ACTA ARITHMETICA
XCIV.1 (2000)

On the diophantine equation $f(a^m, y) = b^n$

by

Pietro Corvaja (Udine) and Umberto Zannier (Venezia)

Introduction. A classical type of exponential diophantine equation takes the form $f(X, Y) = b^n$, where b is a fixed integer and f is a homogeneous polynomial. It is well known since the work of Thue and Mahler how to solve it. On the other hand, removing the assumption that f is homogeneous leads to problems which presently cannot be overcome. Here we show that something may be done by restricting another variable to be of exponential type.

Let $f(X, Y)$ be a polynomial with rational coefficients and let a, b be positive integers. Our aim is to prove that the equation in the title has only finitely many solutions in integers m, n, y, apart from "trivial" cases which can be classified. We achieve this goal under suitable technical hypothesis, labelled (i), (ii), (iii) in the Main Theorem below.

It seems that our method and results, in their general form, do not fall into known treatments of diophantine equations (see e.g. [ST] and the general theorems by M. Laurent [Lau] on mixed polynomial-exponential equations). This happens only in very special situations, like e.g. when f is homogeneous with respect to suitable weights.

Main Theorem. *Let $f(X, Y) = a_0(X)Y^d + a_1(X)Y^{d-1} + \ldots + a_d(X)$ be a polynomial with rational coefficients, of degree $d \geq 2$ in Y; let $a > 1, b > 1$ be integers. Suppose that*

(i) *a_0 is constant,*
(ii) *the equation $f(0, Y) = 0$ has no repeated roots,*
(iii) *a and b are not relatively prime.*

If the equation

$$f(a^m, y) = b^n \tag{1}$$

has an infinite sequence of solutions $(m, n, y) \in \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$, such that $\min\{|m|, |n|\} \to \infty$, then there exist an integer $h \neq 0$ and a polynomial

2000 *Mathematics Subject Classification*: Primary 11D61.

$p(X) \in \mathbb{Q}[X]$ such that $f(X^h, p(X))$ has only one term. Also, a, b are multiplicatively dependent.

REMARK 1. As mentioned above, standard methods lead to the conclusion when (i), (ii) and (iii) are replaced by the assumption that there exist nonzero integers $r,\ s$ such that $f(X^r, Y^s)$ is homogeneous.

Omission of the assumption $\min\{|m|, |n|\} \to \infty$ leads to the consideration of several possibilities. We have preferred to avoid them, since it is well known how to deal with the cases when m or n is fixed.

Finally, we remark that it is easy to generalize the arguments and to deal with certain equations $f(a_m, y) = b_n$, where a_m, b_n are S-units satisfying suitable assumptions which would replace (iii) above.

It is possible, thanks to the Proposition below, to find effectively all the data $h, p(X)$ appearing in the statement. Hence, given a polynomial $f(X, Y)$ satisfying the above assumptions, there exists an effective procedure to decide whether equation (1) has infinitely many solutions. Unfortunately, when there are only finitely many solutions our method does not yield effective results (but only bounds for the number of solutions) since it rests on the ineffective Subspace Theorem.

From our Main Theorem we can deduce the following

COROLLARY. *Let $d \geq 2$ be an integer, $u(X) \in \mathbb{Q}[X]$ be any polynomial, $a > 1$ be an integer. If the equation*

$$y^d = u(a^m) + a^n$$

has infinitely many integral solutions (m, n, y) with $\min\{|m|, |n|\} \to \infty$, there exist a nonzero integer h and a polynomial $p(X) \in \mathbb{C}[X]$ such that $p(X)^d - u(X^h)$ has just one term.

As for the Main Theorem, it is possible to test effectively the existence of such an integer h and polynomial p. We express these facts in the following Proposition, which represents the *functional* counterpart of the theorem.

PROPOSITION. (a) *Let f satisfy* (i) *and* (ii) *of the Theorem. There exist finitely many polynomials $p_i(X)$, which can be found effectively, such that $f(X^h, p(X))$ has only one term if and only if we have $p(X) = p_i(X^{h/e})$ for some i and some integer $e \mid d!$.*

(b) *All the identities $p(X)^d - u(X^h) = cX^k$ are obtained from finitely many of them, where $h = d$, by means of a monomial substitution $X \mapsto X^l$. Further, if $cX^k + u(X^h)$ is a dth power in $\mathbb{C}[X]$ (where $c \neq 0$), then $k \leq \frac{d}{d-1} h \deg u$.*

REMARK 2. The proof of the theorem shows in fact that, for every sequence of solutions satisfying our assumptions, almost all of them can be obtained by a substitution $X = a^r$ in some identity of the type $f(X^h, p(X)) =$

cX^k. The Proposition just stated allows us to find all the infinite families of solutions. (As observed in Remark 1, the families where some variable is bounded can be dealt with by standard techniques.)

As an amusing and simple application, one may classify the infinite families of squares having at most three non-zero digits in the decimal expansion: let $a_0 10^{n_0} + a_1 10^{n_1} + a_2 10^{n_2} = y^2$, $a_i \in [1, 9]$, $n_0 < n_1 < n_2$; we may absorb 10^{n_0} in the right side, and so assume $n_0 = 0$. Our theorem, applied with $f(X, Y) = (Y^2 - a_0 - a_1 X)/a_2$, gives identities $p(X)^2 = a_0 + a_1 X^h + a_2 X^k$. Simple considerations lead to $k = 2h$ and the classification is easily completed. We do not know how to classify squares with a given number $d \geq 4$ of non-zero digits.

