INTERSECTIONS OF POLYNOMIAL ORBITS, AND A DYNAMICAL MORDELL-LANG CONJECTURE

DRAGOS GHIOCA, THOMAS J. TUCKER, AND MICHAEL E. ZIEVE

Abstract. We prove that if nonlinear complex polynomials of the same degree have orbits with infinite intersection, then the polynomials have a common iterate. We also prove a special case of a conjectured dynamical analogue of the Mordell-Lang conjecture.

1. Introduction

One of the main topics in complex dynamics is the study of orbits of polynomial maps: namely, for $f \in \mathbb{C}[X]$ and $x_0 \in \mathbb{C}$, the set $\mathcal{O}_f(x_0) := \{x_0, f(x_0), f(f(x_0)), \dots\}$. We prove the following result about intersections of orbits.

Theorem 1.1. *Let $x_0, y_0 \in \mathbb{C}$ and $f, g \in \mathbb{C}[X]$ with $\deg(f) = \deg(g) > 1$. If $\mathcal{O}_f(x_0) \cap \mathcal{O}_g(y_0)$ is infinite, then f and g have a common iterate.*

The pairs of complex polynomials with a common iterate were determined by Ritt [19]; in Proposition 6.3 we state Ritt's result in the above case $\deg(f) = \deg(g)$.

Our motivation comes from arithmetic geometry. Fundamental progress in this subject has been driven by the Mordell-Lang conjecture on intersections of subgroups and subvarieties of algebraic groups. This conjecture was proved by Faltings [8] and Vojta [25]:

Theorem 1.2. *Let G be a semiabelian variety over $\mathbb{C}$, let V be a subvariety, and let Γ be a finitely generated subgroup of $G(\mathbb{C})$. Then $V(\mathbb{C}) \cap \Gamma$ is a finite union of cosets of subgroups of Γ.*

Recall that a semiabelian variety (over $\mathbb{C}$) is an extension of an abelian variety by a torus $(\mathbb{G}_m)^k$. Theorem 1.2 has the following consequence [11]: if ϕ is an endomorphism of G of degree > 1, then any orbit of ϕ has finite intersection with a subvariety $V \subset G$, unless V contains a positive dimensional subvariety which is periodic under ϕ. In the case $G = \mathbb{G}_m^k$ (which was first treated by Laurent [14]), this implies that if a subvariety $V \subset \mathbb{G}_m^k$ contains no positive dimensional subvariety which is periodic under the map

Date: October 8, 2007.

1991 *Mathematics Subject Classification.* Primary 14G25; Secondary 37F10, 11C08.

The authors thank Robert Benedetto for discussions about polynomial dynamics, and also thank the referee for suggesting several improvements to the exposition. The second author was partially supported by National Security Agency Grant 06G-067.

$\psi : (X_1, \ldots, X_k) \mapsto (X_1^{e_1}, \ldots, X_k^{e_k})$ (with $e_i \in \mathbb{Z}$ and $e_i \geq 2$), then V contains at most finitely many points of any ψ-orbit in $\mathbb{A}^k$.

It is natural to ask whether a similar conclusion holds for any polynomial action on $\mathbb{A}^k$. The first two authors have proposed the following conjecture:

Conjecture 1.3. *Let $f_1, \ldots, f_k$ be polynomials in $\mathbb{C}[X]$, and let V be a subvariety of $\mathbb{A}^k$ which contains no positive dimensional subvariety that is periodic under the action of $(f_1, \ldots, f_k)$ on $\mathbb{A}^k$. Then $V(\mathbb{C})$ has finite intersection with each orbit of $(f_1, \ldots, f_k)$ on $\mathbb{A}^k$.*

This conjecture fits into Zhang's far-reaching system of dynamical conjectures [27]. Zhang's conjectures include dynamical analogues of the Manin-Mumford and Bogomolov conjectures for abelian varieties (now theorems of Raynaud [17, 18], Ullmo [24], and Zhang [26]), as well as a conjecture about the Zariski density of orbits of points under fairly general maps from a projective variety to itself. The latter conjecture is related to our Conjecture 1.3, though neither conjecture contains the other.

A p-adic version of Conjecture 1.3 has been proved in certain special cases [10]. Also, an analogue of Conjecture 1.3 has been proved in positive characteristic, for the additive group under the action of an additive polynomial (Drinfeld module) [9]. This result is a special case of a more general conjecture proposed by Denis [7], in which orbits are replaced with arbitrary submodules under the action of a Drinfeld module.

The techniques of Laurent [14], Faltings [8], and Vojta [25] require conditions that are not implied by the hypotheses of Conjecture 1.3. Laurent's proof uses the fact that the torsion points on a torus are defined over a cyclotomic field; the fields of definition of preperiodic points of general polynomials admit no such simple description. Vojta's proof (which generalizes that of Faltings) relies on the fact that integral points on semiabelian varieties satisfy a strong diophantine property, which does not hold for the points in Conjecture 1.3. Specifically, if z is an S-integral point on $\mathbb{G}_m^k$, then the coordinates of z^n are S-units for all n, whereas the coordinates of points in an orbit of $(f_1, \ldots, f_k)$ need not be S-units. Finally, one crucial difference between the polynomial maps of Conjecture 1.3 and the maps that arise for semiabelian varieties and Drinfeld modules is that the maps in Conjecture 1.3 are not étale in general.

In the present paper we use a new approach to prove the first non-monomial cases of Conjecture 1.3, when the variety V is a line in the affine plane. Our result is as follows, where we write f^n for the n^{th} iterate of the polynomial f.

Theorem 1.4. *Let K be a field of characteristic zero, let $f, g \in K[X]$, and let $x_0, y_0 \in K$. If the set*

$$\{(f^n(x_0), g^n(y_0)) : n \in \mathbb{N}\}$$

has infinite intersection with a line L in $\mathbb{A}^2$ defined over K, then L is periodic under the action of (f, g) on $\mathbb{A}^2$.

Using interpolation (for instance), one can construct examples in which this intersection is finite but larger than any prescribed bound.

Along the lines of Theorem 1.4, we will prove the following generalization of Theorem 1.1.

Theorem 1.5. *Let K be a field of characteristic zero, let $\alpha, \beta, x_0, y_0 \in K$ with $\alpha \neq 0$, and let $f, g \in K[X]$ with $\deg(f) = \deg(g) > 1$. If infinitely many points of $\mathcal{O}_f(x_0) \times \mathcal{O}_g(y_0)$ lie on the line $Y = \alpha X + \beta$, then $g^k(\alpha X + \beta) = \alpha f^k(X) + \beta$ for some positive integer k.*

This result is neither stronger nor weaker than Theorem 1.4: only Theorem 1.4 applies to polynomials of distinct degrees, but if $\deg(f) = \deg(g) > 1$ then Theorem 1.5 strengthens Theorem 1.4 by replacing $\mathcal{O}_{(f,g)}((x_0, y_0))$ with $\mathcal{O}_f(x_0) \times \mathcal{O}_g(y_0)$.

In the simple case that $f(X) = \alpha X$ and $g(X) = \beta X$ with $\alpha, \beta \in K^*$, Theorem 1.4 says that, for any $u, v, w \in K$ that are not all zero, if $u\alpha^n + v\beta^n = w$ for infinitely many n then α or β is a root of unity. Already the result is nontrivial in this case: it is a consequence of Siegel's theorem on integral points of curves, or it could be proved directly using the techniques from Siegel's proof.

One consequence of Theorem 1.4 is that if f and g have distinct degrees then $\mathcal{O}_{(f,g)}((x_0, y_0))$ has finite intersection with any line. We do not know whether the analogous result is true for $\mathcal{O}_f(x_0) \times \mathcal{O}_g(y_0)$ (for lines which are neither horizontal nor vertical, and for polynomials f, g with no common iterate).

Our proofs of Theorems 1.4 and 1.5 involve arguments of several flavors. For general K, we will prove there is a partially-defined map ('specialization') from K to a number field K_0 which allows us to deduce the results for K as a consequence of the results for K_0. Our proof of this fact relies on Ritt's classification of polynomials with a common iterate, as well as a dynamical analogue of a result of Silverman (from [21]) on specialization of nontorsion elements of abelian varieties over function fields.

We reduce the number field case of Theorem 1.4 to the corresponding case of Theorem 1.5 as follows. First, by comparing Weil heights of $f^n(x_0)$ and $g^n(x_0)$, we conclude that f and g must have the same degree if $\mathcal{O}_{(f,g)}((x_0, y_0))$ contains infinitely many points on some line. Next we use Siegel's theorem on integral points to prove Theorem 1.4 when f and g are linear.

