BOUNDED HEIGHT IN PENCILS OF FINITELY GENERATED SUBGROUPS.

F. AMOROSO, D. MASSER AND U. ZANNIER

Abstract. In this paper we prove a general bounded height result for specializations in finitely generated subgroups varying in families which complements and sharpens the toric Mordell-Lang Theorem by replacing *finiteness* by *emptyness*, for the intersection of varieties and subgroups, all moving in a pencil, except for bounded height values of the parameters (and excluding identical relations).

More precisely, an instance of the results is as follows. Consider the torus scheme $\mathbb{G}_{\mathrm{m}/\mathcal{C}}^r$ over a curve $\mathcal{C}$ defined over $\overline{\mathbb{Q}}$, and let Γ be a subgroup-scheme generated by finitely many sections (satisfying some necessary conditions). Further, let V be any subscheme. Then there is a bound for the height of the points $P \in \mathcal{C}(\overline{\mathbb{Q}})$ such that, for some $\gamma \in \Gamma$ which does not generically lie in V, $\gamma(P)$ lies in the fiber V_P.

We further offer some direct diophantine applications, to illustrate once again that the results implicitly contain information absent from the previous bounds in this context.

1. Introduction

Let $\mathcal{C}$ be a projective smooth curve defined over $\overline{\mathbb{Q}}$, with function field denoted $\mathbb{F} := \overline{\mathbb{Q}}(\mathcal{C})$. In 1999 Bombieri and the second and third authors [7] proved a bounded height result in the multiplicative group $\mathbb{G}_{\mathrm{m}}^r$.

Theorem 1.1 ([7], Theorem 1'). *Let $\Gamma \subset \mathbb{G}_{\mathrm{m}}(\mathbb{F})$ be a finitely generated subgroup of non-zero rational functions on $\mathcal{C}$ such that the only constants in Γ are roots of unity. Then the height of the points $P \in \mathcal{C}(\overline{\mathbb{Q}})$, such that for some $\mathbf{x} \in \Gamma \setminus \{1\}$ we have $\mathbf{x}(P) = 1$, is bounded above.*

A significant special case is

$$t^n(1-t)^m = 1$$

provided only n, m are not both zero.

In this paper we prove a general bounded height result for specializations in finitely generated subgroups varying in families. This vastly extends the previously treated constant case and involves entirely different, and more delicate, techniques.

Before stating our results, note that Theorem 1.1 may be also phrased as a kind of toric analogue of Silverman's Specialization Theorem ([24], Theorem C). To illustrate this link, let us consider the 'trivial' family $\mathbb{G}_{\mathrm{m}/\mathcal{C}} := \mathbb{G}_{\mathrm{m}} \times \mathcal{C}$ and the sections $\gamma_i : \mathcal{C} \to \mathbb{G}_{\mathrm{m}} \times \mathcal{C}$, given by $P \mapsto (g_i(P), P)$, where g_i are generators for Γ, independent modulo constants. The above conclusion then means that *the set of points P where the values of the sections are multiplicatively dependent has bounded height.*

Now, multiplicative dependence at P means that some nontrivial monomial attains the value 1 at P. Then, rather more generally, given a constant family $\pi : \mathbb{G}_{\mathrm{m}/\mathcal{C}}^r := \mathbb{G}_{\mathrm{m}}^r \times \mathcal{C} \to \mathcal{C}$, a subvariety $V \subset \mathbb{G}_{\mathrm{m}}^r \times \mathcal{C}$ and sections $\gamma_i : \mathcal{C} \to \mathbb{G}_{\mathrm{m}/\mathcal{C}}^r$ generating a group Γ, we may ask the following

Question: *What can be said about those points P such that some nontrivial element of the group Γ when specialized at P lies on the fiber $V_P = \pi^{-1}(P) \subset \mathbb{G}_{\mathrm{m}}^r$.*

The previous situation is obtained in the very special case when V_P is constantly equal to the origin, *i. e.* $V = \text{origin} \times \mathcal{C}$. By contrast, we stress that here neither V nor the V_P are assumed to have any kind of group structure; especially this feature heavily prevents the previously known proof-pattern to apply.

The present paper offers in a sense a complete solution to this issue, proving that on the appropriate assumptions we have generally bounded height for *any* proper family of subvarieties. In particular, the said intersection is empty except for a 'sparse' set of points.

For simplicity, we phrase this conclusion in the language of Theorem 1.1. Namely, we consider a power $\mathbb{G}_{\mathrm{m}}^r$ of the multiplicative algebraic group and we let V be a subvariety of $\mathbb{G}_{\mathrm{m}}^r$ defined over $\mathbb{F}$; so we may view V as a family of varieties parameterised by $\mathcal{C}$. Then we denote by V_P, for almost all $P \in \mathcal{C}$, a specialized variety defined e.g. by specializing at P a given system of defining equations for V. Given a subgroup Γ of $\mathbb{G}_{\mathrm{m}}^r$ defined over $\mathbb{F}$ we say that Γ is constant-free if its image Γ' by any surjective homomorphism $\mathbb{G}_{\mathrm{m}}^r \to \mathbb{G}_{\mathrm{m}}$ satisfies the assumption $\Gamma' \cap \overline{\mathbb{Q}}^* = \Gamma'_{\mathrm{tors}}$ of Theorem 1.1.

With such notation, we have the following uniform complement to the toric case of Lang's conjecture.

Theorem 1.2. *Let $\Gamma \subset \mathbb{G}_{\mathrm{m}}^r(\mathbb{F})$ be a finitely generated constant-free subgroup and let V be a subvariety of $\mathbb{G}_{\mathrm{m}}^r$ defined over $\mathbb{F}$. Then the height of the points $P \in \mathcal{C}(\overline{\mathbb{Q}})$, such that for some $\mathbf{x} \in \Gamma \setminus V$ the value $\mathbf{x}(P)$ is defined and lies in V_P, is bounded above.*

Consider a generic situation when $\Gamma \cap V$ is empty and V_P does not contain a coset of positive dimension. The Mordell-Lang Theorem tells us that $\Gamma_P \cap V_P$ is finite for all P. Theorem 1.2 gives the following complement: for P of large height $\Gamma_P \cap V_P$ is *empty*.

As a consequence, we recover Theorem 1.1 of [7], taking $V = \{1\}$. An other known example is with $\mathcal{C}$ the affine line, $\mathbb{F} = \overline{\mathbb{Q}}(t)$ and $\Gamma \subset \mathbb{G}_{\mathrm{m}}^2$ the subgroup generated by $(t, 1-t)$. Let V be the hypersurface of $\mathbb{G}_{\mathrm{m}}^2$ defined by the equation $\alpha_1 x_1 + \alpha_2 x_2 = 1$; here we obtain bounded height for

$$\alpha_1 t^n + \alpha_2 (1-t)^n = 1 \tag{1.1}$$

unless $n = 1$ and $\alpha_1 = \alpha_2 = 1$, a result of Beukers [2].

Theorem 1.2 has also several entirely new consequences. For instance, choose $\mathcal{C}$ the affine line as above, $\Gamma \subset \mathbb{G}_{\mathrm{m}}^3$ generated by $(t, 1, 1)$, $(1-t, 1, 1)$, $(1, t, 1)$ and $(1, 1, 1+t)$. With V of equation $x_1 + x_2 + x_3 = 1$ we get bounded

height for the solutions of

$$t^n(1-t)^m + t^l + (1+t)^p = 1, \tag{1.2}$$

with no proviso on n, m, l, p. More generally, any equation

$$\alpha_1 M_1 + \cdots + \alpha_s M_s = 1 \tag{1.3}$$

for algebraic $\alpha_1, \ldots, \alpha_s$ and monomials $M_1 \ldots, M_s$ in fixed algebraic functions of t usually implies that the absolute height of t is bounded above independently of $M_1, \ldots, M_s$. We can also think of allowing $\alpha_1, \ldots, \alpha_s$ in (1.3) to be not necessarily fixed, for example as fixed algebraic functions of t, or even as numbers varying at most subexponentially in the exponents occurring in $M_1, \ldots, M_s$.

Addendum to Theorem 1.2. For a subvariety V defined over the constant field $\overline{\mathbb{Q}}$ (*i. e.* we have a 'trivial' family with $V_P = V$ for all P), the conclusion of Theorem 1.2 still holds for a subgroup Γ which is not necessarily constant-free, but such that $\Gamma/\Gamma \cap \mathbb{G}_{\mathrm{m}}^r(\overline{\mathbb{Q}})$ is of rank 1.

Note however that some assumption on Γ is needed. As a non-trivial example, we may take as $\mathcal{C}$ the affine line as above, and $\Gamma \subset \mathbb{G}_{\mathrm{m}}^2$ the subgroup generated by $\gamma_1 = (t, 1)$ and $\gamma_2 = (1, 2t)$ (note that these vectors are multiplicatively independent modulo constants). Let V be the hypersurface of $\mathbb{G}_{\mathrm{m}}^2$ defined by the equation $x_1 + x_2 = 0$. Then for $n \in \mathbb{N}$ the element $\mathbf{x}^{(n)} = \mathbf{x}^{(n)}(t) = \gamma_1^{n+1}\gamma_2^n = (t^{n+1}, 2^n t^n) \in \Gamma$ is not generically in V but its specialization at $t = -2^n$ is. Of course the image of Γ under the isogeny sending (x, y) to xy contains 2.

Choosing V a subvariety defined over the constant field $\overline{\mathbb{Q}}$, we obtain a bounded height result for certain 'unlikely intersections'.

Corollary 1.3. *Let $\mathcal{C} \subseteq \mathbb{G}_{\mathrm{m}}^r$ be a curve and let $V \subseteq \mathbb{G}_{\mathrm{m}}^r$ be a subvariety, both defined over $\overline{\mathbb{Q}}$. Then the height of the points $P \in \mathcal{C}(\overline{\mathbb{Q}})$, such that there exists an integer n with*[1] *$[n]\mathcal{C} \not\subseteq V$ and $[n]P \in V(\overline{\mathbb{Q}})$, is bounded above.*

Indeed, take $\Gamma = \langle (g_1, \ldots, g_r) \rangle$ where $g_1, \ldots, g_r$ are coordinate functions on $\mathcal{C}$, and let n be an integer such that $[n]\mathcal{C} \not\subseteq V$. Then $\gamma := (g_1^n, \ldots, g_r^n) \in \Gamma \setminus V$. By the addendum of Theorem 1.2, the height of the points $P \in \mathcal{C}(\overline{\mathbb{Q}})$ such that $[n]P \in V(\overline{\mathbb{Q}})$ is bounded independently of n.

Note that by the Skolem-Mahler-Lech theorem, for a fixed point $P_0 \in \mathbb{G}_{\mathrm{m}}^r$ the set of integers n with $[n]P_0 \in V$ is a union of a finite number of points and arithmetic progressions, and is 'usually' finite (if the set is not finite, V must contain the Zariski-closure of a set $[nq + s]P_0$, $n \in \mathbb{N}$, for suitable integers $q \neq 0, s$, and in particular must contain a coset of an algebraic subgroup of positive dimension unless P_0 is torsion). As in a comment above, the corollary says that when we move P_0 along a curve $\mathcal{C}$ the corresponding set is 'usually' empty, except for points $P_0 \in \mathcal{C}(\overline{\mathbb{Q}})$ of bounded height.

For more remarks on Theorem 1.2 and Corollary 1.3, see section 2.4.

[1] We denote as usual by $[n]\colon \mathbb{G}_{\mathrm{m}}^r \to \mathbb{G}_{\mathrm{m}}^r$ the morphism $\mathbf{x} \mapsto \mathbf{x}^n$ of multiplication by $n \in \mathbb{Z}$.

Let us now discuss proofs. Those of [7] may be said to rely on some form of simultaneous diophantine approximation to rational numbers. But the proof of Beukers for (1.1) is very different. The starting point is an explicit construction of Padé approximants to powers of the linear polynomial $1-t$, which goes back to [3] and involves hypergeometric polynomials. This leads to explicit identities of the form $A(t)t^n + B(t)(1-t)^n = C(t)$ (suitable for (1.1) above).

In our general situation, or even just for $\gamma_1^n+\gamma_2^n = 1$, we cannot hope to use such an explicit construction; just to mention one indication in this respect, results of Bombieri-Cohen (see [5]) suggest that in general the coefficients of these Padé approximants have a height growing much faster (*i. e.* like $\exp(cn^2)$) than in the case of Beukers (when it grows like $\exp(cn)$), and this would destroy the basic estimates necessary for the method to go through.

We use instead Thue's Method for avoiding such explicit constructions; this involves divided derivatives, Siegel's Lemma and a zero estimate based on Wronskians. However, to deal with certain unexpected vanishings we have to introduce a quite intricate descent, whose structure is different with respect to other investigations using Wronskians. This allows us to prove the following explicit version of a special case of Theorem 1.2.

Theorem 1.4. *Let $r \geq 2$ and $f_1, \dots, f_r \in \mathbb{F}$ be non-zero rational functions such that f_i/f_j is non-constant for some i and j. Then there exists a positive real number C depending only on $f_1, \dots, f_r$, having the following properties. Let $\boldsymbol{\alpha} = (\alpha_1 : \cdots : \alpha_r) \in \mathbb{P}^{r-1}(\overline{\mathbb{Q}})$. Consider, for a natural number n, a solution $P \in \mathcal{C}(\overline{\mathbb{Q}})$ of the equation*

$$\alpha_1 f_1(P)^n + \cdots + \alpha_r f_r(P)^n = 0. \tag{1.4}$$

Then, if $n \geq C$ and if there are no proper vanishing subsums, we have

$$h(P) \leq \frac{rh(\boldsymbol{\alpha})}{n} + C. \tag{1.5}$$

Our method provides even more explicit bounds: see Theorem 4.1 in Section 5, which is our main technical result. As also remarked by a referee we note that, given a fixed $\varepsilon > 0$, Theorem 4.1 allows to replace r in formula (1.5) by $r - 1 + \varepsilon$ for fixed arbitrarily small $\varepsilon > 0$, by allowing C to depend mildly on ε. Then remark 4.2, iii) shows that now this is sharp.

We finally remark that the assumption that n is sufficiently large is necessary: we may have $\alpha_1 f_1^n + \cdots + \alpha_r f_r^n = 0$ identically for some small values of n. Reciprocally, our results implies that $\alpha_1 f_1^n + \cdots + \alpha_r f_r^n$ cannot in fact be zero for *infinitely many* n. This last fact may be also derived directly on using the main result of [10] (*i. e.* a general form of the *abc*-inequality over function fields).[2]

Of course the use of Thue's Method in classical diophantine approximation is well-known to lead to results which are usually not effective. By contrast all the results of this paper are effective. For example with $t^n + (1-t)^n + (1+t)^n = 1$ (and $n \geq 0$) Denz [12] gets $H(t) \leq 2^{856}$ for the

[2]It is no coincidence that the proof of this uses Wronskians, which also appear in our arguments; on the other hand, no consideration of heights appears in [10].

standard (non-logarithmic, absolute) height.

1.1. `Plan of the paper.` In Section 3 we first define a simple notion of arithmetic height on rational functions and we prove some elementary estimates for this height, including some explicit bounds for the height of divided higher derivatives. Next we have two lemmas: Lemma 3.6 which is an auxiliary construction through Siegel's Lemma, and Lemma 3.11 which constructs by specialization some approximation forms. This last lemma is a fundamental tool of the inductive proof in section 5.

In Section 4 we state Theorem 4.1, which is a more precise and detailed version of Theorem 1.4. Then, we first illustrate the strategy of the proof in a test-run case. For the general proof, we need to overcome a technical obstacle related to non-vanishing. At this stage we only state a Claim and we deduce the Theorem in full generality from it.

Section 5 is devoted to a proof of the Claim which relies on a somewhat intricate descent argument.

The final Section 6 is devoted to the deduction of Theorem 1.2 from Theorem 4.1, namely the case of arbitrary rank from the case of rank one. This is done using simultaneous diophantine approximation applied to the exponents of the group-generators.

Aknowledgements. We thank the referees for their accurate consideration of the paper and several valuable suggestions. The first author thanks Bruno Anglès who draw his attention to Mirimanoff Polynomials.

2. Applications and further remarks

2.1. `Two simple applications.` We show by means of two simple examples that our results, actually already very special cases of them, are capable of applications to diophantine issues, recovering certain finiteness statements. The result achieved in the first example is known and may be obtained by a variety of techniques, but here we reach it directly as a consequence of the above corollary. The second example appears to be new. Probably it can be sharpened, but our aim here is merely to illustrate possible applications, not to develop them in depth. For this same reason we shall be sketchy in the arguments.

