The Orbit Intersection Problem and Polynomial Functions

by

Andrew O'Desky

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mathematics)
in The University of Michigan
2020

Doctoral Committee:

Professor Michael Zieve, Chair
Associate Professor Wei Ho
Professor Finn Larsen
Professor Kartik Prasanna

Andrew O'Desky

aodesky@umich.edu

ORCID iD: 0000-0003-1068-7013

This dissertation is dedicated to the memory of Earl Van Der Zee Gordon.

ACKNOWLEDGEMENTS

I wish to thank everyone who helped me to write this dissertation. In particular, I am grateful to my advisor, Professor Michael Zieve, for suggesting the problem which became the second chapter of this dissertation, and for everything that he has taught me. I am very fortunate to have had the opportunity to learn from many wonderful peers and teachers at the University of Michigan. Finally I wish to express my gratitude to my friends and family for their constant support and encouragement.

TABLE OF CONTENTS

ABSTRACT

The first part of this thesis considers the problem of reconstructing a rational function f from one of its orbits, $\{a, f(a), f(f(a)), f(f(f(a))), \ldots\}$. Conjecturally, two complex rational functions of degree > 1 possess orbits with infinite intersection if and only if they have a common iterate. This conjecture has been recently verified in the polynomial case, however the rational function case is vastly more difficult. Our first main theorem verifies this conjecture for rational functions of coprime degree, and is the first to address intersections of orbits of rational functions which are not polynomials.

The second part of this thesis considers the problem of characterizing polynomial functions s using only their values on natural numbers, $(s(0), s(1), s(2), \ldots)$. Our second main theorem proves that if the mth divided difference $\delta_m s$ of an arbitrary sequence s of rational numbers is integer-valued, then $s(n)$ is given by a polynomial in n if and only if there is a positive number θ with $s(n) \ll \theta^n$ and $1+\theta < e^{1+\frac{1}{2}+\cdots+\frac{1}{m}}$.

CHAPTER I

Introduction

This thesis presents the results of two directions of the author's graduate research conducted at the University of Michigan. We begin with some background which will provide supporting context for Chapter II. The reader who would like to get quickly to the main results can skip to §1.2.2 for the statement of the Orbit Intersection Problem, §1.2.3 for the statements of our results in Chapter II, and §1.3 for our main results in Chapter III.

1.1 Historical Background

The purpose of this section is to explain the statement of an outstanding question in arithmetic dynamics, the *Dynamical Mordell–Lang Problem*. The Orbit Intersection Problem is a special case of the Dynamical Mordell–Lang Problem. To properly motivate the statement, we first explain the series of antecedent diophantine results from the 20th century which led to its formulation.

1.1.1 Diophantine Geometry

We recall some fundamental constructions. Let n be a non-negative integer. We define complex affine space $\mathbb{A}^n$ to be $\mathbb{C}^n$. We define complex projective space $\mathbb{P}^n$ to be the equivalence classes of $\mathbb{A}^{n+1}\backslash\{0\}$ under the relation $x \sim y$ if $y = rx$ for some

nonzero number $r \in \mathbb{C}$.

An affine algebraic set X is the solution set in complex affine space of a family of equations given by polynomials $f_i \colon \mathbb{C}^n \to \mathbb{C}$, $1 \leq i \leq m$,

$$X = \{x \in \mathbb{A}^n : f_i(x) = 0 \text{ for all } 1 \leq i \leq m\}.$$

A projective algebraic set X is the solution set in complex projective space of a family of homogeneous polynomials $g_i \colon \mathbb{C}^{n+1} \to \mathbb{C}$, $1 \leq i \leq m$,

$$X = \{[x] \in \mathbb{P}^n : g_i(x) = 0 \text{ for all } 1 \leq i \leq m\}.$$

We equip affine (resp. projective) space with the weakest topology for which every affine (resp. projective) algebraic subset is closed, and we equip algebraic sets with the corresponding subspace topologies. A topological space X is irreducible if it cannot be expressed as a union of two nonempty proper closed subsets. For us, a variety will be any open and irreducible subset X of a projective algebraic set.[1] Concretely, X is the subset of $\mathbb{P}^n$ where one family of homogeneous polynomials is zero, another finite family of homogeneous polynomials is nonzero, and X cannot be decomposed into two smaller such subsets. We say that a variety X is defined over a field K if there exist defining polynomials for X which have coefficients in K. A variety X is affine (resp. projective) if it is an affine (resp. projective) algebraic set. From here on out we assume some fundamental notions from algebraic geometry such as dimension, (geometric) genus, smoothness, birationality, and morphisms of varieties. For definitions, see [Har77, §1].

It is very interesting — and very classical — to ask for a description of the subset of points whose coordinates belong to a subring R of $\mathbb{C}$ such as $\mathbb{Z}$ or $\mathbb{Q}$. Precisely, let us write $\mathbb{P}^n(R)$ for the subset of points $[x] \in \mathbb{P}^n$ for which there exists a representative

[1] Strictly speaking, this is the definition of a quasi-projective (irreducible) variety over $\mathbb{C}$.

$(x_0, \ldots, x_n) \in [x]$ with all coordinates $x_0, \ldots, x_n$ in R. We define

$$X(R) = \mathbb{P}^n(R) \cap X.$$

For a finitely-generated subfield $K \subset \mathbb{C}$, the subset $X(K)$ is often called the K-rational points of X, and often $\mathbb{Q}$-rational points are simply referred to as rational points. The study of the K-rational points of a variety X for a finitely-generated subfield $K \subset \mathbb{C}$ is the subject of diophantine geometry. All of the results we will discuss in this section remain true when $\mathbb{Q}$ is replaced with a finitely-generated subfield of $\mathbb{C}$, so to simplify the exposition and some of the history we will usually restrict the discussion to $K = \mathbb{Q}$ unless otherwise stated.

Group varieties are of particular interest in diophantine geometry. These are the analogue of Lie groups in algebraic geometry. A group variety is a variety G (defined over K, say) equipped with a composition morphism $G \times G \to G : (P, Q) \mapsto P \cdot Q$, an inverse morphism $G \to G : P \mapsto P^{-1}$, and a distinguished identity element $e \in G$, all defined over K and satisfying the usual axioms. Group varieties play a special role in diophantine geometry for the simple reason that when P and Q are K-rational points, $P \cdot Q$ and P^{-1} are also K-rational, i.e., $G(K)$ is a subgroup of $G(\mathbb{C})$. Thus even a single K-rational point generally leads to a multiplicity of them. Furthermore, even when a variety is not itself a group variety there is often a natural action of a group variety by symmetries which helps to produce more K-rational points from a single one.

An abelian variety is a group variety whose underlying variety is projective, and an elliptic curve is an abelian variety of dimension one. Abelian varieties play a special role in the theory. A group variety is said to be linear if its underlying variety is affine.[2] Abelian varieties and linear group varieties are the most important examples

[2] It is known that every linear group variety arises as a closed subgroup of GL_n for some n.

of group varieties. For example, we mention a theorem of C. Chevalley[3] which asserts that every group variety G is canonically an extension of an abelian variety A by a linear group variety H, i.e., there is a unique normal linear closed subgroup $H \subset G$ such that G/H is an abelian variety. Two other ubiquitous group varieties are the additive group $\mathbb{G}_a$, which denotes $\mathbb{A}^1$ with additive composition, and the multiplicative group $\mathbb{G}_m$, which denotes $\mathbb{A}^1 \backslash \{0\}$ with multiplicative composition. An algebraic torus is a group variety T which is isomorphic over $\mathbb{C}$ to a product of $\mathbb{G}_m$s.

The important role of group varieties in diophantine geometry can already be seen in the one-dimensional setting. In 1922 L. Mordell made a conjecture [Mor22] which in 1983 became a theorem due to G. Faltings [Fal83].

Theorem (Faltings)**.** *Let C be a projective algebraic curve defined over a finitely-generated field K of characteristic zero. If the genus of C is ≥ 2 then $C(K)$ is finite.*

Using Faltings's theorem one can easily show that the projective curves C with the property that $C(K)$ is infinite for some finitely-generated field K may also be characterized as those projective curves which contain a nontrivial group variety. The point is that the projective curves which contain a nontrivial group variety are precisely the projective curves of genus ≤ 1, and Faltings's theorem implies that the projective curves of genus ≤ 1 are also the projective curves such that $C(K)$ is infinite for some finitely-generated field K.

We mention in passing that this observation conjecturally generalizes to higher dimensions. Let $Sp(X)$ denote the Zariski closure of the union of all images of non-constant rational maps

$$G \dashrightarrow X$$

[3] For a proof in modern scheme-theoretic terminology see [Con02].

over all group varieties G, and let $W = X \backslash Sp(X)$. Then W is defined over a finitely-generated field K_0 containing $\mathbb{Q}$. S. Lang has conjectured that $W(K)$ is finite for any finitely-generated field K containing K_0 [Lan91, (3.2)]. In fact, by Chevalley's theorem any point of a group variety G meets a coset of a linear group variety H. It is well-known that any linear group variety is birational to projective space (possibly after extending the basefield), and so for the definition of $Sp(X)$ it suffices to consider the images of non-constant rational maps of the form $\mathbb{P}^n \dashrightarrow X$ and $A \dashrightarrow X$. Projective space is rationally connected (any two points are joined by a rational curve), and so it even suffices to consider rational maps $\mathbb{P}^1 \dashrightarrow X$, which can then be replaced with $E \dashrightarrow X$ for an elliptic curve E. In sum, all but finitely many rational points of an arbitrary variety X are conjectured to lie in the Zariski closure of the images of non-constant rational maps $A \dashrightarrow X$ from an abelian variety A. For more on the set $Sp(X)$ and its relation to geometric properties of X we refer to [Lan91, §I.3].

1.1.2 The Mordell–Weil theorem

Abelian varieties play a central role in the study of rational points in general so it is worthwhile to consider the structure of their rational points in more detail. The subset of K-rational points $A(K)$ of an abelian variety A defined over K is called the Mordell–Weil group of A. The fundamental statement about $A(K)$ is known as the *Mordell–Weil theorem.*

Theorem (Mordell–Weil)**.** *Let A be an abelian variety defined over a number field K. Then the abelian group $A(K)$ is finitely generated.*

The Mordell–Weil theorem can be used to study rational points on curves by means of the jacobian variety. The situation can be described a priori. Let C be a

projective algebraic curve defined over a finitely-generated field K. We may assume that C is smooth. (If C is singular then there is another projective algebraic curve C' defined over K which is smooth and birational to C and such that $C'(K)$ is infinite if and only if $C(K)$ is infinite. Hence for finiteness statements we may as well assume that C is smooth.) The key is to realize C as a subvariety of an abelian variety, as it turns out this can be done in an essentially canonical way. For curves of genus zero the construction turns out to be trivial, but their rational points are parameterized by a rational map so their rational points are easily understood anyway. Hence we may assume that the genus g of C is ≥ 1.

We explain the classical construction of the jacobian variety of C. By smoothness, C inherits the structure of a complex manifold from its embedding into complex projective space. Let $\omega_1, \ldots, \omega_g$ be a basis of $\Omega^1(C)$, the vector space of holomorphic one-forms on C. Choose a point $P_0 \in C$. The map

$$C \ni P \mapsto \int_{P_0}^{P} \omega_1 \in \mathbb{C}$$

is multi-valued on account of monodromy, i.e., the integral depends on the path chosen between P_0 and P. However, the form ω_1 is closed as $\Omega^2(C) = 0$, and so the ambiguity in the integral $\int_{P_0}^{P} \omega_1$ due to choice of path is only up to homotopy, hence up to a finitely-generated additive subgroup of $\mathbb{C}$, namely $\{\int_\gamma \omega_1 : \gamma \in \pi_1(C, P_0)\}$. In this way we obtain the Abel–Jacobi map,

$$\begin{aligned} C &\to \mathbb{C}^g/\Lambda \\ P &\mapsto \Big(\int_{P_0}^{P} \omega_1, \ldots, \int_{P_0}^{P} \omega_g \Big), \end{aligned}$$

where $\Lambda := \{\int_\gamma \omega_i : \gamma \in \pi_1(C, P_0), 1 \leq i \leq g\}$. It can be shown that Λ is an abelian group of rank $2g$ so that $J_C := \mathbb{C}^g/\Lambda$ is a compact complex manifold of dimension g. One can show that the lattice Λ satisfies the "Riemann period relations", whence J_C

can be given the structure of a complex projective algebraic variety with the help of theta functions, and furthermore that the map $C \to J_C$ so obtained is an embedding of complex varieties defined over the field of definition of the chosen point P_0. This shows that J_C is an abelian variety, the jacobian variety of C. For more details see [Mum07].

1.1.3 Classical Mordell–Lang

With the Abel–Jacobi embedding $C \subset J_C$ in mind it is only natural to look for a generalization of Mordell's conjecture by making use of the ambient jacobian variety. Moreover there is no reason to restrict to abelian varieties arising as the jacobian variety of some algebraic curve.[4] The two important data are the curve C, considered as a subvariety of J_C, and the finitely-generated subgroup $J_C(K)$ of K-rational points (we assume that the chosen point P_0 defining the embedding is defined over K). In the context of the abelian variety J_C, Mordell's conjecture is that $C(K) = C \cap (J_C(K))$ is finite if the genus of C is ≥ 2. Equivalently, $C \cap (J_C(K))$ is infinite only when C is genus one, in which case C is a translate of an abelian subvariety of J_C (of dimension one). As early as 1960 Lang considered the following replacements:

$$J_C \rightsquigarrow \text{any abelian variety } A,$$

$$J_C(K) \rightsquigarrow \text{any finitely-generated subgroup } \Gamma \text{ of } A,$$

$$C \rightsquigarrow \text{any closed subvariety } Y \text{ of } A.$$

This led S. Lang to make the following generalization of Mordell's conjecture in 1960 which was proven by G. Faltings in 1994.

[4] The question of which abelian varieties arise as the jacobian of an algebraic curve is known as the Schottky problem, after F. Schottky. For a nice survey see [Gru12]. It was only fairly recently shown that there exists an abelian variety over $\overline{\mathbb{Q}}$ which is not isogeneous to a jacobian, [Tsi12].

Theorem 1.1.1 (Classical Mordell–Lang conjecture for abelian varieties, [Lan60], [Fal91], [Fal94]). *Let A be an abelian variety defined over a field K that is finitely generated over $\mathbb{Q}$. Let $Y \subset A$ be a closed subvariety. Then $Y(K)$ is a finite union of cosets of subgroups of $A(K)$.*

The above statement holds verbatim for an algebraic torus T in place of the abelian variety. This being known at the time (at least in the case Y is of dimension 1), Lang conjectured that for G *either* an abelian variety or a torus, $Y \subset G$ a closed[5] subvariety, and Γ a finitely generated subgroup of G, the intersection $\Gamma \cap Y$ is equal to a finite union of cosets of subgroups of Γ.[6] Lang also asked whether his conjecture would hold for a group variety S fitting into a short exact sequence of group varieties $1 \to T \to S \to A \to 1$ for an abelian variety A; such group varieties are known as semiabelian varieties. In 1996 this became a theorem of P. Vojta.[7]

Theorem 1.1.2 (Classical Mordell–Lang conjecture for semiabelian varieties, [Voj96]). *Let S be a semiabelian variety over a number field K. Let $Y \subset S$ be a closed subvariety and let $\Gamma \subset S$ be a finitely generated subgroup. Then $\Gamma \cap Y$ is a finite union of cosets of subgroups of Γ.*

There exists another variant of the classical Mordell–Lang conjecture in connection

[5] While Lang does not explicitly state that the subvariety Y is closed this hypothesis is necessary and it is included in [Fal91], [Fal94], and [Voj96]. (For example, let $z \neq 0, 1$, let Y be the twice-punctured plane $\mathbb{C}\backslash\{0, z\}$ inside $\mathbb{G}_m$, and let Γ be any infinite finitely-generated subgroup contained in Y. Then $Y \cap \Gamma = \Gamma$ but the only cosets of group subvarieties of $\mathbb{G}_m$ contained in Y are singletons. A similar idea shows the hypothesis is also necessary for abelian varieties.) For [BGT16] a subvariety is assumed to be closed by convention, cf. (3.1.1).

[6] This formulation is more suitable for our dynamical interpretation, which is why we follow [BGT16] in stating Lang's conjecture this way, but it's worth pointing out that Lang actually states his conjecture differently in [Lan60, p29]. The assertion there is that the intersection $\Gamma \cap Y$ is *contained* in a finite union of cosets of *group subvarieties of G contained in Y*.

The two formulations are seen to be equivalent as follows. If $\Gamma \cap Y$ is equal to a finite union $\cup_i \gamma_i \Gamma_i$ of cosets of subgroups $\Gamma_i \subset \Gamma$ then the Zariski closure H_i of Γ_i is a subgroup variety of G (cf. [Mil17, (1.40)]). Each coset $\gamma_i H_i = \overline{\gamma_i \Gamma_i}$ is contained in Y since Y is closed. This proves the original formulation. Conversely, suppose $\Gamma \cap Y$ is contained in a finite union $\cup_i g_i H_i$ of cosets $g_i H_i \subset Y$ of subgroup varieties H_i of G. We may suppose that each intersection $(g_i H_i) \cap \Gamma$ is nonempty. Fix $\gamma_i \in (g_i H_i) \cap \Gamma$ for every i, and set $\Gamma_i := \Gamma \cap H_i$. Let $\gamma \in \Gamma \cap Y$. By assumption, $\gamma \in g_i H_i$ for some i. Then $\gamma \gamma_i^{-1} \in \Gamma \cap H_i = \Gamma_i$. In particular, $\gamma \in \gamma_i \Gamma_i$. This shows that $\Gamma \cap Y \subset \cup_i \gamma_i \Gamma_i$. However, $\gamma_i \Gamma_i \subset g_i H_i \subset Y$, and also $\gamma_i \Gamma_i \subset \Gamma$, showing that $\cup_i \gamma_i \Gamma_i \subset \Gamma \cap Y$. This shows that $\Gamma \cap Y = \cup_i \gamma_i \Gamma_i$, thereby proving the second formulation.

[7] We have stated Vojta's theorem following [BGT16, (3.4.2.1)], hence in terms of the reformulation of Lang's conjecture (cf. footnote 6), though Vojta's theorem was originally stated in terms of Lang's original formulation.

with a generalization of the conjecture in [Lan60] again due to Lang. In [Lan65] Lang formulated a conjecture which encompassed the earlier conjecture in [Lan60] as well as the Manin-Mumford conjecture. The idea is to replace the finitely generated subgroup Γ with its division hull, i.e., the group Γ' of all $a \in S$ such that $a^n \in \Gamma$ for some positive integer n. (The Manin-Mumford conjecture is obtained by taking $\Gamma = 1$, in which case Γ' is the group of torsion points of S.) This general form of the Mordell–Lang conjecture was proven by M. McQuillan [McQ95] following the earlier works of Faltings and Vojta mentioned above, as well as earlier works of M. Hindry, M. Laurent, and M. Raynaud. This is sometimes called the *full form* of the Mordell–Lang conjecture while the form of the Mordell–Lang conjecture proven by Faltings and Vojta is sometimes called the *Mordellic part* of the Mordell–Lang conjecture. There has not been a dynamical form of Mordell–Lang proposed for the full form of the Mordell–Lang conjecture, so we will focus on the Mordellic part in what follows.

1.2 Dynamical Mordell–Lang

The classical Mordell–Lang conjecture gives much more than a condition for finiteness of $\Gamma \cap Y$. To illustrate, we consider a reformulation of this theorem "in coordinates". Let e be the identity element of S. Write $\Gamma = \langle \gamma_1, \ldots, \gamma_r \rangle$, and assume that the generators have been chosen so that every element g of Γ is of the form $\gamma_1^{n_1} \cdots \gamma_r^{n_r}$ for some non-negative integers $n_1, \ldots, n_r$.

Now we make the first conceptual shift from diophantine geometry to algebraic

dynamics. Define the endomorphisms[8]

$$f_i \colon S \to S \qquad a \mapsto a\gamma_i. \tag{1.1}$$

The semiabelian variety S is commutative (cf. [Iit77, Lemma 4]) so the endomorphisms $f_1, \dots, f_r$ commute under composition. Let $f^{\circ n} = f \circ \cdots \circ f$ denote the n-fold iterate of an endomorphism f under function composition. As $(f_1^{\circ n_1} \circ \cdots \circ f_r^{\circ n_r})(e) = \gamma_1^{n_1} \cdots \gamma_r^{n_r}$, the finitely generated subgroup Γ is simply the forward orbit of e under the endomorphisms $f_1, \dots, f_r$, i.e.,

$$\mathcal{O} := \mathcal{O}_{f_1,\dots,f_r}(e) := \{(f_1^{\circ n_1} \circ \cdots \circ f_r^{\circ n_r})(e) : (n_1, \dots, n_r) \in \mathbb{N}^r)\} = \Gamma.$$

With this shift in emphasis, *we change focus from the finitely generated subgroups Γ of S to the (forward) orbits $\mathcal{O}$ of a commuting family of endomorphisms.*

To keep track of the relationship between the orbit $\mathcal{O}$ and the subvariety Y we can consider the set of indices which iterate the initial point into Y, i.e.,

$$Z(Y, \mathcal{O}) := \{(n_1, \dots, n_r) \in \mathbb{N}^r : (f_1^{\circ n_1} \circ \cdots \circ f_r^{\circ n_r})(e) \in Y\}.$$

We can now state a dynamical reformulation of the classical Mordell–Lang conjecture.

Theorem 1.2.1 (Dynamical form of [Voj96])**.** *Let S be a semiabelian variety, $Y \subset S$ a closed subvariety, and $\mathcal{O} = \mathcal{O}_{f_1,\dots,f_r}(e)$ the forward orbit of a family of commuting endomorphisms $f_1, \dots, f_r$ of S of the form (1.1). Then there exist subgroups $G_1, \dots, G_s$ of $\mathbb{Z}^r$ and elements $\underline{n}_i \in \mathbb{N}^r$ such that*

$$Z(Y, \mathcal{O}) = \bigcup_{i=1}^{s} \big(\underline{n}_i + (G_i \cap \mathbb{N}^r)\big).$$

[8] These are not endomorphisms in the group sense but as self-maps of A as an algebraic variety.

Why shift emphasis from subgroups to orbits? One motivation is to generalize Vojta's theorem. As it turns out, there are many other settings where the dynamical reformulation is true which are unrelated to semiabelian varieties.

Example 1. Consider the endomorphism

$$\Phi\colon \mathbb{A}^1 \to \mathbb{A}^1$$

$$y \mapsto y^3.$$

Let $\mathcal{O}$ be the orbit of $a = i$ under Φ. Then $\Phi^{\circ n}(a) = i^{3^n}$ and $Z(\{i\}, \mathcal{O})$ is the set of n such that $i^{3^n} = i$. As $3^n \equiv (-1)^n \pmod 4$, this shows that

$$Z(\{i\}, \mathcal{O}) = 2\mathbb{N}.$$

However, the following examples show that Z is not always of this form, even when the ambient variety is a semiabelian variety. Both examples are from [BGT16], §3.6.

Example 2. Consider the commuting endomorphisms

$$\Phi_1\colon \mathbb{A}^2 \to \mathbb{A}^2$$

$$(x, y) \mapsto (x + 1, y)$$

and

$$\Phi_2\colon \mathbb{A}^2 \to \mathbb{A}^2$$

$$(x, y) \mapsto (x, y^2).$$

Let Δ be the diagonal subvariety of $\mathbb{A}^2$ and let $\mathcal{O}$ be the orbit of $a = (1, 2)$ under Φ_1, Φ_2. Then $\Phi_1^{\circ m}\Phi_2^{\circ n}(a) = (m + 1, 2^{2^n})$ and $Z(\Delta, \mathcal{O})$ consists of the set of (m, n) such that $m + 1 = 2^{2^n}$. This shows that

$$Z(\Delta, \mathcal{O}) = \{(2^{2^m} - 1, m) : m \in \mathbb{N}\}.$$

Example 3. Consider the commuting endomorphisms

$$\Phi_1 \colon \mathbb{G}_m^3 \to \mathbb{G}_m^3$$
$$(x, y, z) \mapsto (x^2 y^{-1}, y^2 z^{-2}, z^2)$$

and

$$\Phi_2 \colon \mathbb{G}_m^3 \to \mathbb{G}_m^3$$
$$(x, y, z) \mapsto (x^2 y^2, y^2 z^4, z^2).$$

Let Y be the subvariety of $\mathbb{G}_m^3$ defined by $x = 1$, and let $\mathcal{O}$ be the orbit of $(1, 1/3, 9)$ under Φ_1, Φ_2. Then

$$Z(\Delta, \mathcal{O}) = \{(12m^2 \pm 6m, 6m^2) : m \in \mathbb{N}\}.$$

1.2.1 The Dynamical Mordell–Lang Problem

Now we make the second conceptual shift from diophantine geometry to algebraic dynamics: *The semiabelian variety S and its commuting set of endomorphisms are replaced with a general commutative dynamical system.* We may now state the *Dynamical Mordell–Lang Problem.*

1.2.1 (Dynamical Mordell–Lang Problem)**.** *Let X be a variety and let $f_1, \ldots, f_r$ be a commuting family of endomorphisms of X. Let Y be a closed subvariety of X, and choose an initial point $a \in X$. Let $\mathcal{O}$ be the forward orbit of a under the endomorphisms $f_1, \ldots, f_r$. Under what conditions — on X, $Y \subset X$, $f_1, \ldots, f_r$, and the initial point a — do there exist subgroups $G_1, \ldots, G_s$ of $\mathbb{Z}^r$ and elements $\underline{n}_1, \ldots, \underline{n}_s \in \mathbb{N}^r$ such that*

$$Z(Y, \mathcal{O}) = \bigcup_{i=1}^{s} \big(\underline{n}_i + (G_i \cap \mathbb{N}^r)\big)?$$

In this generality, this problem first appeared in a paper of D. Ghioca, T. Tucker, and M. Zieve, [GTZ12]. In the case of a single endomorphism, the Dynamical Mordell–Lang Problem asks for conditions ensuring that

$$Z(Y, \mathcal{O}) = \{n \in \mathbb{N} : f^{\circ n}(a) \in Y\}$$

is a finite union of arithmetic progressions. This special case is sometimes called the *cyclic case of the Dynamical Mordell–Lang Problem*, or also the *Dynamical Mordell–Lang Conjecture*. It is conjectured by Ghioca-Tucker [GT09] to hold without any further assumptions on X, Y, f, or a.

1.2.2 The Orbit Intersection Problem

We are interested in a natural specialization of the Dynamical Mordell–Lang Problem. For the ambient variety previously denoted with X we will take the product variety X^r. This has the natural subvariety given by the diagonal,

$$\Delta \subset X^r.$$

For this special case of the Dynamical Mordell–Lang Problem it is easy to construct commuting families of endomorphisms. We take endomorphisms $F_1, \ldots, F_r \in \operatorname{End} X$ (which do not necessarily commute), and use the commuting endomorphisms

$$f_i = (\mathrm{id}_X, \mathrm{id}_X, \ldots, F_i, \ldots, \mathrm{id}_X) \in \operatorname{End}(X^r).$$

This situation leads to the *Orbit Intersection Problem*. The reason for the terminology is that

$$Z(X^r, \Delta) = \{(n_1, \ldots, n_r) \in \mathbb{N}^r : F_1^{\circ n_1}(a) = F_2^{\circ n_2}(a) = \cdots = F_r^{\circ n_r}(a)\}.$$

The Dynamical Mordell–Lang Problem now takes the following form.