The strategy of our proof of Theorem 1.5 for number fields K is as follows, where we simplify the discussion by addressing the case that the line is the diagonal and all polynomials and points are defined over $\mathbb{Z}$. Suppose there are integers x_0, y_0 and polynomials $f, g \in \mathbb{Z}[X]$ such that $\mathcal{O}_f(x_0) \times \mathcal{O}_g(y_0)$ has infinite intersection with the diagonal in $\mathbb{A}^2$. Then, for every m, there are infinitely many integer solutions to the Diophantine equation $f^m(X) = g^m(Y)$. This is an instance of a 'separated variable' Diophantine equation $F(X) = G(Y)$, of which special cases have been studied for many years. The definitive finiteness result for these equations was proved in 2000 by Bilu and Tichy [5]; we will use their result (together with various new results about polynomial decomposition) in order to obtain some information about f and g from the fact that $f^m(X) = g^m(Y)$ has infinitely many integer solutions. Our result will follow upon combining the information deduced for each m.

Although the Bilu-Tichy result has not previously been applied to arithmetic geometry or dynamics, inspection of its proof suggests it fits naturally into both topics. Namely, the two key ingredients in its proof are Siegel's theorem on integral points on curves, and Ritt's results on functional decomposition of complex polynomials.

In more detail, Bilu and Tichy listed five explicit families of 'standard pairs' of polynomials (F_1, G_1) such that, if $F(X) = G(Y)$ has infinitely many integer solutions, then there is a standard pair (F_1, G_1) for which $F = E \circ F_1 \circ a$ and $G = E \circ G_1 \circ b$, where $E, a, b \in \mathbb{Q}[X]$ and $\deg(a) = \deg(b) = 1$. When applying this result to specific polynomials F and G, the main work involved is to determine the various different ways that F and G can be written as compositions of lower-degree polynomials, in order to determine the possibilities for E. In practice, unless F and G are specifically constructed with decomposability in mind, it turns out that any randomly chosen F and G are indecomposable, in which case it is quite simple to apply the Bilu-Tichy criterion (after one has proven this indecomposability). Based on this principle, dozens of recent papers have applied the Bilu-Tichy criterion when F and G come from basically any class of polynomials one can think of: Bernoulli polynomials, falling factorials, power-sum polynomials, Taylor polynomials for e^x, Jacobi polynomials, Laguerre polynomials, Hermite polynomials, Meixner polynomials, Krawtchouk polynomials, etc. (cf., e.g., [3, 4, 23]). In every case, the polynomials were either indecomposable or had just one nontrivial decomposition.

Our situation is quite different, since we are applying Bilu-Tichy to polynomials $F = f^m$ and $G = g^m$, which by their very nature are far from indecomposable. Moreover, we are doing this for arbitrary f and g, which themselves might have various different decompositions. Thus we are forced to prove new results about functional decompositions of polynomials.

The rest of this paper is organized as follows. We begin with some preliminary results about Diophantine equations and functional decomposition. In Section 3 we prove Theorem 1.1 in case K is a number field, modulo the proof of one technical proposition which we give in Section 4. In Section 5 we prove Theorem 1.4 when either K is a number field or the polynomials are linear. Then in Section 6 we prove Theorems 1.4 and 1.5. In the final section we state some conjectures and directions for further research.

Notation. Throughout this paper, f^n denotes the n^{th} iterate of the polynomial f. We also use α^n and X^n for the n^{th} power of a constant or of X itself, but this should not cause confusion. We write $\mathbb{N}$ for the set of positive integers. We write $\overline{K}$ for an algebraic closure of the field K. By a 'nonarchimedean place' of a number field K, we mean a maximal ideal of the ring

$\mathcal{O}_K$ of algebraic integers in K. If S is a finite set of nonarchimedean places of a number field K, then the ring of S-integers of K is the intersection of the localizations of $\mathcal{O}_K$ at all nonarchimedean places outside S.

2. Previous results

In this section we present some known results which will be used in our proof.

2.1. **Diophantine equations.** We will make crucial use of a recent result of Bilu and Tichy [5, Thm. 10.5] describing all $F, G \in \mathbb{Z}[X]$ for which $F(X) = G(Y)$ has infinitely many integer solutions. In fact, they proved a version for S-integers in an arbitrary number field. We state their result in the special case $\deg(F) = \deg(G)$ arising in our proof; in this special case the statement is somewhat simpler than in the general situation.

Theorem 2.1. *Let K be a number field, S a finite set of nonarchimedean places of K, and $F, G \in K[X]$ with $\deg(F) = \deg(G) > 1$. Suppose $F(X) = G(Y)$ has infinitely many solutions in the ring of S-integers of K. Then $F = E \circ F_1 \circ a$ and $G = E \circ G_1 \circ b$, where $E, a, b \in K[X]$ with $\deg(a) = \deg(b) = 1$, and (F_1, G_1) or (G_1, F_1) is one of the following pairs:*

1. (X, X);
2. $(X^2, c \circ X^2)$ *with $c \in K[X]$ linear;*
3. $(D_2(X, \alpha)/\alpha, D_2(X, \beta)/\beta)$ *with $\alpha, \beta \in K^*$;*
4. $(D_n(X, \alpha), -D_n(X \cos(\pi/n), \alpha))$ *with $\alpha \in K$,*

where in the fourth case $n \in \mathbb{N}$ satisfies $\cos(2\pi/n) \in K$.

Here $D_n(X, Y)$ is the unique polynomial in $\mathbb{Z}[X, Y]$ such that $D_n(U + V, UV) = U^n + V^n$. Note that, for $\alpha \in K$, the polynomial $D_n(X, \alpha) \in K[X]$ is monic of degree n. It follows at once from the defining functional equation that $D_n(X, 0) = X^n$ and, for $\alpha \in \mathbb{C}$, we have $\alpha^n D_n(X, 1) = D_n(\alpha X, \alpha^2)$.

We will not need arithmetic information about F_1 and G_1, but instead only need their shape up to composition with linears over an extension of K.

Corollary 2.2. *Let K, S, F, G satisfy the hypotheses of Theorem 2.1. Then $F = \hat{E} \circ H \circ \hat{a}$ and $G = \hat{E} \circ \hat{c} \circ H \circ \hat{b}$ for some $\hat{E} \in \overline{K}[X]$, some linear $\hat{a}, \hat{b}, \hat{c} \in \overline{K}[X]$, and some $H = D_n(X, \hat{\alpha})$ with $\hat{\alpha} \in \{0, 1\}$ and $n \in \mathbb{N}$ satisfying $\cos(2\pi/n) \in K$. In particular, for fixed K, there are only finitely many possibilities for H (even if we vary S, F, G).*

Proof. We consider the four possibilities for (F_1, G_1) in Theorem 2.1. It suffices to show that in each case there is a polynomial H of the desired form such that both F_1 and G_1 are gotten from H by composing on both sides with linears over $\overline{K}$. This is clear in the first two cases (since $D_n(X, 0) = X^n$). For the last two cases, note that if $\gamma \neq 0$ then $D_n(X, \gamma^2) = \gamma^n D_n(X/\gamma, 1)$. Thus, in the third case, F_1 and G_1 are gotten from $D_2(X, 1)$ by composing with linears. And in the fourth case, F_1 and G_1 are gotten from $D_n(X, \hat{\alpha})$ by composing with linears, where $\hat{\alpha} = 1$ if $\alpha \neq 0$ (and $\hat{\alpha} = 0$ otherwise).

Finally, if $\cos(2\pi/n) \in K$ then $[K : \mathbb{Q}] \geq [\mathbb{Q}(\cos(2\pi/n)) : \mathbb{Q}]$; the latter degree equals $\phi(n)/2$ if $n > 2$. Since only finitely many n satisfy $\phi(n) \leq 2[K : \mathbb{Q}]$, there are only finitely many possibilities for H. □

2.2. **Polynomial decomposition.** Our application of Theorem 2.1 relies on results about polynomial decomposition. The fundamental results in this topic were proved by Ritt in the 1920's [20]; for more recent developments, see [16, 22]. Specifically, we will use the following simple but surprising result which shows a type of 'rigidity' of polynomial decomposition.

