EXPLORING RATIONAL FUNCTIONS ON THE PROJECTIVE LINE THROUGH MACHINE LEARNING TECHNIQUES

ESLAM BADR, ELIRA SHASKA, AND TONY SHASKA

Abstract. We apply machine learning to study the moduli space $\mathcal{M}_d^1$ of degree $d \geq 3$ rational functions of the projective line, the automorphism loci in this space, and the distribution of fine moduli points of $\mathcal{M}_d^1$. Both unsupervised and supervised learning techniques are used. As an illustrating example we focus on the case of degree $d = 3$, since $\mathcal{M}_3^1$ is well understood. We consider such space as a weighted projective space $\mathbb{P}_\omega$, for weights $\omega = (2, 2, 3, 3, 4, 6)$. Using a K-Neighbors Classifier we apply weighted projective clustering in this space and are able to detect the automorphism group of rational cubics with an accuracy of 99.9%. To our knowledge this is the first time that machine learning techniques are used in the arithmetic of dynamical systems and the approach provides an effective and exciting direction in the area. There is no reason why such methods can not be extended to higher degree rational functions.

1. Introduction

Let k be an algebraically closed field of characteristic zero and $\mathbb{P}_k^1$ the projective line over k. A degree $d \geq 2$ rational function $\phi : \mathbb{P}^1 \to \mathbb{P}^1$ is given as the ratio of two degree d binary forms, say $\phi(x,y) = \frac{f_0(x,y)}{f_1(x,y)}$ such that the resultant between $f_0(x,y)$ and $f_1(x,y)$ is non-zero. Hence, a rational function is a pair of binary forms of the same degree with no common roots. If we denote

$$f_0(x,y) = \sum_{i=0}^{d} a_i x^i y^{d-i} \quad \text{and} \quad f_1(x,y) = \sum_{i=0}^{d} b_i x^i y^{d-i},$$

then the collection of pairs $[f_0 : f_1]$ can be parametrized by the projective space $\mathbb{P}^{2d+1}$ via

$$[a_d : \cdots : a_0, b_d : \cdots : b_0]$$

such that $\operatorname{Res}(f_0, f_1) \neq 0$. So the parameter space of degree $d > 1$ rational functions on $\mathbb{P}^1$ is the complement of the resultant locus in $\mathbb{P}^{2d+1}$. We will denote this space by $\operatorname{Rat}_d := \mathbb{P}^{2d+1} \setminus V(\operatorname{Res})$. The goal of this paper is to use machine learning to study the space Rat_d following techniques introduced in [1] for binary forms and hyperelliptic curves.

Hence, we will need to create a database of points in Rat_d. In order to order such data we use the concept of the projective height similarly to what we used for polynomials in [1] and [2]. For the rest of this paper this database will be denoted by $\mathcal{P}_d$.

2020 *Mathematics Subject Classification.* 68T07, 68Q32, 37P05, 37P30, 37P45 .

A machine learning model on $\mathcal{P}_d$ would require using equivariant neural networks since there are points in $\mathcal{P}_d$ which will represent equivalent rational functions. Instead we use the approach in [1] where we determine the invariants a priori and the machine learning models are done in the space of invariants using weighted projective clustering as briefly explained below.

The group $\mathrm{SL}_2(k)$ acts on $\mathcal{P}_d$ by conjugation, i.e. for some $M \in \mathrm{SL}_2(k)$

$$\phi \to \phi^M := M^{-1} \circ \phi \circ M.$$

Two rational functions $\phi, \psi \in \mathrm{Rat}_d$ are called *equivalent* if there is an $M \in \mathrm{SL}_2(k)$ such that $\phi = \psi^M$. The moduli space of degree $d > 1$ rational functions (in one variable) is denoted by $\mathcal{M}_d^1$ and can be constructed as a quotient space of this SL_2-action.

For any $\phi(x,y) = \frac{f_0(x,y)}{f_1(x,y)}$ we define a pair of binary forms of degree $d+1$ and $d-1$ (cf. Eq. (15))

$$\mathcal{I}_f := y f_0 - x f_1 \quad \text{and} \quad \mathcal{J}_f := \frac{\partial f_0}{\partial x} + \frac{\partial f_1}{\partial y}$$

Any two degree d rational functions ϕ and ψ are equivalent for some $M \in \mathrm{PGL}_2(k)$ via $\psi = \phi^M$ if and only if $\mathcal{I}_\psi = \mathcal{I}_\phi^M$ and $\mathcal{J}_\psi = \mathcal{J}_\phi^M$; see Thm. 1. Moreover, there is a one to one correspondence between degree d rational functions $\phi(x)$ and points $(f, g) \in V_{d+1} \oplus V_{d-1}$ such that

$$\mathrm{Res}\left(xg + \frac{\partial f}{\partial y}, yg - \frac{\partial f}{\partial x} \right) \neq 0,$$

Hence, determining invariants of rational functions is the same as determining generators for the ring of invariants of $V_{d+1} \oplus V_{d-1}$. Such generators can be determined using a result of Clebsch once generators of the ring of invariants for V_{d+1} and V_{d-1} are known.

Denote by $\mathcal{R}_{(d+1),(d-1)}$ the ring of invariants of $V_{d+1} \oplus V_{d-1}$ and $(\xi_0, \ldots, \xi_n)$ the tuple of generators of this ring with degrees $q_0, \ldots, q_n)$ respectively. Since all $\xi_0, \ldots \xi_n$ are homogenous polynomials then $\mathcal{R}_{(d+1,d-1)}$ is a graded ring and $\mathrm{Proj}\, \mathcal{R}_{(d+1,d-1)}$ is a weighted projective space denoted by $\mathbb{WP}^n_\omega(k)$, where $\omega = (q_0, \ldots, q_n)$ is the set of weights. Thus for each $\phi \in \mathcal{P}_d$ we evaluate its invariants and have a map $\xi : \mathcal{P}_d \to \mathcal{M}_d^1$ via

$$\phi \to [\xi_0(\phi) : \cdots : \xi_n(\phi)]$$

and create a new database $\mathcal{W}_d$ where each entry represents the isomorphism class of some rational function. We can normalize points in $\mathcal{W}_d$ by "dividing" by the weighted gcd as explained in [3]. Then using the weighted projective height we can further order them; see [3] and [4]. Obviously, the data in $\mathcal{W}_d$ doesn't necessarily have to come from $\mathcal{P}_d$, since we can randomly generate points in the weighted projective space $\mathbb{WP}^n_\omega$ similarly how it was done in [1]. The benefit of training the data $\mathcal{W}_d$ instead of $\mathcal{P}_d$ is that there is no redundancy because every data points represents uniquely the equivalence class of a rational function.

The automorphism group of ϕ is defined as $\mathrm{Aut}(\phi) := \{\sigma \in \mathrm{PGL}_2(k) \, : \, \phi^\sigma = \phi\}$. It is a finite subgroup of of $\mathrm{PGL}_2(k)$ then it is isomorphic to one of the following: a cyclic group, a dihedral group, A_4, S_4, or A_5. Determining which one of these groups occur for a fixed degree $d \geq 2$ is known, due to [5, 6] where the loci of such spaces are considered in the moduli space of degree d rational functions.

Our main task in this paper is to design a machine learning model with determines properties of rational functions such that the automorphism group, inclusions among the automorphism loci, a minimal field of definition, etc. A KNN algorithm for example can be used to train the data on determining the automorphism group. However, since our data is in a weighted projective space a weighted projective clustering will be used. This is a new approach first introduced in [1].

As an application to our methods we focus on the case of rational cubics since this is a well known case and we can compare our results with known theoretical facts. Higher degree cases are planned in a subsequent paper. The isomorphism class of cubic rational functions is determined by a set of generators of $V_4 \oplus V_2$, which we explicitly determine as invariants of degree 2, 2, 3, 3, 4, and 6. Hence, the moduli space $\mathcal{M}_3^1$ is the weighted projective space $\mathbb{WP}_\omega(k)$ for the set of weights $\omega = (2, 2, 3, 3, 4, 6)$. A complete list of automorphism groups and the corresponding locus in terms of the invariants are computed.

After experimenting with many methods of unsupervised and supervised learning we were able to detect the automorphism group using K-neighbours Classifier with an accuracy of 99.97 %. There is no reason to believe that this accuracy will go down for higher degree d. Our experiments were run for randomly generated dataset of 5 million points. It seems as with more computational resources such results can be easily extended to larger degree d.

2. Preliminaries

Let k be an algebraically closed field, $\mathbb{P}^N(k)$ the projective space over k, and $k[x, y]$ be the polynomial ring in two variables. By V_d we denote the $(d+1)$-dimensional subspace of $k[x, y]$ consisting of homogeneous polynomials

$$f(x, y) = a_d x^d + a_{d-1} x^d y + \cdots a_1 x y^{d-1} + a_0 y^d,$$

of degree d. Elements in V_d (up to multiplication by a constant) are called *binary forms* of degree d. $\mathrm{GL}_2(k)$ acts as a group of automorphisms on $k[x, y]$ as follows:

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(k), \;\; \textit{then} \quad M \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix} \tag{1}$$

Denote by f^M the binary form

$$f^M(x, y) := f(ax + by, cx + dy).$$

It is well known that $\mathrm{SL}_2(k)$ leaves a bilinear form (unique up to scalar multiples) on V_d invariant.

Consider a_0, a_1, ... , a_d as parameters (coordinate functions on V_d). Then the coordinate ring of V_d can be identified with $k[a_0, ..., a_d]$. For $I \in k[a_0, ..., a_d]$ and $M \in \mathrm{GL}_2(k)$, define

$$I^M \in k[a_0, \dots, a_d] \quad \text{as} \quad I^M(f) := I(f^M),$$

for all $f \in V_d$. Then $I^{MN} = (I^M)^N$ and $I^M(f)$ define an action of $\mathrm{GL}_2(k)$ on $k[a_0, \dots, a_d]$.

A homogeneous polynomial $I \in k[a_0, \dots, a_d, x, y]$ is called a **covariant** of index s if $I^M(f) = \delta^s I(f)$, for all $f \in V_d$, where $\delta = \det(M)$. The homogeneous degree in $a_0, \dots, a_d$ is called the **degree** of I, and the homogeneous degree in x, y is called the **order** of I. A covariant of order zero is called **invariant**. An invariant is a $\mathrm{SL}_2(k)$-invariant on V_d.

From Hilbert's theorem the ring of invariants of binary forms is finitely generated. We denote by $\mathcal{R}_d$ the ring of invariants of the binary forms of degree d. Then, $\mathcal{R}_d$ is a finitely generated graded ring

2.1. **Changing of coordinates.** Let $I_0, \dots, I_n$ be the generators of $\mathcal{R}_d$ with degrees $q_0, \dots, q_n$ respectively. For any two binary forms f and g, $f = g^M$, $M \in \mathrm{GL}_2(k)$, if and only if

$$(I_0(f), \dots I_i(f), \dots, I_n(f)) = (\lambda^{q_0}\, I_0(g), \dots, \lambda^{q_i}\, I_i(g), \dots, \lambda^{q_n}\, I_n(g)), \tag{2}$$

where $\lambda = (\det M)^{\frac{d}{2}}$.

2.2. **Generators of the ring of invariants.** Let V_d be the space of degree $d > 1$ binary forms defined over k, and $\mathcal{R}_d$ the ring of invariants. Below we list the generating set of $\mathcal{R}_d$ for $d \leq 10$. We assume that the binary forms are given in standard form

$$f(x, y) = \sum_{i=0}^{d} \binom{d}{i} a_i x^i y^{d-i}$$

For $f, g \in V_d$ the r-th transvectant of f and g is defined as

$$(f, g)_r := \frac{(m-r)!\,(n-r)!}{n!\,m!} \sum_{k=0}^{r} (-1)^k \binom{r}{k} \cdot \frac{\partial^r f}{\partial x^{r-k}\, \partial y^k} \cdot \frac{\partial^r g}{\partial x^k\, \partial y^{r-k}},$$

While there is no method known to determine a minimal generating set of invariants for any $\mathcal{R}_d$, we display such sets for $3 \leq d \leq 5$ as in [7].

2.2.1. *Cubics.* A generating set for $\mathcal{R}_3$ is $\xi = \{\xi_0\}$, where

$$\xi_0 = ((f, f)_2, (f, f)_2)_2 = -54a_0^2 a_3^2 + 36a_1a_3a_0a_2 - 8a_2^3 a_0 - 8a_1^3 a_3 + 2a_2^2 a_1^2 \tag{3}$$

2.2.2. *Quartics.* A generating set for $\mathcal{R}_4$ is $\xi = [\xi_0, \xi_1]$ with $\omega = (2, 3)$, where

$$\xi_0 = (f, f)_4 \ \text{ and } \ \xi_1 = (f, (f, f)_2)_4 \tag{4}$$

2.2.3. *Quintics.* A generating set for $\mathcal{R}_4$ is $\xi = [\xi_0, \xi_1, \xi_2]$ with $\omega = (4, 8, 12)$, where

$$\xi_0 = (c_1, c_1)_2, \quad \xi_1 = (c_4, c_1)_2, \quad \xi_2 = (c_4, c_4)_2, \tag{5}$$

for $c_1 = (f, f)_4$, $c_2 = (f, f)_2$, $c_3 = (f, c_1)_2$, $c_4 = (c_3, c_3)_2$.