1.2.1 (Orbit Intersection Problem, split case of the Dynamical Mordell–Lang Problem)**.** *Let X be a variety, let $F_1, \dots, F_r$ be (not necessarily commuting) endomorphisms of X, and choose an initial point $a \in X$. Under what conditions — on X, $F_1, \dots, F_r$, and the initial point a — do there exist subgroups $G_1, \dots, G_s$ of $\mathbb{Z}^r$ and elements $\underline{n}_1, \dots, \underline{n}_s \in \mathbb{N}^r$ such that*

$$\{(n_1, \dots, n_r) \in \mathbb{N}^r : F_1^{\circ n_1}(a) = \cdots = F_r^{\circ n_r}(a)\} = \bigcup_{i=1}^{s} \big(\underline{n}_i + (G_i \cap \mathbb{N}^r)\big)?$$

1.2.3 Statements of Main Results in Chapter II

In Chapter II we study the Orbit Intersection Problem for curves. The results in this chapter are joint work with Michael Zieve [OZ]. Our main result concerns the intersection of orbits of endomorphisms of $\mathbb{P}^1$.

Theorem 1.2.1. *Let F and G be endomorphisms of $\mathbb{P}^1$ whose degrees are greater than one. Let $\mathcal{O}_F$ (resp. $\mathcal{O}_G$) be any orbit of F (resp. G). If the degrees of F_1 and F_2 are coprime then $\mathcal{O}_F \cap \mathcal{O}_G$ is finite.*

This has the following corollary for the Orbit Intersection Problem for $X = \mathbb{P}^1$.

Corollary 1.2.2. *Let $F_1, \dots, F_r$ be endomorphisms of $\mathbb{P}^1$ whose degrees are greater than one. Let a be a point of $\mathbb{P}^1$. If the degrees of the endomorphisms are pairwise coprime then*

$$\{(n_1, \dots, n_r) \in \mathbb{N}^r : F_1^{\circ n_1}(a) = \cdots = F_r^{\circ n_r}(a)\}$$

is a finite union of arithmetic progressions.

For curves of positive genus we obtain a characterization for when two endomorphisms have a common iterate.

Theorem 1.2.3. *Let Φ and Ψ be non-invertible endomorphisms of an algebraic curve X of positive genus over a field of characteristic zero. Then there exist orbits of Φ and Ψ with infinite intersection if and only if Φ and Ψ have a common iterate.*

Putting these theorems together, we obtain a solution to the Orbit Intersection Problem for curves which is complete for positive genus.

Corollary. *Let $F_1, \ldots, F_r$ be endomorphisms of a curve X whose degrees are greater than one. Let a be a point of $\mathbb{P}^1$. If the genus of X is zero then further assume that the degrees of the endomorphisms are pairwise coprime. Then*

$$\{(n_1, \ldots, n_r) \in \mathbb{N}^r : F_1^{\circ n_1}(a) = \cdots = F_r^{\circ n_r}(a)\}$$

is a finite union of arithmetic progressions.

1.3 Polynomials with Integral Divided Differences

The second part of this thesis is in the field of number theory. We are interested in a classical problem of interpolation. Let $s(0), s(1), s(2), \ldots$ be a sequence of rational numbers. Under what conditions on s does there exist a polynomial $f \in \mathbb{Q}[x]$ such that $f(n) = s(n)$ for $n = 0, 1, 2, \ldots$?

We approach this general problem by means of the divided differences of s. Divided differences are a fundamental construction of interpolation theory and non-Archimedean analysis. They are of great utility for numerical applications and applied mathematics, but their theoretical applications are much less explored. Let us recall their definition for the unacquainted reader. For more we refer the reader to [MT51] for their use in interpolation theory, and [Sch06] for their use in non-Archimedean analysis.

The mth divided difference $\delta_m s$ of s is the function of distinct nonnegative integers $n_0, \ldots, n_m$ given by

$$\delta_m s(n_0, \ldots, n_m) := \sum_{i=0}^{m} \Big\{ \prod_{j \neq i} (n_i - n_j)^{-1} \Big\} s(n_i).$$

It is a symmetric function on the complement of the union of the diagonal hyperplanes $\{n_i = n_j\}$ in $\mathbb{N}^{m+1}$. Let $\{n_0, n_1, \ldots\} \subset \mathbb{N}$ be any infinite subset. Newton's interpolation formula gives a formal interpolation series for s in terms of its divided differences:[9]

$$s(x) = s(n_0) + \sum_{k=1}^{\infty} \delta_k s(n_0, n_1, \ldots, n_k) \prod_{j=0}^{k-1} (x - n_j) \text{ for all } x \in \{n_0, n_1, \ldots\}. \tag{1.2}$$

This formula should be viewed as a kind of formal Taylor series for the sequence s.

In this chapter we prove the following result which uses the integrality of divided differences to give a criterion for s to be polynomial. Let K be an algebraic number

[9] This equality is to be interpreted in the following formal sense: for each $x \in \mathbb{N}$, the right-hand side of (1.2) is a finite sum that is equal to $s(x)$. For more on Newton's interpolation formula we refer the reader to [MT51, §1].

field of degree d with ring of integers $\mathcal{O}$. We write $r_n \ll q_n$ to mean there is a positive constant C such that $|r_n| \leq C|q_n|$ for all $n \geq 0$.

Theorem 1.3.1. *Let $s\colon \mathbb{N} \to K$. Suppose that*

(i) $\delta_m s$ is $\mathcal{O}$-valued, and

(ii) for each embedding $\sigma\colon K \to \mathbb{C}$, $|\sigma s(n)| \ll \theta_\sigma^n$ for some positive number θ_σ and

$$\prod_{\sigma:K\to\mathbb{C}} (1+\theta_\sigma) < e^{d\left(1+\frac{1}{2}+\cdots+\frac{1}{m}\right)}.$$

Then $s(n)$ is a polynomial in n.

In §3.3 we will prove Theorem 1.3.1 as a very special case of Theorem 3.3.2. The special case of Theorem 1.3.1 when $m = 1$ and $K = \mathbb{Q}$ was discovered independently in 1971 by R. Hall and I. Ruzsa, [Hal71], [Ruz71].

The integrality of divided differences has interesting arithmetic consequences. For simplicity suppose s is valued in $\mathbb{Q}$. The integrality of the zero-th divided difference, $\delta_0 s(n) = s(n)$, simply requires that s itself is integer-valued. The first divided difference of s is equal to

$$\delta_1 s(m,n) := \frac{s(m)-s(n)}{m-n} \quad (m \neq n).$$

The integrality of $\delta_1 s(m,n)$ for all distinct pairs of nonnegative integers m, n is equivalent to requiring that $m-n$ divides $s(m)-s(n)$ for all nonnegative integers m, n, which simply means that s preserves all congruences between nonnegative integers. The integrality of higher divided differences can be interpreted using the subspace topology on the natural numbers $\mathbb{N}$ inherited from its inclusion into the adele ring $\mathbb{A}$ (see §1.3.3).

For the interested reader, we mention that when $K = \mathbb{Q}$ and $m = 1$, Ruzsa conjectured that the second condition could be improved by replacing e with $e+1$

so that the inequality becomes $\theta < e$. Ruzsa's conjecture remains open, though there has been some partial progress, [Zan82], [PZ84], [Zan96]; also see [BN19] for a function field analogue. It is interesting to consider how our inequality might be likewise improved, though we do not venture any guesses here.

1.3.1 Outline of the Proof of Theorem 1.3.1

Our approach to Theorem 1.3.1 uses two new results, one local and one global. The local result concerns the special values of higher divided differences at consecutive integers. Recall that the finite differences c of s are defined by

$$c(n) = \delta_n(0, 1, \ldots, n)n!.$$

We will show that the hypotheses of Theorem 1.3.1 imply that $c(n)$ is eventually zero, thus proving that s is polynomial by Newton's interpolation formula (1.2). To state our results in the local part, we assume some familiarity with non-Archimedean analysis. We refer the reader to the standard text [Sch06] for definitions and terminology.

The local part uses tools from non-Archimedean analysis to show that p-adic integrality of $\delta_m s$ implies p-adic decay of $c(n)$. Let $|\cdot|_p$ denote the usual p-adic norm on $\mathbb{Q}$ and let $\|\delta_m s\|_p$ denote the supremum of $|\delta_m s(n_0, \ldots, n_m)|_p$ over all sets $\{n_0, \ldots, n_m\}$ of nonnegative integers. We will prove the following new result.

$$\|\delta_m s\|_p = \sup_{n \geq m} |c(n)|_p p^{\tau_{m,p}(n)} \tag{1.3}$$

where $\tau_{m,p}(n)$ is the largest possible p-adic valuation of a product of m distinct positive integers $\leq n$, i.e.

$$\tau_{m,p}(n) := \max_{\substack{S \subset \{1,\ldots,n\}, \\ \#S = m}} w_p\Big\{\prod_{s \in S} s\Big\}$$

where w_p denotes the normalized p-adic valuation. Formula (1.3) shows that p-integrality of $\delta_m s$ is equivalent to sufficiently rapid p-adic decay of finite differences, and so integrality of $\delta_m s$ implies local decay of $c(n)$ at all finite primes.

The global part of the argument uses the product formula to obtain an Archimedean growth condition. We recall that the product formula says that for any nonzero element x of K,

$$\prod_{v \in M_K} |x|_v^{n_v} = 1$$

where n_v is the local degree of the place v. If s is not polynomial, there exists a subsequence $c(n_i)$ such that $c(n_i) \neq 0$ for every $i \geq 0$ (this follows from Newton's interpolation formula). Applying the product formula to such a non-vanishing subsequence $c(n_i)$ allows us to combine local decay at all primes to obtain Archimedean growth for c as a necessary condition for the integrality of $\delta_m s$ and non-polynomiality of s. In order to combine local decay rates over all primes it is necessary to study the asymptotic behavior of $\prod_{p \text{ prime}} p^{\tau_{m,p}(n)}$. Our second new result is that

$$\prod_{p \text{ prime}} p^{\tau_{m,p}(n)} = e^{\left(1+\frac{1}{2}+\cdots+\frac{1}{m}\right)n+O(n\exp\{-\alpha(\log n)^{1/2}\}\log n)} \tag{1.4}$$

for some positive constant α. Combining (1.3) with (1.4) forms the basis for Theorem 1.3.1.

1.3.2 Further Discussion and Background

We proceed to discuss the proofs of (1.3) and (1.4) in greater detail and provide some contextual background. Let $\mathbb{C}_p$ be the metric completion of an algebraic closure of the p-adic field $\mathbb{Q}_p$. The use of finite differences in non-Archimedean analysis goes back to a classical result of Mahler [Mah58]. He proved that $s\colon \mathbb{N} \to \mathbb{C}_p$ is the restriction of a continuous function $f\colon \mathbb{Z}_p \to \mathbb{C}_p$ if and only if the finite differences of s converge to zero p-adically. He also showed that when this is the case, the

supremum of f (or s) is equal to the supremum of the finite differences:

$$\|s\|_p = \|c\|_p . \tag{1.5}$$

It is well-known that there are many other Mahler-type characterizations (cf. e.g., [Sch06], §53). For instance, s is the restriction of a Lipschitz continuous function $f\colon \mathbb{Z}_p \to \mathbb{C}_p$ if and only if $|c(n)|_p p^{[\log_p n]}$ is bounded,[10] and the supremum of $|c(n)|_p p^{[\log_p n]}$ is equal to the optimal p-adic Lipschitz constant of s (*loc. cit.*):

$$\|\delta_1 s\|_p = \sup_{n\geq 1} |c(n)|_p p^{[\log_p n]}. \tag{1.6}$$

Our formula (1.3) is an extension of (1.6) to all higher divided differences. In §3.1 we work out the precise Mahler-type criterion obtained from the integrality of higher divided differences (Proposition 3.1.3): If $\delta_{m+1} s$ is p-integral, then s extends to an element of $C^m(\mathbb{Z}_p, \mathbb{C}_p)$, the Banach space of m-times continuously differentiable functions, and $f^{(m)}$ is Lipschitz continuous with constant $|m!|_p$.

For the proof of (1.3) we make use of the Mahler series formula for $\delta_m s$ due to Schikhof. Before we may use this formula, however, we must resolve a technical difficulty which is the fact that $\delta_m s$ (after a minor change of variables to avoid the diagonal hyperplanes) inhabits the larger of the two Banach spaces

$$C(\mathbb{Z}_p^{m+1}, \mathbb{C}_p) \subset \ell_p^\infty(\mathbb{N}^{m+1}).$$

Schikhof stated this formula as a convergent Mahler series for $\delta_m s$ under the assumption that $\delta_m s$ is continuous, and while it still gives a sensible expression when $\delta_m s$ is only bounded, it is a divergent series for the topology of $\ell_p^\infty(\mathbb{N}^{m+1})$. Computing the norm of this divergent series requires some care. A related difficulty is the fact that there is no canonical orthonormal basis of $\ell_p^\infty(\mathbb{N}^{m+1})$, not even when $m = 0$.

[10] This is typically stated in terms of the boundedness of $|c(n)|_p n$, however the exact formula for the optimal Lipschitz constant justifies the use of the semi-norm $s \mapsto \sup_{n\geq 1} |c(n)|_p p^{[\log_p n]}$ over $s \mapsto \sup_{n\geq 1} |c(n)|_p n$.

Let $(C_{\underline{j}})_{\underline{j}\in\mathbb{N}^{m+1}}$ be a bounded $\mathbb{C}_p$-valued function. The divergent series we will consider are of the form

$$\sum_{\underline{j}\in\mathbb{N}^{m+1}} C_{\underline{j}}\binom{\underline{x}}{\underline{j}} \quad (\underline{x}\in\mathbb{Z}_p^{m+1}). \tag{1.7}$$

We can formally interpret any such series as a definition for the function $F\colon \mathbb{N}^{m+1}\to\mathbb{C}_p$ whose value at $\underline{n}$ is given by $\sum C_{\underline{j}}\binom{\underline{n}}{\underline{j}}$ as this reduces to a finite sum. We will show that in fact any element of $\ell_p^\infty(\mathbb{N}^{m+1})$ can be uniquely expressed as a divergent series of the form (1.7). Furthermore, the mapping

$$\begin{aligned}\ell_p^\infty(\mathbb{N}^{m+1}) &\to \ell_p^\infty(\mathbb{N}^{m+1})\\ F &\mapsto (C_{\underline{j}})_{\underline{j}\in\mathbb{N}^{m+1}}\end{aligned} \tag{1.8}$$

so obtained is an isometry (Theorem 3.0.1). Taken together, these two statements may be regarded as a generalization of the second part of Mahler's classical result (1.5) to p-adically bounded functions. The fact that (1.8) is an isometry means that the norm of bounded functions $F\colon \mathbb{N}^{m+1}\to\mathbb{C}_p$ given by divergent series of the form (1.7) can be calculated as though the multivariate binomial polynomials did form an orthonormal basis of $\ell_p^\infty(\mathbb{N}^{m+1})$ despite the fact that they do not. This work-around lets us circumvent the absence of a canonical orthonormal basis and prove (1.3) using Schikhof's formula.

In §3.3 we return to the global setting by combining the local decay of the finite differences over all primes. For this discussion we will consider a function s valued in $\mathbb{Q}$ rather than K for simplicity. If $\delta_m s$ is integral then $\delta_m s$ is p-integral for all primes p, so using (1.3) shows that

$$\prod_{p\text{ prime}} |c(n)|_p \le \prod_{p\text{ prime}} p^{-\tau_{m,p}(n)} \quad (n\in\mathbb{N}).$$

Note that both products have only a finite number of terms that differ from unity. By using the product formula we obtain an Archimedean growth condition for non-vanishing finite differences $c(n)$:

$$|c(n)| \geq \prod_{p \text{ prime}} p^{\tau_{m,p}(n)} \quad (c(n) \neq 0). \tag{1.9}$$

It is well-known that s is polynomial if and only if c is eventually zero. It follows that s is not polynomial if and only if there exists a subsequence n_i such that $c(n_i)$ is non-vanishing, in which case we may apply (1.9) to see that

$$\limsup_{n\to\infty} |c(n)|^{1/n} \geq \sup_{i<\infty} |c(n_i)|^{1/n_i} \geq \lim_{n\to\infty} \prod_{p \text{ prime}} p^{\frac{\tau_{m,p}(n)}{n}} \tag{1.10}$$

whenever $\delta_m s$ is integer-valued but s is not polynomial.

In §3.2 we calculate the asymptotic behavior of $\prod_p p^{\tau_{m,p}(n)}$ (cf. (1.4), Theorem 3.0.2). We find that

$$\lim_{n\to\infty} \prod_{p \text{ prime}} p^{\frac{\tau_{m,p}(n)}{n}} = e^{1+\frac{1}{2}+\cdots+\frac{1}{m}}. \tag{1.11}$$

The proof uses the Chebyshev function $\vartheta(x) = \sum_{p\leq n} \log x$. It is well-known that $\vartheta(x) = x + o(x)$ but we will require a smaller error term. For this purpose we employ a useful result of Rosser and Schoenfeld [RS62] (cf. (3.13)). In the final section, §3.3, we use the inequality (1.10) to obtain a growth condition on s which together with (1.11) leads to the two conditions of Theorem 1.3.1.

1.3.3 Interpreting the Integrality of Divided Differences

Divided differences and the Hall–Ruzsa theorem are connected because the condition of congruence-preserving is equivalent to the integrality of $\delta_1 s$. This interpretation generalizes to the integrality of higher divided differences as we now explain.

Roughly speaking, a function whose mth divided difference is integral is "locally" approximated to mth order by polynomials. "Locally" is in reference to the topology

on $\mathbb{N}$ inherited from the ring of adeles, $\mathbb{N} \subset \mathbb{A}$. In this topology the neighborhoods are infinite arithmetic progressions and small neighborhoods are infinite arithmetic progressions with highly divisible periods.

Suppose that $s\colon \mathbb{N} \to \mathbb{Q}$ is a function whose mth divided difference is integral. Then there is a positive integer N such that $N\delta_i s$ is integral for every $i = 0, \ldots, m$ (this is not obvious but it follows from Theorem 3.0.1, cf. Remark 1). We again make use of the Newton interpolation formula (1.2). Let $\{x_0, x_1, \ldots\} \subset \mathbb{Q}$ be a denumerable subset and let $s\colon \{x_0, x_1, \ldots\} \to \mathbb{Q}$ be a function. For all $x \in \{x_0, x_1, \ldots\}$ we have that

$$s(x) = s(x_0) + \sum_{k=1}^{\infty} \delta_k s(x_0, x_1, \ldots, x_k) \prod_{j=0}^{k-1} (x - x_j). \tag{1.12}$$

We consider the restriction of s to a small neighborhood $U = n_0 + \mathbb{N}\varepsilon$, where ε is a nonzero integer and n_0 is arbitrary. Let $x_0, \ldots, x_m$ be chosen from U, where we consider $x_0, \ldots, x_{m-1}$ as fixed and $x := x_m$ as variable. From (1.12) we obtain that[11]

$$s(x) = P(x) + O(\varepsilon^m) \quad \text{for all } x \in U \tag{1.13}$$

where $P(x)$ is a polynomial in x of degree $< m$ whose coefficients are rational numbers with denominators dividing N. When $\delta_m s$ contains arbitrarily large denominators in its values, the implied constant in the asymptotic notation cannot generally be chosen independently of x.

We can interpret the integrality of $\delta_1 s$ in light of the above discussion as follows. On any neighborhood $U \subset \mathbb{N}$ of order ε the function s is approximated to first order in ε by the constant polynomial $P(x) = s(x_0)$. Moreover, the implied constant in the asymptotic notation may be chosen to be unity. We see that $s(x) - s(x_0)$ is divisible by ε for any $x, x_0 \in U$.

[11] The asymptotic notation is to be interpreted in the following sense: there exists a positive integer M such that for any x in U, $M(s(x) - P(x))$ is integral and divisible by ε^m. If $\delta_m s$ is integer-valued then M may be chosen to be unity.

In §3.1 we will prove a second interpretation of the integrality of $\delta_m s$: If a function $s\colon \mathbb{N} \to \mathbb{Q}$ has $\mathbb{Z}$-valued mth divided difference, then for every prime p the function s extends to a p-adic m-times continuously differentiable function $f_p\colon \mathbb{Z}_p \to \mathbb{Q}_p$ and $f_p^{(m)}$ is Lipschitz continuous with constant $|m!|_p$.

CHAPTER II

Orbits of Rational Functions

This chapter studies pairs of rational functions which possess forward orbits with infinite intersection. This hypothesis suggests that we may use tools from dynamical systems related to orbits, such as equilibrium measures, or recent results on equidistribution in orbits — however this is not the case. The real object of study in this chapter is the monodromy of rational functions. Under what circumstances can two rational functions have "compatible" ramification? This is a subtle question, and its ramifications will lead us to the well-known special classes of dynamically affine rational functions: power maps, Chebyshev polynomials, and Lattès rational functions. The results in this chapter are joint work with Michael Zieve [OZ].

2.1 Algebraic Curves

In this chapter we work over an algebraically closed field K of characteristic zero, although no essential generality will be lost if the reader takes K to be the field of complex numbers. Recall that a variety is an open and irreducible subset of a projective algebraic set (always over K; see §1.1.1). We define a curve to be a variety of dimension one. Let $K(t)$ be the field of rational functions in one variable. We fix, once and for all, an algebraic closure $K(t)^a$ of $K(t)$. For fundamental notions about varieties and curves, such as morphisms of varieties, ramification degree, degree, and

(geometric) genus, we refer to the standard text [Har77]. Morphisms are always over the field K.

We will frequently switch between the geometric and algebraic pictures according to suitability and convenience. On the level of curves, this equivalence takes the following form.

Theorem 2.1.1 ([Har77], (I.6.12))**.** *The following two categories are equivalent:*

1. *field extensions of K of transcendence degree* 1*, and K-homomorphisms;*
2. *nonsingular projective curves, and dominant*[1] *morphisms.*

Recall that the function field K_C of a smooth curve C is the set of morphisms $C \to \mathbb{P}^1$ which are not identically equal to ∞. This set naturally forms a field extension of K of transcendence degree 1. The functor from the second category to the first category is given by mapping a nonsingular projective curve C to its function field K_C, and a dominant morphism $f\colon D \to C$ of nonsingular projective curves D and C is mapped to the K-homomorphism $K_C \to K_D\colon \varphi \mapsto \varphi \circ f$. Defining the functor in the other direction takes a bit more work (see [Har77, §I.6.12]).

It often happens that a function field L of dimension 1 over K is presented as a finite extension of $K(t)$. The field of rational functions $K(t)$ is the function field of the projective line $\mathbb{P}^1$. In this case the function field L gives us a curve C_L and a nonconstant morphism $C_L \to \mathbb{P}^1$.

It is worth keeping in mind that Theorem 2.1.1 is only an equivalence of categories and not an isomorphism of categories. If $f\colon D \to C$ is a nonconstant morphism of curves, then it corresponds by Theorem 2.1.1 to two function fields K_D and K_C with a K-homomorphism $K_C \to K_D$. Passing back through the equivalence produces nonsingular projective curves C' and D', a morphism $f'\colon C' \to D'$, and isomorphisms

[1] Note that for curves, a morphism $f\colon D \to C$ is dominant if and only if f is nonconstant.

$C \xrightarrow{\sim} C'$ and $D \xrightarrow{\sim} D'$ such that the diagram,

$$\begin{array}{ccc} C & \xrightarrow{\sim} & C' \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f'} \\ D & \xrightarrow{\sim} & D', \end{array}$$

commutes.[2] This ambiguity manifests in both the geometric and field-theoretic pictures. Geometrically, we may use the bottom isomorphism to identify D and D', however this does not identify the morphisms f and f'. We obtain an isomorphism $\rho\colon C \to C$ such that $f = f' \circ \rho$. When we want to emphasize this ambiguity, we will say that the curve D_L and morphism $f\colon D_L \to C$ associated to a finite extension L/K_C is defined up to an isomorphism over C.

Field-theoretically, if we use the bottom isomorphism to identify $K_{D'}$ with K_D, and ${K_D}^a$ is a given algebraic closure of K_D, then the subfields of ${K_D}^a$ corresponding to the morphisms f and f' are not necessarily the same, they are only conjugate subfields in general. Note however that if the field extension corresponding to f is *Galois* then there is a unique subfield in any algebraic closure corresponding to f.

This brings us to our next definition.

Definition 2.1.2. A nonconstant morphism of curves $\pi\colon D \to C$ is (generically) Galois if K_D/K_C is a Galois extension of fields.

The usual notion of a Galois morphism from the theory of schemes requires that the morphism π is étale (cf. [Mil80, §1.5]), however this is too restrictive for our purposes since we will generally want to allow for ramification. As we have no need for the stronger notion, we will simply say that π is Galois.

[2] Strictly speaking, these isomorphisms are not part of the statement of Theorem 2.1.1, which only asserts that natural isomorphisms of these functors exists, but there is an "obvious" choice for the isomorphisms $C \xrightarrow{\sim} C'$, namely the bijection between a curve C and the set of valuations corresponding to its points.

Galois morphisms arise naturally when taking quotients by a finite group of automorphisms.

Definition 2.1.3. Let C be a curve and let G be a finite subgroup of $\operatorname{Aut} C$. The quotient of C by G is the nonsingular projective curve C/G associated (by Theorem 2.1.1) to the field $(K_C)^G$.

The quotient comes equipped with a natural morphism $\pi\colon C \to C/G$ coming from the inclusion $(K_C)^G \subset K_C$. By Artin's criterion, [Lan02, VI.1.8], π is Galois with group G.

Definition 2.1.4. Let $f\colon D \to C$ be a nonconstant morphism of curves. The Galois closure (N, p) of f is the data of the nonsingular projective curve N associated to the Galois closure K_N of K_D/K_C together with the morphism $p\colon N \to D$ associated to the field inclusion $K_D \subset K_N$.

The Galois closure of a morphism $f\colon D \to C$ is only defined up to a isomorphism over C.

The geometry of a variety and its projective embeddings is often studied by means of its divisors. On the level of curves, the theory of divisors takes an especially simple form.

