SMALLEST REPRESENTATIVES OF $\mathrm{SL}(2,\mathbb{Z})$-ORBITS OF BINARY FORMS AND ENDOMORPHISMS OF $\mathbb{P}^1$

BENJAMIN HUTZ AND MICHAEL STOLL

ABSTRACT. We develop an algorithm that determines, for a given squarefree binary form F with real coefficients, a smallest representative of its orbit under $\mathrm{SL}(2,\mathbb{Z})$, either with respect to the Euclidean norm or with respect to the maximum norm of the coefficient vector. This is based on earlier work of Cremona and Stoll [SC03]. We then generalize our approach so that it also applies to the problem of finding an integral representative of smallest height in the $\mathrm{PGL}(2,\mathbb{Q})$ conjugacy class of an endomorphism of the projective line. Having a small model of such an endomorphism is useful for various computations.

1 Introduction

Let $F = a_0x^n + a_1x^{n-1}y + \ldots + a_ny^n$ be a binary form with real coefficients. We define its *size* to be

$$\|F\| = a_0^2 + a_1^2 + \ldots + a_n^2$$

(this is the squared Euclidean norm of the coefficient vector) and its *height* to be the maximum norm

$$H_\infty(F) = \max\{|a_0|, |a_1|, \ldots, |a_n|\}.$$

If F has coefficients in $\mathbb{Z}$ with $\gcd(a_0, \ldots, a_n) = 1$, then $H_\infty(F) = H(F)$ is the (multiplicative) global height of F in the sense that it is the global height of the coefficient vector of F, considered as a point in projective space $\mathbb{P}^n$. If F has coefficients in $\mathbb{Q}$, then $H(F) = H_0(F)H_\infty(F)$ with

$$H_0(F) = \prod_{p \text{ prime}} \max\{|a_0|_p, |a_1|_p, \ldots, |a_n|_p\}.$$

Since $H_0(F)$ does not change under the action of $\mathrm{SL}(2,\mathbb{Z})$, we can use $H_\infty(F)$ as a proxy for the global height $H(F)$ for our purposes.

Our goal in this note will be to find, for a given F (without multiple factors, say), a smallest representative F_0 in its $\mathrm{SL}(2,\mathbb{Z})$-orbit, in the sense that $\|F_0\|$ is minimal among all forms in the orbit of F, or in the sense that $H_\infty(F_0)$ (or equivalently, $H(F_0)$ when F_0 has coefficients in $\mathbb{Q}$) is minimal within the orbit.

More generally, we may want to consider some kind of geometric object Φ related to the projective line (over $\mathbb{Q}$, say), given by some polynomials with respect to some chosen coordinates on $\mathbb{P}^1$. Then we usually can associate to Φ (in a natural, i.e., coordinate-independent way) a finite set of points on $\mathbb{P}^1$, or equivalently, a binary form $F = F(\Phi)$.

Date: May 22, 2018.
2010 *Mathematics Subject Classification.* 37P05, 37P45, 11C08, 11Y99 .

"Natural" means that this association is compatible with the action of $\mathrm{PGL}(2) = \mathrm{Aut}(\mathbb{P}^1)$ on both sides. In this way, we can set up a "reduction theory" for the objects Φ by selecting a suitable "reduced" representative $F \cdot \gamma$ in the $\mathrm{SL}(2,\mathbb{Z})$-orbit of F and declaring $\Phi \cdot \gamma$ to be the reduced representative in the $\mathrm{SL}(2,\mathbb{Z})$-orbit of Φ. This is the approach taken in [Sto11] in the setting of $\mathbb{P}^n$ for general n. This works reasonably well when we just want to have a way of selecting a canonical representative of moderate size. If we want to find a representative of smallest height, then more work is required: we have to relate the height of Φ to the size of $F(\Phi)$ and determine a bound on the distance we can move from the canonical representative without increasing the height. The application we have in mind is to endomorphisms of $\mathbb{P}^1$; this is discussed in some detail in Section 2 below.

The plan of this paper is as follows. After the discussion of endomorphisms of $\mathbb{P}^1$ in Section 2, we recall the reduction theory for binary forms as developed in [SC03] in Section 3. This provides us with the canonical representative of an $\mathrm{SL}(2,\mathbb{Z})$-orbit of (real) binary forms, as follows. We associate to a binary form F a point $z(F)$ in the upper half-plane; this association is $\mathrm{SL}(2,\mathbb{R})$-equivariant. Then F is the canonical representative if $z(F)$ is in the standard fundamental domain for the action of $\mathrm{SL}(2,\mathbb{Z})$ on the upper half-plane. Section 4 is the heart of the paper. Here we relate the size $\|F\|$ of a binary form F to its "Julia invariant" $\theta(F)$ and the hyperbolic distance of $z(F)$ to the point i (note that we will denote this point mostly by j instead). This allows us to solve the problem of minimizing the size with respect to the Euclidean norm within an $\mathrm{SL}(2,\mathbb{Z})$-orbit of binary forms. Since this size and the height are fairly closely related, we also obtain a way of minimizing the height. Section 5 spells out the resulting algorithm in some detail and gives an example. Section 6 finally applies this method to the problem of finding a model of minimal height of a given endomorphism of $\mathbb{P}^1$; this is demonstrated with another example. Finally, in Section 7, we answer some questions concerning the number of distinct $\mathrm{GL}(2,\mathbb{Z})$-orbits of integral models with the same (minimal) resultant, posed by Bruin and Molnar [BM12].

2 Endomorphisms of the projective line

Let Hom_d be the space of degree d endomorphisms of $\mathbb{P}^1$. There is a natural action of $\mathrm{PGL}(2)$ by conjugation on Hom_d. We define the quotient space for this action as $\mathcal{M}_d = \mathrm{Hom}_d \,/\, \mathrm{PGL}(2)$. Levy [Lev11] proved that $\mathcal{M}_d$ exists as a geometric quotient, which we call the *moduli space of dynamical systems of degree* d *morphisms of* $\mathbb{P}^1$. For $f \in \mathrm{Hom}_d$, we denote the corresponding element of the moduli space as $[f] \in \mathcal{M}_d$. After choosing coordinates, we can write f as a pair of degree d homogeneous polynomials, which we call a *model* for $[f]$. Given $\alpha \in \mathrm{PGL}(2)$, we denote the conjugate as $f^\alpha = \alpha^{-1} \circ f \circ \alpha$. When working over an infinite field, there are infinitely many models for any conjugacy class. The choice of a "good" model can affect a variety of properties and algorithms. For example, the *field of definition* of a model f is the smallest field containing the coefficients of f. Different models of $[f] \in \mathcal{M}_d$ may have different fields of definition. The fields of definition are studied in [HM14, Sil95]. Here we are considering models defined over $\mathbb{Q}$. In this context, an *integral model* of $[f]$ is a model of f consisting of a pair of polynomials with coefficients in $\mathbb{Z}$. Any model over $\mathbb{Q}$ can be scaled to give an integral model with the property that the coefficients of the two polynomials are coprime.

Given an integral model f, we define the *resultant* $\mathrm{Res}(f)$ to be the resultant of its two defining polynomials considered both as having degree d. When working over a global field, reducing modulo primes can yield information on the arithmetic dynamics of f. In particular, this local-global transfer of information is a key piece of an algorithm to determine all rational preperiodic points for a given map [Hut15]. However, not all primes can be used. A *prime of good reduction* is a prime $\mathfrak{p}$ such that reduction modulo $\mathfrak{p}$ commutes with iteration. For $f \in \mathrm{Hom}_d$, the primes of good reduction are the primes that do not divide the resultant. Consequently, the problem of finding an integral model of $[f] \in \mathcal{M}_d$ with smallest (norm of the) resultant is important. Such a model is called a *minimal model for* $[f]$, and an algorithm to determine a minimal model is given by Bruin and Molnar [BM12]. This algorithm is implemented in the computer algebra system Sage [SJ05]. It is important to note that minimal models are, in general, not unique. For example, conjugating by an element of $SL(2, \mathbb{Z})$ leaves the resultant unchanged. Consequently, there are infinitely many minimal models for a given $[f]$. Furthermore, we prove in Proposition 7.4 that there are maps with arbitrarily many distinct $SL(2, \mathbb{Z})$-orbits of minimal models. Choosing a "best" minimal model depends on the application in mind. For example, when working with polynomials, it is conventional to move the totally ramified fixed point to the point at infinity. Similarly, there are models that are convenient for working with critical points [Ing12] or multipliers [Mil93]. The focus of this article is a minimal model of minimal height. For a model $f = [F : G] \in \mathrm{Hom}_d$ with homogeneous polynomials F and G, we define the *height of* f as

$$H(f) = H(F, G)\,,$$

i.e., the height of the concatenated coefficient vectors. Note that this height does not change if we replace $[F : G]$ by $[\lambda F : \lambda G]$.

There are a number of properties and algorithms where bounds are dependent on $H(f)$. For example, in the algorithm to determine all rational preperiodic points [Hut15] an upper bound on the height of a rational preperiodic point is determined that depends on $H(f)$. Furthermore, the minimal height over the models with minimal resultant defines a height on the moduli space $\mathcal{M}_d$.

See [Sil07, Conjecture 4.98] for a dynamical version of Lang's height conjecture related to heights on $\mathcal{M}_d$.

Definition 2.1. Given $[f] \in \mathcal{M}_d$, we call g a *reduced model of* $[f]$ if g is a minimal model for $[f]$ with smallest height $H(g)$.

We will use the ideas of [SC03] together with the new bounds from Section 4 to devise an algorithm that finds a model of smallest height in the $SL(2, \mathbb{Z})$-orbit of a given model. Together with an extension of the algorithm from [BM12] that finds a representative in each $GL(2, \mathbb{Z})$-orbit of minimal models of a given $[f]$, this results in a procedure that produces a reduced model of any given $[f] \in \mathcal{M}_d$. Note that the minimal height in the $GL(2, \mathbb{Z})$-orbit of a given model is the same as the minimal height in its $SL(2, \mathbb{Z})$-orbit, so it is sufficient to apply our reduction algorithm to a representative of each $GL(2, \mathbb{Z})$-orbit of minimal models of $[f]$.

3 Reduction of binary forms

We recall the reduction theory of binary forms, as described in [SC03] (following earlier work of Hermite [Her48] and Julia [Jul17]). In the following, the degree n of the form is always assumed to be at least 3. The space of binary forms of degree n over a ring R is denoted $R[x, y]_n$.

