THE DYNAMICAL MANIN-MUMFORD CONJECTURE AND THE DYNAMICAL BOGOMOLOV CONJECTURE FOR SPLIT RATIONAL MAPS

D. GHIOCA, K. D. NGUYEN, AND H. YE

ABSTRACT. We prove the Dynamical Bogomolov Conjecture for endomorphisms $\Phi : \mathbb{P}^1 \times \mathbb{P}^1 \longrightarrow \mathbb{P}^1 \times \mathbb{P}^1$, where $\Phi(x, y) := (f(x), g(y))$ for any rational functions f and g defined over $\bar{\mathbb{Q}}$. We use the equidistribution theorem for points of small height with respect to an algebraic dynamical system, combined with a theorem of Levin regarding symmetries of the Julia set. Using a specialization theorem of Yuan and Zhang, we can prove the Dynamical Manin-Mumford Conjecture for endomorhisms $\Phi = (f, g)$ of $\mathbb{P}^1 \times \mathbb{P}^1$, where f and g are rational functions defined over an arbitrary field of characteristic 0.

1. INTRODUCTION

We prove the Dynamical Manin-Mumford Conjecture (over $\mathbb{C}$) and the Dynamical Bogomolov Conjecture (over $\bar{\mathbb{Q}}$) for endomorphisms Φ of $\mathbb{P}^1 \times \mathbb{P}^1$ (see [Zha06, Conjectures 1.2.1 and 4.1.7]). Actually, we can prove an even stronger result than the one conjectured in [Zha06] since we do not assume $\Phi := (f_1, f_2) : \mathbb{P}^1 \times \mathbb{P}^1 \longrightarrow \mathbb{P}^1 \times \mathbb{P}^1$ is necessarily polarizable (i.e., f_1 and f_2 might have different degrees), but we exclude the case when the f_i's are conjugate to monomials, $\pm$Chebyshev polynomials, or Lattès maps since in those cases there are counterexamples to a formulation which does not ask that $\deg(f_1) = \deg(f_2)$ (see Remark 1.2). In Theorem 1.3 we prove the appropriately modified statement of the Dynamical Manin-Mumford Conjecture for all polarizable endomorphisms of $\mathbb{P}^1 \times \mathbb{P}^1$.

First, we recall that the Chebyshev polynomial of degree d is the unique polynomial T_d with the property that for each $z \in \mathbb{C}$, we have $T_d(z+1/z) = z^d + 1/z^d$. Similarly, a Lattès map $f : \mathbb{P}^1 \longrightarrow \mathbb{P}^1$ is a rational function coming from the quotient of an affine map $L(z) = az + b$ on a torus $\mathcal{T}$ (elliptic curve), i.e. $f = \Theta \circ L \circ \Theta^{-1}$ with $\Theta : \mathcal{T} \to \mathbb{P}^1$ a finite-to-one holomorphic map; see [Mil04] by Milnor. For two rational functions f and g, we say they are (linearly) conjugate if there exists an automorphism η of $\mathbb{P}^1$ such that $f = \eta^{-1} \circ g \circ \eta$; we note that, by definition, the class of Lattès maps is invariant under conjugation. Any rational map of degree $d > 1$ which is conjugate either to $z^{\pm d}$, or to $\pm T_d(z)$, or to a Lattès map is called

---

exceptional. Finally, we note that two polynomials f and g are linearly conjugate if there exists a linear polynomial η such that $f = \eta^{-1} \circ g \circ \eta$.

Theorem 1.1. *Let f_1 and f_2 be rational functions and let $C \subset \mathbb{P}^1 \times \mathbb{P}^1$ be an irreducible curve defined over $\mathbb{C}$ which projects dominantly onto both coordinates. Assume $d_i := \deg(f_i) > 1$ for $i = 1, 2$, and also that for some $i = 1, 2$ we have that $f_i(z)$ is not exceptional. Then C contains infinitely many points (x_1, x_2) such that x_i is a preperiodic point for f_i for $i = 1, 2$ if and only if there exist positive integers ℓ_1 and ℓ_2 such that the following two conditions are met:*

- (i) $d_1^{\ell_1} = d_2^{\ell_2}$; *and*
- (ii) C *is preperiodic under the action of $(x, y) \mapsto (f_1^{\ell_1}(x), f_2^{\ell_2}(y))$ on $\mathbb{P}^1 \times \mathbb{P}^1$.*

Moreover, if f_1, f_2 and C are defined over $\bar{\mathbb{Q}}$, then there exist positive integers ℓ_1 and ℓ_2 such that conditions (i)-(ii) are met if and only if there exists an infinite sequence of points $(x_{1,n}, x_{2,n}) \in C(\bar{\mathbb{Q}})$ such that $\lim_{n \to \infty} \widehat{h}_{f_1}(x_{1,n}) + \widehat{h}_{f_2}(x_{2,n}) = 0$, where $\widehat{h}_{f_i}$ is the canonical height with respect to the rational function f_i.

Motivated by the classical Manin-Mumford conjecture (proved by Raynaud [Ray83] in the case of abelian varieties and by McQuillan [McQ95] in the general case of semiabelian varieties) and also by the classical Bogomolov conjecture (proved by Ullmo [Ull98] in the case of curves embedded in their Jacobian and by Zhang [Zha98] in the general case of abelian varieties), Zhang formulated dynamical analogues of both conjectures (see [Zha06, Conjecture 1.2.1, Conjecture 4.1.7]) for polarizable endomorphisms of any projective variety. We say that an endomorphism Φ of a projective variety X is *polarizable* if there exists an ample line bundle $\mathcal{L}$ on X such that $\Phi^* \mathcal{L}$ is linearly equivalent to $\mathcal{L}^{\otimes d}$ for some integer $d > 1$. As initially conjectured by Zhang, one might expect that if X is defined over a field K of characteristic 0 and Φ is a polarizable endomorphism of X, then a subvariety $V \subseteq X$ contains a Zariski dense set of preperiodic points if and only if V is preperiodic. We say that an irreducible subvariety $Y \subseteq X$ is *preperiodic* if $\Phi^m(Y) = \Phi^n(Y)$ for some integers $n > m \geq 0$; if $m = 0$, then we say that Y is *periodic*. Furthermore if K is a number field then one can construct the canonical height $\widehat{h}_\Phi$ for all points in $X(\bar{\mathbb{Q}})$ with respect to the action of Φ (see [CS93]) and then Zhang's dynamical version of the Bogomolov Conjecture asks that if a subvariety $V \subseteq X$ is not preperiodic, then there exists $\epsilon > 0$ with the property that the set of points $x \in V(\bar{\mathbb{Q}})$ such that $\widehat{h}_\Phi(x) < \epsilon$ is not Zariski dense in V. Since all preperiodic points have canonical height equal to 0, the Dynamical Bogomolov Conjecture is a generalization of the Dynamical Manin-Mumford Conjecture when the algebraic dynamical system (X, Φ) is defined over a number field.

Besides the case of abelian varieties X endowed with the multiplication-by-2 map Φ (which motivated Zhang's conjectures), there are known only a

handful of special cases of the Dynamical Manin-Mumford or the Dynamical Bogomolov conjectures. All of these partial results are for endomorphisms of $\mathbb{P}^1 \times \mathbb{P}^1$—see [BH05, GT10, GTZ11] (also see [DF16] for a proof of a version of the Dynamical Manin-Mumford Conjecture in the context of polynomial automorphisms of $\mathbb{A}^2$). With the exception of [DF16], all of the above cited papers deal only with the case of linear subvarieties V of $\mathbb{P}^1 \times \mathbb{P}^1$. Our Theorem 1.1 settles the Dynamical Manin-Mumford and the Dynamical Bogomolov conjectures for all subvarieties of $\mathbb{P}^1 \times \mathbb{P}^1$.

Remark 1.2. A direct computation shows that any dominant endomorphism of $\mathbb{P}^1 \times \mathbb{P}^1$ is always of the form $(x_1, x_2) \mapsto (f_1(x_1), f_2(x_2))$ (up to a switch of $f_1(x_1)$ and $f_2(x_2)$) where f_i are rational functions for $i = 1, 2$. On the other hand, for any dominant endomorphism Φ of $\mathbb{P}^1 \times \mathbb{P}^1$, we have that Φ^2 is indeed of the form $(x_1, x_2) \mapsto (f_1(x_1), f_2(x_2))$. Since Φ and Φ^2 share the same set of preperiodic points (and also the same set of points of small height since they induce the same dynamical system on $\mathbb{P}^1 \times \mathbb{P}^1$), then we can restrict our attention only to endomorphisms of the form $(x_1, x_2) \mapsto (f_1(x_1), f_2(x_2))$. Furthermore, such an endomorphism is polarizable if and only if f_1 and f_2 have the same degree $d > 1$.

In the case when the maps f_i are both conjugate to monomials or $\pm$Chebyshev polynomials, then there are counterexamples to a Dynamical Manin-Mumford statement unless $\deg(f_1) = \deg(f_2)$. In the case the maps f_i are both Lattès maps, there are counterexamples to the original formulation of the Dynamical Manin-Mumford Conjecture even when $\deg(f_1) = \deg(f_2)$—see [GTZ11] and the subsequent reformulation of the Dynamical Manin-Mumford Conjecture from [GTZ11, Conjecture 2.4] which takes into account those counterexamples. Intuitively, the problem with any exceptional rational map f lies with the fact that there are *more than usual* maps g commuting with f; this occurs since the dynamical system $(\mathbb{P}^1, f)$ is covered in the three cases of an exceptional rational map by another dynamical system (G, Φ), where G is an algebraic group ($\mathbb{G}_m$ in the case f is a monomial or a $\pm$Chebyshev polynomial, and an elliptic curve when f is a Lattès map), while Φ is an endomorphism of G. However, when the maps f_i are not exceptional, then one expected that the condition on the degrees of the maps f_i might be weakened—this is exactly what we prove in our Theorem 1.1.

Using Theorem 1.1, we prove [GTZ11, Conjecture 2.4] for all endomorphisms of $\mathbb{P}^1 \times \mathbb{P}^1$.

Theorem 1.3. *Let f_1 and f_2 be rational functions of degree $d > 1$ defined over $\mathbb{C}$, let $\Phi : \mathbb{P}^1 \times \mathbb{P}^1 \longrightarrow \mathbb{P}^1 \times \mathbb{P}^1$ be defined by $\Phi(x_1, x_2) = (f_1(x_1), f_2(x_2))$, and let $C \subset \mathbb{P}^1 \times \mathbb{P}^1$ be an irreducible curve. Then C is preperiodic under the action of Φ if and only if there exist infinitely many smooth points $P = (x_1, x_2) \in C$ which are preperiodic under Φ and moreover such that the tangent space of C at P is preperiodic under the induced action of Φ on $\mathrm{Gr}_1\left(T_{\mathbb{P}^1 \times \mathbb{P}^1, P}\right)$, where $T_{\mathbb{P}^1 \times \mathbb{P}^1, P}$ is the tangent space of $\mathbb{P}^1 \times \mathbb{P}^1$ at P.*

During the final stages of preparing our paper, we learned that independently (using a slightly different approach), Romain Dujardin, Charles Favre and William Gignac proved Theorem 1.3 in the special case when each f_i is a polynomial.

We also note that using a simple induction argument, one can immediately extend both Theorems 1.1 and 1.3 to the case of curves contained in $(\mathbb{P}^1)^n$ endowed with the coordinatewise action of one-variable rational functions $f_1, \ldots, f_n$. The idea is that for any curve $C \subset (\mathbb{P}^1)^n$, if we let $C_i \subset (\mathbb{P}^1)^{n-1}$ (for $i = 1, \ldots, n$) be the image of C under the projection map $\pi_i : (\mathbb{P}^1)^n \longrightarrow (\mathbb{P}^1)^{n-1}$ on the $(n-1)$ coordinates of $(\mathbb{P}^1)^n$ with the exception of its i-th coordinate, then C is an irreducible component of the intersection $\cap_i \pi_i^{-1}(C_i)$. The inductive hypothesis yields that each C_i and thus $\pi^{-1}(C_i)$ must be preperiodic under the action of the f_i's and therefore C is preperiodic.