Definition 2.1.5. A divisor D on a curve C is a finite subset of C with $\mathbb{Z}$-multiplicities. We write a divisor as a sum,

$$D = \sum_{P \in C} n_P [P],$$

where $n_P \in \mathbb{Z}$ and all but finitely many n_P are equal to zero. The degree of D is defined by

$$|D| = \sum_{P \in C} n_P.$$

A nonconstant morphism of curves comes equipped with a few natural divisors. To give their definitions we will need some notation. For a nonconstant morphism $f\colon X \to Y$ of curves and any point $P \in X$, let $e_f(P)$ denote the ramification index of f at P. For any point $Q \in Y$ we define the fiber of f over Q to be the subset $\mathcal{F}_f(Q) = \{P \in X : f(P) = Q\}$ (we occasionally also write $f^{-1}(Q)$ for $\mathcal{F}_f(Q)$). We define $r_f(Q) = \#\mathcal{F}_f(Q)$. The lcm-ramification index of f at Q is defined to be

$$\varepsilon_f(Q) := \operatorname{lcm}\{e_f(P) : P \in \mathcal{F}_f(Q)\}.$$

The ramification divisor of f is

$$\mathcal{R}_f := \sum_{P \in X} (e_f(P) - 1)[P].$$

The branching divisor of f is

$$\mathcal{B}_f := \sum_{Q \in Y} (\deg f - r_f(Q))[Q].$$

In general the ramification locus of a dominant morphism of varieties $f\colon X \to Y$ is a (pure) codimension one closed subvariety of X. For curves this implies that the ramification locus is a finite set of points (the points where $e_f(P) > 1$). Hence $\mathcal{R}_f$ is a divisor. The next formula implies that $f(\mathcal{R}_f) = \mathcal{B}_f$, and so $\mathcal{B}_f$ is also a divisor (on Y). For a proof of the proposition see [Har77, II.6.9].

Proposition 2.1.6. *Let $f\colon D \to C$ be a nonconstant morphism of nonsingular projective curves. Let $\mathcal{F}_f(Q)$ be the fiber of f over a point Q of C. Then*

$$\sum_{P \in \mathcal{F}_f(Q)} e_f(P) = \deg f.$$

Let us recall the well-known Riemann–Hurwitz formula, for which we will have much use.

Theorem 2.1.7 (Riemann–Hurwitz; [Har77], IV.2.4)**.** *Let $f\colon D \to C$ be a nonconstant morphism of nonsingular projective curves. Then*

$$2\mathfrak{g}_D - 2 = (\deg f)(2\mathfrak{g}_C - 2) + |\mathcal{R}_f|.$$

The following useful result computes the ramification of points under the Galois closure morphism.

Lemma 2.1.8. *Let $f\colon D \to C$ be a nonconstant morphism of curves and let (N, p) be the Galois closure of f. Let Q be any point of C and let $\mathcal{F}_f(Q)$ be the fiber of f over Q. Then the ramification index $e_{f\circ p}(P)$ of the composite map $f \circ p$ at any point $P \in N$ such that $(f \circ p)(P) = Q$ is equal to the least common multiple ε_Q of $\{e_f(R) : R \in \mathcal{F}_f(Q)\}$.*

Proof. We will utilize the equivalence between the categories of function fields of dimension 1 over K and dominant morphisms of curves over K (Theorem 2.1.1) and give a field-theoretic proof using Galois theory.

Let G be the Galois group of K_N/K_C, and let H be the subgroup corresponding to K_N/K_D. We make use of the double-coset description of primes in intermediate (not-necessarily-Galois) extensions.[3] Let $D_P = \{s \in G : s(P) = P\}$ be the decomposition subgroup at P. The primes (points) of D lying over Q are in bijection with the double cosets in $H\backslash G/D_P$. Since K is algebraically closed and characteristic zero, D_P is cyclic and generated by a single element γ of order $e_{f\circ p}(P)$. There is a natural bijection $H\backslash G/D_P \cong (H\backslash G)/D_P$, and the action of γ on $H\backslash G$ decomposes this set into cycles whose lengths are precisely the ramification indices of points in M lying over Q. Therefore, if γ has order greater than the lcm ε_Q of these cycle lengths, then D_P cannot act faithfully on the set H/G (i.e., some nontrivial power of γ acts by the

[3] See [vdW35] or, for a proof in the number field setting, [Neu99, §I.9].

identity). However, since K_N is the Galois closure of K_D/K_C, the subgroup H cannot contain any nontrivial subgroups which are normal in G; in particular, $\cap_g {}^g H = 1$. This subgroup is also the kernel of the (right-) action of G on $H\backslash G$, which implies that G acts faithfully, so D_P must also. We see that γ has order which is $\leq \varepsilon_Q$. However, by multiplicativity of ramification, we also have that $\varepsilon_Q \leq e_{f\circ p}(P)$. This proves that $\varepsilon_Q = e_{f\circ p}(P)$. □

The next formula computes the genus of the Galois closure of a nonconstant rational function f considered as a morphism $f\colon \mathbb{P}^1 \to \mathbb{P}^1$.

Lemma 2.1.9. *Let $f\colon \mathbb{P}^1 \to \mathbb{P}^1$ be a nonconstant rational function, and let (N, p) be its Galois closure. Then*

$$\mathfrak{g}_N = 1 + (\deg f)(\deg p)\Big(-1 + \tfrac{1}{2}\sum_{Q\in\mathbb{P}^1}\big(1 - \tfrac{1}{\varepsilon_f(Q)}\big)\Big).$$

Proof. Let us apply the Riemann–Hurwitz formula to the composite map $f\circ p\colon N \to \mathbb{P}^1$. For any point $P \in N$, the ramification index $e_{fp}(P)$ of the composite map $f\circ p$ is equal to the lcm-ramification index $\varepsilon_f(Q)$ of f over $Q = (f\circ p)(P)$ (Lemma 2.1.8). We obtain that

$$\begin{aligned}
2\mathfrak{g}_N - 2 &= -2\deg(f\circ p) + |\mathcal{R}_{f\circ p}| \\
&= -2(\deg f)(\deg p) + \sum_{P\in N}\big(e_{fp}(P) - 1\big) \\
&= -2(\deg f)(\deg p) + \sum_{Q\in\mathbb{P}^1}\sum_{P\in\mathcal{F}_{fp}(Q)}\big(e_{fp}(P) - 1\big) \\
&= -2(\deg f)(\deg p) + \sum_{Q\in\mathbb{P}^1}\big((\deg f)(\deg p) - r_{fp}(Q)\big).
\end{aligned}$$

Since $f\circ p$ is Galois, we have that

$$\deg(f\circ p) = \sum_{P\in\mathcal{F}_{fp}(Q)} e_{fp}(P) = \sum_{P\in\mathcal{F}_{fp}(Q)} \varepsilon_f(Q) = \varepsilon_f(Q) r_{fp}(Q).$$

It follows that

$$2\mathfrak{g}_N - 2 = (\deg f)(\deg p)\Big(-2 + \sum_{Q\in\mathbb{P}^1}\big(1 - \tfrac{1}{\varepsilon_f(Q)}\big)\Big).$$

By rearranging terms we obtain the claimed formula,

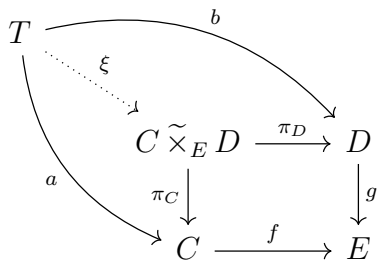
$$\mathfrak{g}_N = 1 + (\deg f)(\deg p)\Big(-1 + \tfrac{1}{2}\sum_{Q\in\mathbb{P}^1}\big(1 - \tfrac{1}{\varepsilon_f(Q)}\big)\Big). \qquad \square$$

The next definition introduces a fundamental construction which will play an essential role in our results.

Definition-Proposition 2.1.10. *Let $f\colon C \to E$ and $g\colon D \to E$ be nonconstant morphisms of curves. The tensor product of fields*

$$K_C \otimes_{K_E} K_D \simeq K_1 \times \cdots \times K_t$$

is isomorphic to a direct product of function fields of dimension 1 over K. We define the smooth fiber product *$C \widetilde{\times}_E D$ to be the finite disjoint union of smooth projective curves associated to the fields $K_1, \ldots, K_t$. The smooth fiber product possesses canonical maps $\pi_C\colon C \widetilde{\times}_E D \to C$ and $\pi_D\colon C \widetilde{\times}_E D \to D$ whose restriction to every connected component is nonconstant. The smooth fiber product $C \widetilde{\times}_E D$ and the maps π_C, π_D satisfy the following universal property: for any smooth projective curve T and nonconstant morphisms $a\colon T \to C$ and $b\colon T \to D$ such that $f \circ a = g \circ b$, there exists a unique morphism $\xi\colon T \to C \widetilde{\times}_E D$ such that the following diagram*



commutes.

Proof. The extension K_D/K_E is finite and separable, so there is an element $\alpha \in K_D$ such that $K_D = K_E(\alpha)$. Let $f(t)$ be the minimal polynomial of α over K_E. Then $K_D \cong \frac{K_E[t]}{(f(t))}$, so we have the isomorphism

$$K_C \otimes_{K_E} K_D \cong \frac{K_C[t]}{(f(t))}.$$

As $f(t)$ is separable, this shows that the tensor product algebra $K_C \otimes_{K_E} K_D$ is a separable K_E-algebra, hence isomorphic to a product of fields $K_1, \ldots, K_t$ which are finite extensions of K_C. Each map $K_C \to K_C \otimes_{K_E} K_D \to K_i$, $1 \leq i \leq t$, corresponds (by Theorem 2.1.1) to a nonconstant morphism $C_i \to C$ where C_i is the curve corresponding to the function field K_i. By putting these maps together, we obtain a canonical map $\pi_C \colon C \widetilde{\times}_E D \to C$ which is nonconstant on every connected component C_i of $C \widetilde{\times}_E D$. Similarly, there is such a map $\pi_D \colon C \widetilde{\times}_E D \to D$.

The universal property for $(C \widetilde{\times}_E D, \pi_C, \pi_D)$ follows immediately from combining the facts that the tensor product over K_E is the coproduct in the category of commutative K_E-algebras, and the functors which realize an equivalence of categories are full and faithful (i.e., they induce bijections on every set of morphisms between two objects). □

Remark. The smooth fiber product is not a fiber product in the category of schemes.

We will have use for the following form of Abhyankar's lemma.

Lemma 2.1.11. *Let $f \colon B \to D$ and $g \colon C \to D$ be nonconstant morphisms of smooth projective curves. Let A be a component of the smooth fiber product $B \widetilde{\times}_D C$, and let $\pi \colon A \to B$ and $\psi \colon A \to C$ be the morphisms associated to the K-homomorphisms $K_B \to B \widetilde{\times}_D C \to K_A$ and $K_C \to B \widetilde{\times}_D C \to K_A$. Let P be a point of A, and set $Q = \pi(P)$ and $R = \psi(P)$. Then $e_{f\circ\pi}(P) = \operatorname{lcm}(e_f(Q), e_g(R))$.*

For a proof see Theorem 3.9.1 of [Sti09].

When the smooth fiber product is irreducible ($\Leftrightarrow$ connected), applying the Riemann–Hurwitz formula results in a useful relation between the genus of the smooth fiber product and the ramification data of f and g.

Lemma 2.1.12. *Let $f\colon C \to E$ and $g\colon D \to E$ be nonconstant morphisms of curves. Suppose that the smooth fiber product $C \mathbin{\widetilde{\times}}_E D$ is irreducible. Then*

$$2\mathfrak{g}_{C \mathbin{\widetilde{\times}}_E D} - 2 = (2\mathfrak{g}_D - 2)\deg f + \sum_{(P,Q)\in F} \big(e_f(P) - \gcd(e_f(P), e_g(Q))\big),$$

where $F = \{(P,Q) \in C \times D : f(P) = g(Q)\}$.

Proof. Let $(P,Q) \in F$ and write $T = C \mathbin{\widetilde{\times}}_E D$. We claim that there are exactly $\gcd(e_f(P), e_g(Q))$ points R of T such that $\pi_C(R) = P$ and $\pi_D(R) = Q$. Let $K_{C,P}$ denote the completion of K_C for the valuation at a point P of C. Then $K_{C,P} \otimes_{K_E} K_{D,Q}$ is isomorphic to the product of the completions $K_{T,R}$ over all points $R \in T$ such that $\pi_C(R) = P$ and $\pi_D(R) = Q$. Let R be such a point of T, and set $a = e_f(P)$, $b = e_g(Q)$. By Abhyankar's lemma (Lemma 2.1.11), the ramification of R through the composite map $f \circ \pi_C$ is equal to $\operatorname{lcm}(a,b)$. The following diagram indicates the ramification degrees:

$$\begin{array}{ccc} R & \xrightarrow{\frac{\operatorname{lcm}(a,b)}{b}} & Q \\ {\scriptstyle \frac{\operatorname{lcm}(a,b)}{a}}\downarrow & & \downarrow{\scriptstyle b} \\ P & \xrightarrow{a} & f(P). \end{array}$$

The dimension of $K_{C,P} \otimes_{K_E} K_{D,Q}$ over $K_{C,P}$ is equal to the dimension of $K_{D,Q}$ over $K_{E,f(P)}$, so the sum over $\operatorname{lcm}(a,b)a^{-1}$ over all points R such that $\pi_C(R) = P$ and $\pi_D(R) = Q$ is equal to the degree of $K_{D,Q}/K_{E,f(P)}$, and this is b since the extension $K_{D,Q}/K_{E,f(P)}$ is totally ramified. Therefore the number of such points R is equal to $b(\operatorname{lcm}(a,b)a^{-1})^{-1} = \gcd(a,b)$, as claimed.

Now we apply the Riemann–Hurwitz formula to the map $\pi_D\colon T \to D$ to obtain

$$2\mathfrak{g}_T - 2 = (2\mathfrak{g}_D - 2)\deg f + \sum_{R\in T}(e_{\pi_D}(R) - 1).$$

We can break up the sum over F to obtain

$$\begin{aligned}
2\mathfrak{g}_T - 2 &= (2\mathfrak{g}_D - 2)\deg f + \sum_{(P,Q)\in F}\sum_{\substack{\pi_C(R)=P,\\ \pi_D(R)=Q}}(e_{\pi_D}(R) - 1) \\
&= (2\mathfrak{g}_D - 2)\deg f + \sum_{(P,Q)\in F}\gcd(e_f(P), e_g(Q))\Big(\tfrac{\operatorname{lcm}(e_f(P),e_g(Q))}{e_g(Q)} - 1\Big) \\
&= (2\mathfrak{g}_D - 2)\deg f + \sum_{(P,Q)\in F}\big(e_f(P) - \gcd(e_f(P), e_g(Q))\big). \qquad \square
\end{aligned}$$

2.2 Dynamically Affine Rational Functions

In this section we study a special class of rational functions arising naturally in the theory of algebraic dynamics and number theory. Let G be a commutative group variety. An affine morphism of G is a functional composition of a nontrivial endomorphism of G as a group variety with a translation.

Recall that the degree of a rational function f is the maximum of the degrees of the numerator and denominator of f in reduced form. The degree is multiplicative under composition and agrees with the usual geometric notion of degree when f is considered as a morphism of curves $f\colon \mathbb{P}^1 \to \mathbb{P}^1$.

Definition 2.2.1. A rational function f of degree > 1 is dynamically affine if there exist a one-dimensional group variety G, an affine morphism $A\colon G \to G$, and a nonconstant morphism $\pi\colon G \to \mathbb{P}^1$ such that $f\pi = \pi A$.

2.2.1 Power Maps, Chebyshev, and Lattès

In this subsection we introduce the standard families of dynamically affine rational functions. For simplicity, we will only consider affine morphisms with trivial translation component. The only group varieties of dimension 1 over an algebraically closed

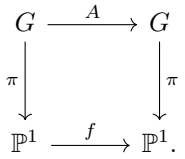


Figure 2.1: A dynamically affine rational function.

field are the additive group $\mathbb{G}_a$, the multiplicative group $\mathbb{G}_m$, and elliptic curves E. The only affine maps of the additive group have degree 1, so they do not give rise to dynamically affine maps. However the multiplicative group and elliptic curves give rise to interesting families of dynamically affine rational functions.

2.2.1.1 The multiplicative group

The endomorphisms of the multiplicative group are of the form x^n for $n \in \mathbb{Z}$. By taking π to be the inclusion map $\mathbb{G}_m \subset \mathbb{P}^1$, we obtain our first family of dynamically affine maps.

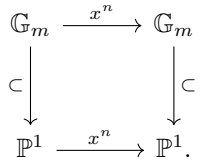


Figure 2.2: Power maps.

The map x^n is totally ramified at $x = 0$ and ∞ with ramification index $|n|$. There are no other critical points or critical values. The ramification divisor of x^n is given by

$$\mathcal{R} = (|n| - 1)[0] + (|n| - 1)[\infty].$$

A more interesting class of examples arises by taking the quotient map π to be

$$\pi \colon \mathbb{G}_m \to \mathbb{P}^1$$

$$x \mapsto x + x^{-1}.$$

Let n be a positive integer.

Definition-Proposition 2.2.1. *The nth Chebyshev polynomial is the unique polynomial $T_n(x) \in \mathbb{Z}[x]$ satisfying*

$$T_n(x + x^{-1}) = x^n + x^{-n}. \tag{2.1}$$

For all positive integers m and n the Chebyshev polynomials satisfy

- $T_n(-x) = (-1)^n T_n(x)$,
- $T_m(T_n(x)) = T_{mn}(x)$.

$$\begin{array}{ccc} \mathbb{G}_m & \xrightarrow{x^n} & \mathbb{G}_m \\ {\scriptstyle x+x^{-1}}\downarrow & & \downarrow{\scriptstyle x+x^{-1}} \\ \mathbb{P}^1 & \xrightarrow{T_n} & \mathbb{P}^1. \end{array}$$

Figure 2.3: Chebyshev polynomials.

For the proof see [Sil07, §6.2].

$$\begin{aligned} T_1 &= x \\ T_2 &= x^2 - 2 \\ T_3 &= x^3 - 3x \\ T_4 &= x^4 - 4x^2 + 2 \\ T_5 &= x^5 - 5x^3 + 5x \\ T_6 &= x^6 - 6x^4 + 9x^2 - 2 \\ T_7 &= x^7 - 7x^5 + 14x^3 - 7x \\ T_8 &= x^8 - 8x^6 + 20x^4 - 16x^2 + 2 \end{aligned}$$

Figure 2.4: The first eight Chebyshev polynomials $T_n(x)$.

From the defining relation (2.1) it is easy to find the critical points and critical values of $T_n(x)$. As $T_n(x)$ is a polynomial of degree n, it is totally ramified at ∞ with ramification index n. The finite critical points of $T_n(x)$ are given by the points of the form

$$z + z^{-1}, \quad z \in \mu_{2n} \backslash \{\pm 1\}.$$

The finite critical values are given by

$$\begin{cases} \varnothing & \text{if } n = 1, \\ \{-2\} & \text{if } n = 2, \\ \{\pm 2\} & \text{if } n \geq 3. \end{cases}$$

The ramification divisor of $T_n(x)$ is given by

$$\mathcal{R} = (n-1)[\infty] + \sum_{k=1}^{n-1} [z^k + z^{-k}]$$

where z is any primitive $2n$-th root of unity in K.

2.2.1.2 Elliptic curves

Let E be an elliptic curve over K. It follows from the Riemann–Roch theorem that E admits a projective model in $\mathbb{P}^2$ whose affine part is given by the Weierstrass equation,

$$y^2 = x^3 + ax + b, \quad a, b \in K,$$

where $\Delta = 4a^3 + 27b^2 \neq 0$. Then E is the closure of this affine curve and possesses a unique point $\mathcal{O}$ at infinity; see [Har77, IV.4] or [Sil09]. The functions x and y, considered as coordinates on the affine part of E, are called Weierstrass coordinates.

Let $L\colon E \to E$ be a morphism. We may express L in Weierstrass coordinates,

$$L(x, y) = (F(x, y), G(x, y)),$$

where $F(x, y)$ and $G(x, y)$ are rational functions in the function field of E and $G(x, y)^2 = F(x, y)^3 + aF(x, y) + b$ for all $(x, y) \in E$.

Let us express the group operation on E additively. It is well-known that the group-theoretic inverse of a point (x, y) is given by $(x, -y)$ (this follows, for instance, from the geometric description of the group law on E). Suppose that $L\colon E \to E$ is

an endomorphism of E as a group variety. The fact that $L(-P) = -L(P)$ gives the relation

$$L(x,-y) = (F(x,-y), G(x,-y)) = (F(x,y), -G(x,y)) = -L(x,y).$$

The Weierstrass equation shows that the function field K_E is a quadratic extension of $K(x)$, so we may write $F(x,y) = F_1(x) + yF_2(x)$ and $G(x,y) = G_1(x) + yG_2(x)$. The above relation now shows that $F_2 = G_1 = 0$, hence that any endomorphism of E takes a special form:

$$L(x,y) = (F(x), yG(x)).$$

Consider the x-coordinate map:

$$x\colon E \to \mathbb{P}^1$$

$$[X:Y:Z] \mapsto \begin{cases} XZ^{-1} & \text{if } Z \neq 0, \\ \infty & \text{if } Z = 0. \end{cases}$$

By using x as our quotient map $\pi\colon E \to \mathbb{P}^1$ we can define a large number of interesting dynamically affine rational functions.

Definition 2.2.2. Let E be an elliptic curve in Weierstrass form, and let $L\colon E \to E$ be an endomorphism of E of degree > 1. The Lattès rational function $\ell(x)$ associated to L is the x-coordinate function of $L(x,y) = (\ell(x), y\tau(x))$.

$$\begin{array}{ccc} E & \xrightarrow{L} & E \\ {\scriptstyle x}\downarrow & & \downarrow{\scriptstyle x} \\ \mathbb{P}^1 & \xrightarrow{\ell} & \mathbb{P}^1. \end{array}$$

Figure 2.5: "flexible" Lattès maps.

Lattès maps which are defined using the x-coordinate projection are sometimes called "flexible" Lattès maps. Indeed the x-coordinate is not the only way to obtain Lattès maps from elliptic curve endomorphisms. The x-coordinate realizes the

"quotient-by-(-1)" map $E \to E/\{\pm 1\} = \mathbb{P}^1$. However, there are two isomorphism classes of elliptic curves with other automorphisms besides $P \mapsto -P$. For these other automorphisms one may use (in Weierstrass coordinates) quotient maps given by y, y^2, or x^2. These produce other examples of Lattès maps, which are sometimes called "rigid" Lattès maps, since the isomorphism class of the associated elliptic curve cannot be deformed without losing the extra automorphisms.

In general, a Lattès map is defined to be any rational function $\ell\colon \mathbb{P}^1 \to \mathbb{P}^1$ which fits into a commuting diagram of the form

$$\begin{array}{ccc} E & \xrightarrow{A} & E \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\pi} \\ \mathbb{P}^1 & \xrightarrow{\ell} & \mathbb{P}^1 \end{array}$$

for some elliptic curve E, affine morphism $A\colon E \to E$, and nonconstant morphism $\pi\colon E \to \mathbb{P}^1$.

Although we have ignored affine morphisms with nontrivial translation components in the case of the multiplicative group, it turns out that all dynamically affine rational functions are essentially accounted for by the standard families above. Recall that two rational functions f and g are said to be conjugate if there exists an automorphism μ of $\mathbb{P}^1$ such that $f = \mu \circ g \circ \mu^{-1}$.

Theorem 2.2.3. *A rational function of degree > 1 is dynamically affine if and only if it is conjugate to one of the following dynamically affine rational functions:*

1. *x^n, with $n \in \mathbb{Z}\backslash\{\pm 1\}$;*
2. *a signed Chebyshev polynomial $\pm T_n(x)$ with $n \geq 2$, or*
3. *a Lattès map $\ell_{E,A,\pi}(x)$ for some elliptic curve E, affine morphism $A\colon E \to E$, and nonconstant morphism $\pi\colon E \to \mathbb{P}^1$.*

We will not prove this theorem, which is standard. Instead we refer the reader

to the standard reference [Sil07]. We remark here that the third item may be optimized by giving quotient maps π that are "reduced"; in Weierstrass coordinates this constrains π to be among a finite set (see Lemma 2.2.5).

2.2.2 Ramification of Dynamically Affine Maps

In the next section we will prove a new characterization of dynamically affine maps in terms of Galois-theoretic conditions (Theorem 2.3.2). In this subsection, we study the ramification of dynamically affine maps and prove the results that we will need for the Galois-theoretic characterization.

The following notion is useful.

Definition 2.2.1. A Galois morphism $\pi\colon D \to C$ of curves is reduced if it has a totally ramified critical value.

We first show that reduced morphisms between curves of low genus are essentially determined by their ramification. Recall that the branching divisor $\mathcal{B}_\pi$ of a nonconstant morphism $\pi\colon D \to C$ is the divisor on C given by $\mathcal{B}_\pi = \sum_{Q\in C}(\deg \pi - r_Q)$ where r_Q is the size of the fiber of π over Q (counted without multiplicity).

Lemma 2.2.2. *Let $\pi\colon D \to C$ and $\varpi\colon D' \to C$ be reduced Galois morphisms. If $\mathfrak{g}_D \leq 1$ and $\mathcal{B}_\pi = \mathcal{B}_\varpi$ then there is an isomorphism $\rho\colon D \xrightarrow{\sim} D'$ such that $\pi = \varpi\rho$.*

Proof. The assumption that $\mathcal{B}_\pi = \mathcal{B}_\varpi$ implies that the same points of C realize the maximal coefficients of these divisors. As the totally ramified points are the points with maximal coefficients, this implies π and ϖ have the same totally ramified branch points. Comparing coefficients at any such point shows that π and ϖ have the same degree. As π and ϖ are Galois and have the same degree, the equality of the branching divisors shows that for any point Q in C, the lcm-ramification indices

$\varepsilon_\pi(Q)$ and $\varepsilon_\varpi(Q)$ are equal. This implies that the normal divisors of π and ϖ are equal, so from Lemma 2.1.9 we see that $\mathfrak{g}_D = \mathfrak{g}_{D'}$.

Consider the smooth fiber product $D \widetilde{\times}_C D'$ (Definition 2.1.10). By Abhyankar's lemma (2.1.11), both projection maps $D \widetilde{\times}_C D' \to D$ and $D \widetilde{\times}_C D' \to D'$ are unramified. Let D'' be an irreducible component of $D \widetilde{\times}_C D'$, say of genus $\mathfrak{g}$. The Riemann–Hurwitz formula applied to the unramified morphism $D'' \to D$ (of degree d, say) is

$$2\mathfrak{g} - 2 = d(2\mathfrak{g}_D - 2). \tag{2.2}$$

If $\mathfrak{g}_D = 0$ then (2.2) implies that $\mathfrak{g} = 0$ and $d = 1$. The same consideration applies to $D'' \to D'$ so the projection maps

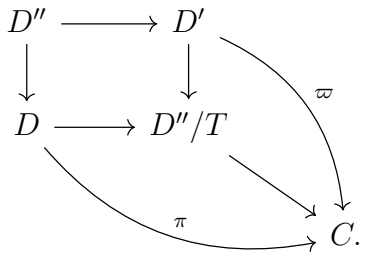
$$D \longleftarrow D'' \longrightarrow D'$$

are isomorphisms. Composing these isomorphisms produces the required isomorphism ρ.

If $\mathfrak{g}_D = 1$ then (2.2) implies that $\mathfrak{g} = 1$. Let $\mathcal{O}$ and $\mathcal{O}'$ be totally ramified points of π and ϖ, respectively, and let $\mathcal{O}''$ be a point of D'' which maps to both $\mathcal{O}$ and $\mathcal{O}'$. We consider D as an elliptic curve E with basepoint $\mathcal{O}$, and likewise for E' and E''. Let $\mathrm{Aut}(D''/C)$ be the set of automorphisms of D'' over C. Let T be the subgroup of $\mathrm{Aut}(D''/C)$ generated by $\mathrm{Aut}(D''/D)$ and $\mathrm{Aut}(D''/D')$, and let D''/T denote the quotient of D'' by T.

