ANNALES

DE

L'INSTITUT FOURIER

Umberto ZANNIER

On the integer solutions of exponential equations in function fields

Tome 54, nº 4 (2004), p. 849-874.

<http://aif.cedram.org/item?id=AIF_2004__54_4_849_0>

Ann. Inst. Fourier, Grenoble
54, 4 (2004), 849–874

ON THE INTEGER SOLUTIONS OF EXPONENTIAL EQUATIONS IN FUNCTION FIELDS

by Umberto ZANNIER

1. Introduction and statements.

It is a rather well-known problem in Number Theory to estimate the number of zeros of linear recurrences of algebraic numbers or, more generally, of polynomial-exponential equations over number fields. We do not pause here on references, since the subject has been widely investigated, but we refer to [Schm] for this and for a general overview.

For recurrences and polynomial-exponential equations over function fields similar results are available. Sometimes they may be reduced to the number-field case by means of specialization arguments, as e.g. in [Schm, §9]. However, for equations which are "truly" defined over function fields other (more elementary) tools seem more efficient (as in the *abc*-theorem - see [BrM]).

For example, in [BMZ, Thm. 2] a method with derivations has been applied to estimate the number of zeros of recurrences in one variable. (See also [Schm, §10] for an analogous approach.) The purpose of the present paper is to carry out a similar analysis for polynomial-exponential equations in several variables. We stress that the alluded proofs do not work automatically for several variables; we briefly outline the main points. The approach in [BMZ] or [Schm, §10] relied in an essential way on viewing the

Keywords: Number theory – Diophantine equations – Function fields.
Math. classification: 11D45 – 11D61 – 11D99.

solutions in question as (integer) zeros of a certain (Wronskian) polynomial in one variable; their number was then bounded by the degree of that polynomial. Naturally, for several variables such a principle is not sufficient. We have succeeded in adapting the approach at the cost of analyzing first the whole set of solutions in complex numbers. The corresponding equations are not even well-defined in that extended context, and so we have found it convenient to embed first the relevant fields into fields of complex-valued algebraic functions in one variable, in order to obtain meaningful notions. (See §2 and Remark 2 below.)

In this way we have obtained bounds which, for a given number of variables, have polynomial growth in the number of terms; this substantially improves on the known estimates over number fields (see [Schm]). As for the dependence on the number of variables, the estimates are simply-exponential, like for number fields (see [ESS]). On the other hand, we have no significant lower bounds so we do not know to what extent the present estimates might possibly be improved.

Among others, we have given applications to the number of zeros of certain recurrences over function fields and to a problem studied in the recent paper [FPT].

Some notation. — We let k be an algebraically closed field, supposed to be embedded in $\mathbb{C}$. We let L/k be a field extension of transcendence degree 1; L is the function field of a certain nonsingular curve over k. Actually, the results below hold for arbitrary finitely generated extensions L/k. However the present case suffices for many applications. For this reason we shall give complete proofs for the case in question and just a hint of the general case in Remark 3 below.

We shall work with vectors $A = (\alpha_1, \dots, \alpha_r) \in (L^*)^r$, using coordinatewise multiplication. For notational convenience we shall use the rule $A^{\mathbf{u}} = \alpha_1^{u_1} \cdots \alpha_r^{u_r}$ for a vector $\mathbf{u} = (u_1, \dots, u_r)$ of integers. (In the course of §2 this notation will be suitably extended to complex vectors $\mathbf{u} \in \mathbb{C}^r$.)

For given $A_1, \dots, A_h \in (L^*)^r$ and $P_1, \dots, P_h \in L[\mathbf{X}] = L[X_1, \dots, X_r]$, rather than dealing directly with equations $\sum_{i=1}^h P_i(\mathbf{m}) A_i^{\mathbf{m}} = 0$, we shall be concerned with the set of integral vectors $\mathbf{m} \in \mathbb{Z}^r$ such that the $P_i(\mathbf{m}) A_i^{\mathbf{m}}$, $i = 1, \dots, h$, are linearly dependent over k. This will turn out more convenient and also will lead to sharper conclusions. Let us then formally state:

DEFINITION 1. — *For $A_1, \dots, A_h \in (L^*)^r$ and $P_1, \dots, P_h \in L[\mathbf{X}] = L[X_1, \dots, X_r]$ we let $S = S(A_1, P_1, \dots, A_h, P_h)$ be the subset of $\mathbb{Z}^r$ made up of the vectors $\mathbf{m} \in \mathbb{Z}^r$ such that the $P_i(\mathbf{m})A_i^{\mathbf{m}}$, $i = 1, \dots, h$, are linearly dependent over k.*

Plainly, when for example all the $A_iA_j^{-1}$ lie in k^r, the question reduces to the dependence of the $P_i(\mathbf{m})$ over k. More generally, it may happen that for a proper subset B of $\{1, \dots, h\}$ the elements $P_i(\mathbf{m})A_i^{\mathbf{m}}$, $i \in B$, are already linearly dependent over k; again, if the quotients $A_iA_j^{-1}$, $i, j \in B$, lie in k^r, the question amounts to the dependence of the polynomial terms. This suggests to group the relevant integral vectors into classes. We do this according to the following definition.

DEFINITION 2. — *For $A_1, \dots, A_h, P_1, \dots, P_h$ as above, let B be a nonempty subset of $\{1, \dots, h\}$. We say that a set $S' \subset \mathbb{Z}^r$ is a class relative to B if the following conditions are verified: (i) For $\mathbf{m} \in S'$ the elements $P_i(\mathbf{m})A_i^{\mathbf{m}}$, $i \in B$, are linearly dependent over k; (ii) for some $\mathbf{m}_0 \in S'$, the set S' consists of all the $\mathbf{m}$ satisfying (i) and such that for $i, j \in B$ we have $(A_iA_j^{-1})^{\mathbf{m}-\mathbf{m}_0} \in k^*$.*

Observe that for given i, j the set of vectors $\mathbf{m} \in \mathbb{Z}^r$ such that $(A_iA_j^{-1})^{\mathbf{m}} \in k^*$ is a subgroup of $\mathbb{Z}^r$; hence, observe that if (ii) is true for some $\mathbf{m}_0 \in S'$ then it is true for any $\mathbf{m}_0 \in S'$. As a further motivation to this definition note that when for instance the P_i's are constant, then if S' is a class relative to B, and $\mathbf{m}_0$ a fixed element of S', the linear dependence over k of the $P_iA_i^{\mathbf{m}}$, $i \in B$, for any $\mathbf{m} \in S'$ is equivalent to the linear dependence over k of the $P_iA_i^{\mathbf{m}_0}$, $i \in B$. Now the classes consist of certain cosets of the alluded subgroups.

We remark that for "general" A_i's in $(L^*)^r$ these subgroups may well be trivial as soon as $i \neq j$. In these cases the classes consist of single points if $\#B > 1$.

Further, note that certainly S may be expressed as a union of classes, but this representation needs not to be unique, nor the relevant classes need to be disjoint.

With these conventions we state our main result.

THEOREM 1. — *Let $A_1, \dots, A_h \in (L^*)^r$ and let $P_1, \dots, P_h \in L[\mathbf{X}]$ satisfy $\deg P_i \leqslant d_i$. Then the above set S may be expressed as a union of no more than $\left(d_1 + \cdots + d_h + \binom{h}{2}\right)^r$ classes.*

In a given class in practice we have to establish the dependence over k of polynomials in $\mathbf{m}$, defined over L. Of course this may lead to any possible diophantine equation, so there is no hope to describe this point in a satisfactory way; however the exponentials disappear from the context. Also, as remarked above, when for instance the P_i's are constant, a class is determined by a single point in it and by certain explicit multiplicative relations among the coordinates of the A_i's.

We shall now present some corollaries. Our choice is however rather special and it will be clear that Theorem 1 allows several other possibilities, not taken into account here.

Naturally, when each class consist of a single point, which "often" holds as remarked above, Theorem 1 gives an estimate for $\#S$. For instance we have the following:

COROLLARY 1. — *Let $P_1, \ldots, P_h, A_1, \ldots, A_h$ be as in Theorem 1. Assume that for all $i \neq j$ the coordinates of $A_i A_j^{-1}$ are multiplicatively independent modulo k^*. Then:*

(a) *There are at most $(d_1 + \cdots + d_h + \binom{h}{2})^r$ solutions $\mathbf{m} \in \mathbb{Z}^r$ to the equation $\sum_{i=1}^h P_i(\mathbf{m}) A_i^{\mathbf{m}} = 0$ with $P_1(\mathbf{m}) \cdots P_h(\mathbf{m}) \neq 0$.*

(b) *If we also assume that the coordinates of each A_i are multiplicatively independent modulo k^*, then there are at most $(d_1 + \cdots + d_h + \binom{h+1}{2})^r$ integer points such that $\sum_{i=1}^h P_i(\mathbf{m}) A_i^{\mathbf{m}}$ is in k and $P_1(\mathbf{m}) \cdots P_h(\mathbf{m}) \neq 0$.*

(c) *Let P_i have at most t_i terms. Then there are at most $\binom{t_1 + \cdots + t_h}{2}^r$ solutions $\mathbf{m} \in \mathbb{Z}^r$ to $\sum_{i=1}^h P_i(\mathbf{m}) A_i^{\mathbf{m}} = 0$ such that not all the $P_i(\mathbf{m})$ vanish.*

(We recall that elements $\lambda_1, \ldots, \lambda_r \in L^*$ are said to be *multiplicatively independent modulo* k^* if no relation $\lambda_1^{a_1} \cdots \lambda_r^{a_r} \in k^*$ holds with integers a_i not all zero.)

Part (b) answers a question of W.M. Schmidt. (For the case of a single variable this was remarked in [BMZ], §5.) We also note that the conclusions in [Schm, §10], or [BMZ, Thm. 2, Thm. 3] are easy corollaries of Theorem 1, with $r = 1$ (see Lemma 2 below for an instance).