Example 1: A family of Thue's equations. Consider the Thue's equation

$$x^3 - (t^3 - 1)y^3 = 1,$$

where $t > 1$ is an integer, to be solved in integers x, y.

Our results easily imply that *there are only finitely many integers t such that the equation has an integer solution with $y \neq 0, 1$.*

In fact, let $u = \sqrt[3]{t^3 - 1}$, e.g. the real determination. It is very easy to see that the group of units of the ring $\mathbb{Z}[u]$ is generated by ± 1 and the unit $t - u$.

Let $(x, y) \in \mathbb{Z}^2$ be a solution, so $x - uy \in \mathbb{Z}[u]$ is a unit of norm 1, whence $= (t-u)^n$ for some integer n, and on taking conjugates over $\mathbb{Q}$ we have

$$(t-u)^n + \omega(t-\omega u)^n + \omega^2(t-\omega^2 u)^n = 0,$$

where ω is a primitive cube root of 1.

If $n > 1$, one may easily show that the left side does not vanish identically. Hence, by the corollary (with $\mathcal{C}$ a curve with function field $\overline{\mathbb{Q}}(z, \sqrt[3]{z^3-1})$ and $r = 3$), the height of t is bounded and finiteness follows.

This particular result (even that there are no t) has been known for nearly a century, but our method allows substantial generalizations; for example it may be applied to the context of a conjecture of Thomas [26] about families of Thue equations, and extended to non-Thue equations such as

$$x^3 - (t^3-1)y^3 + 3(t^3-1)xy + (t^3-1)^2 = 1$$

(which cannot be obtained from a Thue equation by an inhomogeneous linear transformation), and even to more variables such as

$$x^4 + (4t^4-1)y^4 + (4t^4-1)^2 z^4 + 2(4t^4-1)x^2z^2 - 4(4t^4-1)xy^2z = 1.$$

Example 2: Zeros of polynomial recurrences. Consider a linear recurrence sequence $(u_n(t))_{n\in\mathbb{N}}$ of polynomials in t, defined by prescribing polynomial initial data (not all zero) $u_0(t), \dots, u_{r-1}(t) \in \mathbb{Q}[t]$ and imposing

$$u_{n+r} = c_1(t)u_{n+r-1} + \cdots + c_r(t)u_n, \qquad n = 0, 1, \dots,$$

where $c_i(t)$ are polynomials with coefficients say in $\mathbb{Q}$. We assume for simplicity that the characteristic polynomial $Z^r - c_1(t)Z^{r-1} - \cdots - c_r(t)$ has no multiple roots in an algebraic closure of $\mathbb{Q}(t)$.

We assert that: *The set of algebraic numbers ξ such that for some n we have $u_n(t) \neq 0$ but $u_n(\xi) = 0$ has bounded height. In particular, there are only finitely many such ξ having bounded degree over $\mathbb{Q}$.*

Again, this follows rather immediately from the corollary and Northcott's theorem (see [6]), after expressing $u_n(t)$ as a linear combination of n-th powers of the *roots* of the characteristic polynomial of the recurrence.

A rephrasing of the last conclusion is that *for every given D there are only finitely many monic polynomials in $\mathbb{Q}[t]$ of degree $\leq D$ which may divide some $u_n(t) \neq 0$.*

In several cases we may further sharpen this, to say something about the irreducible factors. Assume for instance that $r \geq 3$ and that the characteristic polynomial of the recurrence has Galois group S_r over $\bar{\mathbb{Q}}(t)$ (which is the 'generic' case).[3] Under this assumption, we have:

Apart from a finite set of polynomials, the degrees of the irreducible factors (over $\mathbb{Q}$) of the polynomials u_n tends to infinity with n.

To prove this claim note first that the assumption implies in particular that the recurrence is non-degenerate, meaning that no ratio of its roots is a

[3]A doubly transitive Galois group would suffice for the sharpened conclusion. However it is beyond the scope of these examples to push the analysis further.

root of unity. In turn, by the Skolem-Mahler-Lech theorem, this yields that only finitely many u_n may vanish identically.

Suppose now that an algebraic number ξ_0 is a root of u_n for infinitely many n. Then, again by the Skolem-Mahler-Lech theorem 'several' ratios of the specialised roots become roots of unity, so the n-th powers of the roots of the recurrence collapse and vanishing of $u_n(\xi_0)$ gives rise to linear relations among these roots of 1. If the order of these roots of unity is eventually unbounded, then known theorems on torsion points on curves in $\mathbb{G}_{\mathrm{m}}^r$ (see [27]) imply that the roots of the recurrence may be grouped in subsets of cardinality ≥ 2 such that all the ratios of two roots in the same subset are multiplicatively dependent as algebraic functions; but the assumption on the Galois group then easily yields a contradiction. On the other hand, if the order of the relevant roots of unity is bounded, then we find only finitely many ξ_0, proving the assertion.

Note that some assumption as above is needed for this conclusion; an example in this direction is given by the Chebishev polynomials T_n defined by $T_0(t) = 2, T_1(t) = t$ and $T_{n+2}(t) = tT_{n+1}(t) - T_n(t)$ for $n \geq 0$. It turns out that, for odd m, T_q divides T_{mq} for all odd q, providing an example when the last conclusion is not true. Similarly for polynomials like $u_n(t) := T_n(t)T_{n+h}(t)$, where $r = 4$ but the Galois group is too small.

2.2. `A relation with Unlikely Intersections.` The boundedness of the height in the set of $P \in \mathcal{C}(\overline{\mathbb{Q}})$ such that $[n]P \in V(\overline{\mathbb{Q}})$ is related also to the context of Unlikely Intersections, and more precisely to degenerate cases of the former Bounded Height Conjecture of Bombieri-Masser-Zannier [8], nowadays a theorem of Habegger [13], as in Example 1.3 of [27].

Let $X \subseteq \mathbb{G}_{\mathrm{m}}^r$ be a subvariety. Define X^{oa} as the complement in X of the union of unlikely intersections of positive dimension, namely the components of some positive dimension $\delta > 0$ of some intersection $X \cap B$, where B is a translate of an algebraic subgroup and $\dim B \leq \delta + \mathrm{codim} X - 1$. By the former Habegger Theorem, the Weil height is bounded in the intersection of X^{oa} with the union of algebraic subgroups of dimension $\leq \mathrm{codim} X$.

Let now $X = \mathcal{C} \times V \subset \mathbb{G}_{\mathrm{m}}^r \times \mathbb{G}_{\mathrm{m}}^r$. If $P \in \mathcal{C}(\overline{\mathbb{Q}})$ is such that $[n]P \in V(\overline{\mathbb{Q}})$, then $(P, [n]P)$ is in the intersection of X with the algebraic subgroup $H_n = \{(\mathbf{x}, \mathbf{x}^n),\ \mathbf{x} \in \mathbb{G}_{\mathrm{m}}^r\}$ of dimension $\dim H_n = r \leq \mathrm{codim} X$. Thus it would be tempting to apply the former Bounded Height Conjecture = Habegger's Theorem. Nevertheless, such result gives us no information here, since X is *degenerate*: in the notation of [8] we have $X^{\mathrm{oa}} = \emptyset$.

This failure is not surprising, because in the degenerate cases we cannot hope to have in general bounded height in the whole union $\bigcup_n (X \cap H_n)$. However, our Corollary shows that, at least, bounded height is recovered in the projection of the whole union to the first factor.

In the special case $V = \mathcal{C}$, this kind of problem has been intensively studied by Bays and Habegger in [1], who show (under suitable assumptions) the *finiteness* of the set of $P \in \mathcal{C}$ such that $[n]P \in \mathcal{C}$ for some $n \geq 2$, thus giving a partial answer to a question (with $[n]P \in \mathcal{C}'$ instead for another curve $\mathcal{C}'$) originated by A. Levin (see [27] notes of Chapter 1, p.39). As

an important tool they first observe that the height in this set is bounded, except in the trivial case when $\mathcal{C}$ is a translate of a subtorus (op.cit., Lemma 6). This last result is a rather direct consequence of a generalized Vojta's inequality due to Rémond [20] (which however does not seem to apply when $\mathcal{C}' \neq \mathcal{C}$). This finiteness result provides further evidence of the usefulness of height bounds as in this paper.

2.3. `Abelian analogues?` In this paper we limit ourselves to the toric case, but analogous questions and statements can be naturally formulated in the elliptic and abelian context, thus extending Silverman's very setting.

Note that Silverman's specialization results in fact concern mainly abelian families, actually non-isotrivial, whereas algebraic tori have only isotrivial[4] families; this introduces some differences in the assumptions, as in results by Manin and Demjanenko, see [23]; for instance the assumption of mere *independence* has to be strengthened, as above, in the sense *modulo constants*, in order to obtain bounded height.

Further differences with respect to the abelian case are due to the fact that heights in abelian varieties behave like quadratic forms, so somewhat more 'regularly' than in the toric case, when this lack may introduce difficulties in some parts of the proofs for the toric case.

A specific example of what would be the analogue of our main issue for that case appears in the paper [16], where we find the pencil $\mathcal{J}$ of Jacobians J_t of the curves $H_t : y^2 = x^6 + x + t$, of generic genus 2, parametrized by $t \in \mathbb{A}^1$; we also find the section $\sigma : \mathbb{A}^1 \to \mathcal{J}$, obtained by defining $\sigma(t) \in J_t$ as the class of the divisor $\infty_+ - \infty_-$, difference of the two poles of x on H_t (let us forget here about the values for which H_t has genus < 2). One can now consider the subvariety V of $\mathcal{J}$ obtained as the union of H_t, embedding H_t in J_t e.g. through ∞_+.

The question now is: *What can be said about the points $t_0 \in \mathbb{A}^1$ such that $[n]\sigma(t_0) \in H_{t_0}$, for some $n = n(t_0) > 3$?* (We exclude here $n = 3$ because $[3]\sigma(t) \in H_t$ identically.) Now even to prove that this set of t_0 (necessarily algebraic) has infinite complement in $\overline{\mathbb{Q}}$ is far from being evident; the Appendix by Flynn to [16] in particular achieves this, and much more, providing nontrivial congruence conditions on the suitable t_0. However Flynn's method is not guaranteed to work generally, and moreover the question of whether the set of these numbers has or not bounded height remains open. In fact, we do not know if and how the present methods can be adapted to the abelian context. It seems to us a rather interesting issue to obtain even such a hyperelliptic analogue. Or even a "constant elliptic" analogue; for example the points P in $\mathcal{C}$ such that $[n]P \in \mathcal{C}'$ for some $n > 1$, where $\mathcal{C}$, $\mathcal{C}'$ are defined by $x_1 + x_2 = 1$, $x_1 + x_2 = c$ respectively in the product of $y_1^2 = x_1^3 + x_1 + 1$ and $y_2^2 = x_2^3 + x_2 + 1$ (we thank a referee for pointing out that when $\mathcal{C}' = \mathcal{C}$ this abelian version of Levin's question is covered by Rémond's abelian analogue [19] of [20]).

2.4. `Further remarks.` Let $\mathcal{C} \subseteq \mathbb{G}_{\mathrm{m}}^r$ be a curve and let $V \subseteq \mathbb{G}_{\mathrm{m}}^r$ be a subvariety, both defined over $\overline{\mathbb{Q}}$. Concerning the assumption $[n]\mathcal{C} \not\subseteq V$ in

[4]By 'isotrivial' we mean that the family becomes trivial, *i. e.* a product, after suitable extension of the base.

Corollary 1.3, a relevant issue is to detect the set of integers n such that $[n]\mathcal{C} \subseteq V$. Now, this amounts to $[n]\mathbf{x} \in V$ where the coordinates of $\mathbf{x}$ are the restrictions to $\mathcal{C}$ of the coordinate functions on $\mathbb{G}_{\mathrm{m}}^r$. Classical results describe the Zariski closure of these $[n]\mathbf{x}$ as a finite union of cosets of algebraic subgroups. Thus, if $\mathcal{C}$ is not contained in any translate of a proper subtorus, no proper algebraic subgroup can contain a multiple $[m]\mathbf{x}$ for $m \neq 0$, proving that the said set is finite unless V is the whole space[5]. Actually, using for instance results in [10], it is not difficult to reach directly this finiteness conclusion, moreover determining effectively the set.

A similar remark holds for Theorem 1.2: we may compute 'effectively' the intersection $\mathbf{x} \in \Gamma \cap V$, which however may be infinite in general.

3. Notation and Auxiliary Results.

3.1. **Rational functions.** Given $f_1, \ldots, f_r \in \mathbb{F}$ not all zero, we put[6]

$$\operatorname{div}(f_1, \ldots, f_r) := \sum_P \min_j \operatorname{ord}_P(f_j) P.$$

Note that $\deg(\operatorname{div}(f_1, \ldots, f_r)) = \deg(\operatorname{div}(gf_1, \ldots, gf_r))$ for any nonzero $g \in \mathbb{F}$.

Let us denote $d := -\deg(\operatorname{div}(f_1, \ldots, f_r))$. We remark that $d \geq 0$, since for example if $f_1 \neq 0$ we have $\sum_P \min_j \operatorname{ord}_P(f_j) \leq \sum_P \operatorname{ord}_P(f_1) = 0$. Moreover $d > 0$ if and only if some f_i/f_j is non-constant.

As a special case, let f be a non-zero rational function on $\mathcal{C}$. We define as usual its degree $d(f)$ as the the degree of the polar divisor $\operatorname{div}(f)_\infty = -\operatorname{div}(1, f)$. This is the geometric height of $(1 : f) \in \mathbb{P}^1(\mathbb{F})$.

`An arithmetic height on` $\mathbb{F}$. We define an arithmetic height $h(\cdot)$ of a rational function f on $\mathcal{C}$ as follows. We choose once and for all a non-constant $t \in \mathbb{F} \backslash \overline{\mathbb{Q}}$. Let $F(X, Y) \in \overline{\mathbb{Q}}[X, Y]$ be the irreducible polynomial such that $F(t, f) = 0$ (note that F has degree at most $d(f)$ in X and at most $d(t)$ in Y).

Definition 3.1. *For a function $f \in \mathbb{F}$, we define the height $h(f)$ as the projective Weil height of the vector of the coefficients of F.*

Clearly $h(1/f) = h(f)$. Also, this coincides with the affine height on $\overline{\mathbb{Q}}[t]$ (if $\mathcal{C}$ is the affine line and $f = P(t)$ is a polynomial then $F(X, Y) = P(X) - Y$.)

We shall need the following elementary estimates for this height; we could not find suitable references in the standard literature, and indeed the very definition of $h(f)$ is not quite so standard.

Lemma 3.2. *There is a constant c depending only on $\mathcal{C}$ and t with the following properties. Let $f \in \mathbb{F}$. Then we have*

(i) $h(f^n) \leq nh(f) + cnd(f)$ *for any positive integer* n,
(ii) $h(f') \leq c(h(f) + d(f)), d(f') \leq cd(f)$ *for* $f' = \mathrm{d}f/\mathrm{d}t$,

[5]See also the paper [25] by Silverman and Voloch for more general finiteness results in this direction.

[6]Since we have chosen a smooth projective model of $\mathcal{C}$, the closed points over $\overline{\mathbb{Q}}$ of the curve correspond to the places of its function field $\mathbb{F} = \overline{\mathbb{Q}}(\mathcal{C})$.

(iii) $h(\mathrm{tr} f) \leq h(f) + \log d(t), d(\mathrm{tr} f) \leq d(t)d(f)$ *for the trace tr from* $\mathbb{F}$ *to* $\overline{\mathbb{Q}}(t)$,

(iv) *for* $g \in \mathbb{F}$ *we have*

$$\max\{h(f+g), h(fg)\} \leq c(h(f) + h(g) + d(f) + d(g)),$$

(v) *for any non-constant* $s \in \mathbb{F} \backslash \overline{\mathbb{Q}}$ *there is* C *depending on* s *(and* $\mathcal{C}, t$*) such that the height of* f *with respect to* s *is at most* $C(h(f) + d(f))$.