The projection map $D'' \to D$ defines an isogeny $E'' \to E$ by construction, so $\mathrm{Aut}(D''/D)$ is a subgroup of E'' (namely the kernel of this isogeny; see [Sil09, III.4.10]). This isogeny induces an isomorphism between $E''/\mathrm{Aut}(D''/D)$ and E, showing that it is a Galois morphism of curves. The same consideration applies verbatim to the projection map $D'' \to D'$, and it follows that the quotient map

$D'' \to D''/T$ factors through D and D'. We obtain the following diagram:



The quotient map $D'' \to D''/T$ is unramified (every point has $|T|$ preimages), so by the Riemann–Hurwitz formula the genus of D''/T is equal to 1. However, the fiber of π over $\pi(\mathcal{O})$ contains only $\mathcal{O}$ which shows that $D \to D''/T$ is an isomorphism. The same consideration applies to $D' \to D''/T$ by looking at $\mathcal{O}'$, and composing the isomorphisms $D \to D''/T \leftarrow D'$ produces the required isomorphism ρ.

□

The next lemma gives an explicit description of the ramification of rational functions whose Galois closure has genus zero. For expressing ramification data, it is convenient to work with multisets. If $f\colon D \to C$ is a nonconstant morphism of smooth projective curves, and Q is a point of C, we will write $E_f(Q)$ for the multiset of ramification indices $e_f(P)$ for all points P in the fiber of f over Q. We will denote multiplicities with exponents, e.g., the multiset $\{1,1,1,2,2\}$ is denoted with $[1^3, 2^2]$. We will write E_f for the (disjoint) union of multisets $E_f(Q)$ over all points $Q \in C$.

Lemma 2.2.3. *Suppose that f is a rational function of degree $n > 60$ with a genus zero Galois closure. Then there are automorphisms μ and τ of $\mathbb{P}^1$ such that $(\mu f \tau, E_f)$ is equal to one of the following:*

1. *x^n, $[n]$ over each of two points,*
2. *$T_n(x)$, $[n]$ over one point, $[1^2, 2^{n-1}]$ over a set of two other points,*
3. *n is even, $x^{\frac{n}{2}} + x^{-\frac{n}{2}}$, $[\frac{n}{2}, \frac{n}{2}]$ over one point, $[2^{n/2}]$ over each of two other points.*

Proof. Let (N,p) be the Galois closure of f. By applying the Riemann–Hurwitz formula to the composite morphism $\pi := f \circ p$ and using Lemma 2.1.6, we obtain that

$$-2 = \deg\pi(-2) + \sum_{P\in\mathbb{P}^1}(e_\pi(P)-1) = \deg\pi(-2) + \sum_{Q\in\mathbb{P}^1}(\deg\pi - r_\pi(P)) \tag{2.3}$$
$$= \deg\pi\big(-2 + \sum_{Q\in\mathbb{P}^1}\big(1-\tfrac{1}{\varepsilon_\pi(Q)}\big)\big).$$

This shows that

$$\sum_{Q\in\mathbb{P}^1}\big(1-\tfrac{1}{\varepsilon_\pi(Q)}\big) < 2. \tag{2.4}$$

If $\varepsilon_\pi(Q) > 1$ then $1-\frac{1}{\varepsilon_\pi(Q)} \geq \frac{1}{2}$, so (2.4) shows that π has at most three critical values.

When there are two critical values we may choose an automorphism μ so that μf has critical values at 0 and ∞. Since $\varepsilon_\pi(Q) \leq \deg\pi$ we have that

$$-2 = \deg\pi\big(-2 + (1-\tfrac{1}{\varepsilon_\pi(0)}) + (1-\tfrac{1}{\varepsilon_\pi(\infty)})\big) \leq \deg\pi\big(-\tfrac{2}{\deg\pi}\big) = -2.$$

It follows that we must have equality, and so $\varepsilon_\pi(0) = \varepsilon_\pi(\infty) = \deg\pi$. This implies that these points are totally ramified for μf. Now we choose an automorphism τ sending 0 and ∞ to the unique μf-preimages of 0 and ∞, respectively, so that $\mu f\tau$ has totally ramified fixed points at 0 and ∞. It follows that $\mu f\tau = cx^n$ for some constant $c \in K^\times$, and we may rescale μ to $c^{-1}\mu$ to obtain $\mu f\tau = x^n$.

Now suppose that there are three critical values of f. Let $K(x)$ be the function field of the Galois closure (N,p) of f so that we have containments $K(t) \subset K(x) \subset K(y)$ corresponding to the morphisms $N \xrightarrow{p} \mathbb{P}^1 \xrightarrow{f} \mathbb{P}^1$ with $f(x) = t$ and $g(y) = x$ for some rational function $g(y)$. Let us set $a(y) = f(g(y))$.

From (2.3) we obtain that

$$\sum_{Q\in\mathbb{P}^1}\tfrac{1}{\varepsilon_\pi(Q)} = 1 + \tfrac{2}{\deg\pi} \tag{2.5}$$

which is greater than 1 but less than $1+\frac{2}{60}$ since $\deg\pi \geq \deg f > 60$. If every $\varepsilon_\pi(Q)$ were greater than two then the left-hand side would be too small, so at least one lcm-ramification index must be equal to 2. Suppose that one lcm-ramification index is 2 and the others are greater than 2. Since $\frac{1}{2}+\frac{1}{4}+\frac{1}{4}=1$ at least one of them is less than 4 so must be 3, but for any integer $s \geq 6$ we have that $1 < \frac{1}{2}+\frac{1}{3}+\frac{1}{s} \leq 1$, and so s is either 4 or 5, but then $\frac{1}{2}+\frac{1}{3}+\frac{1}{s} < 1$ which also is a contradiction. We conclude that the three ramification indices are $(2,2,s)$ for some integer s, and from (2.5) we have $s = \frac{\deg\pi}{2}$.

We will now show that $a(y)$ is equal to $y^s + y^{-s}$ up to composing with automorphisms of $\mathbb{P}^1$ on either side. By composing with automorphisms we may assume that a maps 0 and ∞ to ∞ with ramification index s and that the other critical values are 2 and -2. It follows that $a(y)$ has the form $b(y)y^{-s}$ where $b(y)$ is a polynomial of degree $2s$ and has nonzero constant term. Since the preimages of 2 have ramification index 2 it follows that $b(y) - 2y^s = c(y)^2$ and likewise for -2, so that $b(y) + 2y^s = d(y)^2$ for some polynomials $c(y)$ and $d(y)$ of degree s. Putting these together we have $b(y)^2 - 4y^{2s} = e(y)^2$, where $e(y) = c(y)d(y)$. Then $4y^{2s} = (b(y)-e(y))(b(y)+e(y))$, which shows that $b(y) - e(y) = uy^i$ and $b(y) + e(y) = 4u^{-1}y^{2s-i}$ for some nonzero constant u and $0 \leq i \leq 2s$. This shows that $b(y) = \frac{u}{2}y^i + \frac{2}{u}y^{2s-i}$ which has no constant term, and so $i \in \{0, 2s\}$. In either case we have $b(y) = vy^{2s} + v^{-1}$ for some nonzero constant v which implies $a(y) = b(y)y^{-s} = vy^s + (vy^s)^{-1}$. By precomposing with the automorphism $y \mapsto v^{1/s}y$ we may assume $v = 1$. We have shown that $a(y)$ is equal to $y^s + y^{-s}$ up to automorphisms.

The Galois group of $K(y)/K(t)$ is generated by the automorphisms $y \mapsto y^{-1}$ and $y \mapsto \zeta y$ for an s-th root of unity in K. The group they generate will be D_{2s}, the dihedral group of order $2s$. Because $K(y)$ is the Galois closure, the elements

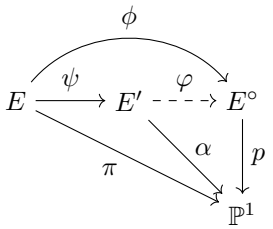
fixing $K(x)$ form a subgroup which cannot contain any normal subgroups. The only possibilities are either a trivial group or a group of order two generated by an involution of the form $\tau\colon y \mapsto \zeta y^{-1}$ for some s-th root of unity in K. In the former case, $g(y)$ is an automorphism of $\mathbb{P}^1$. Since $a(y) = f(g(y))$ we have that $f(x) = x^s + x^{-s}$ after composing with the functional inverse to $g(y)$ and $s = \frac{1}{2}\deg a = \frac{1}{2}\deg f$. In the latter case, $K(x)$ is the subfield of $K(y)$ fixed by τ, which must be $K(x) = K(y + \zeta y^{-1})$. The equality of these fields implies there is an automorphism whose composition with x is equal to $y + \zeta y^{-1}$. After precomposing with the functional inverse to $g(y)$ we have $a(y) = y^s + y^{-s} = f(y + \zeta y^{-1})$. Upon making the substitution $y = \sqrt{\zeta}w$, this becomes $a(w) = -w^s - w^{-s} = f(w\sqrt{\zeta}^{-1} + \sqrt{\zeta}w^{-1})$. By the defining equation for Chebyshev polynomials (2.1), we have that $f(x) = -T_s(x)$. In this case, $\deg a = 2\deg f$ so that $s = n$. Thus $f(x)$ is equal to $T_n(x)$ up to composing with automorphisms.

We now verify the ramification data of these three functions. The rational function $x^s + x^{-s}$ has critical points at 0 and ∞ with ramification degree s. The derivative vanishes when x is an n-th root of unity and so 2 and -2 are critical values; an n-th root of unity lies over 2 or -2 according to whether it is an even or odd power of a primitive n-th root of unity. In either case the ramification degree of any one of these roots of unity is 2. This verifies the third case. For the power map and the Chebyshev polynomial, the multiset descriptions follow easily from the description of their ramification given in §2.2.1.1. □

The next proposition proves that any Galois morphism from a genus one curve to a genus zero curve may be canonically factored into an isogeny and a reduced map.

Proposition 2.2.4. *Let E be a genus one curve and let $\pi\colon E \to \mathbb{P}^1$ be a nonconstant Galois morphism. Let $S \subset \mathbb{P}^1$ be the subset of points Q such that $\varepsilon_\pi(Q)$ is maximal.*

There exists a genus one curve E° such that π factors as $E \xrightarrow{\phi} E^\circ \xrightarrow{p} \mathbb{P}^1$ where p is Galois and totally ramified over any $Q \in S$. Moreover, E° has the following universal property: for any genus one curve E' such that π factors as $E \xrightarrow{\psi} E' \xrightarrow{\alpha} \mathbb{P}^1$, there exists a map $\varphi\colon E' \to E^\circ$ such that $\phi = \varphi\psi$ and $\alpha = p\varphi$.



Proof. Let G be the group of automorphisms of E commuting with π and let ω be a nonzero regular one-form on E. The canonical map $G \to K^\times\colon \sigma \mapsto \omega^\sigma/\omega$ is clearly independent of the choice of ω. Let G° be the kernel of this homomorphism and take $E^\circ = E/G^\circ$ with its quotient map $\phi\colon E \to E^\circ$. Therefore ϕ is unramified and so E° is genus one by the Riemann–Hurwitz formula. Let $q\colon E^\circ \to E/G$ be the map induced by the inclusion $G^\circ \subset G$ and let $i\colon E/G \xrightarrow{\sim} \mathbb{P}^1$ be the map induced by π. We have a factorization $iq\phi = \pi$.

The quotient G/G° is isomorphic to a subgroup of $K^\times$ so it is a cyclic group; let σ be a generator. Any lift of σ to G must have a fixed point in E — G° was precisely the fixed-point-free elements of G — and so σ has a fixed point $P \in E^\circ_K$. As P is fixed by every element of G/G° it is a totally ramified critical point of q. Because ϕ and i are unramified we have $\varepsilon_\pi(i(P)) = \varepsilon_q(P)$ for any $P \in E^\circ_K$. As $\varepsilon_q(P)$ divides $\deg q$ with equality precisely when P is a fixed point of every element of G/G°, it follows that $\varepsilon_q(P) = \deg q$ if and only if $i(P) \in S$. Then $p = iq$ is totally ramified over all the points of S.

Now we prove that E° has the claimed universal property. Suppose π factors

through a genus one curve E' as $E \xrightarrow{\psi} E' \xrightarrow{\alpha} \mathbb{P}^1$, $\pi = \alpha\psi$. We view ψ as an isogeny of elliptic curves by giving E and E' the base-points P and $\psi(P)$, respectively. It follows that $\ker\psi$ — with its natural group structure inherited from (E, P) — is isomorphic to $\mathrm{Gal}(\psi)$ and its natural translation action on E may be identified with the action of $\mathrm{Gal}(\psi)$ on E. It is well-known that any one-form on E is translation-invariant so that $\mathrm{Gal}(\psi) \leq G^\circ$ and it follows that there exists a map $\varphi\colon E' \to E^\circ$ such that $\phi = \varphi\psi$.

Showing that $\alpha = p\varphi$ amounts to an easy diagram chase. Let $Q \in E'_K$ and consider a ψ-preimage P of Q. Then

$$p(\varphi(Q)) = p(\phi(P)) = \pi(P) = \alpha(\psi(P)) = \alpha(Q).$$

As Q was arbitrary this shows that $\alpha = p\varphi$. □

The next lemma proves that there are only four types of reduced maps, and that they may be distinguished on the basis of their ramification. When these maps are used to construct Lattès maps, the degree 2 reduced maps correspond to "flexible" Lattès maps, while the degree 3, 4, and 6 reduced maps correspond to "rigid" Lattès maps (see [Sil07, §6]). Accordingly, for the degree 2 reduced map there is no constraint on the j-invariant, whereas for the other three cases a specific isomorphism class of elliptic curves must be utilized in order to realize the reduced map.

Lemma 2.2.5. *Let E be a genus one curve with j-invariant $j(E)$ and let $p\colon E \to \mathbb{P}^1$ be a reduced morphism. Then $(\deg p, j(E), \mathcal{R}_p, \mathcal{B}_p)$ satisfies one of the rows of the following table: (for each divisor the points are distinct)*

$\deg p$	$j(E)$	$\mathcal{R}_p$	$\mathcal{B}_p$
2	–	$[P_1]+[P_2]+[P_3]+[P_4]$	$[Q_1]+[Q_2]+[Q_3]+[Q_4]$
3	0	$2[P_1]+2[P_2]+2[P_3]$	$2[Q_1]+2[Q_2]+2[Q_3]$
4	1728	$[P_1]+[P_2]+3[P_3]+3[P_4]$	$2[Q_1]+3[Q_2]+3[Q_3]$
6	0	$[P_1]+[P_2]+[P_3]+$ $2[P_4]+2[P_5]+5[P_6]$	$3[Q_1]+4[Q_2]+5[Q_3]$

Proof. The Riemann–Hurwitz formula for p gives the constraint that $\sum_{Q\in\mathbb{P}^1}(1-\varepsilon_p(Q)^{-1})=2$. The finitely many solutions in the positive integers $(\varepsilon_p(Q))_Q$ can be found in an elementary way: they are $(2,2,2,2)$, $(3,3,3)$, $(2,4,4)$, and $(2,3,6)$. This suffices to determine the ramification and branching divisors. For instance, for $(2,3,6)$, the degree of p is 6 as p is reduced. It follows that the branching divisor is

$$\begin{aligned}\mathcal{B}_p &= (6-\tfrac{6}{2})[Q_1]+(6-\tfrac{6}{3})[Q_2]+(6-1)[Q_3]\\ &= 3[Q_1]+4[Q_2]+5[Q_3].\end{aligned}$$

The other three cases are completely similar.

□

The next proposition collects the results we will need about the ramification of rational functions with Galois closure genus one.

Proposition 2.2.6. *Let f be a rational function such that $f^{\circ 2}$ has Galois closure genus one. Suppose that* $\deg f > 24$. *Let (N,p) be the Galois closure of $f\colon\mathbb{P}^1\to\mathbb{P}^1$. Then:*

- ◇ *for any point P, $e_f(P)\in\{1,2,3,4,6\}$,*
- ◇ *for any $Q\in\mathbb{P}^1$, there are at most four preimages of Q which have ramification index $\neq\varepsilon_f(Q)$,*

Proof. Let (N_2, p_2) denote the Galois closure of $f^{\circ 2}\colon \mathbb{P}^1 \to \mathbb{P}^1$, and let $\pi = f^{\circ 2} \circ p_2$. By Proposition 2.2.4 the morphism π factors as $N_2 \xrightarrow{\phi} E^\circ \xrightarrow{q} \mathbb{P}^1$ where q is reduced and E° is a genus one curve. By the Riemann–Hurwitz formula applied to ϕ, we have that

$$0 = 2\mathfrak{g}_{N_2} - 2 = (\deg \phi)(2\mathfrak{g}_{E^\circ} - 2) + |\mathcal{R}_\phi| = |\mathcal{R}_\phi|,$$

which shows that ϕ is unramified.

Let P be any point of $\mathbb{P}^1$. By multiplicativity of ramification indices, it follows that $e_f(P)$ divides $\varepsilon_{f^{\circ 2}}(f(P))$. Let R be any p_2-preimage of P. By Lemma 2.1.8 we see that

$$\varepsilon_{f^{\circ 2}}(f(P)) = e_\pi(R) = e_q(\psi(R)).$$

Since q is reduced, there are only four possibilities for the degree of q. From Lemma 2.2.5 we see that $\deg q \in \{2, 3, 4, 6\}$. From the explicit expressions for the branching divisors it follows that $e_q(Q) \in \{1, 2, 3, 4, 6\}$ for any point Q of N_2. This proves the first claim.

We now prove the second claim. First suppose that N is genus zero. Then f is constrained by Lemma 2.2.3 to one of three possibilities, and in each of them the claim is immediately verified.

Now suppose that N is genus one. We factor p using Proposition 2.2.4 to obtain a genus one curve E and a factorization $N \xrightarrow{\psi} E \xrightarrow{w} \mathbb{P}^1$ where w is reduced. By the same Riemann–Hurwitz argument as before, ψ is an unramified map. Since p is Galois, for any point $P \in \mathbb{P}^1$ the ramification indices $e_p(R)$ of points R which are in the fiber of p over P are all equal (Lemma 2.1.8). Let Q be a point of $\mathbb{P}^1$ and suppose P is an f-preimage of Q such that $e_f(P) \neq \varepsilon_f(Q)$. Let R be any p-preimage

of P. We have from Lemma 2.1.8 that

$$\varepsilon_f(Q) = e_{f\circ p}(R) = e_p(R)e_f(P) = e_{w\circ\psi}(R)e_f(P) = e_w(\psi(R))e_f(P).$$

Since $e_f(P) \neq \varepsilon_f(Q)$, this shows that $e_w(\psi(R)) > 1$, hence that $w(\psi(R)) = p(R) = P$ is a critical value of w. By Lemma 2.2.5, there are at most four critical values of w. This proves the second claim.

□

2.3 Lifting

Let f be a power map, a Chebyshev polynomial, or a Lattès rational function. Although it is not obvious, it turns out that there is almost always a canonical choice for the group variety G which realizes f. Aside from a few exceptions in low degree, it turns out that the Galois closure (N, p) of f contains a group variety which realizes f as a dynamically affine map. This means that there is a morphism $F\colon N \to N$ and a group variety $G \subset N$ with $F(G) \subset G$, such that the diagram

$$\begin{array}{ccc} N & \xrightarrow{F} & N \\ {\scriptstyle p}\downarrow & & \downarrow{\scriptstyle p} \\ \mathbb{P}^1 & \xrightarrow{f} & \mathbb{P}^1 \end{array}$$

commutes.

The above diagram has another property which is also special to dynamically affine maps: this diagram is *cartesian* in the category of smooth projective curves. This means that the smooth fiber product $\mathbb{P}^1 \widetilde{\times}_{f,p} N$ is isomorphic to N over $\mathbb{P}^1$.

The next definition axiomatizes this property to a general morphism $f\colon C \to C$ and Galois morphism $\pi\colon D \to C$. To emphasize the intuition that the left copy of $D \to C$ is the "pullback" of the right copy of $D \to C$, let us write f^*D for the smooth fiber product $C \widetilde{\times}_{f,\pi} D$ and $f^*\pi$ for the projection map $f^*\pi\colon f^*D \to C$.

Definition 2.3.1. Let f be a noninvertible endomorphism of a curve C. Let $\pi\colon D \to C$ be a (generically) Galois morphism and let F be an endomorphism of D. We say that f lifts along π to F if $f \circ (f^*\pi)$ is (generically) Galois, and the smooth fiber product f^*D is isomorphic to D over C.

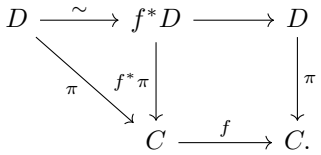


Figure 2.6: Lifting f along a Galois morphism π.

In this section we will prove two theorems. The first theorem is a Galois-theoretic characterization of dynamically affine maps.

Theorem 2.3.2. *Let f be a rational function of degree > 60. The following are equivalent:*

- *f is dynamically affine;*
- *the genus of the Galois closure of $f^{\circ r}$ is bounded independently of r.*

Our second theorem proves that dynamically affine maps admit liftings to their Galois closure.

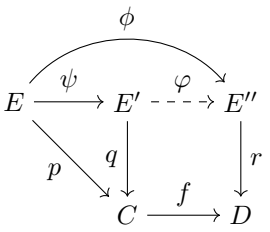
Theorem 2.3.3. *Any dynamically affine rational function of degree > 12 lifts along its Galois closure $p\colon N \to \mathbb{P}^1$ to a morphism $A\colon N \to N$. The étale locus[4] of A is a group variety $G \subset N$, $A(G) \subset G$, and $A|_G$ is an affine morphism.*

We begin by studying rational functions whose iterates have Galois closure ≤ 1.

Lemma 2.3.4. *Let f be a rational function such that the Galois closures of f and $f^{\circ 2}$ have genus one. Let (E, p) denote the Galois closure of f. Then p is reduced.*

[4] i.e., the largest subset of N on which A is unramified.

Proof. Consider f as a morphism of genus zero curves $C \to D$ where C (resp. D) has coordinate x (resp. y). Apply Proposition 2.2.4 to the composite map $fp\colon E \to D$ to get a genus one curve E'' and maps $\phi\colon E \to E''$, $r\colon E'' \to D$. Next apply Proposition 2.2.4 to the map $p\colon E \to C$ to get a genus one curve E' and maps $\psi\colon E \to E'$, $q\colon E' \to C$. Proposition 2.2.4 guarantees the existence of a map $\varphi\colon E' \to E''$ making the diagram commute:



We regard all function fields as subfields of an algebraic closure of K_D. Consider the subfield F of $K_{E'}$ which is fixed by both $\mathrm{Gal}(\varphi)$ and $\mathrm{Gal}(q)$. Then F is a subfield of $K_{E''}$ and K_C inside of $K_{E'}$. We claim that the intersection L of $K_{E''}$) and K_C is K_D. Any critical value of r is a critical value of f, and vice versa. As the Galois closures of f and $f^{\circ 2}$ have genus one, it follows that $\varepsilon_f(Q) = \varepsilon_{f^{\circ 2}}(Q)$ for all $Q \in D$. Let Q be a totally ramified critical value of r and set $R = h(Q)$. Then $\varepsilon_f(Q)$ is maximal among ramification indices of f — and also $f^{\circ 2}$ — so that $\varepsilon_{f^{\circ 2}}(R) \leq \varepsilon_f(Q)$, but the reverse inequality clearly also holds so that $\varepsilon_{f^{\circ 2}}(R) = \varepsilon_f(Q)$. But then $\varepsilon_{f^{\circ 2}}(R) = \varepsilon_f(R) = \varepsilon_r(R)$ is maximal among ramification indices of r, hence R is a totally ramified critical value of r. If every totally ramified critical value of r had no f-unramified f-preimages then we would have $\varepsilon_{f^{\circ 2}}(R) > \varepsilon_f(R)$, which would be a contradiction.

It follows that some totally ramified critical value Q of r has an f-unramified f-preimage. Then the number of places of L over Q is simultaneously equal to both one and $[L : K_D]$, so $L = K_D$. As $K_D \subset F \subset L$ we conclude $F = K_D$. Hence

fq is Galois with Galois group generated by $\mathrm{Gal}(q)$ and $\mathrm{Gal}(\varphi)$ inside of $\mathrm{Aut}(E')$; however, E was the Galois closure of f so that ψ must be an isomorphism. Since q is reduced we conclude p is reduced. □

Proposition 2.3.5. *Let f be a rational function of degree $n > 60$. The following are equivalent:*

(a) $f^{\circ r}$ has genus zero Galois closure for every integer $r \geq 1$,

(b) $f^{\circ r}$ has genus zero Galois closure for some integer $r > 1$,

(c) f is conjugate to $x^{\pm n}$ or $\pm T_n(x)$,

(d) there is a reduced map $p\colon \mathbb{P}^1 \to \mathbb{P}^1$ and an affine morphism $A\colon \mathbb{G}_m \to \mathbb{G}_m$ such that $f \circ p$ is Galois, and the diagram

$$\begin{array}{ccc} \mathbb{G}_m & \xrightarrow{A} & \mathbb{G}_m \\ {\scriptstyle p}\downarrow & & \downarrow{\scriptstyle p} \\ \mathbb{P}^1 & \xrightarrow{f} & \mathbb{P}^1 \end{array}$$

commutes.

Proof. Property (d) implies (c) by the classification of dynamically affine maps (Theorem 2.2.3). We have shown that (c) implies (d) already (in §2.2.1.1) by directly constructing p and A, however we did not show that $f \circ p$ is Galois. We show this now.

Suppose f is conjugate to $x^{\pm n}$. There is no loss of generality in assuming that f is simply equal to $x^{\pm n}$. The field extension induced by $x^{\pm n}$ is $K(t) \to K(x)\colon t \mapsto x^{\pm n}$. For x^n (resp. x^{-n}) this realizes $K(x)$ as the splitting field of $x^n - t \in K(t)[x]$ (resp. $x^n - t^{-1} \in K(t)[x]$). By assumption, K is characteristic zero so this extension is also separable. It follows that $x^{\pm n}$ is Galois and $p = \mathrm{id}$, so $f \circ p = f$ is Galois.