2.2.4. *Sextics.* The case of sextics was studied in detail by XIX-century mathematicians (Bolza, Clebsch, et al.) when char $k = 0$ and by Igusa for char $k > 0$. Let $c_1 = (f, f)_4$, $c_3 = (f, c_1)_4$, $c_4 = (c_1, c_1)_2$. A generating set for $\mathcal{R}_6$ is $\xi = [\xi_0, \xi_1, \xi_2, \xi_3]$ with weights $\omega = (2, 4, 6, 10)$, where

$$\xi_0 = (f, f)_6, \ \xi_1 = (c_1, c_1)_4, \ \xi_2 = (c_4, c_1)_4, \ \xi_3 = (c_4, c_3^2)_4 \tag{6}$$

Usually the invariants of binary sextics are denoted by $[J_2, J_4, J_6, J_{10}]$ with J_{10} being the discriminant of the sextic, but that is not the case here.

2.2.5. *Septics.* A generating set of $\mathcal{R}_7$ is given by $\xi = [\xi_0, \xi_1, \xi_2, \xi_3, \xi_4]$ with weights $\omega = (4, 8, 12, 12, 20)$. We define them as follows. Let

$$c_1 = (f, f)_6, \quad c_2 = (f, f)_4, \quad c_4 = (f, c_1)_2, \quad c_5 = (c_2, c_2)_4, \quad c_7 = (c_4, c_4)_4 \tag{7}$$

$$\begin{aligned} &\xi_0 = (c_1, c_1)_2, && \xi_1 = (c_7, c_1)_2, && \xi_2 = ((c_5, c_5)_2, c_5)_4, \\ &\xi_3 = \left((c_4, c_4)_2, c_1^3\right)_6, && \xi_4 = \left([(c_2, c_5)_4]^2, (c_5, c_5)_2\right)_4. \end{aligned}$$

2.2.6. *Octavics.* A generating set of $\mathcal{R}_8$ is given by $\xi = [\xi_0, \xi_1, \xi_2, \xi_3, \xi_4, \xi_5]$ with weights $\omega = (2, 3, 4, 5, 6, 7)$. We define them as follows. Let

$$c_1 = (f, f)_6, \quad c_2 = (f, c_1)_4, \quad c_3 = (f, f)_4, \quad c_5 = (c_1, c_1)_2. \tag{8}$$

Then the invariants are:

$$\begin{aligned} &\xi_0 = (f, f)_8, && \xi_1 = (f, c_3)_8, && \xi_2 = (c_1, c_1)_4, && \xi_3 = (c_1, c_2)_4, \\ &\xi_4 = (c_5, c_1)_4, && \xi_5 = ((c_1, c_2)_2, c_1)_4. \end{aligned}$$

2.2.7. *Nonics.* A generating set of $\mathcal{R}_9$ is given by $\xi = [\xi_0, \xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6]$ with weights $\omega = (4, 8, 10, 12, 12, 14, 16)$. Let

$$\begin{aligned} &c_1 = (f, f)_8,\ c_2 = (f, f)_6,\ c_4 = (f, f)_2,\ c_5 = (f, c_1)_2,\ c_6 = (f, c_2)_6, \\ &c_7 = (c_2, c_2)_4,\ c_9 = (c_5, c_5)_4,\ c_{21} = (f, c_2)_2,\ c_{25} = (c_4, c_4)_{10},\ c_{27} = (c_6^3, c_6)_3 \end{aligned}$$

$$\begin{aligned} &\xi_0 = (c_1, c_1)_2, && \xi_1 = (c_2, c_6^2)_6, && \xi_2 = (((c_{25}, f)_6, c_{21})_5, c_2)_6\,, \\ &\xi_3 = ((c_7, c_7)_2, c_7)_4\,, && \xi_4 = (c_9, c_1^3)_6, && \xi_5 = ((c_2, c_{27})_3)_6\,, \\ &\xi_6 = \left((c_5, c_5)_2, c_1^5\right)_{10}. \end{aligned}$$

2.2.8. *Decimics.* A generating set of $\mathcal{R}_8$ is given by $\xi = [\xi_0, \xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6, \xi_7, \xi_8]$ with weights $\omega = (2, 4, 6, 6, 8, 9, 10, 14, 14)$. Let

$$\begin{aligned} &c_1 = (f, f)_8, && c_2 = (f, f)_6, && c_5 = (f, c_1)_4, && c_6 = (f, c_2)_8, \\ &c_7 = (c_2, c_2)_6, && c_8 = (c_5, c_5)_4, && c_9 = (c_2, c_7)_4, && c_{10} = (c_1, c_1)_2, \\ &c_{16} = (c_5, c_5)_2, && c_{19} = (c_5, c_1)_1, && c_{25} = (c_7, c_7)_2 \end{aligned}$$

$$\begin{aligned} &\xi_0 = (f, f)_{10}, && \xi_1 = (c_1, c_1)_4, && \xi_2 = (c_5, c_5)_6, \\ &\xi_3 = (c_6, c_6)_2, && \xi_4 = (c_1, c_8)_4, && \xi_5 = (c_{19}, c_1^2)_8), \\ &\xi_6 = (c_{16}, c_1^2)_8, && \xi_7 = (c_{25}, c_9)_4, && \xi_8 = (c_{10}^2, c_{16})_8. \end{aligned}$$

3. Invariants of rational functions

Let k be an algebraically closed field and $\mathbb{P}^1(k)$ be the projective line over k. If one fixes homogeneous coordinates x, y on $\mathbb{P}^1$, then any rational function $\phi : \mathbb{P}^1 \to \mathbb{P}^1$ of degree $d > 1$ can be realized as

$$\phi(x, y) = \frac{f(x, y)}{g(x, y)},$$

where $d = \max\{\deg f, \deg g\}$ and f and g are homogeneous polynomials of degree at most d.

Conversely, a pair of homogeneous polynomials of degree d in x and y, determine a rational function

$$\phi(x, y) = [f(x, y) : g(x, y)] \tag{9}$$

if f, g have no common roots in k. For a fixed degree $d > 1$, the collection of all such pairs of homogeneous polynomials $[f(x, y) : g(x, y)]$, say

$$f = \sum_{i=0}^{d} a_i x^{d-i} y^i \quad \text{and} \quad g = \sum_{i=0}^{d} b_i x^{d-i} y^i,$$

can be naturally parametrized as the projective space $\mathbb{P}^{2d+1}$, via

$$[f : g] \to [a_0 : a_1 : \cdots : a_d : b_0 : \cdots : b_d] \in \mathbb{P}^{2d+1}.$$

We denote the resultant of two binary forms f and g by $\operatorname{Res}(f,g)$. Notice that it's well-defined and a degree 2d homogenous polynomial in $k[a_0, \dots, a_d, b_0, \dots, b_d]$. Let

$$I_{2d}(\phi) := \operatorname{Res}(f,g). \tag{10}$$

$I_{2d}(\phi)$ is an $\mathrm{SL}_2(k)$-invariant of degree $2d$. Moreover, ϕ is a rational function on $\mathbb{P}^1$ if and only if $I_{2d}(\phi) \neq 0$. We can construct the parameter space of rational functions on $\mathbb{P}^1$ as the complement of the vanishing locus $V(I_{2d})$ of I_{2d}. Hence the **rational space of rational functions of degree d on $\mathbb{P}^1$** is defined as

$$\mathrm{Rat}_d^1 := \mathbb{P}^{2d+1} \setminus V(I_{2d})$$

The action of $\mathrm{PGL}_2(k)$ on V_d extends naturally to an action on Rat_d^1. For each $\sigma \in \mathrm{PGL}_2(k)$ we have $\mathrm{PGL}_2(k) \times \mathrm{Rat}_d^1 \to \mathrm{Rat}_d^1$ via

$$(\sigma, \phi(x,y)) \to \phi^\sigma := \sigma^{-1}\phi\sigma,$$

Two rational functions $\phi, \psi \in \mathrm{Rat}_d^1$ are called **k-equivalent** if and only if there exists a matrix $\sigma \in \mathrm{PGL}_2(k)$ such that $\psi = \sigma^{-1} \circ \phi \circ \sigma$. We denote by

$$\phi^\sigma := \sigma^{-1} \circ \phi \circ \sigma \tag{11}$$

Remark 1. *Notice two different uses of notation ϕ^σ for rational functions and f^σ for binary forms.*

The function $\phi^\sigma = \sigma^{-1}\phi\sigma$ is explicitly given as

$$\left(\sigma^{-1}\phi\sigma\right)(x) = \frac{e\, f(ax+b, cx+e) - b\, g(ax+b, cx+e)}{-c\, f(ax+b, cx+e) + a\, g(ax+b, cx+e)} = \frac{e\, f^\sigma - b\, g^\sigma}{-c\, f^\sigma + a\, g^\sigma} \tag{12}$$

Let $\phi(x,y)$ and $\psi(x,y)$ be degree $d \geq 2$ rational functions given by

$$\phi(x,y) = \frac{f_0(x,y)}{f_1(x,y)} \quad \text{and} \quad \psi(x,y) = \frac{g_0(x,y)}{g_1(x,y)}. \tag{13}$$

By Eq. (12), ϕ and ψ are k-equivalent iff there is $\sigma = \begin{pmatrix} a & b \\ c & e \end{pmatrix} \in \mathrm{PGL}_2(k)$ such that

$$g_0 = e f_0^\sigma - b f_1^\sigma \quad \text{and} \quad g_1 = -c f_0^\sigma + a f_1^\sigma. \tag{14}$$

For $\phi(x,y) = \frac{f_0(x,y)}{f_1(x,y)}$ we define its **associated pair of binary forms** as

$$\mathcal{I}_\phi := y f_0 - x f_1 \quad \text{and} \quad \mathcal{J}_\phi := \frac{\partial f_0}{\partial x} + \frac{\partial f_1}{\partial y} \tag{15}$$

Notice that $\mathcal{I}_\phi \in V_{d+1}$ and $\mathcal{J}_\phi \in V_{d-1}$.

Lemma 1. *Let $\phi, \psi \in \mathrm{Rat}_d^1$ and $\sigma \in \mathrm{PGL}_2(k)$. Then $\psi = \phi^\sigma$ if and only if*

$$\mathcal{I}_\psi = \mathcal{I}_\phi^\sigma \quad \text{and} \quad \mathcal{J}_\psi = \mathcal{J}_\phi^\sigma.$$

Proof. Let $\phi = \frac{f_0}{f_1}$ and $\psi = \frac{g_0}{g_1}$. Assume that ϕ and ψ are k-equivalent in Rat_d^1, meaning there exists $\sigma \in \mathrm{PGL}_2(k)$ such that $\psi = \phi^\sigma$.

Substituting the expressions for $\mathcal{I}_\psi = yg_0 - xg_1$ and using the values of g_0 and g_1 as in Eq. (14), we have

$$
\begin{aligned}
\mathcal{I}_\psi &= yg_0 - xg_1 = y(ef_0^\sigma - bf_1^\sigma) - x(-cf_0^\sigma + af_1^\sigma) \\
&= (cx+ey)f_0^\sigma - (ax+by)f_1^\sigma = (yf_0 - xf_1)^\sigma = \mathcal{I}_\phi^\sigma.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\mathcal{J}_\psi &= \frac{\partial g_0}{\partial x} + \frac{\partial g_1}{\partial y} = \frac{\partial(ef_0^\sigma - bf_1^\sigma)}{\partial x} + \frac{\partial(-cf_0^\sigma + af_1^\sigma)}{\partial y} \\
&= e\frac{\partial f_0^\sigma}{\partial x} - b\frac{\partial f_1^\sigma}{\partial x} - c\frac{\partial f_0^\sigma}{\partial y} + a\frac{\partial f_1^\sigma}{\partial y} = \left(\frac{\partial f_0^\sigma}{\partial x} + \frac{\partial f_1^\sigma}{\partial y}\right)^\sigma = \mathcal{J}_\phi^\sigma.
\end{aligned}
$$

Thus, we conclude that $\mathcal{I}_\phi$ and $\mathcal{I}_\psi$ (and similarly, $\mathcal{J}_\phi$ and $\mathcal{J}_\psi$) are k-equivalent via σ as binary forms.

Conversely, suppose that $\mathcal{I}_\phi$ and $\mathcal{I}_\psi$ (respectively, $\mathcal{J}_\phi$ and $\mathcal{J}_\psi$) are k-equivalent via σ as binary forms. This means that $\mathcal{I}_\psi = \mathcal{I}_\phi^\sigma$ and $\mathcal{J}_\psi = \mathcal{J}_\phi^\sigma$. In particular, we have

$$
\begin{aligned}
\mathcal{I}_\psi &= yg_0 - xg_1 = (yf_0 - xf_1)^\sigma = (cx+ey)f_0^\sigma - (ax+by)f_1^\sigma, \\
\mathcal{J}_\psi &= \frac{\partial g_0}{\partial x} + \frac{\partial g_1}{\partial y} = \frac{\partial f_0^\sigma}{\partial x^\sigma} + \frac{\partial f_1^\sigma}{\partial y^\sigma}.
\end{aligned}
$$

From the equation for $\mathcal{I}_\psi$, we obtain

$$
y\left(g_0 - ef_0^\sigma + bf_1^\sigma\right) = x\left(g_1 + cf_0^\sigma - af_1^\sigma\right),
$$

which leads to

$$
\begin{aligned}
g_0 - ef_0^\sigma + bf_1^\sigma &= x \cdot h(x,y), \\
g_1 + cf_0^\sigma - af_1^\sigma &= y \cdot h(x,y),
\end{aligned}
$$

for some $h \in V_{d-1}$. Therefore, we have

$$
\begin{aligned}
\mathcal{J}_\psi &= 2h + \left(x\frac{\partial h}{\partial x} + y\frac{\partial h}{\partial y}\right) + e\frac{\partial f_0^\sigma}{\partial x} + b\frac{\partial f_1^\sigma}{\partial x} - c\frac{\partial f_0^\sigma}{\partial y} + a\frac{\partial f_1^\sigma}{\partial y} \\
&= h + \frac{d-1}{2}h + \frac{\partial g_0}{\partial x} + \frac{\partial g_1}{\partial y}.
\end{aligned}
$$

Thus, we must have $h = 0$, which implies that $\phi^\sigma = \psi$ as claimed. □

Since the pair of binary forms $(\mathcal{I}_\phi, \mathcal{J}_\phi)$ determines the rational function ϕ, we can use the classical theory of binary forms to determine invariants for ϕ. Define

$$
\begin{aligned}
\Phi : \mathrm{Rat}_d^1 &\to V_{d+1} \times V_{d-1}, \\
\phi &\to (\mathcal{I}_\phi, \mathcal{J}_\phi).
\end{aligned}
$$

The inverse of Φ is not well-defined since not every pair $(f,g) \in V_{d+1} \times V_{d-1}$ determines a rational function.

