THE ISOTRIVIAL CASE IN THE MORDELL-LANG CONJECTURE FOR SEMIABELIAN VARIETIES DEFINED OVER FIELDS OF POSITIVE CHARACTERISTIC

DRAGOS GHIOCA

Abstract. Let G be a semiabelian variety defined over a finite subfield of an algebraically closed field K of prime characteristic. We describe the intersection of a subvariety X of G with a finitely generated subgroup of $G(K)$.

1. Introduction

The classical Mordell-Lang conjecture for semiabelian varieties G defined over an algebraically closed field K of characteristic 0 (now a theorem due to Laurent [Lau84] for algebraic tori, to Faltings [Fal91] for abelian varieties, and to Vojta [Voj96] for the general case of semiabelian varieties) says that the intersection of a subvariety X of G with a finitely generated subgroup Γ of $G(K)$ is a finite union of cosets of subgroups of Γ. The statement in positive characteristic (i.e., when K is an algebraically closed field of characteristic p) fails; see Hrushovski [Hru96]. When G is defined over a finite subfield $\mathbb{F}_q$ of K, Moosa and Scanlon [MS04] described the aforementioned intersection $X(K) \cap \Gamma$ under the additional assumption that Γ is mapped into itself by a power of the Frobenius endomorphism F of G corresponding to $\mathbb{F}_q$, i.e., Γ is a finitely generated $\mathbb{Z}[F^\ell]$-submodule of $G(K)$ (for a suitable positive integer ℓ). The precise description of the intersection $X(K) \cap \Gamma$ is a finite union of *F-sets* (see Definition 1.2), as given in Theorem 1.3. In [GY], it was shown that if Γ is not a $\mathbb{Z}[F^\ell]$-submodule (for any $\ell \in \mathbb{N}$), then the structure of the intersection $X(K) \cap \Gamma$ can be quite wild (see [GY, Section 2]); in particular, [GY, Examples 2.1, 2.2, 2.3] show that the intersection $X(K) \cap \Gamma$ is no longer a finite union of F-sets (as opposed to what the author claimed erroneously in an earlier paper [Ghi08]). Furthermore, a geometric description of the intersection $X(K) \cap \Gamma$ was proven in [GY], even without assuming that G is defined over a finite subfield of K, but with the disadvantage that this description is not intrinsic to the subgroup Γ and instead it relies on the geometry of G. In the present paper, we obtain an explicit description (see Theorem 1.9) of the intersection $X(K) \cap \Gamma$, in the spirit of the original description of Moosa-Scanlon [MS04].

1.1. The intersection of a subvariety of a semiabelian variety defined over a finite field with a finitely generated subgroup invariant under a power of the Frobenius endomorphism.

2020 *Mathematics Subject Classification.* Primary: 11G10; Secondary: 14G17.
Key words and phrases. semiabelian varieties; finite fields; Mordell-Lang conjecture.

Notation 1.1. *From this point on, we fix a semiabelian variety G defined over a finite subfield $\mathbb{F}_q$ of an algebraically closed field K. We let F be the Frobenius endomorphism of G corresponding to $\mathbb{F}_q$.*

In order to state the result of Moosa-Scanlon [MS04], we introduce the notion of F-sets.

Definition 1.2. *A* groupless F-set *is any subset of $G(K)$ of the form:*

$$\left\{\alpha_0 + \sum_{i=1}^{r} F^{k_i n_i}(\alpha_i) \colon n_i \in \mathbb{N}\right\}, \tag{1.1.1}$$

for given integers $r \geq 0$ and $k_i \geq 1$, and given points $\alpha_0, \alpha_1, \ldots, \alpha_r \in G(K)$. For any finitely generated subgroup $\Gamma \subset G(K)$, we define a groupless F-set in Γ (based in $G(K)$) as a groupless F-set contained in Γ. Also, an F-set in Γ (based in $G(K)$) is any set of the form $S + B$, where S is a groupless F-set in Γ and B is a subgroup of Γ (as always, for any two subsets A and B of a given group, we have that $A + B$ is the set of all elements $a + b$ where $a \in A$ and $b \in B$).

Next, we state the main result of Moosa-Scanlon [MS04].

Theorem 1.3 (Moosa-Scanlon [MS04])**.** *Let G, K, $\mathbb{F}_q$ and F be as in Notation 1.1. Let $X \subseteq G$ be a subvariety defined over K and let $\Gamma \subset G(K)$ be a finitely generated subgroup with the property that there exists $\ell \in \mathbb{N}$ such that $F^\ell(\Gamma) \subseteq \Gamma$. Then $X(K) \cap \Gamma$ is a finite union of F-sets in Γ.*

1.2. **The intersection of a subvariety of a semiabelian variety with an arbitrary finitely generated subgroup.** In order to state our main result regarding the intersection of a subvariety of G with an arbitrary finitely generated subgroup of G, we introduce a general arithmetic structure associated to any abelian group (see Definition 1.6). But first we define a general class of linear recurrence sequences (see Definition 1.5). We recall that for a linear recurrence equation $\{a_n\}_{n\geq 1}$ given by the recurrence relation:

$$a_{n+m} = \sum_{i=0}^{m-1} c_i a_{n+i},$$

the characteristic equation is $x^m - c_{m-1}x^{m-1} - \cdots - c_0 = 0$; for more details, see [CGSZ21, Section 2.3].

