf{X}_C$ are a set of representative monomials for the equivalence classes and

$$a_C(\mathbf{u}) = \sum_{\mathbf{X}^{\mathbf{i}} \in C} a_{\mathbf{i}} \mathbf{u}^{\mathbf{i}}.$$

This function $g(\mathbf{X})$ converges ν-adically near the origin and verifies $g(\mathbf{x}) = f(\mathbf{u}\mathbf{x})$ for $\mathbf{x} \in H$, $\mathbf{x}$ near the origin.

We put $\mathbf{y}_h := \mathbf{u}^{-1}\mathbf{x}_h$, noting at once that the $\mathbf{y}_h$ verify (1) and (2) of Theorem 1. Also, we have $g(\mathbf{y}_h) = f(\mathbf{x}_h)$ (both series converge for large h), whence in particular, for all positive B we have

$$\big|g(\mathbf{y}_h)\big|_\nu = \mathrm{O}\big(\|\mathbf{y}_h\|_\nu^B\big).$$

The proof now runs quite similarly to what we have already seen, and we shall be a little more sketchy. As before, we fix a sufficiently large positive integer B.

We consider the set of monomials $\mathbf{X}_C$ of partial degrees $\leqslant B$ and such that $a_C(\mathbf{u}) \neq 0$. If this set is empty, then f vanishes on $\mathbf{u}H$ and we are done. We then let $N + 1 > 0$ be the number of such monomials and choose a numbering $\mathbf{M}_0, \ldots, \mathbf{M}_N$ of them. Further, if $\mathbf{M}_r = \mathbf{X}_C$ we put $c_r := a_C(\mathbf{u})$.

For any given (large) positive integer h, we put (dropping the subscript h)

$$w_j := \mathbf{M}_j(\mathbf{y}_h), \quad j = 0, 1, \ldots, N.$$

We go on by proving a ν-adic bound for the sum $c_0 w_0 + \cdots + c_N w_N$.

Since g converges near the origin, we have a bound $|a_C(\mathbf{u})|_\nu \leqslant A_1 \rho^{\deg \mathbf{X}_C}$, where A_1, ρ are suitable positive real numbers. For large enough h, we have

$$|c_0 w_0 + \cdots + c_N w_N|_\nu \leqslant \bigl|g(\mathbf{y}_h)\bigr|_\nu + A_1 \sum_{\|\mathbf{i}\| > B} \bigl(\rho \|\mathbf{y}_h\|_\nu\bigr)^{\|\mathbf{i}\|} \leqslant A_2 \|\mathbf{y}_h\|_\nu^{B/2}, \tag{3.2}$$

for a suitable A_2 independent of h. (For the last inequality we have used that $g(\mathbf{y}_h)$ is small.)

We shall now apply Lemma 2, estimating the various involved heights. In the present application, it will be noted that w_0 has not a special role, since it is automatically an almost S-unit (rather than an almost S-integer).

We fix $\epsilon = 1/BN^2$ and put $\delta = (N+3)\epsilon$. For the rest of the proof, $A_3, A_4, \ldots$ will denote positive numbers independent of B or h.

For large h we have, by (2) in the statement of Theorem 1 (and of the present proposition),

$$H^{-1}(w_0) \geqslant \|\mathbf{y}_h\|_\nu^{A_3}$$

and by (1) of Theorem 1 (again for large h)

$$H_S^{-1}(w_0) \geqslant \|\mathbf{y}_h\|_\nu^{\epsilon}.$$

Further, we get by (2), recalling that $\mathbf{w} = (w_0, w_1, \ldots, w_N)$,

$$\hat{H}^{-1}(\mathbf{w}) \geqslant \|\mathbf{y}_h\|_\nu^{A_4 BN}.$$

Combining all these estimates, we obtain

$$\bigl(H(w_0) H_S(w_0)^{N+1}\bigr)^{-1} \hat{H}(\mathbf{w})^{-\delta} \geqslant \|\mathbf{y}_h\|_\nu^{A_3 + (N+1)\epsilon + A_4 BN\delta}.$$

Recalling our choices of δ and ϵ, we find that the exponent of $\|\mathbf{y}_h\|_\nu$ in the last inequality is bounded above by a number A_5 independent of B or h. Therefore, in view of Eq. (3.2), the assumption (ii) for Lemma 2 (with the present choice of ϵ and δ) is verified for all large h, provided B has been chosen larger than $2A_5$.

Also, assumption (i) of the lemma (with the present choice of ϵ) is verified for all large h: this is in fact an immediate consequence of assumption (1) of Theorem 1.

We conclude that for all large h, a suitable subsum of the $c_j w_j$ vanishes. Since there are only finitely many subsums to consider, we may assume in the sequel that the same subsum occurs for all h.

Write the vanishing subsum in the form $c_{i_1} w_{i_1} + \cdots + c_{i_k} w_{i_k} = 0$. We apply the lemma of Evertse (i.e. Lemma 3) to this equation, writing N in place of k and $\mathbf{w} = (c_{i_1} w_{i_1}, \ldots, c_{i_k} w_{i_k})$.

The assumptions are certainly verified for large h, e.g. by choosing $d = 1/2(N + N^2)$.

In our original notation, we conclude that there exist a finite set Q and indices r, s, with $0 \leqslant r < s \leqslant N$, such that $c_r w_r / c_s w_s \in Q$.

For any fixed choice of $r < s$ and of $q \in Q$, the equation $\mathbf{M}_r(\mathbf{y})/\mathbf{M}_s(\mathbf{y}) = q c_s / c_r$ defines in H a proper sub-coset of H. In fact, note that the relevant monomials are pairwise inequivalent modulo H.