Lemma 2.3. *Let K be a field of characteristic zero. If $A, B, C, D \in K[X] \setminus K$ satisfy $A \circ B = C \circ D$ and $\deg(B) = \deg(D)$, then there is a linear $\ell \in K[X]$ such that $A = C \circ \ell^{-1}$ and $B = \ell \circ D$.*

Proof. Write $F = A \circ B \,(= C \circ D)$. Pick a linear $v \in K[X]$ such that $\hat{B} := v \circ B$ is monic and has no constant term. Then $F = \hat{A} \circ \hat{B}$, where $\hat{A} = A \circ v^{-1}$. We will show that there are unique $\tilde{A}, \tilde{B} \in K[X]$ such that $F = \tilde{A} \circ \tilde{B}$ and $\deg(\tilde{B}) = \deg(B)$ and $\tilde{B}$ is monic with no constant term. Thus $A = \tilde{A} \circ v$ and $B = v^{-1} \circ \tilde{B}$. Since we could have done the same thing with C and D in place of A and B, the result follows.

Let m be the degree of B, and say the leading term of F is αX^{nm}; then the leading term of $\tilde{A}$ is αX^n. Now consider the identity $F = \tilde{A} \circ \tilde{B}$ and equate terms of degrees $nm - 1, nm - 2, \ldots, nm - m + 1$ to uniquely determine, in order, the terms of $\tilde{B}$ of degrees $m - 1, m - 2, \ldots, 1$. Then consider terms of F of degrees $nm - m, nm - 2m, \ldots, 0$ to determine the terms of $\tilde{A}$ of degrees $n - 1, n - 2, \ldots, 0$. □

Remark. This lemma was first proved by Ritt [20] in the case $K = \mathbb{C}$ (using Riemann surface techniques); the proof above is due to Levi [15].

2.3. **Linear relations of polynomials.** The following lemma shows when a polynomial can be gotten from itself by composing with linears.

Lemma 2.4. *Let K be a field of characteristic zero. If $F \in K[X]$ has degree $d > 1$, and $a, b \in K[X]$ are linears such that $a \circ F = F \circ b$, then there exist $\alpha, \beta \in K$, integers $r, s \geq 0$, an element $\gamma \in K^*$ with $\gamma^s = 1$, and a polynomial $\hat{F} \in X^r K[X^s]$ such that $a = -\alpha + \gamma^r(X + \alpha)$, $F = -\alpha + \hat{F}(X - \beta)$, and $b = \beta + \gamma(X - \beta)$. Specifically, if the coefficients of X^d and X^{d-1} in F are θ_d and θ_{d-1}, we can take $\beta = -\theta_{d-1}/(d\theta_d)$ and $\alpha = -F(\beta)$.*

Proof. Putting $\beta = -\theta_{d-1}/(d\theta_d)$ and $\alpha = -F(\beta)$, we see that $\hat{F} := \alpha + F(X + \beta)$ has no terms of degree $d - 1$ or 0. We rewrite $a \circ F = F \circ b$ as $\hat{a} \circ \hat{F} = \hat{F} \circ \hat{b}$, where $\hat{a} := \alpha + a(X - \alpha)$ and $\hat{b} := -\beta + b(X + \beta)$. Since $\hat{F}$ has no term of degree $d - 1$, also $\hat{a} \circ \hat{F}$ (and hence $\hat{F} \circ \hat{b}$) has no such term, so $\hat{b}$ cannot have a term of degree 0. Then $\hat{F} \circ \hat{b}$ has no term of degree 0, so also $\hat{a}$ has no term of degree 0. Writing $\hat{a} = \delta X$ and $\hat{b} = \gamma X$, we have $\delta\hat{F}(X) = \hat{F}(\gamma X)$. Writing $\hat{F} = \sum \hat{\theta}_i X^i$, it follows that $\delta\hat{\theta}_i = \hat{\theta}_i\gamma^i$, so $\delta = \gamma^i$ for every i such that $\hat{\theta}_i \neq 0$. If $\hat{F}$ has terms of distinct degrees i and j, then $\gamma^{i-j} = 1$; letting s be the greatest common divisor of the set of differences between degrees of two terms of $\hat{F}$, it follows that $\gamma^s = 1$, and further $\hat{F} \in X^r K[X^s]$ for some $r \geq 0$ such that $\delta = \gamma^r$. If $\hat{F}(X) = \hat{\theta}_d X^d$ then we take $s = 0$ and $r = d$, so again $\delta = \gamma^r$ and $\gamma^s = 1$ and $\hat{F} \in X^r K[X^s]$. The result follows. □

Remark. The first reference we know for this result is [1] (for $K = \mathbb{C}$).

3. The number field case

In this section we prove the number field version of Theorem 1.1. Our proof relies on Proposition 3.3, which will be proved in the next section. We begin with two lemmas applying the results of the previous section to the present context.

Lemma 3.1. *Let K be a field of characteristic zero. Suppose $F, H, E, \tilde{E} \in K[X] \setminus K$ and linear $a, b, c, \tilde{a}, \tilde{b}, \tilde{c} \in K[X]$ satisfy*

$$\begin{aligned} F &= E \circ H \circ a \\ G &= E \circ c \circ H \circ b \\ F^t &= \tilde{E} \circ H \circ \tilde{a} \\ G^t &= \tilde{E} \circ \tilde{c} \circ H \circ \tilde{b} \end{aligned}$$

for some integer $t > 1$. Then there is a linear $e \in K[X]$ such that $F^{t-1} = G^{t-1} \circ e$.

Proof. We have

$$F^{t-1} \circ E \circ H \circ a = F^t = \tilde{E} \circ H \circ \tilde{a} \quad \text{and}$$
$$G^{t-1} \circ E \circ c \circ H \circ b = G^t = \tilde{E} \circ \tilde{c} \circ H \circ \tilde{b}.$$

By Lemma 2.3, there are linears $\ell_1, \ell_2 \in K[X]$ such that

$$H \circ a = \ell_1 \circ H \circ \tilde{a} \quad \text{and}$$
$$c \circ H \circ b = \ell_2 \circ \tilde{c} \circ H \circ \tilde{b}.$$

Thus

$$F^{t-1} \circ E \circ \ell_1 = \tilde{E} = G^{t-1} \circ E \circ \ell_2.$$

Again using Lemma 2.3, there is therefore a linear $e \in K[X]$ such that

$$F^{t-1} = G^{t-1} \circ e,$$

as desired. □

Lemma 3.2. *Let K be a number field, S a finite set of nonarchimedean places of K, and $f, g \in K[X]$ with $\deg(f) = \deg(g) > 1$. Suppose that, for every $k \in \mathbb{N}$, the equation $f^k(X) = g^k(Y)$ has infinitely many solutions in the ring of S-integers of K. Then there exists $r \in \mathbb{N}$ such that, for both*

$n=1$ and infinitely many other values $n \in \mathbb{N}$, there is a linear $\ell_n \in \overline{K}[X]$ such that $f^{rn} = g^{rn} \circ \ell_n$.

Proof. First we show that there exists $r \in \mathbb{N}$ such that $f^r = g^r \circ \ell$ for some linear $\ell \in \overline{K}[X]$. By Corollary 2.2, for each k we have $f^{2^k} = E_k \circ H_k \circ a_k$ and $g^{2^k} = E_k \circ c_k \circ H_k \circ b_k$ with $E_k \in \overline{K}[X]$, linear $a_k, b_k, c_k \in \overline{K}[X]$, and some $H_k \in \overline{K}[X]$ which comes from a finite set of polynomials. Thus, $H_k = H_s$ for some k and s with $k < s$. Applying Lemma 3.1 with $F = f^{2^k}$ and $G = g^{2^k}$ and $t = 2^{s-k}$, it follows that there is a linear $\ell \in \overline{K}[X]$ such that $F^{t-1} = G^{t-1} \circ \ell$, whence $f^r = g^r \circ \ell$ for $r = 2^s - 2^k$.

Suppose there are only finitely many $n \in \mathbb{N}$ for which there is a linear $\ell_n \in \overline{K}[X]$ with $f^{rn} = g^{rn} \circ \ell_n$. Let N be an integer exceeding each of these finitely many integers n. We get a contradiction by applying the previous paragraph with (f^{rN}, g^{rN}) in place of (f, g). □

In the next section we will prove the following proposition.

Proposition 3.3. *Let K be a field of characteristic zero, and let $F, \ell \in K[X]$ satisfy $\deg(F) = d > 1 = \deg(\ell)$. Suppose that, for infinitely many $n > 0$, there is a linear $\ell_n \in K[X]$ such that $F^n = (F \circ \ell)^n \circ \ell_n$. Then either*

(1) *$F^k = (F \circ \ell)^k$ for some $k \in \mathbb{N}$; or*
(2) *$F = v^{-1} \circ \epsilon X^d \circ v$ and $\ell = v^{-1} \circ \delta X \circ v$ for some linear $v \in K[X]$ and some $\epsilon, \delta \in K^*$.*

We now show that this result implies the number field version of Theorem 1.1. Specifically, we prove the following.