Other applications are to recurrence sequences. We illustrate this with an example which will be useful for Corollary 3. Let $G(n), H(n)$ be *simple* linear recurrences over L, *i.e.* power sums

$$G(n)=\sum_{i=1}^{p}a_i\alpha_i^n,\quad H(n)=\sum_{i=1}^{q}b_i\beta_i^n,\quad p,q\geqslant 1,$$

where $a_i,b_i,\alpha_i,\beta_i\in L^*$, for a field L as above. Then we have:

COROLLARY 2. — *Assume that no α_i or β_j and no ratio α_i/α_j or β_i/β_j, $i\neq j$, lies in k^*. Then:*

(a) *The equation $G(u)=H(v)$ in integers u,v has at most $\binom{p+q}{2}^3$ solutions, unless there are integers u_0,v_0,r,s, with $rs\neq 0$, such that the identity $G(u_0+rm)=H(v_0+sm)$ holds for $m\in\mathbb{Z}$.*

(b) *The equation $G(u)=cH(v)$, $c=c(u,v)\in k^*$ has at most $\binom{p+q}{2}^3$ solutions $(u,v)\in\mathbb{Z}^2$, unless there are integers u_0,v_0,r,s, with $rs\neq 0$, and elements $\xi,\eta\in k^*$ such that the identity $G(u_0+rm)=\eta\xi^m H(v_0+sm)$ holds for $m\in\mathbb{Z}$.*

Plainly, the identity $G(u_0+rm)=H(v_0+sm)$ for $m\in\mathbb{Z}$ in part (a) (respectively $G(u_0+rm)=\eta\xi^m H(v_0+sm)$ for $m\in\mathbb{Z}$ with $\xi,\eta\in k^*$ in part (b)) implies that $p=q$ and that the pairs $(a_i\alpha_i^{u_0},\alpha_i^r)$ coincide in some order with the pairs $(b_i\beta_i^{v_0},\beta_i^s)$ (respectively $(\eta b_i\beta_i^{v_0},\xi\beta_i^s)$ in (b)). The proof also easily shows that in any case there are at most $\binom{p+q}{2}^3$ solutions which do not "come" from such an identity; and it is also rather easily seen that all such identities come from a "minimal" one by substitution $m\mapsto a+bm$. We do not give the easy proofs of these sharpenings since they fall somewhat outside the scope of the paper.

As announced we shall apply the result to the main problem treated in [FPT] (and in previous papers quoted therein); we use in part the notation therein. We let $A_0(X),\dots,A_d(X)\in k[X]$ and we consider a recurrence of polynomials $G_n(X)\in k[X]$ satisfying

$$G_{n+d}(X)=A_{d-1}(X)G_{n+d-1}(X)+\cdots+A_0(X)G_n(X),\quad n\in\mathbb{N}.$$

We assume that d is minimal and as in [FPT] we assume that the characteristic polynomial

$$Q(X,T):=T^d-A_{d-1}(X)T^{d-1}-\cdots-A_0(X)\in k[X][T]$$

has no multiple roots; we let $\alpha_1,\dots,\alpha_d$ be its distinct roots in a fixed finite algebraic extension L of $k(X)$. Then G_n admits a representation

$$G_n(X)=a_1\alpha_1^n+\cdots+a_d\alpha_d^n,$$

where $a_\imath$ are also in L^*. Further, let $P(X) \in k[X]$. We shall deal with the equation in $u, v \in \mathbb{N}$,

$$(*) \qquad G_u = c \cdot G_v \circ P, \quad u, v \in \mathbb{N},$$

where $c = c(u,v) \in k^*$ may depend on u, v. In [FPT] the authors use deep results of Evertse-Györy [EG] and of Evertse-Schlickewei-Schmidt [ESS] to give, under various conditions (see [FPT], Thms. 2.1-2.7) an upper bound for the number of solutions $(u,v) \in \mathbb{N}^2$ of $(*)$.

We do not repeat here the conditions, nor the bounds, which are somewhat complicated to state. As remarked in [FPT] some assumption is necessary for finiteness, in view of the equation $T_{2n}(X) = T_n(2X^2 - 1)$, any $n \in \mathbb{N}$, valid for the Chebyshev polynomials $T_n(X) = \cos(n \arccos X)$; in fact, this equation holds more generally as $T_n \circ T_p = T_{np}$ (see [S]) and there is also a simpler example, obtained by putting $G_n(X) = X^n$, $P(X) = X^p$. Plainly these examples lead to more general similar ones by summation; e.g. we have $G_n(T_p(X)) = G_{np}(X)$ when G_n is a linear combination $\sum c_\jmath T_{\jmath n}$ of Chebyshev polynomials and similarly for the other case. Also, we may change linearly the variable X. Roughly speaking, we shall call these cases and analogous ones the *Chebyshev case* and the *Cyclic case* respectively.

Here, with the present essentially self-contained methods, we shall improve on the paper [FPT] in two respects:

(i) We shall substantially sharpen the doubly exponential bounds of [FPT, Thms. 2.1, 2.3], obtaining a polynomial bound; also, we shall completely eliminate the dependence on $\deg P$ and on the degree of the discriminant of Q (with respect to T).

(ii) We shall show that the Cyclic and Chebyshev cases are the only ones when infinitely many solutions may arise, describing completely the matter.

More precisely we have, with the above notation:

COROLLARY 3. — *Suppose that* $\deg P \geqslant 2$ *and that no* $\alpha_\imath$ *and no ratio* α_i/α_j, $i \neq j$, *lies in* k. *Then if there are only finitely many solutions* (u,v) *for equation* $(*)$, *their number is at most* $\binom{2d}{2}^3 < 8d^6$.

If there are infinitely many solutions then for suitable $r, s \in \mathbb{N}$ *we have an identity*

$$G_{sn+v_0}(P(X)) = \eta \xi^n G_{rn+u_0}(X), \quad n \in \mathbb{N}, \quad |r| = |s| \deg P > 0,$$

for suitable $\xi, \eta \in k^*$, *and two cases may occur.*

Case A: *We are in the "Cyclic case", namely P is of the form $\lambda' \circ X^p \circ \lambda$ for suitable $\lambda, \lambda' \in \mathrm{Aut}(\mathbb{P}_1) = PGL_2(k)$. Also, the α_i are in $k(X)$, of the form $c_i X^{\delta_i} \circ \lambda$, for integers δ_i and $c_i \in k$.*

Case B: *We are in the "Chebyshev case", namely $P(X) = \lambda' \circ T_p \circ \lambda$, λ, λ' as above. The α_i are quadratic over $k(X)$ and of the form $c_i(X \pm \sqrt{X^2-1})^{\delta_i} \circ \lambda$.*

We omit even more explicit formulas that can be written down for $G_n(X)$, in order not to complicate further the already long statement; we believe that the present one is sufficiently illustrative. Also, as in the remarks after Corollary 2 above, the result may be made even sharper, by classifying the infinite families of solutions (as coming from substitutions in an identity of the stated type) and by bounding the number of the remaining ones. Moreover, it is possible to discuss the case $\deg P = 1$: now there may be infinitely many solutions only if P has finite order by composition and the polynomial Q is "essentially" invariant by $X \mapsto P(X)$. We do not pursue however in the task of giving the complete proofs of these further conclusions, since our main purpose for introducing these corollaries is to exemplify some applications of Theorem 1.

Further corollaries of Theorem 1 may be obtained about the solutions of $x_1 + \cdots + x_h = 1$ where the x_i lie in a subgroup of L^* which is finitely generated modulo k^*. However, of course we cannot treat with the present methods the constant solutions, which give rise to an extraordinarily deeper problem, solved to a large extent by Evertse, Schlickewei and Schmidt [ESS], who give quite remarkable estimates; also, Evertse and Győry [EG] have studied a function-field analogue. Here we just give an instance, by considering an equation in two variables, of S-unit type.

COROLLARY 4. — *Let $\lambda, \mu \in L^*$ and let Γ be a subgroup of L^* containing k^* such that Γ/k^* has rank r. Then the equation $\lambda x + \mu y = 1$, $x, y \in \Gamma$, has at most 9^r solutions such that $\lambda x/\mu y \notin k^*$.*

This slightly improves on a result by Evertse [E], who obtained the estimate $2 \cdot 7^{2s}$ when Γ is the group $\mathcal{O}_S^*$ of S-units in L^*, for a set S of places with $\#S = s$ (note that $\mathcal{O}_S^*/k^*$ has rank $\leqslant s-1$). The same method used here for the proof yields analogous results on the structure of the solutions in several variables, with estimates independent of the genus of L/k (see also the Acknowledgment below).

We finally remark that the present methods, though sometimes efficient in estimating the number of solutions, do not give informations about heights of solutions; for that purpose other tools are available, as for instance the multidimensional extension of Mason's *abc*-theorem due to Brownawell and Masser [BrM] (see also [Z]).

Acknowledgment. — I wish to thank Prof. J.-H. Evertse for his kind interest and several most important comments. He has also pointed out a possible application of the methods to a certain generalization of Corollary 4 to linear equations in several unknowns from a group Γ as above; such a result is now the topic of a joint preprint with Evertse [EZ].

It is also a pleasure to thank an anonymous referee for an extremely detailed and helpful review, which led in particular to a substantial improvement in the presentation of the proof of Corollary 2.

2. A complex-numbers analogue.

This section will analyze an analogue problem for complex numbers in place of integers; it will be a crucial tool for the proof of the main results. As a preliminary, we introduce certain conventions about exponentials, to be used only in the proofs of Lemma 1 and Proposition 1.

We let L be a finite extension of $\mathbb{C}(z)$, where z is transcendental over $\mathbb{C}$; we denote by ∂ the derivation d/dz and extend it to L; for notational convenience we put $\alpha^\partial := \partial(\alpha)$ for $\alpha \in L$.

Warning. — This field L need not be the same as that from Theorem 1; however, we shall embed the field from Theorem 1 into one of the shape introduced here; this will be done in the deduction of Theorem 1 from Proposition 1 below.

In the proof of the next Lemma 1 and Proposition 1 we shall need to define expressions α^u for $\alpha \in L^*$ and $u \in \mathbb{C}$; actually, at each time we shall need such definitions only for the α's in a fixed finitely generated subgroup Γ of L^*. We proceed in two steps, as follows. We first express Γ as a direct product of (possibly finite) cyclic groups generated by $\gamma_1, \ldots, \gamma_l$, say. It will then suffice to define γ_i^u and to extend by linearity the definition to all Γ: namely, to put, for $\gamma = \gamma_1^{a_1} \cdots \gamma_l^{a_l} \in \Gamma$, $\gamma^u := \prod \gamma_i^{a_i u}$; of course this is not well defined on a possible finite cyclic factor $\langle \gamma_j \rangle \cong \mathbb{Z}/(m)$, but we may agree to choose the integer exponent a_j in $[0, m-1]$ in that case.

