INTEGRAL POINTS IN THE ORBITS OF RATIONAL FUNCTIONS

PETER POWELL AND TEJO MADHAVARAPU

Abstract. In this paper we prove a theorem by J. Silverman which states that if the orbit of a rational point under a rational function φ contains infinitely many integers and $\deg(\varphi) \geq 2$, then some iterate of φ is a polynomial. First, we introduce some background material about the dynamics of complex functions, focusing mainly on rational maps. We then introduce some important results on Diophantine approximation which are used to prove a special case of Siegel's theorem on integral points. Finally, this is used to study the integral points in the orbits of rational functions.

1. Introduction

Suppose that S is a set and $\varphi : S \to S$ is a function from S to itself. We define the nth iterate of φ as

$$\varphi^n = \underbrace{\varphi \circ \varphi \circ \cdots \circ \varphi}_{n \text{ times}},$$

where φ^0 is the identify map. We can think of φ as describing a discrete time evolution of points in S. For this reason, we make the following definition.

Definition 1. A *(discrete) dynamical system* is an ordered pair (S, φ), where S is a set and $\varphi : S \to S$ is a function.

We now classify points based on how they behave after applying φ to them repeatedly.

Definition 2. If (S, φ) is a dynamical system, then the *orbit* of an element $s \in S$ is the set

$$\mathcal{O}_\varphi(s) = \{s, \varphi(s), \varphi^2(s), \dots\}.$$

Definition 3. We say that an element $s \in S$ is *wandering* if $\mathcal{O}_\varphi(s)$ is infinite, *preperiodic* if $\mathcal{O}_\varphi(s)$ is finite, and *periodic* if $\varphi^n(s) = s$ for some integer $n > 0$. We also say that an element $s \in S$ is a *fixed point* if $\varphi(s) = s$, or equivalently, $\mathcal{O}_\varphi(s) = \{s\}$.

Of course, any periodic point is also preperiodic, but the converse does not necessarily hold unless φ is injective.

Notation 4. The sets of preperiodic and periodic points of φ are denoted by $\mathrm{PrePer}(\varphi, S)$ and $\mathrm{Per}(\varphi, S)$, respectively.

Example 5. Let $\varphi : \mathbb{F}_{13} \to \mathbb{F}_{13}$ denote the map

$$\varphi(z) = z^2 + 1.$$

Then, $(\mathbb{F}_p, \varphi)$ is a dynamical system. Figure 1 illustrates this system. Since $\mathbb{F}_{13}$ is

Date: December 10, 2023.

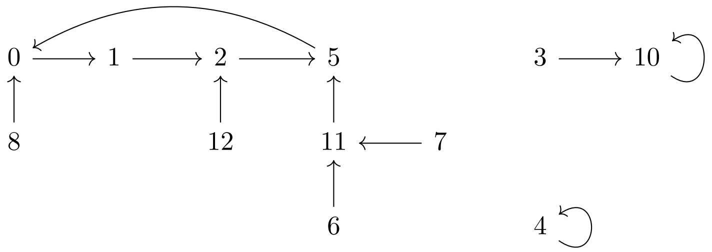

FIGURE 1. The action of $\varphi(z) = z^2 + 1$ on $\mathbb{F}_{13}$.

finite, clearly all points are preperiodic, so $\text{PrePer}(\varphi, S) = \mathbb{F}_{13}$. From Figure 1 we see that 4 and 10 are the fixed points and

$$\text{Per}(\varphi, \mathbb{F}_{13}) = \{0, 1, 2, 4, 5, 10\}$$

The points 0, 1, 2 and 5 have period 4, while 4 and 10 have period 1.

In this paper we will be mostly concerned with the dynamics of rational functions, and in particular, integral points in the orbits of rational functions. We note that if φ is a polynomial with integer coefficients, then clearly there are rational numbers α such that the orbit $\mathcal{O}_\varphi(\alpha)$ contains infinitely many integers (for example, any integer satisfies this). In fact, this is true if any iterate of φ is a polynomial, i.e. φ^n is an integer polynomial for some integer $n \geq 1$. The goal of this paper is to establish an interesting partial converse to this: if φ is a rational function of degree at least 2, and there exists some rational number α such that $\mathcal{O}_\varphi(\alpha)$ contains infinitely many integers, then some iterate of φ is a polynomial with rational coefficients. This does not give a complete characterization of the rational functions which admit orbits of rational points containing infinitely many integers, since a polynomial with rational coefficients might not admit orbits of rational points containing infinitely many integers, as the following example demonstrates.

Example 6. Let $\varphi : \mathbb{Q} \to \mathbb{Q}$ denote the map

$$\varphi(z) = \frac{1}{2}z^2 + 1.$$

Then $(\mathbb{Q}, \varphi)$ is a dynamical system. We note that if $\varphi(z)$ is an integer, then $\frac{1}{2}z^2$ must be an integer and thus z^2 must be an even integer. It follows that $\varphi(z) \in \mathbb{Z}$ if and only if z is an even integer. Thus, if α is a non-integer rational number, then

$$\mathcal{O}_\varphi(\alpha) \cap \mathbb{Z} = \varnothing,$$

and if α is an odd integer then

$$\mathcal{O}_\varphi(\alpha) \cap \mathbb{Z} = \{\alpha\}.$$

If α is an even integer, then $\frac{1}{2}\alpha^2$ is even, so $\varphi(\alpha) = \frac{1}{2}\alpha^2 + 1$ is odd. It follows that $\varphi^2(\alpha)$ is not an integer and hence

$$\mathcal{O}_\varphi(\alpha) \cap \mathbb{Z} = \{\alpha, \varphi(\alpha)\}.$$

In particular, the orbit of any $\alpha \in \mathbb{Q}$ contains finitely many integers.

The theorems and proofs in this paper are mostly based off those in [5] and [1]. We start by studying rational functions as self-maps of the complex projective line.

2. Rational Functions of the Complex Projective Line

We define a rational function φ as a function of the form $\varphi(z) = F(z)/G(z)$, where F and G are polynomials and have no common root. If the coefficients of F and G are complex then we write $\varphi(z) \in \mathbb{C}(z)$. In the special case where these coefficients are required to be rational, we will write $\varphi(z) \in \mathbb{Q}(z)$. We would like to study the dynamics of these rational functions. However, φ is not quite a self-map of $\mathbb{C}$, since the denominator $G(z)$ is not necessarily nonzero. For this reason, we add an extra point at infinity, denoted ∞, and we define

$$\varphi(z) = \infty$$

if $G(z) = 0$. These are the *poles* of φ. Further, we define

$$\varphi(\infty) = \lim_{z\to\infty} \frac{F(z)}{G(z)}.$$

With these definitions, φ is now a self-map of $\mathbb{C} \cup \{\infty\}$, and $(\mathbb{C} \cup \{\infty\}, \varphi)$ is a dynamical system. The set $\mathbb{C} \cup \{\infty\}$ is called the extended complex plane.

We can put a topology on this set as follows. We say a set $U \subset \mathbb{C} \cup \{\infty\}$ is open if U is either an open subset of $\mathbb{C}$ or $U = C \cup \{\infty\} - K$, where K is some closed and bounded subset of $\mathbb{C}$. This is known as the one-point-compactification of $\mathbb{C}$. As the name suggests, it is a compact space. In fact, it is homeomorphic to the sphere S^2 via stereographic projection. For more details, see a textbook on complex analysis, for example [4].

$\mathbb{C} \cup \{\infty\}$ is also homeomorphic to the set of all 1-dimensional subspaces of the complex vector space $\mathbb{C}^2$, which is denoted $\mathbb{P}^1(\mathbb{C})$ and called the complex projective line. If we define the equivalence relation $\sim$ by $(z, w) \sim (z', w')$ if and only if there is some nonzero $\lambda \in \mathbb{C}$ such that $(z', w') = (\lambda z, \lambda w)$, then $\mathbb{P}^1(\mathbb{C})$ is the set of equivalence classes of vectors $(z, w) \in \mathbb{C}^2 - \{0, 0\}$. The equivalence class of (z, w) is denoted $[z, w]$, which is called the homogeneous coordinates, or projective coordinates, of a point in $\mathbb{P}^1(\mathbb{C})$. We define $\pi : \mathbb{C}^2 \to \mathbb{P}^1(\mathbb{C})$ by the projection map $\pi(x, y) = [x, y]$. This induces a topology on $\mathbb{P}^1(\mathbb{C})$ called the quotient topology. A set $U \subset \mathbb{P}^1(\mathbb{C})$ is defined to be open in the quotient topology if and only if the preimage $\pi^{-1}(U)$ is open in $\mathbb{C}^2$.

Now, the map $\mathbb{C} \to \mathbb{C}^2$ defined by $z \to (z, 1)$ is an embedding, and we can compose this with the projection map π to obtain an embedding $\mathbb{C} \to \mathbb{P}^1(\mathbb{C})$. The image of this map is every point in $\mathbb{P}^1(\mathbb{C})$ except $[1, 0]$. Thus, we can extend the embedding $\mathbb{C} \to \mathbb{P}^1(\mathbb{C})$ to a homeomorphism $\mathbb{C} \cup \{\infty\} \to \mathbb{P}^1(\mathbb{C})$ by sending ∞ to the point $[1, 0]$. For this reason, we can identity $\mathbb{C} \cup \{\infty\}$ with $\mathbb{P}^1(\mathbb{C})$ by identifying each point $z \in \mathbb{C} \cup \{\infty\}$ with its image under this homeomorphism. For this reason, we say that the homogeneous coordinates of a point $z \in \mathbb{C}$ are $[z, 1]$ and the homogeneous coordinates of ∞ are $[1, 0]$.

Now, we return to studying rational functions. An important property of rational functions is their degree.