Theorem 3.4. *Let K be a number field, let $x_0, y_0 \in K$, and let $f, g \in K[X]$ satisfy $\deg(f) = \deg(g) > 1$. If $\mathcal{O}_f(x_0) \cap \mathcal{O}_g(y_0)$ is infinite, then $f^k = g^k$ for some $k \in \mathbb{N}$.*

Proof. Let S be a finite set of nonarchimedean places of K such that the ring of S-integers $\mathcal{O}_S$ contains x_0, y_0, and every coefficient of f and g. Then $\mathcal{O}_S$ contains every $f^n(x_0)$ and $g^n(y_0)$ with $n \in \mathbb{N}$.

Our hypotheses imply that x_0 is not preperiodic for f, and y_0 is not preperiodic for g. Moreover, for every $k \in \mathbb{N}$, the equation $f^k(x) = g^k(y)$ has infinitely many solutions $(x, y) \in \mathcal{O}_S \times \mathcal{O}_S$.

By Lemma 3.2, there is some $r \in \mathbb{N}$ such that, for both $n = 1$ and infinitely many $n \in \mathbb{N}$, we have $f^{rn} = g^{rn} \circ \ell_n$ with $\ell_n \in \overline{K}[X]$ linear. Put $F = f^r$ and $\ell = \ell_1^{-1}$; then $g^r = F \circ \ell$, and for infinitely many n we have $F^n = (F \circ \ell)^n \circ \ell_n$. If F and $F \circ \ell$ have a common iterate, then so do f and g. By Proposition 3.3, it remains only to consider the case that $F = v^{-1} \circ \epsilon X^d \circ v$ and $\ell = v^{-1} \circ \delta X \circ v$, where $v \in \overline{K}[X]$ is linear and $\epsilon, \delta \in \overline{K}^*$. Note that $d > 1$.

By hypothesis, the set $\mathcal{M}$ of pairs $(m, n) \in \mathbb{N} \times \mathbb{N}$ satisfying $f^m(x_0) = g^n(y_0)$ is infinite, and (from non-preperiodicity) its projections onto each coordinate are injective. Thus, for some $s_1, s_2 \in \mathbb{N}$, the set $\mathcal{M}$ contains infinitely many pairs $(rm + s_1, rn + s_2)$ with $m, n \in \mathbb{N}$; since the projections are injective, $\mathcal{M}$ contains pairs of this form in which $\min(m,n)$ is arbitrarily large. For any $m,n \in \mathbb{N}$ such that $(rm+s_1, rn+s_2) \in \mathcal{M}$, we have $F^m(x_1) = (F \circ \ell)^n(y_1)$, where $x_1 := f^{s_1}(x_0)$ and $y_1 := g^{s_2}(y_0)$. Thus

$$\begin{aligned} v^{-1}(\epsilon^{(d^m-1)/(d-1)} v(x_1)^{d^m}) &= F^m(x_1) \\ &= (F \circ \ell)^n(y_1) \\ &= v^{-1}((\epsilon\delta^d)^{(d^n-1)/(d-1)} v(y_1)^{d^n}), \end{aligned}$$

so

$$v(x_1)^{d^m} \epsilon^{(d^m-d^n)/(d-1)} = \delta^{d(d^n-1)/(d-1)} v(y_1)^{d^n}. \tag{3.1}$$

We cannot have $v(x_1) = 0$, since otherwise $x_1 = f^{s_1}(x_0)$ is a fixed point of $F = f^r$, contrary to our hypotheses. Likewise $v(y_1) \neq 0$. Now let $\epsilon_1, \delta_1 \in \overline{K}$ satisfy $\epsilon_1^{d-1} = \epsilon$ and $\delta_1^{d-1} = \delta^d$, so (3.1) implies

$$\delta_1 = v(x_1)^{-d^m} \cdot \epsilon_1^{d^n-d^m} \cdot \delta_1^{d^n} \cdot v(y_1)^{d^n}. \tag{3.2}$$

Since (3.2) holds for pairs (m,n) with $\min(m,n)$ arbitrarily large, there are infinitely many $k \in \mathbb{N}$ for which δ_1 is a d^k-th power in the number field $K_0 := \mathbb{Q}(v(x_1), v(y_1), \epsilon_1, \delta_1)$. Letting $\mathcal{O}$ be the ring of algebraic integers in K_0, it follows that the fractional ideal of $\mathcal{O}$ generated by δ_1 is a d^k-th power for infinitely many k; now unique factorization of fractional ideals implies δ_1 is in the unit group U of $\mathcal{O}$. Moreover, δ_1 is a d^k-th power in U for infinitely many k; since U is a finitely generated abelian group, δ_1 must be a root of unity whose order N is coprime to d. Thus $N \mid (d^t - 1)$ for some $t \in \mathbb{N}$. Now $(F \circ \ell)^t = v^{-1} \circ (\epsilon\delta^d)^{(d^t-1)/(d-1)} X^{d^t} \circ v$, and since $\delta^d = \delta_1^{d-1}$ and $\delta_1^{d^t-1} = 1$, it follows that $(F \circ \ell)^t = F^t$, as desired. □

4. Proof of Proposition 3.3

In this section we complete the proof of Theorem 3.4, by proving Proposition 3.3. We consider two cases, depending on whether F is gotten from a monomial by composing with linears on both sides. Our strategy is to show in both cases that there are only finitely many linears $\hat{\ell} \in K[X]$ for which there exists n such that $(F \circ \ell)^n \circ \hat{\ell} = F^n$; after this, we pick two values $n < N$ having the same $\hat{\ell}$, and deduce that $F^{N-n} = (F \circ \ell)^{N-n}$.

Lemma 4.1. *Let K be a field of characteristic zero, and suppose $F \in K[X]$ has the property that $u \circ F \circ v$ has at least two monomial terms whenever $u, v \in K[X]$ are linear. Then the equation $F \circ b = a \circ F$ has only finitely many solutions in linear polynomials $a, b \in K[X]$.*

Proof of Lemma 4.1. Our hypothesis implies $\deg(F) > 1$. Pick $\alpha, \beta \in K$ as in Lemma 2.4, and put $\hat{F} := \alpha + F(X + \beta)$; note that these choices depend only on F. Then $\hat{F} \in X^r K[X^s]$ for some integers $r, s \geq 0$. Our hypothesis implies $s \neq 0$; now choose s to be as large as possible. By Lemma 2.4, if $F \circ b = a \circ F$ with $a, b \in K[X]$ linear, then there is an s^{th} root of unity $\gamma \in K$

such that $b = \beta + \gamma(X - \beta)$ and $a = -\alpha + \gamma^r(X + \alpha)$. Since there are only finitely many possibilities for γ, there are only finitely many possibilities for a and b. $\square$

Remark. Our proof shows that the number of solutions is less than $\deg(F)$ (in fact: the number of solutions is at most the size of the largest group of roots of unity in K of order less than $\deg(F)$).

Lemma 4.2. *Let K be a field of characteristic zero, let $u, v, \ell \in K[X]$ be linear, and let $F = u \circ X^d \circ v$ where $d > 1$. The following are equivalent:*

(1) *The equation*

$$F \circ \ell \circ F \circ b = a \circ F \circ F \tag{4.1}$$

has infinitely many solutions in linears $a, b \in K[X]$.

