A BOGOMOLOV TYPE STATEMENT FOR FUNCTION FIELDS

DRAGOS GHIOCA

Department of Mathematics, University of British Columbia, Vancouver, B.C. V6T 1Z2, Canada.

E-mail: dghioca@math.ubc.ca

Abstract

Let k be a an algebraically closed field of arbitrary characteristic, and we let $h : \mathbb{A}^n(\overline{k(t)}) \longrightarrow \mathbb{R}_{\geq 0}$ be the usual Weil height for the n-dimensional affine space corresponding to the function field $k(t)$ (extended to its algebraic closure). We prove that for any affine variety $V \subset \mathbb{A}^n$ defined over $\overline{k(t)}$, there exists a positive real number $\epsilon := \epsilon(V)$ such that if $P \in V(\overline{k(t)})$ and $h(P) < \epsilon$, then $P \in V(k)$.

1. Introduction

In a paper [6] from 1965, Lang asks the following question: what are the plane irreducible curves C which contain infinitely many points (x, y) where both x and y are roots of unity? It is easy to see that if C is the zero set of an equation of the form $x^m y^n = \zeta$, where $m, n \in \mathbb{Z}$ and ζ is a root of unity, then indeed C contains infinitely many points with both coordinates roots of unity. An old theorem of Ihara-Serre-Tate-Lang says that indeed C must have the above form. Essentially they prove that if C contains infinitely many points where both coordinates are roots of unity, then C must be a (multiplicative) translate of a 1-dimensional torus by a point with both coordinates roots of unity. This result can be extended to higher dimensional varieties, and even to subvarieties of abelian varieties (the latter was formerly known as the Manin-Mumford Conjecture, proven by Raynaud [8]). The following theorem is due to Laurent [7]; we write it in

the language of algebraic groups, more precisely for $\mathbb{G}_m^N$, which is the N-th cartesian power of the multiplicative group.

Theorem 1.1 (Laurent). *Let $V \subset \mathbb{G}_m^N$ defined over $\mathbb{C}$ be an irreducible affine variety which contains a Zariski dense set of torsion points (i.e., points with coordinates roots of unity). Then V is a multiplicative translate of a torus by a torsion point.*

The roots of unity are the points of $\overline{\mathbb{Q}}^*$ which have (naive) Weil height equal to 0. The Weil height of $x \in \overline{\mathbb{Q}}$ is defined as follows (the set $\Omega_{\mathbb{Q}}$ stands for all inequivalent absolute values on $\mathbb{Q}$)

$$\mathrm{h}(x) := \frac{1}{[\mathbb{Q}(x) : \mathbb{Q}]} \cdot \sum_{\sigma : \mathbb{Q}(x) \hookrightarrow \overline{\mathbb{Q}}} \sum_{v \in \Omega_{\mathbb{Q}}} \log^+ |\sigma(x)|_v,$$

where $\log^+(z) := \log \max\{1, z\}$ for any real number z, and for each absolute value of $\mathbb{Q}$, we fix an extension of it to $\overline{\mathbb{Q}}$. Similarly, for any $x_1, \ldots, x_N$ contained in a number field L we define

$$\mathrm{h}((x_1, \ldots, x_N)) := \frac{1}{[L : \mathbb{Q}]} \cdot \sum_{\sigma : L \hookrightarrow \overline{\mathbb{Q}}} \sum_{v \in \Omega_{\mathbb{Q}}} \log \max\{1, |\sigma(x_1)|_v, \ldots, |\sigma(x_N)|_v\}.$$

So, Laurent's result yields that if $V \subset \mathbb{G}_m^N$ contains a Zariski dense subset of points of height equal to 0, then V is a torsion translate of an algebraic subgroup of $\mathbb{G}_m^N$.

The same conclusion holds if one weakens the hypothesis and only asks that V contains a Zariski dense set of points of *small* height; this was initially known as the Bogomolov Conjecture. So, for each $\epsilon \geq 0$, let

$$S_\epsilon := \left\{ P \in \left(\overline{\mathbb{Q}}^*\right)^N \ : \ h(P) \leq \epsilon \right\}.$$

Conjecture 1.2 (Bogomolov). *Let $V \subset \mathbb{G}_m^N$ be an irreducible subvariety (defined over $\overline{\mathbb{Q}}$) such that for each $\epsilon > 0$, we have that $V(\overline{\mathbb{Q}}) \cap S_\epsilon$ is Zariski dense in V. Then V is a torsion translate of an algebraic subgroup of $\mathbb{G}_m^N$.*

The Bogomolov conjecture in the context of abelian varieties was proven by Ullmo [11] for curves V embedded in their Jacobians, and in the general case of any subvariety V of an abelian variety A by Zhang [13]. Both

Ullmo and Zhang proved the Bogomolov conjecture via an equidistribution statement for points of small height on A. An extension of the Bogomolov property to semiabelian varieties was obtained by David and Philippon [3].