Definition 1.4. *A subset $S \subseteq \mathbb{C}^*$ is called* powers-closed *if for any $r \in S$ and any non-negative integer n, we have that $r^n \in S$.*

Definition 1.5. *Let $S \subseteq \mathbb{C}^*$ be a powers-closed set. A linear recurrence sequence $\{a_n\}_{n\geq 1}$ of integers is called an S-arithmetic sequence if the characteristic equation for the linear recurrence sequence $\{a_n\}_{n\geq 1}$ has distinct roots, all contained in S.*

So, with the notation as in Definition 1.5, for an S-arithmetic sequence $\{a_n\}_{n\geq 1}$, there exist distinct numbers $r_1, \ldots, r_m \in S$ and there exist complex numbers $d_1, \ldots, d_m$ such that

$$a_n = \sum_{i=1}^{m} d_i r_i^n \text{ for each } n \geq 1. \tag{1.2.1}$$

Next, we generalize the notion of S-arithmetic sequences inside an arbitrary abelian group.

Definition 1.6. *Let $(\Gamma_0, +)$ be an abelian group, and let $S \subseteq \mathbb{C}^*$ be a powers-closed set. Given an integer $r \geq 1$, elements $P_1, \ldots, P_r \in \Gamma_0$ along with finitely many S-arithmetic sequences*

$$\left\{a_n^{(1)}\right\}_{n\geq 1}, \ \left\{a_n^{(2)}\right\}_{n\geq 1}, \cdots, \ \left\{a_n^{(r)}\right\}_{n\geq 1},$$

we define an S-arithmetic groupless set $\mathcal{U}$ as a set of the following form:

$$\mathcal{U} := \left\{\sum_{i=1}^{r} a_{n_i}^{(i)} \cdot P_i \colon n_1, \cdots, n_r \geq 1\right\}. \tag{1.2.2}$$

Given some finitely generated subgroup $\Gamma \subseteq \Gamma_0$, we say that $\mathcal{U}$ is an S-arithmetic groupless set in Γ (based in Γ_0) if $\mathcal{U} \subseteq \Gamma$. Furthermore, an S-arithmetic set in Γ (based in Γ_0) is defined as a set of the form $\mathcal{U} + B$, where $\mathcal{U}$ is an S-arithmetic groupless set in Γ, while B is a subgroup of Γ.

The notion of S-arithmetic sets is inspired by the definition of F-sets, as shown by the following notation.

Notation 1.7. *With G, K, $\mathbb{F}_q$ and F as in Notation 1.1, then inside the endomorphism ring $\mathrm{End}(G)$, we have that F is integral over $\mathbb{Z}$, i.e., there exists $m \geq 1$ along with integers $c_0, \ldots, c_{m-1}$ (where $c_0 \neq 0$) such that*

$$F^m = \sum_{i=0}^{m-1} c_i F^i \text{ in } \mathrm{End}(G). \tag{1.2.3}$$

Furthermore, the equation

$$x^m - c_{m-1}x^{m-1} - \cdots c_1 x - c_0 = 0 \tag{1.2.4}$$

has distinct complex roots (for more details, see [CGSZ21, Section 2.1]). We let S_F be the subset of $\mathbb{C}^*$ consisting of all complex numbers of the form r^m for integers $m \geq 0$, where r varies among the roots of the equation (1.2.4).

Remark 1.8. Let S_F be as in Notation 1.7. Then using equation (1.2.3) (see also [CGSZ21, Section 3]), we obtain that for any finitely generated subgroup $\Gamma \subset G(K)$, a groupless F-set in Γ (based in $G(K)$) is also an S_F-arithmetic groupless set in Γ (based in $G(K)$). Indeed, there exist m linear recurrence sequences in $\mathbb{Z}$: $\{a_n^{(1)}\}_{n\in\mathbb{N}}, \cdots, \{a_n^{(m)}\}_{n\in\mathbb{N}}$, each one of them satisfying the recurrence relation

$$a_{n+m}^{(i)} = \sum_{k=0}^{m-1} c_k \cdot a_{n+k}^{(i)} \text{ for } n \geq 1,$$

such that for any point $P \in G(K)$, we have

$$F^n(P) = \sum_{i=0}^{m-1} a_n^{(i+1)} \cdot F^i(P).$$

Therefore, an F-set in Γ is an S_F-arithmetic set in Γ. On the other hand, there are many more S_F-arithmetic (groupless) sets in Γ, which are not (groupless) F-sets in Γ. Furthermore, [GY, Example 2.3] shows that arbitrary S_F-arithmetic sets may appear in the intersection of a subvariety $X \subset G$ with some finitely generated subgroup $\Gamma \subset G(K)$; in Theorem 1.9, we prove that *always* $X \cap \Gamma$ must be an S_F-arithmetic set in Γ.

We prove the following main result.

Theorem 1.9. *Let G, K, $\mathbb{F}_q$ and F be as in Notation 1.1, and let S_F be as in Notation 1.7. Let $X \subseteq G$ be a subvariety defined over K and let $\Gamma \subset G(K)$ be a finitely generated subgroup. Then the intersection $X(K) \cap \Gamma$ is a union of finitely many S_F-arithmetic sets in Γ (based in $G(K)$).*

1.3. **Plan for our paper.** We prove Theorem 1.9 as a consequence of a general statement regarding S-arithmetic sets in an abelian group.

Theorem 1.10. *Let $(\Gamma_0, +)$ be an abelian group, let $\tilde{\Gamma}$ and Γ be finitely generated subgroups of Γ_0, and let $S \subseteq \mathbb{C}^*$ be a powers-closed set. Then the intersection of an S-arithmetic set in $\tilde{\Gamma}$ (based in Γ_0) with Γ is a finite union of S-arithmetic sets in Γ (based in Γ_0).*

In Remark 2.5, we show an example that if one were to consider an extension of S-arithmetic sets (see Definition 1.6) in an abelian group which allows for linear recurrence sequences in equation (1.2.2) with characteristic roots of higher multiplicity, then the corresponding problem from Theorem 1.10 leads to some deep unknown questions from classical number theory (such as determining the positive integers whose squares have a given number of nonzero digits when written in base-2).