(2) *$F = v^{-1} \circ \epsilon X^d \circ v$ and $\ell = v^{-1} \circ \delta X \circ v$ for some $\epsilon, \delta \in K^*$.*

Proof of Lemma 4.2. Pick any solution (a, b) to (4.1). By Lemma 2.3, there is a linear $c \in K[X]$ such that

$$\ell \circ F \circ b = c \circ F,$$

which implies

$$F \circ c = a \circ F.$$

For any linears $e_1, e_2 \in K[X]$ such that $e_1 \circ F \circ e_2 = F$, we have

$$X^d = (u^{-1} \circ e_1 \circ u) \circ X^d \circ (v \circ e_2 \circ v^{-1}),$$

so $v \circ e_2 \circ v^{-1} = \gamma X$ and $u^{-1} \circ e_1 \circ u = X/\gamma^d$ for some $\gamma \in K^*$. Thus, there exist $\gamma_1, \gamma_2 \in K^*$ such that

$$\begin{aligned} b &= v^{-1} \circ \gamma_1 X \circ v \\ c^{-1} \circ \ell &= u \circ \frac{X}{\gamma_1^d} \circ u^{-1} \\ c &= v^{-1} \circ \gamma_2 X \circ v \\ a^{-1} &= u \circ \frac{X}{\gamma_2^d} \circ u^{-1}. \end{aligned}$$

We can eliminate c from the second and third equations:

$$u \circ \gamma_1^d X \circ u^{-1} = \ell^{-1} \circ c = \ell^{-1} \circ v^{-1} \circ \gamma_2 X \circ v.$$

Thus,

$$\gamma_1^d X = (u^{-1} \circ \ell^{-1} \circ v^{-1}) \circ \gamma_2 X \circ (v \circ u).$$

Write $\alpha := (v \circ u)(0)$. Since $\gamma_1^d X$ fixes 0, the linear polynomial $h := u^{-1} \circ \ell^{-1} \circ v^{-1}$ must map $\gamma_2 \alpha$ to 0. Since α and h do not depend on a and b, it follows that if $\alpha \neq 0$ then γ_2 (and thus γ_1^d) does not depend on a and b, so there are only finitely many possibilities for a and b. Now assume $\alpha = 0$, so 0 is fixed by both $v \circ u$ and $u^{-1} \circ \ell^{-1} \circ v^{-1}$, whence these two linears have the form ϵX and $\hat{\delta} X$ with $\epsilon, \hat{\delta} \in K^*$. Then $u = v^{-1} \circ \epsilon X$ and $\ell^{-1} = v^{-1} \circ \epsilon \hat{\delta} X \circ v$, so $F = v^{-1} \circ \epsilon X^d \circ v$ and (with $\delta = 1/(\epsilon \hat{\delta})$) we have $\ell = v^{-1} \circ \delta X \circ v$.

It remains only to show that, when $F = v^{-1} \circ \epsilon X^d \circ v$ and $\ell = v^{-1} \circ \delta x \circ v$, the number of solutions of (4.1) is infinite. To this end, pick any $\iota \in K^*$, and note that $b = v^{-1} \circ \iota X \circ v$ and $a = v^{-1} \circ \delta^d \iota^{d^2} X \circ v$ satisfy (4.1). □

Remark. This proof shows that, when the number of solutions to (4.1) is finite, this number is at most d (in fact: at most the number of d^{th} roots of unity in K).

Proof of Proposition 3.3. We have

$$(F \circ \ell)^n \circ \ell_n = F^n \tag{4.2}$$

for every n in some infinite subset $\mathcal{M}$ of $\mathbb{N}$. For $n \in \mathcal{M}$, we apply Lemma 2.3 to (4.2) with $B = F \circ \ell \circ \ell_n$ and $D = F$, to conclude that there is a linear $u_n \in K[X]$ such that

$$F \circ \ell \circ \ell_n = u_n \circ F.$$

By Lemma 4.1, if F is not gotten from a monomial by composing with linears on both sides, then $\{\ell_n : n \in \mathcal{M}\}$ is finite.

Next, for $n \in \mathcal{M}$ with $n > 1$, apply Lemma 2.3 to (4.2) with $B = (F \circ \ell)^2 \circ \ell_n$ and $D = F^2$, to conclude that there is a linear $v_n \in K[X]$ such that

$$(F \circ \ell)^2 \circ \ell_n = v_n \circ F^2.$$

By Lemma 4.2, if F is gotten from a monomial by composing with linears on both sides, then either $\{\ell_n : n \in \mathcal{M}\}$ is finite or conclusion (2) of Proposition 3.3 holds.

Thus, whenever (2) of Proposition 3.3 does not hold, the set $\{\ell_n : n \in \mathcal{M}\}$ is finite, so there exist $n, N \in \mathcal{M}$ such that $\ell_n = \ell_N$ and $n < N$. Then

$$\begin{aligned} F^{N-n} \circ F^n &= F^N \\ &= (F \circ \ell)^N \circ \ell_n \\ &= (F \circ \ell)^{N-n} \circ (F \circ \ell)^n \circ \ell_n \\ &= (F \circ \ell)^{N-n} \circ F^n, \end{aligned}$$

so $F^{N-n} = (F \circ \ell)^{N-n}$, as desired. □

5. Some reductions

In this section we show that it suffices to prove Theorems 1.4 and 1.5 in case K is a finitely generated extension of $\mathbb{Q}$. Moreover, for any such K, it suffices to prove these results in case $\deg(f) = \deg(g) > 1$ and the line is the diagonal, $X = Y$.

We begin with the first reduction. For fixed K, f, g, x_0, y_0, L, only finitely many elements of K occur as coefficients of f or g, as values x_0 or y_0, or in the defining equation for L. Let K_0 be the extension of $\mathbb{Q}$ generated by these finitely many elements. Then Theorem 1.4 holds for (K, f, g, x_0, y_0, L) if it holds for (K_0, f, g, x_0, y_0, L), and likewise for Theorem 1.5.

We next show that we need only consider the case that the line is the diagonal.

Lemma 5.1. *If Theorem 1.4 is true for the line $X = Y$, then it is true for every line.*

Proof. If L has the form $X = \alpha$ then the theorem is obvious: if there are infinitely many n such that $f^n(x_0) = \alpha$, then α is periodic point for f, so $X = \alpha$ is a periodic line for (f, g). Likewise the result is clear if L has the form $Y = \beta$, so we may assume L is $X = \ell(Y)$ with $\ell \in K[Y]$ of degree one. Suppose $\{(f^n(x_0), g^n(y_0)) : n \in \mathbb{N}\}$ has infinite intersection with L. If $f^n(x_0) = \ell(g^n(y_0))$ then $f^n(x_0) = (\ell \circ g \circ \ell^{-1})^n(\ell(y_0))$. Thus, assuming Theorem 1.4 for the line $X = Y$, we conclude that $X = Y$ is periodic under the action of $(f, \ell \circ g \circ \ell^{-1})$; it follows that $X = \ell(Y)$ is periodic under the (f, g)-action. $\square$

The analogous result for Theorem 1.5 follows from a similar argument.

Lemma 5.2. *If Theorem 1.5 is true in case $\alpha = 1$ and $\beta = 0$, then it is true for arbitrary α and β.*

We now prove Theorem 1.4 in case f and g are linear polynomials. As noted in the introduction, Theorem 1.5 fails in this case, so our proof must necessarily distinguish between the equations $f^n(x_0) = g^n(y_0)$ and $f^m(x_0) = g^n(y_0)$.

Proposition 5.3. *Theorem 1.4 holds if $\deg(f) = \deg(g) = 1$ and L is the diagonal.*

Proof. Suppose the hypotheses of Theorem 1.4 hold. As above, we may assume $K \subseteq \mathbb{C}$. By replacing x_0 and y_0 with $f^{n_0}(x_0)$ and $g^{n_0}(y_0)$ (for some $n_0 \in \mathbb{N}$), we may assume $x_0 = y_0$. Let $f(X) = \alpha X + \beta$ and $g(X) = \gamma X + \delta$. Note that α cannot be a root of unity different from 1, for otherwise some iterate of f would be the identity map, contradicting infinitude of $\{f^n(x_0) : n \in \mathbb{N}\}$. Likewise, γ is not a root of unity different from 1. We consider two cases:

Case 1. Neither α nor γ equals 1.

For $n \in \mathbb{N}$, we have $f^n(x_0) = \alpha^n \hat{x}_0 - \frac{\beta}{\alpha-1}$ and $g^n(x_0) = \gamma^n \hat{y}_0 - \frac{\delta}{\gamma-1}$, where $\hat{x}_0 := x_0 + \frac{\beta}{\alpha-1}$ and $\hat{y}_0 := x_0 + \frac{\delta}{\gamma-1}$. Since x_0 is not preperiodic for f or g, both $\hat{x}_0$ and $\hat{y}_0$ are nonzero. By the hypothesis of Theorem 1.4, there are infinitely many $n \in \mathbb{N}$ such that $\alpha^n \hat{x}_0 - \gamma^n \hat{y}_0 = \hat{x}_0 - \hat{y}_0$. If $\hat{x}_0 \neq \hat{y}_0$, we may divide through and obtain infinitely many n such that

$$\hat{a}\alpha^n + \hat{b}\gamma^n = 1$$

for some constants $\hat{a}$ and $\hat{b}$. As noted by Lang [12, p. 28], this is impossible (as can be seen by passing to a curve $\hat{a}\alpha^i t^3 + \hat{b}\gamma^i u^3 = 1$, with $0 \le i \le 2$, and using Siegel's theorem on integral points). Hence, we must have $\hat{x}_0 = \hat{y}_0$, so there are infinitely many $n \in \mathbb{N}$ for which $\alpha^n = \gamma^n$, and $f^n = g^n$ for each such n.