For any $(f,g) \in V_{d+1} \times V_{d-1}$, define the resultant

$$
\Delta_{f,g} = \mathrm{Res}\left(xg + \frac{\partial f}{\partial y},\, yg - \frac{\partial f}{\partial x}\right),
$$

and the *resultant locus* $\mathcal{N}$ as

$$
\mathcal{N} := \{(f,g) \in V_{d+1} \times V_{d-1} \mid \Delta_{f,g} = 0\}.
$$

Then, we have the following result:

Lemma 2. *The map* $\Phi : \mathrm{Rat}_d^1 \to (V_{d+1} \times V_{d-1}) \setminus \mathcal{N}$ *is bijective. Moreover, for any* $(f, g) \in V_{d+1} \times V_{d-1} \setminus \mathcal{N}$,

$$\Phi^{-1}(f, g) = \frac{xg + \frac{\partial f}{\partial y}}{yg - \frac{\partial f}{\partial x}}.$$

Proof. The map Φ is obviously well-defined. Let $\phi, \psi \in \mathrm{Rat}_d^1$ such that $\Phi(\phi) = \Phi(\psi)$. Then, $\mathcal{I}_\phi = \mathcal{I}_\psi$ and $\mathcal{J}_\phi = \mathcal{J}_\psi$ as binary forms in V_{d+1} and V_{d-1}, respectively. Since binary forms are defined up to multiplication by a scalar, there exists a diagonal matrix $\sigma \in \mathrm{PGL}_2(k)$ such that

$$\mathcal{I}_\psi = \mathcal{I}_\phi^\sigma \quad \text{and} \quad \mathcal{J}_\psi = \mathcal{J}_\phi^\sigma.$$

By Thm. 1, we conclude that $\psi = \phi^\sigma = \phi$. Thus, Φ is injective.

Now, assume that $(f, g) \in (V_{d+1} \times V_{d-1}) \setminus \mathcal{N}$. The condition $\Delta_{f,g} \neq 0$ ensures that the preimage

$$\Phi^{-1}(f, g) = \frac{xg - \frac{\partial f}{\partial y}}{yg + \frac{\partial f}{\partial x}}$$

belongs to Rat_d^1. Consequently, the map Φ^{-1} is well-defined.

To show that Φ and Φ^{-1} are inverses of each other, we will demonstrate that $\Phi \circ \Phi^{-1} = \mathrm{id}$ on $(V_{d+1} \times V_{d-1}) \setminus \mathcal{N}$ and $\Phi^{-1} \circ \Phi = \mathrm{id}$ on Rat_d^1.

First, compute $\Phi^{-1} \circ \Phi^{-1}$ on $(V_{d+1} \times V_{d-1}) \setminus \mathcal{N}$: Let $(f, g) \in (V_{d+1} \times V_{d-1}) \setminus \mathcal{N}$. Let $\phi = \Phi^{-1}(f, g) = \frac{xg + \frac{\partial f}{\partial y}}{yg - \frac{\partial f}{\partial x}}$. Then,

$$\begin{aligned}
\mathcal{I}_\phi &= y\left(xg + \frac{\partial f}{\partial y}\right) - x\left(yg - \frac{\partial f}{\partial x}\right) = y\frac{\partial f}{\partial y} + x\frac{\partial f}{\partial x} = (d+1)f = f \text{ in } V_{d+1}, \\
\mathcal{J}_\phi &= \frac{\partial}{\partial x}\left[xg + \frac{\partial f}{\partial y}\right] + \frac{\partial}{\partial y}\left[yg - \frac{\partial f}{\partial x}\right] = 2g + x\frac{\partial g}{\partial x} + y\frac{\partial g}{\partial y} \\
&= 2g + (d-1)g = g \text{ in } V_{d-1}.
\end{aligned}$$

Thus, $\mathcal{I}_\phi = f$ and $\mathcal{J}_\phi = g$, so $\Phi(\phi) = (\mathcal{I}_\phi, \mathcal{J}_\phi) = (f, g)$.

Second, compute $\Phi^{-1} \circ \Phi$ on Rat_d^1: Let $\phi = \frac{f_0}{f_1} \in \mathrm{Rat}_d^1$. Then,

$$\begin{aligned}
\left(\Phi^{-1} \circ \Phi\right)(\phi) &= \Phi^{-1}(\mathcal{I}_\phi, \mathcal{J}_\phi) = \Phi^{-1}\left(yf_0 - xf_1, \frac{\partial f_0}{\partial x} + \frac{\partial f_1}{\partial y}\right) \\
&= \frac{x\left(\frac{\partial f_0}{\partial x} + \frac{\partial f_1}{\partial y}\right) + \frac{\partial}{\partial y}\left[yf_0 - xf_1\right]}{y\left(\frac{\partial f_0}{\partial x} + \frac{\partial f_1}{\partial y}\right) - \frac{\partial}{\partial x}\left[yf_0 - xf_1\right]} \\
&= \frac{x\left(\frac{\partial f_0}{\partial x} + \frac{\partial f_1}{\partial y}\right) + \left(f_0 + y\frac{\partial f_0}{\partial y} - x\frac{\partial f_1}{\partial y}\right)}{y\left(\frac{\partial f_0}{\partial x} + \frac{\partial f_1}{\partial y}\right) - \left(y\frac{\partial f_0}{\partial x} - f_1 - x\frac{\partial f_1}{\partial x}\right)} \\
&= \frac{x\frac{\partial f_0}{\partial x} + y\frac{\partial f_0}{\partial y} + f_0}{x\frac{\partial f_1}{\partial x} + y\frac{\partial f_1}{\partial y} + f_1} = \frac{df_0 + f_0}{df_1 + f_1} = \frac{f_0}{f_1} = \phi.
\end{aligned}$$

Since both compositions $\Phi \circ \Phi^{-1}$ and $\Phi^{-1} \circ \Phi$ are the identity maps, we conclude that Φ and Φ^{-1} are indeed inverse to each other. This completes the proof. $\square$

3.1. **Ring of invariants $\mathcal{R}_{(d+1),(d-1)}$and a theorem of Clebsch.** The action of $\mathrm{GL}_2(k)$ on V_d induces an action of $\mathrm{GL}_2(k)$ in $V_{d+1} \oplus V_{d-1}$. To determine the isomorphism classes of degree d rational functions we have to determine the ring of invariants of $V_{d+1} \oplus V_{d-1}$. This is well known in classical invariant theory; see [8] for a complete treatment. We briefly describe it below.

Let V be an SL_2-module and $\mathcal{O}(V)$ the algebra of polynomial functions on V. $\mathrm{SL}_2(k)$ acts on $\mathcal{O}(V)$ via

$$M \cdot p(f_1, \dots, f_r) \to p(M^{-1} f_1, \dots, M^{-1} f_r),$$

for every $M \in \mathrm{SL}_2(k)$. An invariant of V is an element $\mathcal{T} \in \mathcal{O}(V)$ such that $M\mathcal{T} = \mathcal{T}$, for all $M \in \mathrm{SL}_2(k)$. The set of invariants is denoted by $\mathcal{O}(V)^{\mathrm{SL}_2}$.

A transvectant $(\mathcal{T}, \mathcal{S})_l$ is called **irrelevant** if there exist $\mathcal{T}_1$, $\mathcal{T}_2$, $\mathcal{S}_1$, $\mathcal{S}_2$ and l_1, l_2 such that

$$l = l_1 + l_2, \quad \mathcal{T} = \mathcal{T}_1 \cdot \mathcal{T}_2, \quad \mathcal{S} = \mathcal{S}_1 \cdot \mathcal{S}_2,$$

and $l_1 \leq \operatorname{ord} \mathcal{T}_1, \operatorname{ord} \mathcal{S}_1$, and $l_2 \leq \operatorname{ord} \mathcal{T}_2, \operatorname{ord} \mathcal{S}_2$. A transvectant which is not irrelevant is called **relevant**.

Let V and W be two SL_2-modules whose covariants are finitely generated, and assume

$$\begin{aligned} &\mathcal{T}_1, \dots, \mathcal{T}_r : && \text{are the generators of the covariants of } V \\ &\mathcal{S}_1, \dots, \mathcal{S}_s : && \text{are the generators of the covariants of } W. \end{aligned} \tag{16}$$

Theorem (Clebsch)**.** *The ring of covariants of $V \oplus W$ is also finitely generated. Moreover, a finite generating system can be chosen from the set of all transvectants*

$$(\mathcal{T}, \mathcal{S})_l\,, \quad \textit{for} \quad l \geq 0,$$

where $\mathcal{T}$ is a monomial in the $\mathcal{T}_i$'s and $\mathcal{S}$ a monomial in the $\mathcal{S}_j$'s. In other words, by the relevant transvectants $(\mathcal{T}, \mathcal{S})_l$.

3.2. $\operatorname{Proj} \mathcal{R}_{(d+1,d-1)}$**as a weighted projective space.** Let $\xi_0, \dots, \xi_n$ be a generating system of $\mathcal{R}_{(d+1,d-1)}$. Since all $\xi_0, \dots \xi_n$ are homogenous polynomials then $\mathcal{R}_{(d+1,d-1)}$ is a graded ring and $\operatorname{Proj} \mathcal{R}_{(d+1,d-1)}$ is a weighted projective space.

Let $\omega := (q_0, \dots, q_n) \in \mathbb{Z}^{n+1}$ be a fixed tuple of positive integers called **weights**. Consider the action of $k^\star = k \setminus \{0\}$ on $\mathbb{A}^{n+1}(k)$ as follows

$$\lambda \star (x_0, \dots, x_n) = (\lambda^{q_0} x_0, \dots, \lambda^{q_n} x_n)$$

for $\lambda \in k^*$. The quotient of this action is called a **weighted projective space** and denoted by $\mathbb{WP}^n_\omega(k)$. It is the projective variety $Proj\,(k[x_0, ..., x_n])$ associated to the graded ring $k[x_0, \dots, x_n]$ where the variable x_i has degree q_i for $i = 0, \dots, n$. We will denote a point $\mathfrak{p} \in \mathbb{WP}^n_w(k)$ by $\mathfrak{p} = [x_0 : x_1 : \cdots : x_n]$.

Let $\phi(x, y) \in \mathrm{Rat}^1_d$ given by

$$\phi(x, y) = \frac{f_0(x, y)}{f_1(x, y)},$$

its associated binary forms $\mathcal{I}_\phi \in V_{d+1}$ and $\mathcal{J}_\phi \in V_{d-1}$, and $\xi_0, \xi_1, \dots, \xi_n$ the generators of the ring of invariants $\mathcal{R}_{(d+1),(d-1)}$. The **invariants** of the rational function ϕ are defined as

$$\xi(\phi) := [\xi_0(\mathcal{I}_\phi, \mathcal{J}_\phi), \xi_1(\mathcal{I}_\phi, \mathcal{J}_\phi), \dots, \xi_n(\mathcal{I}_\phi, \mathcal{J}_\phi)] \in \mathbb{WP}^n_\omega(k). \tag{17}$$

Moreover, $\phi = \psi^\sigma$ for $\sigma \in \mathrm{GL}_2(K)$ if and only if $\xi(\phi) = \lambda \star \xi(\psi)$, for $\lambda = (\det \sigma)^{\frac{d}{2}}$.

Next we will determine explicitly invariants of $\mathcal{R}_{d+1,d-1}$. From now on $f \in V_{d+1}$ and $g \in V_{d-1}$ where

$$f = \sum_{i=0}^{d+1} a_i x^i y^{d+1-i}, \quad \text{and} \quad g = \sum_{i=0}^{d-1} b_i x^i y^{d-1-i}, \tag{18}$$

3.3. **Invariants of $V_4 \oplus V_2$.** We take $d = 3$, $f \in V_4$ and $g \in V_2$ as in Eq. (18)

$$\begin{aligned} f(x,y) &= a - 3x^3 + a - 2x^2 y + a - 1xy^2 + a_0 y^3 \\ g(x,y) &= b_2 x^2 + b_1 xy + b_0 y^2 \end{aligned} \tag{19}$$

The generators of covariants of V_4 and V_2 are

$$\begin{aligned} &T = \{f,\, \mathcal{T} = (f,f)_2,\, \mathcal{T}_2 = (f,f)_4,\, \mathcal{T}_3 = ((f,f)_2, f)_4\} \\ &S = \{g,\, S_2 = (G,G)_2\} \end{aligned}$$

respectively. Hence, we are considering all transvectants

$$\left(f^{m_1} \mathcal{T}^{m_2} \mathcal{T}_2^{m_3} \mathcal{T}_3^{m_4}, g^{s_1} \mathcal{S}_2^{s_2}\right)_l,$$

for some $m_1, m_2, m_3, m_4, s_1, s_2$.