We prove Theorem 1.10 in Section 2; then Theorem 1.9 follows as an easy corollary of Theorem 1.10 (see Section 3).

2. Proof of Theorem 1.10

We prove Theorem 1.10 over the next several subsections of Section 2; so, throughout the entire Section 2, we work under the hypotheses and notation from Theorem 1.10.

We start with a very simple lemma, which is used repeatedly in our proofs.

Lemma 2.1. *Let $S \subseteq \mathbb{C}$ be a powers-closed set and let $\{a_n\}_{n\in\mathbb{N}}$ be an S-arithmetic sequence.*

(i) *Then for each $u \in \mathbb{N} \cup \{0\}$ and $v \in \mathbb{N}$, the sequence $\{a_{un+v}\}_{n\in\mathbb{N}}$ is also an S-arithmetic sequence.*
(ii) *Then for each $u, v \in \mathbb{Z}$, the sequence $\{u \cdot a_n + v\}_{n\in\mathbb{N}}$ is also an S-arithmetic sequence.*

Proof. The result follows immediately using the general form (1.2.1) of an element in a linear recurrence sequence whose characteristic roots are non-repeated elements of S. Indeed, let $r_1, \ldots, r_m$ be the characteristic roots of $\{a_n\}_{n\in\mathbb{N}}$.

For part (i), note that the characteristic roots of the linear recurrence sequence $\{a_{un+v}\}$ are r_i^u; since S is powers-closed, then each r_i^u is also in S.

For part (ii), note that adding a nonzero constant to a linear recurrence sequence leads to another linear recurrence sequence whose characteristic roots are distinct; they form the set $\{r_1, \ldots, r_m\} \cup \{1\}$ (also, note that $1 \in S$ because S is a powers-closed set). □

We proceed to proving Theorem 1.10; first, we will sketch the plan for our proof in Section 2.1.

2.1. **Plan for our proof of Theorem 1.10.** The first step in our proof is to make a couple useful reductions:

- in Section 2.2, we show that it suffices to prove Theorem 1.10 when $\Gamma_0 = \tilde{\Gamma}$ is a finitely generated group.
- in Section 2.3, we show that it suffices to assume Γ_0 is torsion-free.

So, having reduced the proof of Theorem 1.10 to the case the ambient group Γ_0 is a torsion-free, finitely generated group, the next step (established in Section 2.4) is to reformulate our problem as a linear algebra question. Using the analysis from Sections 2.5 and 2.6, we conclude our proof of Theorem 1.10 in Section 2.7.

2.2. **It suffices to prove Theorem 1.10 assuming $\Gamma_0 = \tilde{\Gamma}$ is finitely generated.** Indeed, we have an S-arithmetic groupless set $\mathcal{U}$ (contained in $\tilde{\Gamma}$) consisting of all elements of Γ_0 of the form

$$\sum_{i=1}^{m} a_{n_i}^{(i)} \cdot P_i, \text{ as we vary } n_i \in \mathbb{N}, \tag{2.2.1}$$

for some $m \in \mathbb{N}$, for some given elements $P_i \in \Gamma_0$, and for some given S-arithmetic sequences $\{a_n^{(1)}\}_{n\geq 1}, \cdots, \{a_n^{(m)}\}_{n\geq 1}$. Then given a subgroup $H \subseteq \tilde{\Gamma}$, our goal is to show that the S-arithmetic set

$$\mathcal{F} := \mathcal{U} + H \tag{2.2.2}$$

intersects Γ in a finite union of S-arithmetic sets (in Γ). At the expense of replacing $\tilde{\Gamma}$ with a larger subgroup of Γ_0 (but still finitely generated), we may assume both that $\Gamma \subseteq \tilde{\Gamma}$ and that each $P_i \in \tilde{\Gamma}$ for $i = 1, \ldots, m$. Finally, without loss of generality, we may assume $\Gamma_0 = \tilde{\Gamma}$ is finitely generated.

2.3. **It suffices to prove Theorem 1.10 assuming Γ_0 is torsion-free.** We already reduced proving Theorem 1.10 to the special case $\Gamma_0 = \tilde{\Gamma}$ is finitely generated. Therefore, since $(\Gamma_0, +)$ is an abelian group, then Γ_0 is the direct product of a (finitely generated) free subgroup $\Gamma_{0,\text{free}}$ with a finite torsion subgroup $\Gamma_{0,\text{tor}}$. Hence both H and Γ (being subgroups of Γ_0) are finite unions of cosets of subgroups of $\Gamma_{0,\text{free}}$, i.e.,

$$H = \bigcup_{i=1}^{k} (h_i + H_{\text{free}}) \text{ and } \Gamma = \bigcup_{j=1}^{\ell} (\gamma_j + \Gamma_{\text{free}}), \tag{2.3.1}$$

for some $k, \ell \in \mathbb{N}$, some elements h_i and γ_j in Γ_0 and some subgroups H_{free} and Γ_{free} of $\Gamma_{0,\text{free}}$. So, it suffices to prove that for any given $i_0 \in \{1, \ldots, k\}$ and $j_0 \in \{1, \ldots, \ell\}$, the intersection

$$(\mathcal{U} + (h_{i_0} + H_{\text{free}})) \cap (\gamma_{j_0} + \Gamma_{\text{free}}) \tag{2.3.2}$$

is a finite union of S-arithmetic sets in Γ. The following observation will be used repeatedly in our proof.

Remark 2.2. Since a constant sequence (in $\mathbb{Z}$) is itself a linear recurrence sequence whose characteristic root is 1 (this is covered by the case $u = 0$ in Lemma 2.1, either parts (i) or (ii)), we note that for any S-arithmetic groupless set $\mathcal{G}$ and for any $\delta \in \Gamma_0$, we have that also $\delta + \mathcal{G}$

is an S-arithmetic groupless set (note also that S contains 1). Furthermore, if $\mathcal{G}$ is an S-arithmetic set, then also $\delta + \mathcal{G}$ is an S-arithmetic set.