Now, to define γ_i^u, recall that, as is well known (see [Ch, Ch. VII]), we may view L as a field of meromorphic functions of z on a suitable simply connected open set $\Omega \subset \mathbb{C}$; e.g. we may take Ω as a disk not containing any branch point for $L/\mathbb{C}(z)$. By shrinking Ω we may assume that the elements γ_i, $i = 1, \ldots, l$, are holomorphic functions of the variable z, without zeros in Ω. Then we may define the functions $\log \gamma_i$, $i = 1, \ldots, l$, as holomorphic functions on Ω and put $(\gamma_i^u)(z) := \exp(u \log \gamma_i(z))$. [1] Note that with this definition we have $(\gamma^u)^\partial/(\gamma^u) = u(\gamma^\partial/\gamma) \in L$ for any $\gamma \in \Gamma$; also, if $u = m \in \mathbb{Z}$ the definition is consistent with the usual meaning.

Moreover, this definition will not affect the usual properties of exponentials if one restricts to the free part of Γ; in any case the properties are satisfied modulo constants. Namely, we have $\gamma^{u+v} = \gamma^u \gamma^v$ and, up to nonzero constant factors, $(\gamma^u)^m \equiv (\gamma^m)^u \equiv \gamma^{mu}$, $(\gamma\gamma')^u \equiv \gamma^u \gamma'^u$ for $\gamma, \gamma' \in \Gamma$; also, $\gamma^u \in \mathbb{C}^*$ for $\gamma \in \mathbb{C}^* \cap \Gamma$. These properties will amply suffice for our purposes.

Recall also the convention $A^{\mathbf{u}} := \prod_{i=1}^r \alpha_i^{u_i}$ for $A = (\alpha_1, \ldots, \alpha_r)$, $\alpha_i \in \Gamma$, $\mathbf{u} = (u_1, \ldots, u_r)$.

In analogy with Definition 1 above, given $A_1, \ldots, A_h \in \Gamma^r$ and $P_1, \ldots, P_h \in L[\mathbf{X}]$, we let $\mathcal{S}$ be the set of points $\mathbf{u} \in \mathbb{C}^r$ such that the $P_i(\mathbf{u})A_i^{\mathbf{u}}$ are linearly dependent over $\mathbb{C}$. Also, we define a class as a set $\mathcal{S}' \subset \mathbb{C}^r$ similarly to Definition 2 above, namely

Definition 3. — *For $A_1, \ldots, A_h, P_1, \ldots, P_h$ as above, let B be a nonempty subset of $\{1, \ldots, h\}$. We say that a set $\mathcal{S}' \subset \mathbb{C}^r$ is a class relative to B if the following conditions are verified: (i) For $\mathbf{u} \in \mathcal{S}'$ the elements $P_i(\mathbf{u})A_i^{\mathbf{u}}$, $i \in B$, are linearly dependent over k; (ii) there exists $\mathbf{u}_0 \in \mathbb{Q}^r$ such that $\mathcal{S}'$ consists of all the $\mathbf{u}$ satisfying (i) and such that for $i, j \in B$ we have $(A_i A_j^{-1})^{\mathbf{u}-\mathbf{u}_0} \in \mathbb{C}^*$.*

We remark that now the conditions $(A_i A_j^{-1})^{\mathbf{u}} \in \mathbb{C}^*$ define a vector subspace of $\mathbb{C}^r$. (It will be shown in the course of the proof of Lemma 1 that this subspace is defined over $\mathbb{Q}$.) So for instance in this context, when the P_i's are constant, the classes are certain cosets of vector subspaces.

With these preliminaries we have the following crucial:

Lemma 1. — *Let $A_1, \ldots, A_h \in \Gamma^r$ and let $P_1, \ldots, P_h \in L[X_1, \ldots, X_r]$. Then $\mathcal{S}$ is a union of finitely many classes.*

[1] In the preprint [EZ] the definition is given differently, in terms of formal power series, applying to fields other than $\mathbb{C}$.

Proof. — We shall argue by double induction, first on r, then on h. We remark that the inductive assumption will hold for all finite extensions L of $\mathbb{C}(z)$; in fact, at some steps in the proof the field L will have to be enlarged.

For $r = 0$ the statement is true (and in fact empty). Let then $r > 0$ be arbitrary and suppose the statement true up to $r - 1$ (and all h). (The arguments now will not substantially differ from the case $r = 1$.)

Let us now argue by induction on h. For $h = 1$ the statement is again true, since there is only one possible class. Let us then suppose $h > 1$ and the statement true up to $h - 1$.

Let $\mathbf{u} \in \mathcal{S}$ and let B be a minimal nonempty subset of $\{1, \ldots, h\}$ such that the $P_i(\mathbf{u})A_i^{\mathbf{u}}$ are, for $i \in B$, linearly dependent over $\mathbb{C}$. By induction on h we may suppose that $B = \{1, \ldots, h\}$. Let $\mathcal{S}_0$ be the subset of such $\mathbf{u}$'s.

Dividing each term $P_i(\mathbf{u})A_i^{\mathbf{u}}$ by $A_h^{\mathbf{u}}$ has the effect to replace A_i with $A_iA_h^{-1}$. Therefore we may suppose $A_h = (1, \ldots, 1)$. If $P_i(\mathbf{u}) = 0$ then $P_i(\mathbf{u})A_i^{\mathbf{u}}$ is linearly dependent, against the present assumption of minimality. Then $P_i(\mathbf{u}) \neq 0$ for $\mathbf{u} \in \mathcal{S}_0$ and in particular we may divide by $P_h(\mathbf{u})$ obtaining a dependence relation among the $R_i(\mathbf{u})A_i^{\mathbf{u}}$, where $R_i(\mathbf{X}) := P_i(\mathbf{X})P_h^{-1}(\mathbf{X}) \in L(\mathbf{X})$.

Observe that $R_h(\mathbf{u})A_h^{\mathbf{u}} = 1$; hence, differentiating the relation and setting

$$A_i = (\alpha_{i1}, \ldots, \alpha_{ir}), \quad \tilde{R}_i(\mathbf{X}) = R_i^{\partial}(\mathbf{X}) + R_i(\mathbf{X})\Big(\sum_{j=1}^{r} X_j \frac{\alpha_{ij}^{\partial}}{\alpha_{ij}}\Big)$$

we obtain that $\partial(R_i(\mathbf{u})A_i^{\mathbf{u}}) = \tilde{R}_i(\mathbf{u})A_i^{\mathbf{u}}$, whence the $\tilde{R}_i(\mathbf{u})A_i^{\mathbf{u}}$, $i = 1, \ldots,$ $h - 1$, are linearly dependent over $\mathbb{C}$, for $\mathbf{u} \in \mathcal{S}_0$.

Note that no proper subset of the $\tilde{R}_i(\mathbf{u})A_i^{\mathbf{u}}$, $i = 1, \ldots, h - 1$, may be dependent, for otherwise a proper subset of the $R_i(\mathbf{u})A_i^{\mathbf{u}}$ together with $1 = R_h(\mathbf{u})A_h^{\mathbf{u}}$ would be dependent, against the assumption that $\mathbf{u}$ lies in the subset $\mathcal{S}_0$ of $\mathcal{S}$.

Note also that, though the R_i and $\tilde{R}_i$ are rational functions, their denominators divide $P_h^2(\mathbf{X})$ and so do not vanish at $\mathbf{u}$. Therefore we may multiply by P_h^2 and reduce to the polynomial case.

By the inductive assumption (on h) we can then include $\mathcal{S}_0$ in a finite number of classes with respect to $A_1, \ldots, A_{h-1}, P_h^2\tilde{R}_1, \ldots, P_h^2\tilde{R}_{h-1}$. Hence in proving the lemma we may restrict to a single class (with respect to the new data).

Let $\mathcal{S}_1$ be the intersection of $\mathcal{S}_0$ with such a class. Then $\mathcal{S}_1$ is a subset of $\mathcal{S}_0$ such that for some $\mathbf{u}_0 \in \mathbb{Q}^r$ we have $(A_iA_j^{-1})^{\mathbf{u}-\mathbf{u}_0} \in \mathbb{C}^*$ for all $\mathbf{u} \in \mathcal{S}_1$, $i, j \in \{1, \ldots, h-1\}$.

Plainly these last conditions hold precisely if $\mathbf{u} - \mathbf{u}_0$ belongs to a certain vector subspace $V \subset \mathbb{C}^r$. We pause to note that this vector space is defined over $\mathbb{Q}$. In fact, for elements $\alpha_1, \ldots, \alpha_r \in \Gamma$, the fact that $\varphi := \prod_{i=1}^r \alpha_i^{u_i} \in \mathbb{C}^*$ amounts to the vanishing of the differential $\sum u_i(d\alpha_i/\alpha_i)$ (which equals $(\sum u_i(\alpha_i^\partial/\alpha_i)dz = (\varphi^\partial/\varphi)dz)$. This is *a priori* a differential over L with only simple poles. Let V' be the vector space of the $\mathbf{u} \in \mathbb{C}^r$ such that all residues of this differential vanish. Then $V' \supset V$ and V' is defined over $\mathbb{Q}$ because the residues of the $d\alpha_i/\alpha_i$ are integers. On the other hand if $\mathbf{u} \in V' \cap \mathbb{Z}^r$ then plainly $\prod_{i=1}^r \alpha_i^{u_i}$ is constant since it lies in L and has no zeros or poles. Then $V' \cap \mathbb{Z}^r$ is contained in V. Since V' has a basis in $\mathbb{Z}^r$ this proves that $V' = V$ and the claim. (This also follows from a general theorem of Ax [A, Thm. 3].)

Let us then write $\mathbf{u} = \mathbf{u}_0 + \mathbf{v}$, where $\mathbf{v} \in V$. Setting $A := A_1$, $B_i = A_iA_1^{-1}$ for $i = 1, \ldots, h-1$ we rewrite $R_i(\mathbf{u})A_i^{\mathbf{u}}$ as $R_i(\mathbf{u}_0 + \mathbf{v})A_i^{\mathbf{u}_0}A^{\mathbf{v}}B_i^{\mathbf{v}}$.[2] Then, since $B_i^{\mathbf{v}} \in \mathbb{C}^*$ we have that 1 and the $R_i(\mathbf{u}_0 + \mathbf{v})A_i^{\mathbf{u}_0}A^{\mathbf{v}}$, $i = 1, \ldots, h-1$, are linearly dependent over $\mathbb{C}$. Also, no proper subset is linearly dependent. It will suffice to show that such set $\mathcal{S}_1$ of vectors $\mathbf{u}$ is contained in a union of finitely many classes.