Definition 7. The *degree* of a rational function $\varphi(z) = F(z)/G(z)$ is defined as

$$\deg(\varphi) \coloneqq \max\{\deg(F), \deg(G)\}.$$

We can also write rational functions using homogeneous coordinates. It is not difficult to check that a rational function $\varphi(z) = F(z)/G(z)$ takes the form

$$\varphi[X, Y] = [F^*[X, Y], G^*[X, Y]]$$

where $F^*[X,Y] = Y^d F(X/Y)$ and $G^*[X,Y] = Y^d G[X/Y]$, and $d = \deg(\varphi)$. The benefit of this is that the point at infinity no longer needs to be treated specially.

Another important property of rational functions is their zeroes.

Definition 8. A complex function φ has a *zero of order* m at α if φ is complex differentiable at α and φ and its first $m-1$ derivatives vanish at α but its mth derivative does not. If this is the case, we write $\operatorname{ord}_\alpha(\varphi) = m$.

There is another useful way to characterize the zeroes of a function.

Theorem 9. *Suppose f is complex differentiable at α. Then f has a zero of order m if and only if*

$$f(z) = (z-\alpha)^m g(z),$$

where g is complex differentiable at α and $g(\alpha) \neq 0$.

For a proof, see [4], for example. It is then easy to see that for any rational function $\varphi(z) = F(z)/G(z)$ such that F and G share no common roots, $\operatorname{ord}_\alpha(\varphi)$ is the number of times $(z-\alpha)$ divides $F(z)$. Now, for any rational function $\varphi(z)$ and any point $\alpha \neq \infty$ such that $\varphi(\alpha) \neq \infty$, φ has a Taylor series expansion around α:

$$\varphi(z) = \varphi(\alpha) + \varphi'(\alpha)(z-\alpha) + \frac{1}{2}\varphi''(\alpha)(z-\alpha)^2 + \cdots .$$

If $\varphi'(\alpha) = 0$, then we say φ has a *ramification point*, or *critical point* at α. Further, we make the following definition.

Definition 10. If $\varphi \in \mathbb{C}(z)$ and $\alpha \in \mathbb{C}$ is such that $\varphi(\alpha) \neq \infty$, then the *ramification index* of φ at α is defined as

$$e_\alpha(\varphi) = \operatorname{ord}_\alpha(\varphi(z) - \varphi(\alpha)).$$

Essentially, this is the degree of the first nonzero, nonconstant term in the Taylor series of φ at α. Note that $e_\alpha(\varphi) \geq 2$ if and only if α is a ramification point. Also, one can check that if φ is a rational function then we must have $e_\alpha(\varphi) \leq \deg(\varphi)$ for all α. In the case that $e_\alpha(\varphi) = \deg(\varphi)$ we say that φ is *totally ramified* at α, or that α is a *totally ramified* point.

Example 11. The function $\varphi : \mathbb{P}^1(\mathbb{C}) \to \mathbb{P}^1(\mathbb{C})$ defined by

$$\varphi(z) = \frac{z^2-2}{2z+3}$$

is a rational function of degree 2. It has one pole at $z = -\frac{3}{2}$, so

$$\varphi\left(-\frac{3}{2}\right) = \infty.$$

φ has a zero of order one at both $\sqrt{2}$ and at $-\sqrt{2}$, so we write $\operatorname{ord}_{\sqrt{2}}(\varphi) = \operatorname{ord}_{-\sqrt{2}}(\varphi) = 1$. To find the ramification points of φ, we compute the derivative:

$$\varphi'(z) = \frac{2(z+1)(z+2)}{(2z+3)^2}.$$

Thus, the ramification points are -1 and -2. We can then compute

$$e_{-1}(\varphi) = \operatorname{ord}_{-1}(\varphi(z) - \varphi(-1)) = \operatorname{ord}_{-1}\left(\frac{(z+1)^2}{2z+3}\right) = 2,$$

and

$$e_{-2}(\varphi) = \text{ord}_{-2}(\varphi(z) - \varphi(-2)) = \text{ord}_{-1}\left(\frac{(z+2)^2}{2z+3}\right) = 2.$$

In fact, we could have noted that $e_\alpha(\varphi) \leq 2$ for all α because $\deg(\varphi) = 2$. This implies that $e_\alpha(\varphi) = 2$ for both -1 and -2, since $e_\alpha(\varphi) \geq 2$ for ramification points α. Also, for all $\alpha \neq -1, -2, -\frac{3}{2}, \infty$, we have $e_\alpha(\varphi) = 1$. Currently, we have not defined $e_\alpha(\varphi)$ at $\alpha = -\frac{3}{2}$ and $\alpha = \infty$. There is actually a natural way to define $e_\alpha(\varphi)$ in this case, but first we need to introduce something called a Möbius transformation.

3. Möbius Transformations

Definition 12. A *Möbius transformation* is a rational function $f : \mathbb{P}^1(\mathbb{C}) \to \mathbb{P}^1(\mathbb{C})$ of the form

$$f(z) = \frac{az+b}{cz+d},$$

where $a, b, c, d \in \mathbb{C}$ and $ad - bc \neq 0$. As usual for rational functions, if $c \neq 0$ we define $f(-d/c) = \infty$ and

$$f(\infty) = \lim_{z \to \infty} f(z) = \frac{a}{c}.$$

If $c = 0$ we define

$$f(\infty) = \infty.$$

Note that in homogeneous coordinates, we can write a Möbius transformation in the simple form

$$f[X, Y] = [aX + bY, cX + dY].$$

From this, we see that if we associate the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ to the Möbius transformation f, then composition of Möbius transformations corresponds to matrix multiplication of the associated matrices. Note that if we scale a, b, c, and d by some nonzero constant $\lambda \in \mathbb{C}$ then we get the same Möbius transformation. Thus f also has the matrix representation $\begin{pmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{pmatrix}$. Strictly speaking, we should say that the matrix representation of a Möbius transformation f is the equivalence class of 2×2 complex matrices with nonzero determinant under the equivalence relation which equates such matrices if and only if they differ by a nonzero scalar λ. If we define multiplication of these equivalence classes in the obvious way, then this set of equivalence classes becomes a group called the projective linear group of order 2 over $\mathbb{C}$, and is denoted by $PGL(2, \mathbb{C})$.

The reason that these maps are important is that they preserve a lot of important properties. Indeed, from the above we see that any Möbius transformation is induced by the vector space isomorphism

$$(x, y) \to (ax + by, cx + dy)$$

of $\mathbb{C}^2$. Such a map is called an automorphism of projective spaces.

Example 13. The map $f(z) = 1/z$ is a Möbius Transformation with the matrix representation $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. In homogeneous coordinates,

$$f[X, Y] = [Y, X].$$

Note that the matrix representation of $f^2(z) = z$ is

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

as expected.

Next, we prove an important lemma on the existence of certain Möbius transformations.

Lemma 14. *Let $(\alpha, \alpha', \alpha'')$ and (β, β', β'') be two triples of distinct points in $\mathbb{P}^1(\mathbb{C})$. Then there exists a Möbius transformation f such that*

$$f(\alpha) = \beta, \qquad f(\alpha') = \beta', \qquad f(\alpha'') = \beta''.$$

Proof. First consider the case $\alpha = 0$, $\alpha' = 1$, $\alpha'' = \infty$. Suppose that in homogeneous coordinates we have $\beta = [X_1, Y_1]$, $\beta' = [X_2, Y_2]$, and $\beta'' = [X_3, Y_3]$. Define $a, b, c, d \in \mathbb{C}$ by

$$\begin{aligned} a &= Y_1X_2X_3 - X_1Y_2X_3 \\ b &= X_1Y_2X_3 - X_1X_2Y_3 \\ c &= Y_1X_2Y_3 - X_1Y_2Y_3 \\ d &= Y_1Y_2X_3 - Y_1X_2Y_3. \end{aligned}$$

Note that

$$ad - bc = (Y_1X_2 - X_1Y_2)(Y_1X_3 - X_1Y_3)(Y_2X_3 - X_2Y_3),$$

which is nonzero since β, β', and β'' are distinct. Thus we can define a Möbius transformation f by

$$f[X, Y] = [aX + bY, cX + dY].$$

We then have

$$\begin{aligned} f[0,1] &= [b, d] = [X_1(Y_2X_3 - X_2Y_3), Y_1(Y_2X_3 - X_2Y_3)] = [X_1, Y_1] = \beta \\ f[1,1] &= [a + b, c + d] = [X_2(Y_1X_3 - X_1Y_3), Y_2(Y_1X_3 - X_1Y_3))] = [X_2, Y_2] = \beta' \\ f[1,0] &= [a, c] = [X_3(Y_1X_2 - X_1Y_2), Y_3(Y_1X_2 - X_1Y_2)] = [X_3, Y_3] = \beta''. \end{aligned}$$

Thus, f is a Möbius transformation with the desired properties.

Next, consider the case $\beta = 0$, $\beta' = 1$, and $\beta'' = \infty$. By the previous part, we can find a Möbius transformation f such that $f(0) = \alpha$, $f(1) = \alpha'$, and $f(\infty) = \alpha''$. Then, f^{-1} is a Möbius transformation and it satisfies

$$f^{-1}(\alpha) = 0, \qquad f^{-1}(\alpha') = 1, \qquad f^{-1}(\alpha'') = \infty.$$

Finally, consider the general case where $(\alpha, \alpha', \alpha'')$ and (β, β', β'') are arbitrary triples of distinct points in $\mathbb{P}^1(\mathbb{C})$. We can then find Möbius transformations f and g such that

$$f(\alpha) = 0, \qquad f(\alpha') = 1, \qquad f(\alpha'') = \infty$$

and

$$g(0) = \beta, \qquad f(1) = \beta', \qquad g(\infty) = \beta''.$$

The composition $g \circ f$ is then a Möbius transformation which satisfies

$$(g \circ f)(\alpha) = \beta, \qquad (g \circ f)(\alpha') = \beta', \qquad (g \circ f)(\alpha'') = \beta''.$$

□

The following corollary is important for dealing with the point at infinity.