The case of Bogomolov conjecture for any power of the multiplicative group (see Conjecture 1.2) was first proved by Zhang in [12]. Other proofs of the Bogomolov conjecture for $\mathbb{G}_m^n$ were given by Bilu [1] and Bombieri and Zannier [2]. Similarly to Ullmo's and Zhang's proofs of the Bogomolov Conjecture for abelian varieties, Bilu proved that the probability measures supported on Galois orbits of *generic* algebraic points of height tending to 0 converge weakly to the Lebesgue measure μ on $\mathcal{C}^N$, where $\mathcal{C}$ is the complex unit circle. More precisely, if $\{P_n\}_{n\geq 1} \subset \left(\overline{\mathbb{Q}}^*\right)^N$ is a sequence of points with the property that no proper algebraic subgroup of $\mathbb{G}_m^N$ contains infinitely many P_n's, then for each continuous function f, we have

$$\lim_{n\to\infty} \frac{1}{[\mathbb{Q}(P_n):\mathbb{Q}]} \cdot \sum_{\sigma\in\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})} f\left(P_n^\sigma\right) = \int_{\mathcal{C}^N} f d\mu.$$

On the other hand, Bombieri and Zannier's proof [2] of Conjecture 1.2 followed a different path. In [2, Lemma 1], Bombieri and Zannier show that that for any polynomial $f \in \mathbb{Z}[X_1,\ldots,X_N]$, and any point $(x_1,\ldots,x_N) \in \overline{\mathbb{Q}}$, and for any prime number p sufficiently large, there exists a positive real number $\epsilon(p)$ such that either $\mathrm{h}((x_1,\ldots,x_N)) \geq \epsilon(p)$, or $f(x_1^p,\ldots,x_N^p) = 0$. This result allows the authors of [2] to conclude that either the points on the hypersurface $Z(f)$ have a height larger than some absolute positive lower bound, or the hypersurface is invariant under the endomorphism of $\mathbb{G}_m^N$ given by $(X_1,\ldots,X_N) \mapsto (X_1^p,\ldots,X_N^p)$. In the latter case, one can see that this means $Z(f)$ is a finite union of torsion translates of subtori of $\mathbb{G}_m^N$.

The approach of Bombieri and Zannier from [2] inspired the author to extend their [2, Lemma 1] in positive characteristic by applying the Frobenius map to affine subvarieties of $\mathbb{A}^N$ defined over $\overline{\mathbb{F}_p(t)}$. This allowed the author to obtain in [5] a Bogomolov type statement for affine varieties defined over $\overline{\mathbb{F}_p(t)}$. The picture in positive characteristic for the Bogomolov conjecture is much different due to the varieties defined over finite fields. Indeed, if $V \subset \mathbb{G}_m^N$ is any subvariety defined over $\overline{\mathbb{F}_p}$, then V contains a Zariski dense set of points of height 0 (all its points with coordinates in $\overline{\mathbb{F}_p}$). So, it is no longer true that only torsion translates of subtori of $\mathbb{G}_m^N$ contain a Zariski dense set

of points with small height; any *constant* subvariety has this property as well. The group structure of the ambient space $\mathbb{G}_m^n$ disappears from the conclusion of a Bogomolov statement for $\mathbb{G}_m^n$; this motivated our approach from [5] in which the ambient space is simply the affine space, and not an algebraic torus as in [2].

Motivated by a question of Zinovy Reichstein, we consider in this paper the same problem with respect to the height constructed with respect to a function field K/k of arbitrary characteristic. So, let k be an algebraically closed field of arbitrary characteristic, and let $\Omega_{k(t)}$ be all the inequivalent absolute values on $k(t)$. Each $v \in \Omega_{k(t)}$ corresponds either to the place at infinity v_∞, i.e.

$$v_\infty\left(\frac{f}{g}\right) := \deg(g) - \deg(f),$$

for nonzero $f, g \in k[t]$, or to a point $\alpha \in k$, i.e.,

$$v_\alpha\left(\frac{f}{g}\right) := \mathrm{ord}_{t-\alpha}\left(\frac{f}{g}\right),$$

where $\mathrm{ord}_{t-\alpha}(f/g)$ is the order of vanishing at α of the rational function f/g.

For each finite extension K of $k(t)$, we let Ω_K be the set of all (inequivalent) places of K which lie above the places of $k(t)$. We normalize each (exponential) valuation $w \in \Omega_K$ so that the function $w : K \longrightarrow \mathbb{Z}$ is surjective. In other words, for each nonzero $x \in k(t)$, and for each place $w \in \Omega_K$ lying above a place $v \in \Omega_{k(t)}$ we have

$$w(x) = e(w \mid v) \cdot v(x),$$

where $e(w \mid v)$ is the ramification index for $w \mid v$. Then for each $x \in K$ we define its height:

$$\mathrm{h}(x) := \frac{1}{[K : k(t)]} \cdot \sum_{w \in \Omega_K} \max\{0, -w(x)\}.$$

We note that the above definition is independent of the choice of field K containing x because the places of a function field are *coherent*, i.e., for each

finite extensions $k(t) \subset K \subset L$, for each $v \in \Omega_K$, and for each nonzero $x \in K$ we have

$$v(x) = \frac{1}{[L:K]} \cdot \sum_{\substack{w \in \Omega_L \\ w|v}} w(x). \tag{1.1}$$

With the notation from [10], the above condition is that v is *defectless*; this follows from the arguments of [10, Chapter 1, Section 4] (Hypothesis (F) holds for algebras of finite type over fields and so, it holds for localizations of such algebras; for each $v \in \Omega_K$ we apply [10, Propositions 10 and 11] to the local ring of v).