A first case now occurs when $A^{\mathbf{v}} \in \mathbb{C}^*$ for all $\mathbf{v} \in V$. Then this set $\mathcal{S}_1$ is contained in a single class with respect to the original data $A_1, P_1 \ldots, A_h, P_h$ and we are done. So, we suppose in the sequel that $A^{\mathbf{v}} \notin \mathbb{C}^*$ for at least one $\mathbf{v} \in V$.

We may parametrize linearly V, with integer coefficients, without affecting the problem; in fact recall that V is defined over $\mathbb{Q}$, so we may choose a parametrization so that it has integer coefficients, transforming A into another element in Γ^r. In other words we may assume that $V = \mathbb{C}^s$ for a certain $s \leqslant r$ and that $A = (\alpha_1, \ldots, \alpha_s)$, with $\alpha_i \in \Gamma$; note that we may assume that not all the α_i lie in $\mathbb{C}^*$ because otherwise $A^{\mathbf{v}} \in \mathbb{C}^*$ for all $\mathbf{v} \in V$.

Further, put $T_i(\mathbf{v}) = P_i(\mathbf{u}_0 + \mathbf{v})A_i^{\mathbf{u}_0}$. Observe that since $\mathbf{u}_0 \in \mathbb{Q}^r$ the

[2] This holds in fact only up to nonzero constants, but this does not affect our conclusions.

T_i will have coefficients in a certain finite radical extension of L. We may then enlarge L and suppose that the T_i are actually defined over L. In the new notation, we have, for the $\mathbf{v}$ in question, the linear dependence over $\mathbb{C}$ of $T_1(\mathbf{v})A^{\mathbf{v}},\dots,T_{h-1}(\mathbf{v})A^{\mathbf{v}},P_h(\mathbf{u}_0+\mathbf{v})$.

Recall that such set of $\mathbf{v}$, which we denote $\mathcal{S}_2$ consists of those $\mathbf{v}\in\mathbb{C}^s$ such that the above elements, but no proper subset of them, are linearly dependent over $\mathbb{C}$.

Observe also that if it happens that $s<r$ we can use the inductive assumption with respect to r. This need not be the case however, so we shall lower the dimension by another procedure.

Since not all the α_i lie in $\mathbb{C}$ there exists a pole π for at least one of them. (We view here π as a point of the complex nonsingular curve corresponding to the function field L.) Let $a_i=\operatorname{ord}_\pi\alpha_i$ so $a_1,\dots,a_s$ are integers not all zero. Also, let $\xi\in L^*$ be a uniformizer at π. We may write expansions

$$T_i(\mathbf{Y})=\sum_{j=e}^{\infty}U_{ij}(\mathbf{Y})\xi^j,\quad i=1,\dots,h-1$$

where e is some integer and U_{ij} are polynomials in $\mathbf{Y}=(Y_1,\dots,Y_s)$ with coefficients in $\mathbb{C}$.

For a purpose which will be soon clear, we now introduce certain algebraic subvarieties of $\mathbb{C}^s$. Precisely, for any integer $l\geqslant e$ we let W_l be the algebraic variety consisting of the $\mathbf{v}\in\mathbb{C}^s$ such that the rank of the matrix $(U_{ij}(\mathbf{v}))$, $i=1,\dots,h-1$, $j=e,\dots,l$, is less than $h-1$. For $l<e$ we agree that $W_l=\mathbb{C}^s$. We have that these varieties form a descending chain, so the chain stabilizes and we shall have $W=W_l$ for all l larger than a certain l_0, where W is a certain algebraic subvariety of $\mathbb{C}^s$.

Let $\mathbf{v}\in W$. Then there exist $c_1,\dots,c_{h-1}\in\mathbb{C}$, not all zero, such that $\sum_{i=1}^{h-1}c_iU_{ij}(\mathbf{v})=0$ for all $j\geqslant e$. But then the above expansions show that $\sum_{i=1}^{h-1}c_iT_i(\mathbf{v})=0$, so that the $T_i(\mathbf{v})$, and also the $T_i(\mathbf{v})A^{\mathbf{v}}$, are linearly dependent over $\mathbb{C}$. This shows that $W\cap\mathcal{S}_2$ is empty. In particular, for $\mathbf{v}\in\mathcal{S}_2$ the order at π of a nontrivial linear combination $\sum_{i=1}^{h-1}c_iT_i(\mathbf{v})$ must be at most l_0 and has therefore finitely many possibilities, independently of $\mathbf{v}\in\mathcal{S}_2$ and of the complex coefficients c_i, not all zero.

A similar and simpler argument shows that the order at π of $P_h(\mathbf{u}_0+\mathbf{v})$ has finitely many possibilities independently of $\mathbf{v}\in\mathcal{S}_2$.

Let now $\mathbf{v}\in\mathcal{S}_2$ and write a nontrivial vanishing linear combination in the form

$$A^{\mathbf{v}}\Big(\sum_{i=1}^{h-1} c_i T_i(\mathbf{v})\Big) = cP_h(\mathbf{u}_0 + \mathbf{v}).$$

Taking the logarithmic derivative of both sides, multiplying by dz and taking residues at π of the resulting differentials of L/k we obtain a relation

$$\sum_{i=1}^{s} v_i a_i = \operatorname{ord}_\pi(P_h(\mathbf{u}_0 + \mathbf{v})) - \operatorname{ord}_\pi\Big(\sum_{i=1}^{h-1} c_i T_i(\mathbf{v})\Big).$$

We have just shown that the right side has a finite number of possibilities, independently of the c_i's or of $\mathbf{v} \in \mathcal{S}_2$. Therefore we may partition $\mathcal{S}_2$ in a finite number of subsets according to the value of the right side. Each of these subsets will be defined in $\mathcal{S}_2$ by a linear equation

$$\sum_{i=1}^{s} v_i a_i = a,$$

where a is a fixed integer (depending on the subset); recall that the a_i are integers not all zero, so this equation defines a translate of a proper vector subspace of $\mathbb{C}^s$, defined over $\mathbb{Q}$. We may now write parametrizations

$$v_i = b_i + \sum_{j=1}^{s-1} b_{ij} t_j, \quad i = 1, \dots, s,$$

with integer coefficients b_{ij} and rationals b_i. Substituting for $\mathbf{v}$ in the terms $P_h(\mathbf{u}_0 + \mathbf{v})$ and $T_i(\mathbf{v})A^{\mathbf{v}}$ we see that we may view $\mathcal{S}_2$ as embedded in a space $\mathbb{C}^{s-1}$ of strictly lower dimension than $\mathbb{C}^r$. (Note that if the b_i do not all lie in $\mathbb{Z}$ but just in $\mathbb{Q}$, we may have again to enlarge the field L and go to a finite radical extension.) By the inductive assumption on the dimension we may infer that $\mathcal{S}_2$ is contained in a union of finitely many classes, relative to the "new context"; however it is immediate to realize that the each "new" class corresponds to an "old" one: in fact, we may write $A^{\mathbf{v}} = A^{\mathbf{b}}B^{\mathbf{t}}$ for $\mathbf{b} = (b_1, \dots, b_s) \in \mathbb{Q}^s$, for a suitable $B \in (L^*)^{s-1}$ and for $\mathbf{t} = (t_1, \dots, t_{s-1}) \in \mathbb{C}^{s-1}$; then, for $\mathbf{t}$ in a "new" class we have, for a suitable $\mathbf{t}_0 \in \mathbb{Q}^{s-1}$ depending only on the class, $B^{\mathbf{t}-\mathbf{t}_0} \in \mathbb{C}^*$. Defining now $\mathbf{v}_0$ with the above equations, with $\mathbf{t}_0$ in place of $\mathbf{t}$, we have $\mathbf{v}_0 \in \mathbb{Q}^s$ and $A^{\mathbf{v}-\mathbf{v}_0} = (A^{\mathbf{b}}B^{\mathbf{t}})(A^{\mathbf{b}}B^{\mathbf{t}_0})^{-1} = B^{\mathbf{t}-\mathbf{t}_0} \in \mathbb{C}^*$, concluding the proof.

Remark 1. — This lemma is purely qualitative; however, in the present approach, it represents a necessary tool for a quantitative version of itself, which we state as the next result. The arguments for Lemma 1 can be quantified, but this leads to estimates inferior to the sought ones.

Proposition 1. — *Let $A_1,\ldots,A_h \in \Gamma^r$ and let $P_1,\ldots,P_h \in L[X_1,\ldots,X_r]$ have degrees at most $d_1,\ldots,d_h$ respectively. Then $\mathcal{S}$ may be expressed as a union of no more than $\left(d_1+\cdots+d_h+\binom{h}{2}\right)^r$ classes.*

Proof. — We use at once Lemma 1 and express the set $\mathcal{S}$ as a union of finitely many classes $\mathcal{S}_1,\ldots,\mathcal{S}_l$. We start by proving that each $\mathcal{S}_k$ is a closed algebraic subvariety of $\mathbb{C}^r$.

In fact such a class $\mathcal{S}_k$, say relative to a subset $B = B_k \subset \{1,\ldots,h\}$, is defined by two conditions: the first condition is that for $\mathbf{u} \in \mathcal{S}_k$ the $P_\imath(\mathbf{u})A_\imath^{\mathbf{u}}$, $i \in B$, are linearly dependent over $\mathbb{C}$; the second condition is that for $\mathbf{u} \in \mathcal{S}_k$ we have $(A_\imath A_\jmath^{-1})^{\mathbf{u}-\mathbf{u}_0} \in \mathbb{C}^*$ for all $i,j \in B$. Now, by the second condition the first one is equivalent to the fact that the $P_i(\mathbf{u})A_\imath^{\mathbf{u}_0}$, $i \in B$, are linearly dependent. By the Wronskian criterion this amounts to the vanishing of the Wronskian determinant of the elements in question. Plainly this determinant is the value at $\mathbf{X} = \mathbf{u}$ of a certain polynomial in $L[\mathbf{X}]$, and this gives a first algebraic condition on $\mathbf{u}$.[3]

As to the second condition, we have already remarked that it defines the translation by $\mathbf{u}_0$ of a vector subspace of $\mathbb{C}^r$. (This vector space is actually defined over $\mathbb{Q}$, as we have shown in the course of the proof of Lemma 1; this is however immaterial now.) In conclusion, $\mathcal{S}_\jmath$ is the intersection of the above pair of algebraic subsets, proving the claim.

Next, we show that the whole $\mathcal{S}$ is a closed algebraic subset of $\mathbb{C}^r$ defined by equations over $\mathbb{C}$, each of degree $\leqslant d_1+\cdots+d_h+\binom{h}{2}$.