Corollary 15. *Given any rational function φ and any point $\alpha \in \mathbb{P}^1(\mathbb{C})$ such that $\alpha = \infty$ or $\varphi(\alpha) = \infty$, we can find some Möbius transformation f such that $f(\alpha) \neq \infty$ and $f(\varphi(\alpha)) \neq \infty$.*

Proof. If $\alpha = \varphi(\alpha) = \infty$ let f be any Möbius transformation $f(z) = \frac{az+b}{cz+d}$ such that $c \neq 0$. Then,

$$f(\alpha) = f(\varphi(\alpha)) = f(\infty) = a/c \neq \infty.$$

Now suppose $\alpha \neq \varphi(\alpha)$. Choose $\beta, \beta' \in \mathbb{P}^1(\mathbb{C})$ such that $\beta, \beta' \neq \infty$. Also, choose $\alpha'' \neq \alpha, \varphi(\alpha)$, and choose $\beta'' \neq \beta, \beta'$. Then, by Lemma 14 there exists a Möbius transformation f such that

$$f(\alpha) = \beta, \qquad f(\varphi(\alpha)) = \beta', \qquad f(\alpha'') = \beta''.$$

Since $\beta, \beta'' \neq \infty$, f has the desired properties. □

Given a rational map $\varphi : \mathbb{P}^1(\mathbb{C}) \to \mathbb{P}^1(\mathbb{C})$ and a Möbius transformation f, we define the *linear conjugate* of φ by f to be the map

$$\varphi^f = f^{-1} \circ \varphi \circ f.$$

Since f is an automorphism, we think of this map essentially as a change of variables. We imagine two copies of $\mathbb{P}^1(\mathbb{C})$, related by the map f. Given a map φ acting on the first copy of $\mathbb{P}^1(\mathbb{C})$, its linear conjugate φ^f can then be thought of naturally as φ acting on the second copy of $\mathbb{P}^1(\mathbb{C})$. This is demonstrated in the below commutative diagram.

$$\begin{tikzcd}
\mathbb{P}^1(\mathbb{C}) \arrow[r, "\varphi^f"] \arrow[d, "f"] & \mathbb{P}^1(\mathbb{C}) \arrow[d, "f"] \\
\mathbb{P}^1(\mathbb{C}) \arrow[r, "\varphi"] & \mathbb{P}^1(\mathbb{C})
\end{tikzcd}$$

The point of this is that we can now study the behavior of φ at a point α by studying the behavior of φ^f at $f^{-1}(\alpha)$. This is because f is an automorphism, so it turns out that essentially all of the properties of φ we care about are preserved under conjugation. In particular, if $\alpha = \infty$ or $\varphi(\alpha) = \infty$, by Corollary 15 we can take a linear conjugate of φ by a map f so that $f^{-1}(\alpha) \neq \infty$ and $\varphi^f(f^{-1}(\alpha)) = f^{-1}(\varphi(\alpha)) \neq \infty$. In this way, we can study φ at α using the standard techniques of $\mathbb{C}$.

For example, one can verify that the ramification index of φ at α is preserved under conjugation; i.e. if f is a Möbius transformation and $f^{-1}(\alpha) = \beta$, then

$$e_\alpha(\varphi) = e_\beta(\varphi^f),$$

assuming that $\alpha, \beta, \varphi(\alpha), \varphi^f(\beta) \neq \infty$ (since we only defined the ramification index when α and $\varphi(\alpha)$ are not ∞). We can actually use this to define the ramification index at points α such that either $\alpha = \infty$ or $\varphi(\alpha) = \infty$. Given such an α, choose some Möbius transformation f such that $\beta = f^{-1}(\alpha) \neq \infty$ and $\varphi^f(\beta) \neq \infty$. Then, $e_\beta(\varphi^f)$ is defined as before, so we define the ramification index of φ at α to be

$$e_\alpha(\varphi) = e_\beta(\varphi^f).$$

Note that this is well defined, because if f' is a different Möbius transformation such that $\beta' = f'^{-1}(\alpha) \neq \infty$ and $\varphi^{f'}(\beta') \neq \infty$, then

$$\begin{aligned}\varphi^{f'} &= f'^{-1} \circ \varphi \circ f' = f'^{-1} \circ f \circ f^{-1} \circ \varphi \circ f \circ f^{-1} \circ f' \\ &= (f^{-1} \circ f')^{-1} \circ \varphi^f \circ (f^{-1} \circ f') = (\varphi^f)^{f^{-1} \circ f'}.\end{aligned}$$

Since $f^{-1} \circ f'$ is a Möbius transformation it follows that $\varphi^{f'}$ and φ^f are linear conjugates and thus

$$e_\beta(\varphi^f) = e_{\beta'}(\varphi^{f'}).$$

Example 16. Consider the rational function φ from Example 11:

$$\varphi(z) = \frac{z^2 - 2}{2z + 3}.$$

Previously, we computed the ramification index at all points $\alpha \neq -\frac{3}{2}, \infty$. To compute these two ramification indexes, we conjugate by the Möbius transformation

$$f(z) = \frac{1}{z}.$$

This gives

$$\varphi^f(z) = \frac{z(2 + 3z)}{1 - 2z^2}.$$

By definition, $e_\infty(\varphi) = e_0(\varphi^f)$ and $e_{-3/2}(\varphi) = e_{-2/3}(\varphi^f)$. Thus,

$$e_\infty(\varphi) = \mathrm{ord}_0(\varphi^f(z) - \varphi^f(0)) = \mathrm{ord}_0\left(\frac{z(2+3z)}{1-2z^2}\right) = 1$$

and

$$e_{-3/2}(\varphi) = \mathrm{ord}_{-2/3}(\varphi^f(z) - \varphi^f(-2/3)) = \mathrm{ord}_{-2/3}\left(\frac{z(2+3z)}{1-2z^2}\right) = 1.$$

Conjugation is also nice for studying dynamical systems, since it commutes with function iteration:

$$(\varphi^f)^n = (f^{-1} \circ \varphi \circ f) \circ (f^{-1} \circ \varphi \circ f) \circ \cdots \circ (f^{-1} \circ \varphi \circ f) = f^{-1} \circ \varphi^n \circ f.$$

We also note that the degree of a rational function is preserved under linear conjugation, meaning that $\deg(\varphi) = \deg(\varphi^f)$. In fact, this follows from the more general fact that

$$\deg(f \circ g) = \deg(f)\deg(g)$$

for any nonconstant rational functions f and g, which is not too difficult to prove (for example, see [1]).

Example 17 (Lang's Algebra [2] Chapter 1 Problem 55). Let

$$M(z) = \frac{az + b}{cz + d}$$

be a Möbius transformation, and suppose that $M(z)$ has two distinct fixed points not equal to ∞. This means that there are two distinct complex solutions to the equation

$$\frac{az + b}{cz + d} = z.$$

In particular, $c \neq 0$. We will compute $M^k(z)$ for integers $k \geq 1$. Let $W = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ and $W' = \begin{pmatrix} w_1' \\ w_2' \end{pmatrix}$ be the eigenvectors of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with eigenvalues λ and λ' respectively. We also have $\lambda, \lambda' \neq 0$ since $ad - bc \neq 0$. We note that

$$\begin{pmatrix} \lambda w_1 \\ \lambda w_2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} aw_1 + bw_2 \\ cw_1 + dw_2 \end{pmatrix}.$$

This gives the equations $\lambda w_1 = aw_1 + bw_2$ and $\lambda w_2 = cw_1 + dw_2$. This second equation implies $w_2 \neq 0$ because otherwise we would have $cw_1 = 0$ so either $c = 0$ or $w_1 = 0$, both of which would be impossible. Now, if we let $w = w_1/w_2$, we have

$$M(w) = \frac{a\frac{w_1}{w_2} + b}{c\frac{w_1}{w_2} + d} = \frac{aw_1 + bw_2}{cw_1 + dw_2} = \frac{\lambda w_1}{\lambda w_2} = w.$$

So w is a fixed point of M. Similarly, one can show $w_2' \neq 0$ from which it follows that $w' = w_1'/w_2'$ is also a fixed point of M. Note that $\frac{1}{w_2}W = \begin{pmatrix} w \\ 1 \end{pmatrix}$ and $\frac{1}{w_2'}W' = \begin{pmatrix} w' \\ 1 \end{pmatrix}$ are also eigenvectors of M, so

$$aw + b = \lambda w, \qquad cw + d = \lambda, \qquad aw' + b = \lambda' w', \qquad \text{and} \qquad cw' + d = \lambda'.$$

Also, $w \neq w'$ since W and W' are distinct eigenvectors. Next, we define a Möbius transformation S by

$$S(z) = \frac{wz + w'}{z + 1},$$

and we compute the conjugate $M^S = S^{-1} \circ M \circ S$. This is a Möbius transformation given by the matrix

$$\begin{aligned}
\begin{pmatrix} w & w' \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} w & w' \\ 1 & 1 \end{pmatrix} &= \begin{pmatrix} w & w' \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} aw + b & aw' + b \\ cw + d & cw' + d \end{pmatrix} \\
&= \begin{pmatrix} w & w' \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \lambda w & \lambda' w' \\ \lambda & \lambda' \end{pmatrix} \\
&= \frac{1}{w - w'} \begin{pmatrix} 1 & -w' \\ -1 & w \end{pmatrix} \begin{pmatrix} \lambda w & \lambda' w' \\ \lambda & \lambda' \end{pmatrix} \\
&= \frac{1}{w - w'} \begin{pmatrix} \lambda w - \lambda w' & 0 \\ 0 & \lambda' w - \lambda' w' \end{pmatrix} \\
&= \begin{pmatrix} \lambda & 0 \\ 0 & \lambda' \end{pmatrix}.
\end{aligned}$$

It follows that

$$(S^{-1} \circ M \circ S)(z) = \frac{\lambda}{\lambda'} z.$$

Now, we can use the fact that conjugation commutes with function iteration to see that

$$S^{-1} \circ M^k \circ S = (S^{-1} \circ M \circ S)^k = \left(\frac{\lambda}{\lambda'} \right)^k \mathrm{id},$$

where id is the identity map. Composing both sides with S^{-1} on the right and S on the left, we obtain

$$M^k(z) = \frac{\lambda^k w(z - w') - \lambda'^k w'(z - w)}{\lambda^k (z - w') - \lambda'^k (z - w)}.$$

This illustrates how conjugation is useful for studying dynamical systems.