We re-write the intersection from (2.3.2) as

$$\gamma_{j_0} + (((-\gamma_{j_0} + h_{i_0}) + \mathcal{U} + H_{\text{free}}) \cap \Gamma_{\text{free}}). \tag{2.3.3}$$

In light of Remark 2.2, it suffices to prove that

$$((-\gamma_{j_0} + h_{i_0}) + \mathcal{U} + H_{\text{free}}) \cap \Gamma_{\text{free}} \text{ is an } S\text{-arithmetic set in } \Gamma. \tag{2.3.4}$$

Now, for $i \in \{1, \dots, m\}$, we write each P_i (see (2.2.1)) as $P_{i,\text{tor}} + P_{i,\text{free}}$ with $P_{i,\text{tor}} \in \Gamma_{0,\text{tor}}$ and $P_{i,\text{free}} \in \Gamma_{0,\text{free}}$. Also, we write $-\gamma_{j_0} + h_{i_0} = \eta_{\text{tor}} + \eta_{\text{free}}$ for some $\eta_{\text{tor}} \in \Gamma_{0,\text{tor}}$ and $\eta_{\text{free}} \in \Gamma_{0,\text{free}}$.

We let $M := |\Gamma_{0,\text{tor}}|$. Because each linear recurrence sequence of integers is preperiodic modulo any given integer (and thus, in particular, preperiodic modulo M), we obtain that the set of tuples $(n_1, \dots, n_m)$ of positive integers satisfying

$$\eta_{\text{tor}} + \sum_{i=1}^{m} a_{n_i}^{(i)} \cdot P_{i,\text{tor}} = 0 \text{ in } \Gamma_0 \tag{2.3.5}$$

is a finite union of sets of the form

$$\{(k_1 n_1 + \ell_1, k_2 n_2 + \ell_2, \dots, k_m n_m + \ell_m) \colon \text{ for arbitrary } n_1, \dots, n_m \geq 1\}$$

for some given $2m$-tuples of integers:

$$(k_1, \ell_1, k_2, \ell_2, \cdots, k_m, \ell_m) \text{ with } k_i \geq 0 \text{ and } \ell_i \geq 1. \tag{2.3.6}$$

Therefore, at the expense of replacing each linear recurrence sequence $\left\{a_n^{(i)}\right\}_{n\geq 1}$ with $\left\{a_{k_i n+\ell_i}^{(i)}\right\}_{n\geq 1}$ (which is still an S-arithmetic sequence, as shown in Lemma 2.1 (i)), we reduce our problem to proving that the intersection with Γ_{free} of the S-arithmetic set

$$\mathcal{F}_1 := \mathcal{S}_1 + H_{\text{free}}, \tag{2.3.7}$$

where $\mathcal{S}_1$ is the S-arithmetic groupless set given by

$$\mathcal{S}_1 := \left\{ \eta_{\text{free}} + \sum_{i=1}^{m} a_{n_i}^{(i)} \cdot P_{i,\text{free}} \colon n_1, \dots, n_m \geq 1 \right\} \tag{2.3.8}$$

is a finite union of S-arithmetic sets in Γ.

Therefore, from now on, we work under the assumption that Γ_0 is torsion-free (see also equations (2.3.7) and (2.3.8)).

2.4. **Reduction to a linear algebra question.** So, we work under the assumption that Γ_0 is a finitely generated, free abelian group; hence it is isomorphic to $\mathbb{Z}^r$ and we let $Q_1, \dots, Q_r \in \Gamma_0$ be some fixed generators for Γ_0.

We have a subgroup $\Gamma \subseteq \Gamma_0$ and also, we have an S-arithmetic set $\mathcal{F} \subseteq \Gamma_0$. Furthermore, $\mathcal{F} = \mathcal{U} + H$ for an S-arithmetic groupless set $\mathcal{U}$ and for a subgroup $H \subseteq \Gamma_0$. Since H is a subgroup of Γ_0, then it is also a finitely generated, free abelian group; thus, we let $\{R_1, \dots, R_s\}$

be a given $\mathbb{Z}$-basis for H (where $s \leq r$). For each $i \in \{1, \ldots, s\}$, we write R_i in terms of the basis $\{Q_1, \ldots, Q_r\}$ of Γ_0 as

$$R_i := \sum_{j=1}^{r} b_{i,j} Q_j \text{ for some } b_{i,j} \in \mathbb{Z}. \tag{2.4.1}$$

The S-arithmetic groupless set $\mathcal{U}$ consists of all elements of the form

$$\sum_{i=1}^{m} a_{n_i}^{(i)} \cdot P_i, \text{ as we vary } n_i \in \mathbb{N}, \tag{2.4.2}$$

for some given elements $P_1, \ldots, P_m \in \Gamma_0$, where $\{a_n^{(i)}\}_{n \in \mathbb{N}}$ are S-arithmetic sequences (for $1 \leq i \leq m$). Then for each $i \in \{1, \ldots, m\}$, we write P_i as

$$P_i := \sum_{j=1}^{r} c_{i,j} Q_j \text{ for some } c_{i,j} \in \mathbb{Z}. \tag{2.4.3}$$

Therefore, a point in $\mathcal{F} = \mathcal{S} + H$ is of the form

$$\sum_{k=1}^{s} y_k \cdot R_k + \sum_{i=1}^{m} a_{n_i}^{(i)} \cdot P_i, \tag{2.4.4}$$

for some arbitrary integers y_k and arbitrary positive integers n_i. Hence, using equations (2.4.1) and (2.4.3), we write any point in $\mathcal{F}$ as:

$$\sum_{j=1}^{r} \left(\sum_{k=1}^{s} b_{k,j} y_k + \sum_{i=1}^{m} c_{i,j} a_{n_i}^{(i)} \right) \cdot Q_j, \tag{2.4.5}$$

for some arbitrary $y_k \in \mathbb{Z}$ and some arbitrary $n_i \in \mathbb{N}$.