For a natural number ℓ define inductively polynomials $P_{i\ell}$ as follows. Set $P_{\imath 0}(\mathbf{X}) = P_\imath(\mathbf{X})$ and, for $\ell \geqslant 0$ put

$$P_{\imath,\ell+1}(\mathbf{X}) = P_{i\ell}^{\partial}(\mathbf{X}) + P_{i\ell}(\mathbf{X})\left(\sum_{j=1}^{r} X_j \frac{\alpha_{ji}^{\partial}}{\alpha_{j\imath}}\right).$$

The definition is given so that for $\mathbf{u} \in \mathbb{C}^r$ the formula $P_{i,\ell+1}(\mathbf{u})A_i^{\mathbf{u}} = (P_{\imath\ell}(\mathbf{u})A_i^{\mathbf{u}})^{\partial}$ holds.

Consider the matrix $(P_{i\ell}(\mathbf{X}))$, for $i = 1,\ldots,h$, $\ell = 0,\ldots,h-1$, and its determinant $\Delta \in L[\mathbf{X}]$. Then it is immediately checked that $\Delta(\mathbf{u})$ equals, up to $\prod_{i=1}^{h} A_\imath^{\mathbf{u}} \neq 0$, the Wronskian determinant of the $P_\imath(\mathbf{u})A_i^{\mathbf{u}}$.

[3] Note that this condition is over L and corresponds to several conditions defined over $\mathbb{C}$.

We write, as we may, $\Delta(\mathbf{X}) = \gamma_1\Delta_1(\mathbf{X}) + \cdots + \gamma_q\Delta_q(\mathbf{X})$, where $\gamma_i \in L^*$ are linearly independent over $\mathbb{C}$ and where $\Delta_j(\mathbf{X}) \in \mathbb{C}[\mathbf{X}]$ for $j = 1, \ldots, q$ (possibly $q = 0$).

Plainly, $\deg \Delta \leqslant d_1 + \cdots + d_h + \binom{h}{2}$, so the same estimate holds for the degrees of the Δ_j.

Observe now that, by the Wronskian criterion again, for a $\mathbf{u} \in \mathbb{C}^r$, we have $\mathbf{u} \in \mathcal{S}$ if and only if $\Delta(\mathbf{u}) = 0$; in turn this amounts to the equations $\Delta_j(\mathbf{u}) = 0$ for $j = 1, \ldots, q$. This proves our contention about the variety $\mathcal{S}$.

Further, each irreducible component of $\mathcal{S}$ is contained in some class $\mathcal{S}_j$, because the $\mathcal{S}_j$ are (finitely many!) algebraic varieties (not necessarily irreducible) in $\mathbb{C}^r$ whose union contain $\mathcal{S}$. For each component of $\mathcal{S}$, let us pick one class containing it; at the end of the process let us omit the classes, if any, which we have not met so far; the union of the chosen classes will contain the union of the irreducible components of $\mathcal{S}$, hence it will continue to contain the whole $\mathcal{S}$. Also, the number of such classes will not exceed the number of components of $\mathcal{S}$.

To conclude the argument, it will then suffice to estimate suitably the number of components of $\mathcal{S}$. We shall prove first the following claim from elementary algebraic geometry:

There exists an algebraic variety W in $\mathbb{C}^r$ defined by at most r equations of degree $\leqslant d_1 + \cdots + d_h + \binom{h}{2}$ such that each component of $\mathcal{S}$ is a component of W.

The argument below is certainly well known but missing a reference we describe it.

We prove by induction on $p = 0, 1, \ldots, r$ that there exists a variety W_p containing $\mathcal{S}$ and defined by at most p equations of degree $\leqslant d_1 + \cdots + d_h + \binom{h}{2}$ such that each component of $\mathcal{S}$ of dimension s is contained in a component of W_p of dimension $\leqslant \max(s, r - p)$.

For $p = 0$ just take $W_p = \mathbb{C}^r$. Assume now the existence of W_p for a positive $p < r$; we shall construct a suitable W_{p+1}, concluding the induction. Let V be a component of $\mathcal{S}$. By induction, V will be contained in a component V' of W_p of dimension $\leqslant \max(\dim V, r - p)$. If $V' \neq V$ (hence $\dim V' > \dim V$ and so $\dim V' \leqslant r - p$) there exists some polynomial Δ_j, as constructed above, such that Δ_j vanishes on V but not on the whole V'. Let Δ^V denote such a Δ_j. Observe that Δ^V has degree $\leqslant d_1 + \cdots + d_h + \binom{h}{2}$. Choose such a polynomial for each V and form a linear combination $\Lambda := \sum_V c_V \Delta^V$, with coefficients $c_V \in \mathbb{C}$. Plainly, since there are only finitely many components V, for a "general" choice of the c_V, the polynomial Λ will not vanish identically, on any V' such that $V' \neq V$. For such V', the polynomial Λ will define in V' a proper subvariety, hence of dimension $\leqslant \dim V' - 1 \leqslant r - p - 1$, still containing V. Defining then W_{p+1} as the variety determined by Λ in W_p concludes the induction step.

Putting now $W = W_r$ gives immediately the above claim.

Finally, by the generalized Bezout Theorem (see e.g. [D, Ch. III, 2.1]) we have that the sum of the degrees of all the components of such a variety W does not exceed $(d_1 + \cdots + d_h + \binom{h}{2})^r$; Since each component of $\mathcal{S}$ is a component of W, the same bound follows for the number of components of $\mathcal{S}$ and, as remarked above, for the number of relevant classes; this completes the proof.

Remark 2.

(i) In the present paper we are interested in the integer points $\mathbf{u}$, rather than the complex points appearing in the statements proved so far. By the way, the complex points complicate the whole thing, since the meaning of expressions like α^u for complex u is not always well defined, and in fact we had to introduce some preliminaries for that reason. However, we needed to use the whole complex points in question, in order to better investigate the structure of the variety $\mathcal{S}$; of course, if we had Lemma 1 just for the integer points, no quantitative conclusion could be reached since the integer points in the classes of $\mathcal{S}$ could be *a priori* not Zariski dense on any component of the variety determined by the Wronskian. (The only exception occurs in the one-variable case.)

(ii) Note also that in the case when the $P_\imath$'s are constant, the classes are cosets of vector spaces, so Lemma 1 implies that all the components of the variety $\mathcal{S}$ are linear (and over $\mathbb{Q}$).

3. Proofs of main results.

We start by proving Theorem 1. For this we want to suitably embed the relevant field and elements in a finite extension of $\mathbb{C}(z)$, so to apply Proposition 1. (See also [Ch, Ch. V] for a detailed theory of extension of the field of constants.)

The field k is already supposed to be embedded in $\mathbb{C}$. Since L has transcendence degree 1 over k, we may view L as a finite extension of $k(t)$, where t is transcendental over k. Let L' be the quotient field of the ring $L \otimes_k \mathbb{C}$; it is easily verified that in fact this ring is a domain, since k is algebraically closed. Then k, L are embedded in L'; let ϕ denote this embedding. Since t is transcendental over k, $z := \phi(t)$ is transcendental over $\mathbb{C}$ and L' is a finite extension of $\mathbb{C}(z)$. For A_i, P_i as in the theorem, define now $A_i' := \phi(A_i)$, $P_i' := \phi(P_\imath)$. We now apply Proposition 1 to L' and the A_i', P_i'. We let $\mathcal{S}$ be as in Proposition 1 and express $\mathcal{S}$ as the union of classes $\mathcal{S}_1, \dots, \mathcal{S}_N$, where $N \leqslant \left(d_1 + \cdots + d_h + \binom{h}{2}\right)^r$.

Let now S be the set of Theorem 1 and let $\mathbf{m} \in S$, so the $P_i(\mathbf{m})A_i^{\mathbf{m}}$ are linearly dependent over k. Then the $P_i'(\mathbf{m}){A_i'}^{\mathbf{m}}$ are linearly dependent over $\mathbb{C}$. In particular, $\mathbf{m}$ lies in $\mathcal{S}$ and hence in one of the classes, say it lies in $\mathcal{S}_l$. If $\mathcal{S}_l$ is relative to a set $B_l \subset \{1, \dots, h\}$ this means that the $P_i'(\mathbf{m}){A_i'}^{\mathbf{m}}$, $i \in B_j$, are linearly dependent over $\mathbb{C}$ and that, if $\mathbf{m}_0 \in \mathbb{Z}^r$ is any element of $\mathcal{S}_l \cap \mathbb{Z}^r$ we have $(A_i'/A_j')^{\mathbf{m}-\mathbf{m}_0} \in \mathbb{C}$ for all $i, j \in B_l$.

Now, if $x_1, \dots, x_h \in L$ are such that $\phi(x_1), \dots, \phi(x_h)$ are linearly dependent over $\mathbb{C}$, then $x_1, \dots, x_h$ must in fact be linearly dependent over k. This follows immediately by the fact that a basis for the tensor product of vector spaces over k consists of the products of basis elements for the vector spaces.

In particular, if $x \in L$ is such that $\phi(x) \in \mathbb{C}$, then $x \in k$; namely, $\phi^{-1}\mathbb{C} = k$.

Therefore we deduce that the $P_i(\mathbf{m})A_i^{\mathbf{m}}$, $i \in B_l$, are linearly dependent over k and that $(A_i/A_j)^{\mathbf{m}-\mathbf{m}_0} \in k$ for $i, j \in B_l$.

This means that $\mathbf{m}, \mathbf{m}_0$ lie in the same class relative to k, L. In other words, the set S is partioned into classes in the same way the set $\mathcal{S}$ is partitioned into the classes $\mathcal{S}_l$. This completes the proof of Theorem 1.

Remark 3. — If one develops the arguments of §2 for arbitrary function fields L, namely for finite extensions of $\mathbb{C}(z_1, \dots, z_q)$, Theorem 1 follows for all finitely generated extensions L/k. This program needs few changes with respect to the present treatment. One has however to use the Wronskian criterion for several variables; this can be found e.g. in [C, p. 112] for the case of rational functions, but the extension to algebraic functions needs no substantial change.