4. Riemann-Hurwitz Formula

In this section we will prove the Riemann-Hurwitz formula, which is a very important relationship between the global property of the degree of a rational function φ and the local property of the ramification index of φ at some point. First, we prove a lemma.

Lemma 18. *Suppose $\varphi(z) \in \mathbb{C}(z)$ and*

$$\varphi(z) = \frac{F(z)}{G(z)} = \frac{a_0 + a_1 z + \cdots + a_{d-1} z^{d-1}}{b_0 + b_1 z + \cdots + b_d z^d}$$

with $b_d \neq 0$. Then, ∞ is a ramification point of φ if and only if $a_{d-1} = 0$.

Proof. Let g be the Möbius transformation $g(z) = \frac{z+1}{z}$. Then, by definition, $e_\infty(\varphi) = e_0(\varphi^g)$. Note that $0 \neq \infty$ and $\varphi^g(0) = -1 \neq \infty$, so we can compute $e_0(\varphi^g)$ as normal. Suppose that

$$\varphi^g(z) = \frac{f(z)}{g(z)},$$

where f and g share no common roots. Note that $g(0) \neq 0$. We then have

$$(\varphi^g)'(z) = \frac{f'(z)g(z) - f(z)g'(z)}{(g(z))^2}.$$

It follows that $e_0(\varphi^g) \geq 2$ if and only if

$$f'(0)g(0) = f(0)g'(0).$$

We now compute $f(0)$, $f'(0)$, $g(0)$, and $f'(0)$. Note that

$$\begin{aligned}\varphi^g(z) &= \frac{1}{\varphi((z+1)/z) - 1} \\ &= \frac{G((z+1)/z)}{F((z+1)/z) - G((z+1)/z)} \\ &= \frac{z^d G((z+1)/z)}{z^d F((z+1)/z) - z^d G((z+1)/z)}.\end{aligned}$$

The numerator and denominator of this last fraction share no common root; 0 is easily seen to not be a common root and any other common root would then mean that F and G share a common root. Thus $z^d G((z+1)/z) = \lambda f(z)$ and $z^d F((z+1)/z - z^d G((z+1)/z)) = \lambda g(z)$ for some nonzero λ. Writing out the first couple terms, we see that

$$z^d G((z+1)/z) = b_d + (d b_d + b_{d-1})z + \cdots$$

and

$$z^d F((z+1)/z) - z^d G((z+1)/z)) = -b_d + (a_{d-1} - d b_d - b_{d-1})z + \cdots .$$

Thus,

$$\lambda f(0) = b_d, \quad \lambda f'(0) = d b_d + b_{d-1}, \quad \lambda g(0) = -b_d, \quad \lambda g'(0) = a_{d-1} - d b_d - b_{d-1}.$$

Then, we can compute

$$\lambda^2 (f'(0)g(0) - f(0)g'(0)) = -b_d a_{d-1}.$$

Since $\lambda \neq 0$, $e_0(\varphi^g) \geq 2$ if and only if $b_d a_{d-1} = 0$. Because $b_d \neq 0$, this means $e_0(\varphi^g) \geq 2$ if and only if $a_{d-1} = 0$ and thus ∞ is a ramification point of φ if and only if $a_{d-1} = 0$. □

Theorem 19 (Riemann-Hurwitz Formula)**.** *If $\varphi(z) \in \mathbb{C}(z)$ and $\deg(\varphi) = d \geq 1$, then*

$$2d - 2 = \sum_{\alpha \in \mathbb{P}^1(\mathbb{C})} (e_\alpha(\varphi) - 1).$$

Proof. By the previous discussion, both sides of this equation are unchanged if we replace φ with one of its linear conjugates. Now, let $x \in \mathbb{C}$ be some point such that $e_x(\varphi) = 1$, $\varphi(x) \neq x$, and $\varphi(\alpha) \neq x$ for any ramification point α. This is possible since there are only finitely many ramification points and finitely many solutions to $\varphi(x) = x$. It follows by Lemma 14 that we can find a Möbius transformation f such that $f(\infty) = x$ and $f(0) = \varphi(x)$.

Now consider the conjugate of φ by f, φ^f. First, we have

$$\varphi^f(\infty) = f^{-1}(\varphi(f(\infty))) = f^{-1}(\varphi(x)) = 0.$$

Also, since conjugation preserves the ramification index, we have

$$e_\infty(\varphi^f) = e_{f(\infty)}(\varphi) = e_x(\varphi) = 1.$$

Lastly, note that $\varphi^f(\alpha') \neq \infty$ for any point α' such that $e_{\alpha'}(\varphi^f) \geq 2$, since this would imply

$$f^{-1}(\varphi(f(\alpha'))) = \infty \implies \varphi(f(\alpha')) = x,$$

and $f(\alpha')$ is a ramification point of φ since $e_{f(\alpha')}(\varphi) = e_{\alpha'}(\varphi^f) \geq 2$.

Thus, by replacing φ with φ^f, we can assume without loss of generality that ∞ is not a ramification point or the image of a ramification point, and that $\varphi(\infty) = 0$. Now, let $\varphi(z) = F(z)/G(z)$, where F and G are polynomials that share no common roots. Since $\varphi(\infty) = 0$, we have $\deg(F) < \deg(G)$ and thus $\deg(G) = d$. Suppose now that

$$\varphi(z) = \frac{F(z)}{G(z)} = \frac{a_0 + a_1 z + \cdots + a_{d-1} z^{d-1}}{b_0 + b_1 z + \cdots + b_d z^d}$$

with $b_d \neq 0$. It follows from Lemma 18 that $a_{d-1} \neq 0$.

Now, note that $e_\alpha(\varphi) - 1 = 0$ whenever α is not a ramification point. Thus, if we let S be the set of ramification points of φ, we have

$$\sum_{\alpha \in \mathbb{P}^1(\mathbb{C})} (e_\alpha(\varphi) - 1) = \sum_{\alpha \in S} (e_\alpha(\varphi) - 1).$$

Next, note that for all $\alpha \neq \infty$ such that $\varphi(\alpha) \neq \infty$, we have

$$\varphi(z) = \varphi(\alpha) + (z - \alpha)^{e_\alpha(\varphi)} \psi(z)$$

for some rational function $\psi(z)$ such that $\psi(\alpha) \neq 0, \infty$, by Theorem 9. Taking the derivative of this, we obtain

$$\begin{aligned} \varphi'(z) &= e_\alpha(\varphi)(z - \alpha)^{e_\alpha(\varphi) - 1} \psi(z) + (z - \alpha)^{e_\alpha(a)} \psi'(z) \\ &= (z - \alpha)^{e_\alpha(\varphi) - 1}(e_\alpha(\varphi)\psi(z) + (z - \alpha)\psi'(z)). \end{aligned}$$

Since $\psi(\alpha) \neq 0$, it follows that $\mathrm{ord}_\alpha(\varphi'(z)) = e_\alpha(\varphi) - 1$, also by Theorem 9. Therefore, we have

$$\sum_{\alpha \in S} (e_\alpha(\varphi) - 1) = \sum_{\alpha \in S} \mathrm{ord}_\alpha(\varphi'(z)),$$

because we have $\alpha \neq \infty$ and $\varphi(\alpha) \neq \infty$ for all $\alpha \in S$. However, it also follows from Theorem 9 that $\text{ord}_\alpha(\varphi'(z))$ is the number of times α is a root of the numerator of $\varphi'(z)$ (we are using the fact that α is not a root of the denominator of $\varphi'(\alpha)$ since $\varphi(\alpha) \neq \infty$). Thus,

$$\sum_{\alpha \in S} \text{ord}_\alpha(\varphi'(z)) = \deg(G(z)F'(z) - F(z)G'(z)).$$

Then, we note that the term with the largest power of z in $G(z)F'(z) - F(z)G'(z)$ is

$$(d-1)b_d a_{d-1} z^{2d-2} - d a_{d-1} b_d z^{2d-2} = -b_d a_{d-1} z^{2d-2} \neq 0.$$

Thus, $\deg(G(z)F'(z) - F(z)G'(z)) = 2d - 2$ from which it follows that

$$\sum_{\alpha \in \mathbb{P}^1(\mathbb{C})} (e_\alpha(\varphi) - 1) = 2d - 2.$$

□

Corollary 20 (Weak Riemann-Hurwitz)**.** *Let $\varphi : \mathbb{P}^1(\mathbb{C}) \to \mathbb{P}^1(\mathbb{C})$ be a rational map of degree $d \geq 1$.*

(1) If $\alpha \in \mathbb{P}^1(\mathbb{C})$ then

$$\sum_{\beta \in \varphi^{-1}(\alpha)} e_\beta(\varphi) = d.$$

(2)

$$2d - 2 = \sum_{\alpha \in \mathbb{P}^1(\mathbb{C})} (d - |\varphi^{-1}(\alpha)|).$$