We take a more general approach than in the introduction and consider a binary form with complex (instead of real) coefficients

$$F = a_0 x^n + a_1 x^{n-1} y + \ldots + a_n y^n \in \mathbb{C}[x, y]_n \,.$$

In this context, the *size* $\|F\|$ of F is defined as

$$\|F\| = |a_0|^2 + |a_1|^2 + \ldots + |a_n|^2 = \int_0^1 \left|F(e^{2\pi i \phi}, 1)\right|^2 \, d\phi \,.$$

The groups $\mathrm{GL}(2, \mathbb{C})$ and $\mathrm{SL}(2, \mathbb{C})$ act on the space $\mathbb{C}[x, y]_n$ of binary forms of degree n via linear substitution of the variables; this is an action on the right. Concretely,

$$F(x, y) \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = F(ax + by, cx + dy) \,.$$

We write $\mathcal{H}_3 = \mathbb{C} \times \mathbb{R}_{>0}$ for three-dimensional hyperbolic space in the upper half-space model; we write $t + uj$ for the point $(t, u) \in \mathbb{C} \times \mathbb{R}_{>0} = \mathcal{H}_3$. We identify the hyperbolic upper half-plane $\mathcal{H}$ with the subset $\mathbb{R} \times \mathbb{R}_{>0}$ of $\mathcal{H}_3$. There is a standard left action of $\mathrm{SL}(2, \mathbb{C})$ on $\mathcal{H}_3$ that restricts to the standard action of $\mathrm{SL}(2, \mathbb{R})$ on $\mathcal{H}$ via Möbius transformations. The standard fundamental domain of the action of $\mathrm{SL}(2, \mathbb{Z})$ on $\mathcal{H}$ is

$$\mathcal{F} = \{t + uj \in \mathcal{H} : |t| \le \tfrac{1}{2}, t^2 + u^2 \ge 1\} \,.$$

The main idea followed in [SC03] is to set up a map

$$z \colon \mathbb{C}[x, y]'_n \longrightarrow \mathcal{H}_3$$

(where $\mathbb{C}[x, y]'_n$ is a suitable subset of $\mathbb{C}[x, y]_n$ that contains the squarefree forms) that is covariant with respect to the $\mathrm{SL}(2, \mathbb{C})$-actions on both sides, in the sense that

$$z(F \cdot \gamma) = \gamma^{-1} \cdot z(F) \qquad \text{for all } F \in \mathbb{C}[x, y]'_n \text{ and all } \gamma \in \mathrm{SL}(2, \mathbb{C}).$$

We also require z to be compatible with complex conjugation in the sense that if $z(F) = t + uj$, then $z(\bar{F}) = \bar{t} + uj$. This ensures that $z(F) \in \mathcal{H}$ when F has real coefficients. For such F we then say that F is *reduced* when $z(F) \in \mathcal{F}$. Since $\mathcal{F}$ is a fundamental domain for the $\mathrm{SL}(2, \mathbb{Z})$-action, there will be a reduced representative F_0 in each $\mathrm{SL}(2, \mathbb{Z})$-orbit. If $z(F_0)$ is in the interior of $\mathcal{F}$, then F_0 is uniquely determined (up to sign when n is odd, since $-I_2 \in \mathrm{SL}(2, \mathbb{Z})$ acts trivially on $\mathcal{H}$). When $z(F_0)$ is on the boundary of $\mathcal{F}$, there is some ambiguity, which can be resolved by removing part of the boundary. However, for practical purposes, this ambiguity is usually not a problem. We also note that it is impossible to find out whether a numerically computed point is exactly on the boundary.

There are many choices for this map z when $n \ge 5$ (for $n = 3$ or 4 there is only one choice, which is forced by the symmetries of cubics and quartics; see [SC03, Prop. 3.4]).

The choice made by Julia and in [SC03] can be described in the following way. We write $F \in \mathbb{C}[x,y]_n$ as

$$F(x,y) = \prod_{k=1}^{n} (\beta_k x - \alpha_k y)$$

with $\alpha_k, \beta_k \in \mathbb{C}$. Then we define, for $t + uj \in \mathcal{H}_3$,

$$R(F, t+uj) = \prod_{k=1}^{n} \frac{|\alpha_k - \beta_k t|^2 + |\beta_k|^2 u^2}{u} \in \mathbb{R}_{\geq 0}\,;$$

we note that $R(F, t+uj) > 0$ when $F \neq 0$. It can be checked that R is invariant under the $\mathrm{SL}(2,\mathbb{C})$-action: $R(F \cdot \gamma, z) = R(F, \gamma \cdot z)$.

Now we take $\mathbb{C}[x,y]_n' \subset \mathbb{C}[x,y]_n$ to be the subset of stable forms, where a form F is *stable* when no linear factor of F has multiplicity $\geq n/2$. It is shown in [SC03, Prop. 5.1] that for $F \in \mathbb{C}[x,y]_n'$, the function $\mathcal{H}_3 \to \mathbb{R}_{>0}$, $z \mapsto R(F,z)$, has a unique minimizer $z(F)$; we define the *Julia invariant* of F to be the minimal value,

$$\theta(F) = R(F, z(F)) > 0\,.$$

There is a nice geometric interpretation of $z(F)$. Namely, $z(F)$ is the unique point in $\mathcal{H}_3$ with the property that the sum of the unit tangent vectors pointing in the direction of the n roots of F (in $\mathbb{P}^1(\mathbb{C})$, considered as the ideal boundary of $\mathcal{H}_3$) vanishes; see [SC03, Cor. 5.4].

4 Bounds for the size of a binary form

Our goal in this section is to relate $\theta(F)$ and the size $\|F\|$ of a form $F \in \mathbb{C}[x,y]_n'$. The first step is a comparison between $\|F\|$ and $R(F,j)$. Note that for $F = \prod_{k=1}^{n}(\beta_k x - \alpha_k y)$ we have that

$$R(F,j) = \prod_{k=1}^{n} (|\alpha_k|^2 + |\beta_k|^2)\,.$$

Proposition 4.1. *Let* $F \in \mathbb{C}[x,y]_n$. *Then*

$$2^{1-n} R(F,j) \leq \|F\| \leq 2^{-n} \binom{2n}{n} R(F,j)\,.$$

If $F \in \mathbb{C}[x,y]_n'$ *with* $z(F) = j$, *then* $\|F\| \leq R(F,j)$.

Proof. We begin with an observation related to the size of F. Define $\tilde{F}(x,y) = \bar{F}(y,x)$; then clearly $\|F\|$ is the coefficient of $x^n y^n$ in $F(x,y)\tilde{F}(x,y)$. Now assume that $F = GH$. Then

$$F\tilde{F} = GH\tilde{G}\tilde{H} = G\tilde{H} \cdot \widetilde{G\tilde{H}}\,,$$

which implies that $\|F\| = \|GH\| = \|G\tilde{H}\|$. It is clear from the definition that

$$R(GH, j) = R(G,j) \cdot R(H,j) = R(G,j) \cdot R(\tilde{H}, j) = R(G\tilde{H}, j)\,.$$

Writing $F = \prod_{k=1}^{n}(\beta_k x - \alpha_k y)$, this allows us to replace each factor $\beta_k x - \alpha_k y$ with $|\alpha_k| > |\beta_k|$ by its reverse conjugate $-\bar{\alpha}_k x + \bar{\beta}_k y$. Since scaling F by a constant λ clearly scales both $\|F\|$ and $R(F,j)$ by $|\lambda|^2$, we can in this way assume without loss of generality that

$$F = (x - \alpha_1 y) \cdots (x - \alpha_n y) \qquad \text{with} \qquad |\alpha_1|, \ldots, |\alpha_n| \leq 1.$$

We now show the first inequality. We have that

$$\|F\| \geq 1 + \prod_{k=1}^{n} |\alpha_k|^2 \qquad \text{and} \qquad 2^{1-n} R(F, j) = 2^{1-n} \prod_{k=1}^{n} (1 + |\alpha_k|^2) .$$

Write $x_k = |\alpha_k|^2 \in [0, 1]$ and observe that the difference of $2^{1-n}(1 + x_1) \cdots (1 + x_n)$ and $1 + x_1 \cdots x_n$ is an affine-linear function in each of the x_k. This implies that the extrema of this difference must occur at some vertices of the unit cube $[0, 1]^n$. But it is easy to see that

$$1 + x_1 \cdots x_n \geq 2^{1-n}(1 + x_1) \cdots (1 + x_n)$$

whenever $x_1, \ldots, x_n \in \{0, 1\}$. This proves the first inequality.

We now turn to the second inequality. We write $e(\phi) = e^{2\pi i \phi}$ for $\phi \in \mathbb{R}$. Fix $\alpha \in \mathbb{C}$ and set $\beta = 2\alpha/(1 + |\alpha|^2)$. Then

$$|e(\phi) - \alpha|^2 = \frac{1 + |\alpha|^2}{2} \bigl(2 - \beta e(-\phi) - \bar{\beta} e(\phi)\bigr) .$$

So with F as above (note, though, that we do not need to assume that all roots have absolute value ≤ 1), we have that

$$\begin{aligned}
\|F\| &= \int_0^1 |F(e(\phi), 1)|^2 \, d\phi \\
&= \prod_{k=1}^{n} \frac{1 + |\alpha_k|^2}{2} \int_0^1 \prod_{k=1}^{n} \bigl(2 - \beta_k e(-\phi) - \bar{\beta}_k e(\phi)\bigr) \, d\phi \\
&= 2^{-n} R(F, j) \int_0^1 \prod_{k=1}^{n} \bigl(2 - \beta_k e(-\phi) - \bar{\beta}_k e(\phi)\bigr) \, d\phi
\end{aligned}$$

with $\beta_k = 2\alpha_k/(1 + |\alpha_k|^2)$. Note that each factor in the product is a real number in $[0, 4]$:

$$2 - \beta_k e(-\phi) - \bar{\beta}_k e(\phi) = 2 - 2 \operatorname{Re}(\beta_k e(-\phi))$$

and $|\beta_k| \leq 1$. So we can apply the AGM inequality to the product. This gives

$$\|F\| \leq 2^{-n} R(F, j) \int_0^1 \bigl(2 - \beta e(-\phi) - \bar{\beta} e(\phi)\bigr)^n \, d\phi ,$$

where $\beta = \sum_{k=1}^{n} \beta_k / n$ is the arithmetic mean of the β_k. Now one can check that β_k is the vertical projection to the unit disk of α_k, viewed as a point on the Riemann sphere, which forms the ideal boundary of the Poincaré ball model of $\mathcal{H}_3$. When $z(F) = j$, then the sum of the points on the Riemann sphere corresponding to the α_k vanishes by the geometric characterization of $z(F)$, so in this case, we have that $\beta = 0$, which implies $\|F\| \leq R(F, j)$. In the general case, expanding the product and integrating gives

$$\int_0^1 \bigl(2 - \beta e(-\phi) - \bar{\beta} e(\phi)\bigr)^n \, d\phi = \sum_{\ell=0}^{\infty} 2^{n-2\ell} \binom{n}{2\ell} \binom{2\ell}{\ell} |\beta|^{2\ell} ,$$

which shows that the maximum is attained when $|\beta|$ is maximal, so for $|\beta| = 1$. It is clear that the integral does not depend on the argument of β, so we can take $\beta = -1$. Then the integrand simplifies to

$$\bigl(2 + e(\phi) + e(-\phi)\bigr)^n = \bigl(e(\tfrac{1}{2}\phi) + e(-\tfrac{1}{2}\phi)\bigr)^{2n}$$

and so the value of the integral is the constant term of $(X + X^{-1})^{2n}$, which is $\binom{2n}{n}$. This proves the second inequality in the general case. $\square$

We note that both inequalities are sharp in the general case: the lower bound is attained for $F = x^n - y^n$ (which has $z(F) = j$), and the upper bound is attained for $F = (x - y)^n$. The upper bound in the case $z(F) = j$ is sharp for n even, in the sense that

$$\sup_{F\colon z(F)=j} \frac{\|F\|}{R(F, j)} = 1\,;$$

this is attained as F gets close to $(xy)^{n/2}$. For odd n, we cannot balance the roots in this way while they tend to zero or infinity, so the bound will not be sharp, but it will not be too far off when n is large.