The idea of our proof is as follows. Using the equidistribution theorem for points of small height on a variety (see [CL06] for the case of curves and [Yua08] for the case of a higher dimensional variety, and also [BR06] and [FRL06] for the case of $\mathbb{P}^1$), we prove that if $C \subset \mathbb{P}^1 \times \mathbb{P}^1$ contains infinitely many preperiodic points under the action of (f_1, f_2) (or points of small canonical height in the case the dynamical system is defined over $\bar{\mathbb{Q}}$), then the measures $\tilde{\mu}_i$ induced on C by the invariant measures corresponding to the dynamical systems $(\mathbb{P}^1, f_i)$ are equal. Using a careful study of the local analytic maps which preserve (locally) the Julia set of a rational map which is not Lattès, monomial, or Chebyshev polynomial, we obtain the conclusion of Theorem 1.1. Even though all of the previously known results on the Dynamical Manin-Mumford Conjecture (see [BH05, GT10, GTZ11]) used the powerful equidistribution theorem, since those results were valid only when C is a line, one always reduced the question to two rational functions which share the same Julia set. The classification of such pair of rational functions was simple (see [Lev90]). In our case, there is a much weaker consequence of the equidistribution that we obtain: for each point (x_1, x_2) on the curve C, we have that x_1 is preperiodic under the action of f_1 if and only if x_2 is preperiodic under the action of f_2. This consequence was known for quite some time (see [Mim13] which publishes the findings of Mimar's PhD thesis [Mim97] from almost 20 years ago). The novelty of our approach is the exploit of the local symmetries of the Julia set which allows us to settle completely the Dynamical Manin-Mumford and the Dynamical Bogomolov Conjectures for all endomorphisms of $\mathbb{P}^1 \times \mathbb{P}^1$.

Theorem 1.1 yields that if a curve C contains infinitely many points with both coordinates preperiodic under the action of f_1, respectively f_2, then C must be preperiodic. In turn this yields that there are *transversal* (i.e., not horizontal nor vertical) curves in $\mathbb{P}^1 \times \mathbb{P}^1$ which are fixed by the action of $(f_1^{m_1}, f_2^{m_2})$. Hence, the dynamics of f_1 and f_2 are *non-orthogonal* (see [MS14] for the definition of orthogonality in this context). In other words, f_1 and f_2 are related dynamically; our next result shows that in the family

of unicritical polynomials which are not conjugate to a monomial or to a Chebyshev polynomial, each map is non-orthogonal only to maps which are conjugate to itself.

Theorem 1.4. *Let $d_1, d_2 \geq 2$ be integers, let $c_1, c_2 \in \mathbb{C}$, let $f_i(z) = z^{d_i} + c_i$ be polynomials such that $c_i \neq 0$ and also $c_i \neq -2$ if $d_i = 2$, and let Φ be the endomorphism of $\mathbb{P}^1 \times \mathbb{P}^1$ induced by the coordinatewise action of f_1 and f_2. Then there exists a curve $C \subset \mathbb{P}^1 \times \mathbb{P}^1$ which projects dominantly onto both coordinates and also such that C contains infinitely many preperiodic points under the action of Φ if and only if $d_1 = d_2$ and there exists a $(d_1 - 1)$-st root of unity ζ such that $c_2 = \zeta \cdot c_1$.*

We observe that since polynomials f and g of degree $d \geq 2$ in *normal form* (i.e. they are monic and the coefficient of x^{d-1} in both f and g equals 0) are conjugate if and only if there exists a $(d-1)$-st root of unity ζ such that $g(x) = \zeta^{-1} f(\zeta x)$, then a polynomial of the form $x^d + c$ is conjugate to x^d if and only if $c = 0$, and it is conjugate to $\pm T_d(x)$ if and only if $d = 2$ and $c = -2$. This is an immediate consequence of the fact that the coefficient of x^{d-2} in $T_d(x)$ is always nonzero.

The proof of Theorem 1.4 employs the results of [MS14] along additional techniques stemming from the study of polynomial decomposition (see also [Ngu15]). Essentially, now due to our Theorem 1.1 and the classification of invariant curves under polynomial actions done by [MS14], one can relatively easy determine for any two polynomials f_1 and f_2 whether there exists a transversal curve $C \subset \mathbb{P}^1 \times \mathbb{P}^1$ containing infinitely many points (x_1, x_2) with each x_i preperiodic under the action of f_i. Our Theorem 1.4 is an example in this direction that we considered to be interesting by itself and also for its proof, since the family of unicritical polynomials is one of the most studied families of polynomials from the point of view of complex dynamics; e.g. see [AKLS09]. More importantly, lots of the dynamical phenomenons of the family of unicritical polynomials are inherited by general families. We state next another result, which is a consequence of our Theorem 1.3 coupled with [MS14, Theorem 6.24].

Theorem 1.5. *Let $f \in \mathbb{C}[x]$ be a polynomial of degree $d \geq 2$ which is not conjugate to x^d or to $\pm T_d(x)$, and let $\Phi : \mathbb{P}^1 \times \mathbb{P}^1 \longrightarrow \mathbb{P}^1 \times \mathbb{P}^1$ be defined by $\Phi(x, y) = (f(x), f(y))$. Let $C \subset \mathbb{P}^1 \times \mathbb{P}^1$ be a curve defined over $\mathbb{C}$ which projects dominantly onto both coordinates. Then C contains infinitely many preperiodic points under the action of Φ if and only if C is an irreducible component of the zero locus of an equation of the form $\tilde{f}^n(x) = L(\tilde{f}^m(y))$, where $m, n \geq 0$ are integers, L is a linear polynomial commuting with an iterate of f, and $\tilde{f}$ is a non-linear polynomial of minimal degree commuting with an iterate of f.*

The plan of our paper is as follows. In Sections 2 and 3 we setup our notation, state basic properties for the Julia set of a rational function, construct the heights associated to an algebraic dynamical system and define

adelic metrized line bundles which are employed in the main equidistribution result (Theorem 3.1), which we will then use in our proof. In Section 4 we prove that under the hypothesis of Theorem 1.1, the measures induced on the curve C from the two dynamical systems $(\mathbb{P}^1, f_i)$ (for $i = 1, 2$) are equal. In Section 5 we show how to use the equality of these two measures to infer the preperiodicity of the curve. In Section 6 we finalize the proofs for Theorems 1.1 and 1.3. We conclude our paper with Section 7 in which we prove Theorems 1.4 and 1.5.

Acknowledgments. We thank Laura DeMarco, Charles Favre, Holly Krieger, Matt Satriano, Joseph Silverman, Thomas Tucker, Junyi Xie and Shouwu Zhang for useful discussions. We are grateful to both referees for their many useful comments and suggestions which improved our presentation. The first and third authors are partially supported by NSERC and the second author is partially supported by a UBC and PIMS postdoctoral fellowship.

2. Dynamics of a rational function

Let $f : \mathbb{P}^1 \to \mathbb{P}^1$ be a rational function defined over $\mathbb{C}$ of degree $d \geq 2$. As always in dynamics, we denote by f^n the n-th compositional iterate of f.

2.1. **The Julia set.** A family of maps on U is *normal* if for any sequence of maps in this family, we can always pick a subsequence converging locally uniformly on U. The *Julia set* J_f is the set of points on $\mathbb{P}^1_{\mathbb{C}}$, where $\{f^n\}_{n \geq 1}$ is not normal restricted on any of their neighbourhoods. The Julia set J_f is closed, nonempty and invariant under f. Let x be a periodic point in a cycle of exact period n; then the *multiplier* λ of this cycle (or of the periodic point x) is the derivative of f^n at x. A cycle is *repelling* if its multiplier has absolute value greater than 1. All but finitely many cycles of f are repelling, and repelling cycles are in the Julia set J_f. Locally, at a repelling fixed point x with multiplier λ, we can conjugate f to the linear map $z \to \lambda \cdot z$ near $z = 0$ (note that $\lambda \neq 0$ since the point is assumed to be repelling). For more details about the dynamics of a rational function, we refer the reader to Milnor's book [Mil00].

2.2. **The invariant measure.** There is a probability measure μ_f on $\mathbb{P}^1$ associated to f, which is the unique f-invariant measure achieving maximal entropy $\log d$; see [Bro65, Lyu83, FLM83, Man83]. Also μ_f is the unique measure satisfying

$$(1) \qquad\qquad \mu_f(f(A)) = d \cdot \mu_f(A)$$

for any Borel set $A \subset \mathbb{P}^1_{\mathbb{C}}$ with f injective restricted on A. The support of μ_f is J_f, and $\mu_f(x) = 0$ for any $x \in \mathbb{P}^1_{\mathbb{C}}$.

2.3. **Measures on a curve associated to a dynamical system.** Let $C \subset \mathbb{P}^1_\mathbb{C} \times \mathbb{P}^1_\mathbb{C}$ be an irreducible curve projecting dominantly onto both coordinates, i.e., the canonical projections $\pi_i : \mathbb{P}^1 \times \mathbb{P}^1 \longrightarrow \mathbb{P}^1$ (for $i = 1, 2$) restrict to dominant morphisms $(\pi_i)|_C : C \longrightarrow \mathbb{P}^1$. By abuse of notation, we denote the restriction $(\pi_i)_C$ also by π_i. We define probability measures $\tilde{\mu}_{i,f}$ (for $i = 1, 2$) on C corresponding to the dynamical system $(\mathbb{P}^1_\mathbb{C}, f)$. For each $i = 1, 2$, we pullback μ_f by π_i to get a measure $\pi_i^* \mu_f$ on C in the following sense

$$\pi_i^* \mu_f(A) := \mu_f(\pi_i(A))$$

for any Borel set $A \subset C$. Hence we get probability measures on C

$$\tilde{\mu}_{i,f} := \pi_i^* \mu_f / \deg(\pi_i) \text{ for } i = 1, 2.$$

When there is no confusion on what projection π_i we used to construct $\tilde{\mu}_{i,f}$ we drop the index i from our notation and simply use $\tilde{\mu}_f$ to denote the corresponding probability measure on C induced by the dynamical system $(\mathbb{P}^1, f)$.

2.4. **Symmetries of the Julia set.** Let H be a meromorphic function on some disc $B(a, r)$ of radius r centred at a point $a \in J_f$. We say that H is a *symmetry* on J_f if it satisfies

- $x \in B(a, r) \cap J_f$ if and only if $H(x) \in H(B(a, r)) \cap J_f$.
- When J_f is either a circle, line segment, or entire sphere, there is a constant $\alpha > 0$ such that for any Borel set A where $H|_A$ is injective, one has $\mu_f(H(A)) = \alpha \cdot \mu_f(A)$.

A family $\mathcal{H}$ of symmetries of J_f on $B(a, r)$ is said to be *nontrivial* if $\mathcal{H}$ is normal on $B(a, r)$ and no limit function of $\mathcal{H}$ is a constant. A rational function is *post-critically finite* (sometimes called critically finite), if each of its critical points has finite forward orbit, i.e. all critical points are preperiodic. According to Thurston [Thu85, DH93], there is an orbifold structure on $\mathbb{P}^1$ corresponding to each post-critically finite map. A rational function is post-critically finite with *parabolic* orbifold if and only if it is exceptional; or equivalently its Julia set is smooth (circle, line segment or entire sphere) with smooth maximal entropy measure on it; see [DH93]. Actually, according to Hamilton [Ham95], a Julia set which is a one-dimensional topological manifold must be either a circle, closed line segment (up to an automorphism of $\mathbb{P}^1$) or of Hausdorff dimension greater than one. By a theorem of Zdunik [Zdu90], a rational function f is Lattès if and only if J_f is $\mathbb{P}^1$ and μ_f is absolutely continuous with respect to Lebesgue measure on $\mathbb{P}^1$. In [Lev90, Theorem 1], it was shown that there exists an infinite nontrivial family of symmetries on J_f if and only if f is post-critically finite with parabolic orbifold.

3. Adelic metrized line bundles and the equidistribution of small points

In this section, we setup the height functions and state the equidistribution theorem for points of small height, which would be used later in proving the main theorems of this article. The main tool we use here is the arithmetic equidistribution theorem for points with small height on algebraic curves (see [CL06, Yua08]).