Now suppose f is conjugate to ϵT_n, $\epsilon \in \{-1, 1\}$. Again, there is no loss of generality in assuming that f is simply equal to ϵT_n. Recall the defining equation of

the Chebyshev polynomials:

$$T_n(x + x^{-1}) = x^n + x^{-n}. \tag{2.6}$$

The field extension induced by ϵT_n is $K(t) \to K(x)\colon t \mapsto \epsilon T_n$, and we identify $K(t)$ with its image under this embedding. Consider the quadratic extension $K(y)$ over $K(x)$ given by adjoining a root of $y^2 - xy + 1 \in K(x)[y]$. Using (2.6) we obtain that $\epsilon t = T_n(x) = T_n(y+y^{-1}) = y^n + y^{-n}$, or equivalently that $y^{2n} - \epsilon t y^n + 1 = 0$. The roots of this equation in $K(y)$ are $y^{\pm 1}\omega^i$, $i = 0, \ldots, n-1$, where ω is a primitive n-th root of unity in K. This shows that $K(y)$ is the splitting field of $y^{2n} - \epsilon t y^n + 1 \in K(t)[y]$ over $K(t)$. It follows that the function field of $N_{\epsilon T_n}$ is isomorphic to $K(y)$. The morphism $p\colon N_{\epsilon T_n} \to \mathbb{P}^1$ corresponding to the inclusion $K(x) \to K(y) : x \mapsto y + y^{-1}$ is Galois, and has a totally ramified critical point at 1. We have shown that (c) and (d) are equivalent.

Clearly (a) implies (b) so we will show that (c) implies (a) and that (b) implies (c). Assume f satisfies (c). We have shown that f then satisfies (d), which proves that f has a genus zero Galois closure. However, the set of power maps and (signed) Chebyshev polynomials is closed under iteration (Proposition 2.2.1), so we have already shown (a).

Now assume that f satisfies (b). The function field of the Galois closure (N_r, p_r) of the iterate $f^{\circ r}$ contains the function field of the Galois closure of the second iterate $f^{\circ 2}$ (both function fields considered in the algebraic closure of $K(t)$). This produces a canonical map $N_r \to N_2$. By the Riemann–Hurwitz formula, the genus of N_2 must be $\leq$ the genus of N_r, which is zero, and so $f^{\circ 2}$ must also have genus zero Galois closure. Applying Lemma 2.2.3 to $f^{\circ 2}$ constrains the ramification multi-set of $f^{\circ 2}$ to one of the following possibilities:

1. $[n^2]$ over each of two points P and Q,

2. $[n^2]$ over one point P, $[1^2, 2^{n^2-1}]$ over a set of two other points $\{Q, R\}$,

3. $[\frac{n^2}{2}, \frac{n^2}{2}]$ over one point, $[2^{n^2/2}]$ over each of two other points.

Case (3) cannot actually occur. Indeed ramification indices are multiplicative over compositions of dominant maps, so any ramification index for the composite $f^{\circ 2}$ must be realized as a product of ramification indices from a single case of Lemma 2.2.3. However the integer $\frac{n^2}{2}$ cannot be realized as a product of two integers taken from any one of the sets $\{n\}$, $\{n, 1, 2\}$, or $\{\frac{n}{2}, 2\}$.

It is well-known that the automorphism group of the projective line acts 3-transitively.[5] By replacing f with a conjugate we may thus assume that $(P, Q) = (\infty, 0)$ in case (1) and $(P, Q, R) = (\infty, 2, -2)$ in case (2). In the first case, f will have ramification index n over each point of $\{0, \infty\}$. In order for $f^{\circ 2}$ to achieve a ramification index of n^2, f must stabilize the set $\{0, \infty\}$. Let $p(x) := f(1/x)$ if case (1) is satisfied and $f(0) = \infty$, and let $p(x) := f(x)$ otherwise. Then in the first case we have that $p(0) = 0$, while in either case we have that $p(\infty) = \infty$. By Lemma 2.2.3 there exist automorphisms μ and ν of $\mathbb{P}^1$ such that p equals either $\mu \circ x^n \circ \nu$ or $\mu \circ T_n \circ \nu$.

Suppose p is $\mu \circ x^n \circ \nu$. There are two critical values of x^n and both are fixed and totally ramified so we are in the first case. Either μ and ν both fix 0 and ∞, or they both permute 0 and ∞. Hence $(\mu, \nu) = (ax^\epsilon, bx^\epsilon)$ for constants $a, b \in K$ and $\epsilon \in \{-1, 1\}$, so $f = cx^{\pm n}$ for some constant $c \in K$. If $d \in K$ is any solution of $c = d^{-1 \mp n}$, then the automorphism $\tau(x) = dx$ satisfies $\tau \circ f \circ \tau^{-1} = x^{\pm n}$.

Suppose p is $\mu \circ T_n \circ \nu$. The polynomial T_n has only one totally ramified critical value so we are in the second case. In order for $f^{\circ 2}$ to achieve a ramification index

[5] i.e., $\operatorname{Aut} \mathbb{P}^1$ acts transitively on $\{(P, Q, R) : P, Q, R \in \mathbb{P}^1, P, Q, R \text{ distinct}\}$.

of n^2 over ∞, $p\ (= f)$ must fix ∞ as it is the unique totally ramified critical point of f and has ramification index n. Then ν must also fix ∞ for ∞ is the unique totally ramified critical point of T_n, and it follows that μ fixes ∞ also. From the defining equation (2.6) of the Chebyshev polynomials it is easily determined that the finite critical values of T_n are precisely $\{2, -2\}$. It follows that μ stabilizes $\{2, -2\}$. As we are in the second case, the set of the two unramified pre-images of f over the set $\{2, -2\}$ must be $\{2, -2\}$, since otherwise the second iterate of f would have a ramification index of 4. Hence ν stabilizes $\{2, -2\}$ also. It is easily found that the only automorphisms of $\mathbb{P}^1$ fixing ∞ and stabilizing $\{2, -2\}$ are $\{x, -x\}$. It follows that f is conjugate to $\pm T_n(x)$. □

The next proposition is the complement of Proposition 2.3.5 for genus one Galois closure.

Proposition 2.3.6. *Let f be a nonconstant rational function of degree > 12. The following conditions are equivalent:*

(a) $f^{\circ r}$ has genus one Galois closure for every integer $r \geq 1$,

(b) $f^{\circ r}$ has genus one Galois closure for some integer $r > 1$,

(c) f is a Lattès map,

(d) there is a reduced map $p\colon E \to \mathbb{P}^1$ and an affine morphism $A\colon E \to E$ such that $f \circ p$ is Galois and the diagram

$$\begin{array}{ccc} E & \xrightarrow{A} & E \\ {\scriptstyle p}\downarrow & & \downarrow{\scriptstyle p} \\ \mathbb{P}^1 & \xrightarrow{f} & \mathbb{P}^1 \end{array}$$

commutes.

Proof. Condition (a) clearly implies (b). We will show that $(b) \implies (d) \implies (c) \implies (a)$.

Assume f satisfies (b). Let $F\colon X \to Y$ be any nonconstant morphism of curves with Galois closure (N, p), and let R be a point in Y. Recall that the ramification index $e_{F\circ p}(P)$ under $F \circ p$ of any point $P \in N$ such that $F(p(P)) = R$ is equal to $\varepsilon_F(R)$ (Lemma 2.1.8). It follows by multiplicativity of ramification indices that $e_p(P) = \varepsilon_F(R)/e_F(p(P))$. Since p is Galois we have that $r_p(p(P)) = \deg p/e_p(P)$, and therefore

$$\mathcal{B}_p = \sum_{Q\in X}(\deg p - r_p(Q)) = \sum_{Q\in X}(\deg p)\left\{1 - \frac{e_F(Q)}{\varepsilon_F(F(Q))}\right\}[Q]. \tag{2.7}$$

Let (E_i, p_i) be the Galois closure of $f^{\circ i}$ $(i = 1, 2)$ and let $\mathcal{B}_i$ be the branching divisor of p_i. Using (2.7) we have that

$$\mathcal{B}_i = \sum_{Q\in\mathbb{P}^1}(\deg p_i)\left\{1 - \frac{e_{f^{\circ i}}(Q)}{\varepsilon_{f^{\circ i}}(f^{\circ i}(Q))}\right\}[Q] \quad \text{for } i = 1, 2. \tag{2.8}$$

If P is any preimage of f over Q, then $e_{f^{\circ 2}}(P) = e_f(P)e_f(Q)$ divides $\varepsilon_{f^{\circ 2}}(f^{\circ 2}(P))$. As this holds for any preimage P of Q, it follows that $\varepsilon_f(Q)e_f(Q)$ divides $\varepsilon_{f^{\circ 2}}(f^{\circ 2}(P))$ also. This shows that

$$\varepsilon_f(Q)e_f(Q) = \varepsilon_f(Q)\Big(\frac{e_{f^{\circ 2}}(P)}{e_f(P)}\Big) \text{ divides } \varepsilon_{f^{\circ 2}}(f^{\circ 2}(P)) \text{ for all } P \in \mathbb{P}^1. \tag{2.9}$$

Upon multiplying (2.9) on both sides by $e_{f^{\circ 2}}(P)^{-1}$ we obtain that

$$\frac{\varepsilon_f(f(P))}{e_f(P)} \text{ divides } \frac{\varepsilon_{f^{\circ 2}}(f^{\circ 2}(P))}{e_{f^{\circ 2}}(P)} \text{ for all } P \in \mathbb{P}^1. \tag{2.10}$$

Note that both quantities in (2.10) are integers.

Let us write $\mathcal{B}_1 := \sum_{Q\in\mathbb{P}^1} a_Q[Q]$, $\mathcal{B}_2 := \sum_{Q\in\mathbb{P}^1} b_Q[Q]$, and $d_i := \deg p_i$. From (2.8) we have that

$$d_1\Big(1 - \frac{e_f(Q)}{\varepsilon_f(f(Q))}\Big) = a_Q,$$

and by rearranging we obtain

$$(1 - a_Q d_1^{-1})^{-1} = \frac{\varepsilon_f(f(Q))}{e_f(Q)}.$$

Likewise we get that

$$(1 - b_Q d_2^{-1})^{-1} = \frac{\varepsilon_{f^{\circ 2}}(f^{\circ 2}(Q))}{e_{f^{\circ 2}}(Q)}.$$

From (2.10) we have that

$$(1 - a_Q d_1^{-1})^{-1} \text{ divides } (1 - b_Q d_2^{-1})^{-1} \text{ for all } Q \in \mathbb{P}^1 \tag{2.11}$$

where both quantities in (2.11) are integers. We claim that (2.11) suffices to show $\mathcal{B}_1 = \mathcal{B}_2$.

For any positive integer r there is a canonical map

$$i_r \colon N_{f^{\circ(r+1)}} \to N_{f^{\circ r}}$$

coming from the field inclusion of the respective function fields inside the algebraic closure of $K(t)$. By the Riemann–Hurwitz formula, it follows that $N_{f^{\circ 2}}$ has genus $\mathfrak{g}_2$ which is $\leq \mathfrak{g}_r = 1$. If the Galois closure of $f^{\circ 2}$ were genus zero, then Proposition 2.3.5 would imply that $f^{\circ r}$ would have genus zero Galois closure for $r \gg 0$, which is a contradiction. Hence the genus of $N_{f^{\circ 2}}$ is one. By Proposition 2.2.6, every ramification index of f is ≤ 6. If the Galois closure of f were genus zero, then Lemma 2.2.3 implies that f would have some ramification index which is at least $\deg f/2$. As $\deg f/2 > 6$, this is a contradiction, and so the Galois closure of f is genus one.

It follows from Lemma 2.3.4 that p_1 and p_2 are reduced morphisms from genus one curves to genus zero curves. Therefore there are only four possibilities which are tabulated by Lemma 2.2.5. First consider the case that $d_2 = 2$. By Lemma 2.2.5 we have that $\mathcal{B}_2 = [Q_1] + [Q_2] + [Q_3] + [Q_4]$ for distinct points Q_i. As N_1 is reduced, there are only four possibilities for $\mathcal{B}_1$. From (2.11) we have the constraint that $(1 - a_{Q_i} d_1^{-1})^{-1}$ divides $(1 - b_{Q_i} d_2^{-1})^{-1} = (1 - \frac{1}{2})^{-1} = 2$. Considering a as an integer variable in the equation $(1 - a d_1^{-1})^{-1} = 2$ where d_1 ranges over the four possibilities

$\{2,3,4,6\}$ of Lemma 2.2.5, we find that $d_1 = 3$ is disallowed (corresponds to the non-integral solution $a = 3/2$). The remaining possible cases are

$$(1 - ad_1^{-1})^{-1} = i \iff \begin{cases} a = 0, & i = 1, \\ a = 1, & i = 2, \quad d_1 = 2, \\ a = 2, & i = 2, \quad d_1 = 4, \\ a = 3, & i = 2, \quad d_1 = 6. \end{cases}$$

A direct case-by-case inspection of the branching divisors in Lemma 2.2.5 shows that the only possibility is if $d_1 = 2$ and $\mathcal{B}_1 = \mathcal{B}_2$. Similar arguments prove that $\mathcal{B}_1 = \mathcal{B}_2$ for the other three cases of Lemma 2.2.5 when $d_2 = 3, 4$ and 6.

By Lemma 2.2.2 there is an isomorphism $\rho\colon E_1 \xrightarrow{\sim} E_2$ such that $p_1 = p_2\rho$. As E_1 is the Galois closure of f, there is a map $\psi_0\colon E_2 \to E_1$ such that $p_1\psi_0 = fp_2$. Set $\psi = \psi_0\rho$. Any nonconstant morphism $\psi\colon E \to E$ of degree > 1 for an elliptic curve E is an affine morphism. Indeed if $\mathcal{O} \in E$ is the identity section of E, then we may postcompose ψ with a translation T to obtain a morphism $T\psi$ such that $(T\psi)(\mathcal{O}) = \mathcal{O}$, and it is well-known that any morphism $E \to E$ sending $\mathcal{O}$ to $\mathcal{O}$ is automatically a homomorphism of groups, [Sil09, III.4.8]. This shows that f is a Lattès map for the genus one curve E with affine morphism ψ. This proves that (b) implies (d).

Assume f satisfies (d). Then f is a Lattès map by definition. By [Sil12, Theorem 3.26], the Galois closure of any iterate $f^{\circ r}$ is genus one. This shows that (c) implies (a). □

We may now prove Theorem 2.3.2 and Theorem 2.3.3.

Proof of Theorem 2.3.2. For any positive integer r there is a canonical map

$$i_r\colon N_{f^{\circ(r+1)}} \to N_{f^{\circ r}}$$

coming from the field inclusion of the respective function fields inside the algebraic closure of $K(t)$, and the degree of i_r is the degree of the corresponding field extension. For sufficiently large r the genera on both sides of this map are independent of r, say equal to $\mathfrak{g}$. Applying the Riemann–Hurwitz formula to i_r shows that $2\mathfrak{g} - 2 = \deg i_r(2\mathfrak{g} - 2) + |\mathcal{R}_{i_r}|$. If $\mathfrak{g} > 1$ then $\deg i_r = 1$, implying that the corresponding function fields are equal. This cannot happen since the function field of $N_{f^{\circ r}}$ over $K(t)$ contains the function field corresponding to $f^{\circ r}$, so its degree over $K(t)$ must go to infinity.

Hence the genus of the Galois closure of $f^{\circ r}$ is ≤ 1 for all $r \gg 0$, and f satisfies the hypotheses of either Proposition 2.3.5 or Proposition 2.3.6. In either case we see that f is dynamically affine. □

Proof of Theorem 2.3.3. Let f be a dynamically affine rational function of degree n and let (N, p) denote its Galois closure. We have already shown that dynamically affine maps fit into commutative diagrams of the form

$$\begin{array}{ccc} N & \xrightarrow{A} & N \\ {\scriptstyle p}\downarrow & & \downarrow{\scriptstyle p} \\ \mathbb{P}^1 & \xrightarrow{f} & \mathbb{P}^1, \end{array}$$

and that N contains a group variety G ($G = \mathbb{G}_m$ for power maps and Chebyshev, $G = N$ for Lattès maps) such that $A|_G$ is an affine morphism. In both cases, G is precisely the étale locus of A, and $A(G) \subset G$.

What remains is to show that this diagram is cartesian. There is no loss of generality in assuming that f is equal to either $x^{\pm n}$, $\pm T_n(x)$, or a Lattès map, since the property of the diagram being cartesian is preserved when f is replaced by a conjugate of itself.

If f is $x^{\pm n}$ then p is an isomorphism and the diagram is automatically cartesian.

In the remaining two cases, we claim that there exists a totally ramified point Q of p such that f has an unramified preimage over Q. When f is $\pm T_n(x)$ the Galois closure map p is $x+x^{-1}$ (this was proven in Proposition 2.3.5). The map p has totally ramified critical values at $x = \pm 2$. However the critical points of the nth Chebyshev polynomial are at $x = \infty$ and $\{x + x^{-1} : x \in \mu_{2n} \backslash \{\pm 1\}\}$. In particular, $\pm T_n(x)$ is unramified at the points ± 2. For any $\varepsilon \in \{\pm 1\}$ we have that $\varepsilon T_n(2) = \varepsilon 2$, showing that 2 is an unramified preimage of the point $Q = \varepsilon 2$ which is totally ramified for p.

Now suppose f is a Lattès map. The proof of Proposition 2.3.6 showed that f and $f^{\circ 2}$ have Galois closures which are isomorphic over $\mathbb{P}^1$. Therefore we have a diagram,

$$\begin{array}{ccccc} N & \xrightarrow{F} & N & \xrightarrow{F} & N \\ {\scriptstyle p}\downarrow & & {\scriptstyle p}\downarrow & & \downarrow{\scriptstyle p} \\ \mathbb{P}^1 & \xrightarrow{f} & \mathbb{P}^1 & \xrightarrow{f} & \mathbb{P}^1, \end{array}$$

in which the left-most morphism p is the Galois closure of $f^{\circ 2}$. By the Riemann–Hurwitz formula (in the form of Lemma 2.1.9), we have the equalities

$$\sum_{Q \in \mathbb{P}^1} \left(1 - \tfrac{1}{\varepsilon_f(Q)}\right) = \sum_{Q \in \mathbb{P}^1} \left(1 - \tfrac{1}{\varepsilon_{f^{\circ 2}}(Q)}\right) = 0. \tag{2.12}$$

However $\varepsilon_f(Q) \leq \varepsilon_{f^{\circ 2}}(Q)$ for any point Q in $\mathbb{P}^1$. As the quantity $1 - \frac{1}{\varepsilon}$ is monotone-increasing in the variable ε, it follows from (2.12) that we must have equality $\varepsilon_f(Q) = \varepsilon_{f^{\circ 2}}(Q)$ for any point Q in $\mathbb{P}^1$.

Suppose that R is a totally ramified critical value of p. Let P be a point of N such that $(f \circ p)(P) = R$ and let $Q = f(R)$. Then

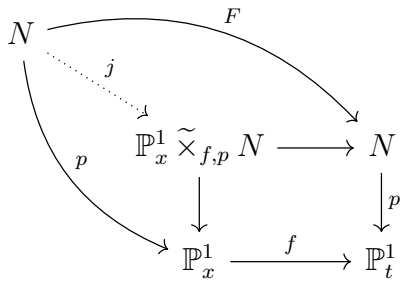
$$\deg p = \varepsilon_f(Q) = \varepsilon_{f^{\circ 2}}(Q) = \varepsilon_{pF^{\circ 2}}(Q) = \varepsilon_p(Q),$$

showing that Q is a totally ramified critical value of p. Since $\varepsilon_{fp}(Q) = \varepsilon_p(Q)$, we must have that $e_f(R) = 1$. This proves the existence of a point Q which is a totally ramified critical value for p such that f has an unramified preimage over it.

Let us choose coordinates x and t for $\mathbb{P}^1$ so that we have the diagram

$$\begin{array}{ccc} N & \xrightarrow{F} & N \\ {\scriptstyle p}\downarrow & & \downarrow{\scriptstyle p} \\ \mathbb{P}^1_x & \xrightarrow{f} & \mathbb{P}^1_t. \end{array}$$

We must show that the smooth fiber product $\mathbb{P}^1_x \widetilde{\times}_{f,p} N$ is isomorphic to N over $\mathbb{P}^1_x$. For this it suffices to show that $\mathbb{P}^1_x \widetilde{\times}_{f,p} N$ is irreducible. Indeed, the universal property of the smooth fiber product implies there is a map $j\colon N \to \mathbb{P}^1_x \widetilde{\times}_{f,p} N$ such that the diagram



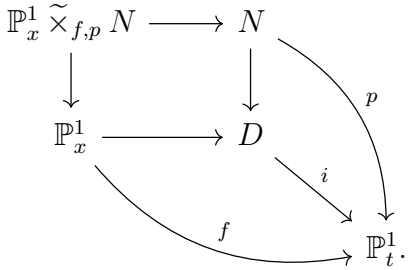
commutes. Once we have shown that $\mathbb{P}^1_x \widetilde{\times}_{f,p} N$ is irreducible, then the diagram shows that the degree of j is 1. This will show that j is an isomorphism of $\mathbb{P}^1_x \widetilde{\times}_{f,p} N$ and N over $\mathbb{P}^1_t$. The existence of the isomorphism j will also verify the other condition for f to lift to p, namely that the composite map

$$\mathbb{P}^1_x \widetilde{\times}_{f,p} N \to \mathbb{P}^1_x \xrightarrow{f} \mathbb{P}^1_t$$

is Galois. Hence to finish the proof we need only show that $\mathbb{P}^1_x \widetilde{\times}_{f,p} N$ is irreducible.

It is the same to show that the function fields of $\mathbb{P}^1_x$ and N are linearly disjoint over the function field of $\mathbb{P}^1_t$. Recall that we have shown the existence of a totally ramified point Q of p such that f has an unramified preimage P, say, over Q. We proceed by contradiction. Let $K(x)$ and K_N be the function fields of $\mathbb{P}^1_x$ and N, resp., considered as subfields of $K(t)^a$. Suppose that $K(x)$ and K_N are not linearly disjoint over $K(t)$. Since p is Galois, the failure of linear disjointness is equivalent to the existence of a subfield $K(t) \subset L \subset (K(x) \cap K_N)$ which has degree > 1 over

$K(t)$. Let $i\colon D \to \mathbb{P}^1_t$ be the curve D and nonconstant morphism i corresponding to the extension $L/K(t)$. We have that $\deg i = [L : K(t)] > 1$. We obtain the following diagram:



As Q is totally ramified for p, it has a single preimage $i^{-1}(Q)$ under i. Since $\deg i > 1$, Q is thus a critical value of i. However, by multiplicativity of ramification indices, the ramification index $e_i(i^{-1}(Q))$, which is greater than 1, must divide $e_f(P)$ which is equal to 1. This contradiction completes the proof. □

2.4 Irreducible Pairs

In this section we introduce the notion of an irreducible pair of rational functions. Let f and g be rational functions over K of degree > 1. Recall that the smooth fiber product $\mathbb{P}^1 \widetilde{\times}_{f,g} \mathbb{P}^1$ is defined to be the smooth, possibly disconnected curve associated to the tensor product of the function fields (Definition 2.1.10).

Let $\mathcal{C}^{rs}$ denote the smooth fiber product $\mathbb{P}^1 \widetilde{\times}_{f^{\circ r}, g^{\circ s}} \mathbb{P}^1$ of the iterates $f^{\circ r}$ and $g^{\circ s}$.

Definition 2.4.1. If $\mathcal{C}^{rs}$ is irreducible for all positive integers r and s, we say that f and g form an irreducible pair.

Theorem 2.4.2. *Let f and g be rational functions of degree > 60. Suppose f and g form an irreducible pair and that the genus of $\mathcal{C}^{rs}$ is bounded independently of r and s. Then there exist a one-dimensional group variety G and a Galois morphism $\pi\colon G \to \mathbb{P}^1$ such that f and g both lift along π to affine morphisms ψ and φ of G.*

In particular, it follows that the group variety G and the quotient map $\pi\colon G \to \mathbb{P}^1$ yield a *simultaneous* realization for f and g as dynamically affine rational functions.

First we will prove a useful result which lets us bound the genus of the Galois closures of the rational functions in an irreducible pair when the smooth fiber products of the iterates have low genus.

Lemma 2.4.3. *Suppose $\mathcal{C}^{11}$ is irreducible and has genus $\mathfrak{g}$. If $\deg g \gg \deg f + \mathfrak{g}$ then the Galois closure of f is genus zero or one.*

Proof. Using Lemma 2.1.12 shows that

$$\sum_{\substack{S\in C,\\ \varepsilon_f(S)>1}} \sum_{\substack{Q\in D,\\ f(Q)=S}} \sum_{\substack{R\in E,\\ g(R)=S}} \big(e_f(Q) - \gcd(e_f(Q), e_g(R))\big) \le 2(\deg f + \mathfrak{g}). \tag{2.13}$$

For any $S \in C$ let T_S be the subset of points $R \in g^{-1}(S)$ such that $\varepsilon_f(S)$ does not divide $e_g(R)$. Then for any $R \in T_S$ there is a f-preimage Q lying over S such that $e_f(Q)$ does not divide $e_g(R)$, and then $e_f(Q) - \gcd(e_f(Q), e_g(R)) \ge 1$. From (2.13) it follows that the size of T_S is at most $2(\deg f + \mathfrak{g})$. Then $r_g(S) = \#T_S + \#T_S^c \le 2(\deg f + \mathfrak{g}) + (\deg g)\varepsilon_f(S)^{-1}$. Applying the Riemann–Hurwitz formula to g gives

$$\begin{aligned}
-2 &\ge -2\deg g + \sum_{\substack{S\in C,\\ \varepsilon_f(S)>1}} (\deg g - r_g(S)) \\
&\ge -2\deg g + \sum_{\substack{S\in C,\\ \varepsilon_f(S)>1}} (\deg g - 2(\deg f + \mathfrak{g}) - (\deg g)\varepsilon_f(S)^{-1}) \\
&= (\deg g)\big\{-2 + \sum_{S\in C}(1 - \varepsilon_f(S)^{-1})\big\} - 2(\deg f + \mathfrak{g}) \cdot \#\{S \in C : \varepsilon_f(S) > 1\}.
\end{aligned}$$

This shows that if $\deg g \gg \deg f + \mathfrak{g}$ then we must have

$$\sum_{S\in C}(1 - \varepsilon_f(S)^{-1}) \le 2.$$

Using the formula for the genus of the Galois closure (Lemma 2.1.9) shows that $\mathfrak{g}_{N_f} \le 1$. □

We are ready to prove Theorem 2.4.2.

Proof of Theorem 2.4.2. We will prove that f and g have the same Galois closure N (up to isomorphism), and that f and g lift along their common Galois closure morphism $p\colon N \to \mathbb{P}^1$ to endomorphisms ψ and φ of N. Then we will show that N contains an open group variety G such that $\psi(G) \subset G$ and $\varphi(G) \subset G$, and the restrictions of ψ and φ to G are affine morphisms.

By Lemma 2.4.3 the genus of the Galois closure of the iterate $f^{\circ r}\colon \mathbb{P}^1 \to \mathbb{P}^1$ is ≤ 1 for sufficiently large r. The Galois-theoretic classification of dynamically affine maps (Theorem 2.3.2) implies f is dynamically affine, and the same consideration holds for g. Let (N_f, p_f) (resp. (N_g, p_g)) denote the Galois closure of f (resp. g). Theorem 2.3.3 implies that f lifts along p_f and g lifts along p_g.