Proofs. We begin with a lemma.

Lemma. *Let $g \in \mathbb{Q}[X, Y]$ be nonzero and suppose that for some integers $a, b > 1$ the equation $g(a^m, b^n) = 0$ has infinitely many solutions (m, n) in integers. Then there exist a nonzero rational number η and integers h, k, not both zero, such that $g(T^h, \eta T^k) = 0$. If $\min\{|m|, |n|\} \to \infty$ for a subsequence of solutions, then a, b are multiplicatively dependent and h, k may be chosen both nonzero. Let $\varepsilon, \delta \in \{\pm 1\}$ be such that εm and δn are both positive for an infinity of solutions (m, n). Then we can choose h, k such that $\varepsilon h, \delta k > 0$.*

Proof. It suffices to prove the assertions for the irreducible factors of g, and so let us assume that g is irreducible (over $\mathbb{Q}$). We apply Theorem 7.3, p. 207 of [La]. This implies that infinitely many solutions correspond to an absolutely irreducible factor of g of the form $\varrho X^k - Y^h$ or $\varrho - X^h Y^k$, for some nonzero complex number ϱ and for natural numbers h, k, not both zero. Since however $a, b \in \mathbb{Q}$, also ϱ must be rational (for otherwise such a factor would correspond to no solution of the relevant type). Since however g is irreducible over $\mathbb{Q}$, g itself must be a constant multiple of such a factor. In case g has the factor $\varrho X^k - Y^h$, the first claim follows by putting $\eta = \varrho$ and similarly for the other factor.

The assumption on $\min\{|m|, |n|\}$ forces h, k to be both nonzero and then the proof is easily completed as follows. Say that g is a constant multiple of $\varrho X^k - Y^h$. Taking two distinct solutions (m_i, n_i), $i = 1, 2$, we get the equations $\varrho a^{hm_i} = b^{kn_i}$, whence $a^{h(m_1 - m_2)} = b^{k(n_2 - n_1)}$, whence our conclusion about the multiplicative dependence. Also, the equation forces hm and kn to have the same sign for almost all the solutions. Similarly if g is a multiple of $\varrho - X^h Y^k$.

Proof of Main Theorem. Suppose that equation (1) has a sequence of solutions (m_i, n_i, y_i) as in the statement. The sequence m_i/n_i admits a convergent subsequence in $\mathbb{R} \cup \{\pm\infty\}$, so we can suppose that m_i/n_i converges. For notational convenience we drop the index i.

The proof splits into cases according to the limit of m/n. (It would be possible to argue without assuming convergence, according to the *order of magnitude* of m/n, suitably specified in terms of the relevant data of the problem. Such an approach would make it possible to estimate the number of solutions, provided there are only finitely many of them. Our presentation of the arguments is motivated by the simplification in the exposition.)

First case: $m/n \to \alpha \leq 0$. Consider first the case $n < 0$. If $m < 0$ for an infinite subsequence, then $\alpha = 0$. Also, the equation then forces $|m| \gg |n|$: it suffices to compare the denominators of both sides of it. Since this contradicts $\alpha = 0$, we can assume $m \geq 0$. But now the equation forces n to be bounded (compare again denominators), contrary to assumptions.

Therefore we assume $n \geq 0$. Let Δ be a common denominator for the coefficients of $a_1(X)/d$. By a variable change, replacing Y by $Y/\Delta - a_1(X)/d$ (which does not affect the assumptions), we can suppose that $a_1(X) = 0$. Then equation (1) reads

$$b^n - a_0 y^d = a_2(a^m)y^{d-2} + \ldots + a_d(a^m). \tag{2}$$

Pick an integer $r \in \{0, \ldots, d-1\}$ and an infinite subsequence of solutions such that $n \equiv r \pmod{d}$ for all of them. Observe that for large n we must have $y \neq 0$ (in fact b^n is much larger than a^m). Then we can write $n = r + ld$ and (2) takes the form

$$\left(\frac{b^l}{y}\right)^d - \frac{a_0}{b^r} = b^{-r}[a_2(a^m)y^{-2} + \ldots + a_d(a^m)y^{-d}].$$

Observe that $m/l \to d\alpha \leq 0$. Therefore the right side of the last displayed equation is $\ll_\varepsilon b^{\varepsilon l}$ for every positive ε. In turn this implies that $\log|y| \gg l$, whence for every $\delta > 0$,

$$\left|\frac{a_0}{b^r} - \left(\frac{b^l}{y}\right)^d\right| \ll |y|^{-2+\delta}$$

where the constant involved in the symbol $\ll$ depends only on the polynomial f and on δ.