Since $\mathcal{T}_2, \mathcal{T}_3$ and $\mathcal{S}_2$ are invariants, their exponents most be zero, otherwise we get reducible invariants. Hence, $\mathcal{T}_2$, $\mathcal{T}_3$, $\mathcal{S}_2$ are part of the generating set and further we only consider $\left(f^{m_1} \mathcal{T}^{m_2}, g^s\right)_l$. Then the relevant transvectants are $\mathcal{S}_2$, $\mathcal{T}_2$, $\mathcal{T}_3$, and

$$R_3 := (\mathcal{T}, g^2)_4, \ R_4 := (f, g^2)_4, \quad R_6 := (g^3, (f, \mathcal{T})_1)_6,$$

Hence, the set of invariants is $\xi(\phi) = (\xi_0, \ldots, \xi_5)$, where

$$\xi_0 = (g,g)_2, \ \xi_1 = (f,f)_4, \ \xi_2 = (\mathcal{T}, f)_4, \ \xi_3 = R_3, \ \xi_4 = R_4, \ \xi_5 = R_6$$

with weights $(2,2,3,3,4,6)$ respectively. Let

$$\begin{aligned} f &= a_4 x^4 + a_3 x^3 y + a_2 x^2 y^2 + a_1 x y^3 + a_0 y^4 \\ g &= b_2 x^2 + b_1 xy + b_0 y^2 \end{aligned}$$

We have the following expressions for invariants:

$$\begin{aligned} \xi_0 &= \frac{1}{2}\left(4b_0 b_2 - b_1^2\right) \\ \xi_1 &= \frac{1}{6}\left(a_2^2 - 3a_1 a_3 + 12 a_0 a_4\right) \\ \xi_2 &= \frac{1}{72}\left(-2a_2^3 + 9(a_1 a_3 + 8 a_0 a_4) a_2 - 27\left(a_4 a_1^2 + a_0 a_3^2\right)\right) \\ \xi_3 &= \frac{1}{6}\left(6 a_4 b_0^2 - 3 a_3 b_1 b_0 + 2 a_2 b_2 b_0 + a_2 b_1^2 + 6 a_0 b_2^2 - 3 a_1 b_1 b_2\right) \end{aligned}$$

$$\xi_4 = -\frac{1}{72}\left(2a_2^2b_1^2 + 4a_2^2b_0b_2 - 24a_4a_2b_0^2 - 24a_0a_2b_2^2 - 6a_3a_2b_0b_1\right.$$
$$-6a_1a_2b_1b_2 + 9a_3^2b_0^2 - 3a_1a_3b_1^2 - 24a_0a_4b_1^2 + 9a_1^2b_2^2 + 36a_1a_4b_0b_1 - 6a_1a_3b_0b_2$$
$$\left.-48a_0a_4b_0b_2 + 36a_0a_3b_1b_2\right)$$
$$\xi_5 = -\frac{1}{32}\left(a_3^3b_0^3 + 8a_1a_4^2b_0^3 - 4a_2a_3a_4b_0^3 - a_2a_3^2b_1b_0^2\right.$$
$$-2a_1a_3a_4b_1b_0^2 + a_1a_3^2b_2b_0^2 - 4a_1a_2a_4b_2b_0^2 + 8a_0a_3a_4b_2b_0^2 + a_1a_3^2b_1^2b_0$$
$$-4a_1a_2a_4b_1^2b_0 + 8a_0a_3a_4b_1^2b_0 - a_1^2a_3b_2^2b_0 + 4a_0a_2a_3b_2^2b_0 - 8a_0a_1a_4b_2^2b_0$$
$$-6a_0a_3^2b_1b_2b_0 + 6a_1^2a_4b_1b_2b_0 - a_0a_3^2b_1^3 + a_1^2a_4b_1^3 - a_1^3b_2^3 + 4a_0a_1a_2b_2^3 - 8a_0^2a_3b_2^3$$
$$-4a_0a_2^2b_1b_2^2 + a_1^2a_2b_1b_2^2 + 2a_0a_1a_3b_1b_2^2 + 16a_0^2a_4b_1b_2^2 - a_1^2a_3b_1^2b_2 + 4a_0a_2a_3b_1^2b_2$$
$$\left.-8a_0a_1a_4b_1^2b_2 - 16a_0a_4^2b_1b_0^2 + 4a_2^2a_4b_1b_0^2\right)$$

The ring of invariants $\mathcal{R}_{4,2}$ is generated by $\xi_0, \xi_1, \xi_2, \xi_3, \xi_4, \xi_5$ and a relation between invariants $\xi_0, \dots, \xi_5$ is computed in [8] which satisfy the suzuki

$$\xi_5^2 = \frac{1}{108}\xi_0^3\xi_1^3 - 18\xi_0^3\xi_2^2 - \frac{1}{24}\xi_0\xi_1^2\xi_3^2 - \frac{1}{6}\xi_2\xi_3^3 + \frac{1}{2}\xi_0\xi_2\xi_3\xi_4 + \frac{1}{4}\xi_1\xi_3^2\xi_4 - \frac{1}{4}\xi_0\xi_1\xi_4^2 - \frac{1}{2}\xi_4^3 \tag{20}$$

In other words, the moduli space $\mathcal{F}_3// \mathrm{SL}_2(k)$ is identified inside the weighted projective space $\mathbb{P}^5_{(2,2,3,3,4,6)}$ with the 4-dimensional variety above. Next we will verify the correctness of this equation and introduce absolute invariants.

3.4. **Invariants of rational cubics.** Finally, given a rational cubic $\phi(x)$ we want to compute explicitly its invariants in terms of its coefficients. We assume that

$$\phi(x) = \frac{f_0(x)}{f_1(x)} = \frac{c_0x^3 + c_1x^2 + c_2x + c_3}{c_4x^3 + c_5x^2 + c_6x + c_7}$$

This choice of notation makes it easier for us that the rational function ϕ is represented in our database as a list of coefficients

$$\mathfrak{p} = [c_0, c_1, \dots, c_7]$$

which as previously explained is considered a point in the projective space $\mathbb{P}^7$. We are assuming that the degree $\deg \phi = 3$, hence $c_0 \neq 0$ or $c_4 \neq 0$. Furthermore we are assuming that it can not be further reduced hence their resultant $I_6 = \mathrm{Res}\,(f_0, f_1)$ is nonzero. Recall that this was the homogenious degree $2d$ polynomial defined in Eq. (10). In the case of $d = 3$ we have:

(21)
$$I_6 = c_3^3c_4^3 - c_0^3c_7^3 + c_3c_0^2c_6^3 - c_2^3c_7c_4^2 + c_1^3c_7^2c_4 - c_3^2c_2c_5c_4^2 - c_3^2c_0c_5^3 - 2c_3^2c_1c_6c_4^2 + c_3c_1^2c_6^2c_4$$
$$+ c_3c_2^2c_6c_4^2 + c_3^2c_1c_5^2c_4 - 3c_3^2c_0c_7c_4^2 - c_1^2c_0c_7^2c_5 + c_1c_0^2c_7^2c_6 + 2c_2c_0^2c_7^2c_5 - c_2c_0^2c_7c_6^2$$
$$+ 3c_3c_0^2c_7^2c_4 - c_2^2c_0c_7c_5^2 + 3c_3^2c_0c_6c_5c_4 - 2c_3c_2c_0c_6^2c_4 + c_3c_2c_0c_6c_5^2 - 2c_3c_1^2c_7c_5c_4$$
$$+ 3c_3c_2c_1c_7c_4^2 + 2c_3c_1c_0c_7c_5^2 - c_3c_1c_0c_6^2c_5 - c_2c_1^2c_7c_6c_4 - 3c_2c_1c_0c_7^2c_4 - 3c_3c_0^2c_7c_6c_5$$
$$+ c_2^2c_1c_7c_5c_4 + 2c_2^2c_0c_7c_6c_4 - c_3c_2c_1c_6c_5c_4 - c_3c_2c_0c_7c_5c_4 + c_3c_1c_0c_7c_6c_4 + c_2c_1c_0c_7c_6c_5 \neq 0$$

The pair of binary forms $\mathcal{I}_\phi$ and $\mathcal{J}_\phi$ associated to ϕ are

$$\begin{aligned} \mathcal{I}_\phi &= c_3x^3y + c_2x^2y^2 + c_1xy^3 + c_0y^4 - c_7x^4 - c_6x^3y - c_5x^2y^2 - c_4xy^3 \\ \mathcal{J}_\phi &= 3c_3x^2 + 2c_2xy + c_1y^2 + c_6x^2 + 2c_5xy + 3c_4y^2 \end{aligned} \tag{22}$$

Next we evaluate the following transvectants

$$
\begin{aligned}
&\xi_0 = (\mathcal{J}_\phi, \mathcal{J}_\phi)_2\,, \quad \xi_1 = (\mathcal{I}_\phi, \mathcal{I}_\phi)_4\,, \qquad\qquad \xi_2 = ((\mathcal{I}_\phi, \mathcal{I}_\phi)_2, \mathcal{I}_\phi)_4\,,\\
&\xi_3 = \left(\mathcal{I}_\phi, \mathcal{J}_\phi^2\right)_4, \quad \xi_4 = \left((\mathcal{I}_\phi, \mathcal{I}_\phi)_2, \mathcal{J}_\phi^2\right)_4, \quad \xi_5 = \left(\mathcal{J}_\phi^3, \left(\mathcal{I}_\phi, (\mathcal{I}_\phi, \mathcal{I}_\phi)_2\right)_1\right)_6
\end{aligned} \tag{23}
$$

and display their expressions in Appendix A. Their resultant $\Delta_{\mathcal{I}_\phi, \mathcal{J}_\phi} = \operatorname{Res}(\mathcal{I}_\phi, J_\phi)$ is homogenous polynomial of degree six in terms of $c_0, \dots, c_7$ and is also displayed in Appendix A. Hence, we will denote it by $J_6 := \operatorname{Res}(\mathcal{I}_\phi, J_\phi)$.

Thus, there are three invariants of degree 6 in the case of cubics, namely I_6, J_6, and ξ_5. We would like to express I_6 and J_6 in terms of $\xi_0, \dots, \xi_5$. The expression of I_6 was computed in [8, pg. 38]. It seems as there are a couple of typos in the printed version of [8, pg. 38] and we could not verify it. It can be easily noticeable that it is incorrect since it is not a homogenous polynomial of degree 6; see for example the monomials ξ_4^2, ξ_1, ξ_3, $\xi_0\xi_3$ which are not of degree 6. The relation between invariants $\xi_0, \dots, \xi_5$ was also computed in [8]

$$
\xi_5^2 = \frac{1}{108}\xi_0^3\xi_1^3 - 18\xi_0^3\xi_2^2 - \frac{1}{24}\xi_0\xi_1^2\xi_3^2 - \frac{1}{6}\xi_2\xi_3^3 + \frac{1}{2}\xi_0\xi_2\xi_3\xi_4 + \frac{1}{4}\xi_1\xi_3^2\xi_4 - \frac{1}{4}\xi_0\xi_1\xi_4^2 - \frac{1}{2}\xi_4^3 \tag{24}
$$

and again we could not verify it. In our approach below we independently compute such relations via a computational algebra approach.

Lemma 3. *The ring of invariants $\mathcal{R}_{4,2}$ is generated by $\xi_0, \xi_1, \xi_2, \xi_3, \xi_4, I_6$.*

The weighted projective space $\mathbb{P}_{(2,2,3,3,4,6)}$ can be embedded into the projective space $\mathbb{P}^5$ via Veronese embedding as

$$[\xi_0, \xi_1, \xi_2, \xi_3, \xi_4, I_6] \to [\xi_0^6, \xi_1^6, \xi_2^4, \xi_3^4, \xi_4^3, I_6^2]$$

Since $I_6 \neq 0$ we can divide by I_6^2 and represent each point as

$$\left[\frac{\xi_0^6}{I_6^2} : \frac{\xi_1^6}{I_6^2} : \frac{\xi_2^4}{I_6^2} : \frac{\xi_3^4}{I_6^2} : \frac{\xi_4^3}{I_6^2} : 1\right] \in \mathbb{P}^5$$

This motivates the definition of the following invariants

$$i_1 = \frac{\xi_0^6}{I_6^2},\ i_2 = \frac{\xi_1^6}{I_6^2},\ i_3 = \frac{\xi_2^4}{I_6^2},\ i_4 = \frac{\xi_3^4}{I_6^2},\ i_5 = \frac{\xi_4^3}{I_6^2},$$

which are $\mathrm{GL}_2(k)$-invariants and are defined everywhere in the moduli space Rat_3^1. We call such invariants $i_1, \dots, i_5$ **absolute invariants** of the rational function $\phi(x, y)$. Hence, we have:

Lemma 4. *Two degree three rational functions are equivalent if and only if they have the same absolute invariants.*

Assume $\phi : \mathbb{P}^1 \to \mathbb{P}^1$ is a degree 3 rational function. From the Riemann - Hurwitz formula it has four branch points, which we can assume are 0, 1, ∞ and t. In the fiber of each of these branch points must be one ramified points. We fix a coordinate on $\mathbb{P}^1$ such that 0,1, ∞ are the ramified points in the fibers of 0,1, ∞ respectively. Hence, our function is

$$\phi(z) = \lambda \frac{x^2(x-a)}{x-b}$$

Then we have

$$\phi(z) - 1 = \lambda \frac{(x-1)^2(x-c)}{x-b} \qquad \text{and} \qquad \phi(z) - t = \frac{(x-d)^2(x-r)}{x-b}$$