From now on, we assume that F is stable, i.e., $F \in \mathbb{C}[x, y]'_n$; then $z(F)$ and $\theta(F)$ are defined.

We want to relate $\|F\|$ and $\theta(F)$. Proposition 4.1 tells us that

$$2^{1-n} \le \frac{\|F\|}{\theta(F)} \le 1$$

when $z(F) = j$, since then $\theta(F) = R(F, z(F)) = R(F, j)$. This leads us to expect that $\|F\|$ can be bounded in terms of $\theta(F)$ and the (hyperbolic) distance between $z(F)$ and j: we have that

$$\frac{\|F\|}{\theta(F)} = \frac{\|F\|}{R(F, z(F))} = \frac{\|F\|}{R(F, j)} \cdot \frac{R(F, j)}{R(F, z(F))}\,;$$

the first factor is bounded above and below by Proposition 4.1, and since $R(F, z)$ attains its minimum when $z = z(F)$, the second factor should grow when $z(F)$ moves away from j.

Let $\gamma \in SL(2, \mathbb{C})$ be such that $F_0 = F \cdot \gamma^{-1}$ satisfies $z(F_0) = j = \gamma \cdot z(F)$. By the invariance of R, we then have that

$$\frac{R(F, j)}{R(F, z(F))} = \frac{R(F_0 \cdot \gamma, j)}{R(F_0 \cdot \gamma, z(F))} = \frac{R(F_0, \gamma \cdot j)}{R(F_0, \gamma \cdot z(F))} = \frac{R(F_0, \gamma \cdot j)}{R(F_0, j)}\,.$$

If $\operatorname{dist}(z, z')$ denotes hyperbolic distance in $\mathcal{H}_3$, then, since distance is invariant under the $SL(2, \mathbb{C})$-action, $\operatorname{dist}(z(F), j) = \operatorname{dist}(\gamma \cdot j, j)$. So it is enough to bound $R(F_0, z)/R(F_0, j)$ in terms of $\operatorname{dist}(z, j)$ when $z(F_0) = j$.

Now for $z = t + uj$, the distance to j is given by

$$\cosh \operatorname{dist}(z, j) = \frac{|t|^2 + u^2 + 1}{2u}\,. \tag{4.1}$$

In particular, $\operatorname{dist}(e^{\delta} j, j) = |\delta|$. We can assume that $F_0 = \prod_{k=1}^{n}(x - \alpha_k y)$. Let $\phi_k \in \mathbb{S}^2$ be the point on the Riemann sphere that corresponds to α_k under stereographic projection (in coordinates in $\mathbb{R}^3 = \mathbb{C} \times \mathbb{R}$, $\phi_k = \left(\frac{2\alpha_k}{|\alpha_k|^2+1}, \frac{|\alpha_k|^2-1}{|\alpha_k|^2+1}\right)$; the first component is β_k in the proof of Proposition 4.1). The condition $z(F_0) = j$ then is equivalent to $\sum_k \phi_k = 0$. We derive a formula for the quotient $R(F_0, z)/R(F_0, j)$.

Lemma 4.2. *Let $0 \neq F_0 \in \mathbb{C}[x,y]_n$ and $z \in \mathcal{H}_3$. We set $\delta = \operatorname{dist}(z, j)$ and denote the unit tangent vector at j in the direction of z by $\phi \in \mathbb{S}^2$ (ϕ is arbitrary when $z = j$). Let $\phi_k \in \mathbb{S}^2$ be as in the preceding paragraph. Then*

$$\frac{R(F_0, z)}{R(F_0, j)} = \prod_{k=1}^{n} (\cosh \delta + \langle \phi, \phi_k \rangle \sinh \delta)\,,$$

where $\langle \cdot, \cdot \rangle$ denotes the standard inner product on $\mathbb{R}^3$ (recall that $\phi, \phi_k \in \mathbb{S}^2 \subset \mathbb{R}^3$).

Proof. Both sides do not change when we replace F_0 and z by $F_0 \cdot \gamma$ and $\gamma^{-1} \cdot z$, respectively, for some $\gamma \in SU(2)$. This allows us to assume that ϕ points upward; then $z = e^{\delta} j$ and $\langle \phi, \phi_k \rangle = (|\alpha_k|^2 - 1)/(|\alpha_k|^2 + 1)$. By definition of R and with $z = e^{\delta} j$, we obtain that

$$\begin{aligned}
\frac{R(F_0, z)}{R(F_0, j)} &= \prod_{k=1}^{n} \frac{|\alpha_k|^2 e^{-\delta} + e^{\delta}}{|\alpha_k|^2 + 1} \\
&= \prod_{k=1}^{n} \left(\frac{e^{\delta} + e^{-\delta}}{2} + \frac{|\alpha_k|^2 - 1}{|\alpha_k|^2 + 1} \cdot \frac{e^{\delta} - e^{-\delta}}{2} \right) \\
&= \prod_{k=1}^{n} (\cosh \delta + \langle \phi, \phi_k \rangle \sinh \delta)\,. \qquad \square
\end{aligned}$$

We can use this to deduce bounds for the R-quotient. We will derive an upper bound assuming that $z(F_0) = j$. Regarding a lower bound, note that the factor $(\cosh \delta + \langle \phi, \phi_k \rangle \sinh \delta)$ in the product above can be as small as $e^{-\delta}$ when $\phi_k = -\phi$. So if lots of roots are in the direction opposite to z, then the product can get quite small. To get a reasonable bound, we assume that F_0 is squarefree. Then the directions ϕ_k cannot get too close to one another, and so at most one factor can get really small. This approach leads to the lower bound below.

Proposition 4.3. *Let $F_0 \in \mathbb{C}[x,y]_n'$ be squarefree and such that $z(F_0) = j$. There is a constant $\varepsilon(F_0) > 0$ such that for all $z \in \mathcal{H}_3$, we have that*

$$\varepsilon(F_0)(\cosh \delta)^{n-2} \leq \frac{R(F_0, z)}{R(F_0, j)} \leq (\cosh \delta)^n\,,$$

where $\delta = \operatorname{dist}(z, j)$.

Proof. By Lemma 4.2,

$$\frac{R(F_0, z)}{R(F_0, j)} = \prod_{k=1}^{n} (\cosh \delta + \langle \phi, \phi_k \rangle \sinh \delta)\,.$$

Since $z(F_0) = j$, we have that $\sum_{k=1}^{n} \phi_k = 0$. Therefore, by the AGM inequality,

$$\prod_{k=1}^{n} (\cosh \delta + \langle \phi, \phi_k \rangle \sinh \delta) \leq \Big(\cosh \delta + \Big\langle \phi, \frac{1}{n} \sum_{k=1}^{n} \phi_k \Big\rangle \sinh \delta\Big)^n = (\cosh \delta)^n\,.$$

By assumption, the ϕ_k are pairwise distinct, which implies that

$$\eta = 1 - \max_{k \neq k'} \sqrt{\frac{\langle \phi_k, \phi_{k'} \rangle + 1}{2}} > 0 .$$

Then for any $\phi \in \mathbb{S}^2$, it follows that $\langle \phi, \phi_k \rangle > 1 - \eta$ for at most one k, say k_0 ($1-\eta$ is the cosine of half the angle between the closest two ϕ_k). Applying this to $-\phi$, we see that for all $k \neq k_0$,

$$\cosh\delta + \langle \phi, \phi_k \rangle \sinh\delta \geq \cosh\delta + (\eta - 1)\sinh\delta = e^{-\delta} + \eta \sinh\delta \geq \eta \cosh\delta .$$

For k_0 we have that

$$\cosh(\delta) - \langle \phi, \phi_{k_0} \rangle \sinh(\delta) \geq e^{-\delta} \geq \frac{1}{2\cosh(\delta)} .$$

This gives that

$$\frac{R(F_0, z)}{R(F_0, j)} \geq \frac{\eta^{n-1}}{2} (\cosh\delta)^{n-2} .$$

So the lower bound holds with $\varepsilon(F_0) = \eta^{n-1}/2$. $\square$

It is fairly clear that the value given for $\varepsilon(F_0)$ in the proof is unlikely to be optimal. Writing $\rho = \tanh\delta$, the optimal value we can take is

$$\varepsilon(F_0) = \inf_{\phi \in \mathbb{S}^2, 0 \leq \rho < 1} \frac{\prod_{k=1}^{n} (1 + \langle \phi, \phi_k \rangle \rho)}{1 - \rho^2} . \tag{4.2}$$

When $n = 3$, the only way for $\phi_1 + \phi_2 + \phi_3$ to vanish is that the ϕ_k are the vertices of an equilateral triangle. Without loss of generality, we can take them to be $1, \omega, \omega^2 \in \mathbb{S}^1 = \mathbb{S}^2 \cap (\mathbb{C} \times \{0\})$, where $\omega = e(1/3)$. If the projection of ϕ to the complex plane is $\lambda e(\varphi)$, with $0 \leq \lambda \leq 1$, then the expression under the infimum works out as

$$\frac{1 - \frac{3}{4}\lambda^2\rho^2 + \frac{1}{4}\cos(6\pi\varphi)\lambda^3\rho^3}{1 - \rho^2} .$$

It is clear that, for fixed ρ, the expression will be minimal when $\cos(6\pi\varphi) = -1$ and $\lambda = 1$. The expression then simplifies to

$$\frac{(1 + \frac{1}{2}\rho)^2}{1 + \rho} = 1 + \frac{\rho^2}{4(1 + \rho)} ,$$

which is minimal for $\rho = 0$. This gives the following.