Using the factorization $\alpha^d - \beta^d = (\alpha - \beta)(\alpha^{d-1} + \alpha^{d-2}\beta + \ldots + \beta^{d-1})$ with $\alpha = (a_0/b^r)^{1/d}$, $\beta = b^l/y$, for a suitable determination of the root we obtain

$$\left|\left(\frac{a_0}{b^r}\right)^{1/d} - \frac{b^l}{y}\right| \ll |y|^{-2+\delta}.$$

Put $\delta = 1/2$, say. Since $(a_0/b^r)^{1/d}$ is an algebraic number and its rational approximant b^l/y is a rational number of a special kind (its numerator is a power of a fixed integer b), we can apply Ridout's Theorem (one-dimensional Subspace Theorem) to deduce that either y is bounded or b^l/y is eventually constant (equal to $(a_0/b^r)^{1/d}$). In the first case l and n would also be

bounded (since $\log|y| \gg l$), a contradiction. In the second case we could write $y = cb^l$ where c is a nonzero constant. Put now

$$g(X, Y) := f(X, cY) - b^r Y^d.$$

If $g = 0$ identically, then $f(X, Y)$ would be a monomial in Y of degree ≥ 2, contrary to assumption (ii). Hence $g \neq 0$ and $g(a^m, b^l) = 0$ has an infinite sequence of integer solutions satisfying the assumptions of the above Lemma, whose conclusion gives what we want.

Second case: $m/n \to \alpha > 0$ where α is a positive number (not ∞). Suppose first that both $m, n \leq 0$. Then assumption (i) shows that $|y|$ is bounded. In fact, either $|y| \leq 1$ or

$$|a_0 y| \leq \sum_{i=1}^{d} \frac{|a_i(a^m)|}{|y|^{i-1}} + \frac{|b^n|}{|y|^{d-1}} \ll 1.$$

Therefore we may in fact work as if y were fixed and we may apply again the Lemma, which leads to the desired conclusion.

From now on we assume that $m, n \in \mathbb{N}$. Choose a prime number p dividing both a and b. We can certainly suppose that y converges p-adically to a solution y_0 of the equation $f(0, Y) = 0$. By hypothesis (ii), the derivative $\frac{\partial f}{\partial Y}(0, y_0)$ does not vanish. By the p-adic implicit function theorem (see [Se, Theorem 1, p. 83]) the solution $Y(x, z)$ to the equation $f(x, Y) = z$ in a neighborhood of $(0, y_0, 0)$ is an analytic function of x and z, admitting a Taylor series, converging in a neighborhood of $(0, 0)$,

$$Y(x, z) = \sum_{i,j} c_{i,j} x^i z^j$$

where $c_{i,j} \in \mathbb{Q}_p$ are p-adic algebraic numbers.

We fix a real number $H > 0$ which is sufficiently large to justify the subsequent arguments. Define

$$\mathcal{A}_H := \{(i, j) : 0 \leq i \leq H,\ 0 \leq j \leq H\}$$

and, for a solution (m, n, y), put

$$S(m, n) = S_H(m, n) = \sum_{(i,j) \in \mathcal{A}_H} c_{i,j} a^{mi} b^{nj}.$$

The above Taylor expansion and the fact that $m \gg n \ll m$ in the sequence under consideration give an inequality

$$|y - S(m, n)|_p < \theta^{nH} \tag{3}$$

for large n, where $\theta < 1$ is a suitable positive number depending only on a, b, f, α.

Let $\mathcal{B}_H$ be a minimal subset of $\mathcal{A}_H$ with the following property: *There exist p-adic algebraic numbers $\lambda_{i,j}$, for $(i,j) \in \mathcal{B}_H$ such that for an infinite subsequence of the solutions we are considering we have*

$$S(m,n) = \sum_{(i,j)\in\mathcal{B}_H} \lambda_{i,j} a^{mi} b^{nj}.$$

Observe that $\mathcal{A}_H$ trivially has the stated property (take $\lambda_{i,j} = c_{i,j}$), hence $\mathcal{B}_H$ exists.

From now on we shall restrict our attention to the subsequence of solutions satisfying the displayed equation relative to $\mathcal{B}_H$.

Observe that $|y|$ is bounded by a fixed power of $\exp(n)$. In fact y is a solution of a monic equation of degree d whose coefficients are so bounded, since we are in the second case.

Suppose first that $\mathcal{B}_H$ is empty. Then the mentioned bound for $|y|$ and (3) show that y must vanish for large enough n, provided H has been chosen large. In this case we are dealing with the equation $f(a^m, 0) = b^n$ and the Lemma immediately proves what we want. Therefore we assume that $\mathcal{B}_H$ is nonempty.

Let $\#\mathcal{B}_H = N > 0$ and, for notational convenience, fix once and for all an enumeration of the N pairs in $\mathcal{B}_H$. For $h = 1, \dots, N$ and for a given solution, let $\beta_h = \beta_h(m,n) = a^{mi} b^{nj}$ if (i,j) is the hth pair in $\mathcal{B}_H$.

By the defining property of $\mathcal{B}_H$ and by (3) we have

$$|y - \lambda_1\beta_1 - \ldots - \lambda_N\beta_N|_p < \theta^{nH} \tag{4}$$

for suitable p-adic algebraic numbers $\lambda_1, \dots, \lambda_N$.

We apply the Subspace Theorem to exploit the fact that the integer y is well approximated by a linear combination of S-units. Let S be the set of valuations of $\mathbb{Q}$ containing the infinite one and those dividing a, b. Let P be the point in $\mathbb{P}^N(\mathbb{Q})$ of homogeneous coordinates $P = (y : \beta_1 : \ldots : \beta_N)$. We denote by T_i, $i = 0, 1, \dots, N$, the ith homogeneous coordinate in $\mathbb{P}^N$.