Case 2. Either α or γ equals 1.

Without loss of generality, we may assume $\alpha = 1$. If also $\gamma = 1$, then since $f^n(x_0) = g^n(x_0)$ for some $n \in \mathbb{N}$, we must have $\beta = \delta$, so $f = g$ as desired. Now assume $\gamma \neq 1$. Then $g^n(x_0) = \gamma^n \left(x_0 + \frac{\delta}{\gamma-1}\right) - \frac{\delta}{\gamma-1}$. Since $\{g^n(x_0) : n \in \mathbb{N}\}$ is infinite, we must have $x_0 \neq -\delta/(\gamma - 1)$. By hypothesis, there are infinitely many $n \in \mathbb{N}$ such that

$$x_0 + n\beta = \gamma^n \left(x_0 + \frac{\delta}{\gamma - 1}\right) - \frac{\delta}{\gamma - 1}. \tag{5.1}$$

This is not possible if $|\gamma| > 1$, since then the absolute value of the right side exceeds that of the left side for sufficiently large n. Thus $|\gamma| \leq 1$, so the right side is bounded independently of n, whence also $x_0 + n\beta$ is bounded. This implies $\beta = 0$, so f is the identity map, contradicting the hypothesis that $\{f^n(x_0) : n \in \mathbb{N}\}$ is infinite. □

Remark. We note that the argument used in Case 2 above does not generalize to the setting of Theorem 1.1, since we used in a crucial way that we have only one variable n in (5.1), so the orders of growth of the two sides of (5.1) are different. In fact, the conclusion of Theorem 1.1 is not generally true if f is a monic linear polynomial. For example, let $f(X) = X + 1$ and let $g(X)$ be any nonconstant polynomial with positive integer coefficients. Then for any positive integers x_0 and y_0 such that $g(y_0) > y_0$, the intersection $\mathcal{O}_f(x_0) \cap \mathcal{O}_g(y_0)$ is infinite, since $\mathcal{O}_f(x_0)$ contains every sufficiently large integer. On the other hand, the argument from Case 1 generalizes at once to show that the conclusion of Theorem 1.1 holds when f and g are non-monic linear polynomials.

The remainder of this section is devoted to proving Theorem 1.4 in case K is a number field and $\deg(f) \neq \deg(g)$. We recall some standard terminology: a *global field* is either a number field or a function field of transcendence degree 1 over another field. Any global field E comes equipped with a set M_E of normalized absolute values $||\cdot||_v$ which satisfy a product formula[1]:

$$\prod_{v \in M_E} ||x||_v = 1 \quad \text{for every } x \in E^*.$$

If E is a global field, the logarithmic Weil height of $x \in \overline{E}$ is defined as

$$h(x) = \frac{1}{[E(x) : E]} \cdot \sum_{v \in M_E} \sum_{\substack{w|v \\ w \in M_{E(x)}}} \log \max\{||x||_w, 1\}.$$

We will use the following easy consequence of these definitions (cf. [13, p. 77]).

[1] A 'normalized absolute value' is a power of an absolute value, but might not be an absolute value itself since it might fail the triangle inequality.

Lemma 5.4. *Let E be a global field, and let $\ell \in E[X]$ be a linear polynomial. Then there exists $c_\ell > 0$ such that $|h(\ell(x)) - h(x)| \le c_\ell$ for all $x \in \overline{E}$.*

Definition 5.5. *Let E be a global field, let $f \in E[X]$ with $\deg(f) > 1$, and let $z \in \overline{E}$. The canonical height $\widehat{h}_f(z)$ of z with respect to the morphism $f : \mathbb{P}^1 \longrightarrow \mathbb{P}^1$ is*

$$\widehat{h}_f(z) = \lim_{k\to\infty} \frac{h(f^k(z))}{\deg(f)^k}.$$

This definition is due to Call and Silverman, who proved the existence of the above limit in [6, Thm. 1.1] by using boundedness of $|h(f(x)) - (\deg f)h(x)|$ and a telescoping series argument due to Tate. We will use the following properties of the canonical height.

Lemma 5.6. *Let E be a global field, let $f \in E[X]$ be a polynomial of degree greater than 1, and let $z \in \overline{E}$. Then*

(a) *for each $k \in \mathbb{N}$, we have $\widehat{h}_f(f^k(z)) = \deg(f)^k \cdot \widehat{h}_f(z)$;*
(b) *$|h(z) - \widehat{h}_f(z)|$ is uniformly bounded independently of $z \in \overline{E}$;*
(c) *if E is a number field, z is preperiodic if and only if $\widehat{h}_f(z) = 0$.*

Proof. Part (a) is clear; for (b) and (c) see [6, Thm. 1.1 and Cor. 1.1.1]. ☐

Part (c) of Lemma 5.6 is not true if E is a function field with constant field E_0, since $\widehat{h}_f(z) = 0$ whenever $z \in E_0$ and $f \in E_0[X]$. But these are essentially the only counterexamples in the function field case (cf. Lemma 6.7).

Lemma 5.7. *Let K be a number field, let $f, g \in K[X]$ and let $x_0, y_0 \in K$. If $\mathcal{O}_{(f,g)}((x_0, y_0))$ has infinitely many points on the diagonal, then $\deg(f) = \deg(g) > 0$.*

Proof. The hypothesis implies x_0 (resp., y_0) is not preperiodic for f (resp., g). Thus f and g are nonconstant. Suppose $\deg(f) > \deg(g)$.

Since $\widehat{h}_f(x_0) > 0$ (by Lemma 5.6), there exists $\delta > 0$ such that every sufficiently large k satisfies

$$h(f^k(x_0)) > (\deg f)^k \delta.$$

If $\deg g = 1$, by Lemma 5.4 there exists $c_g > 0$ such that

$$h(g^k(y_0)) \le kc_g + h(y_0)$$

for every k, and for sufficiently large k we have $(\deg f)^k\delta > kc_g + h(y_0)$. If $\deg g > 1$, there exists $\epsilon > 0$ such that every k satisfies

$$h(g^k(y_0)) < (\deg g)^k \epsilon,$$

and since $\deg f > \deg g$ we have $(\deg f)^k\delta > (\deg g)^k\epsilon$ for k sufficiently large. Hence, in either case, for k sufficiently large we have $h(f^k(x_0)) > h(g^k(y_0))$ and thus $f^k(x_0) \ne g^k(y_0)$. ☐

Remark. This proof does not work for function fields, since it relies on Lemma 5.6 (c). However, one can use a different argument to show that Lemma 5.7 is valid for any field K (of any characteristic). In characteristic zero, this is a consequence of Theorem 1.4. One can prove this for general K using arguments similar to those in this paper; the key intermediate result is that, for any $f \in K[X]$ with $\deg(f) > 1$, and any $z \in K$ non-preperiodic for f, there is an absolute value v of K such that $\lim_{n\to\infty} |f^n(z)|_v = +\infty$.

6. The function field case

In this section we prove Theorems 1.4 and 1.5. Our strategy is to 'specialize' every transcendental generator of K to an element of a number field, and then deduce these results from the number field version proved previously (Theorem 3.4). We begin by proving that Theorem 1.4 follows from the existence of a suitable specialization homomorphism.

Proof of Theorem 1.4, assuming existence of a suitable specialization. From the results of the previous section, it suffices to prove Theorem 1.4 in case K is a finitely generated extension of $\mathbb{Q}$, the line L is the diagonal, and $\deg(f) \geq 2$. We will prove Theorem 1.4 by induction on the transcendence degree of $K/\mathbb{Q}$. The base case is Theorem 3.4 and Lemma 5.7. For the inductive step, let E be a subfield of K such that $\operatorname{tr.deg}(K/E) = 1$ and $E/\mathbb{Q}$ is finitely generated. Suppose in addition that the diagonal is not periodic under the (f, g) action (i.e., there is no $k \in \mathbb{N}$ for which $f^k = g^k$), and that the set $\{(f^n(x_0), g^n(y_0)) : n \in \mathbb{N}\}$ has infinite intersection with the diagonal. Assume there is a subring R of K, a finite extension E' of E, and a homomorphism $\alpha : R \to E'$, such that

(1) R contains x_0, y_0, and every coefficient of f and g, but the leading coefficients of f and g have nonzero image under α;
(2) $f_\alpha^k \neq g_\alpha^k$ for each $k \in \mathbb{N}$;
(3) $x_{0,\alpha}$ is not preperiodic for f_α.