Similarly, we define for any $n \in \mathbb{N}$, the height of $(x_1, \ldots, x_n) \in \mathbb{A}^n(K)$ be

$$\mathrm{h}((x_1, \ldots, x_n)) := \frac{1}{[K:k(t)]} \cdot \sum_{w \in \Omega_K} \max\{0, -w(x_1), \cdots - w(x_n)\}.$$

For each $\epsilon \geq 0$ we let

$$S_\epsilon := \{P \in \mathbb{A}^n(\overline{k(t)}) : \mathrm{h}(P) \leq \epsilon\}.$$

Then our main result is the following.

Theorem 1.3. *Let $V \subset \mathbb{A}^n$ be an affine subvariety defined over $\overline{k(t)}$. Let $W \subseteq V$ be the Zariski closure of $V(k)$. Then there exists $\epsilon > 0$ such that for all $P \in (V \setminus W)(\overline{k(t)})$, we have $\mathrm{h}(P) \geq \epsilon$.*

The following result is an alternative reformulation.

Theorem 1.4. *Let $V \subset \mathbb{A}^n$ be an affine subvariety defined over $\overline{k(t)}$. If for each $\epsilon > 0$, the subset $V(\overline{k(t)}) \cap S_\epsilon$ is Zariski dense in V, then V is defined over k.*

Indeed, Theorem 1.4 follows immediately from Theorem 1.3 since then (by Theorem 1.3) there exists $\epsilon > 0$ such that $h(P) \geq \epsilon$ for all $P \in (V \setminus W)(\overline{k(t)})$. But if V is not defined over k, then W is a proper subvariety of V, and therefore the Zariski closure of $V(\overline{k(t)}) \cap S_{\frac{\epsilon}{2}}$ is contained in W; thus Theorem 1.4 holds. Conversely, using Theorem 1.4 yields that if V is not defined over k then there exists $\epsilon_1 > 0$ such that

$$V_1 := \overline{V(\overline{k(t)}) \cap S_{\epsilon_1}} \neq V.$$

Now, if V_1 is defined over k, then Theorem 1.3 is proven. Assume V_1 is not defined over k; then again applying Theorem 1.4 there exists a positive ϵ_2 (less than ϵ_1) such that

$$V_2 := \overline{V_1(\overline{k(t)}) \cap S_{\epsilon_2}} \neq V_1.$$

Now if V_2 is defined over k, then Theorem 1.3 holds with $\epsilon := \epsilon_2$. If not, then we continue the above process. At one moment this process must end with finding a subvariety V_m defined over k, and then Theorem 1.3 holds with $\epsilon := \epsilon_m$. Otherwise we would have constructed an infinite descending chain of varieties which contradicts the Noetherian property of the Zariski topology; so indeed the above theorems are equivalent.

Remark 1.5. The result of Theorem 1.3 (and its reformulation) extends to any closed projective subvariety V of a projective space $\mathbb{P}^n$. Indeed, we cover $\mathbb{P}^n$ by finitely many open affine spaces $\{U_i\}_i$, and then apply Theorem 1.3 to each $V \cap U_i$ (which is a closed subvariety of the affine space U_i).

We prove Theorem 1.3 using the same strategy employed in [5], only that this time we replace the Frobenius endomorphism by a suitable automorphism σ of $\overline{k(t)}$. We show that for any point $P \in V(\overline{k(t)})$, either $\mathrm{h}(P)$ is uniformly bounded from below away from 0, or $P^\sigma \in V(\overline{k(t)})$. If the latter occurs generically, then V is invariant under σ and therefore it is defined over the fixed field of σ, which is k. There are technical difficulties in our setting, compared to [5], since our automorphism σ has a more complicated action on $\overline{k(t)}$ compared to the action on $\overline{\mathbb{F}_p(t)}$ by the classical Frobenius.

Theorem 1.3 yields a similar result for an arbitrary (finite) transcendence degree function field. Indeed, let k be an algebraically closed field, and let K/k be a finite transcendence degree function field. Let $t_1, \ldots, t_r \in K$ be algebraically independent elements such that $K/k(t_1, \ldots, t_r)$ is a finite extension. For each $i = 1, \ldots, r$ we let K_i be the algebraic closure of $k(t_1, \ldots, t_i)$ in K. Then $K_r = K$; we also let $K_0 := k$. Let $V \subset \mathbb{A}^n$ be an affine variety defined over K, and assume it is not defined over k (otherwise Theorem 1.3 and its consequences hold trivially). Then there exists a smallest (positive) integer i such that V is defined over K_i (but it is not defined over K_{i-1}). We let h be the Weil height constructed with respect to the function field K_i/K_{i-1} (which is a function field of transcendence degree equal to 1). Then Theorem 1.3 applied to $\overline{K_{i-1}}K_i/\overline{K_{i-1}}$ yields that there exists a positive real

number $\epsilon := \epsilon(V)$ such that if $P \in V(\overline{K_i})$ and $h(P) < \epsilon$, then $P \in W(\overline{K_i})$, where W is the largest subvariety of V defined over $\overline{K_{i-1}}$.