By clearing the denumerators and equaling the coefficients of polynomials we get

...

$$
\begin{aligned}
I_6 &= -\frac{1}{8}\xi_1^3 - \frac{1}{384}\xi_0^2\xi_1 + \frac{3}{4}\xi_2^2 - \frac{3}{16}\xi_1\xi_4 - \frac{1}{256}\xi_3^2 + \frac{3}{16}\xi_2\xi_3 + \frac{1}{64}\xi_0\xi_4 - \frac{1}{8}\xi_5 \\
J_6 &= \xi_3^2 - 4\xi_4\xi_0 + \frac{2}{3}\xi_0^2\xi_1
\end{aligned} \tag{25}
$$

Next we have the following:

Theorem 1. *The ring of invariants $\mathcal{R}_{4,2}$ is generated by $\xi_0, \xi_1, \xi_2, \xi_3, \xi_4, I_6$ which satisfy the locus $\mathcal{S}$ below*

$$
\begin{aligned}
&1769472I_6^2 + 9216I_6\xi_0^2\xi_1 - 55296I_6\xi_0\xi_4 + 442368I_6\xi_1^3 + 663552I_6\xi_1\xi_4 - 2654208I_6\xi_2^2 \\
&- 663552I_6\xi_2\xi_3 + 13824I_6\xi_3^2 + 12\xi_0^4\xi_1^2 - 256\xi_0^3\xi_1^3 - 144\xi_0^3\xi_1\xi_4 + 1536\xi_0^3\xi_2^2 + 1152\xi_0^2\xi_1^4 \\
&+ 1728\xi_0^2\xi_1^2\xi_4 - 6912\xi_0^2\xi_1\xi_2^2 - 1728\xi_0^2\xi_1\xi_2\xi_3 + 36\xi_0^2\xi_1\xi_3^2 + 432\xi_0^2\xi_4^2 - 6912\xi_0\xi_1^3\xi_4 + 13824\xi_4^3 \\
&+ 1152\xi_0\xi_1^2\xi_3^2 - 3456\xi_0\xi_1\xi_4^2 + 41472\xi_0\xi_2^2\xi_4 - 3456\xi_0\xi_2\xi_3\xi_4 - 216\xi_0\xi_3^2\xi_4 + 27648\xi_1^6 \\
&+ 82944\xi_1^4\xi_4 - 331776\xi_1^3\xi_2^2 - 82944\xi_1^3\xi_2\xi_3 + 1728\xi_1^3\xi_3^2 - 497664\xi_1\xi_2^2\xi_4 - 124416\xi_1\xi_2\xi_3\xi_4 \\
&- 1440\xi_1\xi_3^2\xi_4 + 62208\xi_1\xi_4^2 + 995328\xi_2^4 + 497664\xi_2^3\xi_3 + 51840\xi_2^2\xi_3^2 + 2016\xi_2\xi_3^3 + 27\xi_3^4 = 0
\end{aligned}
$$

Moreover, the moduli space Rat_3^1 of degree three rational function is identified with the 4-dimensional variety $\mathcal{S} \setminus \{J_6 \neq 0\}$ inside the weighted projective space $\mathbb{P}^5_{(2,2,3,3,4,6)}$.

Proof. ☐

4. Automorphisms

In this section we will define and study the automorphism groups of rational functions. The approach will be similar to studying the automorphism groups of hyperelliptic curves; see [9, 10]. Such groups have long been the focus of many authors in arithmetic dynamics; see [5,8,11,12]. First we recall some preliminaries.

Let G be a finite subgroup of $\mathrm{PGL}_2(k)$. Therefore G is isomorphic to one of the following C_n, D_n, A_4, S_4, or A_5 with branching indices of the corresponding cover $\mathbb{P}^1_x \to \mathbb{P}^1/G$ given respectively by

$$(n,n),\ (2,2,n),\ (2,3,3),\ (2,4,4),\ (2,3,5).$$

We fix a coordinate z in $\mathbb{P}^1/G$. Thus, G is the monodromy group of a cover $\psi : \mathbb{P}^1_x \to \mathbb{P}^1_z$. We denote by $q_1, \dots, q_r$ the corresponding branch points of ψ. For each $q_1, \dots, q_r$ we have a corresponding permutation $\sigma_1, \dots, \sigma_r \in S_n$. The tuple $\bar{\sigma} := (\sigma_1, \dots, \sigma_r)$ is the signature of G. Thus,

$$G = \langle \sigma_1, \dots, \sigma_r \rangle, \quad \text{and} \quad \sigma_1 \cdots \sigma_r = 1.$$

Since each of the above groups is embedded in $\mathrm{PGL}_2(\mathbb{C})$ then we can have these generating systems $\sigma_1, \dots, \sigma_r$ as matrices in $\mathrm{PGL}_2(\mathbb{C})$. Below we display all the cases:

$$\begin{aligned}
&i) && C_n \cong \left\langle \begin{bmatrix} \zeta_n & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \zeta_n^{n-1} & 0 \\ 0 & 1 \end{bmatrix} \right\rangle \\
&ii) && D_n \cong \left\langle \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} \zeta_n & 0 \\ 0 & 1 \end{bmatrix} \right\rangle \\
&iii) && A_4 \cong \left\langle \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \right\rangle \\
&iv) && S_4 \cong \left\langle \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \right\rangle \\
&v) && A_5 \cong \left\langle \begin{bmatrix} \omega & 1 \\ 1 & -\omega \end{bmatrix}, \begin{bmatrix} \omega & \xi_5^4 \\ 1 & -\xi_5^4\omega \end{bmatrix} \right\rangle
\end{aligned} \tag{26}$$

where $\omega = \frac{-1+\sqrt{5}}{2}$, ζ_n is a primitive n^{th} root of unity, ξ_5 is a primitive 5^{th} root of unity, and i is a primitive 4^{th} root of unity.

Remark 2. *In each case above, there is $\sigma \in G$ which fixes 0 and ∞. The proof is elementary. In the first two cases the Mobius transformation $\sigma(x) = \xi_m x$ will fix 0 and ∞. In the next two cases $\sigma(x) = -x$ will do that, and in the last case $\sigma(x) = wx$ is in the group and will fix 0 and ∞.*

4.1. **Fixed fields of the reduced automorphism groups.**

The group G given above acts on $k(x)$ via the natural way. The fixed field is a genus 0 field, say $k(z)$. Thus, z is a degree $|G|$ rational function in x, say $z = \psi(x)$.

Let H be a finite subgroup of $PGL_2(k)$. Let us identify each element of H with the corresponding Moebius transformation and let s_i be the i-th elementary symmetric polynomial in the elements of H, $i = 1, \dots, |H|$. Then any non-constant s_i generates $k(z)$. The fixed field for each of the groups is generated respectively by

$$
\begin{aligned}
\psi(x) &= x^n, \\
\psi(x) &= x^n + \frac{1}{x^n}, \\
\psi(x) &= \frac{x^{12} - 33x^8 - 33x^4 + 1}{x^2(x^4-1)^2}, \\
\psi(x) &= \frac{(x^8 + 14x^4 + 1)^3}{108\left(x(x^4-1)\right)^4}, \\
\psi(x) &= \frac{\left(-x^{20} + 228x^{15} - 494x^{10} - 228x^5 - 1\right)^3}{1728\left(x(x^{10} + 11x^5 - 1)\right)^5}.
\end{aligned}
$$

The branch points of each of the above functions are i) $\{0, \infty\}$, ii) $\{-2, 2, \infty\}$, iii) $\{\infty, -6i\sqrt{3}, 6i\sqrt{3}\}$, iv) $\{0, 1, \infty\}$, v) $\{0, 1728, \infty\}$. The above facts were known to Klein. For $\sigma \in \mathrm{PGL}_2(k)$ we denote by $\mathbf{Fix}(\sigma)$ the **set of fixed points** of σ. Notice that any σ different from identity can fix at most two points.

4.2. **Automorphisms of rational maps.** Let $\phi \in \mathrm{Rat}_d^1$ and as above $\phi \in \mathrm{Rat}_d^1$. A point $[\alpha_3 : \alpha_2] \in \mathbb{P}^1$ is called a **fixed point** for ϕ if $\phi(\alpha_3, \alpha_2) = [\alpha_3, \alpha_2]$. Let $t = x_1/x_0$. Hence, ϕ can be taken as a rational function in t, say

$$\phi(t) = \frac{F(t)}{G(t)}.$$

Then t is a fixed point if $\phi(t) = t$, which implies that

$$S(t) := F(t) - t\,G(t) = 0$$

which is at most a degree $(d+1)$ equation in t. Hence, a degree d rational function has at most $(d+1)$ fixed points.

We denote the set of fixed point of ϕ by $\mathbf{Fix}(\phi)$. Notice that if

$$\mathbf{Fix}(\phi) = \{w_1, \dots, w_{d+1}\}$$

is known, then we can uniquely determine the rational function ϕ by solving the linear system $F(w_i) - w_i\,G(w_i) = 0$, for $i = 1, \dots, d+1$. A function ϕ has less than $d+1$ exactly when the discriminant $\Delta(S, t)$ vanishes.

An **automorphism** of ϕ is called any $\sigma \in \mathrm{PGL}_2(\mathbb{C})$ such that $\phi \circ \sigma = \sigma \circ \phi$.

$$
\begin{tikzcd}
\mathbb{P}^1_x \arrow[r, "\sigma"] \arrow[d, "\phi"'] & \mathbb{P}^1_x \arrow[d, "\phi"] \\
\mathbb{P}^1_z \arrow[r, "\sigma"'] & \mathbb{P}^1_z
\end{tikzcd}
$$

The set of automorphisms of ϕ is denoted by

$$\mathrm{Aut}(\phi) := \{\sigma \in \mathrm{PGL}_2(k) \,:\, \phi^\sigma = \phi\}.$$

It forms a group. For any $\sigma \in \mathrm{Aut}(\phi)$ by $\mathbf{Fix}(\sigma)$ we denote its set of fixed points.

Remark 3. *As in the case of curves, there is some confusion in the literature on what is called the* automorphism group *of ϕ. Throughout this paper by $G := \mathrm{Aut}(\phi)$ we mean the **full automorphism group** of $\phi(x)$ over the algebraic closure of k and not simply some $G \hookrightarrow \mathrm{Aut}(\phi)$ as it is used frequently by many authors.*

Lemma 5. *Let $\phi \in \mathrm{Rat}_d^1$ and $\sigma \in \mathrm{Aut}(\phi)$. Then*
(i) *If $p \in \mathbf{Fix}(\sigma)$ then $\phi(p) \in \mathbf{Fix}(\sigma)$.*
(ii) *If $w \in \mathbf{Fix}(\phi)$ then $\sigma(w) \in \mathbf{Fix}(\phi)$.*

Hence, $\mathrm{Aut}(\phi)$ acts on $\mathbf{Fix}(\phi)$ by permutation. Moreover if $\sigma \in \mathrm{Aut}(\phi)$ is an automorphism of order m, then m divides the cardinality of $\mathbf{Fix}(\phi) \setminus \mathbf{Fix}(\sigma)$. The dimension of the corresponding locus $\delta = s - 1$, where s is the number of orbits on $\mathbf{Fix}(\phi) \setminus \mathbf{Fix}(\sigma)$.

Proof. For any $p \in \mathbf{Fix}(\phi)$, $\sigma(p) \in \mathbf{Fix}(\phi)$ since

$$\phi(\sigma(p)) = \sigma(\phi(p)) = \sigma(p),$$

which implies that $\sigma(p) \in \mathbf{Fix}(\phi)$. If $w \in \mathbf{Fix}(\phi)$ then

$$\sigma(w) = \sigma(\phi(w)) = \phi(\sigma(w)),$$

which implies that $\sigma(w) \in \mathbf{Fix}(\phi)$.

Since $\langle \sigma \rangle$ has no fixed points in $\mathbf{Fix}(\phi) \setminus \mathbf{Fix}(\sigma)$, then it acts transitively on $\mathbf{Fix}(\phi) \setminus \mathbf{Fix}(\sigma)$. Hence, $|\sigma|$ divides its cardinality. We have fixed 0 and ∞ on $\mathbb{P}_x^1$. Hence, the dimension of the family of rational functions $\phi(x)$ is one less than the number of roots of F and G. Hence, this number is exactly $s - 1$. □

Lemma 6. *Let $\sigma \in \mathrm{Aut}(\phi)$ such that $|\sigma| = m$. Then $H := \langle \sigma \rangle$ acts on $\phi^{-1}(0)$ and $\phi^{-1}(\infty)$. Hence, $\phi(x)$ can be written as*

$$\phi(x) = x\, \psi(x^m), \qquad \textit{or equivalently} \qquad \phi(x) = \frac{1}{x}\, \psi(x^m),$$

where $\psi(x)$ is a rational function. Moreover, for $G \cong A_4, S_4, A_5$ then $m = 2, 4, 5$.