3.1. The height functions.
Let K be a number field and $\overline{K}$ be the algebraic closure of K. The number field K is naturally equipped with a set Ω_K of pairwise inequivalent nontrivial absolute values, together with positive integers N_v for each $v \in \Omega_K$ such that

- for each $\alpha \in K^*$, we have $|\alpha|_v = 1$ for all but finitely many places $v \in \Omega_K$.
- every $\alpha \in K^*$ satisfies the *product formula*

$$(2) \qquad \prod_{v \in \Omega_K} |\alpha|_v^{N_v} = 1$$

For each $v \in \Omega_K$, let K_v be the completion of K at v, let $\overline{K}_v$ be the algebraic closure of K_v and let $\mathbb{C}_v$ denote the completion of $\overline{K}_v$. We fix an embedding of $\overline{K}$ into $\mathbb{C}_v$ for each $v \in \Omega_K$; hence we have a fixed extension of $|\cdot|_v$ on $\overline{K}$. When v is archimedean, then $\mathbb{C}_v \cong \mathbb{C}$. Let $f \in K(z)$ be a rational function with degree $d \geq 2$. There is a *canonical height* $\widehat{h}_f$ on $\mathbb{P}^1(\overline{K})$ given by

$$(3) \qquad \widehat{h}_f(x) := \frac{1}{[K(x):K]} \lim_{n \to \infty} \sum_{\tilde{y} \in \mathrm{Gal}(\overline{K}/K) \cdot \tilde{x}} \sum_{v \in \Omega_K} N_v \cdot \frac{\log \|F^n(\tilde{y})\|_v}{d^n}$$

where $F : K^2 \to K^2$ and $\tilde{x}$ are homogenous lifts of f and $x \in \mathbb{P}^1(\overline{K})$, and $\|(z_1, z_2)\|_v := \max\{|z_1|_v, |z_2|_v\}$. By product formula (2), the height $\widehat{h}_f$ does not depend on the choice of the homogenous lift F and therefore it is well-defined. As proven in [CS93], $\widehat{h}_f(x) \geq 0$ with equality if and only if x is preperiodic under the iteration of f.

3.2. Adelic metrized line bundle.
Let $\mathcal{L}$ be an ample line bundle of an irreducible projective curve C over a number field K. As in Subsection 3.1, K is naturally equipped with absolute values $|\cdot|_v$ for $v \in \Omega_K$. A *metric* $\|\cdot\|_v$ on $\mathcal{L}$ is a collection of norms, one for each $x \in X(K_v)$, on the fibres $\mathcal{L}(x)$ of the line bundle, with

$$\|\alpha s(x)\|_v = |\alpha|_v \|s(x)\|_v$$

for any section s of $\mathcal{L}$. An *adelic metrized line bundle* $\overline{\mathcal{L}} = \{\mathcal{L}, \{\|\cdot\|_v\}_{v \in \Omega_K}\}$ over $\mathcal{L}$ is a collection of metrics on $\mathcal{L}$, one for each place $v \in \Omega_K$, satisfying certain continuity and coherence conditions; see [Zha95a, Zha95b].

There are various adelic metrized line bundles; the simplest such adelic metrized line bundle is the line bundle $\mathcal{O}_{\mathbb{P}^1}(1)$ equipped with metrics $\|\cdot\|_v$ (for each $v \in \Omega_K$), which evaluated at a section $s := u_0 X_0 + u_1 X_1$ of $\mathcal{O}_{\mathbb{P}^1}(1)$

(where u_0, u_1 are scalars and X_0, X_1 are the canonical sections of $\mathcal{O}_{\mathbb{P}^1}(1)$) is given by

$$\|s([x_0 : x_1])\|_v := \frac{|u_0 x_0 + u_1 x_1|_v}{\max\{|x_0|_v, |x_1|_v\}}.$$

Furthermore, we can define other metrics on $\mathcal{O}_{\mathbb{P}^1}(1)$ corresponding to a rational function f of degree $d \geq 2$ defined over K. We fix a homogenous lift $F : K^2 \to K^2$ of f with homogenous degree d. For $n \geq 1$, write $F^n = (F_{0,n}, F_{1,n})$. For each place $v \in \Omega_K$, we can define a metric on $\mathcal{O}_{\mathbb{P}^1}(1)$ as

$$\|s([x_0 : x_1])\|_{v,F,n} := \frac{|u_0 x_0 + u_1 x_1|_v}{\max\{|F_{0,n}(x_0, x_1)|_v, |F_{1,n}(x_0, x_1)|_v\}^{1/d^n}},$$

where $s = u_0 X_1 + u_1 X_1$ with u_0, u_1 scalars and X_0, X_1 canonical sections of $\mathcal{O}_{\mathbb{P}^1}(1)$. Hence $\{\mathcal{O}_{\mathbb{P}^1}(1), \{\|\cdot\|_{v,F,n}\}_{v \in \Omega_K}\}$ is an adelic metrized line bundle over $\mathcal{O}_{\mathbb{P}^1}(1)$.

A sequence $\{\mathcal{L}, \{\|\cdot\|_{v,n}\}_{v \in \Omega_K}\}_{n \geq 1}$ of adelic metrized line bundles over an ample line bundle $\mathcal{L}$ on a curve C is convergent to $\{\mathcal{L}, \{\|\cdot\|_v\}_{v \in \Omega_K}\}$, if for all n and all but finitely many $v \in \Omega_K$, $\|\cdot\|_{v,n} = \|\cdot\|_v$, and if $\left\{\log \frac{\|\cdot\|_{v,n}}{\|\cdot\|_v}\right\}_{n \geq 1}$ converges to 0 uniformly on $C(\mathbb{C}_v)$ for all $v \in \Omega_K$. The limit $\{\mathcal{L}, \{\|\cdot\|_v\}_{v \in \Omega_K}\}$ is an adelic metrized line bundle.

A typical example of a convergent sequence of adelic metrized line bundles is $\{\{\mathcal{O}_{\mathbb{P}^1}(1), \{\|\cdot\|_{v,F,n}\}_{v \in \Omega_K}\}\}_{n \geq 1}$ which converges to the metrized line bundle denoted by $\{\mathcal{O}_{\mathbb{P}^1}(1), \{\|\cdot\|_{v,F}\}_{v \in \Omega_K}\}$ (see [BR06] and also see [Zha95b, Theorem 2.2] for the more general case of a polarizable endomorphism f of a projective variety).

Let f be a rational function of degree $d \geq 2$ defined over the number field K, let C be a curve equipped with a non-constant morphism $\psi : C \to \mathbb{P}^1$ defined over K, and $\mathcal{L}_\psi := \psi^* \mathcal{O}_{\mathbb{P}^1}(1)$ be the pullback line bundle by ψ. For a fixed homogenous lift F of f with degree $d \geq 2$ and places $v \in \Omega_K$, the metrics $\|\cdot\|_{v,\psi,F,n}$ on $\mathcal{L}_\psi$ are defined as $\|\cdot\|_{v,\psi,F,n} = \psi^* \|\cdot\|_{v,F,n}$. As the sequence $\{\mathcal{O}_{\mathbb{P}^1}(1), \{\|\cdot\|_{v,F,n}\}_{v \in \Omega_K}\}_{n \geq 1}$ is convergent, we get a sequence of convergent adelic metrized line bundles $\{\mathcal{L}_\psi, \{\|\cdot\|_{v,\psi,F,n}\}_{v \in \Omega_K}\}_{n \geq 1}$ and use

$$\overline{\mathcal{L}}_{\psi,F} := \{\mathcal{L}_\psi, \{\|\cdot\|_{v,\psi,F}\}_{v \in \Omega_K}\}$$

to denote its limit, which is an adelic metrized line bundle over $\mathcal{L}_\psi$ such that

$$(4) \qquad \|\cdot\|_{v,\psi,F} = \psi^* \|\cdot\|_{v,F}.$$

For sections s of $\mathcal{L}_\psi$ of the form $\psi^*(u_0 X_0 + u_1 X_1)$ we have that

$$(5) \quad \|s(t)\|_{v,\psi,F,n} := \frac{|u_0 \psi_0(t) + u_1 \psi_1(t)|_v}{\max\{|F_{0,n}(\psi_0(t), \psi_1(t))|_v, |F_{1,n}(\psi_0(t), \psi_1(t))|_v\}^{1/d^n}},$$

where $\psi(t) = [\psi_0(t) : \psi_1(t)]$, and so,

$$\|s(t)\|_{v,\psi,F} = \frac{|u_0 \psi_0(t) + u_1 \psi_1(t)|_v}{\lim_{n \to \infty} \max\{|F_{0,n}(\psi_0(t), \psi_1(t))|_v, |F_{1,n}(\psi_0(t), \psi_1(t))|_v\}^{1/d^n}}.$$

3.3. **Equidistribution of small points.** For a semipositive line bundle $\overline{\mathcal{L}}$ on a (irreducible) curve C defined over a number field K, the height for $x \in C(\overline{K})$ is given by

$$(6) \qquad \widehat{h}_{\overline{\mathcal{L}}}(x) = \frac{1}{|\operatorname{Gal}(\overline{K}/K) \cdot x|} \sum_{y \in \operatorname{Gal}(\overline{K}/K) \cdot x} \sum_{v \in \Omega_K} -N_v \log \|s(y)\|_v$$

where $|\operatorname{Gal}(\overline{K}/K) \cdot x|$ is the number of points in the Galois orbits of x, and s is any meromorphic section of $\mathcal{L}$ with support disjoint from $\operatorname{Gal}(\overline{K}/K) \cdot x$. A sequence of points $x_n \in C(\overline{K})$ is *small*, if $\lim_{n \to \infty} \widehat{h}_{\overline{\mathcal{L}}}(x_n) = \widehat{h}_{\overline{\mathcal{L}}}(C)$; see [Zha95b] for more details on constructing the height for any irreducible subvariety Y of C (which is denoted by $\widehat{h}_{\overline{\mathcal{L}}}(Y)$). We use the following equidistribution result due to Chambert-Loir [CL06] (in the case of curves) and to Yuan [Yua08] (in the case of an arbitrary projective variety).

Theorem 3.1. [Yua08, Theorem 3.1] *Suppose C is a projective curve over a number field K, and $\overline{\mathcal{L}}$ is a metrized line bundle over C such that $\mathcal{L}$ is ample and the metric is semipositive. Let $\{x_n\}$ be a non-repeating sequence of points in $C(\overline{K})$ which is small. Then for any $v \in \Omega_K$, the Galois orbits of the sequence $\{x_n\}$ are equidistributed in the analytic space $C_{\mathbb{C}_v}^{an}$ with respect to the probability measure $d\mu_v = c_1(\overline{\mathcal{L}})_v / \deg_{\mathcal{L}}(C)$.*

When v is archimedean, $C_{\mathbb{C}_v}^{an}$ corresponds to $C(\mathbb{C})$ and the curvature $c_1(\overline{\mathcal{L}})_v$ of the metric $\| \cdot \|_v$ is given by $c_1(\overline{\mathcal{L}})_v = \frac{\partial \overline{\partial}}{\pi i} \log \| \cdot \|_v$. For non-archimedean place v, $C_{\mathbb{C}_v}^{an}$ is the Berkovich space associated to $C(\mathbb{C}_v)$, and Chambert-Loir [CL06] constructed an analog of curvature on $C_{\mathbb{C}_v}^{an}$. The precise meaning of the equidistribution above is that

$$\lim_{n \to \infty} \frac{1}{|\operatorname{Gal}(\overline{K}/K) \cdot x_n|} \sum_{y \in \operatorname{Gal}(\overline{K}/K) \cdot x_n} \delta_y = \mu_v,$$

where δ_y is point mass probability measure supported on $y \in C_{\mathbb{C}_v}^{an}$, and the limit is the weak limit for probability measures on the compact space $C_{\mathbb{C}_v}^{an}$.

For a dynamical system $(\mathbb{P}^1, f)$, at an archimedean place v, it is well known that the curvature of the limit of metrized line bundles

$$\{\mathcal{O}_{\mathbb{P}^1}(1), \{\| \cdot \|_{v,F,n}\}_{v \in \Omega_K}\}_{n \geq 1}$$

is a $(1,1)$-current given by $d\mu_f$, which is independent of the choice of F. We conclude this section by noting that in the case of the metrized line bundle $\overline{\mathcal{L}}_{\psi,F}$ on a curve C associated to a morphism $\psi : C \longrightarrow \mathbb{P}^1$, then at an archimedean place v we have that

$$(7) \qquad c_1(\overline{\mathcal{L}}_{\psi,F})_v = \psi^* d\mu_f,$$

where μ_f is the invariant measure on $\mathbb{P}^1_{\mathbb{C}_v}$ associated to the rational function f (see Subsection 2.2). Note that since v is archimedean, then $\mathbb{C}_v = \mathbb{C}$ and so, the equality from (7) follows simply by taking $\frac{\partial \overline{\partial}}{\pi i} \log \| \cdot \|_{\psi,F,v}$ which yields $\psi^* d\mu_f$ by the definition of the metric $\| \cdot \|_{\psi,F,v}$.