Now, following [Ghi08, Definition 3.5], we define **C**-subsets, **L**-subsets and **CL**-sets of $\mathbb{Z}^r$; their notation comes from *congruence* equation (for **C**-subset), respctively *linear* equation (for **L**-subset).

Definition 2.3. *A* **C***-subset of* $\mathbb{Z}^m$ *is a set* $\mathbf{C}(d_1, \ldots, d_r, D)$*, where* $d_1, \ldots, d_r, D \in \mathbb{Z}$ *(with* $D \neq 0$*), containing all solutions* $(x_1, \ldots, x_r) \in \mathbb{Z}^r$ *of the congruence equation* $\sum_{i=1}^{r} d_i x_i \equiv 0 \pmod{D}$*.*

An **L***-subset of* $\mathbb{Z}^r$ *is a set* $\mathbf{L}(d_1, \ldots, d_r)$*, where* $d_1, \ldots, d_r \in \mathbb{Z}$*, containing all solutions* $(x_1, \ldots, x_r) \in \mathbb{Z}^r$ *of the linear equation* $\sum_{i=1}^{r} d_i x_i = 0$*.*

A **CL***-subset of* $\mathbb{Z}^r$ *is either a* **C***-subset or an* **L***-subset of* $\mathbb{Z}^r$*.*

Remark 2.4. We note that our **C**-subsets, **L**-subsets and **CL**-subsets are slightly simpler than the ones defined in [Ghi08] because we can absorb any coset of a subgroup in our definition of S-arithmetic groupless subsets of Γ_0 since the constant sequence is a linear recurrence sequence with unique characteristic root equal to 1 (see also Lemma 2.1 and Remark 2.2).

As proven in [Ghi08, Subclaim 3.6], there exist finitely many **C**-subsets C_i of $\mathbb{Z}^r$ (with the index i varying in some finite set I) and there exist finitely many **L**-subsets L_j of $\mathbb{Z}^r$ (with

the index j varying in some finite set J) such that for any $(x_1, \ldots, x_r) \in \mathbb{Z}^r$, we have

$$\sum_{i=1}^{r} x_i Q_i \in \Gamma \tag{2.4.6}$$

if and only if

$$(x_1, \ldots, x_r) \in \left(\bigcap_{i\in I} C_i\right) \bigcap \left(\bigcap_{j\in J} L_j\right). \tag{2.4.7}$$

Combining equations (2.4.5), (2.4.6) and (2.4.7), our problem reduces to analyzing for which integers $y_1, \ldots, y_s$ and for which positive inttegers $n_1, \ldots, n_m$, we have that

$$\left(\sum_{k=1}^{s} b_{k,j} y_k + \sum_{i=1}^{m} c_{i,j} a_{n_i}^{(i)}\right)_{1\le j\le r} \in \left(\bigcap_{i\in I} C_i\right) \bigcap \left(\bigcap_{j\in J} L_j\right). \tag{2.4.8}$$

We analyze the conditions imposed on the tuples $(y_1, \ldots, y_s, n_1, \ldots, n_m)$ so that the left-hand side from equation (2.4.8) belongs to some C_i, respectively some L_j. We split our analysis for these two cases over the next two Sections 2.5 and 2.6; there are significant differences between these two cases, one of them being that the case of **C**-subsets is easier than the case of **L**-subsets and it can be treated one congruence equation at a time.

2.5. **The case of C-subsets of $\mathbb{Z}^r$.** In this Section 2.5, we analyze the condition that the left-hand side of equation (2.4.8) belongs to some given **C**-subset $C \subseteq \mathbb{Z}^r$. Hence, for given integers $d_1, \ldots, d_r$ and nonzero integer D, we analyze the equation

$$\sum_{j=1}^{r} d_j \cdot \left(\sum_{k=1}^{s} b_{k,j} y_k + \sum_{i=1}^{m} c_{i,j} a_{n_i}^{(i)}\right) \equiv 0 \pmod{D}. \tag{2.5.1}$$

We note that in equation (2.5.1), the *unknowns* are the y_k's and the n_i's, while D, the d_j's, the $b_{k,j}$'s and the $c_{i,j}$'s are given integers. Since any linear recurrence sequence of integers (such as each $\{a_n^{(i)}\}_{n\ge 1}$) is preperiodic modulo any given integer (such as D), we obtain that there exist finitely many tuples of non-negative integers

$$(u_1, v_1, u_2, v_2, \cdots, u_{s+m}, v_{s+m})$$

such that the set of tuples $(y_1, \ldots, y_s, n_1, \ldots, n_m)$ satisfying equation (2.5.1) is a finite union of sets of the form

$$\{(u_1 y_1 + v_1, \cdots, u_s y_s + v_s, u_{s+1} n_1 + v_{s+1}, \cdots, u_{s+m} n_m + v_{s+m}) \colon \text{for } y_i \in \mathbb{Z} \text{ and } n_i \in \mathbb{N}\}. \tag{2.5.2}$$

