A SPARSITY RESULT FOR THE DYNAMICAL MORDELL-LANG CONJECTURE IN POSITIVE CHARACTERISTIC

DRAGOS GHIOCA, ALINA OSTAFE, SINA SALEH,
AND IGOR E. SHPARLINSKI

ABSTRACT. We prove a quantitative partial result in support of the Dynamical Mordell-Lang Conjecture (also known as the *DML conjecture*) in positive characteristic. More precisely, we show the following: given a field K of characteristic p, given a semiabelian variety X defined over a finite subfield of K and endowed with a regular self-map $\Phi : X \longrightarrow X$ defined over K, given a point $\alpha \in X(K)$ and a subvariety $V \subseteq X$, then the set of all non-negative integers n such that $\Phi^n(\alpha) \in V(K)$ is a union of finitely many arithmetic progressions along with a subset S with the property that there exists a positive real number A (depending only on N, Φ, α, V) such that for each positive integer M, we have

$$\#\{n \in S \colon n \le M\} \le A \cdot (1 + \log M)^{\dim V}.$$

1. INTRODUCTION

1.1. **Notation.** Throughout this paper, we let $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ denote the set of nonnegative integers. As always in arithmetic dynamics, we denote by Φ^n the n-th iterate of the self-map Φ acting on some ambient variety X. For each point x of X, we denote its orbit under Φ by

$$\mathcal{O}_\Phi(x) := \{\Phi^n(x) \colon n \in \mathbb{N}_0\}.$$

Also, for us, an arithmetic progression is a set $\{an + b\}_{n \in \mathbb{N}_0}$ for some $a, b \in \mathbb{N}_0$; in particular, we allow the possibility that $a = 0$, in which case, the above set is a singleton.

1.2. **The Dynamical Mordell-Lang Conjecture.** The Dynamical Mordell-Lang Conjecture (see [GT09]) predicts that for an endomorphism Φ of a quasiprojective variety X defined over a field K of characteristic 0, given a point $\alpha \in X(K)$ and a subvariety $V \subseteq X$, the

2010 *Mathematics Subject Classification.* Primary 11B37, Secondary 11G25, 37P55.

Key words and phrases. linear recurrence sequences, Dynamical Mordell-Lang Conjecture.

set

$$\mathcal{S}(\Phi,\alpha;V) := \{n \in \mathbb{N}_0 \colon \Phi^n(\alpha) \in V(K)\} \tag{1.1}$$

is a finite union of arithmetic progressions; for a comprehensive discussion of the Dynamical Mordell-Lang Conjecture, we refer the reader to the book [BGT16].

When the field K has positive characteristic, then under the same setting as above, the return set $\mathcal{S}$ from (1.1) is no longer a finite union of arithmetic progressions, as shown in [Ghi19, Examples 1.2 and 1.4]; instead, the following conjecture is expected to hold.

Conjecture 1.1 (Dynamical Mordell-Lang Conjecture in positive characteristic)**.** *Let X be a quasiprojective variety defined over a field K of characteristic p. Let $\alpha \in X(K)$, let $V \subseteq X$ be a subvariety defined over K, and let $\Phi : X \longrightarrow X$ be an endomorphism defined over K. Then the set $\mathcal{S}(\Phi,\alpha;V)$ given by (1.1) is a union of finitely many arithmetic progressions along with finitely many sets of the form*

$$\left\{\sum_{j=1}^{m} c_j p^{a_j k_j} \colon k_j \in \mathbb{N}_0 \text{ for each } j = 1,\dots m\right\}, \tag{1.2}$$

for some given $m \in \mathbb{N}$, some given $c_j \in \mathbb{Q}$, and some given $a_j \in \mathbb{N}_0$ (note that in (1.2), the parameters c_j and a_j are fixed, while the unknowns k_j vary over all non-negative integers, $j = 1,\dots,m$).

In [CGSZ20], Conjecture 1.1 is proven for regular self-maps Φ of tori assuming one of the following two hypotheses are met:

(A) $\dim V \le 2$;

or

(B) $\Phi : \mathbb{G}_m^N \longrightarrow \mathbb{G}_m^N$ is a group endomorphism and there exists no nontrivial connected algebraic subgroup G of $\mathbb{G}_m^N$ such that an iterate of Φ induces an endomorphism of G that equals a power of the usual Frobenius.