Proof. We use resultants. Let $F(X, Y)$ be the irreducible polynomial for f over $\overline{\mathbb{Q}}(t)$: $F(t, f) = 0$. For (i) we can assume $d(f) \geq 1$, and we take the resultant of $F(X, Y)$ and $Y^n - Z$ with respect to Y to get a non-zero polynomial $F_n(X, Z)$ with $F_n(t, f^n) = 0$. Its degrees in X, Z are at most $nd(f), d(t)$ respectively. And its height is at most $n(h(f)+c)$; here one must be careful to avoid a factorial in the number N of terms in the Sylvester determinant, but it is easy to see that $N \leq (d(t) + 1)^n 2^{d(t)} \leq c^n$. Now F_n might not be irreducible, but by well-known estimates the height of any factor is at most

$$n(h(f) + c) + c(nd(f) + d(t)) \leq nh(f) + cnd(f);$$

and (i) follows.

A similar argument works with (v), now taking the resultant of $F(X, Y)$ and $S(X, Z)$ with respect to X, where $S(t, s) = 0$. Also with (iv): now say $G(t, g) = 0$ and then for $f + g$ we take the resultant of $F(X, Y)$ and $G(X, Z-Y)$ with respect to Y; then we do g/f with $F(X, Y)$ and $G(X, ZY)$ and deduce fg using $h(1/f) = h(f)$.

And (iii) is rather easy: if

$$F(X, Y) = F_0(X)Y^e + F_1(X)Y^{e-1} + \cdots + F_e(X)$$

then $\mathrm{tr} f = -\frac{d(t)}{e}\frac{F_1(t)}{F_0(t)}$. Finally for (ii) we can also assume $d(f) \geq 1$ and then we note that

$$f' = -\frac{F_0'(t)f^e + \cdots + F_e'(t)}{eF_0(t)f^{e-1} + \cdots + F_{e-1}(t)}$$

so that we can use (iv).

□

As at the beginning, we choose once and for all a system of Weil's functions associated to a divisor of degree 1 and a corresponding height h on $\mathcal{C}(\overline{\mathbb{Q}})$. By a well-known result of Néron (see [18]), the height of $P \in \mathcal{C}(\overline{\mathbb{Q}})$ differs from $h(t(P))/d(t)$ by an error term bounded by a constant multiple of $1 + h(P)^{1/2}$.

We need the following functorial bound for the arithmetic height associated to values of $(f_1, \ldots, f_r)$, which is an easy consequence of "Weil's Height Machine":

Lemma 3.3. *For* $r \geq 2$ *let* $f_1, \ldots, f_r \in \mathbb{F}$ *and* $P \in \mathcal{C}(\overline{\mathbb{Q}})$*, not a pole or a common zero of* $f_1, \ldots, f_r$*. Put* $d := -\deg \mathrm{div}(f_1, \ldots, f_r)$*. Then the projective Weil height*

$$h(f_1(P) : \cdots : f_r(P)) = dh(P) + O(1 + h(P)^{1/2})$$

where the implicit constant in the big-O may depend on $f_1, \dots, f_r$ but not on P.

Proof. Let $E' = -\mathrm{div}(f_1, \dots, f_r)$. We may assume $f_1, \dots, f_r$ linearly independent over $\overline{\mathbb{Q}}$; indeed, we may select a basis, say $f_1, \dots, f_s$ ($s \geq 2$), of the $\overline{\mathbb{Q}}$-vector space generated by $f_1, \dots, f_r$ and observe that $\mathrm{div}(f_1, \dots, f_s) = \mathrm{div}(f_1, \dots, f_r)$ and $h(f_1(P) : \cdots : f_r(P)) = h(f_1(P) : \cdots : f_s(P)) + O(1)$.

Given a divisor D, we denote by h_D the height on $\mathcal{C}(\overline{\mathbb{Q}})$ associated to it, which is defined up the addition of a bounded term: see [14], Part B for details.

Let $\phi\colon \mathcal{C} \to \mathbb{P}^{r-1}$ be the morphism $P \mapsto (f_1(P) : \ldots : f_r(P))$ and H be an hyperplane of $\mathbb{P}^{r-1}$. Then $\phi^* H$ is linearly equivalent to E' (indeed, for $i = 1, \dots, r-1$, set $g_i = f_i/f_r$; then $\phi^*\{x_r = 0\} = -\mathrm{div}(g_1, \dots, g_{r-1}, 1) \sim E'$).

Thus $h(f_1(P) : \cdots : f_r(P)) = h_{E'}(P) + O(1)$. Here and in the rest of this proof, the big-O depend on the divisors.

We now apply Theorem B.5.9 of [14]) (which goes back to Néron), taking as the ample divisor the divisor D of degree 1 such that $h = h_D + O(1)$ and as the divisor equivalent to zero the divisor $E = dD - E'$. We obtain $h_E(P) \leq c(1 + h_D(P)^{1/2})$. Moreover, by Theorem B.3.2 (d) of [14], $h_E = dh - h_{E'} + O(1)$. Thus

$$\begin{aligned} h(f_1(P) : \cdots : f_r(P)) &= h_{E'}(P) + O(1) \\ &= dh(P) - dh_E(P) + O(1) \\ &= dh(P) + O(1 + h(P)^{1/2}). \end{aligned}$$

□

Uniformity. The question of how the implicit constants in the O-terms in the lemma depend on the functions is a subtle one, and has been treated in number of papers, which control this dependence in data as the degree and heights of the functions. Here we can prove (at the expense of extra complication) the upper bound

$$dh(P) + O\big((d + \max h(f_i))(1 + h(P)^{1/2})\big)$$

with an implicit constant in the big-O which depends only on $\mathcal{C}$ and t. If some sort of refined Height Machine could deliver the analogous lower bound, even for $r = 2$, then it would imply at once some significant cases of our Theorem 1.4. For example with fixed different $F_1, \dots, F_r$ in $\overline{\mathbb{Q}}[t]$ of degree $p \geq 1$ and fixed sufficiently general $\alpha_1, \dots, \alpha_r$ in $\overline{\mathbb{Q}}$ we would have $d = np$ for

$$f(t) = \frac{\alpha_1 F_1(t)^n + \cdots + \alpha_{r-1} F_{r-1}(t)^n}{\alpha_r F_r(t)^n}.$$

Thus when $f(t) = -1$ we would deduce

$$0 = h(f(t)) \geq dh(t) - O\big(n(1 + h(t)^{1/2})\big)$$

so $h(t) = O(1)$.

However such a lower bound is false in general, as the example $f(t) = (t - 2^d)t^{d-1} - 1$ with $f(2^d) = -1$ shows; the lower bound would be

$$dh(2^d) - O\big(d(1 + h(2^d)^{1/2})\big) = d^2 \log 2 - O(d^{3/2}),$$

a contradiction for sufficiently large d.

We shall need also some good bounds for the values of higher derivatives $f^{(l)} = (\mathrm{d}/\mathrm{d}t)^l f$; or rather those of the divided derivatives $\delta_l f = f^{(l)}/l!$ essential to the success of Thue's Method. Iterating part (ii) of Lemma 3.2 does not suffice. In fact we have to consider certain monomial expressions whose curious weighting will soon be justified.

Lemma 3.4. *For any $f \in \mathbb{F}$ there is c, depending only on f and t, with the following property. Suppose f and t are regular at some $P \in \mathcal{C}(\overline{\mathbb{Q}})$ with $dt(P) \neq 0$. For any non-negative integer L let $\boldsymbol{\delta}$ be any vector with components*

$$\delta_0 f(P)^{a_0} \delta_1 f(P)^{a_1} \cdots \delta_l f(P)^{a_l}$$

for non-negative exponents satisfying

$$a_0 + a_1 + \cdots + a_l \leq L, \quad a_1 + \cdots + l a_l \leq L.$$

Then the affine Weil height $h(\boldsymbol{\delta})$ is at most $cL(h(P)+1)$.

Proof. This falls into the circle of Eisenstein-related ideas. We can assume that f is not constant. With $F(t,f) = 0$ as above and $\alpha_j = \delta_j f(P)$ the power series $y = \sum_{j=0}^{\infty} \alpha_j x^j$ satisfies $F_0(x,y) = 0$ with $F_0(X,Y) = F(t(P)+X, Y)$. We may therefore apply Theorem 1 (p.162) of Schmidt's paper [21]. He needs a number field k over which F_0 is defined. As he notes, the α_j lie in an extension K of k of relative degree at most the degree $e \geq 1$ of F_0 in Y; thus $[K:k] \leq d(t)$. We find for each valuation v on k some $A_v \geq 1$, with $A_v = 1$ for all but finitely many v, such that

$$|\alpha_j|_w \leq A_v^{m+j} \qquad (j = 0, 1, 2, \ldots)$$

for any valuation w on K over v, where $m \geq 1$ is the degree of F_0 in X; thus $m \leq d(f)$. Thus

$$|\alpha_0^{a_0} \alpha_1^{a_1} \cdots \alpha_l^{a_l}|_w \leq A_v^{m(a_0+a_1+\cdots+a_l)+(a_1+\cdots+la_l)} \leq A_v^{2mL}.$$

It follows for the non-logarithmic height

$$H(\boldsymbol{\delta})^{[K:\mathbb{Q}]} \;\leq\; \prod_v A_v^{2mL[K:k]}.$$

The v are split into two sets $S_{\infty 1}, S_2$ with

$$\prod_{v \in S_{\infty 1}} A_v \;\leq\; ((m+1)(e+1)\sqrt{e})^{(2e+1)[k:\mathbb{Q}]} H(F_0)^{2e[k:\mathbb{Q}]} \;\leq\; (2H(F_0))^{c[k:\mathbb{Q}]},$$

$$\prod_{v \in S_2} A_v \;\leq\; (16m)^{11e^3[k:\mathbb{Q}]} H(F_0)^{(2e^3+2e)[k:\mathbb{Q}]} \;\leq\; (2H(F_0))^{c[k:\mathbb{Q}]},$$

where $H(F_0)$ is still projective (and absolute). Thanks to the crucial (but nowadays natural) linear dependence on $[k:\mathbb{Q}]$ in the exponents we deduce $H(\boldsymbol{\delta}) \leq (2H(F_0))^{4cmL}$. Finally $H(F_0) \leq cH(t(P))^m$, and the result we want follows by relating $h(t(P))$ to $h(P)$ as described above.

We remark that the extra precision of [21] (especially concerning the set S_2) is not really necessary for us; thus by putting harmless additional restrictions on the point P we could have got ourselves into the "non-singular" situation, where the proofs are much easier (as for example in [6] p.360).

□

Given a divisor D we denote by $L(D)$ the finite-dimensional $\overline{\mathbb{Q}}$-vector space

$$L(D) = \{f \in \mathbb{F}^*,\ \mathrm{div}(f) + D \geq 0\} \cup \{0\}.$$

and by $l(D)$ its dimension. We shall need a good basis of $L(NQ)$ for fixed Q and large N. It is convenient to talk also of $L(\infty Q) = \sum_{N=1}^{\infty} L(NQ)$ the vector space of $f \in \mathbb{F}$ which are regular outside Q.

Lemma 3.5. *For any $Q \in \mathcal{C}(\overline{\mathbb{Q}})$ there is a positive integer Δ and real c together with $g, g_0, g_1, \ldots, g_{\Delta-1}$ in $L(\infty Q)$, depending only on Q and t, such that the following hold for any $N \geq 1$.*

(i) *We have $d(g) = \Delta$, and the elements*

$$g_j g^k \quad (j = 0, 1, \ldots, \Delta - 1, \quad d(g_j) + kd(g) \leq N)$$

form a basis for $L(NQ)$, with

$$N - c \leq l(NQ) \leq N + 1.$$

(ii) *For any f in $L(NQ)$ we have*

$$f = \sum \alpha_{jk} g_j g^k$$

with affine height

$$h(\ldots \alpha_{jk} \ldots) \leq c(h(f) + N),$$

as well as

$$h(f) \leq c(h(\ldots \alpha_{jk} \ldots) + N).$$

Proof. It is well known, for example by the Riemann-Roch Theorem, that as $f \neq 0$ varies over $L(\infty Q)$ the $-\mathrm{ord}_Q(f)$ (which are none other than the degrees $d(f)$) take all sufficiently large values. Let $\Delta \geq 1$ be the smallest positive value, attained by some g in $L(\infty Q)$. If $\Delta = 1$ we are done, as a standard argument of killing poles shows (which will be repeated below).

So we may and shall assume $\Delta \geq 2$. For $j = 1, \ldots, \Delta - 1$ pick g_j in $L(\infty Q)$ with $n_j = -\mathrm{ord}_Q(g_j) \equiv j$ modulo Δ and also as small as possible; here $n_j > 0$ is automatic and even $n_j > \Delta$. We define $g_0 = 1$ and $n_0 = 0$. We show by induction on N that these do the trick in (i).

Pick any $f \neq 0$ in $L(NQ)$, so that $n = -\mathrm{ord}_Q(f) \leq N$. If $n \equiv j$ modulo Δ $(j = 0, 1, \ldots, \Delta - 1)$ then $n \geq n_j$ and we can find α in $\overline{\mathbb{Q}}$ with $f - \alpha g_j g^{(n-n_j)/\Delta}$ in $L((N-1)Q)$. As

$$d(g_j) + \frac{n - n_j}{\Delta} d(g) = n \leq N$$

this shows by induction that the elements in (i) span $L(NQ)$. They are certainly linearly independent, as the

$$-\mathrm{ord}_Q(g_j g^k) = n_j + k\Delta \quad \left(j = 0, 1, \ldots, \Delta - 1, \quad k = 0, 1, \ldots, \left[\frac{N - n_j}{\Delta}\right]\right)$$

are all different (even for all k). In particular

$$l(NQ) = \sum_{j=0}^{\Delta-1} \left(1 + \left[\frac{N - n_j}{\Delta}\right]\right),$$

which leads easily to the required estimates using $n_j \geq \Delta$ for $j > 0$. So (i) is proved. Incidentally it is not difficult to estimate the constants so far solely in terms of the genus of $\mathcal{C}$.

For (ii) we note that this last argument even shows that $g_0, g_1, \ldots, g_{\Delta-1}$ are linearly independent over $\overline{\mathbb{Q}}(g)$. As $\Delta = d(g) = [\mathbb{F} : \overline{\mathbb{Q}}(g)]$ it follows that they form a basis of $\mathbb{F}$ over $\overline{\mathbb{Q}}(g)$. Now we can write

$$f = \sum_{j=0}^{\Delta-1} g_j G_j$$

with $G_j = \sum_k \alpha_{jk} g^k$ in $\overline{\mathbb{Q}}[g]$. In the standard way we multiply by $g_0, g_1, \ldots, g_{\Delta-1}$ and take the trace from $\mathbb{F}$ to $\overline{\mathbb{Q}}(g)$. The resulting equations can be solved for $G_0, G_1, \ldots, G_{\Delta-1}$. It follows easily from Lemma 3.2 (iii) (with g not t),(iv),(v) that

$$\sum_{j=0}^{\Delta-1} h(G_j) \leq c(h(f) + N).$$

But the affine height in (ii) is at most the analogous sum with heights taken with respect to g. So the first of the two required inequalities follows with another appeal to Lemma 3.2 (v). The second is similar but easier.

□

Before we go further we record the following identity for divided derivatives. Namely

$$\delta_l(f^n) = f^{n-l} \sum_{\mathbf{a}} C(\mathbf{a}) \delta_0(f)^{a_0} \delta_1(f)^{a_1} \cdots \delta_l(f)^{a_l} \tag{3.1}$$

where the sum is taken over all $\mathbf{a} = (a_0, a_1, \ldots, a_l)$ with non-negative coordinates satisfying

$$|\mathbf{a}| = a_0 + a_1 + \cdots + a_l = l, \quad a_1 + \cdots + l a_l = l$$

(see earlier) and the $C(\mathbf{a})$ are non-negative integers. Some version for undivided derivatives is attributed to the Blessed Francesco Faà di Bruno (who even has $\phi(f)$ instead of f^n), but in this divided form we get an immediate proof by formally writing $\tilde{f} = \sum_{m=0}^{\infty} \delta_m(f) T^m$. Note that, formally, $\tilde{f}(x) = f(x + T)$.

Taking this into account one may then go ahead by picking out the coefficient of T^l in $\tilde{f}^n$. We take $\delta_0(f) = f$ in a_0' of the factors $\tilde{f}$, and then $\delta_1(f)$ in a_1 of the factors, and so on. Then $a_0' + a_1 + \cdots + a_l = n$ and $a_1 + \cdots + l a_l = l$ making it clear that $a_0 = a_0' - (n - l) \geq 0$. We need also good estimates for the $C(\mathbf{a})$, but it is similarly clear that their sum is majorized by the coefficient of T^l in $(1 + T + T^2 + \cdots)^n = (1 - T)^{-n}$, which is

$$(-1)^l \binom{-n}{l} = \binom{n + l - 1}{l} \leq 2^{n+l}$$

and in particular factorial-free.