We now show that N_f and N_g are isomorphic over $\mathbb{P}^1$. First we prove that f and g have similar ramification. Precisely, we will prove that for any point $Q \in \mathbb{P}^1$ either

1. Q is a totally ramified critical value of f and g, or
2. $\varepsilon_f(Q) = \varepsilon_g(Q)$.[6]

Let $\mathcal{C}$ denote the smooth fiber product $\mathbb{P}^1 \widetilde{\times}_{f,g} \mathbb{P}^1$, and let π and ϖ denote the projection maps:

$$\begin{array}{ccc} \mathcal{C} & \xrightarrow{\varpi} & \mathbb{P}^1 \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle g} \\ \mathbb{P}^1 & \xrightarrow{f} & \mathbb{P}^1. \end{array}$$

To prove the first claim, suppose Q is a totally ramified critical value of f which is not a totally ramified critical value of g. From this assumption we will derive a contradiction. As g is dynamically affine, its ramification indices at points which are not totally ramified are bounded from above by 6. Indeed, when g is conjugate to a power map x^n or a Chebyshev polynomial $\pm T_n(x)$ then this follows from the explicit

[6] Recall that $\varepsilon_f(Q)$ is defined to be the least common multiple of the set $\{e_f(P) : f(P) = Q\}$.

description of the ramification in §2.2.1.1. Suppose g is a Lattès map. We have already seen that g lifts to an endomorphism $\psi\colon N_g \to N_g$, and that N_g is reduced (Theorem 2.3.3 and Proposition 2.3.6). As g has degree > 60, the Galois closure N_g is genus one by Proposition 2.3.6. By the Riemann–Hurwitz formula applied to ψ, we see that

$$2\mathfrak{g}_{N_g} - 2 = (\deg \psi)(2\mathfrak{g}_{N_g} - 2) + |\mathcal{R}_\psi|,$$

which shows that $|\mathcal{R}_\psi| = 0$, i.e., ψ is unramified. Therefore for any point P in $\mathbb{P}^1$ we have that

$$e_g(P) \leq \varepsilon_g(g(P)) = \varepsilon_{g\circ p_g}(g(P)) = \varepsilon_{p_g\circ\psi}(g(P)) = \varepsilon_{p_g}(g(P)).$$

By Lemma 2.2.5 we see that $\varepsilon_{p_g}(Q)$ is ≤ 6 for any point $Q \in \mathbb{P}^1$. This proves that the ramification indices of g are ≤ 6. It follows that $r_g(Q) \geq 6^{-1} \deg g$.

Suppose P is a point of $\mathcal{C}$ such that $g(\varpi(P)) = Q$. Since $g(\varpi(P)) = f(\pi(P))$, we see that

$$e_\varpi(P) = \frac{e_\pi(P) e_f(\pi(P))}{e_g(\varpi(P))} \geq \frac{e_\pi(P)(\deg f)}{6} \geq \tfrac{1}{6} \deg f.$$

As this lower bound holds for any point P lying over $\varpi(P)$, this shows that

$$r_\varpi(\varpi(P)) = \#\mathcal{F}_\varpi(\varpi(P)) \leq (\deg \varpi)(\tfrac{1}{6} \deg f)^{-1} = 6.$$

Applying the Riemann–Hurwitz formula to ϖ obtains

$$\begin{aligned}
2\mathfrak{g}_{\mathcal{C}} - 2 &= (2\mathfrak{g}_{\mathbb{P}^1} - 2)\deg\varpi + \sum_{P\in\mathcal{C}}(e_\varpi(P) - 1) \\
&= (-2)\deg f + \sum_{R\in\mathbb{P}^1}(\deg\varpi - r_\varpi(R)) \\
&\geq -2\deg f + \sum_{R\in\mathbb{P}^1}(\deg f - 6) \\
&\geq -2\deg f + \sum_{R\in g^{-1}(Q)}(\deg f - 6) \\
&\geq -2\deg f + (\tfrac{1}{6}\deg g)(\deg f - 6) \\
&= (\deg f)(\tfrac{1}{6}\deg g - 2) - \deg g \\
&\geq (60)(\tfrac{1}{6}60 - 2) - 60 = 420.
\end{aligned}$$

The right-hand side is positive but this would imply $\mathfrak{g}_{\mathcal{C}} > 1$, a contradiction. This proves the first claim.

We also prove the second claim by contradiction. Suppose that a point $R \in \mathbb{P}^1$ is not a totally ramified critical value of either f or g, and that $\varepsilon_f(R) \neq \varepsilon_g(R)$. Set $r := \varepsilon_f(R)$ and $s := \varepsilon_g(R)$. By switching f and g if necessary, we may assume that $r > s$.

By the description of the ramification in Lemma 2.2.3 and Proposition 2.2.6, $\varepsilon_f(R)$ is less than or equal to 6, and all but at most 4 f-preimages of R have ramification index equal to $\varepsilon_f(R)$. Let P be an f-preimage of R. If $e_f(P) \neq r$ then it is a proper divisor of r, hence no greater than $r/2$. It follows that the sum of $e_f(P)$ over $\{P \in \mathcal{F}_f(R) : e_f(P) \neq r\}$ is at most $4(r/2) = 2r$. The sum of $e_f(P)$ over $\mathcal{F}_f(R)$ is equal to $\deg f$ (Lemma 2.1.6), so the sum of $e_f(P)$ over $\{P \in \mathcal{F}_f(R) : e_f(P) = r\}$ is at least $\deg f - 2r$. Therefore, the number of points in the subset $\{P \in \mathcal{F}_f(R) : e_f(P) = r\}$ is at least $\frac{1}{r}(\deg f - 2r)$.

With the help of Lemma 2.1.12 we get that

$$2\mathfrak{g}_{\mathcal{C}} - 2 = -2\deg f + \sum_{(P,Q)\in F} \big(e_f(P) - \gcd(e_f(P), e_g(Q))\big), \tag{2.14}$$

where $F = \{(P,Q) \in \mathbb{P}^1 \times \mathbb{P}^1 : f(P) = g(Q)\}$. Let Q be any g-preimage of R. Then for any f-preimage P of R we have that $(P,Q) \in F$, so the sum in (2.14) is greater than or equal to

$$\begin{aligned}
&\sum_{Q\in\mathcal{F}_g(R)} \sum_{P\in\mathcal{F}_f(R)} \big(e_f(P) - \gcd(e_f(P), e_g(Q))\big) \\
&\qquad \geq \sum_{Q\in\mathcal{F}_g(R)} \sum_{\substack{P\in\mathcal{F}_f(R),\\ e_f(P)=r}} \big(e_f(P) - \gcd(e_f(P), e_g(Q))\big) \\
&\qquad \geq \sum_{Q\in\mathcal{F}_g(R)} \tfrac{1}{r}(\deg f - 2r)\big(r - \gcd(r, e_g(Q))\big) \\
&\qquad \geq \sum_{Q\in\mathcal{F}_g(R)} \tfrac{1}{r}(\deg f - 2r)\big(r - \gcd(r, s)\big).
\end{aligned}$$

There are at least $\frac{1}{s}\deg g$ points in $\mathcal{F}_g(R)$, so we get that

$$\sum_{(P,Q)\in F} \big(e_f(P) - \gcd(e_f(P), e_g(Q))\big) \geq (\tfrac{1}{s}\deg g)\tfrac{1}{r}(\deg f - 2r)\big(r - \gcd(r,s)\big).$$

Since $r > s$, we have that $\gcd(r,s) \geq r/2$. Recall that $r \leq 6$, and since $s < r$, the next largest possibility for s is 4 (Proposition 2.2.6). Therefore

$$\begin{aligned}
\sum_{(P,Q)\in F} \big(e_f(P) - \gcd(e_f(P), e_g(Q))\big) &\geq (\tfrac{1}{s}\deg g)\tfrac{1}{r}(\deg f - 2r)(\tfrac{r}{2}) \\
&\geq (\tfrac{1}{s}\deg g)\tfrac{1}{2}(\deg f - 2r) \\
&\geq (\tfrac{1}{4}\deg g)\tfrac{1}{2}(\deg f - 12) \\
&= (\tfrac{1}{8}\deg g)(\deg f - 12).
\end{aligned}$$

Returning to (2.14) we see that

$$\begin{aligned} 2\mathfrak{g}_{\mathcal{C}} - 2 &= -2\deg f + \sum_{(P,Q)\in F} \big(e_f(P) - \gcd(e_f(P), e_g(Q))\big) \\ &\geq -2\deg f + (\tfrac{1}{8}\deg g)(\deg f - 12) \\ &\geq -2\deg f + \tfrac{15}{2}(\deg f - 12) \\ &= \tfrac{11}{2}\deg f - 90 \geq \tfrac{11}{2}60 - 90 = 240 > 0. \end{aligned}$$

This would show that $\mathfrak{g}_{\mathcal{C}} > 1$, a contradiction. This proves the second claim.

Now we will use this description of the ramification of f and g to show that N_f and N_g are isomorphic over $\mathbb{P}^1$, i.e, there is an isomorphism $\rho\colon N_f \xrightarrow{\sim} N_g$ such that $p_g\rho = p_f$.

We have already shown that f and g are dynamically affine. In particular, f and g are conjugate to either a power map, a Chebyshev polynomial, or a Lattès map. We will show that f and g are simultaneously conjugate to functions with the same Galois closure. This will show that f and g also have the same Galois closure. Suppose that μ is an automorphism of $\mathbb{P}^1$ such that $\mu f\mu^{-1}$ is either a power map, a Chebyshev polynomial, or a Lattès map, and let us write $h = \mu g\mu^{-1}$. The description of the ramification of dynamically affine maps in Proposition 2.3.5 and Proposition 2.3.6 shows that g has at most three totally ramified critical values, hence the same is true of h.

If h has two totally ramified critical values then h is equal to cx^n for some $c \in K^\times$ and $n \in \mathbb{Z}\backslash\{0\}$. Maps of the form cx^n (as above) are automatically Galois, so this shows that f and g are both Galois, hence $N_f = N_g$ and we may take $\rho = \mathrm{id}$ in this case. Then $\mu f\mu^{-1}$ and h have the same étale closure, namely $\mathbb{G}_m$, and they are both affine morphisms. This proves the theorem when h has two totally ramified critical values.

Suppose h has a single totally ramified critical value. Then this must also be the case with $\mu f \mu^{-1}$, and so $\mu f \mu^{-1}$ must be $\pm T_m$ for some positive integer m. As h and $\pm T_m$ each have a single totally ramified critical value at infinity, and h is itself conjugate to a dynamically affine map, h is equal to a polynomial of the form $\mu \circ (\pm T_n) \circ \mu^{-1}$ for some $\mu(x) = ax + b$, $a \in K^\times$, $b \in K$, and positive integer n. We make use of the description of ramification of Chebyshev polynomials in §2.2.1.1 to determine the possibilities for μ. Since the critical values of $\pm T_m$ are at $\{\pm 2\}$, and these are not totally ramified critical values, the critical values of h are also at $\{\pm 2\}$. Then μ must stabilize the set $\{\pm 2\}$ since $\{\pm 2\}$ is the set of critical values for $\pm T_n$. There are two possibilities depending on whether μ acts trivially on this set or not: we have either $\mu = x$ or $-x$. In either case, we see that $\mu f \mu^{-1} = \pm T_n$ and $\mu g \mu^{-1} = \pm T_m$. All Chebyshev maps of degree > 2 have the same Galois closure (this was shown in the proof of Proposition 2.3.5), so this shows that f and g have the same Galois closure if h has a single totally ramified critical value. Hence $N_f = N_g$ and we may take $\rho = \mathrm{id}$ in this case also. By Theorem 2.3.3 both f and g lift to morphisms ψ and φ of N_f. Up to a simultaneous conjugation, ψ and φ are both power maps, which shows that their étale loci are both equal to $\mathbb{G}_m$. This proves the theorem when h has a single totally ramified critical value.

If h has no totally ramified critical values then f and g are Lattès maps by the classification of dynamically affine rational functions (Theorem 2.2.3). We have shown that any Lattès map of degree > 12 has a reduced realization given by its Galois closure (Proposition 2.3.6):

$$\begin{array}{ccc} N_f & \longrightarrow & N_f \\ {\scriptstyle p_f}\downarrow & & \downarrow{\scriptstyle p_f} \\ \mathbb{P}^1 & \xrightarrow{f} & \mathbb{P}^1. \end{array}$$

As N_f has genus one, the map $N_f \to N_f$ is unramified by the Riemann–Hurwitz

formula. Recall that if P is a point in the Galois closure N_f, then under the composite map fp_f it has ramification index $\varepsilon_f(Q)$ where $Q = f(p_f(P))$. It follows from commutativity of the diagram that $\varepsilon_{p_f}(Q) = \varepsilon_f(Q)$ for any $Q \in \mathbb{P}^1$. The same consideration for g now shows that

$$\varepsilon_{p_f}(Q) = \varepsilon_f(Q) = \varepsilon_g(Q) = \varepsilon_{p_g}(Q).$$

As p_f and p_g are both reduced, comparing ramification at any totally ramified critical value shows that p_f and p_g have the same degree, and it follows that they have the same branching divisors also. Applying Lemma 2.2.2 obtains an isomorphism $\rho\colon N_f \xrightarrow{\sim} N_g$ such that $p_f = p_g\rho$, as required. Since f and g are dynamically affine, Theorem 2.3.3 shows that such maps lift to endomorphisms ψ and φ of their Galois closure. The ramification loci of ψ and φ are empty, so they are equal, and any endomorphism of degree > 1 of an elliptic curve is automatically an affine morphism, [Sil09, III.4.8]. This concludes the proof of Theorem 2.4.2. □

2.5 Proofs of Main Theorems

In this section we prove the main theorems of this chapter: Theorem 1.2.1 and Theorem 1.2.3.

2.5.1 Proof of Theorem 1.2.3

We recall the statement of Theorem 1.2.3.

Theorem (Theorem 1.2.3)**.** *Let C be a curve of positive genus and let f and g be endomorphisms of C of degree greater than one. Then there exist orbits of f and g with infinite intersection if and only if f and g have a common iterate.*

For the proof we will need a theorem of Lang.

Theorem 2.5.1 ([Lan60], p. 320)**.** *Let Γ be a subgroup of finite type of $K^\times$. Then the curve $ax + by = 1$ with $a, b \in K$ and $ab \neq 0$ has only a finite number of points with $x, y \in \Gamma$.*

Proof of Theorem 1.2.3. The Riemann–Hurwitz formula for f shows that

$$2\mathfrak{g}_C - 2 = (\deg f)(2\mathfrak{g}_C - 2) + |\mathcal{R}_f|.$$

If $\mathfrak{g}_C > 1$ then both sides are positive, which can only occur if $|\mathcal{R}_f| = 0$ and $\deg f = 1$. As we have assumed f has degree > 1, this cannot occur. Hence we are reduced to the case when $\mathfrak{g}_C = 1$.

Because f is not a degree one map it has a fixed point $\mathcal{O}$. Let E be the elliptic curve $(C, \mathcal{O})$ and let $\operatorname{End} E$ denote the set of endomorphisms of C to itself which send $\mathcal{O}$ to $\mathcal{O}$; $\operatorname{End} E$ is a commutative ring. We have that $f(X) = r(X)$ and $g(X) = s(X)+Q$ for some $r, s \in \operatorname{End} E$ and $Q \in E$. Because $\deg s = \deg g$ the endomorphism $s - 1$ is nonconstant and thus surjective; suppose $(s - 1)(R) = Q$. Then for any positive integer b we have

$$g^{\circ b}(X) = s^b(X) + (1 + s + \cdots + s^{b-1})(Q) = s^b(X) + (s^b - 1)(R).$$

Thus for any positive integers a and b, the condition that $f^{\circ a} = g^{\circ b}$ is equivalent to the conditions that $r^a = s^b$ and $(s^b - 1)(R) = \mathcal{O}$. Pick orbits of f and g having infinite intersection, and let P be any point in the intersection; then the orbits $\mathcal{O}_f(P)$ and $\mathcal{O}_g(P)$ also have infinite intersection, so there are infinitely many pairs (a, b) of positive integers such that $r^a(P) = s^b(P) + (s^b - 1)(R)$. Let (c, d) be another such pair of positive integers. Then

$$(r^a - s^b)(P) = (s^b - 1)(R) \tag{2.15}$$

and

$$(r^c - s^d)(P) = (s^d - 1)(R). \tag{2.16}$$

Apply $s^d - 1$ to (2.15) and $s^b - 1$ to (2.16) to obtain $(s^d - 1)(r^a - s^b)(P) = (s^b - 1)(r^c - s^d)(P)$. As P cannot be a torsion point we must have $(s^d - 1)(r^a - s^b) = (s^b - 1)(r^c - s^d)$, or equivalently

$$(1 - s^b)r^a + (r^c - 1)s^b = r^c - s^d. \tag{2.17}$$

Embed $\operatorname{End} E$ as a subring of $\mathbb{C}$ and let $R \in \mathbb{C}$ (resp. $S \in \mathbb{C}$) correspond to r (resp. s). We see from (2.17) that if $R^c \neq S^d$ then the equation

$$\left(\frac{1 - S^b}{R^c - S^d}\right) X + \left(\frac{R^c - 1}{R^c - S^d}\right) Y = 1$$

has infinitely many solutions in the multiplicative subgroup $\Gamma \subset \mathbb{C}^\times$ generated by R and S, but this is a contradiction by the theorem of Lang (Theorem 2.5.1). We see that $R^c = S^d$, which implies that $r^c = s^d$. By (2.16) one has $(s^d - 1)(R) = \mathcal{O}$ so that $f^c = g^d$. This concludes the proof of Theorem 1.2.3. □

2.5.2 Lifting Lemma

Let f and g be endomorphisms of degree > 1 of a curve C and let $\pi\colon D \to C$ be a (generically) Galois morphism. Recall we say that f lifts along π to an endomorphism F of D if $f \circ (f^*\pi)$ is (generically) Galois, and the diagram

$$\begin{array}{ccc} D & \xrightarrow{F} & D \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\pi} \\ C & \xrightarrow{f} & C \end{array}$$

is cartesian, i.e., that the smooth fiber product $C \widetilde{\times}_{f,\pi} D$ is isomorphic to D over C.

Suppose that f and g *both* lift along π to endomorphisms F and G of D. We now prove a result which lets us lift orbits of f and g with infinite intersection to orbits

of F and G with infinite intersection. This lets us reduce the proof of Theorem 2.5.1 to Theorem 1.2.3, once we have shown the existence of a lifting for f and g.

Lemma 2.5.1. *Suppose that f and g both lift along π to endomorphisms F and G of D. If there are $c, d \in C$ such that $\mathcal{O}_f(c) \cap \mathcal{O}_g(d)$ is infinite then there exist $a \in \pi^{-1}(c)$ and $b \in \pi^{-1}(d)$ such that $\mathcal{O}_F(a) \cap \mathcal{O}_G(b)$ is infinite.*

Proof. Let $S \subset \mathbb{N} \times \mathbb{N}$ be an infinite subset such that $f^{\circ r}(c) = g^{\circ s}(d)$ for any $(r, s) \in S$. Choose $a \in \pi^{-1}(c)$ and $b \in \pi^{-1}(d)$ arbitrarily. As the orbit of c contains an infinite set, it is itself infinite, so it must be disjoint from the set of preperiodic points for f in C. The same consideration holds for $\mathcal{O}_g(d)$, and also $\mathcal{O}_F(a)$ and $\mathcal{O}_G(b)$ since π maps F-preperiodic points into f-preperiodic points, and likewise for G.

For any $(r, s) \in S$ we have that $\pi F^{\circ r}(a) = \pi G^{\circ s}(b)$ so that $F^{\circ r}(a) \cdot u = G^{\circ s}(b)$ for some $u \in \operatorname{Gal} \pi$ (depending on r and s). Let $S_u := \{(r, s) \in S : F^{\circ r}(a) \cdot u = G^{\circ s}(b)\}$ so that $S = \bigcup_{u \in \operatorname{Gal} \pi} S_u$; as S is infinite one of the S_u must be infinite also.

We claim that F is $\operatorname{Gal} \pi$-equivariant. Once we have shown this we are done, since then $F^{\circ r}(a) \cdot u = F^{\circ r}(a \cdot u)$, showing that $F^{\circ r}(a \cdot u) = G^{\circ s}(b)$ for any $(r, s) \in S_u$, which implies that $\mathcal{O}_F(a \cdot u) \cap \mathcal{O}_G(b)$ is infinite since neither of these orbits contain any preperiodic points.

Let L denote the function field of C over K, and let L^a be an algebraic closure of L. From the assumption that f lifts along π, we have a cartesian diagram:

$$\begin{array}{ccc} D & \xrightarrow{F} & D \\ {\scriptstyle \pi}\downarrow & & \downarrow{\scriptstyle \pi} \\ C & \xrightarrow{f} & C. \end{array}$$

There are unique subfields L', M, M' of L^a corresponding to the diagram above, say

$$\begin{array}{ccc} M' & \longleftarrow & M \\ \uparrow & & \uparrow \\ L' & \longleftarrow & L, \end{array}$$

where arrows denote inclusions. Let us choose a generator α for M over L, so that $M = L(\alpha)$. Since the diagram is cartesian, M and L' are linearly disjoint over L, and M' is equal to the compositum of M and L' inside of L^a. It follows that $M' = L' \otimes_L M = L' \otimes_L L(\alpha) = L'(\alpha)$.

As M/L is Galois, linear disjointness of M and L' implies that $M \cap L' = L$. Therefore restriction, $\sigma \mapsto \sigma|_M$, gives an isomorphism between the Galois groups of M'/L' and M/L (cf. e.g., [Lan02, (1.12)]). This means that for any automorphism $\sigma \in \mathrm{Gal}(M'/L')$, we have a commuting diagram

$$\begin{array}{ccc} M' & \longleftarrow & M \\ {\scriptstyle\sigma}\downarrow & & \downarrow{\scriptstyle\sigma|M} \\ M' & \longleftarrow & M. \end{array} \tag{2.18}$$

There is a canonical isomorphism between $\rho\colon \mathrm{Gal}\,\pi \to \mathrm{Gal}(M/L)$ given by

$$\rho(u)\psi = \psi \circ u^{-1} \quad \text{for } \psi \in K_D.$$

It follows that for any $\sigma \in \mathrm{Gal}(M/L)$ there is a unique $u \in \mathrm{Gal}\,\pi$ such that $\sigma\alpha = \alpha \circ u^{-1}$. There is an isomorphism $\rho'\colon \mathrm{Gal}\,\pi \to \mathrm{Gal}(M'/L')$ which is defined in the same way as ρ, and since restriction $\sigma \mapsto \sigma|_M$ defines an isomorphism between $\mathrm{Gal}(M'/L')$ and $\mathrm{Gal}(M/L)$, it follows that for any $\sigma \in \mathrm{Gal}(M'/L')$ such that $\rho(u) = \sigma$, we have that

$$(\sigma|_M)\alpha = \alpha \circ u^{-1} = \rho(u)\alpha.$$

From this it follows that the diagram of curves which corresponds to (2.18) is given by

$$\begin{array}{ccc} D & \xrightarrow{F} & D \\ {\scriptstyle u}\uparrow & & \uparrow{\scriptstyle u} \\ D & \xrightarrow{F} & D. \end{array}$$

This shows that F is $\mathrm{Gal}\,\pi$-equivariant as was to be shown. □

2.5.3 Proof of Theorem 1.2.1

We recall the statement of Theorem 1.2.1.

Theorem (Theorem 1.2.1)**.** *For rational functions f, g of coprime degree and degree greater than one, the orbits $\mathcal{O}_f(c)$ and $\mathcal{O}_g(d)$ have finite intersection for any $c, d \in K$.*

Proof. Let $K(t)$ be the function field of $\mathbb{P}^1$. Let $K(x), K(y) \subset K(t)^a$ be the field extensions of $K(t)$ corresponding to the morphisms f and g, i.e., $x, y \in K(t)^a$ and we have that $f(x) = g(y) = t$. As the degrees of f and g are coprime, the field extensions $K(x)/K(t)$ and $K(y)/K(t)$ have coprime degree, hence are linearly disjoint. It follows that $K(x) \otimes_{K(t)} K(y)$ is a field, and so the smooth fiber product $\mathbb{P}^1 \widetilde{\times}_{f,g} \mathbb{P}^1$ is a curve (irreducible). The same consideration holds for $\mathbb{P}^1 \widetilde{\times}_{f^{\circ r}, g^{\circ s}} \mathbb{P}^1$, and it follows that f and g form an irreducible pair. Theorem 1.2.1 therefore follows from Theorem 2.5.1. □

Theorem 2.5.1. *Suppose that f and g form an irreducible pair. Then the orbits $\mathcal{O}_f(c)$ and $\mathcal{O}_g(d)$ have finite intersection for any elements c, d of K.*

Proof. Suppose to the contrary that the intersection $\mathcal{O}_f(c) \cap \mathcal{O}_g(d)$ is infinite for some $c, d \in K$. Let $K_0 \subset K$ be the field generated over $\mathbb{Q}$ by c, d, and the coefficients of f and g. Consider the (possibly singular) curve $\mathcal{C}'$ over K_0 defined by the numerator of the bivariate rational function $f(x) - g(y)$ and let $\pi \colon \mathcal{C} \to \mathcal{C}'$ be its normalization.[7] By hypothesis there is an infinite subset $S \subset \mathbb{N}^{>0} \times \mathbb{N}^{>0}$ such that $f^{\circ m}(c) = g^{\circ n}(d)$ for any $(m, n) \in S$. This establishes a map of sets

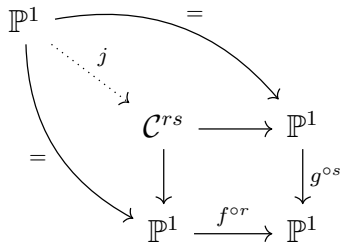
$$S \to \mathcal{C}'(K_0)$$
$$(m, n) \mapsto (f^{\circ(m-1)}(c), g^{\circ(n-1)}(d)).$$

[7] i.e., $\mathcal{C}$ is smooth and π induces an isomorphism between the function fields of $\mathcal{C}$ and $\mathcal{C}'$.

This map is injective as the orbits of c and d are necessarily infinite, so $\mathcal{C}'$ has infinitely many points defined over K_0. As π is a birational isomorphism, it induces an isomorphism on a nonempty open subset, i.e., for some nonempty open subsets $U \subset \mathcal{C}$ and $U' \subset \mathcal{C}'$ we have $\pi|_U \colon U \xrightarrow{\sim} U'$. As the complement of U' in $\mathcal{C}'$ is finite, we see that $U'(K_0)$ is infinite, and so $U(K_0)$ is infinite also. We conclude that $\mathcal{C}$ has infinitely many points defined over K_0. By applying Faltings's theorem as generalized to finitely generated fields, [Fal84, Theorem 3], we see that the genus of $\mathcal{C}$ is zero or one.

For any positive integers r and s, we have that $\mathcal{O}_f(c) = \cup_{i=0}^{r-1} \mathcal{O}_{f^{\circ r}}(f^{\circ i}(c))$. Therefore if $\mathcal{O}_f(c)$ and $\mathcal{O}_g(d)$ have infinite intersection, then for some positive integers i and j, the orbits $\mathcal{O}_{f^{\circ r}}(f^{\circ i}(c))$ and $\mathcal{O}_{g^{\circ s}}(g^{\circ j}(d))$ have infinite intersection also. The pair $(f^{\circ r}, g^{\circ s})$ is clearly still irreducible, so we may again apply Faltings's theorem just as above to conclude that the genus of $\mathcal{C}^{rs}$ is zero or one for any positive integers r and s.