The proof from [CGSZ20] employs various techniques from Diophantine approximation (in characteristic 0), to combinatorics over finite fields, to specific tools akin to semiabelian varieties defined over finite fields; in particular, the deep results of Moosa & Scanlon [MS04] are essential in the proof. Actually, the Dynamical Mordell-Lang Conjecture in positive characteristic turns out to be even more difficult than the classical Dynamical Mordell-Lang Conjecture since even the case of group endomorphisms of $\mathbb{G}_m^N$ leads to deep Diophantine questions in *characteristic* 0, as shown in [CGSZ20, Theorem 1.4]. More precisely, [CGSZ20, Theorem 1.4] shows that solving Conjecture 1.1 just in the case of group endomorphisms of tori is *equivalent* with solving the following polynomial-exponential equation: given any linear recurrence sequence $\{u_n\}$, given a power q of the prime number p, and given positive integers $c_1, \dots, c_m$ such that

$$\sum_{i=1}^{m} c_i < \frac{q}{2},$$

then one needs to determine the set of all $n \in \mathbb{N}_0$ for which we can find $k_1, \dots, k_m \in \mathbb{N}_0$ such that

$$u_n = \sum_{i=1}^{m} c_i q^{k_i}. \tag{1.3}$$

The equation (1.3) remains unsolved for general sequences $\{u_n\}$ when $m > 2$; for more details about these Diophantine problems, see [CZ13] and the references therein.

1.3. **Statement of our results.** Before stating our main result, we recall that a semiabelian variety is an extension of an abelian variety by an algebraic torus; for more details on semiabelian varieties, we refer the reader to [CGSZ20, Section 2.1] and the references therein.

We prove the following result towards Conjecture 1.1.

Theorem 1.2. *Let K be a field of characteristic p, let X be a semiabelian variety defined over a finite subfield of K, let Φ be a regular self-map of X defined over K. Let $V \subseteq X$ be a subvariety defined over K and let $\alpha \in X(K)$. Then the set $\mathcal{S}(\Phi, \alpha; V)$ defined by (1.1) is a union of finitely many arithmetic progressions along with a set $S \subseteq \mathbb{N}_0$ for which there exists a constant A depending only on X, Φ, α and V such that for all $M \in \mathbb{N}$, we have*

$$\#\{n \in S \colon n \le M\} \le A \cdot (1 + \log M)^{\dim V}. \tag{1.4}$$

Our result strengthens [BGT15, Corollary 1.5] for the case of regular self-maps of semiabelian varieties defined over finite fields since in [BGT15] it is shown that the set S (as in the conclusion of Theorem 1.2) is of Banach density zero; however, the methods from [BGT15] cannot be used to obtain a sparseness result as the one from (1.4).

We establish Theorem 1.2 by combining [CGSZ20, Theorem 3.2] with [Lau84, Théorème 6].

2. Proof of Theorem 1.2

2.1. **Dynamical Mordell-Lang conjecture and linear recurrence sequences.** First, since X is defined over a finite field $\mathbb{F}_q$ of q elements of characteristic p, we let $F : X \longrightarrow X$ be the Frobenius endomorphism corresponding to $\mathbb{F}_q$. We let $P \in \mathbb{Z}[x]$ be the minimal polynomial with integer coefficients such that $P(F) = 0$ in $\mathrm{End}(X)$; according to [CGSZ20, Section 2.1], P is a monic polynomial and it has simple roots $\lambda_1, \dots, \lambda_\ell$, each one of them of absolute value equal to q or $\sqrt{q}$.