Proof. We first prove (1). First, by conjugating by a Möbius transformation, we can assume $\alpha \neq \infty$. Next, we can find a Möbius transformation f such that $f(\alpha) = \alpha$ and $f(\infty) \notin \varphi^{-1}(\alpha)$, by Lemma 14. Then, the conjugate φ^f satisfies $\infty \notin (\varphi^f)^{-1}(\alpha)$ since otherwise we would have

$$\varphi(f(\infty)) = f(\alpha) = \alpha,$$

which is impossible. Thus, by replacing φ with φ^f we can assume without loss of generality that $\alpha \neq \infty$ and $\infty \notin \varphi^{-1}(\alpha)$. Now, suppose that $\varphi(z) = F(z)/G(z)$ where F and G share no roots and that $\varphi^{-1}(\alpha) = \{r_1, r_2, \cdots, r_n\}$. This means that $r_1, r_2, \ldots, r_n$ are the solutions to $\varphi(z) = \alpha$, so

$$\frac{F(r_i) - \alpha G(r_i)}{G(r_i)} = 0,$$

since $\alpha \neq \infty$. Since $r_i \neq \infty$, it follows that that r_i is a root of $F(z) - \alpha G(z)$ for all i. Also, note that any root x of $F(z) - \alpha G(z)$ must be in $\varphi^{-1}(\alpha)$. Indeed, if $F(x) - \alpha G(x) = 0$ then x cannot be a root of G since otherwise F and G would have a common root. Thus, we have

$$\frac{F(x) - \alpha G(x)}{G(x)} = 0 \implies \varphi(x) = \alpha.$$

It follows that

$$F(z) - \alpha G(z) = c(z - r_1)^{e_1}(z - r_2)^{e_2} \cdots (z - r_n)^{e_n}$$

for positive integers $e_1, \ldots, e_n$ and a constant c. However, note that

$$e_{r_i}(\varphi) = \operatorname{ord}_{r_i}(\varphi(z) - \varphi(r_i)) = \operatorname{ord}_{r_i}(\varphi(z) - \alpha) = \operatorname{ord}_{r_i}\left(\frac{F(z) - \alpha G(z)}{G_z}\right).$$

By Theorem 9, this is clearly just the number of times $z - r_i$ divides $F(z) - \alpha G(z)$, which is e_i. Thus, we have

$$F(z) - \alpha G(z) = c(z - r_1)^{e_{r_1}(\varphi)}(z - r_2)^{e_{r_2}(\varphi)} \cdots (z - r_n)^{e_{r_n}(\varphi)}.$$

Now, since $\varphi(\infty) \neq \alpha$, we have $\deg(F(z) - \alpha G(z)) = d$ which one can see by simply writing out F and G as polynomials. Thus, taking the degree of both sides, we see that

$$\sum_{\beta \in \varphi^{-1}(\alpha)} e_\beta(\varphi) = d.$$

To prove the second part, we simply use part (1) and the Riemann-Hurwitz formula:

$$\begin{aligned} 2d - 2 &= \sum_{\beta \in \mathbb{P}^1(\mathbb{C})} (e_\beta(\varphi) - 1) \\ &= \sum_{\alpha \in \mathbb{P}^1(\mathbb{C})} \sum_{\beta \in \varphi^{-1}(\alpha)} (e_\alpha(\varphi) - 1) \\ &= \sum_{\alpha \in \mathbb{P}^1(\mathbb{C})} (d - |\varphi^{-1}(\alpha)|). \end{aligned}$$

In the second line we are using the fact that the preimages $\varphi^{-1}(\alpha)$ partition $\mathbb{P}^1(\mathbb{C})$. □

An immediate consequence of part (1) of this is the following.

Corollary 21. *A point $\alpha \in \mathbb{P}^1(\mathbb{C})$ is totally ramified if and only if $\varphi^{-1}(\varphi(\alpha))$ consists of only one point.*

Proof. By part (1) of Corollary 20, we have the inequality

$$e_\alpha(\varphi) \leq \sum_{\beta \in \varphi^{-1}(\varphi(\alpha))} e_\beta(\varphi) = d.$$

Since $e_\beta(\varphi) \geq 1$ for all β, equality holds if and only if $e_\alpha(\varphi)$ is the only term in this sum. In other words, $e_\alpha(\varphi) = d$ if and only if $\varphi^{-1}(\varphi(\alpha))$ consists of only one point. □

In particular, if α is a fixed point, meaning that $\varphi(\alpha) = \alpha$, α is totally ramified if and only if $\varphi^{-1}(\alpha) = \{\alpha\}$.

5. Diophantine Approximation

We now take a detour to establish some results from Diophantine approximation which will be important in studying the integral points in the orbits of rational functions. Diophantine approximation is about how well we can approximate irrational numbers with rational ones. Clearly, any irrational number can be approximated arbitrarily well with rational numbers, since rational numbers are dense in $\mathbb{R}$. We thus study how well we can approximate irrational numbers using rational numbers with a small denominator. A simple result in this is the following.

Theorem 22 (Dirichlet's Approximation Theorem)**.** *Given any irrational number α, there are infinitely many rational numbers x/y such that $x, y \in \mathbb{Z}$ and*

$$\left|\frac{x}{y} - \alpha\right| < \frac{1}{y^2}.$$

Proof. The proof is an application of the pigeonhole principle. For a fixed integer $n \geq 1$, consider the numbers $\{0 \cdot \alpha\}, \{\alpha\}, \{2\alpha\}, \ldots, \{n\alpha\}$, where $\{x\} = x - \lfloor x \rfloor$ is the fractional part of x. This is a sequence of $n+1$ real numbers in the interval $[0,1)$. Consider splitting this interval up into the n intervals $\left[\frac{i}{n}, \frac{i+1}{n}\right)$ for $0 \leq i \leq n-1$. By the pigeonhole principle, two of our numbers must lie in the same interval, say $\{i\alpha\}$ and $\{j\alpha\}$ with $i > j$. It thus follows that the distance from $(i-j)\alpha$ to the nearest integer is less than $1/n$, so there exists an integer x such that

$$|x - (i-j)\alpha| < \frac{1}{n}.$$

Letting $y = i - j \leq n$ and dividing by y, we get

$$\left|\frac{x}{y} - \alpha\right| < \frac{1}{yn} \leq \frac{1}{y^2}.$$

We now show there are infinitely many such rational numbers x/y. Suppose for the sake of contradiction that there are only finitely many such numbers. Note that by the above, for any integer $n \geq 0$ we can find some rational number x/y with $y \leq n$ such that

$$\left|\frac{x}{y} - \alpha\right| < \frac{1}{yn} \leq \frac{1}{y^2}.$$

However, by assumption, there are only finitely many rational numbers x/y which satisfy this. Thus, there must be some rational number x/y such that

$$\left|\frac{x}{y} - \alpha\right| < \frac{1}{yn}$$

for infinitely many values of n, which is clearly a contradiction. □

The next theorem shows that for algebraic numbers (numbers which are the root of some nonzero polynomial with rational coefficients) Dirichlet's approximation theorem is about the best we can do.

Theorem 23 (Roth's Theorem)**.** *Suppose β is an irrational algebraic number and suppose $\epsilon > 0$ is a real number. Then there is some constant $c > 0$, depending on β and ϵ, such that*

$$\left|\frac{a}{b} - \beta\right| \geq \frac{c}{|b|^{2+\epsilon}},$$

for all $a/b \in \mathbb{Q}$.

Unfortunately, the proof of this is too long and difficult to be included here. See [3] for the original proof. We now use this to prove the following theorem.

Theorem 24 (Thue)**.** *Suppose $G(x,y)$ is a homogeneous, integer polynomial of degree d and suppose B is an integer. If $G(x,y)$ has at least three distinct roots in $\mathbb{P}^1(\mathbb{C})$ then the equation*

$$G(x,y) = B$$

has finitely many integer solutions.

Proof. Suppose that G factors over $\mathbb{Q}$ as

$$G(x,y) = k(G_1(x,y))^{e_1}(G_2(x,y))^{e_2}\cdots(G_n(x,y))^{e_n},$$

where k is a constant and each G_i is a homogeneous polynomial which is irreducible over $\mathbb{Q}$. Since each polynomial G_i has rational coefficients, there exists an integer c_i such that $G'_i = c_iG_i$ has integer coefficients. If we let $c = c_1^{e_1}c_2^{e_2}\cdots c_n^{e^n}$, then $G(x,y) = B$ if and only if

$$cB = cG(x,y) = k(G'_1(x,y))^{e_1}(G'_2(x,y))^{e_2}\cdots(G'_n(x,y))^{e_n}.$$

It thus suffices to prove there are only finitely many solutions to this equation. Since each polynomial G'_i has integer coefficients, we must have $G'_1(x,y) \mid cB$ for any integer solution x, y. Therefore it suffices to prove the theorem in the case $G(x,y)$ is an integer polynomial which is irreducible over $\mathbb{Q}$, since this would imply there are only finitely many solutions to $G'_1(x,y) = d$ for each divisor d of cB, and thus there would only be finitely many solutions to $G'_1(x,y) \mid cB$.