Proof of Corollary 1. — We start with part (a). Let $S_1 \subset \mathbb{Z}^r$ be the set of solutions as in the statement. Then S_1 is contained in the set S consisting of the $\mathbf{m} \in \mathbb{Z}^r$ such that the $P_i(\mathbf{m})A_i^{\mathbf{m}}$ are linearly dependent over k. By Theorem 1 we obtain that S, and hence S_1, may be included in the union of at most $\left(d_1 + \cdots + d_h + \binom{h}{2}\right)^r$ classes. Suppose first that $\mathbf{m} \in S_1$ belongs to a class corresponding to the subset $B = \{i\}$; then $P_i(\mathbf{m}) = 0$, a contradiction. If $\#B \geqslant 2$ and the class contains two distinct elements $\mathbf{m}_0, \mathbf{m}_1$ of S, then we have, for i, j distinct elements of B, $(A_i A_j^{-1})^{\mathbf{m}_1 - \mathbf{m}_0} \in k^*$. However we are assuming that the coordinates of $A_i A_j^{-1}$ are multiplicatively independent modulo k^* for $i \neq j$; therefore we have a contradiction. This shows that the class contains at most one element. We deduce that S_1 has no more elements than there are classes and the sought estimate follows.

Part (b) is obtained in the same way, but by applying Theorem 1 to the $P_i(\mathbf{m})A_i^m$, $i = 1, \ldots, h$, together with another pair $P_{h+1} = 1$, $A_{h+1} = (1, \ldots, 1)$.

As to part (c), write $P_i(\mathbf{X}) = \sum_{l=1}^{u_\imath} \pi_{il} Q_{\imath l}(\mathbf{X})$, where, for each i, the $\pi_{il} \in L^*$, $l = 1, \ldots, u_\imath$ are linearly independent over k and the $Q_{\imath\jmath}(\mathbf{X})$ lie in $k[\mathbf{X}]$. Since P_i has $\leqslant t_i$ terms, we may take $u_\imath \leqslant t_i$. Now, if $\mathbf{m}$ is a solution to the equation and not all the $P_i(\mathbf{m})$ vanish, then not all the $Q_{\imath\jmath}(\mathbf{m})$ vanish. Then the elements $\pi_{il} A_i^{\mathbf{m}}$ $(i = 1, \ldots, h,\ l = 1, \ldots, u_\imath)$ are linearly dependent over k. We then apply Theorem 1, with $\sum u_\imath$ in place of h, with the same A_i's, except that now $A_\imath$ is counted $u_\imath$ times, and with the $\pi_{\imath\jmath}$ in place of the $P_\imath$, in the appropriate order. Let Ω be a class, relative to the set B of indices (taken now among the pairs (i, l), $i = 1, \ldots, h$, $l = 1, \ldots, u_\imath$). Since the $A_i A_j^{-1}$ have multiplicatively independent mod k^* coordinates for $i \neq j$, we see that the class cannot contain two distinct elements, unless all the indices (i, l) in the set B all have the same "i". But then the elements π_{il}, $l = 1, \ldots, u_i$, cannot be dependent over k, a contradiction. This proves that each class has at most one element, and the required estimate again follows.

Proof of Corollary 2. — Let us assume that there are no integers u_0, v_0, r, s with $rs \neq 0$ such that $G(u_0 + mr) = H(v_0 + ms)$ for all $m \in \mathbb{Z}$ (in the proof of part (a)) and no integers u_0, v_0, r, s with $rs \neq 0$ and $\xi, \eta \in k^*$ such that $G(u_0 + mr) = \eta \xi^m H(v_0 + ms)$ for all $m \in \mathbb{Z}$ (in the proof of part (b)). On these assumptions we shall prove the stated bounds for the number of relevant solutions.

We define vectors $A_i \in (L^*)^2$, for $i = 1, \ldots, p + q$ by setting

$A_i = (\alpha_i, 1)$ for $1 \leqslant i \leqslant p$ and $A_i = (1, \beta_{i-p})$ for $p+1 \leqslant i \leqslant p+q$. Similarly, we define constants $P_i \in L^*$ by $P_i = a_i$ for $1 \leqslant i \leqslant p$ and $P_i = b_{i-p}$ for $p+1 \leqslant i \leqslant p+q$.

A solution $\mathbf{m} = (u,v) \in \mathbb{Z}^2$ of $G(u) = H(v)$ (or of $G(u) = cH(v)$) gives an element of the set S made up of the integral vectors $\mathbf{m} \in \mathbb{Z}^2$ such that the $P_i A_i^{\mathbf{m}}$, $i = 1, \dots, p+q$, are linearly dependent over k. The set S is, by Theorem 1, contained in the union of no more than $\binom{p+q}{2}^2$ classes as in Definition 2. This gives a corresponding partition of the set of solutions (in both cases (a) and (b) of Corollary 2).

Let Ω be a class, corresponding to the subset $B = B_\Omega \subset \{1, \dots, p+q\}$ and let us estimate the number M_Ω of solutions belonging to that class.

The set B cannot contain a single element because no term vanishes. Suppose first that B contains two distinct integers i, j in $[1, p]$ and let $\mathbf{m}_0 = (u_0, v_0), \mathbf{m} = (u,v) \in \Omega$. Then $(A_i A_j^{-1})^{\mathbf{m}-\mathbf{m}_0} = (\alpha_i/\alpha_j)^{u-u_0} \in k^*$. Since $\alpha_i/\alpha_j \notin k$, we see that $u = u_0$. Therefore for the $(u,v) \in \Omega$ we have $G(u_0) = H(v)$ (resp. $G(u_0) = cH(v)$); if $G(u_0) = 0$ we have that $H(v) \in k$ while if $G(u_0) \neq 0$ we have that $G^{-1}(u_0)H(v)$ lies in k. In both cases, by Corollary 1(b) there are at most $\binom{q+1}{2} \leqslant \binom{p+q}{2}$ such integers v. The same argument works if B contains two distinct integers in $[p+1, p+q]$, proving that in these cases $M_\Omega \leqslant \binom{p+q}{2}$

Suppose now that B consists precisely of two elements $i_0, j_0 + p$ with $1 \leqslant i_0 \leqslant p$, $1 \leqslant j_0 \leqslant q$.

Setting $\mathbf{m} = (u,v)$, $\mathbf{m}_0 = (u_0, v_0)$ we have $\alpha_{i_0}^{u-u_0} \beta_{j_0}^{v_0-v} \in k^*$. This holds precisely if $(u,v) - (u_0,v_0)$ runs over a certain subgroup of $\mathbb{Z}^2$. This subgroup cannot be trivial if we have at least two distinct solutions in Ω, as we assume, and it cannot have rank 2 because otherwise $\alpha_{i_0}, \beta_{j_0}$ would lie in k^*. Therefore the subgroup has rank 1 and is generated by a vector $(r,s) \in \mathbb{Z}^2$, $rs \neq 0$.

Then the class Ω consists of the vectors $(u,v) = (u_0 + mr, v_0 + ms)$, $m \in \mathbb{Z}$. Our solutions in the class Ω then correspond to integers m such that $G(u_0 + mr) = H(v_0 + ms)$ (resp. $G(u_0 + mr) = cH(v_0 + ms)$). This again reduces our problem to the one-variable case.

We could now appeal to the results in [Schm, §10]; however, for completeness we reprove what we need as a simple consequence of Theorem 1, stating it as a lemma.

LEMMA 2. — *Let $\gamma_1, \dots, \gamma_l \in L^*$ be such that $\gamma_i/\gamma_j \notin k^*$ for $i \neq j$.*

For $i = 1, \dots, l$, let $c_{ij} \in L^*$, $j = 1, \dots, j_i$. Suppose that for each i, $c_{i1}, \dots, c_{ij_i}$ are linearly independent over k Then there are at most $\binom{\sum j_i}{2}$ integers m such that the $c_{ij}\gamma_i^m$ are linearly dependent over k.

The proof is very similar to that for Cor. 1 (c); by Theorem 1 (with $r = 1$, $h = \sum_{i=1}^{l} j_i$, with the P_i's equal to the c_{ij} and the A_i's equal to the γ_i's counted j_i times) such integers m can be grouped into at most $\binom{h}{2}$ classes. If a class which contains two distinct elements m_0, m_1 corresponds to the set B we have in particular that $(A_i/A_j)^{m_1 - m_0} \in k^*$ for $i, j \in B$. Then A_i and A_j cannot be distinct, because two distinct ones among the γ_i's cannot have their ratio in k^*. Then B must be such that $A_i = A_j$ for $i, j \in B$ and therefore for some ℓ we would have $A_i = \gamma_\ell$ for all $i \in B$. But then the elements $c_{\ell,1}, \dots, c_{\ell,j_\ell}$ would be linearly dependent over k, a contradiction.

Therefore each class can contain at most one relevant integer, whence the result.

Let us now go back to the proof of Corollary 2 and consider the solutions (u, v) in the class Ω. As observed above, we may express such solutions as $(u_0 + mr, v_0 + ms)$ with u_0, v_0, r, s fixed, and $m \in \mathbb{Z}$. Thus we have to consider the set of integers m for which there is $c \in k^*$ such that

$$G(u_0 + mr) - cH(v_0 + ms) = 0 \tag{1}$$

where in part (a) of Corollary 2 we assume $c = 1$. Writing as above $G(u) = \sum_{\mu=1}^{p} a_\mu \alpha_\mu^u$, $H(v) = \sum_{\nu=1}^{q} b_\nu \beta_\nu^v$, (1) can be rewritten as

$$\sum_{\mu=1}^{p} a'_\mu (\alpha_\mu^r)^m - c \sum_{\nu=1}^{q} b'_\nu (\beta_\nu^s)^m = 0,$$

where $a'_\mu = a_\mu \alpha_\mu^{u_0}$, $b'_\nu = b_\nu \beta_\nu^{v_0}$. We partition the set of α_μ^r ($\mu = 1, \dots, p$) and β_ν^s ($\nu = 1, \dots, q$) into groups such that two elements from this set belong to the same group if and only if their quotient is in k^*. By our assumption on the α_μ, β_ν, a group consists either of one element, or of one α_μ^r and one β_ν^s. This implies that (after reindexing) (1) can be rewritten as

$$\sum_{\mu=1}^{\ell} (a'_\mu - c\delta_\mu^m b'_\mu)(\alpha_\mu^r)^m - \sum_{\mu=\ell+1}^{p} a'_\mu (\alpha_\mu^r)^m - \sum_{\nu=\ell+1}^{q} cb'_\nu (\beta_\nu^s)^m = 0,$$

where $0 \leqslant \ell \leqslant \min(p, q)$, the δ_μ are in k^* and where no two distinct elements of $\{\alpha_\mu^r \ (\mu = 1, \dots, p), \beta_\nu^s (\nu = \ell + 1, \dots, q)\}$ are in k^*.

If $p > \ell$ or $q > \ell$ then by Lemma 2 there are at most $\binom{p+q}{2}$ integers m such that (1) with some $c \in k^*$ holds. Assume henceforth that $\ell = p = q$.