For each $v \in S$, define $N+1$ independent linear forms $L_{v,i}(T_0, \dots, T_N)$ as follows. Put $L_{v,i}(T) = T_i$ if $i \neq 0$ or $v \neq p$ and $L_{p,0}(T) = T_0 - \lambda_1 T_1 - \ldots - \lambda_N T_N$. The coefficients of the linear forms $L_{v,i}$ lie in $\mathbb{Q}_v$ and are algebraic over $\mathbb{Q}$.

Let us estimate the double product

$$\prod_{v\in S}\prod_{i=0}^{N} \frac{|L_{v,i}(P)|_v}{\|P\|_v}$$

$$= |L_{p,0}(P)|_p \Big(\prod_{v\in S} \|P\|_v^{-(N+1)} \Big) \Big(\prod_{v\in S\setminus\{p\}} |y|_v \Big) \Big(\prod_{i=1}^{N}\prod_{v\in S} |\beta_i|_v \Big),$$

where $\|P\|_v$ is the v-adic maximum of the coordinates of P. The first factor is bounded by θ^{nH} by (4). The last factor is equal to 1 since each β_i is an S-unit. Since y is a rational integer we also have $\prod_{v\in S\setminus\{p\}} |y|_v \le |y|$. Then

$$\prod_{v\in S}\prod_{i=0}^{N}\frac{|L_{v,i}(P)|_v}{\|P\|_v} \ll \theta^{nH} H(P)^{-(N+1)}|y|.$$

By the Subspace Theorem [Schm, Theorem 1D′, p. 178], for every ε we also have a lower bound

$$\prod_{v\in S}\prod_{i=0}^{N}\frac{|L_{v,i}(P)|_v}{\|P\|_v} \gg H(P)^{-(N+1+\varepsilon)}$$

holding outside a certain finite union of rational hyperplanes (depending on ε).

It follows that, outside such an exceptional set,

$$1 \ll H(P)^{\varepsilon}|y|\theta^{nH}. \tag{5}$$

On the other hand, recalling that y is bounded by a fixed power of $\exp(n)$, we may estimate $H(P)$ directly, obtaining

$$H(P) \ll \Theta^{nH},$$

where $\Theta > 1$ is a suitable positive number depending only on the equation and on α. This estimate however contradicts (5) for small enough ε.

This means that a linear relation of the kind

$$A_0 y = A_1\beta_1 + \ldots + A_N\beta_N$$

with rational coefficients $A_0, \ldots, A_N$ not all zero, holds for an infinite subsequence of solutions. As before, from now on we shall consider only these solutions.

Suppose first that $A_0 = 0$. Then some other coefficient, say A_1, is not zero and we may write β_1 as a linear combination of β_i, $i > 1$, with fixed rational coefficients. Substituting for β_1 in the right side of the equation $S(m,n) = \sum_{(i,j)\in\mathcal{B}_H}\lambda_{i,j}a^{mi}b^{nj}$, we obtain a contradiction with the minimality of $\mathcal{B}_H$.

Therefore $A_0 \neq 0$. This means that for infinitely many solutions, y is a fixed polynomial function of a^m and b^n. Hence there exists a polynomial $q(X,Z) \in \mathbb{Q}[X,Z]$ such that, putting $g(X,Z) = f(X, q(X,Z)) - Z$, the equation $g(a^m, b^n) = 0$ has infinitely many solutions (m,n) with $\min\{m,n\} \to \infty$. Observe that $g(X,Z)$ is not identically zero. In fact, if $\deg_Z q \ge 1$, then $\deg_Z g \ge d \ge 2$, while, if q does not depend on Z, we have $\deg_Z g = 1$. By applying the Lemma we obtain the conclusions of the theorem. (Since the integers h, k in the statement of the Lemma may be chosen to be positive, $q(X^h, \eta X^k)$ is a polynomial.)

Third case: $m/n \to \infty$. If $m,\ n$ are both negative we argue as at the beginning of the second case (namely we deduce that $|y|$ is bounded and apply the Lemma). Therefore we assume that m, n are both nonnegative.

We factor $f(X,Y)$, viewed as a polynomial in Y, in an algebraic closure of $\mathbb{Q}(X)$, writing

$$f(X,Y) = a_0(Y - \alpha_1(X)) \dots (Y - \alpha_d(X))$$

where the algebraic functions $\alpha_1, \dots, \alpha_d$ admit a Puiseux expansion at infinity of the form

$$\alpha_i(X) = \sum_{j=-h}^{\infty} a_{i,j} X^{-j/e}, \tag{6}$$

converging for sufficiently large X. Here $a_{i,j}$ are algebraic complex numbers and e is a positive integer. We reduce to the case where $e = 1$ in a rather standard way: since there are infinitely many solutions there exists r such that for infinitely many of them $n \equiv r \pmod e$. Replacing X by $a^r X^e$ we obtain infinitely many solutions for an equation of the form $f_1(a^m, y) = b^n$ where $f_1 \in \mathbb{Q}[X,Y]$ is a polynomial such that the solutions to $f_1(X,Y) = 0$ at $X = \infty$ are Laurent series in X^{-1}. In conclusion we may argue as if $f_1 = f$, i.e. we may suppose $e = 1$ in (6).

In the sequel, for a solution (m,n,y) we write $x = a^m$ and we assume that all the Puiseux series converge at $X = x$.