By Theorem 2.4.2, there exist a curve N and a Galois morphism $\pi \colon N \to \mathbb{P}^1$ such that f and g both lift along π to endomorphisms F and G of N. By Lemma 2.5.1 there are orbits of F and G with infinite intersection. If $\mathfrak{g}_N > 0$ then Theorem 1.2.3 implies that F and G have a common iterate. It follows that f and g then have a common iterate also, i.e., $f^{\circ r} = g^{\circ s}$ for some positive integers r and s. By the universal property of the smooth fiber product we obtain a morphism $j \colon \mathbb{P}^1 \to \mathcal{C}^{rs}$ such that the diagram



commutes. As $\mathcal{C}^{rs}$ is irreducible, the degree of the composite $\mathbb{P}^1 \xrightarrow{j} \mathcal{C}^{rs} \to \mathbb{P}^1$ is $(\deg j)(\deg g)^s$ but this cannot equal one, so we have obtained a contradiction in the case that $\mathfrak{g}_N = 1$.

Suppose that $\mathfrak{g}_N = 0$. By Theorem 2.4.2, f and g lift along $\pi\colon N \to \mathbb{P}^1$ to affine morphisms ψ and φ of an open group variety $G \subset N$. As ψ and φ both have degree > 1 (since f and g were assumed to have degree > 1), G cannot be isomorphic to the additive group, so it is isomorphic to $\mathbb{G}_m$ (there are no twists since K is algebraically closed). Therefore ψ and φ are simultaneously conjugate to affine morphisms of $\mathbb{G}_m$. It follows that $\psi^{\circ 2}$ and $\varphi^{\circ 2}$ are simultaneously conjugate to polynomials (note that $(cx^n)^{\circ 2}$ is a polynomial). By Lemma 2.5.1 there are orbits of ψ and φ with infinite intersection, and it follows that $\psi^{\circ 2}$ and $\varphi^{\circ 2}$ also have orbits with infinite intersection. It follows that $\psi^{\circ 2}$ and $\varphi^{\circ 2}$ are simultaneously conjugate to polynomials S and T which have orbits with infinite intersection. We have thus reduced the problem to the main theorem of a paper of Ghioca-Tucker-Zieve, [GTZ12], which asserts that S and T have a common iterate. It follows that $\psi^{\circ 2}$ and $\varphi^{\circ 2}$ have a common iterate, hence that ψ and φ have a common iterate, and finally that f and g have a common iterate. This concludes the proof of Theorem 2.5.1. $\square$

CHAPTER III

Polynomials with Integral Divided Differences

This chapter studies functions with integral divided differences. We prove the following characterization of polynomial functions whose mth divided difference is integer-valued. Let K be an algebraic number field of degree d with ring of integers $\mathcal{O}$.

Theorem (Theorem 1.3.1)**.** *Let $s\colon \mathbb{N} \to K$. Suppose that*[1]

(i) $\delta_m s$ is $\mathcal{O}$-valued, and

(ii) for each embedding $\sigma\colon K \to \mathbb{C}$, $|\sigma s(n)| \ll \theta_\sigma^n$ for some positive number θ_σ and

$$\prod_{\sigma : K \to \mathbb{C}} (1 + \theta_\sigma) < e^{d\left(1+\frac{1}{2}+\cdots+\frac{1}{m}\right)}.$$

Then $s(n)$ is a polynomial in n.

Along the way to proving Theorem 1.3.1 we prove two other results, one local and one global. Both results concern the following elementary number-theoretic function:

$$\tau_{m,p}(n) := \max_{\substack{S \subset \{1,\ldots,n\}, \\ \#S = m}} w_p\Big\{ \prod_{k \in S} k \Big\}.$$

In other words, $\tau_{m,p}(n)$ is the maximal p-adic valuation of a product of m distinct positive integers that are $\leq n$. We will prove the following new results concerning $\delta_m s$ and $\tau_{m,p}$.

[1] We write $r_n \ll q_n$ to mean there is a positive constant C such that $|r_n| \leq C|q_n|$ for all $n \geq 0$.

Theorem 3.0.1. *Let $s\colon \mathbb{N} \to \mathbb{C}_p$. Let $\|\delta_m s\|_p$ denote the supremum of $\delta_m s$ over $\mathbb{N}^{m+1} \setminus \cup_{i<j} \{n_i = n_j\}$. Let $c(n)$ denote the nth finite difference of s (cf. (3.3)). Then*

$$\|\delta_m s\|_p = \sup_{n \geq m} |c(n)|_p p^{\tau_{m,p}(n)}. \tag{3.1}$$

Theorem 3.0.2.

$$\prod_{p \text{ prime}} p^{\tau_{m,p}(n)} = \exp\left\{\left(1 + \tfrac{1}{2} + \cdots + \tfrac{1}{m}\right)n + O(n \exp\{-\alpha(\log n)^{1/2}\} \log n)\right\}$$

for some positive constant α.

These theorems are proved in §3.1 and §3.2.

Notation: A place v of K is an equivalence class of isometric embeddings $\sigma\colon K \to \mathbb{C}_p$ with $p \in \{2, 3, 5, \ldots, \infty\}$ where $\mathbb{C}_p$ is the completion of an algebraic closure of the p-adic field $\mathbb{Q}_p$ and $\mathbb{C}_\infty := \mathbb{C}$. M_K denotes the set of all places of K. d_v denotes the local degree at v. $|x|_v := |\sigma(x)|_p$ is the norm corresponding to a place v and a representative embedding σ, and w_p is the (additive) p-adic valuation. $[\cdot]$ denotes the floor function and $H_m = 1 + \frac{1}{2} + \cdots + \frac{1}{m}$ is the mth harmonic number, $H_0 := 0$. $\vartheta(n) := \sum_{p \leq n} \log p$ is the Chebyshev function and $\pi(n)$ equals the number of rational primes $\leq n$.

3.1 Divided Differences

In this section p always denotes a finite rational prime. The results of this section are in the local setting so we often omit the subscript p from norms for brevity. The goal of this section is to prove Theorem 3.0.1. In §3.3 we will combine the local estimates (3.1) using the product formula to obtain a condition for the Archimedean growth of the finite differences of a sequence whose mth divided difference is integral.

Let us briefly recall some necessary background from difference calculus. Let $s\colon \mathbb{N} \to K$ be a sequence and let m be a non-negative integer. The mth divided

difference of s is the function $\delta_m s\colon X_m \to K$ given by

$$\delta_m s(n_0, \ldots, n_m) := \sum_{i=0}^{m} \Big\{ \prod_{j \neq i} (n_i - n_j)^{-1} \Big\} s(n_i) \tag{3.2}$$

where

$$X_m := \{(n_0, \ldots, n_m) \in \mathbb{N}^{m+1} : n_i \text{ all distinct}\}.$$

We mention without proof that the sequence s can be reconstructed using values of its divided differences by means of Newton's interpolation formula (cf. [MT51], §1). The nth finite difference of s is defined by[2]

$$c(n) := \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} s(k). \tag{3.3}$$

Recall the following classical result of difference calculus. We only sketch a proof.

Lemma 3.1.1. *Let $s\colon \mathbb{N} \to K$ be a sequence and let $c\colon \mathbb{N} \to K$ be its sequence of finite differences. Then s is polynomial if and only if c is eventually zero.*

Proof. Let S be the forward shift operator on sequences defined by $(Ss)(n) := s(n+1)$ for all non-negative integers n. Then for any non-negative integer ℓ,

$$\{(S - \mathrm{id})^n s\}(\ell) = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} s(\ell + k),$$

and in particular, $\{(S - \mathrm{id})^n s\}(0) = c(n)$. We have that $(S - \mathrm{id})(n^d) = dn^{d-1} + O(n^{d-2})$, and so the restriction of $S - \mathrm{id}$ to the space of polynomial sequences is nilpotent. This shows that c is eventually zero if s is polynomial.

Conversely, assume that c is eventually zero. It is easy to verify that the inverse relation of (3.3) is given by

$$s(n) = \sum_{k=0}^{n} \binom{n}{k} c(k), \tag{3.4}$$

[2] Strictly speaking, this is the sequence obtained by evaluating the finite differences of s at zero. We will not have use for the usual finite differences, so we refer to the sequence defined by (3.3) as the finite differences of s for the sake of brevity.

and this shows that s is given by a polynomial of degree $\leq N$ if $c(n) = 0$ for $n > N$. □

Let m be a non-negative integer and p a prime. Let E be a non-Archimedean Banach space over $\mathbb{C}_p$ and let $\ell_p^\infty(\mathbb{N}^{m+1}, E)$ denote the Banach space of bounded functions $F\colon \mathbb{N}^{m+1} \to E$ equipped with the norm given by

$$\|F\| := \sup_{\underline{n} \in \mathbb{N}^{m+1}} \|F(\underline{n})\|.$$

Define $\underline{x} = (x_0, \ldots, x_m)$, $\underline{j} = (j_0, \ldots, j_m)$, and

$$\binom{\underline{x}}{\underline{j}} := \binom{x_0}{j_0}\binom{x_1}{j_1}\cdots\binom{x_m}{j_m} \in \mathbb{Z}[x_0, \ldots, x_m].$$

The following proposition generalizes Mahler's theorem to bounded functions (cf. (1.5)).

Proposition 3.1.2. *Let $F\colon \mathbb{N}^{m+1} \to E$. There exist unique $C_{\underline{j}} \in E$ such that for all $\underline{n} \in \mathbb{N}^{m+1}$*

$$F(\underline{n}) = \sum_{\underline{j} \in \mathbb{N}^{m+1}} C_{\underline{j}} \binom{\underline{n}}{\underline{j}} \quad \text{(finite sum)}. \tag{3.5}$$

The Mahler coefficients C have the properties that

(i) F is bounded if and only if C is bounded,

(ii) the mapping $F \mapsto C$ is a self-isometry of $\ell_p^\infty(\mathbb{N}^{m+1}, E)$, and

(iii) F extends to a continuous function $\mathbb{Z}_p^{m+1} \to E$ if and only if C goes to zero.[3]

Note that the proposition does not imply that the $\binom{\underline{n}}{\underline{j}}$ form an orthonormal basis for $\ell_p^\infty(\mathbb{N}^{m+1}, E)$ as the sum $\sum_{\underline{j} \in \mathbb{N}^{m+1}} C_{\underline{j}} \binom{\underline{x}}{\underline{j}}$ does not necessarily converge.

[3] i.e., $\lim_{N \to \infty} \sup_{j_0 + \cdots + j_m > N} \left\|C_{\underline{j}}\right\| = 0$.

Proof. We proceed by induction on m. If $m = 0$ then we take C_n to be the nth finite difference of F given by $C_n := \sum_{k=0}^{n} \binom{n}{k}(-1)^{n-k}F(k)$. The inverse relation is given by (3.4) which proves existence for (3.5), and uniqueness follows from bijectivity of this mapping. If F is bounded then C is bounded by the ultrametric inequality and vice versa. To see that $F \mapsto C$ is an isometry when F is bounded it suffices to observe that the relation (3.3) and its inverse (3.4) are both defined over $\mathbb{Z}$ and to apply the ultrametric inequality. The third property is Mahler's theorem [Mah58].

Now suppose m is a positive integer. Fix a natural number n_m and define the function

$$G_{n_m} \colon \mathbb{N}^m \to E$$

$$(n_0, \ldots, n_{m-1}) \mapsto F(n_0, \ldots, n_{m-1}, n_m).$$

By the inductive hypothesis there are uniquely defined coefficients $D_{\underline{i}} = D_{\underline{i}}(n_m) \in E$ for $\underline{i} \in \mathbb{N}^m$ such that for all $(n_0, \ldots, n_{m-1}) \in \mathbb{N}^m$

$$F(n_0, \ldots, n_{m-1}, n_m) = \sum_{\underline{i} \in \mathbb{N}^m} D_{\underline{i}}(n_m) \binom{n_0}{i_0} \binom{n_1}{i_1} \cdots \binom{n_{m-1}}{i_{m-1}}.$$

We may also express $D_{\underline{i}}(n_m)$ as a function of n_m in terms of its finite differences, $c_n(\underline{i}) \in E$, to obtain

$$F(n_0, \ldots, n_{m-1}, n_m) = \sum_{\underline{i} \in \mathbb{N}^m} \sum_{k \geq 0} c_k(\underline{i}) \binom{n_m}{k} \binom{n_0}{i_0} \binom{n_1}{i_1} \cdots \binom{n_{m-1}}{i_{m-1}}. \tag{3.6}$$

Setting $C_{j_0,\ldots,j_m} := c_{j_m}(j_0, \ldots, j_{m-1})$ proves (3.5), and uniqueness follows from that of $c_k(\underline{i})$ and $D_{\underline{i}}(n_m)$.

If F is bounded then G_{n_m} is bounded for all $n_m \in \mathbb{N}$, and by the inductive hypothesis $\|G_{n_m}\| = \|D(n_m)\|$. Similarly, $\sup_{n_m \in \mathbb{N}} |D_{\underline{i}}(n_m)| = \sup_{n_m \in \mathbb{N}} |c_k(\underline{i})|$ for all $\underline{i} \in \mathbb{N}^m$. This proves that C is bounded. If C is bounded then, by means of the ultrametric inequality, (3.5) shows that F is bounded.

If F is bounded we observe that

$$\begin{aligned}\|F\| = \sup_{n_m\in\mathbb{N}}\sup_{\underline{i}\in\mathbb{N}^m}|F(n_0,\ldots,n_{m-1},n_m)| &= \sup_{n_m\in\mathbb{N}}\|G_{n_m}\| = \sup_{n_m\in\mathbb{N}}\|D(n_m)\| \\ &= \sup_{\underline{i}\in\mathbb{N}^m}\sup_{n_m\in\mathbb{N}}|D_{\underline{i}}(n_m)| \\ &= \sup_{\underline{i}\in\mathbb{N}^m}\sup_{n_m\in\mathbb{N}}|c_k(\underline{i})| = \|C\|\,.\end{aligned}$$

This proves that $F \mapsto C$ is a self-isometry of $\ell_p^\infty(\mathbb{N}^{m+1}, E)$. If F extends to a continuous function $\mathbb{Z}_p^{m+1} \to E$ then by Corollaire 1, §2.7, [Ami64], the coefficients C go to zero. □

We can now prove that the p-adic supremum of higher divided differences is given by the p-adic supremum of the finite differences relative to $p^{-\tau_{m,p}(n)}$. As before, let

$$\tau_{m,p}(n) := \max_{0<i_1<\cdots<i_m\le n}(w_p(i_1)+\cdots+w_p(i_m)).$$

We recall that the finite differences of a function s are defined by

$$c(n) := \sum_{k=0}^{n}\binom{n}{k}(-1)^{n-k}s(k) \quad (n\in\mathbb{N}).$$

Theorem (Theorem 3.0.1)**.** *Let $s\colon \mathbb{N}\to\mathbb{C}_p$. Then*

$$\|\delta_m s\|_p = \sup_{n\ge m}|c(n)|_p p^{\tau_{m,p}(n)}. \tag{3.7}$$

In particular, if $\delta_m s(\underline{n})$ is integral for all $\underline{n}\in X_m$ then $|c(n)|_p \le p^{-\tau_{m,p}(n)}$ for all $n\ge m$.

We will show that (3.7) holds even if $\|\delta_m s\|_p$ is infinite.

Remark 1. In view of the theorem, and the fact that $\tau_{m,p}(n)$ is monotonically increasing in m, we have the bound ($m\ge 1$):

$$\|\delta_{m-1}s\|_p \le \max\left\{|c(m-1)|_p|(m-1)!|_p^{-1}, \|\delta_m s\|_p\right\}.$$

This shows that if $\delta_m s$ is $\mathbb{Z}$-valued then there is a positive integer N such that $N\delta_k s$ is $\mathbb{Z}$-valued for all $k \leq m$.

Proof. The claim is clearly true when $m = 0$ so suppose $m \geq 1$. Let $(n_0, \ldots, n_m) \in X_m$ and set $\underline{\ell} = (n_{i_0}, n_{i_1}, \ldots, n_{i_m})$ where the indices have been reindexed so that $n_{i_0} > n_{i_1} > \cdots > n_{i_m}$. Now set $\underline{\ell} = (x_1 + \cdots + x_m + y, x_1 + \cdots + x_{m-1} + y, \ldots, x_1 + x_2 + y, x_1 + y, y)$ where the $x_1, \ldots, x_m$ are positive integers as the integers $n_0, \ldots, n_m$ are all distinct. We make use of the formula for the Mahler series for the mth divided difference due to Schikhof, [Sch06], Theorem 54.1:

$$\delta_m s(\underline{n}) = \sum_{j \geq 0} \sum_{k_1, \ldots, k_m \geq 1} \frac{c(j + k_1 + \cdots + k_m)}{k_m(k_m + k_{m-1}) \cdots (k_m + \cdots + k_1)} \binom{y}{j} \prod_{i=1}^{m} \binom{x_i - 1}{k_i - 1}. \tag{3.8}$$

Note that (3.8) is always a finite sum.

The dependence of the $x_1, x_2, \ldots, x_m, y$ on the entries of $\underline{\ell}$ is clearly invertible, and as the $x_1, x_2, \ldots, x_m$ vary over all positive integers, and y varies over all non-negative integers, the corresponding $\underline{\ell}$ will vary over all strictly decreasing tuples in X_m. As $\delta_m s$ is a symmetric function, the right-hand side of (3.8) will therefore achieve all values of $\delta_m s$ as $x_1, \ldots, x_m$ vary over all positive integers and y varies over all non-negative integers. Now setting $(x_1, x_2, \ldots, x_m, y) = (a_1 + 1, a_2 + 1, \ldots, a_m + 1, a_{m+1})$, $\underline{a} := (a_1, a_2, \ldots, a_m, a_{m+1})$, and letting $\underline{n}_{\underline{a}}$ be the corresponding element of X_m, we see the right-hand side of (3.8) gives a well-defined function

$$F \colon \mathbb{N}^{m+1} \to \mathbb{C}_p$$

$$(a_1, \ldots, a_{m+1}) \mapsto \delta_m s(\underline{n}_{\underline{a}}),$$

and that moreover $\|F\|_p = \|\delta_m s\|_p$.

By reindexing with $i_1 = k_m$, $i_2 = k_m + k_{m-1}$, $\ldots$, $i_m = k_m + k_{m-1} + \cdots + k_1$ and

$n = j + k_1 + \cdots + k_m$, we get that

$$\sup_{j\geq 0, k_1,\ldots,k_m\geq 1} \left| \frac{c(j+k_1+\ldots+k_m)}{k_m(k_m+k_{m-1})\cdots(k_m+\cdots+k_1)} \right|_p \tag{3.9}$$

$$= \sup_{0<i_1<\cdots<i_m\leq n} \left| \frac{c(n)}{i_1 i_2 \cdots i_m} \right|_p = \sup_{n\geq m} |c(n)|_p p^{\tau_{m,p}(n)}$$

where $\tau_{m,p}(n) := \max_{0<i_1<\cdots<i_m\leq n}(w_p(i_1)+\cdots+w_p(i_m))$.

If $\|F\|_p$ is infinite, then (3.8) and (3.9) show that

$$\sup_{j\geq 0, k_1,\ldots,k_m\geq 1} \left| \frac{c(j+k_1+\ldots+k_m)}{k_m(k_m+k_{m-1})\cdots(k_m+\cdots+k_1)} \right|_p = \sup_{n\geq m} |c(n)|_p p^{\tau_{m,p}(n)} = \infty,$$

for if this were finite then the second part of Proposition 3.1.2 would imply that $F \in \ell_p^\infty(\mathbb{N}^{m+1})$.

If $\|F\|_p$ is finite, then it follows from (3.8), (3.9), and Proposition 3.1.2 that

$$\|F\|_p = \|\delta_m s(\underline{n})\|_p = \sup_{n\geq m} |c(n)|_p p^{\tau_{m,p}(n)}.$$

This concludes the proof. □

We have already remarked that congruence-preservation is equivalent to integrality of $\delta_1 s$ and now we offer a third interpretation. Integrality of $\delta_1 s$ implies that for all primes p and integers $m, n \in \mathbb{N}$,

$$|s(m) - s(n)|_p \leq |m-n|_p.$$

In other words, $\delta_1 s$ is $\mathbb{Z}$-valued if and only if s is simultaneously Lipschitz continuous with Lipschitz constant 1 for every p-adic metric on $\mathbb{N}$. It is natural to ask for a similar interpretation for the integrality of higher divided differences. The next proposition provides such an interpretation though we will not have use for it.

Proposition 3.1.3. *Let $s\colon \mathbb{N} \to \mathbb{C}_p$ and let m be a positive integer. Suppose that $\|\delta_m s\|_p \leq M$. Then s extends to an element f of $C^{m-1}(\mathbb{Z}_p, \mathbb{C}_p)$ and $f^{(m-1)}$ is Lipschitz continuous with Lipschitz constant $M|(m-1)!|_p$.*

Proof. By the recursive definition of divided differences (cf. [MT51], §1),

$$|\delta_{m-1}s(x_0,\cdots,x_{m-1}) - \delta_{m-1}s(x_1,\cdots,x_m)| \le M|x_0 - x_m| \tag{3.10}$$

for all $x = (x_0,\cdots,x_m) \in X_m$. Since $\delta_m s$ is a symmetric function, from (3.10) we obtain the inequalities

$$|\delta_{m-1}s(x_i;y_{ij}) - \delta_{m-1}s(x_j;y_{ij})| \le M|x_i - x_j| \tag{3.11}$$

where $0 \le i < j \le m$ and $y_{ij} := (x_0,\ldots,\widehat{x_i},\ldots,\widehat{x_j},\ldots,x_m) \in X_{m-2}$.

We equip $\mathbb{Z}_p^m$ with the metric given by

$$d_m(x,y) := \max_{1\le i\le m} |x_i - y_i|_p,$$

and will show that $\delta_{m-1}s$ is Lipschitz continuous for this metric on the dense subset $X_{m-1} \subset \mathbb{Z}_p^m$. By a limiting argument, it suffices to show the Lipschitz condition for $\delta_{m-1}s$ on elements $x = (x_1,\ldots,x_m)$ and $y = (y_1,\ldots,y_m)$ in X_{m-1} such that $\{x_1,\ldots,x_m\} \cap \{y_1,\ldots,y_m\} = \varnothing$. Let $z_0 = x$, $z_m = y$, and

$$z_i := (x_1, x_2, \ldots, x_{m-i}, y_{m-i+1}, \ldots, y_m) \quad \text{for} \quad 0 < i < m.$$

Then $d_m(z_i, z_{i+1}) = |x_{m-i} - y_{m-i}|_p$ for $0 \le i < m$. Because z_i and z_{i+1} differ by only one coordinate, and since we assumed none of the x_i coincided with any of the y_j, we may apply (3.11). We conclude that $|\delta_{m-1}s(z_i) - \delta_{m-1}s(z_{i+1})|_p \le M|x_{m-i} - y_{m-i}|_p = Md_m(z_i, z_{i+1})$ for $0 \le i < m$. Using

$$\delta_{m-1}s(z_0) - \delta_{m-1}s(z_m) = \sum_{i=0}^{m-1} (\delta_{m-1}s(z_i) - \delta_{m-1}s(z_{i+1}))$$

together with the ultrametric inequality shows that

$$\begin{aligned} |\delta_{m-1}s(x) - \delta_{m-1}s(y)|_p &\le M \max_{0\le i<m} |\delta_{m-1}s(z_i) - \delta_{m-1}s(z_{i+1})|_p \\ &\le M \max_{0\le i<m} |x_{m-i} - y_{m-i}|_p \\ &= Md_m(x,y). \end{aligned}$$

We have shown that $\delta_{m-1}s$ is Lipschitz continuous on a dense subset of $\mathbb{Z}_p^m$ so we may extend its domain from X_{m-1} to $\mathbb{Z}_p^m$ to obtain a Lipschitz continuous extension $\hat{\delta}_{m-1}s\colon \mathbb{Z}_p^m \to \mathbb{C}_p$ of $\delta_{m-1}s$. It follows that s extends to a $(m-1)$-times continuously differentiable function $f\colon \mathbb{Z}_p \to \mathbb{C}_p$ and that $f^{(m-1)}(a) = (m-1)!\hat{\delta}_{m-1}s(x_a)$ where $x_a := (a, a, \ldots, a)$ (cf. [Sch06], §29). By Lipschitz continuity of $\hat{\delta}_{m-1}s$ we get that

$$|f^{(m-1)}(a) - f^{(m-1)}(b)| \leq M|(m-1)!|_p d_m(x_a, y_b) = M|(m-1)!|_p|a-b|_p. \qquad \square$$

Corollary. *Let $s\colon \mathbb{N} \to \mathbb{Q}$ and suppose that $\delta_{m+1}s$ is $\mathbb{Z}$-valued. Then for every prime p, the function s extends to a p-adic m-times continuously differentiable function $f_p\colon \mathbb{Z}_p \to \mathbb{Q}_p$ and $f_p^{(m)}$ is Lipschitz continuous with constant $|m!|_p$.*

The corollary extends in the obvious way to functions valued in number fields. We now derive an explicit formula for $\tau_{m,p}(n)$ which will be needed in the next section. Recall that for $n \geq m$ we define

$$\tau_{m,p}(n) := \max_{0<i_1<\cdots<i_m\leq n}(w_p(i_1) + \cdots + w_p(i_m)).$$

Lemma 3.1.4. *Let m be a non-negative integer, p a prime $\geq m$, n an integer $\geq m$, and $a_p(n) := [np^{-[\log_p n]}]$. Then*

$$\tau_{m,p}(n) = \begin{cases} m[\log_p n] & \text{if } a_p(n) > m, \\ m[\log_p n] + a_p(n) - m & \text{if } a_p(n) \leq m. \end{cases} \tag{3.12}$$

The formula generally fails if $p < m$, e.g., $\tau_{p+1,p}(p^2) = p+1$ whereas (3.12) gives $p+2$.

Proof. If m is zero the formula clearly holds so suppose that m is positive. Let $t := [\log_p n]$, and suppose

$$n = a_0 + a_1 p + a_2 p^2 + \cdots + a_t p^t$$

where $0 \leq a_i < p$ for $0 \leq i \leq t$ and $a_t \neq 0$. We will calculate a set of integers $1 \leq i_1 \leq \cdots \leq i_m \leq n$ that realize the maximum p-adic valuation.