Now comes our basic "auxiliary polynomial".

Lemma 3.6. *Let $f_1, \ldots, f_r$ be in $\overline{\mathbb{Q}}(\mathcal{C})$, let Q in $\mathcal{C}(\overline{\mathbb{Q}})$ be not a pole or zero of $f_1, \ldots, f_r$ also with $dt(Q) \neq 0$, and write as before $d = -\deg \operatorname{div}(f_1, \ldots, f_r) \geq$*

0. *Then there are c_0, c depending only on $\mathcal{C}, f_1, \ldots, f_r, t$ and Q with the following property. For any non-negative integers $n, M_1, \ldots, M_r$ with $S > M + nd + c_0$ for*

$$S = M_1 + \cdots + M_r, \quad M = \max\{M_1, \ldots, M_r\}$$

define the "Dirichlet exponent"

$$\varrho = \frac{M + dn}{S - M - dn - c_0}.$$

Then there are $A_1 \in L(M_1 Q), \ldots, A_r \in L(M_r Q)$, not all zero and with heights at most $c(\varrho + 1)(M + n)$, such that $A_1 f_1^n + \cdots + A_r f_r^n = 0$.

Proof. By Lemma 3.5 (i) we can take

$$A_i = \sum \alpha_{ijk} g_j g^k \quad (i = 1, \ldots, r) \tag{3.2}$$

with j, k satisfying

$$j = 0, 1, \ldots, \Delta - 1, \quad d(g_j) + kd(g) \leq M_i$$

and algebraic numbers α_{ijk} to be determined. We first find them such that if $\phi = \sum_{i=1}^r A_i f_i^n \neq 0$ then

$$\mathrm{ord}_Q \phi > T \tag{3.3}$$

where the integer T is nearly as large as linear algebra allows.

If U is the number of unknowns α_{ijk} then their vector $\boldsymbol{\alpha}$ must lie in a certain subspace V of $\overline{\mathbb{Q}}^U$. Here $U = \sum_{i=1}^r l(M_i Q)$ so Lemma 3.5(i) gives

$$S - c \leq U \leq S + r. \tag{3.4}$$

We have $E = T + 1$ equations, so the dimension D of V satisfies

$$D \geq U - E \geq S - T - c_0. \tag{3.5}$$

Thus we assume $T < S - c_0$ for solvability. But if T is too near S then as in Thue's Method we would lose control of the heights. To regain this we use the version of the Absolute Siegel Lemma proved by David and Philippon [11]; for example taking $\epsilon = 1$ in the estimate at the bottom of page 523 we find non-zero $\boldsymbol{\alpha} \in V$ with

$$h(\boldsymbol{\alpha}) \leq \frac{h(V)}{D} + \frac{1}{2} \log D + 1 \tag{3.6}$$

where $h(V)$ is the euclidean height.

To estimate $h(V)$ we note that it is defined by certain equations, and by Hadamard's inequality for determinants (here the factorials don't matter) we get

$$h(V) \leq cE(\log E + h_{\mathrm{eq}}) \tag{3.7}$$

where h_{eq} is an upper bound for the logarithmic euclidean height of each equation. These are $\delta_l \phi(Q) = 0$ or more explicitly

$$\sum_{i=1}^r \sum_{j,k} \alpha_{ijk} \beta_{ijkl} = 0 \quad (l = 0, 1, \ldots, T)$$

with $\beta_{ijkl} = \delta_l(g_j g^k f_i^n)(Q)$. Now

$$\delta_l(g_j g^k f_i^n) = \sum \delta_s(g_j)\delta_p(g^k)\delta_q(f_i^n)$$

taken over all non-negative integers s, p, q with $s + p + q = l$. By (3.1)

$$\delta_p(g^k) = g^{k-p} \sum_{\mathbf{a}} C(\mathbf{a})\delta_0(g)^{a_0}\delta_1(g)^{a_1} \cdots \delta_p(g)^{a_p},$$

$$\delta_q(f_i^n) = f_i^{n-q} \sum_{\mathbf{b}} C(\mathbf{b})\delta_0(f_i)^{b_0}\delta_1(f_i)^{b_1} \cdots \delta_q(f_i)^{b_q}.$$

Now we see without difficulty thanks to Lemma 3.4 that $h_{\mathrm{eq}} \le c(T+M+n)$. Then (3.6) and (3.7) lead to

$$h(\boldsymbol{\alpha}) \le c\frac{(T+M+n)(T+1)}{S-T-c} + \log(S+c)$$

because $S - T - c_0 \le D \le U \le S + r$ by (3.4) and (3.5).

We now choose T so large that the condition (3.3) forces after all $\phi = 0$ in the sense that $\mathrm{ord}_Q \phi = \infty$. In fact (3.3) holds also for $\tilde{\phi} = \phi/f_r^n$ because $f_r(Q) \ne 0$, and since $A_1, \ldots, A_r$ are in $L(MQ)$ it is easy to see that $d(\tilde{\phi}) \le M + dn$. So $T = M + dn$ will do, leading to

$$h(\boldsymbol{\alpha}) \le c\varrho(M + dn) + \log(S + c).$$

Finally Lemma 3.5 (ii) gets us to $h(A_i)$ by (3.2); and then we use $S \le rM$.

□

3.2. **Orthogonal spaces and key lemma.**
Let $\mathbf{w}$ be a fixed vector of $\mathbb{C}^n$ with all entries non-zero. Given a subset Λ of $\{1, \ldots, r\}$ we consider the vector space

$$V_\Lambda = V_{\Lambda,\mathbf{w}} = \{\mathbf{v} \in \mathbf{w}^\perp \mid \forall j \notin \Lambda,\ \mathrm{v}_j = 0\}.$$

Thus $V_\emptyset = \{0\}$ and $\dim V_\Lambda = |\Lambda| - 1$ if $\Lambda \ne \emptyset$. We clearly have $V_{\Lambda_1} \cap V_{\Lambda_2} = V_{\Lambda_1 \cap \Lambda_2}$.

Remark 3.7. *Let Λ_1, Λ_2 be non-empty subsets of $\{1, \ldots, r\}$. If $\Lambda_1 \cap \Lambda_2 \ne \emptyset$ we have*

$$V_{\Lambda_1} + V_{\Lambda_2} = V_{\Lambda_1 \cup \Lambda_2}$$

while $V_{\Lambda_1} + V_{\Lambda_2}$ is a subspace of $V_{\Lambda_1 \cup \Lambda_2}$ of codimension 1 if $\Lambda_1 \cap \Lambda_2 = \emptyset$.

Proof. Let us assume $\Lambda_1 \cap \Lambda_2 \ne \emptyset$. The displayed formula follows from the trivial inclusions $V_{\Lambda_i} \subseteq V_{\Lambda_1 \cup \Lambda_2}$ and from the equality of dimensions:

$$\begin{aligned}\dim(V_{\Lambda_1} + V_{\Lambda_2}) &= \dim(V_{\Lambda_1}) + \dim(V_{\Lambda_2}) - \dim(V_{\Lambda_1} \cap V_{\Lambda_2}) \\ &= \dim(V_{\Lambda_1}) + \dim(V_{\Lambda_2}) - \dim(V_{\Lambda_1 \cap \Lambda_2}) \\ &= (|\Lambda_1| - 1) + (|\Lambda_2| - 1) - (|\Lambda_1 \cap \Lambda_2| - 1) \\ &= |\Lambda_1 \cup \Lambda_2| - 1 = \dim(V_{\Lambda_1 \cup \Lambda_2}).\end{aligned}$$

The last assertion follows similarly.

□

In order to generalise this simple remark, we introduce the following definition.

Definition 3.8. *Let $\Gamma = \{\Lambda_1, \dots, \Lambda_s\}$ be a collection of subsets of $\{1, \dots, r\}$. We say that a subset C of $\{1, \dots, r\}$ is a connected component of Γ, if, after renumbering $\Lambda_1, \dots, \Lambda_s$ if necessary, there exists an integer k with $1 \le k \le s$ such that*

i) $C = \Lambda_1 \cup \dots \cup \Lambda_k$;
ii) *for* $j = 1, \dots, k-1$ *we have* $\Lambda_{j+1} \cap (\Lambda_1 \cup \dots \cup \Lambda_j) \neq \emptyset$;
iii) *for* $j = k+1, \dots, s$ *we have* $C \cap \Lambda_j = \emptyset$.

We say that Γ is connected, if it has only one connected component.

We may also rephrase this definition, as follows: consider first the graph on $\{1, \dots, s\}$ defined by joining i, j if and only if $\Lambda_i \cap \Lambda_j \neq \emptyset$. Then a connected component in our sense is a union $C = \bigcup_{i \in U} \Lambda_i$, where U is a connected component, in the usual sense, of the graph just defined.

As an example, $\Gamma = \{\{1,2\}, \{3,4,5\}, \{2,5\}, \{6,7,8\}, \{7,8,9\}\}$ has two connected components, $\{1,2,3,4,5\}$ and $\{6,7,8,9\}$.

By Remark 3.7 we easily see that:

Remark 3.9. Let $\Gamma = \{\Lambda_1, \dots, \Lambda_s\}$ be a collection of subsets of $\{1, \dots, r\}$ and let $C_1, \dots, C_p$ be the connected components of Γ. Then,

i) $V_{\Lambda_1} + \dots + V_{\Lambda_s} = V_{C_1} + \dots + V_{C_p}$.
ii) $\dim(V_{\Lambda_1} + \dots + V_{\Lambda_s}) = |\Lambda_1 \cup \dots \cup \Lambda_s| - p$.

The following definition is crucial for our purposes.

Definition 3.10. *Let V be a $\overline{\mathbb{Q}}$-vector space, $v_1, \dots, v_s$ be $s \ge 2$ vectors of V. Let $a_1 v_1 + \dots + a_s v_s = 0$ be a non-trivial linear relation. We say that this relation is minimal if there are no non-trivial relations $\sum b_i v_i = 0$ over a proper non-empty subset of $\{1, \dots, s\}$.*

We remark that the relation $a_1 v_1 + \dots + a_s v_s = 0$ is minimal if and only if $a_1, \dots, a_s \in \overline{\mathbb{Q}}^*$ and $\dim\langle v_1, \dots, v_s \rangle = s - 1$.

We also remark that, given $v_1, \dots, v_s \in V$ linearly dependent and not all zero, there exists a subset $\Lambda \subseteq \{1, \dots, r\}$ such that $\{v_i\}_{i \in \Lambda}$ satisfy a minimal linear relation.

We now agree on some conventions which will be followed in the rest of this section and in the next section.

We fix as above $r \ge 2$ rational functions $f_1, \dots, f_r \in \mathbb{F} \backslash \{0\}$ and a non-constant $t \in \mathbb{F}$. We define S_0 as the finite set consisting of all zeros and poles of $f_1, \dots, f_r, \mathrm{d}t$. We fix a point $Q \in \mathcal{C}(\overline{\mathbb{Q}}) \backslash S_0$ and we define $S = S_0 \cup \{Q\}$. We choose once and for all a positive integer Δ, a real c and $g, g_0, g_1, \dots, g_{\Delta-1} \in L(\infty Q)$ depending only on Q and t and satisfying the statement of Lemma 3.5. The implicit constants in the big-O below will depend only on these data.

The next lemma is the main tool in our construction.

Lemma 3.11. *Let $\Lambda \subseteq \{1,\dots,r\}$ be of cardinality ≥ 2 and $(M_i')_{i\in\Lambda}$ be positive integers with maximum M'. Define*

$$d_\Lambda = -\deg \operatorname{div}(f_i)_{i\in\Lambda}, \qquad \Theta = \max\Big(1, \sum_{i\in\Lambda} M_i' - (M' + nd_\Lambda)\Big).$$

Let also $P \in \mathcal{C}(\overline{\mathbb{Q}})\backslash S$, $n \in \mathbb{N}$ and put

$$\mathbf{w} = (f_1(P)^n, \dots, f_r(P)^n).$$

Let finally $\{A_i\}_{i\in\Lambda} \subset \mathbb{F}$ not all zero, such that $A_i \in L(M_i'Q)$. Let us assume that $\{A_i f_i^n\}_{i\in\Lambda}$ satisfy a minimal linear relation. Then there exists a basis of algebraic vectors $\mathbf{v}_1, \dots, \mathbf{v}_{|\Lambda|-1}$ of $V_{\Lambda,\mathbf{w}}$ which satisfies

$$h(\mathbf{v}_j) \leq M'h(P) + O\left(\Theta(h(P)+1) + (n+M')(1+h(P)^{1/2}) + \max h(A_i)\right).$$

Proof. In the proof we use a Wronskian argument. Let us first recall some basic facts on it. The derivative $\frac{d}{dt}$ on $\overline{\mathbb{Q}}(t)$ can be uniquely extended to $\mathbb{F}$. For $F \in \mathbb{F}$ and $R \in \mathcal{C}(\overline{\mathbb{Q}})$ we have

$$\begin{cases} \operatorname{ord}_R(\mathrm{d}F/\mathrm{d}t) = \operatorname{ord}_R(F) - 1 - \operatorname{ord}_R(\mathrm{d}t), & \text{if } \operatorname{ord}_R(F) \neq 0; \\ \operatorname{ord}_R(\mathrm{d}F/\mathrm{d}t) \geq -\operatorname{ord}_R(\mathrm{d}t), & \text{if } \operatorname{ord}_R(F) = 0. \end{cases} \tag{3.8}$$

The (normalized) *Wronskian* of $F_1, \dots, F_k \in \mathbb{F}$ with respect to t is the determinant

$$W(F_1,\dots,F_k) = \det\left(\frac{1}{j!}\frac{\mathrm{d}^j F_i}{\mathrm{d}t^j}\right)_{\substack{i=1,\dots,k\\ j=0,\dots,k-1}}.$$

It is well known that $W = 0$ if and only if the F_i's are linearly dependent over $\overline{\mathbb{Q}}$.

Let us now go on with the proof of our lemma. We may assume $\Lambda = \{1,\dots,s\}$. For short we put $F_i = A_i f_i^n$ and $W_i = W(F_1,\dots,F_{i-1},F_{i+1},\dots,F_s)$ for $i = 1,\dots,s$.

For later reference, we remark that $F_1,\dots,F_s$ are S-integers (as elements of the function field $\mathbb{F}$): indeed the $\operatorname{div}(f_i)$ are supported in S and the A_i are also S-integers since $A_i \in L(M'Q)$. Moreover the zeros and the poles of dt are in S as well. Thus $\mathrm{d}^l F_i/\mathrm{d}t^l$ are S-integers (*cf* (3.8)), and so also W_s.

By assumption, we have a minimal linear relation $a_1F_1 + \dots + a_sF_s = 0$. Thus $F_1,\dots,F_{s-1}$ are linearly independent and $a_1,\dots,a_s \neq 0$. This proves that $W_s \neq 0$.

Let $1 \leq i \leq s-1$. Since $a_i \neq 0$ we can replace F_i by

$$-\frac{a_1}{a_i}F_1 - \dots - \frac{a_{i-1}}{a_i}F_{i-1} - \frac{a_{i+1}}{a_i}F_{i+1} - \dots - \frac{a_s}{a_i}F_s$$

in W_s. This shows that $W_s = \pm(a_s/a_i)W_i$.

We want to obtain a suitable upper bound for $m_0 := \operatorname{ord}_P(W_s)$. For this, we shall use the fact that W_s has already a big multiplicity at the zeros of F_i, since these functions are essentially n-th powers.