If at least one pair a'_μ, b'_μ is linearly independent over k then again by Lemma 2 there are at most $\binom{p+q}{2}$ integers m for which there is $c \in k^*$ such that (1) holds. Now, suppose that all pairs a'_μ, b'_μ ($\mu = 1, \dots, \ell$) are linearly dependent over k. By Lemma 2 (but now with $\sum j_i \leqslant \ell$) there are at most $\binom{\ell}{2}$ integers m for which there is $c \in k^*$ such that (1) holds, and at least one of the coefficients $a'_\mu - c\delta_\mu^m b'_\mu$ is non-zero. We show below that there is at most one m such that these coefficients are all 0. Thus there are at most $1 + \binom{\ell}{2} \leqslant \binom{p+q}{2}$ possibilities for m. Then it follows that each class Ω has at most $\binom{p+q}{2}$ solutions, and so in view of the upper bound obtained above for the number of classes, our original equation $G(u) = H(v)$ (in case (a)) or $G(u) = cH(v)$ with $c = c(u,v) \in k^*$ (in case (b)) in $u, v \in \mathbb{Z}$ has at most $\binom{p+q}{2}^3$ solutions.

Now suppose that there are at least two distinct m's for which there exists $c \in k^*$ such that all coefficients $a'_\mu - c\delta_\mu^m b'_\mu$ ($\mu = 1, \dots, \ell$) are 0. We have to distinguish between the cases (a) and (b) of Corollary 2. First consider the most difficult case (b). Then there are integers $m_1 < m_2$ such that there are $c_1, c_2 \in k^*$ with $a'_\mu = c_j \delta_\mu^{m_j} b'_\mu$ for $\mu = 1, \dots, \ell$, $j = 1, 2$. It follows that $\delta_\mu^{m_2 - m_1} = c_1/c_2$ for $\mu = 1, \dots, \ell$. Taking $\eta = c_1$, $\xi = c_2/c_1$, it follows that for every $t \in \mathbb{Z}$ we have $a'_\mu = \eta\xi^t \delta_\mu^{m_1 + t(m_2 - m_1)} b'_\mu$ for $\mu = 1, \dots, \ell$. But then tracing back it follows that $G(u_0 + (m_1 + t(m_2 - m_1))r) = \eta\xi^t H(v_0 + (m_1 + t(m_2 - m_1))s)$ for all $t \in \mathbb{Z}$, which was excluded by the assumption made in the beginning of our proof of Corollary 2.

Now consider case (a); thus we consider those m such that (1) holds with $c = 1$. We can repeat the argument from above, and arrive at the same conclusion, but with $\xi = \eta = 1$. Thus, $G(u_0 + (m_1 + t(m_2 - m_1))r) = H(v_0 + (m_1 + t(m_2 - m_1))s)$ for all $t \in \mathbb{Z}$, which was again excluded by the assumption at the beginning of our proof.

Proof of Corollary 3. — The polynomials $G_n(P(X))$ will of course satisfy a recurrence obtained by substituting $P(X)$ in place of X in the recurrence for $G_n(X)$. Accordingly, we shall have formulas

$$(2) \qquad G_n(X) = a_1\alpha_1^n + \cdots + a_d\alpha_d^n, \quad G_n(P(X)) = b_1\beta_1^n + \cdots + b_d\beta_d^n.$$

For a nonconstant polynomial $R(X) \in k[X]$ we let $L_{R(X)}$ be the splitting field of $Q(R(X), T)$ over $k(R(X))$; these fields are all isomorphic over k. Then the $a_i, \alpha_i \in L_X$, while the $b_\imath, \beta_i$ lie in $L_{P(X)}$. (The notation $\alpha_j(P(X))$ for the β_j's is is not well-defined *a priori*; this can be done only locally. This is why we shall work with complete sets of roots rather than single elements.)

Now, consider the equation $G_u(X) = cG_v(P(X))$, $u, v \in \mathbb{Z}$, for $c = c(u,v) \in k^*$. In view of the present assumptions we can apply Corollary 2(b) to $G(u) := G_u(X)$, $H(v) := G_v(P(X))$. Now we have $p = q = d$ and we obtain that there are at most $\binom{2d}{2}^3$ solutions (u,v) unless for some $\eta, \xi \in k^*$ and integers u_0, v_0, r, s with $rs \neq 0$, we have

$$G_{u_0+rm}(X) = \eta\xi^m G_{v_0+sm}(P(X)) \tag{3}$$

identically in $m \in \mathbb{Z}$. Plainly if equation (3) holds we have infinitely many solutions; in particular, this proves the first part of Corollary 3. In the sequel we shall assume equation (3).

We shall use the notation $\alpha \sim \beta$ if the ratio $\alpha/\beta \in k^*$. We extend this to finite sets on putting $\{\alpha_1, \dots, \alpha_d\} \sim \{\beta_1, \dots, \beta_d\}$ if there exists a permutation σ of the indices so that $\alpha_i \sim \beta_{\sigma(i)}$ for $i = 1, \dots, d$.

Further, we shall use the notation $R_\ell(X)$ (for a rational function R) to indicate its ℓ-th iterate under composition $R \circ \cdots \circ R$.

Let $\alpha_{i,\ell}$ denote the roots of $Q(P_\ell(X), T)$, so $L_{P_\ell}(X) = k(P_\ell(X), \alpha_{1,\ell}, \dots, \alpha_{d,\ell})$. Take an integer $\ell \geqslant 1$ and substitute in (3) $X \mapsto P_{\ell-1}(X)$; we have in particular (as remarked after Corollary 2)

$$\{\alpha^s_{i,\ell} : i = 1, \dots, d\} \sim \{\alpha^r_{i,\ell-1} : i = 1, \dots, d\}, \tag{4}$$

whence by induction we get

$$\{\alpha^{s^\ell}_{i,\ell} : i = 1, \dots, d\} \sim \{\alpha_i^{r^\ell} : i = 1, \dots, d\}.$$

Now, for a fixed i, let $\alpha_i^{r^\ell} \sim \alpha^{s^\ell}_{j,\ell}$. The degree $[L_X(\alpha_{j,\ell}) : k(\alpha^{s^\ell}_{j,\ell})]$ is divisible by $[k(\alpha_{j,\ell}) : k(\alpha^{s^\ell}_{j,\ell})]$ and hence by s^ℓ. Therefore s^ℓ divides $[L_X : k(\alpha^{s^\ell}_{j,\ell})][L_X(\alpha_{j,\ell}) : L_X]$. The factor on the right is bounded by $[k(X, \alpha_{j,\ell}) : k(X)] \leqslant d$, while the factor on the left equals $[L_X : k(\alpha_i^{r^\ell})] = [L_X : k(\alpha_i)]r^\ell$. We deduce that s^ℓ divides r^ℓ times a non-zero factor independent of ℓ. Letting ℓ grow, we deduce that s divides r; let us put $r = qs$; then (4) for $\ell = 1$ entails

$$\{\beta_i : i = 1, \dots, d\} \sim \{\alpha_i^q : i = 1, \dots, d\}. \tag{5}$$

In particular $L_{P(X)}$ (denoted L_P from now on) is contained in L_X. Let us show that L_X/L_P is somewhere ramified. Consider the point at infinity of $k(P(X))$, denoted ∞_P. This point lifts to the point ∞ of $k(X)$ and is totally ramified below it, with index $[k(X) : k(P)] = \deg P = p$. Let now M be the maximum ramification index above ∞_P in L_P. Then M is as well the maximum ramification index above ∞ in L_X. Suppose that L_X/L_P were unramified everywhere. Then the maximum ramification index above ∞_P in L_X would continue to be M. On the other hand this ramification index is at least $M[k(X):k(P)] = Mp$. Since we are assuming p $(= \deg P) > 1$ we have a contradiction.

Now, the fields L_X and L_P are isomorphic over k, hence have the same genus g. Since the extension L_X/L_P is somewhere ramified, and thus has a degree > 1, the Hurwitz genus formula yields $2g-2 > [L_X : L_P](2g-2)$; we deduce that the genus is zero.$^{(4)}$

Then $L_X = k(t)$ for some $t \in L_X$. There exists a field-isomorphism $\varphi : L_X \to L_P$ over k such that $X^\varphi = P(X)$ and $\alpha_i^\varphi = \beta_{\tau(i)}$ for some permutation τ. Let $\pi = t^\varphi$. Since $L_P \subset L_X = k(t)$ we have $\pi = \pi(t) \in k(t)$ and $L_P = k(\pi(t))$. Also, if $X = B(t)$ then $P(B(t)) = P(X) = X^\varphi = B(t^\varphi) = B(\pi(t))$, so $P \circ B = B \circ \pi$; in particular, we have $p = \deg P = \deg \pi$.

Further, $\alpha_i = A_i(t)$ and $\beta_{\tau(i)} = \alpha_i^\varphi = A_i(\pi(t))$. Moreover, the β_j are up to constants a permutation of the α_i^q, whence for some permutation σ of $\{1,\dots,d\}$ we have

$$A_i(\pi(t)) \sim A^q_{\sigma(i)}(t).$$

Let us now iterate this equation obtaining $A_i(\pi_\ell(t)) \sim A^{q^\ell}_{\sigma^\ell(i)}(t)$ for each positive integer ℓ. We may take ℓ to be a multiple of the order of σ. Also, these equations show that π permutes the set $\mathcal{Z}$ of zeros/poles of all the functions A_i and also that $\pi^{-1}(\mathcal{Z}) \subset \mathcal{Z}$. Hence we may take ℓ to be a multiple of the order of the corresponding permutation, so π_ℓ fixes each element in such set. We obtain in particular that

$$A_i(\pi_\ell(t)) \sim A_i^{q^\ell}(t), \quad i = 1,\dots,d$$

and then, comparing degrees, $p = |q|$. Now, since π_ℓ fixes $\mathcal{Z}$ and $\pi_\ell^{-1}(\mathcal{Z}) \subset \mathcal{Z}$, $\pi_\ell : \mathbb{P}_1 \to \mathbb{P}_1$ is a rational map of degree $p^\ell > 1$, totally ramified above each $z \in \mathcal{Z}$. The Hurwitz genus formula gives immediately that $\mathcal{Z}$ contains at most two elements, hence precisely two elements.