Acknowledgments

We thank the Institute of Mathematics Academia Sinica for its hospitality, and we thank the organizers Liang-Chung Hsia and Tzu-Yueh Julie Wang of the conference on "Diophantine Problems and Arithmetic Dynamics" from Taipei for creating a stimulating research environment where the results of this paper were disseminated. We also thank the referee and Liang-Chung Hsia for their useful comments.

2. Proof of Our Main Result

Our proof follows the strategy from [5]; when the proof is identical with the one from [5] we refer to our earlier paper, otherwise we present the argument entirely. Unless otherwise stated, all our subvarieties are closed; we continue with our notation from Theorem 1.3. We start with a definition.

Definition 2.1. We call *reduced* a non-constant polynomial

$$f \in k[t][X_1, \ldots, X_n],$$

whose coefficients a_i have no non-constant common divisor in $k[t]$. We define the *height* $h(f)$ of the polynomial f as the maximum of the degrees of the coefficients $a_i \in k[t]$ of f.

For some integer $M > 1$, let $\sigma := \sigma_M$ be an automorphism of $\overline{k(t)}$ which fixes the elements of k, and maps t into t^M.

Lemma 2.2. *For each $x \in \overline{k(t)}$, we have $h(\sigma(x)) = M \cdot h(x)$.*

Proof. Let

$$L := \bigcup_{n \geq 1} k\left(t^{\frac{1}{M^n}}\right).$$

We claim that σ restricts to an automorphism of L. First of all, it is clear that for each positive integer n, there exists an M^n-th root of unity $\zeta_n \in k$ such that $\sigma(t^{1/M^n}) = \zeta_n t^{1/M^{n-1}}$. So, $\sigma(\zeta_n^{-1} t^{1/M^n}) = t^{1/M^{n-1}}$ showing that

indeed σ restricts to an automorphism of L (note that k is algebraically closed).

Clearly, we may assume $x \neq 0$. Let $f \in L[z]$ be a nonzero polynomial of minimal degree such that $f(x) = 0$. Since f has finitely many coefficients (say, $\deg(f) = d \geq 1$), then there exists $N \in \mathbb{N}$ such that f has all its coefficients in $k\left(t^{1/M^N}\right)$. For the sake of simplifying the notation, we let $T := t^{1/M^N}$. Since $k[T]$ is a PID, we may assume $f \in k[T][z]$ and moreover, the coefficients of f are all relatively prime. Furthermore, f is irreducible in $k[T][z]$; also let $D := \mathrm{h}_T(f)$ be the maximum of the degrees (in T) of the coefficients of f. So, applying [4, Lemma 2.1], we conclude that

$$\mathrm{h}(x) = \frac{D}{dM^N}.$$

An observation regarding our formula above and [4, Lemma 2.1]: because our height is defined relative to $k(t)$, while in [4, Lemma 2.1] the height is computed relative to the field $k(T) = k\left(t^{1/M^N}\right)$, the factor M^N appears in the denominator of our formula.

On the other hand, we claim that $f^\sigma \in L[z]$ is also irreducible, where f^σ is the polynomial obtained by applying σ to each coefficient of f. Indeed, if f^σ were reducible over L, then there exist nonconstant polynomials $g, h \in L[z]$ such that $f^\sigma = g \cdot h$. But then $f = g^{\sigma^{-1}} \cdot h^{\sigma^{-1}}$, and $g^{\sigma^{-1}}, h^{\sigma^{-1}} \in L[z]$ which thus contradicts the hypothesis that f is irreducible in $L[z]$. Moreover, the coefficients of f^σ are relatively prime. Indeed, because the coefficients $\{a_i\}_{0 \leq i \leq d}$ of f are relatively prime there exist $b_i \in k[T]$ such that $\sum_{i=0}^{d} a_i b_i = 1$, and so, $\sum_{i=0}^{d} \sigma(a_i)\sigma(b_i) = 1$ showing that also the coefficients $\sigma(a_i)$ of f^σ are relatively prime. Hence, applying again [4, Lemma 2.1] we conclude that

$$\mathrm{h}(\sigma(x)) = \frac{M \cdot D}{dM^N},$$

since $\deg_T(\sigma(a_i)) = M \cdot \deg_T(a_i)$ for each i. Thus $\mathrm{h}(\sigma(x)) = M \cdot \mathrm{h}(x)$, as desired. $\qquad \square$

Corollary 2.3. *For each $x_1, \ldots, x_n \in \overline{k(t)}$ we have*

$$\mathrm{h}((\sigma(x_1), \ldots, \sigma(x_n))) \leq nM \cdot \mathrm{h}((x_1, \ldots, x_n)).$$