Lemma 4.4. *If $F_0 \in \mathbb{C}[x, y]_3'$ satisfies $z(F_0) = j$, then we can take $\varepsilon(F_0) = 1$.*

If we would use the value $\eta^2/2$ from the proof of Proposition 4.3, then we would obtain $\varepsilon(F_0) = \frac{1}{8}$ instead. In general, we can at least in principle compute (an approximation to) the optimal $\varepsilon(F_0)$ by solving the optimization problem in (4.2). We remark here that when working over $\mathbb{R}$ instead of $\mathbb{C}$, we can restrict ϕ to run over the circle $\mathbb{S}^2 \cap (\mathbb{R} \times \mathbb{R})$ (inside $\mathbb{R}^3 = \mathbb{C} \times \mathbb{R}$). This simplifies the computation and can also result in a better bound. A further improvement (which can also be used to extend the applicability of the resulting algorithm from squarefree forms to stable forms) is based on the following result.

Lemma 4.5. *Let $F_0 \in \mathbb{C}[x,y]'_n$ be stable with $z(F_0) = j$. Then*

$$\varepsilon_{F_0} \colon \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}_{\geq 1}\,, \qquad \delta \longmapsto \min_{z\colon\, \mathrm{dist}(z,j)=\delta} \frac{R(F_0,z)}{R(F_0,j)}$$

is a strictly monotonically increasing bijection.

Proof. We know that the unique minimum of $R(F_0, z)$ is attained when $z = z(F_0) = j$. We now fix ϕ in Lemma 4.2 and consider $R(F_0,z)/R(F_0,j)$ as a function of δ. Note that

$$\cosh\delta + \langle \phi, \phi_k \rangle \sinh\delta = \frac{1 + \langle \phi, \phi_k \rangle}{2} e^{\delta} + \frac{1 - \langle \phi, \phi_k \rangle}{2} e^{-\delta}$$

is a nonnegative linear combination of e^{δ} and $e^{-\delta}$. This implies that $R(F_0,z)/R(F_0,j)$, which is the product of these expressions over all $1 \leq k \leq n$, is a nonnegative linear combination of terms $e^{m\delta}$ (with $-n \leq m \leq n$), which are all convex from below. For $m \neq 0$, they are even strictly convex. The only possibility that results in a constant product is that half of the ϕ_k equal ϕ and the other half equal $-\phi$, but this would imply that F_0 is not stable. So the quotient, considered as a function of δ for fixed ϕ, is strictly convex from below with minimum value 1 at $\delta = 0$. In particular, $\delta \mapsto R(F_0,z)/R(F_0,j)$ is strictly increasing as δ grows from 0 to ∞. It follows that ε_{F_0} is strictly increasing as well; also $\varepsilon_{F_0}(0) = 1$. It remains to show that $\varepsilon_{F_0}(\delta)$ tends to infinity with δ. Note that

$$\cosh\delta + \langle \phi, \phi_k \rangle \sinh\delta \geq \max\Bigl\{\frac{1 + \langle \phi, \phi_k \rangle}{2} e^{\delta}, e^{-\delta}\Bigr\}\,.$$

Since F_0 is stable by assumption, the maximal multiplicity m of a root of F_0 is strictly less than $n/2$. Defining η similarly as in the proof of Proposition 4.3, but restricting to pairs k, k' with $\phi_k \neq \phi_{k'}$, we deduce that

$$\frac{R(F_0,z)}{R(F_0,j)} \geq \eta^{n-m} e^{(n-2m)\delta} \geq \eta^{n-m} e^{\delta}\,,$$

which finishes the proof. □

Definition 4.6. Let $F \in \mathbb{C}[x,y]_n$ be stable. Let F_0 be a form in the $\mathrm{SL}(2,\mathbb{C})$-orbit of F satisfying $z(F_0) = j$. Then we define $\varepsilon_F(\delta) = \varepsilon_{F_0}(\delta)$ with ε_{F_0} as in Lemma 4.5. If F is squarefree, then we define $\varepsilon(F) = \varepsilon(F_0)$ with $\varepsilon(F_0)$ as in (4.2).

We remark that $\varepsilon(F_0 \cdot \gamma) = \varepsilon(F_0)$ for $\gamma \in \mathrm{SU}(2)$ (such γ induce rotations of $\mathbb{S}^2$ and so do not change the geometry of the situation), which implies that ε_F (or $\varepsilon(F)$) does not depend on the choice of F_0. When working over $\mathbb{R}$, we take F_0 in the $\mathrm{SL}(2,\mathbb{R})$-orbit of F; the previous remark then applies with $\mathrm{SO}(2)$ in place of $\mathrm{SU}(2)$.

We now combine the results obtained so far.

Theorem 4.7. *Let $F \in \mathbb{C}[x,y]_n$ be stable; we write $\delta = \mathrm{dist}(z(F), j)$. Then*

$$2^{1-n}\varepsilon_F(\delta) \leq \frac{\|F\|}{\theta(F)} \leq 2^{-n}\binom{2n}{n}(\cosh\delta)^n\,.$$

For $\gamma \in \mathrm{SL}(2,\mathbb{C})$ write $\gamma^{-1} \cdot z(F) = t + uj$. If

$$\frac{|t|^2 + u^2 + 1}{2u} > \cosh \varepsilon_F^{-1}\Bigl(2^{n-1}\frac{\|F\|}{\theta(F)}\Bigr)\,,$$

then $\|F \cdot \gamma\| > \|F\|$.

If F *is squarefree, then we can replace the condition on* γ *by*

$$\frac{|t|^2 + u^2 + 1}{2u} > 2\left(\frac{2\|F\|}{\varepsilon(F)\theta(F)}\right)^{1/(n-2)} .$$

Proof. Recall that $\theta(F) = R(F, z(F))$. Choose $\gamma_0 \in \mathrm{SL}(2, \mathbb{C})$ such that $\gamma_0 \cdot z(F) = j$; then $F_0 = F \cdot \gamma_0^{-1}$ satisfies $z(F_0) = j$ and $\varepsilon(F) = \varepsilon(F_0)$. By the invariance of R,

$$\frac{\|F\|}{\theta(F)} = \frac{\|F\|}{R(F, j)} \cdot \frac{R(F, j)}{R(F, z(F))} = \frac{\|F\|}{R(F, j)} \cdot \frac{R(F_0, \gamma_0 \cdot j)}{R(F_0, j)} .$$

Combining the definition of ε_F with the bounds from Propositions 4.1 and 4.3 and using that $\mathrm{dist}(\gamma_0 \cdot j, j) = \delta$, we get the bounds on $\|F\|/\theta(F)$. For the second statement, we apply the lower bound for $F \cdot \gamma$ in place of F. Since $\theta(F \cdot \gamma) = \theta(F)$ and $\varepsilon_{F\cdot\gamma} = \varepsilon_F$, this results in

$$\|F \cdot \gamma\| \geq 2^{1-n} \varepsilon_F\big(\mathrm{dist}(\gamma^{-1} \cdot z(F), j)\big)\theta(F) .$$

By (4.1), $\cosh \mathrm{dist}(\gamma^{-1} \cdot z(F), j) = (|t|^2 + u^2 + 1)/(2u)$, so the stated condition is equivalent to the right hand side being $> \|F\|$.

For squarefree F, we use the estimate

$$\varepsilon_F(\delta) \geq \varepsilon(F)(\cosh \delta)^{n-2} ,$$

which implies that any γ satisfying the last condition also satisfies the previous one. □

The last part of the proof shows that using ε_F will in general result in better bounds than using $\varepsilon(F)$. On the other hand, inverting ε_F may be algorithmically more involved.

Note that for the second statement, we only need the lower bound, based on the lower bound in Proposition 4.1, which is sharp for some F.

Now consider a squarefree (or stable) $F \in \mathbb{R}[x, y]_n$. Then Theorem 4.7 gives us a finite subset of the $\mathrm{SL}(2, \mathbb{Z})$-orbit of F that is guaranteed to contain the representative of minimal size. We will turn this into an algorithm in the next section.

Remark 4.8. If we want to replace $\mathbb{Q}$ by a number field K with ring of integers $\mathcal{O}_K$, then we would have to work with the (diagonal) action of $\mathrm{SL}(2, \mathcal{O}_K)$ on a product of upper half planes and spaces (one upper half plane for each real embedding of K and one upper half space for each pair of complex embeddings). For each embedding, we get a covariant point of $F \in K[x, y]_n'$ in the corresponding half plane or space. By taking products, we can extend the definitions of $\|F\|$ and $\theta(F)$ to this situation, and we should be able to prove a version of Theorem 4.7 that applies to it. The major problem, however, will be the enumeration of the points in the $\mathrm{SL}(2, \mathcal{O}_K)$-orbit of $z(F)$ (which is now the tuple of covariant points associated to F) that have bounded distance from $(j, j, \ldots, j)$. To our knowledge, there are no good general algorithms for this so far. In some special cases (like imaginary quadratic fields of class number 1), an approach similar to that described in Section 5 should be workable, however.

Remark 4.9. With applications beyond binary forms in mind, we note that we can generalize Theorem 4.7 to the following situation. Assume we consider some kind of objects Φ associated to $\mathbb{P}^1$ and that we have a PGL(2)-equivariant way of associating to Φ a binary form F (up to scaling). The object Φ will be given in terms of coordinates of points or coefficients of binary forms, so there is actually an action of GL(2); we assume that the map $\Phi \mapsto F$ is in fact GL(2)-equivariant. Let $s(\Phi)$ denote some measure of the "size" or "height" of Φ; we need the further assumption that there are constants $C, k > 0$ such that $\|F\| \le Cs(\Phi)^k$. Since the coefficients of F will usually be given as homogeneous polynomials in the coordinates or coefficients describing our "model" of Φ, such a bound will be easily established. Then, for Φ defined over $\mathbb{C}$, we can deduce in the same way as in the proof of Theorem 4.7 that when $\gamma^{-1} \cdot z(F) = t + uj$ and

$$\frac{|t|^2 + u^2 + 1}{2u} > \cosh \varepsilon_F^{-1}\Big(2^{n-1} \frac{Cs(\Phi)^k}{\theta(F)}\Big),$$

we have that $s(\Phi \cdot \gamma) > s(\Phi)$. We will apply this to endomorphisms of $\mathbb{P}^1$ in Section 6.