Since $n/m \to 0$, for every positive δ we have

$$|f(a^m, y)| \ll |a|^{\delta m} = x^{\delta},$$

where the implied constant depends on δ, whence

$$|y - \alpha_1(x)| \dots |y - \alpha_d(x)| \ll x^{\delta}. \tag{7}$$

Let k be the minimal index such that the coefficients $a_{1,k}, \dots, a_{d,k}$ are not all equal. We remark at once that $k \le 0$. In fact, assume the contrary and put $P(X) = \sum_{j=-h}^{0} a_{1,j} X^{-j} \in \mathbb{C}[X]$. Then all the Puiseux series at $X = \infty$ of the polynomial $F(X,Y) := f(X, Y - P(X))$ are $O(1/X)$. Therefore all the coefficients of Y^j in F, for $j < d$, must be zero, for they are polynomials vanishing at ∞. So f would be a perfect dth power in $\mathbb{C}[X,Y]$, contrary to our assumption (ii).

Since for each j the set $\{a_{1,j}, \dots, a_{d,j}\}$ is invariant under conjugation over $\mathbb{Q}$, the algebraic numbers $a_{i,j}$ with $j < k$ are all rational, with common denominator D, say. So the substitution

$$Y \mapsto \frac{Y}{D} + \sum_{j=-h}^{k-1} a_{i,j} X^{-j}$$

does not affect the assumptions and we can suppose that $k = -h$. This is like assuming that the coefficients of X^{-h} in $\alpha_1, \dots, \alpha_d$ are not all equal.

Let us suppose, as we may, that for an infinite subsequence of solutions we have

$$|\alpha_1(x) - y| = \min_i\{|\alpha_i(x) - y|\}.$$

Renumbering indices we may also assume that $a_{1,-h} = \dots = a_{t,-h}$ and $a_{i,-h} \neq a_{1,-h}$ for $t < i \leq d$. Since we are in the case $k = -h$, we have $t < d$. Then for $i > t$ we have

$$|y - \alpha_i(x)| \geq |\alpha_i(x) - \alpha_1(x)| - |\alpha_1(x) - y| \geq |\alpha_i(x) - \alpha_1(x)| - |\alpha_i(x) - y|,$$

whence

$$|y - \alpha_i(x)| \geq \frac{|\alpha_i(x) - \alpha_1(x)|}{2} \gg |x|^h \quad \text{if } i > t.$$

From (7) we then get

$$|y - \alpha_1(x)|^t \ll x^{\delta - h(d-t)},$$

whence

$$|y - \alpha_1(x)| \ll x^{(\delta - h(d-t))/t}. \tag{8}$$

Assume first that $h \leq 0$. Then the coefficients of $f(X, Y)$, viewed as a polynomial in Y, are polynomials in X which are bounded at ∞. Therefore they are all constant and $f(X, Y) = f(Y)$ in fact does not depend on X. In particular, if (m, n, y) is a solution of (1), then (m', n, y) is also a solution, for all $m' \in \mathbb{Z}$. Therefore in this case we may conclude the proof by putting either $m' = [\sqrt{n}]$, falling in the first case, or $m' = n$, say, falling in the second case, both cases being treated above ([1]).

From now on we shall assume $h > 0$.

Let N be an integer with $N - h > h(d-t)/t$. Then, by Taylor expansion of α_1 up to the $(N-h)$th term and by (8), we get

$$\left| y - \sum_{j=-h}^{N-h-1} a_{1,j} x^{-j} \right| \ll x^{(\delta - h(d-t))/t} + x^{-(N-h)} \ll x^{(\delta - h(d-t))/t}. \tag{9}$$

We now shall proceed as in the second case, by applying the Subspace Theorem in the projective space $\mathbb{P}^N$, S now being the set containing the prime valuations dividing a and the infinite valuation.

The linear forms $L_{v,j}$ are defined by $L_{v,j} = T_j$ if $v \neq \infty$ or $j \neq 0$ and $L_{\infty,0} = T_0 - \sum_{j=-h}^{N-h-1} a_{1,j} T_j$.

([1]) Recall that we are presently assuming $\min\{m, n\} \to \infty$. This implies that actually the present case cannot occur at all, by known results on equations $f(y) = b^n$.

Put $P = (y : x^h : \ldots : x^{-(N-h-1)})$ and consider the double product

$$\prod_{v\in S}\prod_{i=0}^{N}\frac{|L_{v,i}(P)|_v}{\|P\|_v}$$

$$= |L_{\infty,0}(P)|_\infty \Big(\prod_{v\in S}\|P\|_v^{-(N+1)}\Big)\Big(\prod_{v\in S\setminus\{\infty\}}|y|_v\Big)\Big(\prod_{j=-h}^{N-h-1}\prod_{v\in S}|x^{-j}|_v\Big).$$

Since every power of x is an S-unit (recall that $x = a^m$) the last product is 1. Also, since y is an integer, the factor $\prod_{v\in S\setminus\{\infty\}}|y|_v$ is ≤ 1.

By (9) we thus get

$$\prod_{v\in S}\prod_{i=0}^{N}\frac{|L_{v,i}(P)|_v}{\|P\|_v} \ll x^{(\delta-h(d-t))/t}H(P)^{-N-1}.$$

The Subspace Theorem asserts that, for every $\varepsilon > 0$, we have the inequality

$$x^{(\delta-h(d-t))/t}H(P)^{-N-1} \gg H(P)^{-(N+1+\varepsilon)}$$

provided P lies outside a certain finite union of rational hyperplanes. Then for such P we get

$$1 \ll H(P)^{\varepsilon}x^{(\delta-h(d-t))/t}.$$

Since y is bounded by a fixed power of x, depending only on f, for large enough N the height of P is $\le |y|x^{N-h} \ll x^{2N}$. Therefore, since h and $d-t$ are positive the last displayed inequality cannot hold for small enough ε and δ.