For $i = 1,\dots,s$ and for $l = 0,\dots,s-1$ we have (*cf* (3.8))

$$\operatorname{ord}_R\left(\frac{\mathrm{d}^l F_i}{\mathrm{d}t^l}\right) \geq \operatorname{ord}_R(F_i) - l(1+\operatorname{ord}_R(\mathrm{d}t)) = \operatorname{ord}_R(A_i) + n\operatorname{ord}_R(f_i) + O(1)$$

for any $R \in \mathcal{C}(\overline{\mathbb{Q}})$. Thus, for $i = 1, \dots, s$,

$$\begin{aligned}\operatorname{ord}_R(W_s) = \operatorname{ord}_R(W_i) &\geq \sum_{j \neq i} \Big(\operatorname{ord}_R(A_j) + n \operatorname{ord}_R(f_j) + O(1) \Big) \\ &= \sum_{j=1}^{s} n \operatorname{ord}_R(f_j) - \lambda_{R,i} + O(1).\end{aligned}$$

where $\lambda_{R,i} := n \operatorname{ord}_R(f_i) - \sum_{j \neq i} \operatorname{ord}_R(A_j)$. We deduce:

$$\begin{aligned}\operatorname{ord}_R(W_s) &\geq \max_{i=1,\dots,s} \left(\sum_{j=1}^{s} n \operatorname{ord}_R(f_j) - \lambda_{R,i} + O(1) \right) \\ &= \sum_{j=1}^{s} n \operatorname{ord}_R(f_j) - \lambda_R + O(1)\end{aligned}$$

where we have defined, for $R \in S$,

$$\lambda_R := \min_{i=1,\dots,s} \lambda_{R,i} = \min_{i=1,\dots,s} \Big\{ n \operatorname{ord}_R(f_i) - \sum_{j \neq i} \operatorname{ord}_R(A_j) \Big\}.$$

We shall use this inequality, for $R \in S$, in the functional (*i. e.* in $\mathbb{F}$) product formula, applied to W_s, namely the formula $\sum_R \operatorname{ord}_R(W_s) = 0$.

In this formula, for $R = P \notin S$ we find the quantity $\operatorname{ord}_P(W_s)$ that we have to estimate, whereas for R outside $S \cup \{P\}$ we use the trivial bound $\operatorname{ord}_R(W_s) \geq 0$. Also, since the $\operatorname{div}(f_i)$ are supported in $S_0 \subseteq S$, we have[7] $\sum_{R \in S} \operatorname{ord}_R(f_i) = \deg(\operatorname{div}(f_i)) = 0$. Moreover $P \notin S$. Thus

$$0 = \sum_{all\ R} \operatorname{ord}_R(W_s) \geq \operatorname{ord}_P(W_s) + \sum_{R \in S} \operatorname{ord}_R(W_s) \geq \operatorname{ord}_P(W_s) - \sum_{R \in S} \lambda_R + O(1).$$

We now recall that $A_i \in L(M_i' Q)$ and that the f_i are supported in S_0. Thus

$$\lambda_Q = \min_j \Big\{ -\sum_{i \neq j} \operatorname{ord}_Q(A_i) \Big\} \leq \min_j \sum_{i \neq j} M_i' = \sum_{i=1}^{s} M_i' - M'$$

and

$$\sum_{R \in S_0} \lambda_R \leq \sum_{R \in S_0} \min_i \{ n \operatorname{ord}_R(f_i) \} = n \sum_R \min_i \{ \operatorname{ord}_R(f_i) \} = -n d_\Lambda.$$

Collecting together these last three inequalities, we get the following sought upper bound for $m_0 = \operatorname{ord}_P(W_s)$:

$$m_0 \leq \sum_{R \in S} \lambda_R + O(1) \leq \sum_{i \in \Lambda} M_i' - (M' + n d_\Lambda) + O(1) = \Theta + O(1). \tag{3.9}$$

We can now construct the desired basis $\mathbf{v}_1, \dots, \mathbf{v}_{|\Lambda|-1}$ of $V_\Lambda = V_{\Lambda,\mathbf{w}}$. For a non-negative integer ρ we put as before

$$\delta_\rho = \frac{1}{\rho!} \frac{\mathrm{d}^\rho}{\mathrm{d}t^\rho}$$

[7]We use the fact that S contains not only the poles but also the zeros of the f_i's.

Given a vector of non-negative integers $\boldsymbol{\rho} = (\rho_1, \dots, \rho_{s-1})$ we let

$$W_{\boldsymbol{\rho}} = \det(\delta_{\rho_j} F_i)_{i,j=1,\dots,s-1}.$$

Thus $W_s = W_{(0,1,\dots,s-1)}$. It is also easily seen that

$$\delta_{m_0} W_s \in \sum_{|\boldsymbol{\rho}|=e} \mathbb{Z} W_{\boldsymbol{\rho}} \tag{3.10}$$

where $|\boldsymbol{\rho}| = \rho_1 + \cdots + \rho_{s-1}$ and $e = m_0 + 1 + \cdots + (s-2)$. By (3.9) we have

$$e \le \Theta + O(1). \tag{3.11}$$

Since $P \notin S$ and the zeros of dt are in S, we have $\mathrm{ord}_P(dt) = 0$. By definition of m_0 and by (3.8) we have $\mathrm{ord}_P(\delta_{m_0} W_s) = 0$. Again by (3.8) and since F_i are S-integers, for all $\boldsymbol{\rho}$ we have $\mathrm{ord}_P(\delta_{m_0} W_{\boldsymbol{\rho}}) \ge 0$. By (3.10), this implies that there exists $\boldsymbol{\rho}'$ with $|\boldsymbol{\rho}'| = e$ such that $\mathrm{ord}_P(W_{\boldsymbol{\rho}'}) = 0$. For $i = 1, \dots, s$ and $j = 1, \dots, s-1$, let

$$B_{ij} = a_i f_i^{-n} \delta_{\rho'_j} F_i. \tag{3.12}$$

Recall that: $P \notin S$, the f_i's are supported in S, the F_i have all their poles in S, the zeros of dt are in S. By (3.8) we see that $\mathrm{ord}_P(B_{ij}) \ge 0$. Thus $B_{ij}(P) \in \overline{\mathbb{Q}}$.

Since $a_1 F_1 + \cdots + a_s F_s = 0$, we have $B_{1j} f_1^n + \cdots + B_{sj} f_s^n = 0$ for $j = 1, \dots, s-1$. Thus, for $i = 1, \dots, s$,

$$\mathbf{v}_j = (B_{1j}(P), \dots, B_{sj}(P), 0, \dots, 0) \in V_{\{1,\dots,s\}}.$$

The important fact that we have achieved so far is that *since P is not a zero of $W_{\boldsymbol{\rho}'}$ the vectors $\mathbf{v}_1, \dots, \mathbf{v}_{s-1}$ are linearly independent and form a basis of $V_{\{1,\dots,s\}}$.*

By Lemma 3.2, $h(F_i) = O(n + M'_i + h(A_i))$. In order to deduce an upper bound for the height of B_{ij} we still need a bound for the height of the coefficients $a_1, \dots, a_s$ of the minimal linear relation $a_1 F_1 + \cdots + a_s F_s = 0$. Obviously, we may assume $a_s = -1$. We differentiate the relation up to order $s-2$. Since the rational functions $F_1, \dots, F_{s-1}$ are linearly independent over $\overline{\mathbb{Q}}$, their Wronskian is not zero and so we can solve the resulting system for $a_1, \dots, a_{s-1}$. Using Cramér's Rule and $h(F_i) = O(n + M'_i + h(A_i))$ together with Lemma 3.2 especially (i) and (ii) we find without difficulty that $h(a_i) = O(n + M' + \max h(A_i))$.

Now to simplify the notation we write $\mathbf{v}_j$ as $(B_1(P), \dots, B_s(P), 0, \dots, 0)$ for $B_i = a_i f_i^{-n} \delta_\rho(A_i f_i^n)$ as in (3.12); here $\rho \le e \le \Theta + O(1)$ by (3.11). As before

$$f_i^{-n} \delta_\rho(A_i f_i^n) = f_i^{-n} \sum_{l+m=\rho} \delta_l(A_i) \delta_m(f_i^n)$$

which by (3.1) is

$$\sum_{l+m=\rho} \delta_l(A_i) f_i^{-m} \sum_{\mathbf{a}} C(\mathbf{a}) \delta_0(f_i)^{a_0} \delta_1(f_i)^{a_1} \cdots \delta_m(f_i)^{a_m}$$

(note the changed power of f_i). As in Lemma 3.5(ii) we write

$$A_i = \sum_{j,k} \alpha_{ijk} g_j g^k \qquad (i = 1, \dots, s)$$

and it suffices here to take $k \le M_i'/\Delta \le M'/\Delta$. Again (3.1) for $\delta_q(g^k)$ gives

$$f_i^{-n}\delta_\rho(A_i f_i^n) = \sum_{l+m=\rho} f_i^{-m} \sum_{p+q=l} \sum_{j,k} \alpha_{ijk}\delta_p(g_j) g^{k-q} E_{qm},$$

where

$$E_{qm} = \sum_{\mathbf{b}} C(\mathbf{b})\delta_0(g)^{b_0}\delta_1(g)^{b_1}\cdots\delta_q(g)^{b_q} \sum_{\mathbf{a}} C(\mathbf{a})\delta_0(f_i)^{a_0}\delta_1(f_i)^{a_1}\cdots\delta_m(f_i)^{a_m}. \tag{3.13}$$

We evaluate all this at P and we want a main term $M'h(P)$ "uniformly in i", so that when we bundle the coordinates into the vector we don't get $sM'h(P)$.

In fact the terms g^{k-q} with $k \ge q$ already give

$$h\Big(g^{k-q}(P)\Big) = (k-q)h(g(P)) \le \frac{M'}{\Delta}h(g(P)) \le \frac{M'}{\Delta}\Big(\Delta h(P) + O(1+\sqrt{h(P)})\Big)$$

by the upper bound in Lemma 3.3 with just two functions. The right-hand side is

$$M'h(P) + O\big(M'(1+\sqrt{h(P)})\big)$$

so we already have the main term, clearly uniformly. Thus the rest had better be small. The point here is

$$b_0 + b_1 + \cdots + b_q = b_1 + \cdots + qb_q = q \le l \le \rho \le \Theta + O(1),$$
$$a_0 + a_1 + \cdots + a_m = a_1 + \cdots + ma_m = m \le \rho \le \Theta + O(1).$$

So if $k < q$ in g^{k-q} then $|k-q| \le q$ so we get

$$h\Big(g^{k-q}(P)\Big) = |k-q|h(g(P)) \le qh(g(P)) = O\left(\Theta(h(P)+1)\right).$$

And by Lemma 3.4 we get for the δ-terms in (3.13), as well as the f_i-term, a height of order at most

$$(q+m)(h(P)+1) \le \rho(h(P)+1) \le (\Theta + O(1))(h(P)+1),$$

also uniformly. The C-terms contribute logarithmically to order at most

$$k + q + n + m \le k + n + \rho \le M' + n + \Theta + O(1).$$

So this deals uniformly with the E_{qm}. As $p \le \rho$ the $\delta_p(g_j)$ $(j = 0, 1, \ldots, \Delta)$ give nothing new, and by Lemma 3.5(ii) the α-terms contribute $O(M' + \max h(A_i))$.

Collecting everything up, we get

$$h(\mathbf{v}_j) \le M'h(P) + O\left(\Theta(h(P)+1) + n + \max h(A_i) + M'(1+h(P)^{1/2})\right)$$

for $j = 1, \ldots, s-1$, slightly better than required.

□

4. Proof of Theorem 1.4.

The next theorem is a refined version of Theorem 1.4.

Theorem 4.1. *Let $r \geq 2$ and $f_1, \dots, f_r \in \mathbb{F}$ be non-zero rational functions such that f_i/f_j is non-constant for some $i \neq j$. Let $d = -\deg \operatorname{div}(f_1, \dots, f_r)$, $K > 0$ be sufficiently large with respect to $f_1, \dots, f_r$ and $\boldsymbol{\alpha} = (\alpha_1 : \cdots : \alpha_r) \in \mathbb{P}^{r-1}(\overline{\mathbb{Q}})$. Consider, for a natural number n, a solution $P \in \mathcal{C}(\overline{\mathbb{Q}})$ of the equation*

$$\alpha_1 f_1(P)^n + \cdots + \alpha_r f_r(P)^n = 0.$$

Then, if $n \geq K$ and if there are no proper vanishing subsums, we have

$$h(P) \leq \left(\frac{r-1}{d} + O(1/K) \right) \frac{h(\boldsymbol{\alpha})}{n} + O(K^2) \tag{4.1}$$

where the implicit constant in the big-O depends only on $f_1, \dots, f_r$.

Theorem 1.4 easily follows from Theorem 4.1. Indeed, choosing K sufficiently large, we have $(r-1)/d + O(1/K) \leq (r-1)/d + 1 \leq r$.

Remark 4.2.
i) In the proof we shall show:

$$\frac{h(\boldsymbol{\alpha})}{n} \geq \frac{dh(P)}{r-1} + O\left(\frac{1}{K} h(P) + h(P)^{1/2} + K \right). \tag{4.2}$$

which immediately implies (4.1), since either $h(P) \leq K^2$ or $h(P)^{1/2} + K \leq 2h(P)/K$. Note also that we will not use the assumption $f_i/f_j \neq$ constant in the proof of (4.2). This assumption is equivalent to $d \neq 0$ and (4.2) is trivially satisfied if $d = 0$.

ii) Let $\boldsymbol{\beta} = (f_1(P) : \cdots : f_r(P)) \in \mathbb{P}^{r-1}(\overline{\mathbb{Q}})$. By Weil's Height Machine Lemma 3.3 (and since $d > 0$), inequality (4.2) is equivalent to

$$\frac{h(\boldsymbol{\alpha})}{n} \geq \frac{h(\boldsymbol{\beta})}{r-1} + O\left(\frac{1}{K} h(\boldsymbol{\beta}) + h(\boldsymbol{\beta})^{1/2} + K \right) \tag{4.3}$$

(which in turn implies $h(\boldsymbol{\beta}) \leq (r - 1 + O(1/K)) \frac{h(\boldsymbol{\alpha})}{n} + O(K^2)$).

iii) A standard application of Siegel's lemma to the linear equation

$$\alpha_1 \beta_1^n + \cdots + \alpha_r \beta_r^n = 0 \tag{4.4}$$

in the unknowns $\alpha_1, \dots, \alpha_r$, shows that there exists a solution with

$$\frac{h(\boldsymbol{\alpha})}{n} \leq \frac{h(\boldsymbol{\beta})}{r-1} + O(1).$$

Thus inequality (4.3) is sharp. More precisely, Theorem 4.1 gives a lower bound for the first minimum (with respect to the height) of the linear equation (6.1), and hence shows that the successive minima are close to each other for large n.

iv) If we ask that f_i/f_j is non-constant for *all* $i \neq j$, the assumption on vanishing subsums can be easily removed (by induction on r).

v) We finally remark that we have a result even if $d = 0$ (*i. e.* f_i/f_j constant for all i, j), but now the lower bound $n \geq K$ must depend also on $\boldsymbol{\alpha}$. Indeed, if $f_i = c_i f_1$ with c_i constants, our equation becomes $(\alpha_1 c_1^n + \cdots + \alpha_r c_r^n) f_1(P)^n = 0$. By the Skolem-Mahler-Lech Theorem, if $f_1(P) \neq 0$ then n is bounded by a constant depending on $\alpha_1, \ldots, \alpha_r$ and on $c_1, \ldots, c_r$.

Strategy of the proof of Theorem 4.1. Let $f_1, \ldots, f_r \in \mathbb{F}$ be as in the statement of the theorem. We recall that we have chosen a non-constant rational function $t \in \mathbb{F}$ and that S is the finite set consisting of all zeros and poles of $f_1, \ldots, f_r, dt$ and of an extra point Q (which is neither a zero nor a pole of $f_1, \ldots, f_r, dt$).

We fix algebraic numbers $\alpha_1, \ldots, \alpha_r$, not all zero. In order to prove (4.1) we may suppose that P does not lie in any prescribed finite set of points. We thus choose $P \in \mathcal{C}(\overline{\mathbb{Q}}) \backslash S$ satisfying our equation

$$\alpha_1 f_1(P)^n + \cdots + \alpha_r f_r(P)^n = 0$$

for some $n \geq K$. We shall also assume $h(P) \geq 1$.

Put now $\mathbf{w} = (f_1^n(P), \ldots, f_r^n(P))$. Thus $\boldsymbol{\alpha} \in \mathbf{w}^\perp$, the orthogonal space of $\mathbf{w}$. Our strategy is the following:

We shall first construct a basis of function-vectors (with controlled heights) for the orthogonal of the vector $(f_1^n, \ldots, f_r^n) \in \mathbb{F}^r$. Then we shall specialise at P, in order to obtain a basis of $\mathbf{w}^\perp$, again with controlled heights. All of this shall involve an induction, necessary to take into account certain unexpected linear relations, *i. e.* relations with certain special properties in addition to those imposed by the construction.

At this stage we shall get a new basis, on replacing one of the vectors of the previous basis with $\boldsymbol{\alpha}$. By well-known facts, $\mathbf{w}$ and $\mathbf{w}^\perp$ have the same height. Lemma 3.3 gives a lower bound for the height of $\mathbf{w}$. The height of $\mathbf{w}^\perp$ is bounded from above by the sum of the heights of the vectors of our new basis. Comparing these bounds, we shall get the desired conclusion.