Another application is when we want to use the maximum norm $H_\infty(F)$ instead of $\|F\|$ to measure how large F is. We have that $\|F\| \le (n+1)H_\infty(F)^2$, which leads to the condition

$$\frac{|t|^2 + u^2 + 1}{2u} > \cosh \varepsilon_F^{-1}\Big(2^{n-1} \frac{(n+1)H_\infty(F)^2}{\theta(F)}\Big)$$

that guarantees that $H_\infty(F \cdot \gamma) > H_\infty(F)$.

5 The algorithm

The basic outline of an algorithm implementing Theorem 4.7 is clear. We assume that $F(x, y) \in \mathbb{R}[x, y]_n$ is stable (so in particular, $n \ge 3$). In practice, we may have to assume that F is squarefree, since some implementations require this to be able to compute $z(F)$ and $\theta(F)$.

Algorithm 5.1.

1. Compute $z(F)$ and $\theta(F)$.
2. Determine $\gamma_0 \in SL(2, \mathbb{Z})$ such that $\gamma_0^{-1} \cdot z(F) \in \mathcal{F}$.
   Replace $z(F)$ by $\gamma_0^{-1} \cdot z(F)$ and F by $F \cdot \gamma_0$.
3. Compute an upper bound c for $\cosh \varepsilon_F^{-1}(2^{n-1}\|F\|/\theta(F))$.
4. Let Γ be the set of all $\gamma \in SL(2, \mathbb{Z})$ such that $\cosh \operatorname{dist}(\gamma^{-1} \cdot z(F), j) \le c$.
5. Determine $\gamma \in \Gamma$ with $\|F \cdot \gamma\|$ minimal and return $\gamma_0\gamma$ and $F \cdot \gamma$.

It is explained in [SC03] how one can compute $z(F)$ and $\theta(F)$; implementations are available in Magma [BCP97] and Sage [SJ05]. Let $z(F) = t_0 + u_0 j$ (with $t_0 \in \mathbb{R}$, since F has real coefficients). To compute c, we construct a suitable F_0, for example, by first replacing F by $F_1(x, y) = F(x + t_0 y, y)$ and then setting $F_0(x, y) = F_1(u_0^{1/2} x, u_0^{-1/2} y)$ (or $F_1(x, y/u_0)$; scaling does not affect ε_{F_0}). From the roots of F_0 we can deduce the collection of $\phi_k \in \mathbb{S}^2$. Then we can use Lemma 4.2 to evaluate $\varepsilon_{F_0}(\delta)$ for a collection of values δ; the monotonicity of $\cosh \circ \varepsilon_{F_0}$ then gives us a suitable c. Alternatively, when F is squarefree, we can use (4.2) to find a lower bound ε for $\varepsilon(F) = \varepsilon(F_0)$ and set

$$c = 2\Big(\frac{2\|F\|}{\varepsilon\theta(F)}\Big)^{1/(n-2)}.$$

(If $n = 3$, we can simply set $c = 4\|F\|/\theta(F)$; see Lemma 4.4.) As already mentioned, using ε_F will give better bounds and therefore lead to a smaller search space than using $\varepsilon(F)$, but inverting ε_F may require some additional work.

The least straight-forward step is step 4 above, which comes down to enumerating all points in the $\mathrm{SL}(2,\mathbb{Z})$-orbit of $z(F)$ whose hyperbolic distance from j does not exceed $\operatorname{arcosh} c$. One possibility to deal with this is to think backwards: given a point z in the orbit of $z(F)$, where we assume that $z(F) \in \mathcal{F}$, we get from z to $z(F)$ by the usual "shift-and-invert" procedure (add $m \in \mathbb{Z}$ to z so that $|\operatorname{Re}(z+m)| \le \frac{1}{2}$; if $|z+m| < 1$, replace by $-1/(z+m)$ and repeat), where the shifts decrease the distance from j and the inversions do not change it. So we get all points in the orbit with bounded distance from j by constructing a rooted tree with nodes labeled by $z \in \mathcal{H}$ and edges labeled by S (inversion), T (shift by $+1$) or T^{-1} (shift by -1), as follows.

1. The root is $z(F)$. It has three children: $-1/z(F)$ (edge labeled S), $z(F)+1$ (edge labeled T) and $z(F)-1$ (edge labeled T^{-1}).
2. Let $z = t + uj$ be a non-root node. Let ℓ be the label of the edge connecting the node to its parent. If $(t^2+u^2+1)/(2u) > c$, then remove this node. Otherwise:
   a. If $\ell \ne S$ and $(|t|-1)^2 + u^2 \ge 1$, add a child $-1/z$ (edge labeled S).
   b. If $\ell \ne T$, add a child $z-1$ (edge labeled T^{-1}).
   c. If $\ell \ne T^{-1}$, add a child $z+1$ (edge labeled T).

The condition $(|t|-1)^2 + u^2 \ge 1$ is equivalent to $|\operatorname{Re}(-1/z)| \le \frac{1}{2}$; this means that z can have resulted from an inversion step in the shift-and-invert procedure. If $z(F)$ is close to $\pm\frac{1}{2} + \frac{\sqrt{3}}{2}j$, then using the condition as stated can lead to cycling in the algorithm. To avoid this, we can use any representative point r in the translate of $\mathcal{F}$ under γ, for example the points in the orbit of $2j$, to test the condition. We then store $-1/r$, $r-1$ or $r+1$ in addition to $-1/z$, $z-1$ or $z+1$ in the child node.

Since we can update the bound c when we have found a new temporary minimum of $\|F\cdot\gamma\|$, the most efficient way to perform the tree search is to order the nodes by their distance from j and always expand the node closest to j that has not yet been dealt with.

Of course, we also keep track of $F\cdot\gamma$: a shift by ± 1 in z corresponds to substituting $(x \mp y, y)$ for (x, y), and an inversion corresponds to substituting $(y, -x)$ for (x, y).

Example 5.2. Consider the cubic form

$$F(x,y) = -2x^3 + 2x^2y + 3xy^2 + 127y^3$$

with $\|F\| = 16\,146$. This form is reduced in the sense of [SC03], since

$$z(F) \approx 0.17501 + 3.99543j \in \mathcal{F}.$$

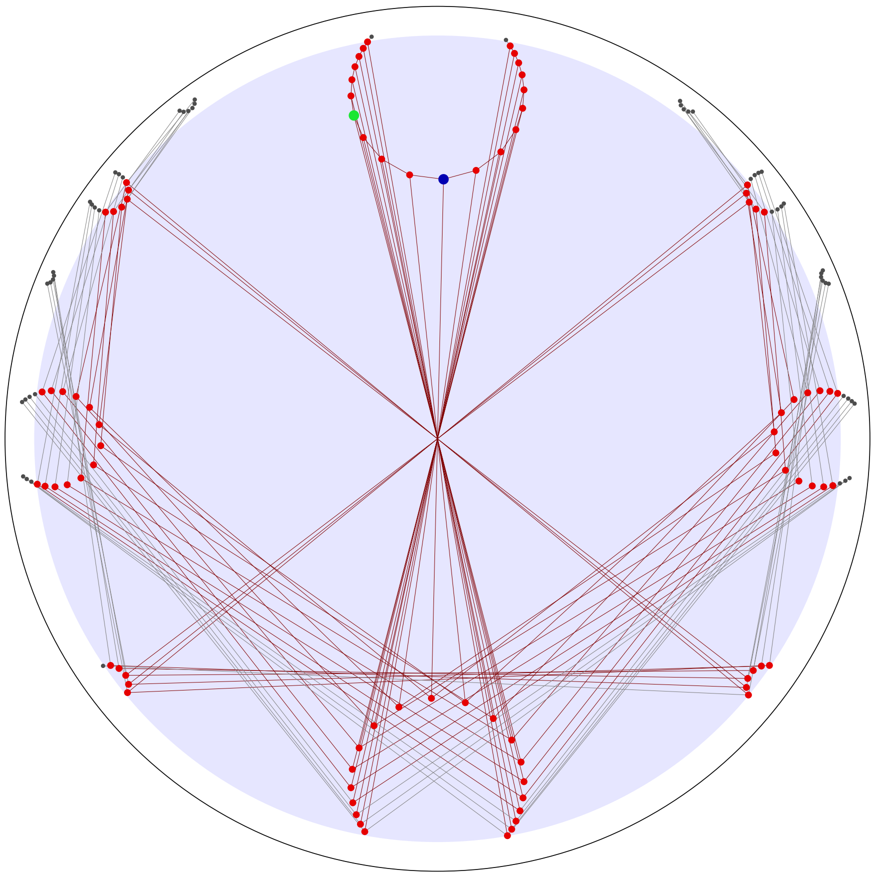

FIGURE 1. The tree that is traversed during the computation of the representative of minimal size in Example 5.2, in the Poincaré disk model. The light blue disk bounds the search region. The larger blue dot represents $z(F)$ and the larger green dot is where the minimum is attained. The red dots mark the nodes that are expanded; the gray dots are nodes that are discarded.

We then need to find the $\gamma \in \mathrm{SL}(2, \mathbb{Z})$ with $\gamma^{-1} \cdot z(F) = t + uj$ satisfying

$$\frac{t^2 + u^2 + 1}{2u} \leq c = \cosh \varepsilon_F^{-1}\Big(\frac{4\|F\|}{\theta(F)}\Big) \approx 28.0049\,.$$

For comparison, if we use $\varepsilon(F) = 1$ instead, the initial bound we obtain is $c \approx 31.5022$. We use the tree search explained above. In the course of the search, the bound gets reduced to ≈ 14.3415; there are then 88 points (marked red in Figure 1) in the $\mathrm{SL}(2, \mathbb{Z})$-orbit of $z(F)$ that have to be considered. A representative of smallest size is obtained for $\gamma = \left(\begin{smallmatrix} 1 & 4 \\ 0 & 1 \end{smallmatrix}\right)$ with

$$\|F \cdot \gamma\| = \|F(x + 4y, y)\| = \|{-2x^3} - 22x^2y - 77xy^2 + 43y^3\| = 8\,266\,.$$

If instead we minimize the height, then the bound c is ≈ 111.891, which gets reduced to ≈ 23.3403 during the search. There are 140 points in the search region, and a smallest representative is

$$F(4x - 5y, x - y) = 43x^3 - 52x^2y - 47xy^2 + 58y^3 \qquad \text{of height } 58$$

(compared to $H_\infty(F) = 127$). The covariant point of this form is $\approx 1.12502 + 0.13060j$; the cosh of its distance to j is ≈ 8.73973.