4. Equal measures

We work under the assumption from the direct implication in Theorem 1.1, under the additional assumption that the rational functions f_i and the curve C are all defined over $\bar{\mathbb{Q}}$.

So, f_1 and f_2 are rational functions each of degree $d_i > 1$ defined over a number field K. Also, $C \subset \mathbb{P}^1 \times \mathbb{P}^1$ is a (irreducible) curve defined over K, which projects dominantly onto both coordinates. We assume C contains infinitely many points of small height, i.e. there exists an infinite sequence of points $(x_{1,n}, x_{2,n}) \in C(\overline{K})$ such that $\lim_{n \to \infty} \widehat{h}_{f_1}(x_{1,n}) + \widehat{h}_{f_2}(x_{2,n}) = 0$.

We let $\pi_i : C \longrightarrow \mathbb{P}^1$ for $i = 1, 2$ be projection maps of $\mathbb{P}^1 \times \mathbb{P}^1$ (restricted on C) onto both coordinates. For each $i = 1, 2$, we let $\overline{\mathcal{L}}_i := \overline{\mathcal{L}}_{\pi_i, F_i}$ be the adelic metrized line bundle on $\mathcal{L}_i := \pi_i^* \mathcal{O}_{\mathbb{P}^1}(1)$ constructed as in Subsection 3.2 with respect to the rational function f_i. Then for each point $t \in C(\bar{\mathbb{Q}})$, using the product formula (2) and equations (3, 4, 6) we have

$$(8) \qquad \widehat{h}_{\overline{\mathcal{L}}_i}(t) = \widehat{h}_{f_i}(\pi_i(t)).$$

Using [Zha95b, Theorem (1.10)] along with (8), also coupled with the fact that C projects dominantly onto each factor of $\mathbb{P}^1 \times \mathbb{P}^1$ and the fact that there are infinitely many preperiodic points for f_i (which thus have canonical height 0), we conclude that

$$(9) \qquad \widehat{h}_{\overline{\mathcal{L}}_i}(C) = 0 \text{ for each } i = 1, 2.$$

We let μ_i be the invariant measure on $\mathbb{P}^1$ corresponding to the rational function f_i, for $i = 1, 2$; then we let $\tilde{\mu}_i := (\pi_i^* \mu_i) / \deg(\pi_i)$.

Proposition 4.1. *With the above notation, $\tilde{\mu}_1 = \tilde{\mu}_2$.*

Proof. This is an immediate consequence of Theorem 3.1 applied to the sequence of points $t_n := (x_{1,n}, x_{2,n})$ with respect to the adelic metrized line bundles $\overline{\mathcal{L}}_1$ and $\overline{\mathcal{L}}_2$. Indeed, using (8) and (9) and the assumption on the points $(x_{1,n}, x_{2,n})$ we get that

$$\lim_{n \to \infty} \widehat{h}_{\overline{\mathcal{L}}_i}(t_n) = \widehat{h}_{\overline{\mathcal{L}}_i}(C),$$

for each $i = 1, 2$. Therefore we can apply Theorem 3.1 and conclude that the points $\{t_n\}$ are equidistributed on C with respect to both measures $\tilde{\mu}_1$ and $\tilde{\mu}_2$ (see also (7) which yields that $\tilde{\mu}_i$ are indeed the probability measures on $C(\mathbb{C})$ with respect to which the points $\{t_n\}$ are equidistributed). This concludes the proof of Proposition 4.1. $\square$

5. From equal measures to preperiodic curves

In this section we pick up exactly where Proposition 4.1 led us: we have a curve C projecting dominantly onto each coordinate of $\mathbb{P}^1 \times \mathbb{P}^1$, and we have two rational functions of degree greater than 1 with the property that $\tilde{\mu}_f = \tilde{\mu}_g$, where $\tilde{\mu}_f$ and $\tilde{\mu}_g$ are the measures induced on C by the dynamical systems $(\mathbb{P}^1, f)$ respectively $(\mathbb{P}^1, g)$ through the canonical projection maps

of C onto each of the two coordinates of $\mathbb{P}^1 \times \mathbb{P}^1$. The results of this section are valid when C, f and g are defined over $\mathbb{C}$ (thus not necessarily over a number field), as long as $\tilde{\mu}_f = \tilde{\mu}_g$. We prove the following crucial result, which will be used later in the proof of Theorem 1.1. We thank one of the referees for pointing out that we needed a strengthening of our previous version of the following result in order to apply it correctly in the proof of Theorem 1.1.

Theorem 5.1. *Let f and g be rational functions defined over $\mathbb{C}$ with degrees greater than one, and assume $f(x)$ is not exceptional. Let $C \subset \mathbb{P}^1 \times \mathbb{P}^1$ be an irreducible curve defined over $\mathbb{C}$ which projects dominantly onto both coordinates, and let $\tilde{\mu}_f$ and $\tilde{\mu}_g$ be the induced measures on C by the dynamical systems $(\mathbb{P}^1, f)$ and $(\mathbb{P}^1, g)$. Suppose that $\tilde{\mu}_f = \tilde{\mu}_g$ and also suppose that there is a point $(x_0, y_0) \in C$ such that:*

- *x_0 is a repelling fixed point of f and y_0 is fixed by g; and*
- *there is a non-constant holomorphic germ $h(x)$, with $h(x_0) = y_0$ and $(x, h(x)) \in C$ for all x in a small neighbourhood of x_0.*

Then C is fixed by the endomorphism $(x, y) \mapsto (f^n(x), g^m(y))$ of $\mathbb{P}^1 \times \mathbb{P}^1$ for some positive integers n and m, and in particular $\deg(f)^n = \deg(g)^m$.

Proof. We prove this theorem following an idea of Levin from [Lev90], where he studied the symmetries of the Julia set of a rational function. The following proposition is central to our proof.

Proposition 5.2. *With the same assumptions as in Theorem 5.1, there exist integers $n, m \geq 1$, such that $f^n = h^{-1} \circ g^m \circ h$ as germs at x_0.*

Proof of Proposition 5.2. As f is repelling at x_0, the multiplier $\lambda_1 := f'(x_0)$ have absolute value greater than one, i.e., x_0 is in the support of μ_f hence (x_0, y_0) is in the support of $\tilde{\mu}_f = \tilde{\mu}_g$. Consequently y_0 is in the Julia set J_g (the support of μ_g), and then we have $|\lambda_2| \geq 1$ for the multiplier $\lambda_2 := g'(y_0)$. As $h(x)$ is not a constant, we define the germ $\tilde{g}$ at x_0 as a branch of $\tilde{g}(x) := h^{-1} \circ g \circ h(x)$, which has multiplier $\sqrt[j]{\lambda_2} = \tilde{g}'(x_0)$ where j is the local degree of $h(x)$ at x_0.

First, we show that $|\lambda_2| > 1$. Suppose this is not true, then we have $|\lambda_2| = 1$. Fix an integer j_0 with $\deg(g)^{j_0} > \deg(f)$. As $\tilde{\mu}_f = \tilde{\mu}_g$, since $\mu_f(x_0) = 0$ and from (1), one has $\mu_f(\tilde{g}(B)) = \deg(g) \cdot \mu_f(B)$ for any measurable set B in a small neighbourhood of x_0 with $\tilde{g}$ injective on B. Then for any small neighbourhood A of x_0 with both f and $\tilde{g}$ injective on A, one has

(10) $\mu_f(f(A)) = \deg(f) \cdot \mu_f(A)$ and $\mu_f(\tilde{g}(A)) = \deg(g) \cdot \mu_f(A)$

Moreover, since $|\tilde{g}'(x_0)| = 1$ and $|f'(x_0)| > 1$, we can pick a very small neighbourhood A of x_0 with $\tilde{g}^{j_0}(A) \subset f(A)$. Then one has

$$\mu_f(\tilde{g}^{j_0}(A)) = \deg(g)^{j_0} \cdot \mu_f(A) \leq \mu_f(f(A)) = \deg(f) \cdot \mu_f(A).$$

Because $\deg(g)^{j_0} > \deg(f)$, we have $\mu_f(A) = 0$, which contradicts the fact that x_0 is in the support of μ_f. So we have $|\lambda_2| > 1$.

As both $|\lambda_1|$ and $|\lambda_2|$ are greater than one, there are invertible holomorphic maps h_1 and h_2 from a neighbourhood of x_0 to $\mathbb{C}$ with $h_1(x_0) = h_2(x_0) = 0$, which also satisfy

$$h_1 \circ f \circ h_1^{-1}(z) = \lambda_1 \cdot z, \ \ h_2 \circ \tilde{g} \circ h_2^{-1}(z) = \sqrt[j]{\lambda_2} \cdot z,$$

i.e., in a neighbourhood of x_0, both f and $\tilde{g}$ are conjugate to linear functions.

We pick a sequence of pairs of positive integer $\{(n_\ell, m_\ell)\}_{\ell \geq 1}$ with $n_\ell, m_\ell \to \infty$ as $\ell \to \infty$ such that

$$\lim_{\ell \to \infty} \frac{\lambda_1^{n_\ell}}{\sqrt[j]{\lambda_2}^{m_\ell}} = 1.$$

Hence the germs

$$f^{n_\ell} \circ \tilde{g}^{-m_\ell}(x) = [h_1^{-1} \circ (h_1 \circ f \circ h_1^{-1})^{n_\ell} \circ h_1] \circ [h_2^{-1} \circ (h_2 \circ \tilde{g} \circ h_2^{-1})^{-m_\ell} \circ h_2(x)]$$

$$= h_1^{-1} \circ (h_1 \circ f \circ h_1^{-1})^{n_\ell} \circ (h_1 \circ h_2^{-1}) \circ (h_2 \circ \tilde{g} \circ h_2^{-1})^{-m_\ell} \circ h_2(x)$$

$$= h_1^{-1} \left(\lambda_1^{n_\ell} \cdot h_1 \circ h_2^{-1} \left(\frac{h_2(x)}{\sqrt[j]{\lambda_2}^{m_\ell}} \right) \right)$$

are well defined on $B(x_0, r)$ for some small $r > 0$ and all $\ell \geq 1$. We can pick a sufficiently small $r > 0$ such that $f^{n_\ell} \circ \tilde{g}^{-m_\ell}|_{B(x_0,r)}$ is injective and $f^{n_\ell} \circ \tilde{g}^{-m_\ell}(B(x_0, r))$ is contained in a small neighbourhood of x_0 for all $\ell \geq 1$. Hence $\mathcal{H} := \{f^{n_\ell} \circ \tilde{g}^{-m_\ell}\}_{\ell \geq 1}$ is a normal family on $B(x_0, r)$. Moreover, as $\tilde{\mu}_f = \tilde{\mu}_g$, then $f^{n_\ell} \circ \tilde{g}^{-m_\ell}$ not only preserves J_f but also μ_f (up to a scalar multiple) restricted on $B(x_0, r)$, i.e. it is a symmetry on J_f. The multiplier $(f^{n_\ell} \circ \tilde{g}^{-m_\ell})'(x_0) = \lambda_1^{n_\ell} / \sqrt[j]{\lambda_2}^{m_\ell}$ of $f^{n_\ell} \circ \tilde{g}^{-m_\ell}$ at the fixed point x_0 tends to 1 as ℓ tends to infinity; thus any limit function of $\mathcal{H}$ would have derivative 1 at x_0 and therefore it will not be a constant function. Consequently, $\mathcal{H}$ contains only finitely many distinct functions on $B(x_0, r)$, otherwise it would contradict to the fact that f is not a post-critically finite map with parabolic orbifold; see [Lev90, Theorem 1]. We pick integers $\ell_1, \ell_2 > 1$ with

$$n_{\ell_1} > n_{\ell_2}, \ m_{\ell_1} > m_{\ell_2}, \ f^{n_{\ell_1}} \circ \tilde{g}^{-m_{\ell_1}} = f^{n_{\ell_2}} \circ \tilde{g}^{-m_{\ell_2}}.$$

Hence on a neighbourhood of x_0, one has

$$f^{n_{\ell_2}} \circ f^{n_{\ell_1} - n_{\ell_2}} = f^{n_{\ell_2}} \circ \tilde{g}^{m_{\ell_1} - m_{\ell_2}},$$

and then post-compose the germ $f^{-n_{\ell_2}}$ at x_0 fixing the point x_0 on both sides of the above equation. This concludes the proof of Proposition 5.2. $\square$

Now consider the analytic equation

$$h(x) - y = 0$$

on a neighbourhood of $(x_0, y_0) \in \mathbb{P}_\mathbb{C}^1 \times \mathbb{P}_\mathbb{C}^1$. The zero set of this equation is an analytic curve passing though the point (x_0, y_0). For x close to x_0, $(x, h(x))$ are points on C. Let n, m be integers in Proposition 5.2; then $f^n(x) = h^{-1} \circ g^m \circ h(x)$ for all x near x_0, i.e. $f^n(x) = h^{-1}(g^m(h(x)))$.