Therefore we may assume that P belongs to one of the finitely many exceptional hyperplanes. In fact, we may assume that for an infinite subsequence of solutions P lies in a fixed hyperplane, say of equation $A_0T_0 + A_1T_1 + \ldots + A_NT_N$. This means that

$$A_0y + A_1x^h + \ldots + A_Nx^{h-N+1} = 0.$$

Now, A_0 cannot vanish, for otherwise we would obtain infinitely many roots for the nonzero rational function $A_1X^h + \ldots + A_NX^{h-N+1}$.

Hence we may divide by $-A_0$ and suppose that in fact $A_0 = -1$, so $y = A_1x^h + \ldots + A_Nx^{h-N+1}$. Since y is an integer and since $x = a^m$ is an integer tending to ∞, all the coefficients A_i with $h - i + 1 < 0$ must vanish. In conclusion, for our sequence of solutions, y is given by a certain polynomial function $p(x)$, where $p \in \mathbb{Q}[X]$.

Substituting this value for y into (1) and recalling that $x = a^m$, we see that the (nontrivial) equation

$$f(X, p(X)) = Z$$

has a sequence of solutions $X = a^m$, $Z = b^n$ with $\min\{m, n\} \to \infty$. Again, the above Lemma gives the required conclusion.

REMARK 3. Inspection of the proof (using the full force of the Lemma) in fact shows that the cases $\alpha \in \{0, \pm\infty\}$ cannot occur.

REMARK 4. The above arguments need some modification in case we work with a general number field K replacing $\mathbb{Q}$. While it is straightforward to deal with the first case and the second case, some problem may appear in the third and last case. In fact, the presence of several archimedean absolute values introduces extra factors in the double product, and some of the factors may be quite large (depending on the height of y). To compensate these contributions we need to improve on (8) and (9). For this we may need expansions in two variables, similarly to the second case. Taking this remark into account, it seems however that everything works nicely, at least assuming the analogue of (ii) at $X = \infty$.

Proof of the Corollary. If one of $m,\ n$ is negative, the other must also be negative. Then y must be bounded and we conclude e.g. by applying the Lemma. So, as before we assume that m, n are both nonnegative.

We can write $u(X) = X^e q(X)$ with $q(0) \neq 0$. Then the equation becomes

$$y^d = a^{me} q(a^m) + a^n. \tag{10}$$

Suppose first that q is a nonzero constant. Then the conclusion of the Corollary is trivially true (taking $h = d$). If q is not constant, then write $m = 2m' + r$, where r is 0 or 1. We may assume that the same value of r occurs infinitely often and so, replacing m' with m and $q(X)$ with $q(a^r X^2)$, we reduce to the case $\deg q \geq 2$.

We proceed by considering separately two possibilities, according as $n \geq me$ or not.

(1): $n \geq me$. Now, from (10) we find that a^{me} divides y^d. Hence we obtain a finite number of equations of the form

$$a' z^d = a^{n-me} + q(a^m)$$

where $a' \in \{1, a, \ldots, a^{d-1}\}$ (so a' may be assumed to be fixed). Suppose that $n - me$ is bounded, say equal to s for all solutions. Then, standard results show that $a^s + q(X^k)$ must be a dth power in $\mathbb{C}[X]$ for a suitable integer k (equivalently, one sees that all the roots of $a^s + q(X)$, with the possible exception of 0, have multiplicity divisible by d). This concludes the proof in this case. If $n - me$ is unbounded, we may assume that it tends to ∞ and apply the Main Theorem, by setting $f(X, Y) = a' Y^d - q(X)$. That conclusion easily leads to what we need.

(2): $n < me$. We again reduce to a finite number of equations of the kind

$$a' y^d = a^{me-n} q(a^m) + 1$$

which we rewrite as

$$y^d = c_k a^{mk-n} + \ldots + c_0 a^{me-n} + c,$$

where $c_k, \ldots, c_0, c$ are rational numbers and $k = e + \deg q$ is the degree of u. We can treat this case again by an application of the Taylor formula and the Subspace Theorem, in a similiar way to the second and third cases of the proof of the theorem. The technique is also very similar to the one used in [CZ]. The right term is

$$c_k a^{mk-n}\left(1 + \frac{c_{k-1}}{c_k}a^{-m} + \frac{c_{k-2}}{c_k}a^{-2m} + \ldots + \frac{c_0}{c_k}a^{-(k-e)m} + \frac{ca^{-mk+n}}{c_k}\right)$$

and, by expanding the dth root by Taylor's formula for $(1+X)^{1/d}$ and approximating it by a power sum plus a remainder term, we will get an approximation for an integer by a sum of S-units. This will allow an application of the Subspace Theorem as before (or as in [CZ], especially Lemma 2).

Actually, we can also give a complete proof by using Theorem 3 in [CZ]. For the reader's convenience we state an immediate corollary of it, sufficient for our purposes.