Proof. Let $\sigma \in G$ and $t \in \mathbf{Fix}(\sigma)$. For each $\alpha \in \phi^{-1}(t)$ we have $\phi(\sigma(\alpha)) = \sigma(\phi(\alpha)) = \sigma(t) = t$. Then $\langle \sigma \rangle$ acts on the fiber $\phi^{-1}(t)$. From Remark 2 there is $\sigma \in G$ which fixes 0 and ∞. Then $\langle \sigma \rangle$ acts on $\phi^{-1}(0)$ and $\phi^{-1}(\infty)$. Then points in $\phi^{-1}(0)$ and $\phi^{-1}(\infty)$ are

$$\begin{aligned} &\alpha_1, \xi\alpha_1, \dots, \xi^{m-1}\alpha_1,\ \alpha_2, \xi\alpha_2, \dots, \xi^{m-1}\alpha_2,\ \dots, \alpha_r, \xi\alpha_r, \dots, \xi^{m-1}\alpha_r, \\ &\beta_1, \xi\beta_1, \dots, \xi^{m-1}\beta_1,\ \beta_2, \xi\beta_2, \dots, \xi^{m-1}\beta_2,\ \dots, \beta_r, \xi\beta_r, \dots, \xi^{m-1}\beta_r, \end{aligned}$$

where $r = \frac{d-1}{m}$ and $\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_r \in k \setminus \{0, 1, \infty\}$. Hence,

$$\begin{aligned} \mathbf{F}(x) &= \prod_{j=1}^{r} \prod_{i=0}^{m-1} \left(x - \xi_m^i \alpha_j\right) = \prod_{j=1}^{r} (x^m - \alpha_j^m), \\ \mathbf{G}(x) &= \prod_{j=1}^{r} \prod_{i=0}^{m-1} \left(x - \xi_m^i \beta_j\right) = \prod_{j=1}^{r} (x^m - \beta_j^m) \end{aligned}$$

Hence, $\phi(x)$ can be written as

$$\phi(x) = x\, \frac{\mathbf{F}(x^n)}{\mathbf{G}(x^n)} = x\, \frac{x^{rm} + a_{r-1} x^{(r-1)m} + \cdots a_1 x^m + a_0}{x^{rm} + b_{r-1} x^{(r-1)m} + \cdots b_1 x^m + b_0}$$

This completes the proof. □

4.3. **Automorphism groups of rational cubics.** The automorphism groups for cubic rational functions and their parametric families were determined in [11] and in [8]. Here we give a brief treatment to sort out some mistakes and furthermore characterize such families in terms of generators of the ring of invariants $\mathcal{R}_{4,2}$ defined in Eq. (3).

Lemma 7. *Let $\phi \in \mathrm{Rat}_3^1$. Then the following hold:*
 (i) *Elements of $\mathrm{Aut}(\phi)$ have orders at most 4.*
 (ii) $\mathrm{Aut}(\phi)$ *is isomorphic to one of the following C_2, C_3, D_4, D_6, A_4, or S_4.*

Proof. Suppose that $\sigma \in \mathrm{Aut}(\phi)$ of order n. There is no loss of generality to assume that $\sigma = \begin{pmatrix} 1 & 0 \\ 0 & \zeta_n \end{pmatrix}$, where ζ_n denotes a fixed primitive nth root of unity in k. Then, $\phi^\sigma = \phi$ if and only if ϕ commutes with σ. If so, we can construct the pair (f, g), defining ϕ, from the set of pairs of monomials $(x^{r_0} y^{3-r_0}, x^{r_1} y^{3-r_1})$ such that $n \mid r_0 - r_1 - 1$. Since $|r_0 - r_1 - 1| \leq 4$, then $n \leq 4$.

As a result, we only need to characterize cubic rational functions that have an automorphism of order 2 or 3 in terms of their invariants (Recall that if $\sigma \in \mathrm{Aut}(\phi)$ has order 4, then $\sigma^2 \in \mathrm{Aut}(\phi)$ has order 2).

From i) we can eliminate C_n and D_{2n} for $n \geq 5$, as well as A_5, from consideration as possible values for $\mathrm{Aut}(\phi)$. □

For the rest of this section we assume $\phi \in \mathrm{Rat}_3^1$ given by

$$\phi(z) = \frac{c_0 z^3 + c_1 z^2 + c_2 z + c_3}{c_4 z^3 + c_5 z^2 + c_6 z + c_7} \tag{27}$$

or equivalently by a point $\mathfrak{p} \in \mathbb{P}^7$, such that $\mathfrak{p} = [c_0, c_1, \ldots, c_7]$, for which $I_6(\mathfrak{p}) \neq 0$.

Let $W := \mathbf{Fix}(\phi)$ be the set of fixed points, $\sigma \in \mathrm{Aut}(\phi)$ is a non-trivial automorphism of order n. By Lem. 7, we know that $n = 2, 3$, or 4. We now consider each case separately:

4.4. **Involutions.** We assume that $|\sigma| = 2$. We can pick a coordinate in $\mathbb{P}^1$ such that $\sigma(z) = -z$. Then $\mathbf{Fix}(\sigma) = \{0, \infty\}$. From Lem. 5 we have that $\phi(0), \phi(\infty) \in \mathbf{Fix}(\sigma)$. Hence, there are two cases:
 i) ϕ fixes points of $\mathbf{Fix}(\sigma)$
 ii) ϕ permutes points of $\mathbf{Fix}(\sigma)$
We consider each one of them in detail.

4.4.1. *ϕ fixes points of* $\mathbf{Fix}(\sigma) = \{0, \infty\}$. In this case $\phi(z)$ fixes both points in $\mathbf{Fix}(\sigma)$. Hence, $\phi(0) = 0$ which implies that $\phi(z) = z\psi(z^2)$, say for some $a, b \in k$

$$\phi(z) = z\psi(z^2) = z\,\frac{az^2 + b}{cz^2 + d} = z\frac{z^2 + t}{sz^2 + 1}$$

and equivalently $t, s \in k$, such that $I_6(\phi) = t^2 s^2 - 2ts + 1 = (ts - 1)^2 \neq 0$. Hence,

$$\boxed{\phi(z) = \frac{z^3 + tz}{sz^2 + 1}, \qquad \mathfrak{p} = [1 : 0 : t : 0 : 0 : s : 0, 1]} \tag{28}$$

for some $t, s \in k$ such that $ts \neq 1$. Computing the invariants $\xi(\phi)$ we get

$$\begin{aligned}
&\xi_0 = 2(t+3)(s+3), \quad \xi_1 = \frac{1}{2}(t-1)(s-1), \quad \xi_2 = \xi_3 = 0 \\
&\xi_4 = -\frac{1}{3}\left(t^2s^2 + 2t^2s + 3t^2 + 2ts^2 - 8ts - 6t + 3s^2 - 6s + 9\right) \\
&\xi_5 = -(t-s)^2(ts+t+s-3)
\end{aligned}$$

Let $u = t+s$ and $v = ts$. From the resultant locus we get that $v \neq 1$. Then we get

$$\begin{aligned}
&\xi_0 = 2(v+3u+9), \quad \xi_1 = \frac{1}{2}(v-u+1), \quad \xi_2 = \xi_3 = 0 \\
&\xi_4 = -\frac{1}{3}\left(v^2 + 3u^2 + 2uv - 14v - 6u + 9\right), \quad \xi_5 = (u^2-4v)(3-u-v).
\end{aligned}$$

Eliminating u, v we get

$$u = \frac{1}{8}(\xi_0 - 4\xi_1 - 16), \qquad v = \frac{1}{8}(\xi_0 + 12\xi_1 - 24)$$

and the relations among the invariants

$$\begin{cases}
\xi_0\left(3\xi_0^2 + 8\xi_0\xi_1 - 192\xi_0 + 48\xi_1^2 - 768\xi_1 + 96\xi_4 + 3072\right) = 0, \\
\xi_0\left(\xi_0^3 - 4\xi_0^2\xi_1 - 96\xi_0^2 - 16\xi_0\xi_1^2 - 256\xi_0\xi_1 + 3072\xi_0 + 64\xi_1^3\right. \\
\left.-1536\xi_1^2 + 12288\xi_1 + 256\xi_5 - 32768\right) = 0
\end{cases}$$

If $\xi_0 = 0$, then $t = -3$ or $s = -3$, which will be discussed in **??**. Assuming we have $\xi_0 \neq 0$ we have

$$\mathcal{L}_1 : \quad \begin{cases}
3\xi_0^2 + 8\xi_0\xi_1 - 192\xi_0 + 48\xi_1^2 - 768\xi_1 + 96\xi_4 + 3072 = 0, \\
\xi_0^3 - 4\xi_0^2\xi_1 - 96\xi_0^2 - 16\xi_0\xi_1^2 - 256\xi_0\xi_1 + 3072\xi_0 + 64\xi_1^3 \\
-1536\xi_1^2 + 12288\xi_1 + 256\xi_5 - 32768 = 0
\end{cases} \tag{29}$$

4.4.2. *ϕ permutes points of* $\mathbf{Fix}(\sigma)$. Assume that $\mathbf{Fix}(\sigma) \cap \mathbf{Fix}(\phi) = \{0, \infty\}$. From Lemma 2, i) we have that $\phi(0), \phi(\infty) \in \mathbf{Fix}(\sigma)$.

Assume that ϕ permutes 0 and ∞. From Lem. 6 we have that $\phi(z) = \frac{1}{z}\,\psi(z^2)$, for some degree two rational function. Thus

$$\phi(z) = \frac{1}{z}\,\frac{tz^2 + r}{pz^2 + q}$$

with $r \neq 0$ otherwise $\deg \phi < 3$. We can simplify by r and have

$$\boxed{\phi(z) = \frac{t\,z^2 + 1}{z^3 + sz} : \qquad \mathfrak{p} = [0 : t : 0 : 1 : 1 : 0 : s : 0]} \tag{30}$$

for some $t, s \in k$ such that $ts \neq 1$. The resultant in this case is

$$I_6(\phi) = 2ts - 1 \neq 0$$

By computing $\xi(\phi)$ for ϕ given in Eq. (30) we get

$$\begin{aligned}
&\xi_0 = -2\,(t+s)^2, \quad \xi_1 = \frac{1}{6}\left((t-s)^2 - 12\right), \quad \xi_2 = -\frac{1}{36}(t-s)\left((t-s)^2 + 36\right) \\
&\xi_3 = \frac{2}{3}(t-s)(t+s)^2, \quad \xi_4 = -\frac{1}{9}(t+s)^2\left((t-s)^2 + 12\right), \quad \xi_5 = 0
\end{aligned}$$

and letting $u = (t+s)^2$ and $v = t - s$ we obtain a birational parametrization

$$\mathfrak{p} = \left[-2u, \frac{1}{6}\left(v^2 - 12\right), -\frac{1}{36} v \left(v^2 + 36\right), \frac{2}{3} uv, -\frac{1}{9} u(v^2+12), 0 \right]$$

Assume $u \neq 0$. Eliminating u and v we find $u = -\frac{1}{2}\xi_0$ and$v = -3\frac{\xi_3}{\xi_0}$ and the equations for the $\mathcal{L}_2$-locus:

$$\mathcal{L}_2 : \quad \begin{cases} x_1^3 + 18x_1^2 + 96x_1 - 6x_2^2 + 128 = 0 \\ 2x_0^2 x_1 + 4x_0^2 - 3x_3^2 = 0 \\ x_0 x_1 + 4x_0 - 3x_4 = 0 \end{cases} \tag{31}$$

4.5. **An automorphism of order 3.** Let σ be of order 3 taken as in Eq. (26), so $\sigma(z) = \zeta_3 z$. Then by Lem. 6 we have $\langle \sigma \rangle$ acts on the fiber $\phi^{-1}(0)$ which implies that the numerator of $\phi(z)$ is a polynomial $p(z) = z^3 - t$, for some $t \in k^\star$. We can pick 0 and ∞ in $\phi^{-1}(\infty)$. Since $\phi^{-1}(\infty)$ is an orbit of $\langle \sigma \rangle$ and σ fixes them, then one of them must have multiplicity 2. Thus, we can take

$$\boxed{\phi(z) = \frac{z^3 - t}{z}, \quad \mathfrak{p} = [1:0:0:-t:0:0:1:0]} \tag{32}$$

for some $t \in k^*$. The resultant is $I_6(\phi) = -t^3$ which gives the condition that $t \neq 0$. Next, we compute the invariants $\xi(\phi)$, and we find:

$$\xi(\phi) = \left[-2,\ \frac{1}{6},\ \frac{27t+2}{72},\ -\frac{27t+2}{3},\ \frac{54t-1}{9},\ -\frac{t\,(27t-16)}{4} \right]$$

Since $t \neq 0$ then $\xi_0 \neq 0$. By eliminating t from these equations, we obtain

$$\mathcal{L}_3 : \begin{cases} 24\xi_2 + \xi_3 = 0 \\ 144\xi_2 - 9\xi_4 - 5 = 0 \\ 144\xi_2^2 - 40\xi_2 + 3\xi_5 + 1 = 0 \end{cases} \tag{33}$$

4.6. **Extra involutions.** Similarly to [?deg-2] we can assume that there is another involution $\tau(z) = \frac{1}{z}$.