Hence for points x close to x_0, the points $(f^n, g^m)(x, h(x))$, which are points on $(f^n, g^m)(C)$ are on the zero set of $h(x) - y = 0$. Finally, as

both C and $(f^n, g^m)(C)$ share an analytic curve in a neighbourhood of (x_0, y_0), they must be identical. So C is fixed by the endomorphism $(x, y) \mapsto (f^n(x), g^m(y))$ of $\mathbb{P}^1 \times \mathbb{P}^1$.

We are left to show that $\deg(f)^n = \deg(g)^m$. There are various ways of proving this last statement, including (as suggested by one of the referees) by exploiting the equality of degrees for various maps between C and $\mathbb{P}^1$. We include (as an alternative proof) an analytic argument in the spirit of the rest of our proof of Theorem 5.1. So, we pick a sufficiently small $r > 0$, and let $A = B(x_0, r)$. Then both $f^n|_A$ and $\tilde{g}^m|_A$ are injective. From (10), we have $\mu_f(f^n(A)) = \deg(f)^n \cdot \mu_f(A)$, and $\mu_f(\tilde{g}^m(A)) = \deg(g)^m \cdot \mu_f(A)$. Moreover, as f^n and $\tilde{g}^m$ are identical as germs at x_0 and $\mu_f(A)$ is not zero (x_0 is in J_f and the support of μ_f is J_f), we have $\deg(f)^n = \deg(g)^m$. $\square$

6. Conclusion of the proof of Theorems 1.1 and 1.3

In this section we prove both Theorems 1.1 and 1.3. We first prove the converse implication in Theorems 1.1 and 1.3; i.e., we prove the following result.

Lemma 6.1. *Let f and g be rational functions of same degree $d > 1$ defined over $\mathbb{C}$, and let $C \subset \mathbb{P}^1_{\mathbb{C}} \times \mathbb{P}^1_{\mathbb{C}}$ be a curve which projects dominantly onto both coordinates. If C is preperiodic under the endomorphism Φ of $\mathbb{P}^1 \times \mathbb{P}^1$ given by $\Phi(x, y) = (f(x), g(y))$, then C contains infinitely many preperiodic points.*

We note that the hypothesis that C is preperiodic under the endomorphism Φ (and also C is neither horizontal nor vertical) automatically yields the equality of degrees for the two rational functions f and g. Also, we remark that an extension of Lemma 6.1 to polarizable endomorphisms of projective varieties was proven by Fakhruddin [Fak03].

Proof of Lemma 6.1. As Φ projects dominantly and $f : \mathbb{P}^1 \to \mathbb{P}^1$ has infinitely many preperiodic points on $\mathbb{P}^1$, there are infinitely many points (x_1, x_2) on the curve C with x_1 preperiodic under the action of f. We want to show that x_2 is preperiodic under g. As x_1 and C is preperiodic under f and respectively Φ, there are only finitely many distinct points $f^n(x_1)$ and respectively curves $C_n := \Phi^n(C)$ for $n \geq 1$. Hence there are only finitely many distinct points $(f^n(x_1), g^n(x_2)) \in \pi_1^{-1}|_{C_n}(f^n(x_1))$ for $n \geq 1$, i.e. $\{g^n(x_2)\}_{n \geq 1}$ is a finite set. Consequently, x_2 is preperiodic under g. $\square$

Proof of Theorem 1.1. First, we note that the converse implication was already proven in Lemma 6.1.

Secondly, we prove the direct implication in Theorem 1.1 under the assumption that the curve C and the rational functions f_1, f_2 are defined over $\bar{\mathbb{Q}}$. Using [Mim13, Theorem 1.8], we get that for each point $(x_1, x_2) \in C(\bar{\mathbb{Q}})$, we have that x_1 is preperiodic for f_1 if and only if x_2 is preperiodic for f_2. In particular, C contains infinitely many preperiodic points.

Since all but finitely many periodic points of a rational map are repelling, and also, at all but finitely many points the projection map on the first coordinate $C \to \mathbb{P}^1$ is unramified, then we can find a point $(x_0, y_0) \in C$ such that

(a) x_0 is a periodic repelling point for f_1; and
(b) there is a non-constant holomorphic germ $h_0(x)$, with $h_0(x_0) = y_0$ and $(x, h_0(x)) \in C$ for all x in a small neighborhood of x_0.

Note that since x_0 is fixed by f_1, then y_0 must be preperiodic under f_2. At the expense of replacing f_1 by $F_1 := f_1^{r_1}$ and f_2 by $F_2 := f_2^{r_2}$ for some positive integers r_1, r_2, and also replacing C by $C_1 := \Phi_1(C)$ where Φ_1 is the endomorphism of $\mathbb{P}^1 \times \mathbb{P}^1$ given by $\Phi_1(x_1, x_2) := (F_1(x_1), F_2(x_2))$, we can assume that C_1 contains a fixed point $(x_{0,1}, x_{0,2})$ of Φ_1 such that $x_{0,1} := F_1(x_0)$ is a repelling fixed point for F_1 (while $x_{0,2} := F_2(y_0)$ is fixed by F_2), and also ensure that condition (b) above is satisfied for an analytic germ in a neighborhood of $(x_{0,1}, x_{0,2}) \in C_1$, i.e., for the analytic germ $h := F_2 \circ h_0 \circ F_1^{-1}$ at $x_{0,1}$, we have $h(x_{0,1}) = x_{0,2}$ and $(x, h(x)) \in C_1$ for all x in a small neighbourhood of $x_{0,1}$.

Clearly, also C_1 contains infinitely many preperiodic points under the action of Φ_1. For $i = 1, 2$, we let $\pi_i : C_1 \longrightarrow \mathbb{P}^1$ be the corresponding projection of $\mathbb{P}^1 \times \mathbb{P}^1$ onto each coordinate, restricted to C_1. We let μ_i be the invariant measure on $\mathbb{P}^1(\mathbb{C})$ corresponding to F_i for $i = 1, 2$; then we let $\tilde{\mu}_i := (\pi_i^* \mu_i)/\deg(\pi_i)$ be the corresponding measures on $C_1(\mathbb{C})$. According to Proposition 4.1, we get that $\tilde{\mu}_1 = \tilde{\mu}_2$. Therefore the hypotheses of Theorem 5.1 are met. Then there exist positive integers s_1, s_2 such that C_1 is fixed by the endomorphism $(x_1, x_2) \mapsto (F_1^{s_1}(x_1), F_2^{s_2}(x_2))$. In particular, $\deg(F_1^{s_1}) = \deg(F_2^{s_2})$.

Now consider the curve C and the rational functions $G_i := f_i^{r_i s_i}$ for $i = 1, 2$; then G_1 and G_2 are of the same degree. Then we run the same argument as above: we replace G_1 and G_2 by iterates of them, and then replace C by an iterate of it under the action of $(x_1, x_2) \mapsto (G_1(x_1), G_2(x_2))$ such that assumptions in Theorem 5.1 are satisfied. Since already G_1 and G_2 have the same degree, we conclude that C is preperiodic under the action of $(x, y) \mapsto (f_1^{r_1 s_1}(x_1), f_2^{r_2 s_2}(x_2))$ and $\deg(f_1^{r_1 s_1}) = \deg(f_2^{r_2 s_2})$.

Finally, we deal with the general case of the direct implication from Theorem 1.1. All we need to prove is that the hypotheses of Theorem 5.1 are met (after perhaps replacing each f_i by an iterate and also replacing C by its image under a map of the form $(x_1, x_2) \mapsto (f_1^{r_1}(x_1), f_2^{r_2}(x_2))$). So, all we need to show is that $\tilde{\mu}_1 = \tilde{\mu}_2$ where the measures $\tilde{\mu}_i$ on $C(\mathbb{C})$ are the probability measures on $C(\mathbb{C})$ induced by the pullbacks of the invariant measures under f_i on $\mathbb{P}^1_\mathbb{C}$ through the projection maps.

In this case, let K be a finitely generated subfield of $\mathbb{C}$ such that f_1, f_2 and C are defined over K, and let $\overline{K}$ be a fixed algebraic closure of K in $\mathbb{C}$. We know there exists an infinite sequence $S := \{(x_{1,n}, x_{2,n})\} \subset C(\mathbb{C})$ such that each $x_{i,n}$ is a preperiodic point for f_i for $i = 1, 2$ and for $n \geq 1$. Then

f_1 and f_2 are base changes of endomorphisms $f_{1,K}$ and $f_{2,K}$ of $\mathbb{P}^1_K$; similarly, S is the base change of a subset $S_K \subset C(\overline{K})$. Without loss of generality, we may assume S_K is closed under the action of $\mathrm{Gal}(\overline{K}/K)$ and therefore we can view S_K as a union of closed points in $\mathbb{P}^1_K \times \mathbb{P}^1_K$. We can further extend $f_{i,K}$ to endomorphisms

$$f_{i,U} : \mathbb{P}^1_U \longrightarrow \mathbb{P}^1_U$$

over a variety U over $\mathbb{Q}$ of finite type and with function field K. For each geometric point $t \in U(\overline{\mathbb{Q}})$, the objects $(f_{1,U}, f_{2,U}, S_U)$ have reductions $(f_{1,t}, f_{2,t}, S_t)$ such that S_t consists of points with both coordinates preperiodic under the action of $f_{1,U}$, respectively of $f_{2,U}$. We also let $\tilde{\mu}_{i,t} := \pi_i^* \mu_{f_{i,t}} / \deg(\pi_i)$ (for $i = 1, 2$) be the probability measures on C_t obtained as pullback through the usual projection map onto each coordinate restricted to C_t of the invariant measures on $\mathbb{P}^1_\mathbb{C}$ corresponding to each $f_{i,t}$. In [YZb, Theorem 4.7] (see also [YZa, Lemma 3.2.3]), Yuan and Zhang prove that the set S_t is still infinite for all the $\overline{\mathbb{Q}}$-points of a dense open subset $U_0 \subseteq U$; we note that the results of [YZa] were recently published in [YZ16], while [YZb] has been updated to [YZc] using slightly different arguments. Thus, as proven in Proposition 4.1, we conclude that $\tilde{\mu}_{1,t} = \tilde{\mu}_{2,t}$ for each $t \in U_0(\overline{\mathbb{Q}})$. Since $U_0(\overline{\mathbb{Q}})$ is dense in $U(\mathbb{C})$ with respect to the usual archimedean topology, while the measures $\tilde{\mu}_{i,t}$ vary continuously with the parameter t (since from the construction, the potential functions of these measures vary continuously with the coefficients of $f_{i,t}$), we conclude that $\tilde{\mu}_{1,t} = \tilde{\mu}_{2,t}$ for all points in $U(\mathbb{C})$ including the point corresponding to original embedding $K \subset \mathbb{C}$. Thus we have shown that the measures $\tilde{\mu}_1, \tilde{\mu}_2$ induced on C by the action of f_1 and f_2 on $\mathbb{P}^1_\mathbb{C}$ are equal. Arguing identically as before, i.e., replacing f_i by iterates and also replacing C by its image under a suitable map $(x_1, x_2) \mapsto (f_1^{r_1}(x_1), f_2^{r_2}(x_2))$, we can guarantee that the hypotheses of Theorem 5.1 are met and thus conclude the proof of Theorem 1.1. $\square$

Proof of Theorem 1.3. First we note that the converse implication was proven in Lemma 6.1; note also that the tangent subspaces at *each* preperiodic point on the preperiodic curve C is also preperiodic in the corresponding Grassmanian.