After changing t with $\lambda(t)$ for a suitable $\lambda \in PGL_2(k)$ (replacing correspondingly π with $\lambda \circ \pi \circ \lambda^{-1}$) we may assume that $\mathcal{Z} = \{0, \infty\}$. Then we have that each $A_\imath$ is proportional to a power of t and the same holds for π_ℓ and π. Namely, for certain integers δ_i, elements $c_\imath, c \in k^*$ and a suitable choice of the sign we have

$$A_\imath(t) = c_\imath t^{\delta_\imath}, \quad \pi(t) = ct^{\pm p}. \tag{6}$$

(4) This argument is of course well known.

If the sign is negative, we can change t into $1/t$ (and change consequently π in $1/\pi$ and $B(u)$ in $B(1/u)$) so to assume that the plus sign holds in (6). We now distinguish two cases:

First case: $L_X = k(X)$. Then $k(X) = k(t)$, whence $\deg B = 1$ and $B \in PGL_2(k)$; then the equation $P(B(t)) = B(\pi(t)) = B(ct^p)$ easily implies (on comparing denominators and recalling that P is a polynomial) that $B(t) = a + bt^{\pm 1}$. Then $P(a + bt^{\pm 1}) = a + bct^{\pm p}$. Also, $\alpha_i = A_i(t) = A_i(B^{-1}(X)) = c_i(B^{-1}(X))^{\delta_i}$ (where B^{-1} denotes the inverse map).

Recall that $G_n(X) = \sum_{i=1}^{d} a_i(X)\alpha_i(X)^n$, [5]whence $G_n(B(X)) = \sum_{i=1}^{d} a_i(B(X))c_i X^{\delta_i n}$. We have just seen that $\lambda'(P(B(t)))$ is the cyclic polynomial t^p for a λ' in $PGL_2(k)$; all of this proves that we are now reduced to the cyclic case. We remark that it is then not difficult to deduce from (3) that the $a_i(B(X))$ are proportional to appropriate powers of X. We leave this further verification to the interested reader.

Second case: $[L_X : k(X)] > 1$. Recall that L_X is a normal extension of $k(X)$ (it is a splitting field). The Galois group, viewed as a group of automorphism of $k(t)$, is naturally a finite subgroup of $PGL_2(k)$. Such a Galois group permutes the $A_i(t) = c_i t^{\delta_i}$; therefore each element sends t to $\xi t^{\pm 1}$, for a root of unity $\xi \in k^*$ and some choice of the sign. Also, $t \mapsto \xi t$ implies $\xi^{\delta_i} = 1$ for all i: in fact, $c_i \xi^{\delta_i} t^{\delta_i}$ is some A_j; but $A_i/A_j \notin k$ for $i \neq j$ and thus $i = j$ and $\xi^{\delta_i} = 1$. Then $\xi^{\delta} = 1$ for the gcd δ of the δ_i. But $B(\xi t) = B(t)$ (since X is fixed), so B is a rational function of t^l for the order l of ξ, which divides δ. Since $k(t) = L_X = k(B(t), A_1(t), \dots, A_d(t))$ we deduce that $l = 1$, so $\xi = 1$. Hence the Galois group has order 2 and is generated by $t \mapsto \pm t^{-1}$ for some choice of the sign. By changing t into $\sqrt{-1}t$ if necessary we may suppose that the nontrivial automorphism is $t \mapsto t^{-1}$. Since $B(t)$ has then degree 2 and is invariant by $t \mapsto t^{-1}$, it is of the form $b(t + t^{-1})$, $b \in k^*$. Now the equation $P \circ B = B \circ \pi$ entails $P(b(t + t^{-1})) = b(\pi(t) + \pi(t)^{-1}) = b(ct^{\pm p} + c^{-1}t^{\mp p})$. Since this function is invariant by $t \mapsto t^{-1}$ we must have $c = \pm 1$. But then we find that P is, up to transformations in $PGL_2(k)$ the Chebyshev polynomial T_p. Also, from $X = b(t + t^{-1})$ we find $2bt = X \pm \sqrt{X^2 - 4b^2}$; after a linear transformation $X \mapsto 2bX$ we get the stated shape for t and the roots α_i. We see that we fall in the Chebyshev case.

Again, we remark that with these informations it is not difficult to

[5] Now the notation is well defined.

go further and determine the precise shape of the coefficients, but we omit the easy though a little tedious argument.

Proof of Corollary 4. — Note first that we may assume that Γ/k^* is finitely generated. In fact, assume the result true in this special case and let $g_1,\dots,g_r \in \Gamma$ be multiplicatively independent modulo k^*; let also Γ' be the group generated by k^* together with the g_i's. Then Γ is the union of the groups $\Gamma_n := \{g \in \Gamma : g^n \in \Gamma'\}$, each of which satisfies the opening assumption. Since the set E of solutions in question is the union of the sets E_n of solutions with $x, y \in \Gamma_n$, the estimates $\#E_n \leqslant 9^r$ and the inclusions $E_n \subset E_{nm}$, $n, m \in \mathbb{N}$, yield $\#E \leqslant 9^r$, as required.

Now, since Γ/k^* is torsion free and finitely generated, it is free abelian. Let then $\gamma_1,\dots,\gamma_r \in \Gamma$ be representatives for a basis of Γ/k^*. Then we may write for the solutions x, y in question,

$$x = \xi\gamma_1^{a_1}\cdots\gamma_r^{a_r}, \quad y = \eta\gamma_1^{b_1}\cdots\gamma_r^{b_r}, \quad \xi,\eta \in k^*, \quad a_1,\dots,a_r,b_1,\dots,b_r \in \mathbb{Z}.$$

We shall apply Theorem 1 with the following data: $h = 3$, $2r$ in place of r, $P_1 = \lambda, P_2 = \mu, P_3 = 1$, $A_1 = (\gamma_1,\dots,\gamma_r,1,\dots,1)$, $A_2 = (1,\dots,1,\gamma_1,\dots,\gamma_r)$, $A_3 = (1,\dots,1)$.

Note that if (x,y) is a solution of $\lambda x + \mu y = 1$, then $P_1A_1^{\mathbf{m}}$, $P_2A_2^{\mathbf{m}}$, $P_3A_3^{\mathbf{m}}$ are linearly dependent over k, where $\mathbf{m} = (a_1,\dots,a_r,b_1,\dots,b_r) \in \mathbb{Z}^{2r}$. Thus the solutions to our equation give rise to integral exponent vectors $\mathbf{m}$ which fall in at most $\binom{3}{2}^{2r} = 9^r$ classes, in the sense of Definition 2. To conclude the proof it suffices to show that each class can correspond to at most one solution (x,y) such that $\lambda x/\mu y \notin k^*$.

Assume the contrary and let (x_1,y_1), (x_2,y_2) be two distinct such solutions, whose exponent vectors, denoted $\mathbf{m}_1, \mathbf{m}_2 \in \mathbb{Z}^{2r}$, lie in a same class. The class cannot correspond to a set B properly contained in $\{1,2,3\}$, for otherwise two of the three terms of the equation would be linearly dependent over k, against the present assumptions. Hence the class corresponds to the whole set $\{1,2,3\}$, whence $A_i^{\mathbf{m}_2-\mathbf{m}_1} = (A_iA_3^{-1})^{\mathbf{m}_2-\mathbf{m}_1}$, $i = 1, 2$, both lie in k^*. Since the γ_i's are multiplicatively independent modulo k^* by assumption, this means that $\mathbf{m}_1 = \mathbf{m}_2$. Therefore $x_2 = \alpha x_1$, $y_2 = \beta y_1$ for some $\alpha, \beta \in k^*$. Then the equations $\lambda x_1 + \mu y_1 = \lambda x_2 + \mu y_2 = 1$ yield $(\beta - \alpha)\lambda x_1 = \beta - 1$ and $(\beta - \alpha)\mu y_1 = 1 - \alpha$. If $\alpha = \beta$ this implies $\alpha = \beta = 1$ and the solutions would not be distinct. Therefore $\alpha \neq \beta$, but then we deduce that λx_1 and μy_1 both lie in k, a contradiction which completes the argument.

BIBLIOGRAPHY

[A] J. Ax, On Schanuel's conjectures, Ann. Math., 93 (1971), 252–271.

[BMZ] E. Bombieri, J. Müller, U. Zannier, Equations in one variable over function fields, Acta Arith., 99 (2001), 27–39.

[BrM] D. Brownawell, D. Masser, Vanishing sums in function fields, Math. Proc. Camb. Phil. Soc., 100 (1986), 427–434.

[C] J.W.S. Cassels, An Introduction to Diophantine Approximation, Hafner, New York, 1972.

[Ch] C. Chevalley, The Theory of Algebraic Functions of One Variable, American Math. Soc. Math. Monographs, Vol. 6, 1991.

[D] V.I. Danilov, Algebraic Varieties and Schemes, in Algebraic Geometry I, I.R. Shafarevich Ed., Encyclopaedia of Math. Sciences, 23, Springer-Verlag, 1994.

[E] J.-H. Evertse, On equations in two S-units over function fields of characteristic 0, Acta Arith., 47 (1986), 233–253.

[EG] J.-H. Evertse, K. Györy, On the number of solutions of weighted unit equations, Comp. Math., 66 (1988), 329–354.

[ESS] J.H. Evertse, H.P. Schlickewei, W.M. Schmidt, Linear equations in variables which lie in a multiplicative group, Annals of Math., 155 (2002), 807–836.

[EZ] J.H. Evertse, U. Zannier, Linear equations with unknowns from a multiplicative group in a function field, Preprint, University of Leiden Report n° MI 2004-01, January 2004.

[FPT] C. Fuchs, A. Pethö, R.F. Tichy, On the Diophantine Equation $G_n(x) = G_m(P(x))$: Higher Order Recurrences, Transactions of the American Math. Soc., to appear.

[S] A. Schinzel, Polynomials with special regard to reducibility, Encyclopedia of Mathematics and its applications, vol. 77, Cambridge Univ. Press, 2000.

[Schm] W.M. Schmidt, Linear Recurrence Sequences and Polynomial-Exponential Equations, in Diophantine Approximation, F. Amoroso, U. Zannier Eds., Proc. of the C.I.M.E. Conference, Cetraro (Italy), 2000, Springer-Verlag LNM 1819, 2003.

[Z] U. Zannier, Some remarks on the S-unit equation in function fields, Acta Arith., LXIV (1993), 87–98.

Manuscrit reçu le 29 septembre 2003,
accepté le 26 janvier 2004.

Umberto ZANNIER,
Università degli studi di Udine
Dipartimento di Matematica e informatica
Via delle Scienze 206
33100 Udine (Italia).
zannier@dimi.uniud.it