Proof. Let K be a finite extension of $k(t)$ containing each x_i and each $\sigma(x_i)$. Using Lemma 2.2, we obtain

$$\mathrm{h}((\sigma(x_1), \ldots, \sigma(x_n))) = \frac{1}{[K:k(t)]} \sum_{w \in \Omega_K} \max\{0, -w(\sigma(x_1)), \ldots, -w(\sigma(x_n))\}$$

$$\leq \frac{1}{[K:k(t)]} \sum_{w \in \Omega_K} \sum_{i=1}^{n} \max\{0, -w(\sigma(x_i))\}$$

$$= \sum_{i=1}^{n} \mathrm{h}(\sigma(x_i))$$

$$= M \cdot \sum_{i=1}^{n} \mathrm{h}(x_i)$$

$$= \frac{M}{[K:k(t)]} \sum_{w \in \Omega_K} \sum_{i=1}^{n} \max\{0, -w(x_i)\}$$

$$\leq \frac{Mn}{[K:k(t)]} \sum_{w \in \Omega_K} \max\{0, -w(x_1), \ldots, -w(x_n)\}$$

$$= Mn\mathrm{h}((x_1, \ldots, x_n)),$$

as desired. $\qquad\square$

The following result is also an easy corollary of Lemma 2.2.

Lemma 2.4. *The fixed field of σ is k.*

Proof. Let $x \in \overline{k(t)}$ such that $\sigma(x) = x$. Then by Lemma 2.2 we have that $\mathrm{h}(x) = \mathrm{h}(\sigma(x)) = M\mathrm{h}(x)$; so $\mathrm{h}(x) = 0$ (because $M > 1$). Therefore $x \in k$ since they are the only points in $\overline{k(t)}$ of height equal to 0. $\qquad\square$

The following result is key for our proof, and it is similar to [5, Lemma 3.2].

Lemma 2.5. *Let $f \in k[t][X_1, \ldots, X_n]$ be a reduced polynomial of total degree d. Let M be an integer satisfying $M \geq 2 \cdot \max\{1, \mathrm{h}(f)\}$, and let σ be an automorphism of $\overline{k(t)}$ such that σ restricts to the identity morphism on k, and $\sigma(t) = t^M$. If $(x_1, \ldots, x_n) \in \mathbb{A}^n_{\overline{k(t)}}$ satisfies $f(x_1, \ldots, x_n) = 0$, then either*

$$\mathrm{h}((x_1, \ldots, x_n)) \geq \frac{1}{2dn}$$

or

$$f(\sigma(x_1), \ldots, \sigma(x_n)) = 0.$$

Proof. Let $(x_1, \ldots, x_n) \in \mathbb{A}_{\overline{k(t)}}^n$ be a zero of f. We let $f = \sum_i a_i M_i$, where the a_i's are the nonzero coefficients of f and the M_i's are the corresponding monomials of f. For each i, we let $m_i := M_i(x_1, \ldots, x_n)$.

Assume $f(\sigma(x_1), \ldots, \sigma(x_n)) \neq 0$. We let K be a finite extension of $k(t)$ containing each x_i and each $\sigma(x_i)$. If $\zeta = f(\sigma(x_1), \ldots, \sigma(x_n))$, then (because $\zeta \neq 0$) the product formula for $\zeta \in K$ yields

$$\sum_{w \in M_K} w(\zeta) = 0. \tag{2.1}$$

Note that there are no extra factors in (2.1) because k is algebraically closed and thus each place has degree equal to 1.

Because $f(x_1, \ldots, x_n) = 0$, we get $\zeta = \zeta - \sigma(f(x_1, \ldots, x_n))$ and so,

$$\zeta = \sum_i (a_i - \sigma(a_i)) \cdot \sigma(m_i). \tag{2.2}$$

Claim 2.6. *For every* $g \in k[t]$, $(t^M - t) \mid (\sigma(g) - g)$.

Proof. [Proof of Claim 2.6.] Let $g := \sum_{j=0}^m b_j t^j$, with $b_j \in k$. Then $\sigma(g) = \sum_{j=0}^m b_j t^{jM}$. The proof is immediate because $(t^M - t) \mid (t^{jM} - t^j)$. $\square$

Using the result of Claim 2.6 and equation (2.2), we get

$$\zeta = (t^M - t) \cdot \sum_i b_i \sigma(m_i), \tag{2.3}$$

where $b_i = \frac{a_i - \sigma(a_i)}{t^M - t} \in k[t]$. Let S be the set of valuations $w \in M_K$ such that w lies above each place of $k(t)$ corresponding to a root of $t^M - t$. For each $w \in S$,

$$w(\zeta) \geq w(t^M - t) - d \max\{0, -w(\sigma(x_1)), \ldots, -w(\sigma(x_n))\}, \tag{2.4}$$

because for each i, $w(b_i) \geq 0$ (as $b_i \in k[t]$ and w does not lie over v_∞) and the total degree of M_i is at most d.