4.6.1. *First case.* In this case $V_4 \hookrightarrow \mathrm{Aut}(\phi)$. It is straightforward to verify that for this to hold, we must have $t = s$. Thus, we can reduce $\phi(x, y)$ to either

$$\phi_1(z) = \frac{tz^2 + 1}{z^3 + tz},$$

where $t^2 \neq 1$. Thus, we can assume

$$\boxed{\phi_1(z) = \frac{tz^2+1}{z^3+tz}, \quad \mathfrak{p} = [0, t, 0, 1, 1, 0, t, 0]} \tag{34}$$

The resultant here is $I_6(\phi) = t^4 - 2t^2 + 1 = (t^2 - 1)^2 \neq 0$. Computing the invariants $\xi(\phi_\pm)$ we get

$$\xi(\phi) = \left[-8t^2,\ -2,\ 0,\ 0,\ -\frac{16t^2}{3},\ 0 \right]$$

and eliminating t we find

$$\mathcal{L}_4 : \quad \begin{cases} 2\xi_0 - 3\xi_4 = 0 \\ \xi_2 = \xi_3 = \xi_5 = 0 \end{cases} \tag{35}$$

4.6.2. *Second case.*

$$\boxed{\phi_2(z)=\frac{1}{z}\,\frac{tz^2-1}{z^2-t}, \quad \mathfrak{p}=[0,t,0,-1,1,0,-t,0]} \tag{36}$$

Its invariants are

$$\xi(\phi_2)=\left[0,\ \frac{2\left(t^2+3\right)}{3},\ -\frac{2t\,(t-3)\,(t+3)}{9},\ 0,\ 0,\ 0\right]$$

The satisfy the equations

$$\mathcal{L}_5:\qquad \begin{cases}\xi_1^3-18\xi_1^2+96\xi_1-6\xi_2^2-128=0\\ \xi_0=\xi_3=\xi_4=\xi_5=0\end{cases} \tag{37}$$

4.7. **Alternating group:** $\mathrm{Aut}(\phi)\cong A_4$**.** Suppose we can extend such a V_4 subgroup within $\mathrm{Aut}(\phi)$ to an A_4 subgroup. In this case, from Eq. (26) we can take $\sigma(z)=-z$ and $\tau(z)=\frac{z+i}{z-i}$. Then

$$\boxed{\phi(z)=\frac{z^3-3}{-3z^2}, \quad \mathfrak{p}=[1:0:0:-3:0:-3,0,0]} \tag{38}$$

The moduli point is

$$\mathcal{L}_6:\xi(\phi)=[0:0:18:0:0:0]\equiv[0:0:1:0:0:0] \tag{39}$$

4.8. **Dihedral group** D_4**.** Here, we can assume that σ is an automorphism of the form:

$$[x:y]\to[x:\zeta_4 y],$$

where ζ_4 is a fixed primitive 4th root of unity in k, is an automorphism.

Since σ^2 in $\mathrm{Aut}(\phi)$ acts as $[x:y]\to[x:-y]$, a situation we discussed earlier, we can begin with a rational function of the form

$$\phi(x,y)=\frac{tx^2y+y^3}{x^3+sxy^2} \quad \text{or} \quad \frac{x^3+txy^2}{sx^2y+y^3}.$$

$$\begin{aligned}\phi(x,y)&=\frac{tx^2y+y^3}{x^3+sxy^2}, \quad \mathfrak{p}=[0:t:0:1:1:0:s:0] \quad \text{or}\\ \phi(x,y)&=\frac{x^3+txy^2}{sx^2y+y^3}, \quad \mathfrak{p}=[1:0:t:0:0:s:0:1]\end{aligned}$$

Verifying that $\sigma\in\mathrm{Aut}(\phi)$, we can confirm that the second form does not yield any possible rational cubic functions, while in the first form, it must be the case that $t=s=0$. Hence

$$\boxed{\phi=\frac{y^3}{x^3}, \qquad \mathfrak{p}=[0,0,0,1,1,0,0,0]} \tag{40}$$

Its invariants are

$$\xi(\phi)=[0,-2,0,0,0,0]\equiv[0,1,0,0,0,0] \tag{41}$$

Finally, we get

$$\xi_0=18, \quad \xi_1=\frac{1}{2}, \quad \xi_4=-3, \quad \xi_2=\xi_3=\xi_5=0 \tag{42}$$

by substituting $t=s=0$ into **??**

TABLE 1. Automorphism loci of degree 3 rational functions

	G	$\phi(z)$	$\mathfrak{p} \in \mathbb{P}^7$	dim	Eq. $\mathcal{L}_i$	$\xi_i = 0$
$\mathcal{L}_1$	C_2	$\frac{z^3+tz}{sz^2+1}$, Eq. (28)	$[1,0,t,0,0,s,0,1]$	2	(29)	ξ_2, ξ_3
$\mathcal{L}_2$	C_2	$\frac{t\,z^2+1}{z^3+sz}$, Eq. (30)	$[0,t,0,1,1,0,s,0]$	2	(31)	ξ_5
$\mathcal{L}_3$	C_3	$\frac{z^3-t}{z}$, Eq. (32)	$[1,0,0,-t,0,0,1,0]$	1	(33)	
$\mathcal{L}_4$	V_4	$\frac{tz^2+1}{z^3+tz}$, Eq. (34)	$[0,t,0,1,1,0,t,0]$	1	(35)	ξ_2, ξ_3, ξ_5
$\mathcal{L}_5$	V_4	$\frac{tz^2-1}{z^3-tz}$, Eq. (36)	$[0,t,0,-1,1,0,-t,0]$	1	(37)	$\xi_0, \xi_3, \xi_4, \xi_5$
$\mathcal{L}_6$	A_4	$\frac{z^3-3}{-3z^2}$, Eq. (38)	$[1,0,0,-3,0,-3,0,0]$	0	(39)	$\xi_0, \xi_1, \xi_3, \xi_4, \xi_5$
$\mathcal{L}_7$	D_4	$\frac{1}{z^3}, Eq.$ (40)	$[0,0,0,1,1,0,0,0]$	0	(41)	$\xi_0, \xi_2, \xi_3, \xi_4, \xi_5$

5. Datasets of rational functions

We want to create a database of rational functions $\phi \in \mathbb{Z}(x)$ of degree $\deg \phi = n$. Data will be stored in a Python dictionary. We will assume that $\phi(x) = \frac{f(x)}{g(x)}$ such that

$$f(x) = \sum_{i=0}^{d} a_i x^i, \quad g(x) = \sum_{i_0}^{d} b_i x^i, \quad \text{and} \quad I_6(\phi) \neq 0$$

Our points will be points in the projective space $\mathbb{P}_{\mathbb{Q}}^{2d+1}$, i.e. points with integer coordinates

$$P_\phi = [a_d : \cdots c_3 : b_d : \cdots c_7] \in \mathbb{P}_{\mathbb{Q}}^{2d+1}$$

such that $\gcd(a_d, \dots, c_7) = 1$. The height of the function $\phi(x)$ is called the projective height of P_ϕ, denoted by $H(\phi)$.

Let us now generate a dataset of rational functions of degree d with a bounded height h. We will denote the set of such rational by $\mathcal{P}_d^h$. In other words

$$\mathcal{P}_d^h := \left\{ P_\phi \in \mathbb{P}_{\mathbb{Q}}^{2d+1} \;\mid\; H_{\mathbb{Q}}(P_\phi) \leq h \right\}$$

For a degree $d \geq 3$ and height h one can use *Sagemath* and count such points as follows:

```
PP = ProjectiveSpace(d, QQ)
Pts = PP.rational_points(h)
```

For every point q in the set *Pts* above we will compute the following

$$\begin{array}{rl} I_6(\mathfrak{p}) & I_6 \text{ invariant} \\ J_6(\mathfrak{p}) & I_6 \text{ invariant} \\ \mathfrak{p} = [\xi_0, \dots, \xi_n] & \text{Weighted moduli point} \end{array}$$

We delete from the database all points with $I_6 = J_6 = 0$ and all the points which do not satisfy Thm. 1. For the rest of the points we compute

$$\begin{array}{rl} (i_1, \dots, i_5) & \text{tuple of absolute invariants} \\ H(\phi) & \text{Height of } \phi(x) \\ \mathrm{Aut}(\phi) & \text{Automorphism group of } \phi(x) \end{array}$$

Our database is organized in a Python dictionary with keys the tuple of absolute invariants. Since duplicate keys are not allowed in a dictionary, this assures that every point in the database represents uniquely the equivalence class of a degree 3 rational function. At a later stage we intend to extend the database and further compute the following

$$\begin{array}{rl} \mathfrak{h}(\mathfrak{p}) & \text{Weighted moduli height} \\ \tilde{\mathfrak{h}}(\mathfrak{p}) & \text{Absolute weighted moduli height} \\ Mon(\phi) & \text{Monodromy group} \\ \text{Fine/Coarse} & \text{True if a fine point and False otherwise} \end{array}$$

Notice that all points in our database are actually fine points since we start from rational functions with integer coefficients. However, we intend to extend this database by randomly generating points in $\mathfrak{p} = [\xi_0, \dots, \xi_n]$ in the weighted projective space. When we start from such points we are able to compute the automorphism group, but the corresponding rational function $\phi(z)$ is not necessarily defined over $\mathbb{Q}$. Here is the Python code to generate the database.

Listing 1. Rational Functions Python Code

```
def rational_functions(h_min, h_max, L):
    # Define projective space and gather rational points
    P = ProjectiveSpace(7, QQ)
    all_points = P.rational_points(h_max)
    points = [p for p in all_points if h_min < max(abs(coord) for
        coord in p) <= h_max]

    print(f"Total number of rational points of heights in ({h_min}, {
        h_max}] is {len(points)}")

    # Filter points based on both I6 and J6 conditions
    filtered_points = [p for p in points if I6(p) != 0 and J6(p) != 0]
    print(f"Number of points satisfying I6(p) != 0 and J6(p) != 0: {
        len(filtered_points)}")

    num_count = 0
    for p in filtered_points:
        key = absinv(p)

        # Skip if the key is already in the dictionary
        if key in L:
            continue

        # Add the point to the dictionary
        L[key] = [p, inv(p), I6(p), AutGroup(p)]
        num_count += 1

        if num_count % 10000 == 0:
            save(L, f'degree-3-functions')
            print(f"Added {num_count} rational functions, current
                dictionary size: {len(L)}")

    print(f"Final count: {len(L)} rational functions added with degree
         3 and height between {h_min} and {h_max}")
    save(L, f'degree-3-functions')
    return
```

There are 2078697 rational functions with naive height ≤ 4 distributed as follows:

```
Counts of each naive height:
Height 0: 0 points
Height 1: 304878 points
Height 2: 555070 points
Height 3: 213708 points
Height 4: 1005104 points
```

When the height increases the size of the rational points in $\mathbb{P}^7$ increases and the database becomes more difficult to handle. This could be avoided if we use the weighted projective space and weighted projective height instead similar to techniques used in [1].

Remark 4. *We notice the following interesting fact. From 2078697 rational functions with naive height* ≤ 4 *only 63 have non trivial automorphism group. Interestingly enough all of them have automorphism group* C_2 *and are in the locus* $\mathcal{L}_1$*. We display them all in the next table.*

5.1. **Preprocessing.** We convert the dictionary into a pandas DataFrame (df2) and then split the tuple columns into separate columns and convert boolean columns to binary. Then we select specific columns ('xi0', 'xi1', 'xi2', 'xi3', 'xi4', 'xi5', 'Weighted-Height') as the feature matrix X. Before performing a given model, we normalize the data using MinMaxScaler to scale data between 0 and 1.

5.2. **Unsupervised Learning.**

5.2.1. *Autoencoder.* We implemented an autoencoder using TensorFlow and Keras to obtain a latent representation of the data. The autoencoder consists of an input layer with 64 neurons, followed by two hidden layers with 32 and 64 neurons, respectively, using ReLU activation functions. We compiled the model using the Adam optimizer and mean squared error loss. The model was trained on the normalized feature matrix with a specified number of epochs and batch size. Then we added Batch Normalization and Dropout layers to improve training stability and prevent overfitting. We also added early stopping with a patience of 5 epochs to monitor the validation loss and stop training when it stops improving.

Observing the performance throughout epochs, we noted a sudden increase in validation loss compared to training loss suggested potential overfitting and instability in the model's performance. To address this issue, several adjustments were made to the autoencoder. These included increasing model complexity with additional layers and implementing dropout regularization to mitigate overfitting. The learning rate was set to 0.001 to facilitate smoother convergence during training, while the patience for early stopping was increased to 10 epochs for better monitoring of validation loss.

Additionally, the Adam optimizer was explicitly defined with the adjusted learning rate to further optimize the training process. These adjustments aimed to enhance the stability and effectiveness of the autoencoder in capturing essential features of the data while minimizing overfitting.

5.2.2. *K-means Clustering on latent representation.* After obtaining the latent representation from the autoencoder, we applied K-Means clustering to identify inherent patterns in the data. Setting the number of clusters to 2, we fitted the K-Means clustering model and assigned cluster labels to each data point.

We performed hyperparameter tuning for K-Means using grid search (GridSearchCV). This allows us to find the best number of clusters based on the latent representation. After finding the best hyperparameters, we initialize a new K-Means model (kmeans-best) with the optimal number of clusters.

We fit the K-means-best model on the latent representation and obtain cluster labels. Then we apply PCA for further dimensionality reduction to visualize the clusters in 2D. Finally, we plot the clusters in a 2D scatter plot using PCA.

5.2.3. *Gaussian Mixture Model (GMM) on latent representation.* In addition to K-Means, we explored Gaussian Mixture Model (GMM) clustering to further understand the underlying data distribution.