We thus assume that $G(x,y)$ is an integer polynomial which is irreducible over $\mathbb{Q}$. Suppose that $G(x,y)$ factors over $\mathbb{C}$ as

$$G(x,y) = c(x-\alpha_1y)(x-\alpha_2y)\cdots(x-\alpha_dy).$$

Since G is irreducible over $\mathbb{Q}$, all of the roots α_i are distinct. Further, we must have $d \geq 3$ since G has at least three distinct roots. Dividing by cy^d, we want to show that there are finitely many solutions to the equation

$$\left(\frac{x}{y}-\alpha_1\right)\left(\frac{x}{y}-\alpha_2\right)\cdots\left(\frac{x}{y}-\alpha_d\right) = \frac{B}{cy^d}.$$

Now, intuitively, the right-hand side of this is very small for large values of y. However, since the roots $\alpha_1, \alpha_2, \ldots, \alpha_d$ are distinct, at most one term on the left-hand side can be small. The idea is that this should imply that the smallest term on the left-hand side must shrink at least as fast as $1/y^d$. More formally, we should be able to find some constant M such that $M/|y|^d$ is always larger than the absolute value of the smallest term on the left-hand side. After this, Roth's theorem will imply that there are only finitely many solutions. The rest of the proof is just a technical argument to prove this formally. Let

$$N = \frac{2(|B/c|)^{1/d}}{\min_{i\neq j}|\alpha_i-\alpha_j|},$$

where the minimum is taken over all pairs $1 \leq i, j \leq d$ such that $i \neq j$. Then, for all solutions (x,y) such that $|y| > N$, we have

$$\frac{|B|}{|c||y|^d} \geq \min_{1\leq i\leq d}\left|\frac{x}{y}-\alpha_i\right|^d.$$

This implies

$$\min_{1\leq i\leq d}\left|\frac{x}{y}-\alpha_i\right| \leq \frac{(|B/c|)^{1/d}}{|y|} \leq \frac{1}{2}\min_{i\neq j}|\alpha_i-\alpha_j|.$$

Let $1 \leq m \leq d$ be the integer such that $|\frac{x}{y} - \alpha_m|$ is minimal. Thus $\left|\frac{x}{y} - \alpha_m\right| \leq \frac{1}{2}\min_{i \neq j}|\alpha_i - \alpha_j|$. Then, for all $i \neq m$,

$$\begin{aligned} \left|\frac{x}{y} - \alpha_i\right| &\geq |\alpha_i - \alpha_m| - \left|\frac{x}{y} - \alpha_m\right| \\ &\geq |\alpha_i - \alpha_m| - \frac{1}{2}\min_{i \neq j}|\alpha_i - \alpha_j| \\ &\geq \frac{|\alpha_i - \alpha_m|}{2}. \end{aligned}$$

Now, for a fixed value of i, let

$$M_i = \prod_{i \neq j} \frac{|a_i - a_j|}{2},$$

and define

$$M = \frac{|B/c|}{\min(M_i)}.$$

We thus have

$$\frac{|B/c|}{|y|^d} \geq \left|\frac{x}{y} - \alpha_m\right| \prod_{i \neq m} \frac{|\alpha_i - \alpha_m|}{2} = \left|\frac{x}{y} - \alpha_m\right| M_m \geq \left|\frac{x}{y} - \alpha_m\right| \frac{|B/c|}{M}.$$

Rearranging, we obtain

$$\min_{1 \leq i \leq d}\left|\frac{x}{y} - \alpha_i\right| \leq \frac{M}{|y|^d},$$

for all solutions x, y such that $|y| > N$. Now, let $0 < \epsilon < 1$ be a real number. Since each α_i is an algebraic number, by Roth's theorem we can find some constant K_i such that

$$\left|\frac{x}{y} - \alpha_i\right| \geq \frac{K_i}{|y|^{2+\epsilon}}$$

for all x, y. Letting $K = \min(K_i)$, we have

$$\min_{1 \leq i \leq d}\left|\frac{x}{y} - \alpha_i\right| \geq \frac{K}{|y|^{2+\epsilon}}$$

for all rational numbers x/y. It follows that for all integers x, y such that $G(x, y) = B$ and $|y| > N$,

$$\frac{K}{|y|^{2+\epsilon}} \leq \min_{1 \leq i \leq d}\left|\frac{x}{y} - \alpha_i\right| \leq \frac{M}{|y|^d}.$$

Thus,

$$|y|^{d-2-\epsilon} \leq \frac{M}{K}.$$

Since $d \geq 3$, this means $|y| \leq (M/K)^{1/(d-2-\epsilon)}$. Thus, for all solutions x, y we have

$$|y| \leq \max\left(N, (M/K)^{1/(d-2-\epsilon)}\right).$$

Clearly this means there are only finitely many possible values of y. Finally, note that for a fixed value of y, the equation $G(x, y) = B$ is a polynomial of degree d in x and thus has at most d possible solutions. Since there are only finitely many possible values for y, this means that there are only finitely many pairs of integers (x, y) such that $G(x, y) = B$, as desired. □

We can now use this to prove the following important theorem about rational functions.

Theorem 25 (Siegel)**.** *Suppose $\varphi(z) \in \mathbb{Q}(z)$ has at least three distinct poles in $\mathbb{P}^1(\mathbb{C})$. Then there are only finitely many rational numbers α such that $\varphi(\alpha)$ is an integer.*

Proof. Suppose that in homogeneous coordinates, $\varphi[X, Y] = [F(X, Y), G(X, Y)]$ where F and G are homogeneous polynomials of degree d with rational coefficients and with no common root. By scaling F and G by a constant if necessary we can assume F and G are integer polynomials. Note that G has at least three distinct roots, since φ has at least three distinct poles. Any rational number $\frac{a}{b}$ with $\gcd(a, b)$ has homogeneous coordinates $[a, b]$, so

$$\varphi\left(\frac{a}{b}\right) = \frac{F(a, b)}{G(a, b)}.$$

It follows that $\varphi\left(\frac{a}{b}\right)$ is an integer if and only if $G(a, b)$ divides $F(a, b)$. The idea is to show that if $G(a, b)$ divides $F(a, b)$ then $G(a, b)$ must divide some constant, after which we can apply Theorem 24. To do this, we need to construct something called the resultant of F and G.

For each integer $n \geq 1$, let $\mathbb{Q}[X, Y]_n$ denote the set of homogeneous polynomials in two variables X and Y with rational coefficients and degree n. Clearly $\mathbb{Q}[X, Y]_n$ is a vector space over $\mathbb{Q}$ with basis $\{X^n, X^{n-1}Y, \ldots, XY^{n-1}, Y^n\}$. Now, we define a map $\phi : \mathbb{Q}[X, Y]_{d-1} \times \mathbb{Q}[X, Y]_{d-1} \to \mathbb{Q}[X, Y]_{2d-1}$ by

$$\phi(C(X, Y), D(X, Y)) = C(X, Y)F(X, Y) + D(X, Y)G(X, Y).$$

Noting that $\mathbb{Q}[X, Y]_{d-1} \times \mathbb{Q}[X, Y]_{d-1}$ is also a vector space, we see that ϕ is a linear map of vector spaces. In fact, if

$$\begin{aligned} F(X, Y) &= a_0X^d + a_1X^{d-1}Y + \cdots + a_{d-1}XY^{d-1} + a_dY^d, \\ G(X, Y) &= b_0X^d + b_1X^{d-1}Y + \cdots + b_{d-1}XY^{d-1} + b_dY^d, \\ C(X, Y) &= C_0X^{d-1} + C_1X^{d-2}Y + \cdots + C_{d-2}XY^{d-2} + C_{d-1}Y^{d-1}, \end{aligned}$$

and

$$D(X, Y) = D_0X^{d-1} + D_1X^{d-2}Y + \cdots + D_{d-2}XY^{d-2} + D_{d-1}Y^{d-1}$$

then after choosing the obvious bases, ϕ is represented by the matrix multiplication

$$\phi(C(X, Y), D(X, Y)) = \begin{pmatrix} a_0 & 0 & \cdots & 0 & b_0 & 0 & \cdots & 0 \\ a_1 & a_0 & \cdots & 0 & b_1 & b_0 & \cdots & 0 \\ a_2 & a_1 & \cdots & 0 & b_2 & b_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{d-1} & a_{d-2} & \cdots & a_0 & b_{d-1} & b_{d-2} & \cdots & b_0 \\ a_d & a_{d-1} & \cdots & a_1 & b_d & b_{d-1} & \cdots & b_1 \\ 0 & a_d & \cdots & a_2 & 0 & b_d & \cdots & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_d & 0 & 0 & \cdots & b_d \end{pmatrix} \begin{pmatrix} C_0 \\ C_1 \\ \vdots \\ C_{d-1} \\ D_0 \\ D_1 \\ \vdots \\ D_{d-1} \end{pmatrix}.$$

Let M denote this matrix and note that the entries of M are integers. We define the resultant of F and G to be the quantity $\mathrm{Res}(F, G) = R = \det(M)$, which is an integer. We claim that this is nonzero. Indeed, if $R = 0$ then the kernel of ϕ is nontrivial, so we can find polynomials $C(X, Y)$ and $D(X, Y)$ such that $C(X, Y)F(X, Y) + D(X, Y)G(X, Y) = 0$. Rearranging, this becomes

$$C(X, Y)F(X, Y) = -D(X, Y)G(X, Y).$$

The d roots of $F(X, Y)$ must also be roots of the right-hand side. However, $D(X, Y)$ has degree $d-1$ and thus has only $d-1$ roots. It follows that $G(X, Y)$ must share a root with $F(X, Y)$, which is a contradiction. Thus we conclude that $R \neq 0$ and M is invertible. The inverse matrix M^{-1} does not necessarily consist of integer entries. However, the adjoint matrix $M_{\text{adj}} = RM^{-1}$ does consist of integer entries, because M has only integer entries. We can thus let $C_1(X, Y)$ and $D_1(X, Y)$ be polynomials with integer coefficients $C_0, C_1, \ldots, C_{d-1}$ and $D_0, D_1, \ldots, D_{d-1}$ such that

$$\begin{pmatrix} C_0 \\ C_1 \\ \vdots \\ C_{d-1} \\ D_0 \\ D_1 \\ \vdots \\ D_{d-1} \end{pmatrix} = M_{\text{adj}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

and we can let $C_2(X, Y)$ and $D_2(X, Y)$ be polynomials with integer coefficients $C'_0, C'_1, \ldots, C'_{d-1}$ and $D'_0, D'_1, \ldots, D'_{d-1}$ such that

$$\begin{pmatrix} C'_0 \\ C'_1 \\ \vdots \\ C'_{d-1} \\ D'_0 \\ D'_1 \\ \vdots \\ D'_{d-1} \end{pmatrix} = M_{\text{adj}} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Then,

$$\phi(C_1(X, Y), D_1(X, Y)) = M \begin{pmatrix} C_0 \\ C_1 \\ \vdots \\ C_{d-1} \\ D_0 \\ D_1 \\ \vdots \\ D_{d-1} \end{pmatrix} = MM_{\text{adj}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} R \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

and

$$\phi(C_2(X,Y),D_2(X,Y)) = M\begin{pmatrix} C_0' \\ C_1' \\ \vdots \\ C_{d-1}' \\ D_0' \\ D_1' \\ \vdots \\ D_{d-1}' \end{pmatrix} = MM_{\mathrm{adj}}\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ R \end{pmatrix}.$$