For each $w \in M_K \setminus S$, because $\zeta = \sum_i a_i \sigma(m_i)$, we have

$$w(\zeta) \geq -\max\{0, \max_i -w(a_i)\} - d \max\{0, -w(\sigma(x_1)), \ldots, -w(\sigma(x_n))\}. \tag{2.5}$$

Adding all inequalities from (2.4) and (2.5) we obtain

$$0 = \sum_{w \in M_K} \frac{w(\zeta)}{[K:k(t)]} \geq -\mathrm{h}(f) - d\mathrm{h}((\sigma(x_1), \ldots, \sigma(x_n))) + \sum_{\substack{w \in M_K \\ w(t^M - t) > 0}} \frac{w(t^M - t)}{[K:k(t)]}.$$

$$(2.6)$$

Indeed, note that

$$\frac{1}{[K:k(t)]} \cdot \sum_{w \in \Omega_K} \max\{0, -\max_i w(a_i)\} = h\left((a_i)_{0 \leq i \leq \deg(f)}\right)$$

$$= h(f),$$

by the definition of the Weil height on the affine space of dimension $\deg(f)+1$, and by the Definition 2.1. Now, by the coherence of the valuations on $\overline{K}$ (see (1.1)), we have

$$\sum_{\substack{w \in M_K \\ w(t^M - t) > 0}} \frac{w(t^M - t)}{[K:k(t)]} = \sum_{\substack{v \in M_{k(t)} \\ v(t^M - t) > 0}} v(t^M - t) = -v_\infty(t^M - t) = M.$$

Thus, inequality (2.6) yields

$$0 \geq -\mathrm{h}(f) - d\mathrm{h}((\sigma(x_1), \ldots, \sigma(x_n))) + M$$

and so, using Corollary 2.3 we obtain

$$dnM\mathrm{h}((x_1, \ldots, x_n)) \geq M - h(f). \qquad (2.7)$$

Because M was chosen such that $M \geq 2\mathrm{h}(f)$, we conclude that

$$\mathrm{h}((x_1, \ldots, x_n)) \geq \frac{1}{2dn}. \qquad (2.8)$$

as desired.   □

Remark 2.7. The above proof (more precisely Claim 2.6) shows why it is essential to work with a function field of transcendence degree equal to 1; otherwise we would not be able to conclude that there exists a positive contribution for the same set of valuations for all $\sigma(a) - a$, where $a \in k[B]$ (for some transcendence basis B for the function field K/k). If $|B| > 1$, say

$B = \{s, t\}$, then $\sigma(s) - s$ and $\sigma(t) - t$ are relatively prime for an automorphism σ which maps t to t^M and maps s to s^M.

The proof of the next result is identical with the proof of [5, Lemma 3.4]. Essentially, if $f \mid f^{\sigma^{-1}}$, then the hypersurface $Z = Z(f)$ is invariant under σ and thus defined over the fixed field of σ. We also include an alternative elementary proof of this result.

Lemma 2.8. *Let* $f \in \overline{k(t)}[X_1, \ldots, X_n]$ *be a nonzero polynomial, and let* σ *be an automorphism of* $\overline{k(t)}$ *fixing pointwise* k *and mapping* t *into* t^M *(for some integer* $M > 1$*). If* $f(X_1, \ldots, X_n) \mid f^{\sigma^{-1}}(X_1, \ldots, X_n)$, *then there exists a nonzero* $a \in \overline{k(t)}$ *such that* $a \cdot f \in k[X_1, \ldots, X_n]$.

Proof. Let $\tau := \sigma^{-1}$. Since f and f^τ are polynomials of same degree in each variable, then $f \mid f^\tau$ if and only if there exists a nonzero $\alpha \in \overline{k(t)}$ such that $f^\tau = \alpha \cdot f$. So, for each nonzero coefficient c_i of f, we have $\tau(c_i) = \alpha \cdot c_i$. Hence for any two coefficients c_i and c_j we obtain $\tau\left(\frac{c_j}{c_i}\right) = \frac{c_j}{c_i}$, i.e. $\frac{c_j}{c_i} \in k$ (see Lemma 2.4). Therefore Lemma 2.8 holds with $a := c_i$ (for any nonzero coefficient of f). □

The proof of the next result is similar to [5, Lemma 3.5], but there are some technical complications due to our general setting for the automorphism σ as opposed to the Frobenius endomorphism.

Lemma 2.9. *Let* $V \subset \mathbb{A}^n$ *be a proper affine* $\overline{k(t)}$*-subvariety. Then there exists a positive constant* C, *depending only on* V, *and there exists a proper affine* k*-subvariety* $Z \subset \mathbb{A}^n$, *which also depends only on* V, *such that for every* $P \in V(\overline{k(t)})$, *either* $P \in Z(\overline{k(t)})$ *or* $\mathrm{h}(P) \geq C$.

Remark 2.10. The only difference between Lemma 2.9 and Theorem 1.3 is that we do not require Z be contained in V.

Proof. [Proof of Lemma 2.9.] We proceed by induction on n. The case $n = 1$ is obvious, because any subvariety of $\mathbb{A}^1$, different from $\mathbb{A}^1$, is a finite union of points. Thus we may take $Z = V(k)$, (which is also a finite union of points) and $C := \min_{P \in (V \setminus Z)(\overline{k(t)})} \mathrm{h}(P)$ (if there are no points in $V(\overline{k(t)}) \setminus V(k)$, then we may take $C = 1$, say). We note that in this case ($n = 1$) we actually proved Theorem 1.3, because the variety Z that we chose is a subvariety of V.