Claim. *Let $p \in \mathbb{Q}[X]$, $d \geq 2$ be an integer, $0 \leq r < d$ and $0 \leq \varrho < 1 - 1/d$. Suppose that the inequality*

$$|y^d - p(a^{r+ld})| \ll |p(a^{r+ld})|^{\varrho}$$

has infinitely many solutions in integers y and $l \geq 0$. Then there exists a power sum of the form $\beta(l) := \sum_{i=1}^{h} c_i b_i^l$, where $c_i \in \mathbb{Q}$, $b_i \in \mathbb{N}$, $b_1 > \ldots > b_h > 0$, such that $y = \beta(l)$ for an infinite subsequence of solutions (y, l).

We deduce at once that in fact we may take

$$\beta(l) = h(a^l) \tag{11}$$

for a suitable polynomial $h \in \mathbb{Q}[X]$. In fact, the inequality displayed in the Claim implies $|\beta(l) - (p(a^{r+ld}))^{1/d}| \ll \theta^l$ for some positive $\theta < 1$. On the other hand, expansion for the dth root of $p(a^{r+ld})$ gives an infinite series of type $\sum_{j=-r}^{\infty} \widetilde{c}_j a^{-jl}$. The preceding inequality shows that the *leading term* in $\beta(l) - \sum_{j=-r}^{\infty} \widetilde{c}_j a^{-jl}$ must be < 1 in absolute value. Since every b_i is a positive integer, we see that it equals some positive power of a, concluding the argument.

To go on, take again equation (10) and divide n by m, obtaining $n = mc + s$, where $0 \leq s < m$. Necessarily $0 \leq c < e$ and we may assume that in fact c is fixed. We divide throughout by a^{mc} and we again obtain an equation of the form

$$a' z^d = a^{m(e-c)} q(a^m) + a^s,$$

where $a' \in \{1, a, \dots, a^{d-1}\}$. Again, we may assume that a' is fixed. Since $\deg q > 2$ and $e - c \geq 1$ we have

$$a^s \ll (a^{m(e-c)} q(a^m))^{1/3}.$$

Therefore we may apply the Claim and its consequence (11), taking $p(X) = x^{e-c} q(X)/a'$. More precisely, for a subsequence of m's congruent to the same integer r modulo d we may write $m = dl + r$ and

$$a' = a^{rc}, \quad y = a^{lc} z, \quad z = h_1(a^l),$$

for some polynomial $h_1 \in \mathbb{Q}[X]$. Substituting into equation (10) and putting $a^l = T$, $h(T) = T^c h_1(T)$, we obtain an identity

$$h^d(T) = u(a^r T^d) + a^{rc+s} T^{dc}.$$

Finally, the substitution $X = a^{r/d} T$ proves what we want.

Proof of Proposition. (a) Suppose we are given an identity $f(X^h, p(X)) = cX^k$. Let e be a common multiple of the ramification indices of the Puiseux series in x of $f(X, Y) = 0$ above $X = 0$ and $X = \infty$. Then e divides $d!$. By substituting X^e in place of X, we may replace $f(X, Y)$ with $f(X^e, Y)$ and we may assume that 0 and ∞ are not ramified. In this way it suffices to prove the conclusion with $e = 1$. We assume that f is not constant in X, a trivial case easily dealt with.

We let $\alpha_1(X), \dots, \alpha_d(X) \in \mathbb{C}[[X]]$ (resp. $\beta_1(X), \dots, \beta_d(X) \in \mathbb{C}((1/X))$) be the Puiseux series of Y above $X = 0$ (resp. $X = \infty$) for the equation $f(X, Y) = 0$.

We begin to show that $p(X)$ is a polynomial in $X^{\pm h}$, so after another substitution we may assume $h = \pm 1$.

We start with the case $h > 0$. Assume the contrary and let X^m be a monomial appearing in $p(X)$, with degree m not divisible by h. In the formal power series ring $\mathbb{C}[[X]]$ we have the identity

$$\prod_{i=1}^{d} (p(X) - \alpha_i(X^h)) = cX^k, \quad c \neq 0. \tag{12}$$

Recall that the constant terms $\alpha_i(0)$ are pairwise distinct. Hence in the left side of (12) all the terms but one have order 0 at $X = 0$, so the remaining term must have order k. On the other hand the term containing X^m appears in every factor, so the order of each factor is $\leq m$. Therefore $k \leq m$.

Consider now the identity at infinity (i.e. in $\mathbb{C}((1/X))$)

$$\prod_{i=1}^{d} (p(X) - \beta_i(X^h)) = cX^k. \tag{13}$$

Even now each factor contains X^m and therefore the order at $X = \infty$ is $\leq -m$ in each factor. We conclude that $k \geq dm$ and comparison with the previous bound gives a contradiction.

On the other hand, assume $h = -h' < 0$. As before, but with a substitution $X \mapsto 1/X$, we get the equation

$$\prod_{i=1}^{d}(p(1/X) - \alpha_i(X^{h'})) = cX^{-k} \tag{14}$$

valid in $\mathbb{C}((X))$ and the equation

$$\prod_{i=1}^{d}(p(X) - \beta_i(X^{-h'})) = cX^{k} \tag{15}$$

valid again in $\mathbb{C}((X))$ (because $h' > 0$).