It follows that

$$C_1(X,Y)F(X,Y) + D_1(X,Y)G(X,Y) = RX^{2d-1}$$

and

$$C_2(X,Y)F(X,Y) + D_2(X,Y)G(X,Y) = RY^{2d-1}.$$

Therefore, if $\frac{a}{b}$ is a rational number in lowest terms such that $\varphi\left(\frac{a}{b}\right) \in \mathbb{Z}$, and thus $G(a,b) \mid F(a,b)$, we can plug in $(X,Y) = (a,b)$ into the above equation to get

$$G(a,b) \mid C_1(a,b)F(a,b) + D_1(a,b)G(a,b) = Ra^{2d-1}$$

and

$$G(a,b) \mid C_2(a,b)F(a,b) + D_2(a,b)G(a,b) = Rb^{2d-1}.$$

It follows that $G(a,b) \mid \gcd(Ra^{2d-1}, Rb^{2d-1})$. Since a and b are relatively prime, this greatest common divisor is just R, so $G(a,b)$ divides R. However, there are only finitely many divisors of R, say $d_1, d_2, \ldots, d_r$. By Theorem 24, the equation

$$G(a,b) = d_i$$

has finitely many integer solutions for all i, since G has at least three distinct roots. There are thus only finitely many pairs of integers (a,b) such that $G(a,b) \mid R$. Finally, it follows that there are only finitely many rational numbers $\frac{a}{b}$ such that $\varphi\left(\frac{a}{b}\right)$ is an integer. □

6. Polynomials and Integral Points in Orbits

Next, we will study rational functions which have an iterate which is a polynomial. However, since we are working in $\mathbb{P}^1(\mathbb{C})$, the most natural properties to study are those preserved under conjugation by Möbius transformations. Polynomials are not necessarily preserved under conjugation by Möbius transformations, so it is often useful to study maps which are either a polynomial or conjugate to a polynomial. There is an important criteria for such functions.

Theorem 26. *A rational map $\varphi : \mathbb{P}^1(\mathbb{C}) \to \mathbb{P}^1(\mathbb{C})$ is a polynomial or a conjugate of a polynomial if and only if φ has a totally ramified fixed point; i.e. there exists some $\alpha \in \mathbb{P}^1(\mathbb{C})$ such that $\varphi(\alpha) = \alpha$ and $e_\alpha(\varphi) = \deg(\varphi) = d$.*

Proof. We first note that φ is a polynomial if and only if ∞ is a totally ramified fixed point. Indeed, the points z such that $\varphi(z) = \infty$ are exactly the roots of the denominator of φ, possibly along with ∞. By Corollary 21, it follows that ∞ is a totally ramified fixed point if and only if there are no roots of the denominator of φ, which is the same as saying that φ is a polynomial. Since the ramification index is preserved under conjugation, it immediately follows that if φ is a conjugate of a polynomial then φ has a totally ramified fixed point.

To prove the converse, suppose that φ has a totally ramified fixed point at α, and let f be a Möbius transformation such that $f(\infty) = \alpha$. We then have

$$e_\infty(\varphi^f) = e_\alpha(\varphi) = d.$$

Also, it is easy to see that $\varphi^f(\infty) = \infty$. Thus, φ^f has a totally ramified fixed point at ∞ and is therefore a polynomial. Since φ is a conjugate of φ^f, it follows that φ is conjugate to a polynomial. □

By Corollary 21, this means that φ is a conjugate of a polynomial if and only if $\varphi^{-1}(\alpha) = \{\alpha\}$ for some point α.

An important type of set in dynamics is one which is invariant under the map φ. As Theorem 28 shows, for rational functions, if such a set is finite it is quite simple.

Definition 27. A set E such that $\varphi^{-1}(E) = E = \varphi(E)$ is called an *invariant set* of φ. If E is also finite, it is called an *exceptional set*.

Theorem 28. *Suppose $\varphi : \mathbb{P}^1(\mathbb{C}) \to \mathbb{P}^1(\mathbb{C})$ is a rational map of degree at least 2 and let E be an exceptional set for φ. Then $|E| \leq 2$.*

Proof. Since $\varphi^{-1}(E) = E$, it follows that φ must permute the elements of E. Since E is finite, this implies that for some integer $n \geq 1$, $\varphi^n(x) = x$ for all $x \in E$. Now, suppose that φ^n has degree d. We must have $d \geq 2$ since $\deg(\varphi) \geq 2$. For each point $z \in E$, we must have $(\varphi^n)^{-1}(z) \subset E$, since E is invariant under φ. However, for all $y \in E$ such that $y \neq x$, $\varphi^n(y) = y \neq x$. Thus, $(\varphi^n)^{-1}(z) = \{z\}$ and in particular, $|(\varphi^n)^{-1}(z)| = 1$. By the Weak Riemann-Hurwitz Theorem (Corollary 20), we have

$$\begin{aligned} 2d - 2 &= \sum_{\alpha \in \mathbb{P}^1(\mathbb{C})} (d - |\varphi^{-1}(\alpha)|) \\ &\geq \sum_{\alpha \in E} (d - |\varphi^{-1}(\alpha)|) \\ &= \sum_{\alpha \in E} (d - 1) \\ &= |E|(d-1). \end{aligned}$$

Since $d \geq 2$, this implies $|E| \leq 2$. □

Theorem 29. *Suppose φ is a rational function of degree $d \geq 2$ and φ^n is conjugate to a polynomial for some integer $n \geq 1$. Then φ^2 is already conjugate to a polynomial.*

Proof. By Theorem 26, φ^n has a totally ramified fixed point, say α. This implies $(\varphi^n)^{-1}(\alpha) = \{\alpha\}$. Now, it follows that $\varphi^{-1}(\varphi^i(\alpha)) = \{\varphi^{i-1}(\alpha)\}$ for all $1 \leq i \leq n$. Indeed, if the preimage $\varphi^{-1}(\varphi^i(\alpha))$ contained some point $x \neq \varphi^{i-1}(\alpha)$, then any point $y \in \varphi^{-(i-1)}(x)$ would satisfy $\varphi^n(y) = \alpha$, and $\alpha \neq y$, which contradicts $(\varphi^n)^{-1}(\alpha) = \{\alpha\}$. It then follows that the set

$$E = \{\alpha, \varphi(\alpha), \varphi^2(\alpha), \ldots, \varphi^{n-1}(\alpha)\}$$

is an exceptional set. Thus, by Theorem 28, $|E| \leq 2$ and hence we have either $\varphi(\alpha) = \alpha$ or $\varphi(\alpha) \neq \alpha$ and $\varphi^2(\alpha) = \alpha$. In the first case, we have $\varphi^{-1}(\alpha) = \{\varphi^{n-1}(\alpha)\} = \{\alpha\}$, so α is a totally ramified fixed point. Thus φ is conjugate to a polynomial by Theorem 26. In the second case, note that $\varphi^n(\alpha) = \alpha$, so n must be even. Then, we have

$$\varphi^{-2}(\alpha) = \varphi^{-1}(\varphi^{-1}(\alpha)) = \varphi^{-1}(\varphi^{n-1}(\alpha)) = \varphi^{-1}(\varphi(\alpha)) = \{\alpha\}.$$

Thus, α is a totally ramified fixed point of φ^2, so φ^2 is conjugate to a polynomial by Theorem 26. □

In fact, by the exact same proof as above but in the special case $\alpha = \infty$, we have the following.

Corollary 30. *Suppose φ is a rational function of degree $d \geq 2$ and φ^n is a polynomial for some integer $n \geq 1$. Then φ^2 is already a polynomial.*

Finally, we can prove the following theorem.

Theorem 31. *If φ is a rational function of degree $d \geq 2$ and φ^2 is not polynomial, then for all $a \in \mathbb{Q}$, $\mathcal{O}_\varphi(a)$ contains finitely many integers.*

Proof. By Corollary 30, no iterate of φ is a polynomial. We first show that $|\varphi^{-4}(\infty)| \geq 3$ and hence φ^4 has at least 3 distinct poles. If $|\varphi^{-3}(\infty)| \geq 3$ then this is true because $|\varphi^{-4}(\infty)| = |\varphi^{-1}(\varphi^{-3}(\infty))| \geq |\varphi^{-3}(\infty)|$ (because $\varphi^{-1}(z)$ is nonempty for all z). Thus suppose $|\varphi^{-3}(\infty)| \leq 2$. We now consider several cases. In each case, we will prove that $d \leq 2$ and then that $|\varphi^{-4}(\infty)| \geq 3$. Note that if $d \leq 2$ and then $d = 2$ by assumption. If this is the case, then we claim that there are at most two points α such that $\varphi^{-1}(\alpha)$ consists of one point. Indeed, by the Weak Riemann-Hurwitz Theorem (Corollary 20),

$$2 = \sum_{\alpha \in \mathbb{P}^1(\mathbb{C})} (2 - |\varphi^{-1}(\alpha)|).$$

So there are at most 2 points α such that $2 - |\varphi^{-1}(\alpha)| > 0 \iff |\varphi^{-1}(\alpha)| = 1$.