Using [CGSZ20, Theorem 3.2], we obtain that the set $\mathcal{S}(\Phi, \alpha; V)$ defined by (1.1) is a finite union of *generalized* F*-arithmetic sequences*, and furthermore, each such generalized F-arithmetic sequence is an intersection of finitely many F*-arithmetic sequences*; see [CGSZ20, Section 3] for exact definitions. Each one of these F-arithmetic sequences consists of all non-negative integers n belonging to a suitable arithmetic progression, for which there exist $k_1, \dots, k_m \in \mathbb{N}_0$ such that

$$u_n = \sum_{i=1}^{m} \sum_{j=1}^{\ell} c_{i,j} \lambda_j^{a_i k_i}, \tag{2.1}$$

for some given linear recurrence sequence $\{u_n\}$ over $\bar{\mathbb{Q}}$, some given $m \in \mathbb{N}_0$, some given constants $c_{i,j} \in \bar{\mathbb{Q}}$ and some given $a_1, \dots, a_m \in \mathbb{N}$. Applying Part (1) of [CGSZ20, Theorem 3.2], we also see that $m \le \dim V$. Furthermore, the linear recurrence sequence $\{u_n\}$ (and the λ_i) along with the constants $c_{i,j}$ and a_i depend solely on X, Φ, α and V.

Moreover, at the expense of further refining to another arithmetic progression, we may assume from now on, that the linear recurrence sequence $\{u_n\}$ is *non-degenerate*, i.e the quotient of any two characteristic roots of this linear recurrence sequence is not a root of unity; furthermore, we may also assume that if one of the characteristic roots is a root of unity, then it actually equals 1. For more details regarding linear recurrence sequences, we refer the reader to [Sch03]. In addition, we know that the characteristic roots of $\{u_n\}$ are all algebraic integers (see part (2) of [CGSZ20, Theorem 3.2]); the characteristic roots of $\{u_n\}$ are either equal to 1 (when Φ contains also a translation besides a group endomorphism) or equal to positive integer powers of the roots of the minimal polynomial of Φ inside $\mathrm{End}(X)$; for more details, see [CGSZ20, Section 3]. So, the equation (2.1) becomes

$$\sum_{r=1}^{s} Q_r(n) \mu_r^n = \sum_{i=1}^{m} \sum_{j=1}^{\ell} c_{i,j} \lambda_j^{a_i k_i}, \tag{2.2}$$

where $\mu_1, \dots, \mu_s$ are the characteristic roots of the sequence $\{u_n\}$ and $Q_1, \dots, Q_s \in \bar{\mathbb{Q}}[x]$.

2.2. **Reduction to the case $s = 1$.** Now, if each polynomial Q_r from the equation (2.2) is constant, then the famous result of Laurent [Lau84] solving the classical Mordell-Lang conjecture (inside an algebraic torus) provides the desired conclusion that the set of all $n \in \mathbb{N}_0$ satisfying an equation of the form (2.2) must be a finite union of arithmetic progressions. So, from now on, we assume that not all of the polynomials Q_r are constant.

Without loss of generality, we assume Q_1 is a non-constant polynomial. According to [Lau84, Section 8, p. 319] (see also [Sch03, Theorem 7.1]) all but finitely many solutions to the equation (2.2) are also solutions to a *subsum* corresponding to the equation (2.2) which contains the term $Q_1(n)\mu_1^n$. More precisely, there exists a subset $1 \in \Sigma_1 \subseteq \{1, \dots, s\}$ and also, there exists a subset $\Sigma_2 \subseteq \{1, \dots, m\} \times \{1, \dots, \ell\}$ such that

$$\sum_{r \in \Sigma_1} Q_r(n)\mu_r^n = \sum_{(i,j) \in \Sigma_2} c_{i,j}\lambda_j^{a_i k_i}. \tag{2.3}$$

Moreover, letting $\pi_1 : \{1, \dots, m\} \times \{1, \dots, \ell\} \longrightarrow \{1, \dots, m\}$ be the projection on the first coordinate, we have $m_1 := \#(\pi_1(\Sigma_2))$; in particular, $m_1 \le m$. Without loss of generality, we assume $\pi_1(\Sigma_2) = \{1, \dots, m_1\}$ (with the understanding that, a priori, m_1 could be equal to 0, even though we show next that this is not the case).