First step of the inductive proof.

Let N_1 be the minimum of the set of integers $m \geq 0$ such that there exist a non-empty $\Lambda \subseteq \{1, \ldots, r\}$ and rational functions $A_i \in \mathbb{F}$ $(i \in \Lambda)$ not all zero, satisfying

$$\begin{cases} A_i \in L(mQ), & \text{for } i \in \Lambda; \\ h(A_i) \leq nK, & \text{for } i \in \Lambda; \\ (A_i f_i^n)_{i \in \Lambda} \text{ are linearly dependent over } \overline{\mathbb{Q}}. \end{cases} \tag{4.5}$$

Fact 4.3.

$$(r-1)N_1 \leq nd + O(n/K). \tag{4.6}$$

Proof. We provide an upper bound for N_1 using Lemma 3.6, as we are going to illustrate. Let $0 < \varepsilon < 1/2$ and for this argument define N as the smallest integer such that

$$(r-1-\varepsilon)N \geq (1+\varepsilon)nd + c_0. \tag{4.7}$$

Then $N = O(n)$ and, with $M_1 = \cdots = M_r = N$, the Dirichlet exponent ϱ of Lemma 3.6 satisfies

$$\varrho = \frac{N + dn}{(r-1)N - dn - c_0} \leq \frac{1}{\varepsilon}.$$

Thus, by that lemma, there exist $A_1, \ldots, A_r \in L(NQ)$ not all zero such that

$$A_1 f_1^n + \ldots + A_r f_r^n = 0 \tag{4.8}$$

and

$$h(A_i) = O(n/\varepsilon + N/\varepsilon) = O(n/\varepsilon).$$

Choosing $\varepsilon = c/K$, where c is a sufficiently large constant to kill the implicit constant in the last $O()$, we see that there exists a non-trivial solution of (4.8) with $A_i \in L(NQ)$ and $h(A_i) \leq nK$. Not all $A_1, \ldots, A_r$ are zero, and we see that the non-zero ones among $A_1 f_1^n, \ldots, A_r f_r^n$ sum up to zero, and so they are linearly dependent, as required by (4.5) (on choosing Λ simply as the set of i such that $A_i \neq 0$).

This shows that $N_1 \leq N$. Since N is the smallest integer satisfying (4.7) and since $\varepsilon = c/K$, we have $(r-1)N = nd + O(n/K)$. Thus (4.6) holds.

□

Among all subsets Λ which realize the minimum defining N_1 in (4.5), we choose a subset which is minimal. We denote by Λ_1 such a set, by $l_1 = |\Lambda_1|$ its cardinality (necessarily $l_1 \geq 2$) and by $\{A_i^{(1)}\}_{i\in\Lambda_1}$ the corresponding rational functions. This implies in particular that $\{A_i^{(1)} f_i^n\}_{i\in\Lambda_1}$ satisfy a minimal linear relation.

To go ahead we want to apply Lemma 3.11 to find a suitable basis of V_{Λ_1}. In that lemma, let us put $M_i' = N_1$ for $i \in \Lambda_1$, so $M' = \max M_i' = N_1 = O(n)$ and

$$\Theta = \max\Big(1, \sum_{i\in\Lambda_1} M_i' - (M' + nd_1)\Big) = \max\Big(1, (l_1-1)N_1 - nd_1\Big)$$

with $d_1 = d_{\Lambda_1} = -\deg \operatorname{div}(f_i)_{i\in\Lambda_1}$.

Fact 4.4. $\Theta = O(n/K)$.

Proof. To prove this, we use again Lemma 3.6, this time on the f_i ($i \in \Lambda_1$), with $M_i = N_1 - 1$ for $i \in \Lambda_1$. The Dirichlet exponent ϱ is then

$$\varrho = \frac{N_1 - 1 + nd_1}{(l_1-1)(N_1-1) - nd_1 - c_0}.$$

By Lemma 3.6, there exist rational functions B_i not all zero such that $B_i \in L((N_1-1)Q)$ for $i \in \Lambda_1$, $\sum_{i\in\Lambda_1} B_i f_i^n = 0$ and

$$h(B_i) = O((\varrho+1)n).$$

By the minimality of N_1, we cannot have $\max_i h(B_i) \leq nK$. Thus $\varrho \geq K/c$, where c is a sufficiently large constant to kill the implicit constant in the last $O()$. This implies

$$(l_1 - 1)N_1 - nd_1 \leq \frac{c}{K}(N_1 + nd_1) + O(1) = O(n/K)$$

as required. This concludes the proof of this fact.

□

Using Lemma 3.11 and the inequalities

$$h(P) \geq 1, \quad h(A_i) \leq nK, \quad M_i' = N_1 = O(n), \quad \Theta = O(n/K),$$

we find a basis $\mathbf{v}_1^{(1)}, \ldots, \mathbf{v}_{l_1-1}^{(1)}$ of V_{Λ_1} satisfying

$$\begin{aligned} h(\mathbf{v}_i^{(1)}) &\leq M' h(P) \\ &\quad + O\left(\Theta(h(P)+1) + (n+M')(1+h(P)^{1/2}) + \max h(A_i)\right) \\ &= N_1 h(P) + O\left(\frac{n}{K}h(P) + nh(P)^{1/2} + nK\right). \end{aligned} \tag{4.9}$$

Proof of Theorem 4.1 in a test-run case. Let as now assume $\Lambda_1 = \{1, \ldots, r\}$, which is in essence the generic case. In this case we shall obtain directly the desired conclusion of Theorem 4.1, as we now show.

Recall that $\mathbf{w} = (f_1^n(P), \ldots, f_r^n(P))$. Since $\boldsymbol{\alpha}$ is a non-zero vector in $\mathbf{w}^\perp$ we may assume (reordering $\mathbf{v}_1^{(1)}, \ldots, \mathbf{v}_{r-1}^{(1)}$ if necessary) that

$$\boldsymbol{\alpha}, \mathbf{v}_1^{(1)}, \ldots, \mathbf{v}_{r-2}^{(1)}$$

is a basis of $\mathbf{w}^\perp$. Let us denote by h_2 the logarithmic euclidean height (defined on choosing the L_2-norm at the infinite places). By well-known facts on the height of subspaces (see [9] and [22]) and by the previous upper bounds for the height of these vectors.

$$\begin{aligned} h(\mathbf{w}) \leq h_2(\mathbf{w}) = h_2(\mathbf{w}^\perp) &\leq h_2(\boldsymbol{\alpha}) + \sum_{i=1}^{r-2} h_2(\mathbf{v}_i^{(1)}) + \log(r-1) \\ &\leq (r-2)N_1 h(P) + h(\boldsymbol{\alpha}) + O\left(\frac{n}{K}h(P) + nh(P)^{1/2} + nK\right). \end{aligned}$$

Moreover, by the functorial lower bound for the height Lemma 3.3, we have:

$$h(\mathbf{w}) \geq \left(h(P) + O(h(P)^{1/2})\right) nd.$$

Thus

$$0 \leq \lambda h(P) + \frac{h(\boldsymbol{\alpha})}{n} + O\left(\frac{1}{K}h(P) + h(P)^{1/2} + K\right)$$

with

$$\lambda = (r-2)N_1/n - d.$$

By (4.6) we have

$$\lambda \leq (r-2)\frac{d}{r-1} - d + O(1/K) = -\frac{d}{r-1} + O(1/K).$$

Inequality (4.2) follows.

Inductive construction. The obstacle in the approach of the test-run case is that Λ_1 may be smaller than $\{1,\dots,r\}$. If this happens, we can somewhat take advantage of the fact that we have an 'unexpected' dependence relation. To exploit this, let us sketch how we intend to argue by an induction procedure. The following claim resumes the inductive construction we shall do in the next section.

Claim 4.5. *There exists an integer s with $1 \le s \le r$, positive integers $N_1,\dots N_s$ and non-empty subsets $\Lambda_1,\Lambda_2,\dots,\Lambda_s$ of $\{1,\dots,r\}$ of cardinalities $l_1,l_2\dots,l_s$ satisfying:*

i) $N_1 \le N_2 \le \dots \le N_s$.

ii) *For $j=2,\dots,s$, the set Λ_j is contained in no connected component of $\{\Lambda_1,\dots,\Lambda_{j-1}\}$.*

iii) *The collection $\{\Lambda_1,\dots,\Lambda_s\}$ is connected and its union is the full set $\{1,\dots,r\}$.*

iv) *Let $t_1=\dim(V_{\Lambda_1})$ and, for $j=2,\dots,s$ let*

$$t_j=\dim(V_{\Lambda_1}+\dots+V_{\Lambda_j})-\dim(V_{\Lambda_1}+\dots+V_{\Lambda_{j-1}}).$$

Then for $j=1,\dots,s$ we have

$$t_1N_1+\dots+t_{j-1}N_{j-1}+\Big(r-1-\sum_{i=1}^{j-1}t_i\Big)N_j \le nd+O(n/K).$$

v) *For $j=1,\dots,s$, there exists a basis $\mathbf{v}_1^{(j)},\dots,\mathbf{v}_{l_j-1}^{(j)}$ of V_{Λ_j} satisfying*

$$\max_i h(\mathbf{v}_i^{(j)}) \le N_j h(P)+O\left(\frac{n}{K}h(P)+nh(P)^{1/2}+nK\right).$$

We shall explain how to perform this construction in the next section, proving the claim. For the moment, we pause to show how this claim allows us to conclude the proof of Theorem 4.1.

We state at once a general elementary lemma.

Lemma 4.6. *Let τ, ρ, $a_1,\dots,a_s$ be positive real numbers such that $a_s\ge 1$ and $a_1+\ldots+a_s=\rho$. Let also $x_1\le\ldots\le x_s$ be positive real numbers such that*

$$\sum_{i=1}^{s} a_ix_i \le \tau.$$

Then

$$\sum_{j=1}^{s-1} a_jx_j+(a_s-1)x_s-\tau \le -\tau/\rho.$$

Proof. Set $\sigma:=\sum_{i=1}^s a_ix_i$, so $\sigma\le\tau$ and also $\sigma\le\rho x_s$. Hence $\sigma-x_s\le\sigma(1-\frac{1}{\rho})\le\tau(1-\frac{1}{\rho})$, since $\rho\ge a_s\ge 1$.

Now, on subtracting τ from both sides we obtain $\sigma-x_s-\tau\le-\frac{\tau}{\rho}$, as required.

□

Deduction of Theorem 4.1 from Claim 4.5. We remark that

$$V_{\Lambda_1} + \cdots + V_{\Lambda_s} = \mathbf{w}^{\perp}. \quad \text{In particular, } t_1 + \cdots + t_s = r - 1, \tag{4.10}$$

by Claim 4.5 iii) and by Remark 3.9 i). Thus, taking into account i) of Claim 4.5, we may assume, after reordering for each j the vectors $\mathbf{v}_1^{(j)}, \ldots, \mathbf{v}_{l_j-1}^{(j)}$ and possibly omitting some of them, that

$$\mathbf{v}_1^{(1)}, \ldots, \mathbf{v}_{t_1}^{(1)}, \ldots, \mathbf{v}_1^{(s-1)}, \ldots, \mathbf{v}_{t_{s-1}}^{(s-1)}, \mathbf{v}_1^{(s)}, \ldots, \mathbf{v}_{t_s}^{(s)}$$

is a basis of $\mathbf{w}^{\perp}$.

We also remark that $t_s \geq 1$. Otherwise $\dim(V_{\Lambda_1} + \cdots + V_{\Lambda_{s-1}}) = t_1 + \cdots + t_{s-1} = r - 1$ by (4.10). By Remark 3.9 ii), this implies that $\{\Lambda_1, \ldots, \Lambda_{s-1}\}$ is connected and its union is the full set $\{1, \ldots, r\}$, which contradicts Claim 4.5 ii) with $j = s$.

By assumption there are no proper vanishing subsums in $\alpha_1 f_1(P)^n + \cdots + \alpha_r f_r(P)^n = 0$. This implies that $\boldsymbol{\alpha} \notin V_{\Lambda_1} + \cdots + V_{\Lambda_{s-1}}$. Thus we may assume that

$$\mathbf{v}_1^{(1)}, \ldots, \mathbf{v}_{t_1}^{(1)}, \ldots, \mathbf{v}_1^{(s-1)}, \ldots, \mathbf{v}_{t_{s-1}}^{(s-1)}, \boldsymbol{\alpha}, \mathbf{v}_1^{(s)}, \ldots, \mathbf{v}_{t_s-1}^{(s)}$$

is a basis of $\mathbf{w}^{\perp}$. Arguing as we did before, we deduce that

$$0 \leq \lambda h(P) + \frac{h(\boldsymbol{\alpha})}{n} + O\left(\frac{1}{K} h(P) + h(P)^{1/2} + K\right) \tag{4.11}$$

with

$$\lambda = t_1 \frac{N_1}{n} + \cdots + t_{s-1} \frac{N_{s-1}}{n} + (t_s - 1) \frac{N_s}{n} - d.$$

To go ahead, we recall that $t_1 + \cdots + t_s = r - 1$ (see (4.10)). Thus, Claim 4.5 iv) for $j = s$ reads

$$t_1 \frac{N_1}{n} + \cdots + t_{s-1} \frac{N_{s-1}}{n} + t_s \frac{N_s}{n} \leq d + O(1/K).$$

We apply the lemma 4.6 with

$$a_j = t_j, \quad \rho = r - 1, \quad x_j = N_j/n \quad \text{and} \quad \tau = d + O(1/K).$$

We find

$$\lambda = t_1 \frac{N_1}{n} + \cdots + t_{s-1} \frac{N_{s-1}}{n} + (t_s - 1) \frac{N_s}{n} - d \leq -d/(r-1) + O(1/K).$$

Thus, by (4.11),

$$0 \leq -\frac{d}{r-1} h(P) + \frac{h(\boldsymbol{\alpha})}{n} + O\left(\frac{1}{K} h(P) + h(P)^{1/2} + K\right).$$

Inequality (4.2) follows. This concludes the proof of Theorem 4.1, assuming the truth of Claim 4.5.

5. Proof of Claim 4.5

In this section, as promised, we detail our inductive process, verifying all the assertions of Claim 4.5.
We construct by induction an integer s with $1 \le s \le r$ and, for each $j = 1, \dots, s$,

- a positive integer N_j;
- a subset Λ_j of $\{1, \dots, r\}$ of cardinality denoted $l_j := |\Lambda_j|$;
- a subset J_j of $\Lambda_1 \cup \dots \cup \Lambda_{j-1}$ such that:
  - $|J_j \cap C| = 1$ for each connected component of $\{\Lambda_1, \dots, \Lambda_{j-1}\}$
  - J_j is disjoint from $\Lambda_1 \cup \dots \cup \Lambda_{j-2} \setminus J_{j-1}$ if $j \ge 2$.
- rational functions $\{A_i^{(j)}\}_{i \in \Lambda_j}$.
- a function $\varphi_j \colon \{1, \dots, r\} \to \{1, \dots, j\}$;

The role of the functions φ_j shall appear along the discussion.

For $j = 1$, we let N_1 be the minimum of the set of integers $N \ge 0$ such that there exist $\Lambda \subseteq \{1, \dots, r\}$ and rational functions $A_i \in \mathbb{F}$ $(i \in \Lambda)$ not all zero, of height $\le nK$, with $(A_i f_i^n)_{i \in \Lambda}$ linearly dependent and such that

$$A_i \in L(NQ), \text{ for } i \in \Lambda.$$

Among all subsets Λ which realize the minimum defining N_1, we choose a subset which is minimal. We denote Λ_1 such set, $l_1 = |\Lambda_1|$ its cardinality and $\{A_i^{(1)}\}_{i \in \Lambda_1}$ the corresponding rational functions. This implies in particular that $\{A_i^{(1)} f_i^n\}_{i \in \Lambda_1}$ satisfy a minimal linear relation. As we have already shown earlier (see (4.6) and (4.9)), we have

$$(r-1)N_1 \le nd + O(n/K)$$

and there exists a basis $\mathbf{v}_1^{(1)}, \dots, \mathbf{v}_{l_1-1}^{(1)}$ of V_{Λ_1} satisfying

$$h(\mathbf{v}_i^{(1)}) \le N_1 h(P) + O\left(\frac{n}{K} h(P) + n h(P)^{1/2} + nK\right).$$

We also set $J_1 = \emptyset$ and $\varphi_1(i) = 1$ for $i = 1, \dots, r$.