As before (see Lemma 2.1 (i)), replacing each linear recurrence sequence $\{a_n^{(i)}\}_{n\in\mathbb{N}}$ by $\{a^{(i)}_{u_{s+i}n+v_{s+i}}\}_{n\in\mathbb{N}}$ (for $i = 1, \ldots, m$) leads to another m linear recurrence sequences with simple characteristic roots that all live in the set S. Furthermore, replacing each y_k by $u_k y_k + v_k$ leads to replacing the subgroup H with a coset $\delta_1 + H_1$ of a subgroup $H_1 \subseteq H$. Then, once again using Remark 2.2, we conclude that replacing each y_k by $u_k y_k + v_k$ (for $1 \le k \le s$) and replacing each n_i by $u_{s+i} n_i + v_{s+i}$ (for $1 \le i \le m$) leads to replacing the S-arithmetic subset $\mathcal{F} = \mathcal{S} + H$ by

another S-arithmetic set $\mathcal{S}_1 + H_1$, where $\mathcal{S}_1$ is the S-arithmetic groupless set

$$\mathcal{S}_1 := \left\{ \sum_{k=1}^{s} v_k R_k + \sum_{i=1}^{m} a^{(i)}_{u_{s+i}n_i+v_{s+i}} \cdot P_i \colon n_i \in \mathbb{N} \right\},$$

while H_1 is the subgroup of H spanned by $u_1R_1, \cdots, u_kR_k$. Therefore, from now on, we may assume that in the right-hand side of the equation (2.4.8) we only have **L**-subsets L_j.

2.6. **The case of L-subsets of $\mathbb{Z}^r$.** So, with the notation as in equation (2.4.8), letting $|J| = u$, we need to find $(y_1, \ldots, y_s, n_1, \ldots, n_m) \in \mathbb{Z}^s \times \mathbb{N}^m$ such that

$$\left(\sum_{k=1}^{s} b_{k,j} y_k + \sum_{i=1}^{m} c_{i,j} a^{(i)}_{n_i} \right)_{1 \le j \le r} \in \bigcap_{1 \le h \le u} L_h, \tag{2.6.1}$$

where for each $h \in \{1, \ldots, u\}$, the subset $L_h \subseteq \mathbb{Z}^r$ is cut out by the linear equation

$$\sum_{j=1}^{r} d_{h,j} x_j = 0, \tag{2.6.2}$$

for some given integers $d_{h,j}$. Combining equation (2.6.1) with the linear equations (2.6.2) (for $1 \le j \le u$), we obtain a system of u linear equations:

$$\sum_{k=1}^{s} \left(\sum_{j=1}^{r} b_{k,j} d_{h,j} \right) \cdot y_k = -\sum_{i=1}^{m} \left(\sum_{j=1}^{r} d_{h,j} c_{i,j} \right) \cdot a^{(i)}_{n_i} \text{ for } 1 \le h \le u. \tag{2.6.3}$$

For each $h = 1, \ldots, u$ and for each $k = 1, \ldots, s$, we let

$$e_{h,k} := \sum_{j=1}^{r} b_{k,j} d_{h,j};$$

also, for each $h = 1, \ldots, u$ and for each $i = 1, \ldots, m$, we let

$$f_{h,i} := -\sum_{j=1}^{r} d_{h,j} c_{i,j}.$$

Clearly, $e_{h,k} \in \mathbb{Z}$ (for each $1 \le h \le u$ and $1 \le k \le s$) and $f_{h,i} \in \mathbb{Z}$ (for each $1 \le h \le u$ and $1 \le i \le m$); also, we recall that each $a^{(i)}_{n_i} \in \mathbb{Z}$ for $1 \le i \le m$ and each $n_i \in \mathbb{N}$. So, we have a linear system of u equations with unknowns $y_1, \ldots, y_s$:

$$\sum_{k=1}^{s} e_{h,k} y_k = \sum_{i=1}^{m} f_{h,i} a^{(i)}_{n_i} \text{ for } 1 \le h \le u. \tag{2.6.4}$$

In order for the system (2.6.4) be solvable for some $y_1, \ldots, y_s \in \mathbb{Q}$, there are finitely many linear relations to be satisfied by the right-hand side terms from (2.6.4); since each $e_{h,k}$ is an integer, these linear relations will have integer coefficients as well. Hence, there are finitely many equations, say w equations (for some $w \ge 0$), and there are some given integers $g_{\ell,h}$ with $1 \le \ell \le w$ and $1 \le h \le u$ such that:

$$\sum_{h=1}^{u} g_{\ell,h} \cdot \sum_{i=1}^{m} f_{h,i} a^{(i)}_{n_i} = 0, \tag{2.6.5}$$

which need to be satisfied (for $1 \leq \ell \leq w$) by the positive integers n_i in order to find a solution $(y_1, \ldots, y_s)$ (even over the rationals) for the system (2.6.4). Letting

$$z_{\ell,i} := \sum_{h=1}^{u} g_{\ell,h} \cdot f_{h,i} \text{ for each } \ell \in \{1, \ldots, w\} \text{ and for each } i \in \{1, \ldots, m\},$$

then equations (2.6.5) translate to equations:

$$\sum_{i=1}^{m} z_{\ell,i} \cdot a_{n_i}^{(i)} = 0 \text{ for } \ell = 1, \ldots, w. \tag{2.6.6}$$

Since each sequence $\{a_n^{(i)}\}_{n\in\mathbb{N}}$ is a linear recurrence sequence with simple characteristic roots $\tilde{r}_{i,1}, \ldots, \tilde{r}_{i,m_i} \in S$ (for some $m_i \in \mathbb{N}$), then the equations (2.6.6) translate into equations:

$$\sum_{i=1}^{m} \sum_{j=1}^{m_i} \tilde{z}_{\ell,i,j} \cdot \tilde{r}_{i,j}^{n_i} = 0 \text{ for } \ell = 1, \ldots, w, \tag{2.6.7}$$

for some constants $\tilde{z}_{\ell,i,j} \in \mathbb{C}$. The famous theorem of Laurent [Lau84] (which solves the classical Mordell-Lang conjecture for algebraic tori) yields that the set of tuples $(n_1, \ldots, n_m) \in \mathbb{N}^m$ satisfying the equations (2.6.7) is a union of finitely many sets $\tilde{S}_k$ of the following form. Each set $\tilde{S}_k$ consists of all tuples $(n_1, \ldots, n_m) \in \mathbb{N}^m$ satisfying finitely many equations of the form:

$$n_j = \tilde{n}_{0,j} \text{ for some given } \tilde{n}_{0,j} \in \mathbb{N}, \tag{2.6.8}$$

or

$$n_j \equiv \tilde{n}_{0,j} \pmod{\tilde{N}_{0,j}} \text{ for some given } \tilde{n}_{0,j}, \tilde{N}_{0,j} \in \mathbb{N}, \tag{2.6.9}$$

or

$$\tilde{n}_{0,j} \cdot n_j = \tilde{n}_{0,j_1} \cdot n_{j_1} \text{ for some given } j \neq j_1 \text{ and } \tilde{n}_{0,j}, \tilde{n}_{0,j_1} \in \mathbb{N}. \tag{2.6.10}$$