Since w_j equals $\mathbf{M}_j(\mathbf{y}_h)$, we have proved that the sequence $\{\mathbf{y}_h\}$ is contained in a finite union of proper sub-cosets of H, whence $\{x_h\}$ is contained in a finite union of proper sub-cosets of $\mathbf{u}H$. This contradicts the minimality of $\mathbf{u}H$ with respect to $\{\mathbf{x}_h\}$ and concludes the proof of the first part of the proposition (which suffices to imply Theorem 1).

To prove the proposition completely, we note that, enlarging the cosets if necessary, we may assume that they are maximal for f in the sense of Definition 4. Now it suffices to apply Theorem 2. □

4. Proof of remaining statements

Proof of corollary. – Since $\mathcal{O}_S^*$ is finitely generated we may write $x_1 = z^d\xi$ where $z \in \mathcal{O}_S^*$ and where $\xi \in \mathcal{O}_S^*$ has finitely many possibilities; hence in what follows we assume that ξ is fixed. Writing $x_i z^d$ in place of x_i we easily see that it suffices to prove the conclusion assuming that $x_1 = \xi$ is fixed. Note that this substitution leaves condition (i) unchanged.

We shall apply Theorem 1 with $f(\mathbf{X}) = \sqrt[d]{\xi + X_2 + \cdots + X_n}$, as a Taylor series at the point ξ, for some given determination of the dth root of ξ. Note that for any of the points $(x_2, \ldots, x_n)$ in question, the value of f is in $\mathcal{O}_S$, for a suitable determination of the dth root of ξ (because then the value coincides with $y_{\mathbf{x}}$). We may assume that the same determination holds for all of the points.

Conditions (1), (3), (4) of Theorem 1 are automatically verified. Also, condition (2) follows from the present condition (i). Let then $\mathbf{u}H$ be one of the finitely many translates in the conclusion of Theorem 1, whose union contains all the points $(x_2, \ldots, x_n)$; note that here we are working in $\mathbf{G}_m^{n-1}$. Then $f(\mathbf{X}) = P(\mathbf{X})$ on $\mathbf{u}H$, where $P \in K[X_2, \ldots, X_n]$. Because of the definition of f, this means that on $\mathbf{u}H$ the identity $P^d(X_2, \ldots, X_n) = \xi + X_2 + \cdots + X_n$ holds. (Recall that the "restriction" is interpreted formally.)

To read the identity in terms of the original coordinates, replace now X_j by $X_j\xi/X_1$. The coset $\mathbf{u}H$ becomes a coset $\mathbf{u}'H'$ in $\mathbf{G}_m^n$; if $\mathbf{u}H$ contains $(x_2, \ldots, x_n)$ then $\mathbf{u}'H'$ contains $(x_1, \ldots, x_n)$. Also, multiplying the identity by X_1 we get

$$X_1 + \cdots + X_n = \xi^{-1} X_1 P^d\left(\frac{X_2\xi}{X_1}, \ldots, \frac{X_n\xi}{X_1}\right),$$

valid on $\mathbf{u}'H'$. This proves the corollary. □

Proof of Theorem 3. – It is immediately seen that this follows from Proposition 3. In fact, assumption (1) of Theorem 1 is automatic, because we are dealing with S-unit points. Assumption (2) appears in the statement, and finally the growth condition of $f(\mathbf{x}_h)$ is also implied by the vanishing of these numbers, which is assumed in Theorem 3. Now, the conclusion of Proposition 3 yields the conclusion of Theorem 3, since the finitely many exceptions may be viewed as translates of the trivial 0-dimensional torus. □

REFERENCES

[1] BOMBIERI E., ZANNIER U., Algebraic points on subvarieties of $\mathbf{G}_m^n$, *Internat. Math. Res. Notices* **7** (1995) 333–347.

[2] BOMBIERI E., ZANNIER U., Heights of algebraic points on subvarieties of abelian varieties, *Ann. Scuola Norm Sup. Pisa Cl. Sci.* **XXIII** (4) (1996) 4.

[3] CORVAJA P., ZANNIER U., Some new applications of the subspace theorem, *Compositio Math.* **131** (2002) 319–340.

[4] CORVAJA P., ZANNIER U., On the diophantine equation $f(a^m, y) = b^n$, *Acta Arith.* **94** (1) (2000) 25–40.
[5] LAURENT M., Équations diophantiennes exponentielles, *Inv. Math.* **78** (1984) 299–327.
[6] MASSER D., A vanishing theorem for power series, *Inv. Math.* **67** (1982) 275–296.
[7] NISHIOKA K., Algebraic independence by Mahler's method and S-unit equations, *Compositio Math.* **92** (1994) 87–110.
[8] SCHMIDT W.M., Heights of points on subvarieties of $\mathbf{G}_m^n$, in: David S. (Ed.), *Number Theory 1993–1994*, Cambridge Univ. Press, LMS, 1996, pp. 199–213.
[9] ZANNIER U., Some Applications of Diophantine Approximation to Diophantine Equations, Editrice Forum, Udine, 2003.

(Manuscrit reçu le 11 décembre 2003 ;
accepté le 21 septembre 2004.)

Pietro CORVAJA
Dip. di Matematica e Informatica,
Via delle Scienze,
33100 Udine, Italy
E-mail: corvaja@dimi.uniud.it

Umberto ZANNIER
Scuola Normale Superiore,
Piazza dei Cavalieri, 7,
56100 Pisa, Italy
E-mail: u.zannier@sns.it