For $j \ge 2$ we go ahead similarly, but modifying somewhat the requirements for the A_j, taking into account the previous steps. More precisely, let $j \ge 2$ and assume to have already constructed $N_1, \dots, N_{j-1}$, $\Lambda_1, \dots, \Lambda_{j-1}$ and J_{j-1}, φ_{j-1}. If $\{\Lambda_1, \dots, \Lambda_{j-1}\}$ is connected and $\Lambda_1 \cup \dots \cup \Lambda_{j-1} = \{1, \dots, r\}$ we put $s = j-1$ and we stop here the process.

Otherwise, we choose a subset J_j of $\Lambda_1 \cup \dots \cup \Lambda_{j-1}$ such that $|J_j \cap C| = 1$ for each connected component of $\{\Lambda_1, \dots, \Lambda_{j-1}\}$. We need to show that we can choose J_j disjoint from $\Lambda_1 \cup \dots \cup \Lambda_{j-2} \setminus J_{j-1}$. If the set Λ_{j-1} does not intersect $\Lambda_1 \cup \dots \cup \Lambda_{j-2}$, the connected components of $\{\Lambda_1, \dots, \Lambda_{j-1}\}$ are the connected components of $\{\Lambda_1, \dots, \Lambda_{j-2}\}$ plus the set Λ_{j-1} itself. Thus we may choose $J_j = J_{j-1} \cup \{i_0\}$ where i_0 is any element of Λ_{j-1}. If otherwise Λ_{j-1} intersects $\Lambda_1 \cup \dots \cup \Lambda_{j-2}$, each connected component of $\{\Lambda_1, \dots, \Lambda_{j-1}\}$ contains at least one connected component of $\{\Lambda_1, \dots, \Lambda_{j-2}\}$. Thus we may choose J_j as a suitable subset of J_{j-1}.

Then we let N_j be the minimum of the set of integers $N \ge 0$ such that there exist $\Lambda \subseteq \{1, \dots, r\}$ and rational functions $A_i \in \mathbb{F}$ $(i \in \Lambda)$ not all zero,

satisfying[8]

$$(5.1)\quad \begin{cases} A_i \in L((N_{\varphi_{j-1}(i)}-1)Q), & \text{if } i \in (\Lambda_1 \cup \cdots \cup \Lambda_{j-1} \backslash J_j) \cap \Lambda, \\ A_i \in L(NQ), & \text{if } i \in \Lambda \text{ and } i \notin \Lambda_1 \cup \cdots \cup \Lambda_{j-1} \backslash J_j, \\ h(A_i) \le nK, & \text{for } i \in \Lambda, \\ (A_i f_i^n)_{i\in\Lambda} \text{ are linearly dependent.} \end{cases}$$

We remark that the set of such N is indeed not empty, as we easily see since $\{1,\dots,r\}\backslash(\Lambda_1 \cup \cdots \cup \Lambda_{j-1}\backslash J_j)$ has cardinality ≥ 2 (for otherwise $\{\Lambda_1,\dots,\Lambda_{j-1}\}$ would be connected and $\Lambda_1 \cup \cdots \cup \Lambda_{j-1} = \{1,\dots,r\}$).

We select a minimal set Λ_j among all sets Λ which realize the minimum defining N_j. We denote by $\{A_i^{(j)}\}_{i\in\Lambda_j}$ the corresponding rational functions. Thus $\{A_i^{(j)} f_i^n\}_{i\in\Lambda_j}$ satisfy a minimal linear relation.

We finally set

$$(5.2)\quad \varphi_j(i) = \begin{cases} \varphi_{j-1}(i) & \text{if } i \in \Lambda_1 \cup \cdots \cup \Lambda_{j-1} \backslash J_j, \\ j & \text{otherwise, } i.\,e. \text{ if } i \notin \Lambda_1 \cup \cdots \cup \Lambda_{j-1} \text{ or } i \in J_j. \end{cases}$$

Thus, by the minimality of N_j, there are no subsets Λ and rational functions $(A_i)_{i\in\Lambda}$ not all zero such that

$$\begin{cases} A_i \in L((N_{\varphi_j(i)}-1)Q), & \text{for } i \in \Lambda, \\ h(A_i) \le nK, & \text{for } i \in \Lambda, \\ (A_i f_i^n)_{i\in\Lambda} \text{ are linearly dependent.} \end{cases}$$

We notice that for $j = 2,\dots,s$ we have

$$(5.3)\qquad \varphi_{j-1}(i) = j-1 \text{ for } i \in J_j,$$

since J_j is disjoint from $\Lambda_1 \cup \cdots \cup \Lambda_{j-2} \setminus J_{j-1}$.

Proof of Claim 4.5 i). Since at each step we have added conditions, we have

$$(5.4)\qquad N_1 \le N_2 \le \cdots \le N_s$$

as required.

Proof of Claim 4.5 ii) and iii) We first showthat for $j = 2,\dots,s$, the set Λ_j is contained in no connected component of $\{\Lambda_1,\dots,\Lambda_{j-1}\}$.

To verify this assertion, let us assume by contradiction that Λ_j is contained in a connected component C of $\{\Lambda_1,\dots,\Lambda_{j-1}\}$. Let $C \cap J_j = \{i_0\}$. Then for $i \in \Lambda_j$, $i \ne i_0$, we have $A_i^{(j)} \in L((N_{\varphi_{j-1}(i)}-1)Q)$. Using the minimal linear relation $\sum_{i\in\Lambda_j} a_i A_i^{(j)} f_i^n = 0$ and the fact that Q is not in the support of any of the divisors $\operatorname{div}(f_i)$ we see that $A_{i_0}^{(j)} \in L(N^*Q)$ with $N^* = \max_i(N_{\varphi_{j-1}(i)}-1)$. By (5.4) we have $N^* \le N_{j-1}$ and, by (5.3), $j-1 = \varphi_{j-1}(i_0)$. Thus $A_i^{(j)} \in L((N_{\varphi_{j-1}(i)}-1)Q)$ for *all* $i \in \Lambda_j$. This contradicts the minimality in the definition of N_{j-1} (see the remark after (5.2)).

[8]Note that we are prescribing somewhat more stringent conditions than before on the indices inside the subsets previously defined.

Thus for $j = 2, \dots, s$ the set Λ_j is contained in no connected component of $\{\Lambda_1, \dots, \Lambda_{j-1}\}$. This proves assertion ii) of Claim 4.5 and ensures that the inductive process ends somewhere. Thus Claim 4.5 iii) also holds by inductive construction.

We still have to check assertions iv) and v) of Claim 4.5. To prove assertion iv) we first need the following lemma. Let, as in Claim 4.5 iv), $t_1 = \dim(V_{\Lambda_1})$ and

$$t_j = \dim(V_{\Lambda_1} + \cdots + V_{\Lambda_j}) - \dim(V_{\Lambda_1} + \cdots + V_{\Lambda_{j-1}}).$$

for $j = 2, \dots, s$.

Lemma 5.1.

i) *For $j = 2, \dots, s$ we have:*

$$|\Lambda_1 \cup \cdots \cup \Lambda_{j-1}| - |J_j| = \dim(V_{\Lambda_1} + \cdots + V_{\Lambda_{j-1}}) = \sum_{j'=1}^{j-1} t_{j'}$$

ii) *Let j, j' be two integers with $j \geq 2$ and $1 \leq j' \leq j-1$. Then the set of $i \in \{1, \dots, r\}$ such that $\varphi_j(i) = j'$ has cardinality $t_{j'}$.*

Proof. The first assertion follows from Remark 3.9 ii), since J_j has cardinality equal to the number of connected components of $\Lambda_1 \cup \cdots \cup \Lambda_{j-1}$.

We prove the second assertion by induction on j. For $j = 2$ we have by construction $\#\{i \mid \varphi_2(i) = 1\} = |\Lambda_1| - 1 = \dim(V_{\Lambda_1}) = t_1$.

Let $j > 2$ and assume that $\#\{i \mid \varphi_{j-1}(i) = j'\} = t_{j'}$ for $j' = 1, \dots, j-2$. We want to show that $\#\{i \mid \varphi_j(i) = j'\} = t_{j'}$ for $j' = 1, \dots, j-1$.

Assume first $j' \leq j-2$. Let i be such that $\varphi_j(i) = j'$. Then $\varphi_j(i) \neq j$, thus by (5.2) $\varphi_j(i) = \varphi_{j-1}(i)$. This shows that $\{i \mid \varphi_j(i) = j'\} \subseteq \{i \mid \varphi_{j-1}(i) = j'\}$. On the other hand, let i such that $\varphi_{j-1}(i) = j'$. Then $\varphi_{j-1}(i) \neq j-1$ and (5.2) (with j replaced by $j-1$) shows that

$$i \in \Lambda_1 \cup \cdots \cup \Lambda_{j-2} \backslash J_{j-1} \subseteq \Lambda_1 \cup \cdots \cup \Lambda_{j-1}.$$

Moreover $i \notin J_j$ by (5.3). Thus, by (5.2), $\varphi_j(i) = \varphi_{j-1}(i)$. This proves that $\{i \mid \varphi_{j-1}(i) = j'\} \subseteq \{i \mid \varphi_j(i) = j'\}$. Putting together the two inclusions we see that

$$\{i \mid \varphi_j(i) = j'\} = \{i \mid \varphi_{j-1}(i) = j'\}.$$

Thus, by induction, $\#\{i \mid \varphi_j(i) = j'\} = \#\{i \mid \varphi_{j-1}(i) = j'\} = t_{j'}$.

Assume now $j' = j-1$. By (5.2)

$$\{i \mid \varphi_j(i) = j-1\} = \{i \in \Lambda_1 \cup \cdots \cup \Lambda_{j-1} \backslash J_j \mid \varphi_{j-1}(i) = j-1\}$$

By (5.3) we have $\varphi_{j-1}(i) = j-1$ on J_j. By (5.2) (again with j replaced by $j-1$) we still have $\varphi_{j-1}(i) = j-1$ outside $\Lambda_1 \cup \cdots \cup \Lambda_{j-2}$ and thus, *a fortiori* outside $\Lambda_1 \cup \cdots \cup \Lambda_{j-1}$. These facts and the first assertion of the present lemma imply

$$\begin{aligned}\#\{i \mid \varphi_j(i) = j-1\} &= \#\{i \mid \varphi_{j-1}(i) = j-1\} - (r - |\Lambda_1 \cup \cdots \cup \Lambda_{j-1}|) - |J_j| \\ &= \dim(V_{\Lambda_1} + \cdots + V_{\Lambda_{j-1}}) - (r - \#\{i \mid \varphi_{j-1}(i) = j-1\}).\end{aligned}$$

By induction,

$$r - \#\{i \mid \varphi_{j-1}(i) = j-1\} = \sum_{j'=1}^{j-2} \#\{i \mid \varphi_{j-1}(i) = j'\} = \sum_{j'=1}^{j-2} t_{j'}$$
$$= \dim(V_{\Lambda_1} + \cdots + V_{\Lambda_{j-2}}).$$

Putting together the last two displayed equations, we get

$$\#\{i \mid \varphi_j(i) = j-1\} = \dim(V_{\Lambda_1} + \cdots + V_{\Lambda_{j-1}}) - \dim(V_{\Lambda_1} + \cdots + V_{\Lambda_{j-2}}) = t_{j-1}$$

as desired.

□

Proof of Claim 4.5 iv). We proceed by induction. For $j = 1$ assertion iv) was already proved in (4.6). Let $j \geq 2$. In a similar way as we have done for N_1, we are going to provide an upper bound for N_j using Lemma 3.6. Let $\varepsilon = c/K$ for a sufficiently large constant c and choose N as the smallest integer such that

$$\sum_{j'=1}^{j-1} t_{j'}(N_{j'} - 1) + \Big(r - 1 - \sum_{j'=1}^{j-1} t_{j'} - \varepsilon\Big)N \geq (1+\varepsilon)dn + c_0. \tag{5.5}$$

We notice that $N = O(n)$.

Fact 5.2. *We have*

$$N_{j-1} \leq N.$$

Proof. By Claim 4.5 iv) with j replaced by $j-1$ (which holds by the present inductive assumption) we have

$$\sum_{j'=1}^{j-2} t_{j'} N_{j'} + \Big(r - 1 - \sum_{j'=1}^{j-2} t_{j'}\Big)N_{j-1} \leq nd + O(n/K).$$

Thus

$$\begin{aligned}
&\Big(r - 1 - \sum_{j'=1}^{j-1} t_{j'} - \varepsilon\Big)(N_{j-1} - N) \\
&\quad = \Big(r - 1 - \sum_{j'=1}^{j-2} t_{j'}\Big)N_{j-1} - t_{j-1}N_{j-1} - \varepsilon N_{j-1} - \Big(r - 1 - \sum_{j'=1}^{j-1} t_{j'} - \varepsilon\Big)N \\
&\quad \leq \Big(nd + O(n/K)\Big) - \sum_{j'=1}^{j-2} t_{j'}N_{j'} - t_{j-1}N_{j-1} - \Big(r - 1 - \sum_{j'=1}^{j-1} t_{j'} - \varepsilon\Big)N \\
&\quad \leq \Big(nd + O(n/K)\Big) - \Big((1+\varepsilon)dn + c_0)\Big) \\
&\quad = \big(-cd + O(1)\big)\frac{n}{K} < 0,
\end{aligned}$$

if c is a sufficiently large constant to kill the implicit constant in the $O(1)$.

□

We are going to apply Lemma 3.6 with $f_1, \dots, f_r$ and with

$$M_i = \begin{cases} N_{\varphi_{j-1}(i)} - 1, & \text{if } i \in (\Lambda_1 \cup \dots \cup \Lambda_{j-1} \backslash J_j); \\ N, & \text{if } i \notin \Lambda_1 \cup \dots \cup \Lambda_{j-1} \backslash J_j \end{cases}$$

$(i = 1, \dots, r)$. Since Q is neither a zero nor a pole of $f_1, \dots, f_r$ and since $N_1 \le \dots \le N_{j-1} \le N$ (by Claim 4.5 i) and by Remark 5.2), we have

$$M = \max_i M_i = N. \tag{5.6}$$

We recall that, by Lemma 5.1 i),

$$|\Lambda_1 \cup \dots \cup \Lambda_{j-1}| - |J_j| = \sum_{j'=1}^{j-1} t_{j'}$$

and, by (5.2) and by Lemma 5.1 ii),

$$\#\{i \in (\Lambda_1 \cup \dots \cup \Lambda_{j-1} \backslash J_j) \mid \varphi_{j-1}(i) = j'\} = \#\{i \mid \varphi_j(i) = j'\} = t_{j'}$$

for $j' = 1, \dots, j-1$. Thus

$$S = \sum_{i=1}^{r} M_i = \sum_{j'=1}^{j-1} t_{j'}(N_{j'} - 1) + \Big(r - \sum_{j'=1}^{j-1} t_{j'}\Big) N. \tag{5.7}$$

By (5.5), (5.6) and (5.7), $S \ge (1+\varepsilon)(N+dn)+c_0$ and the Dirichlet exponent ϱ of Lemma 3.6 satisfies

$$\varrho = \frac{N + dn}{S - N - dn - c_0} \le \frac{1}{\varepsilon}.$$

By that lemma (if c is a sufficiently large constant) there exist rational functions $A_1, \dots, A_r$ not all zero and of height $\le nK$ such that

$$\begin{cases} A_i \in L((N_{\varphi_{j-1}(i)} - 1)Q), & \text{if } i \in (\Lambda_1 \cup \dots \cup \Lambda_{j-1} \backslash J_j); \\ A_i \in L(NQ), & \text{if } i \notin \Lambda_1 \cup \dots \cup \Lambda_{j-1} \backslash J_j \end{cases}$$

satisfying

$$A_1 f_1^n + \ldots + A_r f_r^n = 0.$$

Not all $A_1, \dots, A_r$ are zero, and we see that the non-zero ones among $A_1 f_1^n, \dots, A_r f_r^n$ sum up to zero, and so they are linearly dependent. Choosing Λ as the set of i such that $A_i \ne 0$ we see that (5.1) is satisfied. By minimality of N_j we have $N_j \le N$. Since N is the smallest integer satisfying (5.5) and since $\varepsilon = c/K$, we have

$$\sum_{j'=1}^{j-1} t_{j'} N_{j'} + \Big(r - 1 - \sum_{j'=1}^{j-1} t_{j'}\Big) N_j \le nd + O(n/K).$$

as required. This conclude the proof of assertion iv) of Claim 4.5.