So, from now on, we assume that C contains infinitely many preperiodic points verifying the hypotheses of Theorem 1.3 and we prove next that C is preperiodic. First we note that if the curve does not project dominantly onto both coordinates, then the result is trivial, since in this case it must be that C projects to one coordinate to a preperiodic point under the corresponding map f_i.

Let $d := \deg(f_i)$ (for $i = 1, 2$). If one of the maps f_i is not exceptional, then the conclusion is provided by Theorem 1.1; in this case we do not require the hypothesis regarding the action on the tangent subspaces of the preperiodic points.

Assume now that f_1 is conjugate to $z^{\pm d}$ or $\pm T_d(z)$. The hypothesis that C contains infinitely many preperiodic points (once again we do not require

the extra hypothesis regarding the induced action on the tangent spaces) yields, according to Proposition 4.1, the equality of the following measures on $C(\mathbb{C})$:

$$\text{(11)} \qquad \deg(\pi_2) \cdot \pi_1^* \mu_1 = \deg(\pi_1) \cdot \pi_2^* \mu_2,$$

where the π_i's are the corresponding projection maps restricted on C to each coordinate of $\mathbb{P}^1 \times \mathbb{P}^1$, while μ_i is the invariant measure on $\mathbb{P}^1_{\mathbb{C}}$ corresponding to f_i. Then also f_2 must be conjugate to $z^{\pm d}$ or $\pm T_d(z)$; see Section 2.4.

We assume now that both f_i are conjugate to $T_d(z)$; the other cases follow similarly. So, $f_i = \eta_i^{-1} \circ T_d \circ \eta_i$ for some linear transformations η_i. Then it is sufficient to prove Theorem 1.3 for f_1 and f_2 replaced by T_d and the curve C replaced by $\Psi_1(C)$, where $\Psi_1 : \mathbb{P}^1 \times \mathbb{P}^1 \longrightarrow \mathbb{P}^1 \times \mathbb{P}^1$ is defined by $\Psi_1(x_1, x_2) = (\eta_1(x_1), \eta_2(x_2))$. We let $\nu : \mathbb{P}^1 \longrightarrow \mathbb{P}^1$ be defined by $\nu([a : b]) = [a^2 + b^2 : ab]$, and let Ψ_2 be the endomorphism of $\mathbb{P}^1 \times \mathbb{P}^1$ induced by $\nu \times \nu$. Then an irreducible component C_1 of $\Psi_2^{-1}(C)$ contains infinitely many points (x_1, x_2) where each x_i is a preperiodic point for the map $z \mapsto z^d$ on $\mathbb{P}^1$, i.e. each x_i is a root of unity. Then the classical Manin-Mumford Conjecture for $\mathbb{G}_m^2$ proved by Laurent [Lau84] yields that C_1 is a torsion translate of an algebraic subtorus of $\mathbb{G}_m^2$ and therefore, it is preperiodic under the action of $(x_1, x_2) \mapsto (x_1^d, x_2^d)$ on $\mathbb{P}^1 \times \mathbb{P}^1$. Thus $C = \Psi_2(C_1)$ is preperiodic under the action of (f_1, f_2).

Now, assume f_1 is a Lattès map; again using that $\deg(\pi_2) \cdot \pi_1^* \mu_1 = \deg(\pi_1) \cdot \pi_2^* \mu_2$ yields that also f_2 is a Lattès map; see Section 2.4. So, again at the expense of replacing each f_i by a conjugate of it, and also replacing C by its image under an automorphism of $\mathbb{P}^1 \times \mathbb{P}^1$, we may and do assume that we have elliptic curves E_1 and E_2 defined over $\mathbb{C}$, a coordinatewise projection map $\pi : E_1 \times E_2 \longrightarrow \mathbb{P}^1 \times \mathbb{P}^1$, and we have a polarizable group endomorphism $\tilde{\Phi}$ of $E_1 \times E_2$ (note that $\deg(f_1) = \deg(f_2) > 1$) such that $\Phi \circ \pi = \pi \circ \tilde{\Phi}$, where we recall that $\Phi : \mathbb{P}^1 \times \mathbb{P}^1 \longrightarrow \mathbb{P}^1 \times \mathbb{P}^1$ is given by $\Phi(x_1, x_2) = (f_1(x_1), f_2(x_2))$. Then our hypotheses imply the existence of a curve $C_1 \subset E_1 \times E_2$ (which is a irreducible component of $\pi^{-1}(C)$) such that C_1 contains infinitely many smooth preperiodic points P under the action of $\tilde{\Phi}$, and moreover the tangent space of C_1 at P is preperiodic under the induced action of $\tilde{\Phi}$ on $\text{Gr}_1(T_{E_1 \times E_2, P})$ (here we use the fact that the ramified locus for π is a finite union of divisors of the form $\{a\} \times E_2$ and $E_1 \times \{b\}$, while at any unramified point of π, we get an induced isomorphism π^* on the correspnding tangent spaces). Since $\tilde{\Phi}$ is a polarizable endomorphism of $E_1 \times E_2$, then the result follows from [GTZ11, Theorem 2.1]. $\qquad\square$

7. Proof of Theorems 1.4 and 1.5

In Subsection 7.1, we develop the technical results needed in the proof of Theorem 1.4 whose proof is finalized in Subsection 7.2. We conclude our paper with the proof of Theorem 1.5 in Subsection 7.3.

7.1. **A result in polynomial decomposition.** Throughout this subsection, fix $d \geq 2$ and $f(x) = x^d + c \in \mathbb{C}[x]$ that is not conjugate to x^d or $\pm T_d(x)$. As explained in Section 1, this amounts to $c \neq 0$ and also $c \neq -2$ if $d = 2$. For every positive divisor δ of d, let $f_\delta(x) = x^{d/\delta} \circ (x^\delta + c) = (x^\delta + c)^{d/\delta}$. Note that $f_d(x) = f(x)$ while $f_1(x) = (x + c)^d$ is conjugate to $f(x)$. Fix $n \in \mathbb{N}$. We prove the following:

Theorem 7.1. *Every non-constant solution $A, B \in \mathbb{C}[x]$ of the functional equation:*

$$f^n \circ A = A \circ B$$

has the form $A = f^m \circ (x^\delta + c) \circ L$ and $B = L^{-1} \circ f_\delta^n \circ L$ for some integer $m \geq 0$, positive divisor δ of d, and linear polynomial L.

We start with the following lemmas:

Lemma 7.2. *Let $D \geq 2$ and $2 \leq m \leq D$ be integers. Assume that $P(x)$ has the form $ax^D + bx^{D-m} + \ldots$ with $ab \neq 0$. Then for every $n \in \mathbb{N}$, P^n has the form $\alpha x^{D^n} + \beta x^{D^n - m} + \ldots$ with $\alpha\beta \neq 0$.*

Proof. This can be proved easily by induction on n. $\qquad\square$

Lemma 7.3. *If $A, B \in \mathbb{C}[x]$ are non-constant such that $f^n \circ A = A \circ B$ and $\gcd(d, \deg(A)) = 1$ then A is linear.*

Proof. Assume otherwise, let $\alpha := \deg(A) \geq 2$. By [Ino11, Theorem 8], we have that f^n is linearly conjugate to $x^r(P(x))^\alpha$ for some $P(x) \in \mathbb{C}[x] \setminus x\mathbb{C}[x]$ and $r \equiv d^n$ modulo α. Since f^n is not linearly conjugate to the power map, we have that $\deg(P) > 0$. Write $F(x) = f^{n-1}(x)$, and let $ux + w$ be a linear polynomial that conjugates $f^n = (F(x))^d + c$ to the above polynomial. We have:

$$uF\left(\frac{x-w}{u}\right)^d + uc + w = x^r(P(x))^\alpha.$$

Assume that $uc + w \neq 0$, then the Mason-Stothers theorem [Lan02, p. 194] (an effective version of the *abc*-conjecture over function fields) implies:

(12) $$d^n \leq d^{n-1} + \deg(P).$$

On the other hand, we have $r + \alpha \deg(P) = d^n$. Recall that $\alpha \geq 2$, and $\gcd(\alpha, d) = 1$ so that $\alpha \geq 3$ if $d = 2$. Hence (12) implies that $\alpha \deg(P) \geq d^n$, contradiction. So we must have $uc + w = 0$. Hence $x^r(P(x))^\alpha$ is a d-th power. This is impossible since $\gcd(r, d) = 1$ and $P(0) \neq 0$. $\qquad\square$

Following [GTZ12, Definition 2.3], we say that $g(x, y)$ is a *Siegel factor* of the polynomial $f(x, y)$ if g divides f and the plane curve given by the equation $g(x, y) = 0$ is irreducible, has geometric genus 0, and has at most two points at infinity. The following is an immediate consequence of [GTZ12, Corollary 2.7]:

Lemma 7.4. *Let m and D be coprime integers greater than 1. Let $\tilde{A}$ be a polynomial of degree D. If $x^m - \tilde{A}(y)$ has a Siegel factor then one of the following holds:*

    (I) *There exist $P(x) \in \mathbb{C}[x] \setminus x\mathbb{C}[x]$, $r \in \mathbb{N}$ satisfying $r \equiv D$ modulo m, and linear polynomial $L(x)$ such that $\tilde{A}(x) = (x^r P(x)^m) \circ L(x)$.*

    (I') *There exist linear polynomials $L_1(x)$ and $L(x)$ such that $\tilde{A}(x) = L_1 \circ x^D \circ L$.*

    (II) *There exist linear polynomials $L_2(x)$, $L_1(x)$, and $L(x)$ such that $x^m = L_1 \circ T_m \circ L_2$ and $\tilde{A} = L_1 \circ T_D \circ L$.*

Moreover, case (II) can only happen if $m = 2$ and $L_1(x)$ has the form $\ell_1 x + 2\ell_1$ for some $\ell_1 \neq 0$.

Proof. By [GTZ12, Corollary 2.7] together with the fact that $\gcd(m, D) = 1$, there exist linear polynomials L, L_1, L_2 and polynomials F and G such that $x^m = L_1 \circ F \circ L_2$, $\tilde{A} = L_1 \circ G \circ L$ and one of the following holds:

    (a) *$F(x) = x^m$ and $G(x) = x^r P(x)^m$ with P satisfying the conditions in (I).*

    (a') *$G(x) = x^D$.*

    (b) *$F(x) = T_m(x)$ and $G(x) = T_D(x)$.*

Note that cases (a) and (a') come from [GTZ12, (2.7.1)], while case (b) comes from [GTZ12, (2.7.3)]. It is immediate that (I') and (II) follow from (a') and (b) respectively. Now assume that (a) holds. From $x^m = L_1 \circ x^m \circ L_2$ and $m \geq 2$, we have that L_1 and L_2 has zero constant coefficient. Now we "absorb" the leading coefficient of L_1 into $P(x)$ and (I) follows.

For the moreover part, we have that the functional equation $x^m = L_1 \circ T_m \circ L_2$ could only hold when $m = 2$ by comparing the ramification behavior of x^m and T_m. Write $L_i(x) = \ell_i x + k_i$ for $i = 1, 2$. From $x^2 = \ell_1 T_2(\ell_2 x + k_2) + k_1$, we have that $k_2 = 0$ and $k_1 = 2\ell_1$. Therefore $L_1(x)$ has the form $\ell_1 x + 2\ell_1$. $\square$

The following result of Engstrom [GTZ12, Lemma 3.2] will be used many times:

Lemma 7.5. *Let $A, B, C, D \in \mathbb{C}[x] \setminus \mathbb{C}$ such that $A \circ B = C \circ D$.*

    (a) *If $\deg(A) \mid \deg(C)$ then there exists a polynomial $P(x) \in \mathbb{C}[x]$ such that $C = A \circ P$ and $B = P \circ D$.*

    (b) *If $\deg(B) \mid \deg(D)$ then there exists a polynomial $Q(x) \in \mathbb{C}[x]$ such that $D = Q \circ B$ and $A = C \circ Q$.*

We now prove Theorem 7.1:

Proof of Theorem 7.1. By contradiction, assume (A, B) is a non-constant solution that does not have the desired form. Assume that $\deg(A)$ is *minimal* among all counter-examples.