6 Application to dynamical systems

A *dynamical system on* $\mathbb{P}^1$ is a non-constant endomorphism $f\colon \mathbb{P}^1 \to \mathbb{P}^1$. As described in Section 2, f can be specified by a model $[F : G]$, where F and G are binary forms of the same degree d and without common factors. If f is defined over $\mathbb{Q}$, we can choose $F, G \in \mathbb{Z}[x, y]_d$ with coprime coefficients. We have a natural right action of $SL(2, \mathbb{Z})$ on endomorphisms of degree d by conjugation, which is given explicitly for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL(2, \mathbb{Z})$ by

$$f^\gamma = [F^\gamma : G^\gamma]\,,$$

where

$$\begin{aligned} F^\gamma(x, y) &= dF(ax + by, cx + dy) - bG(ax + by, cx + dy) \qquad \text{and} \\ G^\gamma(x, y) &= -cF(ax + by, cx + dy) + aG(ax + by, cx + dy)\,. \end{aligned}$$

This action can be extended to $\mathrm{Mat}(2, \mathbb{Z}) \cap GL(2, \mathbb{Q})$. Note that this amounts to using the adjoint $\left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)$ in place of the inverse compared to the action by conjugation. Since scaling both forms by a common factor does not change the endomorphism they represent, this still induces the usual action of $PGL(2, \mathbb{Q})$ on endomorphisms by conjugation.

To apply the reduction algorithm of binary forms to dynamical systems, we follow the framework of Remark 4.9. We first associate to each dynamical system a binary form in a covariant way. Let $f\colon \mathbb{P}^1 \to \mathbb{P}^1$ be a dynamical system defined over $\mathbb{Q}$. Choose a model $f = [F : G]$ with two homogeneous polynomials F and G with coprime integral coefficients. We write the mth iterate of f as $f^m = [F_m : G_m]$. Define

$$\Phi_m(f) = yF_m - xG_m \qquad \text{and} \qquad \Phi^*_m(f) = \prod_{k \mid m} (yF_k - xG_k)^{\mu(m/k)}\,,$$

where μ is the Möbius function. The zeros of the form $\Phi_m(f)$ are the points of period m for f. The form $\Phi^*_m(f)$ is called the *dynatomic polynomial* and its zeros, in most cases, are the points of minimal period m for f. See [Sil07, Section 4.1] for properties of dynatomic polynomials in dimension 1. It is easy to check that $\Phi_m(f^\gamma) = \Phi_m(f) \cdot \gamma$ and similarly for Φ^*_m.

We can bound the size of $\Phi_m(f)$ and $\Phi^*_m(f)$ in terms of the height of $f = [F : G]$.

Proposition 6.1. *Given $d \geq 2$ and $m \geq 1$, there exist positive constants $C_{d,m}, C^*_{d,m}$ and $k_{d,m}, k^*_{d,m}$ such that for every morphism $f\colon \mathbb{P}^1 \to \mathbb{P}^1$ of degree $d \geq 2$ as above,*

$$\|\Phi_m(f)\| \leq C_{d,m} H(f)^{k_{d,m}} \qquad \textit{and} \qquad \|\Phi^*_m(f)\| \leq C^*_{d,m} H(f)^{k^*_{d,m}}\,.$$

Proof. The coefficients of $\Phi_m(f)$ and $\Phi^*_m(f)$ are homogeneous polynomials in the coefficients of f. Then the sum $s(f)$ of the squares of the coefficients of $\Phi_m(f)$ or $\Phi^*_m(f)$ is a homogeneous polynomial of some degree $k_{d,m}$ or $k^*_{d,m}$ in the coefficients of f. The bound on $\|\Phi^{(*)}_m(f)\| = s(f)$ is obtained by bounding the coefficients of f by $H(f)$ and applying the triangle inequality. □

For given d and m, we can find suitable constants (at least in principle) by doing the computation mentioned in the proof for a generic f. This gives

$$\begin{array}{llll} C_{d,1} = C^*_{d,1} = 4d+2\,, & k_{d,1} = k^*_{d,1} = 2\,; & C_{2,2} = 322\,, & k_{2,2} = 6\,; \\ C^*_{2,2} = 43\,, & k^*_{2,2} = 4\,; & C^*_{2,3} = 106\,459\,, & k^*_{2,3} = 12\,; \\ C_{3,2} = 18\,044\,, & k_{3,2} = 8\,; & C^*_{3,2} = 1604\,, & k^*_{3,2} = 6\,. \end{array}$$

For most endomorphisms, using $m = 1$ is sufficient to get a squarefree binary form of degree at least 3, but for some maps we need to consider higher order periodic points.

Applying Remark 4.9 now results in the following.

Corollary 6.2. *Let $f = [F : G] \colon \mathbb{P}^1 \to \mathbb{P}^1$ be a morphism of degree d, given by two binary forms $F, G \in \mathbb{Z}[x, y]_d$ with coprime coefficients. Pick some $\Phi = \Phi_m(f)$ or $\Phi^*_m(f)$ such that Φ is stable (in particular, $\deg \Phi \geq 3$). Let $C = C_{d,m}$ or $C^*_{d,m}$ and $k = k_{d,m}$ or $k^*_{d,m}$ be the constants from Proposition 6.1, depending on the choice of Φ.*

Let $\gamma \in \mathrm{SL}(2, \mathbb{Z})$ and write $\gamma^{-1} \cdot z(\Phi) = t + uj$. If

$$\frac{t^2 + u^2 + 1}{2u} > \cosh \varepsilon_\Phi^{-1} \left(2^{\deg \Phi - 1} \frac{CH(f)^k}{\theta(\Phi)} \right) ,$$

then $H(f^\gamma) > H(f)$.

This easily translates into an algorithm that produces a representative of smallest height in a given $\mathrm{SL}(2, \mathbb{Z})$-orbit. We just have to use the bound from Corollary 6.2 in place of c in the algorithm of Section 5 and search for the minimal $H(f^\gamma)$ (reducing the bound whenever possible). However, the problem remains of finding the representative of smallest height over *all* $\mathrm{SL}(2, \mathbb{Z})$-orbits of minimal models. Bruin and Molnar prove that for f defined over $\mathbb{Q}$, there are only finitely many $\mathrm{SL}(2, \mathbb{Z})$-orbits of minimal models; see [BM12, Proposition 6.5]. So to obtain a reduced model, we only have to apply the algorithm of this paper to a representative from each $\mathrm{SL}(2, \mathbb{Z})$-orbit of minimal models. We will discuss in Section 7 how we can obtain such representatives.

The following example shows how to find the smallest height representative for an endomorphism of $\mathbb{P}^1$.

Example 6.3. Consider the endomorphism

$$\begin{aligned} f\colon \qquad \mathbb{P}^1 &\longrightarrow \mathbb{P}^1 \\ (x : y) &\longmapsto (50x^2 + 795xy + 2120y^2 : 265x^2 + 106y^2)\,. \end{aligned}$$

Since f has even degree and the given model is minimal, we need only consider its $\mathrm{SL}(2, \mathbb{Z})$-orbit; see Proposition 7.2.

The binary form defining its fixed points is

$$\Phi_1(f) = \Phi(x, y) = 265x^3 - 50x^2y - 689xy^2 - 2120y^3\,.$$

The bound for the cosh of the distance δ to j for finding the representative of smallest height in the orbit of Φ is ≈ 14.2268. As soon as the optimum is found, this bound is reduced to ≈ 9.6546, which cuts down the number of points in the search space to 54. We find that

$$\Phi(x-y,y) = 265x^3 - 845x^2y + 206xy^2 - 1746$$

has both smallest height and smallest size in the $\mathrm{SL}(2,\mathbb{Z})$-orbit of Φ. This corresponds to the endomorphism

$$(x:y) \longmapsto (-315x^2 - 165xy - 1746y^2 : -265x^2 + 530xy - 371y^2)\,.$$

of height 1746. However, this is not the representative of smallest height among the $\mathrm{SL}(2,\mathbb{Z})$-conjugates of f: running the algorithm with the modifications indicated above, we get an initial bound of ≈ 35.5547 for $\cosh\delta$, which gets reduced to ≈ 19.7017. The algorithm runs through 118 points $\gamma^{-1}\cdot z(\Phi)$, showing that a representative of smallest height is obtained for the conjugate

$$f^\gamma = [480x^2 + 1125xy - 1578y^2 : -265x^2 - 1060xy - 1166y^2]$$

of height 1578 for $\gamma = \left(\begin{smallmatrix}1&2\\0&1\end{smallmatrix}\right)$.

7 Orbits of minimal models

Questions about the structure of the set of minimal models of a dynamical system, including how to calculate a minimal model, have been studied in two different contexts. Bruin-Molnar [BM12], in addressing questions about integer points in orbits, consider the problem over number fields. Rumely [Rum15] approaches the problem over a complete algebraically closed nonarchimedean field and applies Berkovich space methods to solve the problem. We will use input from both publications to devise an algorithm that finds representatives of all $\mathrm{GL}(2,\mathbb{Z})$-orbits of minimal models of a given endomorphism. Furthermore, Rumely shows that the valuation of the resultant factors through a map from the Berkovich projective line $\mathbb{P}^1_{\mathrm{Berk}}$ to $\mathbb{R}$ and proves that for even degree maps, the minimal valuation is achieved at a single point in $\mathbb{P}^1_{\mathrm{Berk}}$. However, this point does not necessarily correspond to a model defined over $\mathbb{Q}_p$. We will use his results to show that also over $\mathbb{Q}_p$, the minimal valuation of the resultant is obtained for a unique $\mathrm{GL}(2,\mathbb{Z}_p)$-orbit of models; see Proposition 7.2 below.