We assume Lemma 2.9 holds for $n-1$ and we prove it for n ($n \geq 2$). Let K be a finite field extension of $k(t)$ (of minimal degree) such that V is defined over K. Let p^m be the inseparable degree of the extension $K/k(t)$ ($m \geq 0$). Let

$$V_1 = \bigcup_\sigma V^\sigma,$$

where σ denotes any field homomorphism $K \to \overline{k(t)}$ which fixes $k(t)$. The variety V_1 is a $k(t^{1/p^m})$-variety (note that k is algebraically closed). Also, V_1 depends only on V. Thus, if we prove Lemma 2.9 for V_1, then our result will hold also for $V \subset V_1$. Hence we may and do assume that V is defined over $k(t^{1/p^m})$.

Assume $m > 0$; then k has positive characteristic. We let F be the Frobenius corresponding to $\mathbb{F}_p$. The variety $V' = F^m V$ is a $k(t)$-variety, which depends only on V. Assume we proved Lemma 2.9 for V' and let C' and Z' be the positive constant and the k-variety, respectively, associated to V', as in the conclusion of Lemma 2.9. Let $P \in V(\overline{k(t)})$. Then $P' := F^m(P) \in V'(\overline{k(t)})$. Thus, either

$$\mathrm{h}(P') \geq C' \text{ or}$$

$$P' \in Z'(\overline{k(t)}).$$

In the former case, because $\mathrm{h}(P) = \frac{1}{p^m}\mathrm{h}(P')$, we obtain a lower bound for the height of P, depending only on V (note that m depends only on V). In the latter case, if we let Z be the k-subvariety of $\mathbb{A}^n$, obtained by extracting the p^m-th roots of the coefficients of a set of polynomials (defined over k) which generate the vanishing ideal for Z', we get $P \in Z(\overline{k(t)})$. By its construction, Z depends only on V and so, we obtain the conclusion of Lemma 2.9.

Thus, from now on in this proof, we assume V is a $k(t)$-variety. We fix a set of defining polynomials for V which contains polynomials $P_i \in k(t)[X_1, \ldots, X_n]$ for which

$$\max_i \deg(P_i)$$

is minimum among all possible sets of defining polynomials for V (where $\deg P_i$ is the total degree of P_i). If all of them have coefficients in k, then

Lemma 2.9 holds with $Z = V$ and C any positive constant. Once again, in this case, Theorem 1.3 holds.

Assume there exists a polynomial $f \notin k[X_1, \ldots, X_n]$ in the fixed set of defining equations for V. Let $L := \bigcup_{\ell \geq 1} k\left(t^{1/\ell!}\right)$, and let $\{f_i\}_i$ be the set of all the L-irreducible factors of f. For each i let H_i be the zero set of f_i. Then V is contained in the finite union $\cup_i H_i$. The polynomials f_i depend only on f. Thus it suffices to prove Lemma 2.9 for each H_i. Hence we may and do assume V is the zero set of an L-irreducible polynomial $f \notin k[X_1, \ldots, X_n]$. Let ℓ be the smallest positive integer such that (after clearing the denominators of the coefficients of f we have) $f \in k\left[t^{1/\ell!}\right][X_1, \ldots, X_n]$; in particular, ℓ depends only on V. For the sake of simplifying our notation, we let $T := t^{1/\ell!}$, and moreover we may assume $f \in k[T][X_1, \ldots, X_n]$ is reduced (over $k[T]$). Let M and σ be as in Lemma 2.5 with respect to the polynomial f defined over $k[T]$ (note that $\sigma(T) = T^M$ also yields $\sigma(t) = t^M$).

Let $P = (x_1, \ldots, x_n) \in V(\overline{k(t)}) = V(\overline{k(T)})$. We apply Lemma 2.5 to f and P and conclude that either

$$\mathrm{h}(P) \geq \frac{1}{2n\ell! \deg(f)} \qquad (2.9)$$

or

$$f(\sigma(x_1), \ldots, \sigma(x_n)) = 0. \qquad (2.10)$$

Note that in (2.9), the factor $\ell!$ comes from the definition of the Weil height and the fact that f is defined over $k[T]$, while $[k(T) : k(t)] = \ell!$.

If (2.9) holds, then we obtained a good lower bound for the height of P (depending only on V).

Assume (2.10) holds. Because f is an irreducible and reduced polynomial, whose coefficients are not all in k, Lemma 2.8 yields that $f(X_1, \ldots, X_n)$ cannot divide $f^{\sigma^{-1}}(X_1, \ldots, X_n)$. Indeed, if $f \mid f^{\sigma^{-1}}$ then there exists $a \in \overline{k(t)}$ (actually in $k(T)$) such that $af \in k[X_1, \ldots, X_n]$. But this contradicts the assumption that f is reduced (with coefficients in $k[T]$), and not all of its coefficients are in k.

We know f has more than one monomial because it is reduced and not all of its coefficients are in k. Without loss of generality, we may assume f has positive degree in X_n. Because f is L-irreducible and f does not divide $f^{\sigma^{-1}}$ (which is also defined over L), we conclude that f and $f^{\sigma^{-1}}$ are

relatively prime. So, the resultant R of the polynomials $f(X_1, \ldots, X_n)$ and $f^{\sigma^{-1}}(X_1, \ldots, X_n)$ with respect to the variable X_n is nonzero. Moreover, R depends only on f and on σ (which ultimately depends on f).