If $\gcd(\deg(A), d) = 1$ then Lemma 7.3 implies that A is linear. Then we can write $A(x) = (x + c) \circ L$ and $B(x) = A^{-1} \circ f^n \circ A = L^{-1} \circ f_1 \circ L$,

violating the assumption that (A, B) is a counter-example. If $d \mid \deg(A)$ then by Lemma 7.5 there is $\tilde{A} \in \mathbb{C}[x]$ such that $A = f \circ \tilde{A}$ and $f^{n-1} \circ A = \tilde{A} \circ B$. Hence $f^n \circ \tilde{A} = \tilde{A} \circ B$. Since (A, B) is a counter-example, the pair $(\tilde{A}, B)$ is also a counter-example. However $\deg(\tilde{A}) < \deg(A)$, contradicting the minimality of $\deg(A)$. Therefore we may assume that $\delta := \gcd(\deg(A), d)$ satisfies $1 < \delta < d$.

Using $f(x) = (x^\delta + c) \circ x^{d/\delta}$ and Lemma 7.5, we have that $A = (x^\delta + c) \circ \tilde{A}$ and

$$(13) \qquad x^{d/\delta} \circ f^{n-1} \circ (x^\delta + c) \circ \tilde{A} = \tilde{A} \circ B$$

for some polynomial $\tilde{A}$. Since (A, B) is a counter-example, we have that $\tilde{A}$ is not linear. Put $D = \deg(\tilde{A}) \geq 2$, we have that $\gcd(D, d/\delta) = 1$ by the definition of δ. Now Lemma 7.4 gives that one of the cases (I), (I') or (II) occurs.

Case (I): there exist nonzero polynomial $P(x)$ such that $P(0) \neq 0$, $r \in \mathbb{N}$ such that $r \equiv D$ modulo d/δ, and linear polynomial L such that $\tilde{A} = (x^r P(x)^m) \circ L$. Replace $\tilde{A}$ by $\tilde{A} \circ L^{-1}$ and B by $L \circ B \circ L^{-1}$, we may assume that $\tilde{A} = x^r P(x)^m$. Now (13) gives that $B(x)^r P(B(x))$ is the (d/δ)-th power of a polynomial. Since $P(0) \neq 0$ and $\gcd(r, d/\delta)$, we have that $B(x) = (\tilde{B}(x))^{d/\delta} = x^{d/\delta} \circ \tilde{B}$ for some polynomial $\tilde{B}$.

Case (I'): there exist linear polynomials $L_1(x)$ and $L(x)$ such that $\tilde{A} = L_1 \circ x^D \circ L$. As before, replacing $\tilde{A}$ and B by $\tilde{A} \circ L^{-1}$ and $L \circ B \circ L^{-1}$ respectively, we may assume that $\tilde{A} = L_1 \circ x^D = \ell_1 x^D + k_1$. Equation (13) gives:

$$(14) \qquad \ell_1 B(x)^D + k_1 = Q(x)^{d/\delta}$$

where $Q(x) = f^{n-1} \circ (x^\delta + c) \circ \tilde{A}$. Now assume $k_1 \neq 0$. Then the Mason-Stothers Theorem gives: $\deg(B) + \deg(Q) - 1 \geq D \deg(B)$. Since $\deg(B) = d^n$ and $\deg(Q) = d^{n-1} \delta D$, the above inequality implies $\frac{d}{\delta} + D > \frac{d}{\delta} D$ which is impossible since $\frac{d}{\delta}$ and D are at least 2. Hence $k_1 = 0$. By (14) and $\gcd(D, d/\delta) = 1$ we have that $B(x)$ is a (d/δ)-th power of a polynomial.

Now both Case (I) and Case (I') give that $B(x) = x^{d/\delta} \circ \tilde{B}$ for some polynomial $\tilde{B}$. Composing $x^{d/\delta}$ to the right of both sides of the functional equation $f^n \circ A = A \circ B$, we have:

$$(15) \qquad f^n \circ A \circ x^{d/\delta} = A \circ x^{d/\delta} \circ \tilde{B} \circ x^{d/\delta}.$$

Since $d \mid \deg(A) d/\delta$, by Lemma 7.5 there is a polynomial $\hat{A}$ such that $A \circ x^{d/\delta} = f \circ \hat{A}$ and $f^{n-1} \circ A \circ x^{d/\delta} = \hat{A} \circ \tilde{B} \circ x^{d/\delta}$. Hence $f^n \circ \hat{A} = \hat{A} \circ \tilde{B} \circ x^{d/\delta}$. Since $\deg(\hat{A}) = \deg(A)/\delta = D < \deg(A)$ and the minimality of $\deg(A)$ among all counter-examples to Theorem 7.1, there exist $m \in \mathbb{N}_0$, a divisor η of d, and a linear polynomial $U(x)$ such that $\hat{A} = f^m \circ (x^\eta + c) \circ U$ and $\tilde{B} \circ x^{d/\delta} = U^{-1} \circ f_\eta^n \circ U$. Since $\deg(\hat{A}) = D$ is coprime to d/δ, we have that $m = 0$ and hence $\eta = D \geq 2$. Now f_η^n is a polynomial in x^η while $\tilde{B} \circ x^{d/\delta}$ is a polynomial in $x^{d/\delta}$, we have that the constant coefficient of $U(x)$ is zero. By

Lemma 7.2, η is the maximum number among all M such that $U^{-1} \circ f_\eta^n \circ U$ is a polynomial in x^M. On the other hand $U^{-1} \circ f_\eta^n \circ U = B \circ x^{d/\delta}$ is a polynomial in $x^{d/\delta}$; we must have that d/δ divides $\eta = D$. This contradicts $\gcd(D, d/\delta) = 1$. It remains to deal with the following:

Case (II): assume that $d/\delta = 2$ and there exist linear polynomials L_1 and L_2 such that $\tilde{A} = L_1 \circ T_D \circ L$; moreover L_1 has the form $\ell_1 x + 2\ell_1$ for some $\ell_1 \neq 0$. As before, replace $\tilde{A}$ by $\tilde{A} \circ L^{-1}$ and B by $L \circ B \circ L^{-1}$, we may assume $\tilde{A} = L_1 \circ T_D = \ell_1 T_D(x) + 2\ell_1$. Since $d/\delta = 2$, equation (13) implies that $\tilde{A} \circ B = \ell_1(T_D(B(x)) + 2)$ is the square of a polynomial. Since $\gcd(D, d/\delta) = \gcd(D, 2) = 1$, we have that -2 is a simple root of $T_D(x) + 2$. Hence we have that $B(x) + 2$ is the square of a polynomial. In other words, there is a polynomial $\tilde{B}$ such that $B(x) = (x^2 - 2) \circ \tilde{B}$. We can proceed as in Case (I) and Case (I') as follows.

Compose $x^2 - 2$ to the right of both sides of $f^n \circ A = A \circ B$, we have:

$$(16) \qquad f^n \circ A \circ (x^2 - 2) = A \circ (x^2 - 2) \circ \tilde{B} \circ (x^2 - 2).$$

Recall that $d/2 = \delta = \gcd(d, \deg(A))$, hence $d \mid \deg(A \circ (x^2 - 2))$. By Lemma 7.5, there is $\hat{A} \in \mathbb{C}[x]$ such that $A \circ (x^2 - 2) = f \circ \hat{A}$ and $f^{n-1} \circ A \circ (x^2 - 2) = \hat{A} \circ \tilde{B} \circ (x^2 - 2)$, hence $f^n \circ \hat{A} = \hat{A} \circ \tilde{B} \circ (x^2 - 2)$. Since $\deg(\hat{A}) = D < \deg(A)$ and the minimality of $\deg(A)$ among all counter-examples to Theorem 7.1, there exist $m \in \mathbb{N}_0$, a divisor η of d, and a linear polynomial $U(x)$ such that $\hat{A} = f^m \circ (x^\eta + c) \circ U$ and $\tilde{B} \circ (x^2 - 2) = U^{-1} \circ f_\eta^n \circ U$. Argue as before, we have that $m = 0$, $\eta = D \geq 2$, the constant coefficient of U is zero, and that $2 = d/\delta$ divides $\eta = D$, contradiction. □

7.2. **Proof of Theorem 1.4.** We have:

Lemma 7.6. *Let $d \geq 2$ be an integer and let $c \in \mathbb{C}$ be such that $x^d + c$ is not linearly conjugate to x^d or $\pm T_d(x)$. Let $\delta > 1$ be a divisor of d and let $g(x) = (x^\delta + c)^{d/\delta}$. We have the following:*

   (i) *$g(x)$ is not linearly conjugate to x^d or $\pm T_d(x)$.*
   (ii) *If $G(x)$ is a polynomial of degree at least 2 having a common iterate with $g(x)$ then $G(x)$ is an iterate of $g(x)$.*

Proof. If $\delta = d$, then $g(x) = x^d + c$ which is not linearly conjugate to x^d or $\pm T_d(x)$, according to our hypothesis. If $\delta < d$, then the conclusion of part (i) follows from the ramification behavior of x^d and $T_d(x)$. It remains to show part (ii).

Let $\tilde{g}$ be a non-linear polynomial of minimal degree such that $\tilde{g}$ commutes with an iterate of g. Let Γ be the group of linear polynomials commuting with an iterate of g. By [Ngu15, Proposition 2.3] (also see [Ngu15, Remark 2.4]), the set of polynomials commuting with an iterate of g is exactly:

$$\{\tilde{g}^m \circ L : \; m \geq 0, \; L \in \Gamma\}.$$

It suffices to show that Γ is trivial and $\tilde{g} = g$.

From [Ngu15, Proposition 2.3(d)], any two polynomials that commute with an iterate of g and have the same degree differ by either a "post-composition" or "pre-composition" by an element of the finite cyclic group Γ. Therefore if $L(x) = ax + b$ is a generator of Γ, we have $g \circ L = L_1 \circ g$ for some $L_1 \in \Gamma$, and $L_1 = L^D$ for some positive integer D. By comparing the coefficients from both sides of $g \circ L = L^D \circ g$, we have that $b = 0$. Hence $g(ax) = a^D g(x)$. This implies that $a^d = a^D$ and $a^{d-\delta} = a^D$, hence a is a δ-th root of unity. This gives $g \circ L = g$, hence $g^m \circ L = g^m$ for every $m \in \mathbb{N}$. Therefore Γ is trivial.

From $g = \tilde{g}^m \circ L$ for some m and $L \in \Gamma$, we have $g = \tilde{g}^m$. We assume that $m \geq 2$ and arrive at a contradiction. Write $\tilde{\delta} = \deg(\tilde{g})$ and $\epsilon = \gcd(\delta, \tilde{\delta})$, we have $\epsilon \geq 2$ since $\tilde{\delta}^m = d$ is divisible by δ. From $\tilde{g}^m = g = x^{d/\delta} \circ (x^{\delta/\epsilon} + c) \circ x^\epsilon$ and Lemma 7.5, we have that $\tilde{g}$ is a polynomial in x^ϵ. Hence there is a maximal $\eta \geq 2$ such that $\tilde{g}$ is a polynomial in x^η. By Lemma 7.2, η is also the maximal number M such that $\tilde{g}^m$ is a polynomial in x^M. Since $\tilde{g}^m = g$ is a polynomial in x^δ, we have that δ divides η. By Lemma 7.5, there is $P \in \mathbb{C}[x]$ such that $\tilde{g} = P \circ (x^\delta + c)$ and $x^{d/\delta} = \tilde{g}^{m-1} \circ P$. By Lemma 7.5 and the fact that $m \geq 2$, we have: $\tilde{g} = x^{\tilde{\delta}} \circ L_1$ for some linear $L_1 \in \mathbb{C}[x]$. Since $\tilde{g}$ is a polynomial in x^δ and $\delta > 1$, we must have that the constant coefficient of L_1 is zero. Hence $\tilde{g}$ is a monomial, so is $g = \tilde{g}^m$ which yields a contradiction. Therefore $m = 1$, $g = \tilde{g}$, and every polynomial G commuting with an iterate of g must be an iterate of g.   □

In view of Theorem 1.1, Theorem 1.4 follows from the following:

Proposition 7.7. *Let $d_1 \geq 2$ and $d_2 \geq 2$ be integers. Let $F(x) = x^{d_1} + c_1$ and $G(x) = x^{d_2} + c_2$ be polynomials that are not linearly conjugate to $\pm T_d(x)$ or x^d. Then there are positive integers ℓ_1 and ℓ_2 such that the self-map $F^{\ell_1} \times G^{\ell_2}$ of $(\mathbb{P}^1)^2$ has a non-trivial periodic curve if and only if $d_1 = d_2$ and $c_1 = \zeta c_2$ for a $(d_1 - 1)$-th root of unity ζ.*

Proof. If $d_1 = d_2$ and $c_1 = \zeta c_2$ for a $(d_1 - 1)$-th root of unity then there is a linear polynomial L such that $F = L \circ G \circ L^{-1}$. The curve $\{(L(x), x) : x \in \mathbb{P}^1\}$ is invariant under $F \times G$. It remains to prove the "only if" part.