To determine a representative from each orbit, we recall the key points of the algorithm of Bruin-Molnar and then adapt it to our purposes. Given a model $[F:G]$ of an endomorphism f, where $F, G \in \mathbb{Q}[x,y]_d$, we can scale F and G by some $\lambda \in \mathbb{Q}^\times$, which results in the model $[\lambda F, \lambda G]$ of the same f, and we can conjugate it by some element γ of $\mathrm{GL}(2,\mathbb{Q})$. In section 6, we defined an action of $\gamma \in \mathrm{GL}(2,\mathbb{Q})$ on f that avoids introducing denominators. As in Bruin-Molnar, these two operations combine to act on $[F:G]$ by $(\lambda, \gamma = \left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right))$ resulting in

$$[\lambda F^\gamma : \lambda G^\gamma] = \Big[\lambda\big(dF(ax+by,cx+dy) - bG(ax+by,cx+dy)\big) : \lambda\big(-cF(\ldots) + aG(\ldots)\big)\Big]\,.$$

Acting by $(\lambda^{-d-1}, \lambda I_2)$ has no effect (here I_2 denotes the 2×2 identity matrix), so we can assume γ to have integral entries. Recall that we are interested in *minimal* models of f, which are integral models with minimal absolute value of the resultant. According to [BM12, Proposition 2.2], we have that

$$\operatorname{Res}(\lambda F^\gamma, \lambda G^\gamma) = \lambda^{2d} \det(\gamma)^{d^2+d} \operatorname{Res}(F,G)\,. \tag{7.1}$$

Each prime can be considered separately, and [BM12, Proposition 6.3] shows that we need only consider primes that divide the resultant of the given integral model. We can therefore make the resultant smaller if, for a prime p dividing the resultant, we can find $\gamma \in \mathrm{Mat}(2,\mathbb{Z}) \cap \mathrm{GL}(2,\mathbb{Q})$ such that the gcd of the coefficients of F^γ and G^γ is α with $2v_p(\alpha) > (d+1)v_p(\det(\gamma))$, where $v_p(\cdot)$ is the (normalized) p-adic valuation.

Further, [BM12, Proposition 2.12] shows that it is sufficient to consider affine transformations. Let p be a prime dividing the resultant and consider the affine transformation and scaling factor

$$\gamma = \begin{pmatrix} p^{e_2} & \beta \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad \lambda = p^{-e_1}\,.$$

In this form, the condition for the power of p dividing the resultant to decrease then becomes

$$2e_1 > (d+1)e_2\,, \tag{7.2}$$

where e_1 is the minimal p-adic valuation of a coefficient of

$$F^\gamma = F(p^{e_2}x + \beta y, y) - \beta G(p^{e_2}x + \beta y, y) \quad \text{or} \quad G^\gamma = p^{e_2} G(p^{e_2}x + \beta y, y)\,.$$

We now make use of Rumely's results from [Rum15]. We write $\mathbb{P}^1_{\mathrm{Berk}}$ for the Berkovich projective line over $\mathbb{C}_p$, the completion of the algebraic closure of $\mathbb{Q}_p$, and we denote the "Berkovich upper half space" $\mathbb{P}^1_{\mathrm{Berk}} \setminus \mathbb{P}^1(\mathbb{C}_p)$ by $\mathcal{H}_{\mathrm{Berk}}$. Fix an endomorphism f of $\mathbb{P}^1$ over $\mathbb{C}_p$. Rumely shows that there is a continuous, piecewise affine with integral slopes (with respect to the logarithmic distance on $\mathcal{H}_{\mathrm{Berk}}$) and convex map $\mathrm{ordRes}_f \colon \mathcal{H}_{\mathrm{Berk}} \to \mathbb{R}_{\geq 0}$ such that $v_p(\mathrm{Res}(f^\gamma)) = \mathrm{ordRes}_f(\gamma \cdot \zeta_G)$ for all $\gamma \in \mathrm{GL}(2,\mathbb{C}_p)$, where $\zeta_G \in \mathbb{P}^1_{\mathrm{Berk}}$ is the Gauss point. Here, $\mathrm{Res}(f^\gamma)$ denotes the resultant of a representative $[F : G]$ of f^γ that is scaled so that F and G have p-adically integral coefficients with one of them a unit (i.e., we take the maximal e_1 in the notation above). The orbit of ζ_G under $\mathrm{GL}(2,\mathbb{Q}_p)$ consists of the set $\mathcal{V}$ of vertices of a subtree $\mathcal{T}$ of $\mathcal{H}_{\mathrm{Berk}}$ whose edges have length 1 in the logarithmic metric; the vertices have degree $p+1$.

The $\mathrm{GL}(2,\mathbb{Z}_p)$-orbits of minimal models of f then correspond to the points in $\mathcal{V}$ in which $\mathrm{ordRes}_f|_{\mathcal{V}}$ takes its minimal value. Note that it is possible that ordRes_f will take on a smaller value at a point of $\mathcal{T}$ which is not a vertex, corresponding to a model defined over an extension of $\mathbb{Q}_p$, see Example 7.3. The restriction of ordRes_f to $\mathcal{T}$ is still piecewise affine and convex. This implies that a point in $\mathcal{V}$ is a minimizer of $\mathrm{ordRes}_f|_{\mathcal{V}}$ if (and only if) it is a local minimum, in the sense that ordRes_f does not take a strictly smaller value in a neighboring vertex. It also implies that at each vertex, there is at most one edge leading to a vertex with strictly smaller value, and following these edges leads to a minimum. As noted by Rumely, this can be used to simplify the Bruin-Molnar algorithm. We obtain the following procedure for finding a p-adically minimal model.

We assume that a *normalized* model $f = [F : G]$ is given, i.e., such that $F, G \in \mathbb{Z}_p[x,y]_d$, where at least one coefficient is a p-adic unit.

Algorithm 7.1.

1. Let $T = \left\{ \left(\begin{smallmatrix} 1 & 0 \\ 0 & p \end{smallmatrix}\right) \right\} \cup \left\{ \left(\begin{smallmatrix} p & a \\ 0 & 1 \end{smallmatrix}\right) : a \in \mathbb{Z},\ 0 \leq a < p \right\}$. Set $\gamma_0 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$.

2. For $\gamma \in T$, compute $f^\gamma = [\lambda F^\gamma : \lambda G^\gamma]$, where λ is chosen so that the resulting model is normalized.
   If we come from step 3, then we leave out the one γ that would bring us back to the $\mathrm{GL}(2, \mathbb{Z}_p)$-orbit of the previously considered model.
3. If $v_p(\mathrm{Res}(f^\gamma)) < v_p(\mathrm{Res}(f))$ for some γ, then replace f with f^γ and γ_0 with $\gamma_0\gamma$.
   If $v_p(\mathrm{Res}(f^\gamma)) \geq d$ for d even or $\geq 2d$ for d odd, then go to step 2.
4. Otherwise, return f and γ_0.

The set T contains representatives γ that map ζ_G to each of its $p+1$ neighbors in $\mathcal{T}$. If we have used $\gamma = \left(\begin{smallmatrix} p & a \\ 0 & 1 \end{smallmatrix}\right)$ to reduce the valuation of the resultant, then we exclude $\left(\begin{smallmatrix} 1 & 0 \\ 0 & p \end{smallmatrix}\right)$ when we carry out step 2 the next time; if we have used $\gamma = \left(\begin{smallmatrix} 1 & 0 \\ 0 & p \end{smallmatrix}\right)$, then we exclude $\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Note that from (7.1), we can deduce that the difference of the values of ordRes_f in neighboring vertices of $\mathcal{T}$ is divisible by d when d is even and by $2d$ when d is odd. So if the value for the model we are considering is smaller than d or $2d$, respectively, then the model must be minimal.

This algorithm is essentially an enumerative approach similar to both the algorithm in Rumely [Rum15] and Bruin-Molnar [BM12]. In the case of Rumely, there is additional logic to limit the number of directions to consider and to compute how far to move in each direction. In the case of Bruin-Molnar, a set of inequalities is solved to determine the a to use in step 2. For reasonably small primes, this enumerative approach is sufficient. For larger primes, a version of the Bruin-Molnar inequality solver can be used to speed up steps 2 and 3.

If the degree d of f is even, then (7.1) shows that the change of ordRes_f along each edge is a linear combination of $d^2 + d$ and $2d$ and so is $\equiv d \bmod 2d$, hence, is never zero. Consequently, if we have found a minimizing vertex, then ordRes_f will strictly increase along all edges emanating from it, which implies that this vertex is the unique minimizer. Since this holds for each prime p, we obtain a proof of the following statement, which answers Question 6.2 in [BM12] in the affirmative in the strongest possible sense. (The argument is already in [Rum15, p. 280], if somewhat implicit.)

Proposition 7.2. *Let $f\colon \mathbb{P}^1 \to \mathbb{P}^1$ be defined over $\mathbb{Q}$. If the degree of f is even, then f has a single $\mathrm{GL}(2, \mathbb{Z})$-orbit of minimal models.*

It is possible that the minimal value of ordRes_f on $\mathcal{H}_{\mathrm{Berk}}$ is not attained on $\mathcal{T}$; see Example 6.2 in [Rum15]. The following example shows that it is also possible that the minimum is attained on $\mathcal{T}$, but not on $\mathcal{V}$.

Example 7.3. Looking more closely at Example 6.1 from Rumely [Rum15], we consider the endomorphism

$$f\colon \quad \mathbb{P}^1 \longrightarrow \mathbb{P}^1$$
$$(x : y) \longmapsto [x^2 - py^2 : xy],$$

where p is a prime. In this case, the minimum value of ordRes_f over the algebraic closure occurs between two vertices of $\mathcal{T}$. Specifically, $\mathrm{Res}(f) = -p$ is minimal over $\mathrm{GL}(2, \mathbb{Q}_p)$, but going to the ramified quadratic extension $\mathbb{Q}_p(\sqrt{p})$, for $\gamma = \left(\begin{smallmatrix} \sqrt{p} & 0 \\ 0 & 1 \end{smallmatrix}\right)$ we have, after normalizing, $\mathrm{Res}(f^\gamma) = -1$. The next vertex of $\mathcal{T}$ that lies in this direction corresponds to conjugating by $\alpha = \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)$ and has $\mathrm{Res}(f^\alpha) = -p^3$. In particular, the function ordRes_f has slope $-d = -2$ for half of the edge containing the minimum f^α and $3d = 6$ for the remaining half, giving a net increase of $d = 2$ along the edge.

Rumely shows that when the degree d is odd, ordRes_f achieves its minimum value at a unique point or along an interval in $\mathcal{H}_{\mathrm{Berk}}$. If this set meets $\mathcal{V}$, then the intersection is the set of minimizers of $\mathrm{ordRes}_f|_{\mathcal{V}}$, and this set is either one point or consists of the vertices in a path in $\mathcal{T}$. If the minimizing set meets $\mathcal{T}$, but not $\mathcal{V}$, then the intersection with $\mathcal{T}$ must be contained in the interior of an edge of $\mathcal{T}$. Otherwise, Proposition 3.5 of [Rum15] implies that $\mathrm{ordRes}_f|_{\mathcal{T}}$ has a unique minimum, which is attained at an interior point of an edge. In both of these last two cases, the set of minimizers of $\mathrm{ordRes}_f|_{\mathcal{V}}$ consists either of one or both of the endpoints of this edge. (We can indeed have two minimizing vertices in this case; this is what happens in Example 6.4 in [Rum15].) So in all cases, the subset of $\mathcal{V}$ corresponding to $\mathrm{GL}(2, \mathbb{Z}_p)$-orbits of p-minimal models of f is either one point or consists of the vertices in a path. This path can have any length. This is demonstrated by the following example, which extends Example 6.1 of [BM12].