Case 1: $|\varphi^{-3}(\infty)| = 1$. This means that there exists exactly one point P such that $\varphi^3(P) = \infty$. Let $\varphi(P) = Q$ and $\varphi(Q) = R$. We thus have $\varphi^{-1}(\infty) = \{R\}$, $\varphi^{-1}(R) = \{Q\}$, and $\varphi^{-1}(Q) = \{P\}$. We note that R, Q, and ∞ are distinct. If $R = \infty$ then we would have $\varphi(\infty) = \infty$, so ∞ would be a totally ramified fixed point. However, φ is not a polynomial so ∞ cannot be a totally ramified fixed point (Theorem 26). $R \neq Q$ since $R = Q$ would imply $\varphi(R) = \varphi(Q) \implies \infty = R$. Similarly, $Q \neq \infty$ since otherwise $\varphi^2(\infty) = \infty$ and hence ∞ is a totally ramified fixed point of φ^2, which is impossible since φ^2 is not a polynomial. Thus, by the Weak Riemann-Hurwitz theorem (Corollary 20)

$$\begin{aligned} 2d - 2 &= \sum_{\alpha \in \mathbb{P}^1(\mathbb{C})} (d - |\varphi^{-1}(\alpha)|) \\ &\geq (d - |\varphi^{-1}(\infty)|) + (d - |\varphi^{-1}(R)|) + (d - |\varphi^{-1}(Q)|) \\ &= 3d - 3. \end{aligned}$$

We thus have $d \leq 1$, which is a contradiction, so this case is impossible.

Case 2: $|\varphi^{-3}(\infty)| = 2$ and $|\varphi^{-2}(\infty)| = 1$. Let $\varphi^{-3}(\infty) = \{P, P'\}$, and define $Q = \varphi(P) = \varphi(P')$ and $R = \varphi(Q)$. Note that $\varphi^{-1}(\infty) = \{R\}$, $\varphi^{-1}(R) = \{Q\}$, and $\varphi^{-1}(Q) = \{P, P'\}$. By the exact same argument as in case 1, P, Q, and ∞ are distinct points. Then, by the Weak Riemann-Hurwitz theorem,

$$\begin{aligned} 2d-2 &= \sum_{\alpha \in \mathbb{P}^1(\mathbb{C})} (d - |\varphi^{-1}(\alpha)|) \\ &\geq (d - |\varphi^{-1}(\infty)|) + (d - |\varphi^{-1}(R)|) + (d - |\varphi^{-1}(Q)|) \\ &= 3d - 4. \end{aligned}$$

So $d \leq 2$. Further, because $\varphi^{-1}(\infty) = \{R\}$ and $\varphi^{-1}(R) = \{Q\}$, there are no other points α such that $|\varphi^{-1}(\alpha)| = 1$. P and P' are not equal to R and not equal to each other, so at least one of them must be distinct from both R and ∞, say P. Then, $|\varphi^{-1}(P)| = 2$. Thus,

$$|\varphi^{-4}(\infty)| = |\varphi^{-1}(\{P, P'\})| = |\varphi^{-1}(P)| + |\varphi^{-1}(P')| \geq 2 + 1 = 3,$$

as desired.

Case 3: $|\varphi^{-3}(\infty)| = |\varphi^{-2}(\infty)| = 2$ and $|\varphi^{-1}(\infty)| = 1$. Let $\varphi^{-3}(\infty) = \{P, P'\}$, and define $Q = \varphi(P)$, $Q' = \varphi(P')$, and $R = \varphi(Q) = \varphi(Q')$. Note that $\varphi^{-1}(\infty) = \{R\}$, $\varphi^{-1}(R) = \{Q, Q'\}$, and $\varphi^{-1}(Q) = \{P\}$. Like the previous cases, $R \neq \infty$. Also, neither Q nor Q' are equal to R, since this would imply $R = \infty$. Since $Q \neq Q'$, at least one of Q and Q' are not equal to infinity. By relabeling if necessary, we may assume that $Q \neq \infty$ and thus Q, R, and ∞ are distinct. By the Weak Riemann-Hurwitz theorem,

$$\begin{aligned} 2d-2 &= \sum_{\alpha \in \mathbb{P}^1(\mathbb{C})} (d - |\varphi^{-1}(\alpha)|) \\ &\geq (d - |\varphi^{-1}(\infty)|) + (d - |\varphi^{-1}(R)|) + (d - |\varphi^{-1}(Q)|) \\ &= 3d - 4. \end{aligned}$$

So $d \leq 2$. Because $\varphi^{-1}(\infty) = \{R\}$ and $\varphi^{-1}(Q) = \{P\}$, for any $\alpha \neq \infty, Q$, $|\varphi^{-1}(\alpha)| \geq 2$. Noting that neither P nor P' can be equal to Q, at least one of $|\varphi^{-1}(P)|$ and $|\varphi^{-1}(P')|$ must be equal to 2. Thus,

$$|\varphi^{-4}(\infty)| = |\varphi^{-1}(\{P, P'\})| = |\varphi^{-1}(P)| + |\varphi^{-1}(P')| \geq 3,$$

as desired.

Case 4: $|\varphi^{-3}(\infty)| = |\varphi^{-2}(\infty)| = |\varphi^{-1}(\infty)| = 2$. Let $\varphi^{-3}(\infty) = \{P, P'\}$, and define $Q = \varphi(P)$, $Q' = \varphi(P')$, $R = \varphi(Q)$, and $R' = \varphi(Q')$. Note that $\varphi^{-1}(\infty) = \{R, R'\}$, $\varphi^{-1}(R) = \{Q\}$, and $\varphi^{-1}(Q) = \{P\}$. Since $|\varphi^{-1}(R)| = 1$ and $|\varphi^{-1}(\infty)| = 2$, $R \neq \infty$. Similarly $R' \neq \infty$. This implies $Q \neq R$ and $Q' \neq R'$. Since $Q \neq Q'$, one of these is not equal to ∞, so we can assume $Q \neq \infty$ by relabeling if necessary. Then, Q, R, and ∞ are distinct. By the Weak Riemann-Hurwitz theorem,

$$\begin{aligned} 2d-2 &= \sum_{\alpha \in \mathbb{P}^1(\mathbb{C})} (d - |\varphi^{-1}(\alpha)|) \\ &\geq (d - |\varphi^{-1}(\infty)|) + (d - |\varphi^{-1}(R)|) + (d - |\varphi^{-1}(Q)|) \\ &= 3d - 4. \end{aligned}$$

So $d \leq 2$. We have $\varphi^{-1}(R) = \{Q\}$ and $\varphi^{-1}(Q) = \{P\}$, so we must have $|\varphi^{-1}(\alpha)| = 2$ for all $\alpha \neq Q, R$. Neither of P and P' can be equal to Q, so at least one of them is not equal to Q or R and thus

$$|\varphi^{-4}(\infty)| = |\varphi^{-1}(\{P, P'\})| = |\varphi^{-1}(P)| + |\varphi^{-1}(P')| \geq 3,$$

as desired.

We have thus shown that in all cases, $|\varphi^{-4}(\infty)| \geq 3$, which is the same as saying that the rational function φ^4 has at least 3 distinct poles. We can now apply Siegel's Theorem (Theorem 25) on φ^4 to see that the set

$$S = \{\alpha \in \mathbb{Q} \mid \varphi^4(\alpha) \in \mathbb{Z}\}$$

is finite. Next, we note that for all $\alpha \in \mathbb{Q}$, if $\varphi^n(\alpha)$ is an integer and $n \geq 4$, then we can write

$$\varphi^n(\alpha) = \varphi^4(\varphi^{n-4}(\alpha)),$$

so $\varphi^{n-4}(\alpha) \in S$. This means that all integers m in the orbit $\mathcal{O}_\varphi(\alpha)$ must either be equal to $\alpha, \varphi(\alpha)$, $\varphi^2(\alpha)$, $\varphi^3(\alpha)$, or $\varphi^4(s)$ for some s in S. Since S is finite, it follows that there can only be finitely many integers in the orbit $\mathcal{O}_\varphi(\alpha)$, which is what we wanted to show. □

Example 32. Consider the orbit of 2 under the rational function

$$\varphi(z) = \frac{2z^2 - 2z + 1}{4z^2 - 4z + 1}.$$

We have

$$\begin{aligned}
\varphi(2) &= \frac{5}{9} \\
\varphi^2(2) &= 41 \\
\varphi^3(2) &= \frac{3281}{6561} \\
\varphi^4(2) &= 21523361 \\
\varphi^5(2) &= \frac{926510094425921}{1853020188851841} \\
\varphi^6(2) &= 1716841910146256242328924544641,
\end{aligned}$$

and so on. It appears as though every other term in this orbit is an integer and thus there are infinitely many integers in the orbit of 2. Indeed, this pattern does hold, which by Theorem 31 means that some iterate of φ is a polynomial. In fact, we have

$$\varphi^2(z) = 8z^4 - 16z^3 + 12z^2 - 4z + 1.$$

By Corollary 30, this is actually the only way this can happen; φ^2 must be a polynomial since φ is not.

References

[1] A.F. Beardon. *Iteration of Rational Functions: Complex Analytic Dynamical Systems*. Graduate Texts in Mathematics. Springer New York, 2000. ISBN: 9780387951515. URL: https://books.google.com/books?id=wAhFqe0AWRQC.

[2] Serge Lang. *Algebra*. Graduate Texts in Mathematics. Springer New York, 2002. ISBN: 978-0-387-95385-4. URL: https://link.springer.com/book/10.1007/978-1-4613-0041-0.

[3] K. F. Roth. “Rational approximations to algebraic numbers”. In: *Mathematika* 2.1 (1955), pp. 1–20. DOI: https://doi.org/10.1112/S0025579300000644. eprint: https://londmathsoc.onlinelibrary.wiley.com/doi/pdf/10.1112/S0025579300000644. URL: https://londmathsoc.onlinelibrary.wiley.com/doi/abs/10.1112/S0025579300000644.

[4] E.B. Saff and A.D. Snider. *Fundamentals of Complex Analysis with Applications to Engineering and Science*. Prentice Hall, 2003. ISBN: 9780139078743. URL: https://books.google.com/books?id=fVsZAQAAIAAJ.

[5] Joseph H. Silverman. *The Arithmetic of Dynamical Systems*. Graduate Texts in Mathematics. Springer New York, 2010. ISBN: 978-0-387-69904-2. URL: https://link.springer.com/book/10.1007/978-0-387-69904-2.