For every positive divisor δ (respectively η) of d_1 (respectively d_2), denote $F_\delta(x) = (x^\delta + c_1)^{d_1/\delta}$ (respectively $G_\eta(x) = (x^\eta + c_2)^{d_2/\eta}$).

By Medvedev-Scanlon [MS14, Proposition 2.34], there is a positive integer n and non-constant polynomials A_1, A_2, and B such that $F^{\ell_1 n} \circ A_1 = A_1 \circ B$ and $G^{\ell_2 n} \circ A_2 = A_2 \circ B$. By Theorem 7.1, there exist positive divisors δ of d_1 and η of d_2 such that $F_\delta^{\ell_1 n}$ and $G_\eta^{\ell_2 n}$ are linearly conjugate to each other. Since $F_1(x) = (x + c_1)^{d_1}$ and $G_1(x) = (x + c_2)^{d_2}$ are linearly conjugate to $F(x) = F_{d_1}(x)$ and $G(x) = G_{d_2}(x)$ respectively, it suffices to prove the following.

Claim: if there exist a positive integer n and positive divisors δ of d_1 and η of d_2 such that $2 \leq \min\{\delta, \eta\}$ and $F_\delta^{\ell_1 n}$ and $G_\eta^{\ell_2 n}$ are linearly conjugate to each other then $d_1 = d_2$ and $c_1 = \zeta c_2$ for some $(d_1 - 1)$-th root of unity ζ.

Let $U(x)$ be a linear polynomial such that $F_\delta^{\ell_1 n} = (U^{-1} \circ G_\eta \circ U)^{\ell_2 n}$. By Lemma 7.6, we have that $U^{-1} \circ G_\eta \circ U$ is an iterate of F_δ and $U \circ F_\delta \circ U^{-1}$ is an iterate of G_η. This gives $d_1 = d_2$ (hence $\ell_1 = \ell_2$) and $F_\delta = U^{-1} \circ G_\eta \circ U$. Write $U(x) = ux + w$ and $d := d_1 = d_2$, we have that $w = 0$; otherwise the coefficient of x^{d-1} in $U^{-1} \circ G_\eta \circ U$ is nonzero while the coefficient of x^{d-1} in F_δ is zero. Hence we have: $(x^\delta + c_1)^{d/\delta} = \frac{1}{u}((ux)^\eta + c_2)^{d/\eta}$. It follows immediately that $\delta = \eta$ and $u^{d-1} = 1$. Compare the coefficients of $x^{d-\delta}$ on both sides and note that $\delta = \eta$, we have:

$$\frac{d}{\delta} c_1 = \frac{1}{u} \frac{d}{\delta} u^{d-\delta} c_2.$$

Since u is a $(d-1)$-th root of unity, we finish the proof of the proposition. $\square$

7.3. **Proof of Theorem 1.5.**

Let $\tilde{f}$ be a non-linear polynomial of minimal degree commuting with an iterate of f and let G be the group of linear polynomials commuting with an iterate of f. It is well-known from Ritt's theory that $\tilde{f}$ and f have a common iterate and

$$\{\tilde{f}^m \circ L \colon m \geq 0,\ L \in G\} = \{L \circ \tilde{f}^m \colon m \geq 0,\ L \in G\}$$

which is exactly the set of polynomials commuting with an iterate of f (see, for example, [Ngu15, Proposition 2.3]).

Now assume that C is an irreducible curve in $\mathbb{P}^1 \times \mathbb{P}^1$ that is neither horizontal nor vertical and contains infinitely many Φ-preperiodic points. By Theorem 1.1, C is preperiodic under Φ, hence under the self-map $\tilde{\Phi}(x, y) := (\tilde{f}(x), \tilde{f}(y))$ of $\mathbb{P}^1 \times \mathbb{P}^1$. Let $N \geq 0$ be such that $\tilde{\Phi}^N(C)$ is periodic under $\tilde{\Phi}$. By Medvedev-Scanlon [MS14, Theorem 6.24], $\tilde{\Phi}^N(C)$ is given by the equation $y = g(x)$ or $x = g(y)$ where g commutes with an iterate of f. Assume we have the equation $x = g(y)$ and write $g = \ell \circ \tilde{f}^M$, then C satisfies the equation $\tilde{f}^N(x) = \ell(\tilde{f}^{M+N}(y))$. Similarly, assume we have the equation $y = g(x)$ for $\tilde{\Phi}^N(C)$ and write $g = \ell \circ \tilde{f}^M$, then C satisfies the equation $\tilde{f}^{M+N}(x) = \ell^{-1}(\tilde{f}^N(y))$.

Conversely, if C satisfies an equation of the form $\tilde{f}^n(x) = L(\tilde{f}^m(y))$ where L commutes with an iterate of f then for every $(\alpha, \beta) \in C$, α is f-preperiodic if and only if β is f-preperiodic. This gives that C contains infinitely many Φ-preperiodic points.

REFERENCES

[AKLS09] A. Avila, J. Kahn, M. Lyubich and W. Shen, *Combinatorial rigidity for uni-critical polynomials*, Ann. of Math. (2) **170** (2009), no. 2, 783–797.

[BH05] M. Baker and L.-C. Hsia, *Canonical heights, transfinite diameters, and polynomial dynamics*, J. reine angew. Math. **585** (2005), 61–92.

[Bro65] H. Brolin, *Invariant sets under iteration of rational functions*, Ark. Mat. **6**(1965), 103–144.

[BR06] M. Baker and R. Rumely, *Equidistribution of small points, rational dynamics, and potential theory*, Ann. Inst. Fourier (Grenoble) **56** (2006), 625–688.

[CS93]   G. Call and J. Silverman, *Canonical heights on varieties with morphisms*, Compositio Math. **89**(1993), 163–205.

[CL06]   A. Chambert-Loir, *Mesures et équidistribution sur les espaces de Berkovich*, J. Reine Angew. Math. **595** (2006), 215–235.

[DF16]   R. Dujardin and C. Favre, *The dynamical Manin-Mumford problem for plane polynomial automorphisms*, to appear in Journal of the European Mathematical Society, European Mathematical Society, 2016.

[DH93]   A. Douady and J. Hubbard, *A proof of Thurston's topological characterization of rational functions*, Acta. Math. **171** (1993) 263–297.

[Fak03]   N. Fakhruddin, *Questions on self maps of algebraic varieties*, J. Ramanujan Math. Soc. **18** (2003), no. 2, 109-122.

[FLM83]   A. Freire, A. Lopes and R. Mañé, *An invariant measure for rational maps*, Bol. Soc. Brasil Mat. **14** (1983) 45-62.

[FRL06]   C. Favre and J. Rivera-Letelier, *Équidistribution quantitative des points de petite hauteur sur la droite projective*, Math. Ann. **355** (2006), 311–361.

[GT10]   D. Ghioca and T. J. Tucker, *Proof of a dynamical Bogolov conjecture for lines under polynomial actions*, Proc. Amer. Math. Soc. **138** (2010), 937–942.

[GTZ11]   D. Ghioca, T. J. Tucker and S. Zhang, *Towards a dynamical Manin-Mumford Conjecture*, Int. Math. Res. Not. IMRN 2011, no. 22, 5109–5122.

[GTZ12]   D. Ghioca, T. J. Tucker and M. Zieve, *Linear relations between polynomial orbits*, Duke Math. J. **161** (2012), no. 7, 1379–1410.

[Ham95]   D. Hamilton, *Length of Julia curves*, Pacific J. Math. **169** (1995), no. 1, 75–93.

[Ino11]   H. Inou, *Extending local analytic conjugacies*, Trans. Amer. Math. Soc. **363** (2011), no. 1, 331–343.

[Lan02]   S. Lang, *Algebra*, Revised Third Edition, Graduate Texts in Mathematics vol. 211, Springer, 2002.

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

[Lev90]   G. M. Levin, *Symmetries on Julia sets*, Mat. Zametki **48** (1990), 72–79, 159.

[Lyu83]   M. Lyubich, *Entropy properties of rational endomorphisms of the Riemann sphere*, Erg. Th. and Dyn. Sys. **3** (1983), 351–385.

[Man83]   R. Mañé, *On the Uniqueness of the maximizing measure for rational maps*, Bol. Soc. Bras. Math. **14.2** (1983), 27–43.

[McQ95]   M. McQuillan, *Division points on semi-abelian varieties*, Invent. Math. **120** (1995), no. 1, 143–159.

[Mil00]   J. Milnor, *Dynamics in one complex variable*, Wiesbaden, Germany: Vieweg. ISBN 3-528-13130-6. (2000).

[Mil04]   J. Milnor, *On Lattès maps*, preprint available online at http://arxiv.org/abs/math/0402147.

[Mim13]   A. Mimar, *On the preperiodic points of an endomorphism of $\mathbb{P}^1 \times \mathbb{P}^1$ which lie on a curve*, Trans. Amer. Math. Soc. **365** (2013), no. 1, 161–193.

[Mim97]   A. Mimar, *On the preperiodic points of an endomorphism of $\mathbb{P}^1 \times \mathbb{P}^1$ which lie on a curve*, Thesis (Ph.D.), Columbia University, 1997.

[MS14]   A. Medvedev and T. Scanlon, *Invariant varieties for polynomial dynamical systems*, Ann. of Math. (2) **179** (2014), no. 1, 81–177.

[Ngu15]   K. Nguyen, *Some arithmetic dynamics of diagonally split polynomial maps*, Int. Math. Res. Not. IMRN, vol. 2015, 1159–1199.

[Ray83]   M. Raynaud, *Sous-variétés d'une variété abélienne et points de torsion*, Arithmetic and geometry, vol. I, Progr. Math., vol. 35, Birkhäuser, Boston, MA, 1983, pp. 327–352.

[Thu85]   W. Thurston, *On the combinatorics of iterated rational maps*, preprint, (1985).

[Ull98]   E. Ullmo, *Positivité et discrétion des points algébriques des courbes*, Ann. of Math. (2) **147** (1998), 167–179.

[Yua08]   X. Yuan, *Big line bundles over arithmetic varieties*, Invent. Math. **173** (2008), 603–649.

[YZa]   X. Yuan and S. Zhang, *Calabi theorem and algebraic dynamics*, preprint, 24 pages.

[YZb]   _____, *Small points and Berkovich metrics*, preprint available online at `http://www.math.columbia.edu/~szhang/papers/Preprints.htm`, 28 pages.

[YZc]   _____, *The arithmetic Hodge index theorem for adelic line bundles II*, arXiv:1304.3539.

[YZ16]   _____, *The arithmetic Hodge index theorem for adelic line bundles*, Math. Ann. (2016), 1–49.

[Zdu90]   A. Zdunik, *Parabolic orbifolds and the dimension of the maximal measure for rational maps*, Inv. Math. **99** (1990) 627–649.

[Zha95a]   S. Zhang, *Positive line bundles on arithmetic varieties*, J. Amer. Math. Soc. **8** (1995), 187–221.

[Zha95b]   _____, *Small points and adelic metrics*, J. Alg. Geom. **4** (1995), 281–300.

[Zha98]   _____, *Equidistribution of small points on abelian varieties*, Ann. Math. (2), **147**(1998), 159–165.

[Zha06]   _____, *Distributions in algebraic dynamics*, Survey in Differential Geometry, vol. 10, International Press, 2006, pp. 381–430.

DRAGOS GHIOCA, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BRITISH COLUMBIA, VANCOUVER, BC V6T 1Z2, CANADA
   E-mail address: `dghioca@math.ubc.ca`

KHOA D. NGUYEN, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BRITISH COLUMBIA, AND PACIFIC INSTITUTE FOR THE MATHEMATICAL SCIENCES, VANCOUVER, BC V6T 1Z2, CANADA
   E-mail address: `dknguyen@math.ubc.ca`

HEXI YE, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BRITISH COLUMBIA, VANCOUVER, BC V6T 1Z2, CANADA
   E-mail address: `yehexi@math.ubc.ca`