Proof of Claim 4.5 v). For $j = 1$ assertion v) was already proved in (4.9). Let $j \ge 2$. As we did for $j = 1$, we apply Lemma 3.11 to find a suitable basis of V_{Λ_j}. Let, for $i \in \Lambda_j$,

$$M_i' = \begin{cases} N_{\varphi_{j-1}(i)} - 1, & \text{if } i \in (\Lambda_1 \cup \dots \cup \Lambda_{j-1} \backslash J_j) \cap \Lambda_j, \\ N_j, & \text{if } i \in \Lambda_j \text{ and } i \notin \Lambda_1 \cup \dots \cup \Lambda_{j-1} \backslash J_j \end{cases}$$

and $M' = \max_{i\in\Lambda_j} M'_i = N_j$. Let also, as in Lemma 3.11,

$$\Theta = \max\Big(1, \sum_{i\in\Lambda_j} M'_i - (M' + nd_j)\Big) = \max\Big(1, \sum_{i\in\Lambda_j} M'_i - (N_j + nd_j)\Big)$$

with $d_j = d_{\Lambda_j} = -\deg\operatorname{div}(f_i)_{i\in\Lambda_j}$. We contend that $\Theta = O(n/K)$. As for $j = 1$, we prove this assertion using Lemma 3.6, now for the f_i ($i \in \Lambda_j$) with M_i the same as M'_i except for $N_j - 1$ in place of N_j, *i. e.*

$$M_i = \begin{cases} N_{\varphi_{j-1}(i)} - 1, & \text{if } i \in (\Lambda_1 \cup \cdots \cup \Lambda_{j-1}\backslash J_j) \cap \Lambda_j, \\ N_j - 1, & \text{if } i \in \Lambda_j \text{ and } i \notin \Lambda_1 \cup \cdots \cup \Lambda_{j-1}\backslash J_j. \end{cases}$$

Let $S = \sum_{i\in\Lambda_j} M_i = \sum_{i\in\Lambda_j} M'_i + O(1)$ and $M = \max_{i\in\Lambda_j} M_i = N_j - 1 = O(n)$. The Dirichlet exponent ϱ is then

$$\varrho = \frac{N_j - 1 + nd_j}{S - N_j - nd_j - c_0 + 1}.$$

By Lemma 3.6, there exist rational functions B_i not all zero such that $B_i \in L(M_iQ)$ for $i \in \Lambda_j$, $\sum_{i\in\Lambda_j} B_i f_i^n = 0$ and

$$h(B_i) = O((\varrho + 1)n).$$

By the minimality of N_j, we cannot have $\max_i h(B_i) \le nK$. Thus $\varrho \ge K/c$, where c is a sufficiently large constant to kill the implicit constant in the last $O()$. This implies

$$\sum_{i\in\Lambda_j} M'_i - (N_j + nd_j) \le \frac{c}{K}(N_j + nd_j) + O(1) = O(n/K)$$

as required.

We apply Lemma 3.11 to the rational functions $\{A_i^{(j)}\}_{i\in\Lambda_j}$, taking into account:

$$h(P) \ge 1, \quad h(A_i) \le nK, \quad M'_i = N_j = O(n), \quad \Theta = O(n/K).$$

By this lemma, there exists a basis $\mathbf{v}_1^{(j)}, \ldots, \mathbf{v}_{l_j-1}^{(j)}$ of V_{Λ_j} satisfying

$$\begin{aligned} h(\mathbf{v}_i^{(j)}) &\le M'h(P) + O\Big(\Theta(h(P)+1) + (n+M')(1+h(P)^{1/2}) + \max h(A_i)\Big) \\ &= N_j h(P) + O\Big(\frac{n}{K}h(P) + nh(P)^{1/2} + nK\Big). \end{aligned}$$

This proves assertion v) of Claim 4.5.

6. Proof of Theorem 1.2

Let $\Gamma \subset \mathbb{G}_{\mathrm{m}}^r(\mathbb{F})$ be a finitely generated constant-free subgroup, and let V be an algebraic subvariety of $\mathbb{G}_{\mathrm{m}}^r$, defined over $\mathbb{F}$. By writing V as an intersection of hypersurfaces, we see that it is enough to prove Theorem 1.2 and its addendum for a hypersurface V. We may further assume that V is a hyperplane. The case of a general hypersurface can indeed be easily deduced using an isogeny. For technical reasons, it is convenient to homogenize the statement in the linear case:

Proposition 6.1. *Let $\Gamma \subset \mathbb{G}_{\mathrm{m}}^r(\mathbb{F})$ be a finitely generated subgroup such that*

$$\forall \mathbf{f} \in \Gamma, \quad \forall i,j = 1,\dots,r, \quad f_i/f_j \in \overline{\mathbb{Q}} \Longrightarrow f_i/f_j \in \overline{\mathbb{Q}}^*_{\mathrm{tors}}. \tag{6.1}$$

Let $\boldsymbol{\gamma} = (\gamma_1,\dots,\gamma_r) \in \Gamma$ and $\theta_1,\dots,\theta_r \in \mathbb{F}$ such that $\theta_1\gamma_1 + \cdots + \theta_r\gamma_r \neq 0$. Then the height of $P \in \mathcal{C}(\overline{\mathbb{Q}})$ such that

$$\theta_1(P)\gamma_1(P) + \cdots + \theta_r(P)\gamma_r(P) = 0 \tag{6.2}$$

is bounded from above in terms only on Γ and $\theta_1,\dots,\theta_r$, independently of $\gamma_1,\dots,\gamma_r$. The same conclusion holds without the assumption (6.1)*, if $\theta_1,\dots,\theta_r \in \overline{\mathbb{Q}}$ and $\Gamma/\Gamma \cap \mathbb{G}_{\mathrm{m}}^r(\overline{\mathbb{Q}})$ is of rank* 1*.*

Proof. Replacing $\{1,\dots,r\}$ by a subset (and Γ by its projection on the coordinates in the subset) we may assume that there are no proper vanishing subsums in (6.2). Dividing (6.2) by $\gamma_1(P)$ (and replacing Γ by $\{(1, f_2/f_1,\dots,f_r/f_1) \mid \mathbf{f} \in \Gamma\}$) we may also assume $\gamma_1 = 1$ and

$$\Gamma \subseteq \{x_1 = 1\}\,. \tag{6.3}$$

Assume first that $\Gamma/\Gamma \cap \mathbb{G}_{\mathrm{m}}^r(\overline{\mathbb{Q}})$ is of rank 1 and $\theta_i \in \overline{\mathbb{Q}}$. Thus there exists $\mathbf{f} = (1, f_2,\dots,f_r) \in \Gamma$ such that $\Gamma = \Gamma \cap \mathbb{G}_{\mathrm{m}}^r(\overline{\mathbb{Q}}) \oplus \langle \mathbf{f} \rangle$. Then $\boldsymbol{\gamma} = (1, c_2 f_2^n,\dots,c_r f_r^n)$ for some $c_i \in \overline{\mathbb{Q}}$. We may assume that there exists some $i > 1$ such that f_i is non constant, since otherwise $\theta_1\gamma_1 + \cdots + \theta_r\gamma_r$ is a non zero constant and equation (6.2) does not have solutions. Thus we can apply Theorem 1.4, and we find that the solutions of (6.2) have bounded height.

Assume now that Γ satisfies (6.1). Then $\Gamma = \Gamma_{\mathrm{tors}} \oplus \Gamma'$ where Γ' is freely generated by, say, $\mathbf{g}_1,\dots,\mathbf{g}_\kappa$. We now use Dirichlet's Theorem in a way which is inspired by a method appearing already in Bombieri's paper [4], especially Lemma 4 therein. There exist $\boldsymbol{\omega} \in \Gamma_{\mathrm{tors}}$ and integers $\lambda_1,\dots,\lambda_\kappa$ such that

$$\boldsymbol{\gamma} = \boldsymbol{\omega}\mathbf{g}_1^{\lambda_1}\cdots\mathbf{g}_\kappa^{\lambda_\kappa}.$$

Let $A = \max|\lambda_j|$. Since the equation (6.2) is not trivial (and since Γ_{tors} is finite), we can obviously assume A unbounded. Let $Q \geq 1$ be an integer which will be fixed later, independently of A. By Dirichlet's Theorem on simultaneous approximation, there exists a positive integer $q \leq Q^\kappa$ and integers p_j such that

$$\left| q\frac{\lambda_j}{A} - p_j \right| < \frac{1}{Q}.$$

Let n be the integral part of A/q. We write $\lambda_j = np_j + r_j$ and we set, for $i = 1,\dots,r$,

$$\boldsymbol{\rho} = \prod_{j=1}^{\kappa} \mathbf{g}_j^{r_j} \in \Gamma, \qquad \mathbf{f} = \prod_{j=1}^{\kappa} \mathbf{g}_j^{p_j} \in \Gamma, \qquad \alpha_i = \omega_i\theta_i(P)\rho_i(P).$$

Since $\gamma_i = \omega_i\rho_i f_i^n$, equation (6.2) can be rewritten as $\alpha_1 f_1(P)^n + \cdots + \alpha_r f_r(P)^n = 0$ (without proper vanishing subsums).

We have

$$|p_j| \leq \left| q\frac{\lambda_j}{A} \right| + \frac{1}{Q} \leq q + Q^{-1} \leq 2Q^\kappa.$$

Thus, $\mathbf{f}$ belongs to a finite set, depending only on the generators $\mathbf{g}_1, \dots, \mathbf{g}_\kappa$ and on Q, as the exponents λ_j vary. Moreover

$$|r_j| = \left| \frac{A}{q} \left(q\frac{\lambda_j}{A} - p_j \right) - \left(n - \frac{A}{q} \right) p_j \right| \leq (n+1)Q^{-1} + 2Q^\kappa.$$

Thus $d(\theta_i \rho_i) = O(n/Q + Q^{r\kappa})$, where the implicit constant in the big-O depends only on $\theta_1, \dots, \theta_r$ and on $\mathbf{g}_1, \dots, \mathbf{g}_\kappa$.
We recall that $\alpha_i = \omega_i \theta_i(P)\rho_i(P)$, with ω_i root of unity. Using Lemma 3.3 (for the two functions 1 and $\theta_i\rho_i$), we see easily that

$$h(\boldsymbol{\alpha}) \leq h(\alpha_1) + \cdots + h(\alpha_r) \leq C_1(n/Q + Q^\kappa)h(P)$$

where the constant C_1 depends only on $\theta_1, \dots, \theta_r$ and on $\mathbf{g}_1, \dots, \mathbf{g}_\kappa$ (and neither on Q nor on A). We now choose $Q = [4rC_1] + 1$.

There is a j such that $\lambda_j = \pm A$ and thus some of the p_j are not zero. Since $g_1, \dots, g_\kappa$ is a basis of Γ' we have $\mathbf{f} \notin \Gamma_{\text{tors}}$. By (6.3) we have $f_1 = 1$. Thus, by (6.1), $f_i = f_i/f_1$ is non-constant for some $i > 1$ and we can apply Theorem 1.4.

Let C be the constant appearing in this theorem, which depends on $f_1, \dots, f_r$ and thus only on $\theta_1, \dots, \theta_r$ and on $\mathbf{g}_1, \dots, \mathbf{g}_\kappa$ (since the rational functions f_i belong to a finite set, depending only on the generators $\mathbf{g}_1, \dots, \mathbf{g}_\kappa$ and on Q and Q has already been fixed as $Q = [4rC_1] + 1$). Since A is unbounded and Q does not depend on A, we may assume $n = [A/q] \geq C$ and $n \geq 4rC_1Q^\kappa$. By Theorem 1.4 and by the inequalities $h(\boldsymbol{\alpha}) \leq C_1(n/Q + Q^\kappa)h(P)$, $Q \geq 4rC_1$ and $n \geq 4rC_1Q^\kappa$,

$$h(P) \leq \frac{rh(\boldsymbol{\alpha})}{n} + C \leq rC_1(1/Q + Q^\kappa/n)h(P) + C \leq h(P)/2 + C\ .$$

We deduce that $h(P)$ is bounded, as claimed.

□

References

[1] M. Bays and P. Habegger, "A note on divisible points of curves". *Trans. Amer. Math. Soc.* **367** (2015), 1313–1328.

[2] F. Beukers, "On a sequence of polynomials." In Algorithms for algebra. *J. Pure Appl. Algebra* **117-118** (1997), 97–103.

[3] F. Beukers and R. Tijdeman, "On the multiplicities of binary complex recurrences". *Compositio Math.* **51** (1984), no. 2, 193–213.

[4] E. Bombieri, "Effective Diophantine approximation on $\mathbb{G}_{\mathrm{m}}$", *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **20** (1993), 61–89.

[5] E. Bombieri and P. B. Cohen, Siegel's lemma, Padé approximations and Jacobians, *Ann. Sc. Normale Sup. Cl. Sci. (4)* **25** (1997), 155–178.

[6] E. Bombieri and W. Gubler, "Heights in diophantine geometry", New Math. Monographs **4**, Cambridge University Press 2006.

[7] E. Bombieri, D. Masser, and U. Zannier, "Intersecting a curve with algebraic subgroups of multiplicative groups", *Internat. Math. Res. Notices* **1999**, no. 20, 1119–1140.

[8] E. Bombieri, D. Masser, and U. Zannier, "Anomalous Subvarieties - Structure Theorems and Applications", *Internat. Math. Res. Notices* **2007**, no. 19, 1–33.

[9] E. Bombieri and J. Vaaler, "On Siegel's lemma", *Invent. Math.*, **73** (1983), 11-32.

[10] W. D. Brownawell and D. W. Masser, "Vanishing sums in function fields". *Math. Proc. Cambridge Philos. Soc.* **100** (1986), no. 3, 427–434.

[11] S. David and P. Philippon, Minorations des hauteurs normalisées des sous-variétés des tores, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* **28** (1999), 489–543.

[12] A. Denz, "Bounding the height of certain algebraic numbers", Master Thesis, University of Basle 2016.

[13] P. Habegger, "On the Bounded Height Conjecture", *Internat. Math. Res. Notices 2009*, no. 5, 860–886.

[14] M. Hindry and J. H. Silverman, "Diophantine geometry. An introduction". GTM 201. Springer-Verlag, Berlin, 2000.

[15] S. Lang, "Introduction to Algebraic and Abelian Functions", Springer-Verlag, GTM **89**, 1982.

[16] D. Masser and U. Zannier, Torsion points on families of simple abelian surfaces and Pell's equation over polynomial rings, with an appendix by V. Flynn, *J. European Math. Soc.* **17** (2015), 2379–2416.

[17] D. Mirimanov, "Sur l'équation $(x+1)^l - x^l - 1 = 0$". *Nouv. Ann. Math.* **3** (1903), 385–397.

[18] A. Néron, "Quasi-fonctions et hauteurs sur les variétés abéliennes", *Ann. of Math. (2)* **82** (1965), 249–331.

[19] G. Rémond, "Inégalité de Vojta en dimension supérieure", *Ann. Sc. Normale Sup. Cl. Sci. (4)* **29** (2000), 101–151.

[20] G. Rémond, "Sur les sous-variétés des tores", *Compositio Math.* **134** (2002), no. 3, 337–366.

[21] W. M. Schmidt, "Eisenstein's theorem on power series expansions of algebraic functions", *Acta Arith.* **56** (1990), 161–179.

[22] W. M. Schmidt, "Diophantine approximations and Diophantine equations". LNM 1467, Springer-Verlag, Berlin, 1991.

[23] J-P. Serre, Lectures on the Mordell-Weil Theorem, 2-nd Ed., Vieweg, 1990.

[24] J. H. Silverman, "Heights and the specialization map for families of abelian varieties", *J. Reine Angew. Math.* **342** (1983), 197–211.

[25] J. H. Silverman and F. Voloch, "Multiple Weierstrass Points", *Compositio Math.*, **79** (1991), 123–134.

[26] E. Thomas, "Solutions to certain families of Thue equations", *J. Number Theory* **43** (1993), 319–369.

[27] U. Zannier, "Some problems of unlikely intersections in arithmetic and geometry", *Annals of Mathematics Studies*, vol. **181**, Princeton University Press, 2012, With appendixes by David Masser.

[28] S. Zhang, "Positive line bundles on arithmetic surfaces", *Ann. Math. (2)* **136** (1992), no. 3, 569–587.