Remark 2.5. This is the only part of our proof which requires that in the definition of S-arithmetic sets the corresponding linear recurrence sequences have distinct characteristic roots. Indeed, otherwise, the problem becomes *very difficult*. For example, consider the case when $\Gamma_0 = \mathbb{Z}^2$, $\Gamma = \mathbb{Z} \times \{0\}$ and $\mathcal{F} = \mathcal{S}$ is the set of all elements of $\mathbb{Z}^2$ of the form:

$$n_0^2 \cdot (1, 1) + \sum_{i=1}^{m} 2^{n_i} \cdot (1, -1) \text{ for arbitrary } n_0, n_1, \ldots, n_m \in \mathbb{N}. \tag{2.6.11}$$

The set (2.6.11) corresponds to the linear recurrence sequences $\{n_0^2\}_{n_0\geq 1}$ along with $\{2^{n_i}\}_{n_i\geq 1}$ (for $1 \leq i \leq m$); their characteristic roots live in the powers-closed set $\{2^s \colon s \geq 0\}$, *but* the first of these linear recurrence sequences has 1 as a *repeated characteristic root*. Then analyzing the counterpart of Theorem 1.10 for this example leads to finding all solutions $(n_0, n_1, \ldots, n_m) \in \mathbb{N}^{m+1}$ for the polynomial-exponential equation:

$$n_0^2 = \sum_{i=1}^{m} 2^{n_i}. \tag{2.6.12}$$

The equation (2.6.12) is a *very deep* Diophantine question, well-beyond the reach of the current known methods; for more details, see [CGSZ21].

Thus, it suffices to prove that for each such set $\tilde{S} := \tilde{S}_k \subseteq \mathbb{N}^m$ (satisfying finitely many equations of the form (2.6.8), (2.6.9) and (2.6.10)), the corresponding set of points (2.4.5) lying in $\mathcal{F} \cap \Gamma$ is an S-arithmetic subset in Γ.

On the other hand, due to the simple form of the equations (2.6.8), (2.6.9) and (2.6.10), the set

$$\left\{ \sum_{i=1}^{m} a_{n_i}^{(i)} \cdot P_i \colon (n_1, \dots, n_m) \in \tilde{S} \right\} \tag{2.6.13}$$

is still an S-arithmetic groupless set (see Lemma 2.1 (i)). Therefore, at the expense of replacing the original S-arithmetic groupless set $\mathcal{U}$ by the set from (2.6.13), we may assume that the linear system of equations (2.6.4) is solvable (in $\mathbb{Q}$) for each $(n_1, \dots, n_m) \in \mathbb{N}^m$.

So, at the expense of re-shuffling our variables $(y_1, \dots, y_s)$ (which amounts to re-labelling the points $R_1, \dots, R_s$), we may assume that $y_1, \dots, y_t$ (with $t \leq s$) are free variables for the linear system (2.6.4) and the general solution in $\mathbb{Q}$ to the system (2.6.4) consists of tuples $(y_1, \dots, y_s)$ with the property that $y_1, \dots, y_t$ are arbitrary rational numbers, while for $j = 1, \dots, s-t$, we have

$$y_{t+j} = \sum_{k=1}^{t} \tilde{a}_{j,k} \cdot y_k + \sum_{\substack{1 \leq i \leq m \\ 1 \leq h \leq u}} \tilde{b}_{j,h} \cdot f_{h,i} \cdot a_{n_i}^{(i)}, \tag{2.6.14}$$

for some given constants $\tilde{a}_{j,k}, \tilde{b}_{j,h} \in \mathbb{Q}$, while the n_i's from (2.6.14) are arbitrary positive integers. For each $j = 1, \dots, s-t$ and each $i = 1, \dots, m$, we let

$$\tilde{c}_{j,i} := \sum_{h=1}^{u} \tilde{b}_{j,h} f_{h,i};$$

so, we can re-write equation (2.6.14) as follows:

$$y_{t+j} = \sum_{k=1}^{t} \tilde{a}_{j,k} \cdot y_k + \sum_{i=1}^{m} \tilde{c}_{j,i} \cdot a_{n_i}^{(i)}. \tag{2.6.15}$$

2.7. **Conclusion for our proof of Theorem 1.10.** Knowing that the constants $\tilde{a}_{j,k}$ and $\tilde{c}_{j,i}$ from equation (2.6.15) are rational numbers, coupled with the fact that each linear recurrence sequence of integeres is preperiodic with respect to any given moduli, then $y_{t+j} \in \mathbb{Z}$ for each $j = 1, \dots, s-t$ in equation (2.6.15) yields that the corresponding set of all tuples $(y_1, \dots, y_t, n_1, \dots, n_m) \in \mathbb{Z}^t \times \mathbb{N}^m$ consists of finitely many sets of the form

(2.7.1)

$$\left\{ (u_1 y_1 + v_1, \cdots, u_t y_t + v_t, u_{t+1} n_1 + v_{t+1}, \cdots, u_{t+m} n_m + v_{t+m}) : (y_1, \dots, y_t, n_1, \dots, n_m) \in \mathbb{Z}^t \times \mathbb{N}^m \right\}$$

for some given integers u_j, v_j for $1 \leq j \leq t+m$. So, we consider now a given choice of integers u_j and v_j as in equation (2.7.1) so that for each $(y_1, \dots, y_t, n_1, \dots, n_m) \in \mathbb{Z}^t \times \mathbb{N}^m$, we have that $y_{t+1}, \dots, y_s$ given as in (2.6.15) provide a solution over the integers to the system (2.6.4).