From equation (14) we obtain $k = d \deg p$ if $\deg p > 0$, as we can assume (if $\deg p = 0$ the claim we are proving is automatic). Let again X^m be a monomial appearing in p, where m is not a multiple of h. Then $m \leq \deg p$. In equation (15) we compare the order at $X = 0$ in both sides. Observe that at least one of the terms has a pole. (In fact, assume the contrary. Then, since f is monic in Y we have

$$f(1/X, Y) = \prod_{i=1}^{d}(Y - \beta_i(X^{-h'})).$$

Now, if no term on the right has a pole at $X = 0$, we have that f is constant in X, against our assumptions.)

The other terms have order $\leq m$, whence the total order is $\leq m(d-1) - h' < k$, a contradiction.

From now on we assume $h = 1$ (the case $h = -1$ being similar and even simpler). By replacing Y with $Y - a_1(X)/d$ we may assume that f, as a polynomial in Y, has vanishing second coefficient. We show first that $D := \deg p$ and k are bounded effectively in terms of f. From (13) with $h = 1$ it follows immediately that if D is large enough, then $k = dD$. If we have $p^d(X) = cX^{dD}$, then it is easy to see that $f(X, p(X))$ has at least two terms for large D, a contradiction. Otherwise we have

$$\deg(p^d(X) - cX^{dD}) \leq \delta + D(d-2),$$

where $\delta := \deg_X f$. On the other hand, $p^d(X) - cX^{dD} = \prod_{\zeta^d=c}(p(X) - \zeta X^D)$, and all terms but one in the product have degree $\geq D$, a contradiction if $D > \delta$.

Now, the arguments used at the beginning prove also that $p(X) = p_1(X) + X^k p_2(X)$ for polynomials $p_1,\ p_2$ such that p_1 has degree $\leq k-1$ and coincides with the first k terms of some α_i. In particular p_1 has finitely

many possibilities which can be computed. Putting $\beta_i^*(X) = \beta_i(X) - p_1(X)$ we have

$$\prod_{i=1}^{d}(X^k p_2(X) - \beta_i^*(X)) = cX^k. \tag{16}$$

If in some factor some term in $X^k p_2(X)$ does not disappear, that factor will have order $\leq -k$ at $X = \infty$. But (16) says that this can happen for at most one factor. Therefore in some other factor all the mentioned terms must disappear. This shows that p_2 will coincide with some *subsum* of the polynomial part of the series of some β_i^*, whence it has finitely many possibilities which can be computed.

(b) Take an equation $p(X)^d = u(X^h) + cX^k$, $c \neq 0$, and substitute in it ζX for X, where $\zeta^h = 1$. Subtracting we obtain

$$\prod_{\theta^d=1}(p(X) - \theta p(\zeta X)) = c(1 - \zeta^k)X^k. \tag{17}$$

A first case is when $1 = \zeta^k$ for all hth roots of unity ζ. This means that $k = mh$ is a multiple of h. In turn, this easily implies $p(X) = X^r p_1(X^h)$ for an integer $r \leq h$ such that $rd = sh$ is a multiple of h (so $s \leq d$). Now, our equation is obtained by monomial substitution from

$$(X^r p_1(X^h))^d = u(X^h) + cX^{md}.$$

A second case occurs when $1 \neq \zeta^k$ for some ζ and $d \geq 3$. Then all the factors on the left side of (17) have exactly one term. Since at most one term can have degree $< \deg p$, we see that p itself has just one term and the proof is easily completed.

The third and last case is when $d = 2$ and $1 \neq \zeta^k$ for some ζ. The same argument as before shows that p has at most two terms, so we write $p(X) = \mu X^E + \nu X^F$. Now, inserting this into the original equation shows that at least two among $2E, E + F, 2F$ are divisible by h. This implies that $2E$ and $2F$ must be divisible by h and we are reduced to the case $h = 2 = d$.

The last assertion of the Proposition is proved simply as follows. Suppose that

$$cX^k + u(X^h) = (c_0 X^{m_0} + c_1 X^{m_1} + \ldots)^d,$$

with nonzero coefficients c_i and decreasing degrees m_i. Assume $k > h \deg u$. Then $k = m_0 d$. Also, the monomial $X^{(d-1)m_0+m_1}$ appears on the right side, so $(d-1)m_0 \leq (d-1)m_0 + m_1 \leq h \deg u$. This concludes the argument.

Acknowledgements. The authors are grateful to the referee for several helpful remarks on a previous draft of the present paper.

Part of this paper was written during a stay at Institut Henri Poincaré in Paris, during the special trimester on Arakelov Theory. The authors wish to

thank the organizers, especially professors Sinnou David and Patrice Philippon, for their kind invitation and the I.H.P. for hospitality and financial support.

References

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

[La] S. Lang, *Fundamentals of Diophantine Geometry*, Springer, 1983.

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

[Schm] W. M. Schmidt, *Diophantine Approximations and Diophantine Equations*, Lecture Notes in Math. 1467, Springer, 1991.

[Se] J.-P. Serre, *Lie Algebras and Lie Groups*, Lecture Notes in Math. 1500, Springer, 1991.

[ST] T. N. Shorey and R. Tijdeman, *Exponential Diophantine Equations*, Cambridge Tracts in Math. 87, Cambridge Univ. Press, 1986.

Dipartimento di Matematica
Università di Udine
via delle Scienze 206
33100 Udine, Italy
E-mail: corvaja@dimi.uniud.it

Istituto Univ. di Architettura, D.C.A.
S. Croce 191
30135 Venezia, Italy
E-mail: zannier@brezza.iuav.unive.it

Received on 19.10.1998
and in revised form on 22.11.1999 (3482)