The nonzero polynomial $R \in k(t^{1/M\ell!})[X_1, \ldots, X_{n-1}]$ (since $f^{\sigma^{-1}}$ is defined over $k(t^{1/M\ell!})$) vanishes on $(x_1, \ldots, x_{n-1})$. Applying the induction hypothesis to the hypersurface $R = 0$ in $\mathbb{A}^{n-1}$, we conclude there exists a proper k-subvariety $Z_1 \subset \mathbb{A}^{n-1}$, depending only on R (and so, only on V) and there exists a positive constant C, depending only on R (and so, only on V) such that either

$$\mathrm{h}(x_1, \ldots, x_{n-1}) \geq C \text{ or} \tag{2.11}$$

$$(x_1, \ldots, x_{n-1}) \in Z_1(\overline{k(t)}). \tag{2.12}$$

If (2.11) holds, then $\mathrm{h}(x_1, \ldots, x_{n-1}, x_n) \geq \mathrm{h}(x_1, \ldots, x_{n-1}) \geq C$ and we have a height inequality as in the conclusion of Lemma 2.9. If (2.12) holds, then $(x_1, \ldots, x_n) \in \left(Z_1 \times \mathbb{A}^1\right)\left(\overline{k(t)}\right)$ and $Z_1 \times \mathbb{A}^1$ is a k-variety, strictly contained in $\mathbb{A}^n$, as desired in Lemma 2.9. This proves the inductive step and concludes the proof of Lemma 2.9. □

The following result is proved in [5, Corollary 2.4]; essentially the proof relies on the fact that if (C_i, Z_i) are two pairs as in the conclusion of Lemma 2.9 (for $i = 1, 2$), then $(\min\{C_1, C_2\}, Z_1 \cap Z_2)$ is another pair satisfying the conclusion of Lemma 2.9.

Corollary 2.11. *Let V be a proper subvariety of $\mathbb{A}^n$ defined over $\overline{k(t)}$. There exists a positive constant C and a proper subvariety $Z \subset \mathbb{A}^n$ defined over k, such that the pair (C, Z) satisfies the conclusion of Lemma 2.9, and moreover Z is minimal with this property (with respect to the inclusion of subvarieties of $\mathbb{A}^n$).*

Then Theorem 1.3 follows from Corollary 2.11 exactly as the proof of [5, Theorem 2.2]; the only difference is that $\mathbb{F}_p$ is replaced by k.

References

1. Y. Bilu, Limit distribution of small points on algebraic tori, *Duke Math. J.*, **89**(1997), no. 3, 465-476.

2. E. Bombieri and U. Zannier, Algebraic points on subvarieties of $\mathbb{G}_m^n$, *Internat. Math. Res. Notices*, **7** (1995), 333-347.

3. S. David and P. Philippon, Sous-variétés de torsion des variétés semi-abéliennes (French) [Torsion subvarieties of semi-abelian varieties], *C. R. Acad. Sci. Paris Sr. I Math.* **331** (2000), no. 8, 587-592.

4. H. Derksen and D. Masser, Linear equations over multiplicative groups, recurrences, and mixing I, *Proc. Lond. Math. Soc.* (3) **104** (2012), 1045-1083.

5. D. Ghioca, Points of small height on varieties defined over a function field, *Canad. Math. Bull.*, **52** (2009), 237-244.

6. S. Lang, Division points on curves, *Ann. Mat. Pura Appl.* (4) **70** (1965), 229-234.

7. M. Laurent, Équations diophantiennes exponentielles (French) [Exponential Diophantine equations], *Invent. Math.*, **78** (1984), 299-327.

8. M. Raynaud, Sous-variétés d'une variété abélienne et points de torsion (French) [Subvarieties of an abelian variety and torsion points], Arithmetic and Geometry, Vol. I, 327–352, Prog. Math., **35**, Birkhauser boston, Boston, MA, 1983.

9. J.-P. Serre, Lectures on the Mordell-Weil theorem [Translated from the French and edited by Martin Brown from notes by Michel Waldschmidt. With a foreword by Brown and Serre], Third edition. Aspects of Mathematics. Friedr. Vieweg & Sohn, Braunschweig, 1997. x+218 pp.

10. J.-P. Serre, Local fields [Translated from the French by Matin Jay Greenberg], Graduate Texts in Mathematics. Springer-Verlag, New York-Berlin, 1979. viii+241 pp.

11. E. Ullmo, Positivité et discrétion des points algébriques des courbes, (French) [Positivity and discreteness of algebraic points of curves] *Ann. Math.* (2) **147** (1998), no. 1, 167-179.

12. S. Zhang, Positive line bundles on arithmetic varieties, *J. Amer. Math. Soc.*, **8** (1995), no. 1, 187-221.

13. S. Zhang, Equidistribution of small points on abelian varieties, *Ann. Math.* (2) **147** (1998), no. 1, 159-165.