Next, we show that the set of points in $\mathcal{F}$ corresponding to a subset $\tilde{E} \subseteq \mathbb{Z}^t \times \mathbb{N}^m$ of the form (2.7.1) is an S-arithmetic set. Indeed, we note first that replacing each n_i (for $1 \leq i \leq m$)

by $u_{t+i} \cdot n_i + v_{t+i}$ yields that the corresponding points in $\mathcal{U}$:

$$\sum_{i=1}^{m} a^{(i)}_{u_{t+i}n_i+v_{t+i}} \cdot P_i \text{ (as we vary } (n_1, \dots, n_m) \in \mathbb{N}^m)$$

still form an S-arithmetic groupless set (once again, see Lemma 2.1 (i)). Second, we claim that the set of all points in H of the form:

$$\sum_{k=1}^{t}(u_k y_k + v_k) \cdot R_k + \sum_{j=1}^{s-t} y_{t+j} \cdot R_{t+j}, \tag{2.7.2}$$

with $y_{t+1}, \dots, y_s$ given as in equation (2.6.15) (where each y_i is replaced by $u_i y_i + v_i$ for $i = 1, \dots, t$) is actually an S-arithmetic set. Indeed, the set from (2.7.2) can be re-written as the sum of an S-arithmetic groupless subset (see also Lemma 2.1 and Remark 2.2):

$$\sum_{k=1}^{t} v_k \cdot R_k + \sum_{j=1}^{s-t} \left(\sum_{k=1}^{t} \tilde{a}_{j,k} v_k + \sum_{i=1}^{m} \tilde{c}_{j,i} a^{(i)}_{u_{t+i}n_i+v_{t+i}} \right) \cdot R_{t+j} \tag{2.7.3}$$

(as we vary $n_1, \dots, n_m$ freely in $\mathbb{N}$) with the subgroup of H consisting of all points:

$$\sum_{k=1}^{t} u_k y_k \cdot R_k + \sum_{j=1}^{s-t} \left(\sum_{k=1}^{t} \tilde{a}_{j,k} u_k y_k \right) \cdot R_{t+j}, \tag{2.7.4}$$

as we vary $y_1, \dots, y_t$ freely in $\mathbb{Z}$.

This concludes our proof of Theorem 1.10.

3. Proof of Theorem 1.9

Theorem 1.9 follows now as an easy corollary of Theorem 1.10.

So, we let G, K, X, Γ, $\mathbb{F}_q$, F and S_F be as in Theorem 1.9 (see also Notations 1.1 and 1.7). We let $\tilde{\Gamma}$ be the finitely generated $\mathbb{Z}[F]$-submodule of $G(K)$ spanned by Γ; in particular, $\tilde{\Gamma}$ is also a finitely generated subgroup of $G(K)$. Using Theorem 1.3, the intersection $X(K) \cap \tilde{\Gamma}$ is a finite union of F-sets in $\tilde{\Gamma}$ (based in $G(K)$). Furthermore, according to Remark 1.8, we have that each F-set in $\tilde{\Gamma}$ is also an S_F-arithmetic set in $\tilde{\Gamma}$ (based in $G(K)$). Finally, using Theorem 1.10 (with $\Gamma_0 := G(K)$) coupled with the fact that

$$X(K) \cap \Gamma = \left(X(K) \cap \tilde{\Gamma} \right) \cap \Gamma,$$

we conclude that $X(K) \cap \Gamma$ is a finite union of S_F-arithmetic sets in Γ (based in $G(K)$), as desired.

References

[CGSZ21] P. Corvaja, D. Ghioca, T. Scanlon, and U. Zannier, *The dynamical Mordell-Lang conjecture for endomorphisms of semiabelian varieties defined over fields of positive characteristic*, J. Inst. Math. Jussieu **20** (2021), no. 2, 669–698.

[Fal91] G. Faltings, *The general case of S. Lang's conjecture*. Barsotti Symposium in Algebraic Geometry (Abano Terme, 1991), 175-182, Perspect. Math., 15, Academic Press, San Diego, CA, 1994.

[Ghi08] D. Ghioca, *The isotrivial case in the Mordell-Lang theorem*, Trans. Amer. Math. Soc. **360** (2008), no. 7, 3839–3856.

[GY] D. Ghioca and S. Yang, *The Mordell-Lang conjecture for semiabelian varieties defined over fields of positive characteristic*, Bull. Aust. Math. Soc., 11 pp., to appear.

[Hru96] E. Hrushovski, *The Mordell-Lang conjecture for function fields*. J. Amer. Math. Soc. **9** (1996), no. 3, 667–690.

[Lau84] M. Laurent, *Équations diophantiennes exponentielles*, Invent. Math. **78** (1984), 299–327.

[MS04] R. Moosa and T. Scanlon, *F-structures and integral points on semiabelian varieties over finite fields*, Amer. J. Math. **126** (2004), 473–522.

[Voj96] P. Vojta, *Integral points on subvarieties of semiabelian varieties. I*, Invent. Math. **126** (1996), no. 1, 133–181.

Department of Mathematics, University of British Columbia, 1984 Mathematics Road, Canada V6T 1Z2

Email address: dghioca@math.ubc.ca