Using [Lau84, Théorème 6], the equation (2.3) has finitely many solutions, unless the following subgroup $G_\Sigma \subseteq \mathbb{Z}^{1+m_1}$ is nontrivial. As described in [Lau84, Section 8, p. 320], the subgroup G_Σ consists of all tuples $(f_0, f_1, \dots, f_{m_1})$ of integers with the property that

$$\mu_r^{f_0} = \lambda_j^{a_i f_i} \text{ for each } r \in \Sigma_1 \text{ and each } (i,j) \in \Sigma_2. \tag{2.4}$$

Since μ_{r_2}/μ_{r_1} is not a root of unity if $r_1 \ne r_2$, we conclude that if Σ_1 contains at least two elements (we already have by our assumption that $1 \in \Sigma_1$), then $f_0 = 0$ in (2.4); furthermore, if $f_0 = 0$, then the equation (2.4) yields that each $f_i = 0$ (since each λ_j has an absolute value greater than 1 and $a_i \in \mathbb{N}$). So, if Σ_1 has more than one element, then the subgroup G_Σ is trivial and thus, [Lau84, Théorème 6] yields that the equation (2.3) (and therefore, also the equation (2.2)) has finitely many solutions, as desired.

2.3. **Concluding the argument.** Therefore, from now on, we may assume that Σ_1 has a single element, i.e., $\Sigma_1 = \{1\}$. In particular, this also means that Σ_2 cannot be the empty set since otherwise the equation (2.3) would simply read

$$Q_1(n)\mu_1^n = 0,$$

which would only have finitely many solutions n (since $\mu_1 \neq 0$ and Q_1 is non-constant). So, we see that indeed Σ_2 is nonempty, which also means that $1 \leq m_1 \leq m$.

We have two cases: either μ_1 equals 1, or not.

Case 1. $\mu_1 = 1$.

Then the equation (2.3) reads:

$$Q_1(n) = \sum_{(i,j)\in\Sigma_2} c_{i,j}\lambda_j^{a_i k_i}. \tag{2.5}$$

Now, for the equation (2.5), the subgroup G_Σ defined above as in [Lau84, Section 8, p. 320] is the subgroup $\mathbb{Z}\times\{(0,\dots,0)\} \subset \mathbb{Z}^{1+m_1}$ since each integer f_i from the equation (2.4) must equal 0 for $i = 1,\dots,m_1$ (note that $\mu_1 = 1$, while each λ_j is not a root of unity). According to [Lau84, Théorème 6, part (b)], there exist positive constants A_1 and A_2 depending only on Q_1, the $c_{i,j}$ and the a_i such that for any solution $(n, k_1,\dots,k_{m_1})$ of the equation (2.5), we have

$$\max\{|k_1|,\dots,|k_{m_1}|\} \leq A_1 \log|n| + A_2. \tag{2.6}$$

So, for each non-negative integer $n \leq M$ (for some given upper bound M) for which there exist integers k_i satisfying the equation (2.5), we have that $|k_i| \leq A_2 + A_1 \log M$, which means that we have at most $A_3\,(1+\log M)^{m_1}$ possible tuples $(k_1,\dots,k_{m_1}) \in \mathbb{Z}^{m_1}$, which may correspond to some $n \in \{0,\dots,M\}$ solving the equation (2.5) (where, once again, A_3 is a constant depending only on the initial data in our problem). Since Q_1 is a polynomial of degree $D \geq 1$, we conclude that the number of solutions $0 \leq n \leq M$ to the equation (2.5) is bounded above by $D\cdot A_3\,(1+\log M)^{m_1}$. Finally, recalling that $m_1 \leq m \leq \dim V$, we obtain the desired conclusion from inequality (1.4).

Case 2. $\mu_1 \neq 1$.

In this case, since we also know that any characteristic root μ_r of the linear reccurence sequence $\{u_n\}_{n\in\mathbb{N}_0}$ is either equal to 1, or not a root of unity, we conclude that μ_1 is not a root of unity.