Proposition 7.4. *Let $0 \neq c \in \mathbb{Z}$ and n a positive integer. Define*

$$\begin{aligned} f\colon \qquad \mathbb{P}^1 &\longrightarrow \mathbb{P}^1 \\ (x : y) &\longmapsto [x^{2n+1} - c^{n+1}y^{2n+1} : x^n y^{n+1}]. \end{aligned}$$

Let $c = \pm \prod_{i=1}^{s} p_i^{e_i}$ be the prime factorization of c. Then f has exactly $\prod_{i=1}^{s}(e_i + 1)$ distinct $\mathrm{GL}(2, \mathbb{Z})$-orbits of minimal models, one for each positive divisor of c.

Proof. Write $c = rs$ with $r, s \in \mathbb{Z}$. Acting on the given model by $(\lambda, \gamma) = \left(r^{-n-1}, \left(\begin{smallmatrix} r & 0 \\ 0 & 1 \end{smallmatrix}\right)\right)$, we obtain the minimal model $[r^n x^{2n+1} - s^{n+1}y^{2n+1} : x^n y^{n+1}]$. The models associated to the factorizations $c = rs$ and $c = r's'$ are in the same $\mathrm{GL}(2, \mathbb{Z})$-orbit if and only if the corresponding matrices differ multiplicatively by a scalar multiple of a matrix in $\mathrm{GL}(2, \mathbb{Z})$, which is the case if and only if $r' = \pm r$. So we obtain a distinct $\mathrm{GL}(2, \mathbb{Z})$-orbit of minimal models for each positive divisor r of c.

To see that these models cover all $\mathrm{GL}(2, \mathbb{Z})$-orbits, we consider an affine transformation $\gamma = \left(\begin{smallmatrix} \alpha & \beta \\ 0 & \delta \end{smallmatrix}\right)$ with $\alpha, \beta, \delta \in \mathbb{Z}$ coprime and $\alpha, \delta > 0$. Since we are interested in orbits under $\mathrm{GL}(2, \mathbb{Z})$, we can assume that $|\beta| < \alpha$. Denoting the polynomials in the given model by F and G, we have that

$$F^\gamma = \delta(\alpha x + \beta y)^{2n+1} - \delta^{2n+2}c^{n+1}y^{2n+1} - \beta(\alpha x + \beta y)^n \delta^{n+1} y^{n+1}$$

and

$$G^\gamma = \alpha(\alpha x + \beta y)^n \delta^{n+1} y^{n+1}.$$

If this is to lead to a minimal model, $\det(\gamma)^{n+1} = \alpha^{n+1}\delta^{n+1}$ must divide all coefficients of F^γ and G^γ. Considering the y^{2n+1} term in G^γ, we see that α must divide β, so that $\beta = 0$. Then α and δ are coprime, and

$$F^\gamma = \alpha^{2n+1}\delta x^{2n+1} - \delta^{2n+2}c^{n+1}y^{2n+1},$$

so δ divides α and α divides c. Since α and δ are coprime, this means that $\delta = 1$ and $\alpha = r$ with $c = rs$ a factorization as above with $r > 0$. □

Remark 7.5. Note that $c = \pm 1$ results in a map with a single orbit, so it is possible for odd degree maps to have a single orbit of minimal models.

To get representatives of *all* $\mathrm{GL}(2, \mathbb{Z}_p)$-orbits of minimal models of a given dynamical system $f = [F : G]$, we modify the algorithm in the following way. Traverse the vertices of $\mathcal{T}$ until the minimal value of ordRes_f is attained. Then, find all vertices of $\mathcal{T}$ that attain that minimum; these will all lie on a path in $\mathcal{T}$, so can be determined one edge at a time. Note that the first minimal model we find could correspond to a vertex in the interior of this path, so we may have to search in two directions to find all vertices in the path. The algorithm below returns a set of pairs (f^γ, γ) with $\gamma \in \mathrm{GL}(2, \mathbb{Q}_p)$ such that the f^γ represent all $\mathrm{GL}(2, \mathbb{Z}_p)$-conjugacy classes of minimal models of f.

Algorithm 7.6.

1. Let $T = \left\{ \left(\begin{smallmatrix} 1 & 0 \\ 0 & p \end{smallmatrix}\right) \right\} \cup \left\{ \left(\begin{smallmatrix} p & a \\ 0 & 1 \end{smallmatrix}\right) : a \in \mathbb{Z},\ 0 \le a < p \right\}$.
2. Use Algorithm 7.1 to find a minimal model $f_0 = f^{\gamma_0}$ of f.
   Set $M \leftarrow \{(f_0, \gamma_0)\}$ and $f \leftarrow f_0$.
3. Determine $S = \{\gamma \in T : v_p(\mathrm{Res}(f^\gamma)) = v_p(\mathrm{Res}(f))\}$.
4. For each element $\gamma \in S$ (there are at most two), call **search**$(f^\gamma, \gamma_0\gamma, \gamma)$.
5. Return M.

search(f, γ_0, γ):

a. Set $M \leftarrow M \cup \{(f, \gamma_0)\}$.
b. Determine $S = \{\gamma' \in T : \gamma\gamma' \notin \mathrm{GL}(2, \mathbb{Z}_p), v_p(\mathrm{Res}(f^{\gamma'})) = v_p(\mathrm{Res}(f))\}$.
c. If $S = \{\gamma'\}$, then set $(f, \gamma_0, \gamma) \leftarrow (f^{\gamma'}, \gamma_0\gamma', \gamma')$; go to step a.
d. Otherwise, $S = \emptyset$. Return.

In the call to **search**, the last argument γ specifies the direction we come from. See the discussion after Algorithm 7.1 for what the condition $\gamma\gamma' \notin \mathrm{GL}(2, \mathbb{Z}_p)$ amounts to in terms of directions in Berkovich space.

To get representatives of all $\mathrm{GL}(2\,\mathbb{Z})$-orbits of minimal models, we run the above algorithm for each prime dividing the resultant of the given model, applying the returned γ_0 to each of the models obtained so far.

We finally note that any $\mathrm{GL}(2, \mathbb{Z})$-orbit splits into at most two $\mathrm{SL}(2, \mathbb{Z})$-orbits and that the action of $\left(1, \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)\right)$ preserves the height of the model. It is therefore sufficient to look at the $\mathrm{SL}(2, \mathbb{Z})$-orbits of the representatives of the $\mathrm{GL}(2, \mathbb{Z})$-orbits when we want to find a reduced model.

References

[BCP97] Wieb Bosma, John Cannon, and Catherine Playoust, *The Magma algebra system. I. The user language*, J. Symbolic Comput. **24** (1997), no. 3-4, 235–265, DOI 10.1006/jsco.1996.0125. Computational algebra and number theory (London, 1993). MR1484478 ↑5

[BM12] Nils Bruin and Alexander Molnar, *Minimal models for rational functions in a dynamical setting*, LMS J. Comput. Math. **15** (2012), 400–417. MR3015733 ↑1, 2, 2, 6, 7, 7, 7, 7

[Her48] C. Hermite, *Note sur la réduction des fonctions homogènes à coefficients entiers et à deux indéterminées*, J. Reine Angew. Math. **36** (1848), 357–364 (French). MR1578622 ↑3

[Hut15] Benjamin Hutz, *Determination of all rational preperiodic points for morphisms of PN*, Math. Comp. **84** (2015), no. 291, 289–308. MR3266961 ↑2

[HM14] Benjamin Hutz and Michelle Manes, *The field of definition for dynamical systems on* $\mathbb{P}^N$, Bull. Inst. Math. Acad. Sin. (N.S.) **9** (2014), no. 4, 585–601. MR3309942 ↑2

[Ing12] Patrick Ingram, *A finiteness result for post-critically finite polynomials*, Int. Math. Res. Not. IMRN **3** (2012), 524–543. MR2885981 ↑2

[Jul17] Gaston Julia, *Étude sur les formes binaires non quadratiques à indéterminées réelles, ou complexes, ou à indéterminées conjuguées*, Mém. Acad. Sci. Inst. France, vol. 55, 1917 (French). Also in Julia's Œuvres, vol. 5. MR3532882 ↑3

[Lev11] Alon Levy, *The space of morphisms on projective space*, Acta Arith. **146** (2011), no. 1, 13–31. MR2741188 ↑2

[Mil93] John Milnor, *Geometry and dynamics of quadratic rational maps*, Experiment. Math. **2** (1993), no. 1, 37–83. With an appendix by the author and Lei Tan. MR1246482 ↑2

[Rum15] Robert Rumely, *The minimal resultant locus*, Acta Arith. **169** (2015), no. 3, 251–290, DOI 10.4064/aa169-3-3. MR3361223 ↑7, 7, 7, 7, 7.3, 7

[Sil07] Joseph H. Silverman, *The arithmetic of dynamical systems*, Graduate Texts in Mathematics, vol. 241, Springer, New York, 2007. MR2316407 ↑2, 6

[Sil95] ———, *The field of definition for dynamical systems on* $\mathbf{P}^1$, Compositio Math. **98** (1995), no. 3, 269–304. MR1351830 ↑2

[SJ05] William Stein and David Joyner, *SAGE: System for Algebra and Geometry Experimentation*, Communications in Computer Algebra (SIGSAM Bull.) **39** (2005), no. 4, 61–64. http://www.sagemath.org. ↑2, 5

[Sto11] Michael Stoll, *Reduction theory of point clusters in projective space*, Groups Geom. Dyn. **5** (2011), no. 2, 553–565, DOI 10.4171/GGD/139. MR2782185 ↑1

[SC03] Michael Stoll and John E. Cremona, *On the reduction theory of binary forms*, J. Reine Angew. Math. **565** (2003), 79–99. MR2024647 ↑(document), 1, 2, 3, 5, 5.2

Department of Mathematics and Statistics, Saint Louis University, St. Louis, MO

E-mail address: benjamin.hutz@slu.edu

Mathematisches Institut, Universität Bayreuth, 95440 Bayreuth, Germany.

E-mail address: Michael.Stoll@uni-bayreuth.de

URL: http://www.mathe2.uni-bayreuth.de/stoll/