(Here f_α, g_α, and $x_{0,\alpha}$ denote the images of f, g, and x_0, respectively, under the homomorphism α.)

Properties (1) and (3) show that $\{(f_\alpha^n(x_{0,\alpha}), g_\alpha^n(y_{0,\alpha})) : n \in \mathbb{N}\}$ has infinite intersection with the diagonal. The inductive hypothesis implies $f_\alpha^k = g_\alpha^k$ for some $k \in \mathbb{N}$, which contradicts property (2). Theorem 1.4 follows. $\square$

The proof of Theorem 1.5 is nearly identical to the proof of Theorem 1.4, the only difference being that we replace the set $\{(f^n(x_0), g^n(y_0)) : n \in \mathbb{N}\}$ with $\mathcal{O}_f(x_0) \times \mathcal{O}_g(y_0)$.

To explain why there exists an α as in the proof of Theorem 1.4, we recall the usual setup for specialization. By replacing E with a finite extension of E, we may assume E is algebraically closed in K. Let C be a smooth projective curve over E whose function field is K, and let $\pi : \mathbb{P}^1_C \to C$ be the natural fibration. Any $z \in \mathbb{P}^1_K$ gives rise to a section $Z : C \to \mathbb{P}^1$ of π, and for $\alpha \in C(\overline{E})$, we let $z_\alpha := Z(\alpha)$, and let $E(\alpha)$ be the residue field of K at the valuation corresponding to α. In the notation of the previous paragraph, R is the valuation ring for this valuation, E' is $E(\alpha)$, and the homomorphism $R \to E'$ is $z \mapsto z_\alpha$. The polynomial $f \in K[X]$ extends to a rational map (of E-varieties) from $\mathbb{P}^1_C$ to itself, whose generic fiber is f, and whose fiber above any $\alpha \in C$ is f_α. Note that f_α is a morphism of degree $\deg(f)$ from the fiber $(\mathbb{P}^1_C)_\alpha = \mathbb{P}^1_{E(\alpha)}$ to itself whenever the coefficients of f have no poles or zeros at α; hence it is a morphism on $\mathbb{P}^1_{E(\alpha)}$ of degree $\deg(f)$ at all but finitely many α (we call these α places of *good reduction* for f).

Intuitively, we will show that most choices of α satisfy conditions (2) and (3) above (obviously all but finitely many α satisfy (1)).

We will first prove the following result about specializations of polynomials.

Proposition 6.1. *For each $r > 0$, there are at most finitely many $\alpha \in C(\overline{E})$ such that $[E(\alpha) : E] \le r$ and $f_\alpha^k = g_\alpha^k$ for some $k \in \mathbb{N}$.*

Next, letting h_C be the logarithmic Weil height on C associated to a fixed degree-one ample divisor, we will prove the following dynamical analogue of Silverman's specialization result for abelian varieties [21, Thm. C].

Proposition 6.2. *There exists $c > 0$ such that, for $\alpha \in C(\overline{E})$ with $h_C(\alpha) > c$, the point $x_{0,\alpha}$ is not preperiodic for f_α.*

We now show that these two results imply the existence of α satisfying (1)–(3), which in turn implies Theorems 1.4 and 1.5. Let $\phi : C \to \mathbb{P}^1_E$ be any nonconstant rational function, and let $r = \deg(\phi)$. By [13, Prop. 4.1.7], there are positive constants c_1 and c_2 such that for all $P \in \mathbb{P}^1(\overline{E})$, the preimage $\alpha = \phi^{-1}(P)$ satisfies $h_C(\alpha) \ge c_1 h(P) + c_2$. Since there are infinitely many $P \in \mathbb{P}^1(E)$ such that $h(P) > (c - c_2)/c_1$, we thus obtain infinitely many $\alpha \in C(\overline{E})$ such that $h_C(\alpha) > c$ and $[E(\alpha) : E] \le r$. Hence, Propositions 6.1 and 6.2 imply there are infinitely many α satisfying (2) and (3), and all but finitely many of these satisfy (1) as well.

6.1. **Polynomials with a common iterate.** In this section we prove Proposition 6.1.

Our proof relies on a classical result of Ritt [19, p. 356] describing the pairs of complex polynomials having a common iterate, i.e., $F^n = G^m$ for some $n, m \in \mathbb{N}$. We only need this for $n = m$, in which case Ritt's result is as follows.

Proposition 6.3. *Let $F, G \in \mathbb{C}[X]$ with $d := \deg(F) > 1$. For $n \in \mathbb{N}$, we have $F^n = G^n$ if and only if $F(x) = -\beta + \gamma H(x + \beta)$ and $G(x) = -\beta + H(x + \beta)$ for some $\gamma \in \mathbb{C}^*$, $\beta \in \mathbb{C}$ and $H \in x^r\mathbb{C}[x^s]$ (with $r, s \ge 0$) such that $\gamma^s = 1$ and $\gamma^{(d^n-1)/(d-1)} = 1$.*

Corollary 6.4. *Let K be a field of characteristic zero, and let N_K be the number of roots of unity in K. Let $F, G \in K[X]$ satisfy $\deg(F) = d > 1$ and $F^k = G^k$ for some $k \in \mathbb{N}$. Then $F^n = G^n$ for some n with $1 \le n \le N_K$.*

Proof of Corollary 6.4. Let K_0 be the subfield of K generated by the coefficients of F and G. Then K_0 is a finitely generated extension of $\mathbb{Q}$, so K_0 is isomorphic to a subfield of $\mathbb{C}$. After identifying K_0 with its image in $\mathbb{C}$, Proposition 6.3 implies that $F = -\beta + \gamma H(x+\beta)$ and $G = -\beta + H(x+\beta)$ for some $\gamma \in \mathbb{C}^*$, $\beta \in \mathbb{C}$, and $H \in x^r\mathbb{C}[x^s]$ (with $r, s \geq 0$) such that $\gamma^s = 1$. Moreover, for $n \in \mathbb{N}$ we have $F^n = G^n$ if and only if $\gamma^{(d^n-1)/(d-1)} = 1$. Since γ is the ratio of the leading coefficients of F and G, we see that $\gamma \in K_0^*$. Since $\gamma^{(d^k-1)/(d-1)} = 1$, the multiplicative order m of γ is coprime to d. Note that $m \leq N_K$.

Let p be a prime factor of m, and let p^t be the maximal power of p dividing m. If $p \nmid (d-1)$ then let q_p be the order of d in $(\mathbb{Z}/p^t)^*$; otherwise, put $q_p = p^t$. Then $n := \prod q_p$ satisfies $n \leq m$ and $m \mid (d^n-1)/(d-1)$, whence $n \leq N_K$ and $F^n = G^n$. □

Proof of Proposition 6.1. Pick a point α on C such that $[E(\alpha) : E] \leq r$ and $f_\alpha^k = g_\alpha^k$ for some $k \in \mathbb{N}$. Let N_α be the number of roots of unity in $E(\alpha)$. By Corollary 6.4, the least $n \in \mathbb{N}$ with $f_\alpha^n = g_\alpha^n$ satisfies $n \leq N_\alpha$. Now, N_α is bounded in terms of the degree $[E(\alpha) \cap \overline{\mathbb{Q}} : \mathbb{Q}]$, which is at most $r \cdot [E \cap \overline{\mathbb{Q}} : \mathbb{Q}]$; since E is finitely generated, the latter number is finite, so there is a finite bound on n which depends only on E and r (and not on α).

For any fixed $n \in \mathbb{N}$, we have $f^n \neq g^n$, so $\deg(f_\alpha^n - g_\alpha^n) = \deg(f^n - g^n) \geq 0$ for all but finitely many $\alpha \in C$. The result follows. □

6.2. **Specialization of non-preperiodic points.** In this section we prove Proposition 6.2.

First note that E is a global field. The key ingredient in our proof is the following result of Call and Silverman [6, Thm. 4.1], which relates h_C to the canonical heights $\widehat{h}_f : \overline{K} \to \mathbb{R}_{\geq 0}$ and $\widehat{h}_{f_\alpha} : \overline{E} \to \mathbb{R}_{\geq 0}$ of f and f_α (cf. Definition 5.5).