The equation (2.3) reads now:

$$Q_1(n)\mu_1^n = \sum_{(i,j)\in\Sigma_2} c_{i,j}\lambda_j^{a_i k_i}. \tag{2.7}$$

We analyze again the subgroup $G_\Sigma \subseteq \mathbb{Z}^{1+m_1}$ containing the tuples $(f_0, f_1,\dots,f_{m_1})$ of integers satisfying the equations (2.4), i.e.,

$$\mu_1^{f_0} = \lambda_j^{a_i f_i} \text{ for each } (i,j) \in \Sigma_2. \tag{2.8}$$

Because μ_1 is not a root of unity and also each λ_j is not a root of unity, while the a_i are positive integers, we conclude that a nontrivial tuple

$(f_0, f_1, \dots, f_{m_1})$ satisfying the equations (2.8) must actually have each entry nonzero (i.e., $f_i \neq 0$ for each $i = 0, \dots, m_1$). Therefore, each $\lambda_j^{a_i}$ is multiplicatively dependent with respect to μ_1 and so, there exists an algebraic number λ (which is not a root of unity), there exists a nonzero integer b such that $\mu_1 = \lambda^b$, and whenever there is a pair $(i, j) \in \Sigma_2$, there exist roots of unity $\zeta_{j,i}$ along with nonzero integers b_i such that

$$\lambda_j^{a_i} = \zeta_{j,i} \cdot \lambda^{b_i}. \tag{2.9}$$

We let E be a positive integer such that $\zeta_{j,i}^E = 1$ for each $(j, i) \in \Sigma_2$; then we let $B_i := E \cdot b_i$ for each $i = 1, \dots, m_1$. We now put each exponent k_i appearing in (2.7) in a prescribed residue class modulo E (just getting E^m possible choices) and use (2.9) along with the fact that $\mu_1 = \lambda^b$. Writing $K_i := \lfloor k_i/E \rfloor$, $i = 1, \dots, m_1$, we obtain that finding $n \in \mathbb{N}_0$ which solves the equation (2.7) (and then, in turn, also (2.3) and (2.2)) reduces to finding $n \in \mathbb{N}_0$ which solves at least one of the at most E^m distinct equations of the form:

$$Q_1(n)\lambda^{bn} = \sum_{i=1}^{m_1} d_i \lambda^{B_i K_i}, \tag{2.10}$$

for some algebraic numbers $d_1, \dots, d_{m_1}$, depending only on E, the c_i, and the $\zeta_{j,i}$, $(i, j) \in \Sigma_2$. So, dividing the equation (2.10) by λ^{bn} yields that

$$Q_1(n) = \sum_{i=1}^{m_1} d_i \lambda^{g_i}, \tag{2.11}$$

for some integers g_i. Then once again applying [Lau84, Théorème 6, part (b)] (see also our inequality (2.6)) yields immediately that any solution $(n, g_1, \dots, g_{m_1})$ to the equation (2.11) must satisfy the inequality:

$$\max\{|g_1|, \dots, |g_{m_1}|\} \leq A_4 \log |n| + A_5,$$

for some constants A_4 and A_5 depending only on the initial data in our problem (X, Φ, α, V). Then once again (exactly as in **Case 1**), we conclude that there exists a constant A_6 such that for any given upper bound $M \in \mathbb{N}$, we have at most $A_6 (1 + \log M)^{m_1}$ possible tuples $(g_1, \dots, g_{m_1}) \in \mathbb{Z}^{m_1}$, which may correspond to some $n \in \{0, \dots, M\}$ solving the equation (2.11). Since Q_1 is a polynomial of degree $D \geq 1$, we conclude that the number of solutions $0 \leq n \leq M$ to the equation (2.11) is bounded above by $D \cdot A_6 (1 + \log M)^{m_1}$. Finally, recalling that $m_1 \leq m \leq \dim V$, we obtain the desired conclusion from inequality (1.4).

This concludes our proof of Theorem 1.2.

3. Comments

Remark 3.1. If in the equation (2.2) there exists at least one characteristic root μ_r of $\{u_n\}$ which is multiplicatively independent with respect to each one of the λ_j, then there is never a subsum (2.3) containing μ_r on its left-hand side for which the corresponding group G_Σ would be nontrivial. So, in this case, the equation (2.2) would have only finitely many solutions. Therefore, with the notation as in Theorem 1.2, arguing as in the proof of [CGSZ20, Theorem 1.3], one concludes that if Φ is a group endomorphism of the semiabelian variety X with the property that each characteristic root of its minimal polynomial (in $\mathrm{End}(X)$) is multiplicatively independent with respect to each eigenvalue λ_j of the Frobenius endomorphism of X, then for each $\alpha \in X(K)$, the set $\mathcal{S}(\Phi, \alpha; V)$ defined by (1.1) is a finite union of arithmetic progressions.