If $m < a_t$ then we take $i_m = a_t p^t$, $i_{m-1} = (a_t - 1)p^t$, $\ldots$, $i_1 = (a_t - m + 1)p^t$. Adding up the valuations we get that $\tau_{m,p}(n) = mt$. If $m \geq a_t$, then we can take $i_m = a_t p^t$, $i_{m-1} = (a_t - 1)p^t$, $\ldots$, $i_{m-a_t+1} = p^t$. Subsequently, we may take $i_{m-a_t} = p^t - p^{t-1}$, $i_{m-a_t-1} = p^t - 2p^{t-1}$, $\ldots$, $i_1 = p^t - (m - a_t)p^{t-1}$. As $p \geq m$ by hypothesis, $m - a_t \leq p - 1$ so that the valuation of $p^t - (m - a_t)p^{t-1}$ is precisely $t - 1$. Putting the valuations together we get that

$$\tau_{m,p}(n) = ta_t + (t-1)(m - a_t) = mt + a_t - m.$$

This finishes the proof of (3.12). □

3.2 Asymptotic Behavior of Certain Sums over Primes

The previous section established a local estimate for the finite differences c of an arbitrary sequence s with integral mth divided difference: for any $s\colon \mathbb{N} \to \mathbb{C}_p$ and any finite prime v of K,

$$\delta_m s \text{ } v\text{-integral} \implies |c(n)|_v \leq p^{-\tau_{m,p}(n)}.$$

To combine this local estimate over all primes we will need to calculate the asymptotic behavior of

$$\sum_{p \leq n} \tau_{m,p}(n) \log p.$$

The goal of this section is to prove that this is $nH_m + o(n)$ (Theorem 3.0.2). The standard bound for the Chebyshev function coming from the prime number theorem, $\vartheta(x) = x + o(x)$, or even $\vartheta(x) = x + O(\frac{x}{\log x})$, is not strong enough to establish the estimates needed for the proof. Instead we will employ the following useful estimate

due to Rosser and Schoenfeld, [RS62], (2.29):

$$\vartheta(x) = x + O(x \exp\{-\alpha(\log x)^{1/2}\}) \tag{3.13}$$

for some positive constant α.

First we prove a simple lemma. Let $[\,\cdot\,]\colon \mathbb{R} \to \mathbb{Z}$ denote the floor function.

Lemma 3.2.1.

$$\sum_{p \le n} [\log_p n] \log p = n + O(n \exp\{-\alpha(\log n)^{1/2}\})$$

for some positive constant α.

Proof. For any prime p in the sum we have that $r := [\log_p n]$ must be positive. We then have that

$$[\log_p n] = r \iff \frac{\log n}{r+1} < \log p \le \frac{\log n}{r} \iff n^{\frac{1}{r+1}} < p \le n^{\frac{1}{r}}.$$

Then

$$\begin{aligned} 0 \le \sum_{p \le n} [\log_p n] \log p &= \sum_{r=1}^{\infty} \sum_{n^{\frac{1}{r+1}} < p \le n^{\frac{1}{r}}} r \log p \\ &\le \sum_{\sqrt{n} < p \le n} \log p \\ &= \vartheta(n) - \vartheta(\sqrt{n}). \end{aligned}$$

The last term is $n + O(n \exp\{-\alpha(\log n)^{1/2}\})$ by (3.13). □

Let m be a non-negative integer, p a prime, n any integer $\ge m$. As before we set

$$\tau_{m,p}(n) := \max_{0 < i_1 < \cdots < i_m \le n} (w_p(i_1) + \cdots + w_p(i_m)).$$

Let H_m be the mth harmonic number and set $H_0 = 0$.

Theorem (Theorem 3.0.2)**.**

$$\sum_{p\le n}\tau_{m,p}(n)\log p = nH_m + O(n\exp\{-\alpha(\log n)^{1/2}\}\log n)$$

for some positive constant α*.*

Proof. This is clear if m is zero as $\tau_{0,p}(n)\equiv 0$, so we suppose m is positive. Suppose $n = c_0 + c_1 p + \cdots + c_t p^t$ where $0 \le c_i < p$ and $c_t \neq 0$. Let $a_p(n) := c_t$. When $n, p \ge m$, Lemma 3.1.4 gives the formula:

$$\tau_{m,p}(n) = \begin{cases} m[\log_p n] & \text{if } a_p(n) > m, \\ m[\log_p n] + a_p(n) - m & \text{if } a_p(n) \le m. \end{cases} \tag{3.14}$$

The asymptotic contribution to $\sum_{p\le n}\tau_{m,p}(n)\log p$ from the logarithmic term in $\tau_{m,p}$ is given by Lemma 3.2.1:

$$\sum_{p\le n} m[\log_p n]\log p = mn + O(n\exp\{-\alpha(\log n)^{1/2}\}). \tag{3.15}$$

The asymptotic contribution from $m[\log_p n] - \tau_{m,p}(n)$ is more difficult to establish. We will show that for some positive constant α

$$\sum_{p\le n}\big(m[\log_p n] - \tau_{m,p}(n)\big)\log p = (m - H_m)n + O(n\exp\{-\alpha(\log n)^{1/2}\}\log n). \tag{3.16}$$

Combining (3.14), (3.15), and (3.16) immediately proves the claim so we now establish (3.16). Let $t \ge 1, a \ge 1, n > 1$ be integers and let p be a prime. Then we have the following equivalences,

$$\begin{aligned} a = [np^{-t}] &\iff a \le np^{-t} < a+1 \\ &\iff an^{-1} \le p^{-t} < (a+1)n^{-1} \\ &\iff (na^{-1})^{\frac{1}{t}} \ge p > (n(a+1)^{-1})^{\frac{1}{t}}. \end{aligned} \tag{3.17}$$

If $a = [np^{-t}]$ then we also claim that

$$1 \le a < p \iff t = [\log_p n], a = a_p(n). \tag{3.18}$$

To see this, let $n = b_0 + b_1 p + \cdots + b_s p^s$ with $0 \le b_i < p$ for $0 \le i \le s$, and s chosen to be $\ge t$. Then

$$[np^{-t}] = [(b_0 + \cdots + b_{t-1}p^{t-1})p^{-t} + b_t + b_{t+1}p + \cdots + b_s p^{s-t}].$$

As $0 \le (b_0 + \cdots + b_{t-1}p^{t-1})p^{-t} < 1$, we have that $a = [np^{-t}] = b_t + b_{t+1}p + \cdots + b_s p^{s-t}$. If $1 \le a < p$ then we must have $t = [\log_p n]$, $b_{t+1} = \cdots = b_s = 0$, and $a = b_t = a_p(n)$. The converse of (3.18) is clear by the definition of $a_p(n)$. This proves the equivalence (3.18), and by putting (3.17) and (3.18) together we get that for any integers $t \ge 1, a \ge 1, n > 1$, and prime p,

$$a = a_p(n), t = [\log_p n] \iff (n(a+1)^{-1})^{\frac{1}{t}} < p \le (na^{-1})^{\frac{1}{t}}, 1 \le a < p. \tag{3.19}$$

We will use (3.19) to sum over triples of integers t, a, p such that p is prime, $a = a_p(n)$, and $t = [\log_p n]$. For integers $a \ge 1, n > 1$, define

$$P_{a,n} := \{p \text{ prime} : (n(a+1)^{-1})^{\frac{1}{t}} < p \le (na^{-1})^{\frac{1}{t}} \text{ for some integer } 1 \le t \le [\log_2 n]\},$$

and consider the sum

$$G(n) := \sum_{a=1}^{m-1} \sum_{p \in P_{a,n}} (m-a) \log p.$$

Using (3.19) shows that

$$G(n) = \sum_{p \le n, a_p(n) \le m} (m - a_p(n)) \log p.$$

From (3.14),

$$\sum_{p \le n} \big(m[\log_p n] - \tau_{m,p}(n)\big) \log p = O(1) + \sum_{p \le n, a_p(n) \le m} (m - a_p(n)) \log p,$$

where the implied constant comes from the primes $\le m$, and so we have that

$$\sum_{p \le n} \big(m[\log_p n] - \tau_{m,p}(n)\big) \log p = G(n) + O(1). \tag{3.20}$$

In view of (3.20) it will suffice to prove (3.16) for $G(n)$. Observe that

$$G(n) = \sum_{t=1}^{[\log_2 n]} \sum_{a=1}^{m-1} (m-a)\{\vartheta((na^{-1})^{\frac{1}{t}}) - \vartheta((n(a+1)^{-1})^{\frac{1}{t}})\}.$$

Let $G_t(n)$ denote the inner sum for $1 \le t \le [\log_2 n]$. Then

$$G_t(n) = m\vartheta\{(\tfrac{n}{1})^{1/t}\} - \vartheta\{(\tfrac{n}{1})^{1/t}\} - \vartheta\{(\tfrac{n}{2})^{1/t}\} - \cdots - \vartheta\{(\tfrac{n}{m})^{1/t}\}. \tag{3.21}$$

By (3.13) there is a positive constant α such that

$$\vartheta\{(\tfrac{n}{a})^{1/t}\} = (\tfrac{n}{a})^{1/t} + O(n\exp\{-\alpha(\log n)^{1/2}\}).$$

With the help of (3.21) we get that

$$G_t(n) = \{mn^{\frac{1}{t}} - n^{\frac{1}{t}} - (n/2)^{\frac{1}{t}} - \cdots - (n/m)^{\frac{1}{t}}\} + O(n\exp\{-\alpha(\log n)^{1/2}\}). \tag{3.22}$$

Let $L_a(n) := \sum_{t=1}^{[\log_2 n]} (\frac{n}{a})^{\frac{1}{t}}$ for $1 \le a \le m$. By summing up (3.22) we get

$$G(n) = mL_1(n) - L_1(n) - L_2(n) - \cdots - L_m(n) + O(n\exp\{-\alpha(\log n)^{1/2}\}\log n). \tag{3.23}$$

Once $n \ge a$ we have that

$$\begin{aligned}(n/a) \le L_a(n) &= (n/a) + (n/a)^{\frac{1}{2}}\big\{1 + (n/a)^{\frac{1}{3}-\frac{1}{2}} + \cdots + (n/a)^{\frac{1}{[\log_2 n]}-\frac{1}{2}}\big\} \\ &< (n/a) + (n/a)^{\frac{1}{2}}(\log_2 n),\end{aligned}$$

and so $L_a(n) = (n/a) + O(\sqrt{n}\log n)$. Finally, from (3.23) we get that

$$G(n) = mn - (n/1) - (n/2) - \cdots - (n/m) + O(n\exp\{-\alpha(\log n)^{1/2}\}\log n)$$

which proves (3.16) in view of (3.20). □

Remark. If the Riemann hypothesis is true then the error terms in Lemma 3.2.1 and Theorem 3.0.2 improve significantly: for any $\varepsilon > 0$,

$$\sum_{p\le n} [\log_p n]\log p = n + O(n^{1/2+\varepsilon}), \quad \sum_{p\le n} \tau_{m,p}(n)\log p = nH_m + O(n^{1/2+\varepsilon}\log n).$$

We will not need these stronger error terms for the applications in §3.3.

3.3 Proof of Theorem 3.3.2

In §3.1 we established a local estimate for the finite differences c of a sequence s with integral mth divided difference. The calculations in §3.2 show that

$$\lim_{n\to\infty} \prod_{p \text{ prime}} p^{\frac{\tau_{m,p}(n)}{n}} = e^{1+\frac{1}{2}+\cdots+\frac{1}{m}}.$$

By combining the local estimates for $\delta_m s$ with this calculation we will obtain a characterization of polynomial sequences in terms of the Archimedean growth of their finite differences. Let K be an algebraic number field of degree d with ring of integers $\mathcal{O}$. For the sake of generality, we work with an arbitrary finite set S of places of K that contains the Archimedean places. Recall that the finite differences of s are defined by

$$c(n) := \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} s(k) \quad (n \in \mathbb{N}).$$

Proposition 3.3.1. *Let $s\colon \mathbb{N} \to K$ and let $S \subset M_K$ be a finite set containing the Archimedean places. Suppose that*

(i) $\delta_m s$ is $\mathcal{O}$-valued, and

(ii) for each v in S there is a positive number ρ_v such that $|c(n)|_v \ll \rho_v^n$.

If

$$\prod_{v\in S} \rho_v^{d_v} < e^{d\left(1+\frac{1}{2}+\cdots+\frac{1}{m}\right)} \tag{3.24}$$

then $s(n)$ is a polynomial in n.

Proof. By Lemma 3.1.1 the conclusion is equivalent to c being eventually zero, so for the sake of contradiction suppose that (3.24) holds and that c has infinitely many nonzero terms.

Let v be a place of K not in S, p_v the rational prime v lies over, σ_v a representative embedding for v, and d_v the local degree of v. The finite differences (3.3) of the sequence $\sigma_v s\colon \mathbb{N} \to \mathbb{C}_p$ are clearly given by $\sigma_v c(n)$. We may apply Theorem 3.0.1 to see that $|c(n)|_v = |\sigma_v c(n)|_{p_v} \leq p_v^{-\tau_{m,p_v}(n)}$ for $n \geq m$, and so

$$\prod_{v \notin S} |c(n)|_v^{d_v} \leq \prod_{v \notin S} p_v^{-d_v \tau_{m,p_v}(n)}. \tag{3.25}$$

Note that both sides amount to finite products ($\tau_{m,p_v}(n) = 0$ if $p_v > n$).

By the definition of $\tau_{m,p}(n)$,

$$\tau_{m,p}(n) \leq \max_{1 \leq i_1 \leq \cdots \leq i_m \leq n} (w_p(i_1) + \cdots + w_p(i_m)) = m \max_{1 \leq i \leq n} w_p(i) = m[\log_p(n)],$$

and in particular $\tau_{m,p}(n) = O(\log n)$. Then, once n is larger than any prime lying under a prime of S, we have that

$$\begin{aligned} \sum_{v \notin S} d_v \tau_{m,p_v}(n) \log p_v &= \sum_{p_v \leq n} d_v \tau_{m,p_v}(n) \log p_v - \sum_{v \in S} d_v \tau_{m,p_v}(n) \log p_v \\ &= d \sum_{p \leq n} \tau_{m,p}(n) \log p + O(\log n). \end{aligned}$$

With the help of Theorem 3.0.2 we see that

$$\sum_{v \notin S} d_v \tau_{m,p_v}(n) \log p_v = dnH_m + o(n).$$

By putting this together with (3.25) we obtain

$$\prod_{v \notin S} |c(n)|_v^{d_v} \leq e^{-ndH_m + o(n)}.$$

Now let n_i be chosen so that $c(n_i) \neq 0$ for all non-negative integers i. With the help of the product formula we obtain

$$\prod_{v \in S} \rho_v^{-d_v} \leq \liminf_{i \to \infty} \left(\prod_{v \in S} |c(n_i)|_v^{-d_v/n_i} \right) \leq e^{-dH_m}.$$

This contradicts (3.24) and concludes the proof. □

We now prove a generalization of Theorem 1.3.1. In addition to the integrality of higher divided differences we will consider the possibility of p-adic analytic interpolation — i.e., the existence of a power series $F(x) \in \mathbb{C}_p[\![x]\!]$ which converges for all $x \in \mathbb{D}_p^{<R} := \{x \in \mathbb{C}_p : |x|_p < R\}$ such that $R > 1$ and $F(n) = \sigma_v s(n)$ for all $n \geq 0$. It is known that p-adic analytic interpolation corresponds to p-adic decay of finite differences (cf. e.g., [Sch06], §54). By combining this decay with the decay coming from the integrality of $\delta_m s$, we obtain a common generalization of the Hall–Ruzsa and Hilliker–Straus theorems, [Hal71], [Ruz71], [HS70], as well as one of the main results from Dwork's work on the rationality of the zeta function over a finite field (cf. Remark 2).

Theorem 3.3.2. *Let $s\colon \mathbb{N} \to K$, let $S \subset M_K$ be a finite set containing the Archimedean places M_K^∞, and let $F \subset M_K$ be another finite set disjoint from S. Suppose that*

(i) $\delta_m s$ is $\mathcal{O}$-valued,

(ii) for each v in S there is a positive number θ_v such that $|s(n)|_v \ll \theta_v^n$, and

(iii) for each v in F there is a number $R_v > 1$ such that $\sigma_v s$ extends to a p-adic analytic function $\mathbb{D}_{p_v}^{<R_v} \to \mathbb{C}_{p_v}$.

If

$$\prod_{v \in M_K^\infty} (1+\theta_v)^{d_v} \prod_{v \in S \setminus M_K^\infty} \max\{1, \theta_v\}^{d_v} \prod_{v \in F} (p_v^{\frac{1}{p_v - 1}} R_v)^{-d_v} < e^{d\left(1 + \frac{1}{2} + \cdots + \frac{1}{m}\right)} \tag{3.26}$$

then $s(n)$ is a polynomial in n.

Proof of Theorem 1.3.1. Take $S = M_K^\infty$, $F = \varnothing$ and apply Theorem 3.3.2, noting that there are two isometric embeddings for every complex place. □

Proof of Theorem 3.3.2. By Proposition 3.3.1 we see that s is polynomial if for each

$v \in S \cup F$ there are positive constants D_v and ρ_v such that

$$|c(n)|_v \leq D_v \rho_v^n \quad \text{for all } n \geq 0 \tag{3.27}$$

and

$$\prod_{v \in S \cup F} \rho_v^{d_v} < e^{dH_m}. \tag{3.28}$$

Suppose that v is an Archimedean place in S. Then $|c(n)|_v \leq \max_{0 \leq k \leq n} |s(k)|_v$ by the ultrametric inequality. As $|s(n)|_v \ll \theta_v^n$ by hypothesis, there is a positive constant D_v such that

$$|c(n)|_v \leq \max_{0 \leq k \leq n} C\theta_v^k = \begin{cases} D_v & \text{if } \theta_v < 1, \\ D_v \theta_v^n & \text{if } \theta_v \geq 1. \end{cases}$$

Hence in either case we may take $\rho_v = \max\{1, \theta_v\}$. Now suppose v is a non-Archimedean place in S. We have that

$$|c(n)|_v \leq \sum_{0 \leq k \leq n} \binom{n}{k} |s(k)|_v.$$

Therefore for some positive constant D_v we have that $|c(n)|_v \leq D_v(1 + \theta_v)^n$, and here we take $\rho_v = 1 + \theta_v$. Then for every v in S (3.27) is satisfied.

Now we consider the places v in F. By hypothesis, for any place v in F there exists an analytic function $f_v(x)$ defined on the closed disk of $\mathbb{C}_{p_v}$ of radius R_v, $R_v > 1$, containing zero such that $\sigma_v s(n) = f_v(n)$ for all $n \geq 0$. Without loss of generality, we may assume that f_v is analytic on a disk of radius strictly larger than R_v for all $v \in F$ since the inequality (3.26) remains valid even if R_v is replaced with a sufficiently close but smaller quantity. Furthermore, by taking a sufficiently small $\varepsilon > 0$ we may assume that for all v in F, f_v is analytic on a disk of radius strictly larger than $R_v + \varepsilon$. Now we make use of a theorem of Iwasawa, [Iwa72], Theorem 3,

to see that

$$\lim_{n\to\infty} |c(n)|_v r^{-n} = 0 \tag{3.29}$$

for any real number r such that

$$p_v^{\frac{-1}{p_v-1}}(R_v+\varepsilon)^{-1} < r. \tag{3.30}$$

On the other hand, (3.29) implies that $|c(n)|_v^{1/n} > r$ for only finitely many n, or that $\limsup_n |c(n)|_v^{1/n} \leq r$. Hence there are positive constants D_v, ρ_v satisfying (3.27) and $\rho_v \leq r$ for all $v \in F$. As r was arbitrary subject to (3.30) this shows that r may be taken to be $\leq p_v^{\frac{-1}{p_v-1}} R_v^{-1}$. Therefore the constants D_v, ρ_v may be chosen to satisfy (3.27) as well as

$$\prod_{v\in F} \rho_v^{d_v} \leq \prod_{v\in F} (p_v^{\frac{1}{p_v-1}} R_v)^{-d_v}. \tag{3.31}$$

Putting (3.31) together with the choices of ρ_v for v in S shows that

$$\prod_{v\in S\cup F} \rho_v^{d_v} \leq \prod_{v\in M_K^\infty} (1+\theta_v)^{d_v} \prod_{v\in S\setminus M_K^\infty} \max\{1,\theta_v\}^{d_v} \prod_{v\in F} (p_v^{\frac{1}{p_v-1}} R_v)^{-d_v}.$$

This inequality shows that (3.26) implies (3.28) and concludes the proof. □

Remark 2. It is well-known that a power series $\sum_{n\geq 0} a_n X^n$ is the expansion of a rational function if and only if there exists an integer N such that

$$c(n) := \det(a_{n+i+j})_{i,j=0}^{N} \tag{3.32}$$

is zero for all sufficiently large n. Theorem 3.3.2 may be applied to the sequence s whose finite differences are given by (3.32) to obtain a generalization of Dwork's Theorems 2 and 3 from his article proving the rationality of the zeta function of a variety over a finite field [Dwo60]. We have not emphasized this application however as hypotheses (1) and (3) of Theorem 3.3.2 do not appear to be natural conditions on power series.

BIBLIOGRAPHY

BIBLIOGRAPHY

[Ami64] Yvette Amice. Interpolation p-adique. *Bull. Soc. Math. France*, 92:117–180, 1964.

[BGT16] Jason P. Bell, Dragos Ghioca, and Thomas J. Tucker. *The dynamical Mordell-Lang conjecture*, volume 210 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2016.

[BN19] Jason P. Bell and Khoa D. Nguyen. An analogue of Ruzsa's conjecture for polynomials over finite fields, 2019.

[Con02] Brian Conrad. A modern proof of Chevalley's theorem on algebraic groups. *J. Ramanujan Math. Soc.*, 17(1):1–18, 2002.

[Dwo60] Bernard Dwork. On the rationality of the zeta function of an algebraic variety. *Amer. J. Math.*, 82:631–648, 1960.

[Fal83] G. Faltings. Endlichkeitssätze für abelsche Varietäten über Zahlkörpern. *Invent. Math.*, 73(3):349–366, 1983.

[Fal84] Gerd Faltings. Complements to Mordell. In *Rational points (Bonn, 1983/1984)*, Aspects Math., E6, pages 203–227. Friedr. Vieweg, Braunschweig, 1984.

[Fal91] Gerd Faltings. Diophantine approximation on abelian varieties. *Ann. of Math. (2)*, 133(3):549–576, 1991.

[Fal94] Gerd Faltings. The general case of S. Lang's conjecture. In *Barsotti Symposium in Algebraic Geometry (Abano Terme, 1991)*, volume 15 of *Perspect. Math.*, pages 175–182. Academic Press, San Diego, CA, 1994.

[Gru12] Samuel Grushevsky. The Schottky problem. In *Current developments in algebraic geometry*, volume 59 of *Math. Sci. Res. Inst. Publ.*, pages 129–164. Cambridge Univ. Press, Cambridge, 2012.

[GT09] D. Ghioca and T. J. Tucker. Periodic points, linearizing maps, and the dynamical Mordell-Lang problem. *J. Number Theory*, 129(6):1392–1403, 2009.

[GTZ12] Dragos Ghioca, Thomas J. Tucker, and Michael E. Zieve. Linear relations between polynomial orbits. *Duke Math. J.*, 161(7):1379–1410, 2012.

[Hal71] R. R. Hall. On pseudo-polynomials. *Mathematika*, 18:71–77, 1971.

[Har77] Robin Hartshorne. *Algebraic geometry*. Springer-Verlag, New York-Heidelberg, 1977. Graduate Texts in Mathematics, No. 52.

[HS70] D. L. Hilliker and E. G. Straus. Some p-adic versions of Polya's theorem on integer valued analytic functions. *Proc. Amer. Math. Soc.*, 26:395–400, 1970.

[Iit77] S. Iitaka. On logarithmic Kodaira dimension of algebraic varieties. In *Complex analysis and algebraic geometry*, pages 175–189. 1977.

[Iwa72] Kenkichi Iwasawa. *Lectures on p-adic L-functions*. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1972. Annals of Mathematics Studies, No. 74.

[Lan60] Serge Lang. Integral points on curves. *Inst. Hautes Études Sci. Publ. Math.*, pages 27–43, 1960.

[Lan65] Serge Lang. Division points on curves. *Ann. Mat. Pura Appl. (4)*, 70:229–234, 1965.

[Lan91] Serge Lang. *Number theory. III*, volume 60 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 1991. Diophantine geometry.

[Lan02] Serge Lang. *Algebra*, volume 211 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 2002.

[Mah58] K. Mahler. An interpolation series for continuous functions of a p-adic variable. *J. Reine Angew. Math.*, 199:23–34, 1958.

[McQ95] Michael McQuillan. Division points on semi-abelian varieties. *Invent. Math.*, 120(1):143–159, 1995.

[Mil80] James S. Milne. *Étale cohomology*, volume 33 of *Princeton Mathematical Series*. Princeton University Press, Princeton, N.J., 1980.

[Mil17] J. S. Milne. *Algebraic groups*, volume 170 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2017. The theory of group schemes of finite type over a field.

[Mor22] L. J. Mordell. On the rational solutions of the indeterminate equation of the third and fourth degrees. *Math. Proc. Cambridge Philos. Soc.*, 21:179–192, 1922.

[MT51] L. M. Milne-Thomson. *The Calculus of Finite Differences*. Macmillan and Co., Ltd., London, 1951.

[Mum07] David Mumford. *Tata lectures on theta. I*. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, 2007. With the collaboration of C. Musili, M. Nori, E. Previato and M. Stillman, Reprint of the 1983 edition.

[Neu99] Jürgen Neukirch. *Algebraic number theory*, volume 322 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999. Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder.

[OZ] Andrew O'Desky and Michael E. Zieve. On the intersection of orbits of rational functions. *In preparation*.

[PZ84] A. Perelli and U. Zannier. On recurrent mod p sequences. *J. Reine Angew. Math.*, 348:135–146, 1984.

[RS62] J. Barkley Rosser and Lowell Schoenfeld. Approximate formulas for some functions of prime numbers. *Illinois J. Math.*, 6:64–94, 1962.

[Ruz71] Imre Ruzsa, Jr. On congruence-preserving functions. *Mat. Lapok*, 22:125–134 (1972), 1971.

[Sch06] W. H. Schikhof. *Ultrametric calculus*, volume 4 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.

[Sil07] Joseph H. Silverman. *The arithmetic of dynamical systems*, volume 241 of *Graduate Texts in Mathematics*. Springer, New York, 2007.

[Sil09] Joseph H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics*. Springer, Dordrecht, second edition, 2009.

[Sil12] Joseph H. Silverman. *Moduli spaces and arithmetic dynamics*, volume 30 of *CRM Monograph Series*. American Mathematical Society, Providence, RI, 2012.

[Sti09] Henning Stichtenoth. *Algebraic function fields and codes*, volume 254 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, second edition, 2009.

[Tsi12] Jacob Tsimerman. The existence of an abelian variety over $\overline{\mathbb{Q}}$ isogenous to no Jacobian. *Ann. of Math. (2)*, 176(1):637–650, 2012.

[vdW35] B. L. van der Waerden. Die Zerlegungs-und Trägheitsgruppe als Permutationsgruppen. *Math. Ann.*, 111(1):731–733, 1935.

[Voj96] Paul Vojta. Integral points on subvarieties of semiabelian varieties. I. *Invent. Math.*, 126(1):133–181, 1996.

[Zan82] U. Zannier. A note on recurrent mod p sequences. *Acta Arith.*, 41(3):277–280, 1982.

[Zan96] Umberto Zannier. On periodic mod p sequences and G-functions: on a conjecture of Ruzsa. *Manuscripta Math.*, 90(3):391–402, 1996.