Lemma 6.5. *For each $z \in K$ we have*

$$\lim_{h_C(\alpha)\to\infty} \frac{\widehat{h}_{f_\alpha}(z_\alpha)}{h_C(\alpha)} = \widehat{h}_f(z). \tag{6.1}$$

We will also use a result about canonical heights of non-preperiodic points for polynomials that are not *isotrivial*.

Definition 6.6. *We say a polynomial $f \in K[X]$ is isotrivial if there exists a finite extension K' of K and a linear $\ell \in K'[X]$ such that $\ell^{-1} \circ f \circ \ell \in \overline{E}[X]$.*

Benedetto proved that a non-isotrivial polynomial can only have canonical height equal to 0 at its preperiodic points [2, Thm. B]:

Lemma 6.7. *Let $f \in K[X]$ with $\deg(f) \geq 2$, and let $z \in \overline{K}$. If f is not isotrivial, then $\widehat{h}_f(z) = 0$ if and only if z is preperiodic for f.*

We need one more preliminary result.

Lemma 6.8. *Let $f \in K[X]$ be isotrivial with $\deg(f) \geq 2$, and let ℓ be as in Definition 6.6. If $z \in \overline{K}$ satisfies $\widehat{h}_f(z) = 0$, then $\ell^{-1}(z) \in \overline{E}$.*

Proof. Put $F := \ell^{-1} \circ f \circ \ell \in K'[X]$, so $F^n(\ell^{-1}(z)) = \ell^{-1}(f^n(z))$. Since $\widehat{h}_f(z) = 0$, Lemma 5.4 implies that $\widehat{h}_F(\ell^{-1}(z)) = 0$. For any $v \in M_{K'(z)}$, we know that every nonzero coefficient γ of F satisfies $||\gamma||_v = 1$ (since $\gamma \in \overline{E}$). Since v is nonarchimedean, if $y \in K'(z)$ satisfies $||y||_v > 1$ then $\log ||F^n(y)||_v = \deg(F)^n \log ||y||_v$, so $\widehat{h}_F(y) > 0$. Thus $||\ell^{-1}(z)||_v \leq 1$ for every $v \in M_{K'(z)}$, so $\ell^{-1}(z) \in \overline{E}$. □

Proof of Proposition 6.2. Put $z = x_0$. If $\widehat{h}_f(z) > 0$ then, by Lemma 6.5, there exists $c > 0$ such that every $\alpha \in C(\overline{E})$ with $h_C(\alpha) > c$ satisfies

$$\frac{\widehat{h}_{f_\alpha}(z_\alpha)}{h_C(\alpha)} > 0.$$

Then $\widehat{h}_{f_\alpha}(z_\alpha) > 0$, so part (a) of Lemma 5.6 implies z_α is not preperiodic for f_α.

If f is not isotrivial, Lemma 6.7 implies $\widehat{h}_f(z) > 0$, so the proof is complete. It remains only to consider the case that f is isotrivial and $\widehat{h}_f(z) = 0$.

Pick a finite extension K' of K and a linear $\ell \in K'[X]$ such that $g := \ell^{-1} \circ f \circ \ell$ is in $\overline{E}[X]$, and put $E' := \overline{E} \cap K'$. Lemma 6.8 implies $w := \ell^{-1}(z)$ is in E'. Moreover, since $\ell^{-1} \circ f^n(z) = g^n(w)$ and z is not preperiodic for f, we see that w is not preperiodic for g. Because $g \in E'[X]$ and $w \in E'$, then for all places α' of K', the reductions of g and w at α' equal g, and respectively w (because E' embeds naturally into the residue field at α'). Hence, for all but finitely many α' (we only need to exclude the places where ℓ does not have good reduction), if α is the place of K lying below α', then z_α is not preperiodic for f_α.

□

7. Further conjectures

We suspect that Theorem 1.5 remains true without the hypothesis that $\deg(f) = \deg(g)$. It might be possible to prove this by methods similar to those in this paper; however, this seems to require substantial effort, since the results of Bilu-Tichy and Ritt which we used became much simpler in our case $\deg(f) = \deg(g)$.

It would be interesting to study Conjecture 1.3 for other curves in the plane. In particular, it may be possible to treat curves of the form $F(X) = G(Y)$ (with F, G polynomials) by methods similar to ours.

References

[1] G. af Hällström, *Über halbvertauschbare Polynome*, Acta Acad. Abo. **21** (1957), no. 2, 20 pp.

[2] R. Benedetto, *Heights and preperiodic points of polynomials over function fields*, Int. Math. Res. Not. **62** (2005), 3855–3866.

[3] Yu. Bilu, B. Brindza, P. Kirschenhofer, Á. Pintér and R. F. Tichy, *Diophantine equations and Bernoulli polynomials*, Compositio Math. **131** (2002), 173–188.

[4] Yu. F. Bilu, M. Kulkarni and B. Sury, *The Diophantine equation* $x(x+1)\dots(x+(m-1))+r=y^n$, Acta Arith. **113** (2004), 303–308.

[5] Y. F. Bilu and R. F. Tichy, *The Diophantine equation* $f(x)=g(y)$, Acta Arith. **95** (2000), 261–288.

[6] G. S. Call and J. H. Silverman, *Canonical heights on varieties with morphisms*, Compositio Math. **89** (1993), 163–205.

[7] L. Denis, *Géométrie diophantienne sur les modules de Drinfeld*, in: The Arithmetic of Function Fields (Columbus, OH, 1991), 285–302, de Gruyter, Berlin, 1992.

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

[9] D. Ghioca and T. J. Tucker, *A dynamical version of the Mordell-Lang conjecture for the additive group*, Compositio Math., to appear, 14 pages. (arXiv:0704.1333 [math.NT])

[10] D. Ghioca and T. J. Tucker, *p-adic logarithms for polynomial dynamics*, submitted for publication, 11 pages, arXiv:0705.4047 [math.NT].

[11] D. Ghioca and T. J. Tucker, *Mordell-Lang and Skolem-Mahler-Lech theorems for endomorphisms of semiabelian varieties*, submitted for publication, 14 pages.

[12] S. Lang, *Integral points on curves*, Publ. Math. IHES **6** (1960), 27–43.

[13] S. Lang, *Fundamentals of Diophantine geometry*, Springer-Verlag, New York, 1983.

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

[15] H. Levi, *Composite polynomials with coefficients in an arbitrary field of characteristic zero*, Amer. J. Math. **64** (1942), 389–400.

[16] P. Müller and M. Zieve, *Ritt's theorems on polynomial decomposition*, in preparation.

[17] M. Raynaud, *Courbes sur une variété abélienne et points de torsion*, Invent. Math. **71** (1983), 207–233.

[18] M. Raynaud, *Sous-variétés d'une variété abélienne et points de torsion*, in: Arithmetic and Geometry, Vol. I, Progr. Math. **35**, 327–352, Birkhauser, Boston, 1983.

[19] J. F. Ritt, *On the iteration of rational functions*, Trans. Amer. Math. Soc. **21** (1920), 348–356.

[20] J. F. Ritt, *Prime and composite polynomials*, Trans. Amer. Math. Soc. **23** (1922), 51–66.

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

[22] A. Schinzel, Polynomials with Special Regard to Reducibility, Cambridge University Press, 2000.

[23] Th. Stoll and R. F. Tichy, *Diophantine equations involving general Meixner and Krawtchouk polynomials*, Quaest. Math. **28** (2005), 105–115.

[24] U. Ullmo, *Positivité et discrétion des points algébriques des courbes*, Ann. of Math. (2) **147** (1998), 167–179.

[25] P. Vojta, *Integral points on subvarieties of semiabelian varieties. I*, Invent. Math. **126** (1996), 133–181.

[26] S. Zhang, *Equidistribution of small points on abelian varieties*, Ann. of Math. (2) **147** (1998), 159–165.

[27] S. Zhang, *Distributions in algebraic dynamics*, A tribute to Professor S.-S. Chern, Survey in Differential Geometry **10** (2006), 381–430.

Dragos Ghioca, Department of Mathematics & Computer Science, University of Lethbridge, 4401 University Drive, Lethbridge, Alberta T1K 3M4, Canada

E-mail address: dragos.ghioca@uleth.ca

Thomas Tucker, Department of Mathematics, Hylan Building, University of Rochester, Rochester, NY 14627, USA

E-mail address: ttucker@math.rochester.edu

Michael E. Zieve, Center for Communications Research, 805 Bunn Drive, Princeton, NJ 08540, USA

E-mail address: zieve@idaccr.org

URL: http://www.math.rutgers.edu