Remark 3.2. We notice that in (2.11), if we deal with a polynomial Q_1 of degree 1, then the conclusion from inequality (1.4) is sharp. More precisely, as a specific example, the number of positive integers $n \le M$ which have precisely m nonzero digits (all equal to 1) in base-p is of the order of $(\log M)^m$, which shows that Theorem 1.2 is tight if the Dynamical Mordell-Lang Conjecture reduces to solving the equation (2.11) when $Q_1(n) = n$, $m_1 = m$, $c_1 = \ldots = c_m = 1$ and $\lambda = p$. As proven in [CGSZ20, Theorem 1.4], there are instances when the Dynamical Mordell-Lang Conjecture reduces *precisely* to such equation.

Now, for higher degree polynomials $Q_1 \in \mathbb{Z}[x]$ appearing in the equation (2.11), one expects a lower exponent than m appearing in the upper bounds from (1.4). One also notices that for any polynomial Q_1, arguments n with k nonzero digits in base-p lead to sparse outputs. Hence, simple combinatorics allows us to obtain a lower bound on the best possible exponent in (1.4). However, finding a more precise exponent replacing m in (1.4) when $\deg Q_1 > 1$ seems very difficult beyond some special cases; the authors hope to return to this problem in a sequel paper.

Acknowledgement

D. G. and S S. were partially supported by a Discovery Grant from NSERC, A. O. by ARC Grants DP180100201 and DP200100355, and I. S. by an ARC Grant DP200100355.

References

[BGT15] J. P. Bell, D. Ghioca, and T. J. Tucker, *The Dynamical Mordell-Lang problem for Noetherian spaces*, Funct. Approx. Comment. Math. **53** (2015), 313–328. 3

[BGT16] J. P. Bell, D. Ghioca, and T. J. Tucker, *The dynamical Mordell-Lang conjecture*, Mathematical Surveys and Monographs, **210**, American Mathematical Society, Providence, RI, 2016, xiii+280 pp. 2

[CGSZ20] 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, to appear. 2, 3, 4, 8

[CZ13] P. Corvaja and U. Zannier, *Finiteness of odd perfect powers with four nonzero binary digits*, Ann. Inst. Fourier (Grenoble) **63** (2013), 715–731. 3

[Ghi19] D. Ghioca, *The dynamical Mordell-Lang conjecture in positive characteristic*, Trans. Amer. Math. Soc. **371** (2019), 1151–1167. 2

[GT09] D. Ghioca and T. J. Tucker, *Periodic points, linearizing maps, and the dynamical Mordell-Lang problem*, J. Number Theory **129** (2009), 1392–1403. 1

[Lau84] M. Laurent, *Équations diophantiennes exponentielles*, Invent. Math. **78** (1984), 299–327. 3, 5, 6, 7

[MS04] R. Moosa and T. Scanlon, *F-structures and integral points on semiabelian varieties over finite fields*, Amer. J. Math. **126** (2004), 473–522. 2

[Sch03] W. Schmidt, *Linear recurrence sequences*, Diophantine Approximation (Cetraro, Italy, 2000), Lecture Notes in Math. 1819, Springer-Verlag Berlin Heidelberg, 2003, pp. 171–247. 4, 5

Dragos Ghioca, Department of Mathematics, University of British Columbia, Vancouver, BC V6T 1Z2, Canada
E-mail address: dghioca@math.ubc.ca

Alina Ostafe, School of Mathematics and Statistics, University of New South Wales, Sydney NSW 2052, Australia
E-mail address: alina.ostafe@unsw.edu.au

Sina Saleh, Department of Mathematics, University of British Columbia, Vancouver, BC V6T 1Z2, Canada
E-mail address: sinas@math.ubc.ca

Igor E. Shparlinski, School of Mathematics and Statistics, University of New South Wales, Sydney NSW 2052, Australia
E-mail address: igor.shparlinski@unsw.edu.au