Table 2. default

(-3 : 2 : 0 : 0 : -2 : 0 : -1 : 1)	[-56, 0, -3/8, -7, 10, -43/4]
(3 : -2 : 0 : 0 : 2 : 0 : -3 : 1)	[4, 5/3, 25/72, 23/3, -178/9, -203/4]
(-3 : 1 : 0 : 0 : -2 : 0 : -1 : 1)	[-54, -5/6, -2/9, 0, -9, -27]
(3 : -2 : 0 : 0 : 2 : 0 : -2 : 1)	[22, 3/2, 3/4, -18, -11, 5]
(3 : -1 : 0 : 0 : 2 : 0 : -1 : 1)	[46, 3/2, 3/4, -18, -17, 4]
(-3 : 1 : 0 : 0 : -2 : 0 : -2 : 1)	[-56, 0, -3/8, -7, -2, -155/4]
(3 : -1 : 0 : 0 : 2 : 0 : -2 : 1)	[36, 5/3, 25/72, -3, -22, -27/4]
(3 : -1 : 0 : 0 : 2 : 0 : -3 : 1)	[22, 13/6, -2/9, 64/3, -325/9, -19]
(-3 : -1 : 0 : 0 : -2 : 0 : -3 : 1)	[-86, -5/6, -2/9, 64/3, -289/9, -35]
(3 : 1 : 0 : 0 : 2 : 0 : -3 : 1)	[46, 25/6, -91/36, 86/3, -649/9, 100]
(-3 : -3 : 0 : 0 : -2 : 0 : -3 : 1)	[-126, -3/2, -3/4, 18, -45, -216]
(-3 : -2 : 0 : 0 : -2 : 0 : -3 : 1)	[-104, -4/3, -29/72, 77/3, -358/9, -283/4]
(-3 : 1 : 0 : 0 : -2 : 0 : -3 : 1)	[-62, 7/6, -37/36, -22/3, 35/9, -100]
(-3 : 2 : 0 : 0 : -2 : 0 : -3 : 1)	[-56, 8/3, -169/72, -71/3, 290/9, -419/4]
(3 : 0 : 0 : 0 : 2 : 0 : -3 : 1)	[36, 3, -9/8, 27, -54, 189/4]
(3 : 2 : 0 : 0 : 2 : 0 : -3 : 1)	[52, 17/3, -331/72, 91/3, -826/9, 365/4]
(3 : 3 : 0 : 0 : 2 : 0 : -3 : 1)	[54, 15/2, -15/2, 36, -117, -27]
(-3 : 3 : 0 : 0 : -2 : 0 : -1 : 1)	[-62, 7/6, -37/36, -22/3, 251/9, 88]
(-3 : -3 : 0 : 0 : -2 : 0 : -2 : 1)	[-104, -4/3, -79/72, 31/3, -466/9, -883/4]
(-3 : -2 : 0 : 0 : -2 : 0 : -2 : 1)	[-86, -3/2, -3/4, 18, -43, -59]
(-3 : -1 : 0 : 0 : -2 : 0 : -2 : 1)	[-72, -4/3, -29/72, 15, -34, -27/4]
(-3 : 3 : 0 : 0 : -2 : 0 : -2 : 1)	[-56, 8/3, -169/72, -71/3, 398/9, 269/4]
(3 : -2 : 0 : 0 : 2 : 0 : -1 : 1)	[36, 5/3, 83/72, -33, -10, 189/4]
(3 : -3 : 0 : 0 : 2 : 0 : -2 : 1)	[4, 5/3, 83/72, -131/3, -70/9, 397/4]
(3 : 1 : 0 : 0 : 2 : 0 : -2 : 1)	[52, 3, -9/8, 11, -54, 101/4]
(3 : 2 : 0 : 0 : 2 : 0 : -2 : 1)	[54, 25/6, -91/36, 18, -75, -27]
(3 : 3 : 0 : 0 : 2 : 0 : -2 : 1)	[52, 17/3, -331/72, 91/3, -934/9, -755/4]
(-3 : 0 : 0 : 0 : -2 : 0 : -1 : 1)	[-56, -4/3, -29/72, 29/3, -226/9, -35/4]
(3 : 0 : 0 : 0 : 2 : 0 : -1 : 1)	[52, 5/3, 25/72, -25/3, -262/9, -19/4]
(-3 : -3 : 0 : 0 : -2 : 0 : -1 : 1)	[-86, -5/6, -23/18, 44/3, -505/9, -227]
(-3 : -2 : 0 : 0 : -2 : 0 : -1 : 1)	[-72, -4/3, -79/72, 21, -46, -243/4]
(-3 : -1 : 0 : 0 : -2 : 0 : -1 : 1)	[-62, -3/2, -3/4, 18, -37, -4]
(3 : -3 : 0 : 0 : 2 : 0 : -1 : 1)	[22, 13/6, 31/18, -172/3, -109/9, 173]
(3 : 1 : 0 : 0 : 2 : 0 : -1 : 1)	[54, 13/6, -2/9, 0, -45, -27]
(3 : 2 : 0 : 0 : 2 : 0 : -1 : 1)	[52, 3, -9/8, 11, -66, -443/4]
(3 : 3 : 0 : 0 : 2 : 0 : 0 : 1)	[36, 3, -9/8, 27, -90, -1539/4]
(3 : 2 : 0 : 0 : 2 : 0 : 0 : 1)	[46, 13/6, -2/9, 16/3, -547/9, -172]
(3 : -1 : 0 : 0 : 2 : 0 : 0 : 1)	[52, 5/3, 83/72, -83/3, -154/9, 5/4]
(3 : -2 : 0 : 0 : 2 : 0 : 0 : 1)	[46, 13/6, 31/18, -124/3, -115/9, 64]
(3 : -3 : 0 : 0 : 2 : 0 : 0 : 1)	[36, 3, 21/8, -63, -18, 837/4]
(-3 : -4 : 0 : 0 : -2 : 0 : -4 : 1)	[-182, -3/2, -3/4, 18, -43, -499]
(-3 : -3 : 0 : 0 : -2 : 0 : -4 : 1)	[-152, -4/3, -29/72, 125/3, -382/9, -899/4]
(-3 : -2 : 0 : 0 : -2 : 0 : -4 : 1)	[-126, -5/6, -2/9, 48, -43, -108]
(-3 : -1 : 0 : 0 : -2 : 0 : -4 : 1)	[-104, 0, -3/8, 41, -38, -403/4]
(-3 : 0 : 0 : 0 : -2 : 0 : -4 : 1)	[-86, 7/6, -37/36, 74/3, -211/9, -155]
(-3 : 1 : 0 : 0 : -2 : 0 : -4 : 1)	[-72, 8/3, -169/72, 3, 2, -891/4]
(-3 : 2 : 0 : 0 : -2 : 0 : -4 : 1)	[-62, 9/2, -9/2, -20, 37, -256]
(-3 : 3 : 0 : 0 : -2 : 0 : -4 : 1)	[-56, 20/3, -551/72, -121/3, 698/9, -827/4]
(-3 : 4 : 0 : 0 : -2 : 0 : -4 : 1)	[-54, 55/6, -431/36, -54, 117, -27]
(3 : -1 : 0 : 0 : 2 : 0 : -4 : 1)	[4, 3, -9/8, 59, -66, -83/4]
(3 : 0 : 0 : 0 : 2 : 0 : -4 : 1)	[22, 25/6, -91/36, 182/3, -787/9, 101]
(3 : 1 : 0 : 0 : 2 : 0 : -4 : 1)	[36, 17/3, -331/72, 57, -106, 837/4]
(3 : 2 : 0 : 0 : 2 : 0 : -4 : 1)	[46, 15/2, -15/2, 52, -123, 256]
(3 : 3 : 0 : 0 : 2 : 0 : -4 : 1)	[52, 29/3, -821/72, 149/3, -1282/9, 773/4]
(3 : 4 : 0 : 0 : 2 : 0 : -4 : 1)	[54, 73/6, -593/36, 54, -171, -27]
(-3 : -4 : 0 : 0 : -2 : 0 : -3 : 1)	[-152, -4/3, -79/72, -17/3, -490/9, -2075/4]
(-3 : 4 : 0 : 0 : -2 : 0 : -3 : 1)	[-56, 20/3, -551/72, -121/3, 806/9, 725/4]
(3 : 4 : 0 : 0 : 2 : 0 : -3 : 1)	[52, 29/3, -821/72, 149/3, -1390/9, -1211/4]
(-3 : -4 : 0 : 0 : -2 : 0 : -2 : 1)	[-126, -5/6, -23/18, -12, -67, -540]
(-3 : 4 : 0 : 0 : -2 : 0 : -2 : 1)	[-62, 9/2, -9/2, -20, 61, 292]

We evaluated the performance of GMM clustering and compared it with the results obtained from K-Means.

5.2.4. *K-means Clustering on original data.* We applied K-Means clustering directly on the original data to explore its clustering performance without using the latent representation obtained from the autoencoder. However, the accuracy of K-Means clustering on the original data was found to be very low, reaching only 14%. This suggests that the model struggled to identify meaningful clusters in the dataset.

5.2.5. *Gaussian Mixture Model (GMM) on original data.* We also investigated the performance of Gaussian Mixture Model (GMM) clustering on the original data. Surprisingly, GMM performed significantly better on the original data compared to the latent representation obtained from the autoencoder. It achieved an accuracy of 94%, indicating its effectiveness in capturing the underlying data distribution directly from the original feature space.

5.2.6. *Evaluation Measure.* To assess the quality of clustering, we employed the adjusted Rand score as our evaluation metric. The adjusted Rand score measures the similarity between true cluster assignments and the clustering results, accounting for chance agreement. Higher scores indicate better clustering performance. We got these results:

- Adjusted Rand score for K-Means using autoencoder: 0.43
- Adjusted Rand score for K-Means on original data: 0.15
- Adjusted Rand score for GMM using autoencoder: 0.44
- Adjusted Rand score for GMM on original data: 0.94

These ARI scores provide insights into the effectiveness of each clustering algorithm in capturing meaningful patterns in the dataset. The significantly higher ARI score for GMM on the original data underscores its superior performance compared to K-Means and even GMM using the latent representation obtained from the autoencoder.

5.3. **Supervised learning.** K-Neighbors Classifier is a non-parametric method used for classification and regression tasks. It operates by assigning the class membership of an instance based on the majority class among its k nearest neighbors in the feature space. The accuracy of K-Neighbors classifier in classifying L 2 points in one group was 99.9%.

Supervised learning models exhibited superior performance in classifying curves into the L 2 space. K-Neighbors Classifier classified the points with the highest accuracy, 99,9%.

To evaluate the performance of supervised models we used the F1 score, which is a measure of a models accuracy that considers both the precision and recall of the model. It is the harmonic mean of precision and recall, and it ranges from 0 to 1, where 1 indicates perfect precision and recall, and 0 indicates the worst possible precision and recall.

F1 Score for Logistic Regression was 0.3123, suggesting that the logistic regression model has relatively low performance in terms of both precision and recall. It may indicate that the model struggles to correctly classify positive instances while minimizing false positives.

F1 Score for Random Forest was 0.7087, which indicates that the random forest model performs reasonably well, achieving a good balance between precision and recall.

F1 Score for Multinomial Naive Bayes was very low, 0.1719. This score suggests that the multinomial Naive Bayes model has poor performance, with both precision and recall being low.

F1 Score for k-Nearest Neighbors was 0.9994. This score indicates very high performance for the k-nearest neighbors' model, with near-perfect precision and recall. It suggests that the model is highly accurate in classifying positive instances while minimizing false positives and false negatives.

References

[1] Elira Shaska and Tony Shaska, *Machine learning for moduli space of genus two curves and an application to isogeny based cryptography*, Journal of Algebraic Combinatorics (2025), available at 2403.17250.

[2] T. Shaska, *Polynomials, galois groups, and machine learning*, Recent advances in mathematics and artificial intelligence, 2025, pp. 1–40.

[3] L. Beshaj, J. Gutierrez, and T. Shaska, *Weighted greatest common divisors and weighted heights*, J. Number Theory **213** (2020), 319–346. MR4091944

[4] Sajad Salami and Tony Shaska, *Local and global heights on weighted projective varieties*, Houston J. Math. **49** (2023), no. 3, 603–636 (English).

[5] Nikita Miasnikov, Brian Stout, and Phillip Williams, *Automorphism loci for the moduli space of rational maps*, Acta Arith. **180** (2017), no. 3, 267–296. MR3709645

[6] Brandon Gontmacher, Benjamin Hutz, Grayson Jorgenson, Srinjoy Srimani, and Simon Xu, *Automorphism loci for degree 3 and degree 4 endomorphisms of the projective line*, New York J. Math. **27** (2021), 1613–1702. MR4359208

[7] Elira Curri, *On the stability of binary forms and their weighted heights*, Albanian J. Math. **16** (2022), no. 1, 3–23. MR4448533

[8] Lloyd William West, *The moduli space of rational maps*, Ph.D. Thesis, 2015.

[9] T. Shaska, *Determining the automorphism group of a hyperelliptic curve*, Proceedings of the 2003 International Symposium on Symbolic and Algebraic Computation, 2003, pp. 248–254. MR2035219

[10] ———, *Subvarieties of the hyperelliptic moduli determined by group actions*, Serdica Math. J. **32** (2006), no. 4, 355–374. MR2287373

[11] Brandon Gontmacher, Benjamin Hutz, Grayson Jorgenson, Srinjoy Srimani, and Simon Xu, *Automorphism loci for degree 3 and degree 4 endomorphisms of the projective line*, New York J. Math. **27** (2021), 1613–1702. MR4359208

[12] Joao Alberto de Faria and Benjamin Hutz, *Automorphism groups and invariant theory on* $\mathbb{P}^N$, J. Algebra Appl. **17** (2018), no. 9, 1850162, 38. MR3846410

Appendix A. Invariants of rational cubics

$$
\begin{aligned}
\xi_0 =& 2(3c_2c_0 + c_2c_5 - c_1^2 - 2c_1c_6 + 9c_0c_7 + 3c_7c_5 - c_6^2)\\
\xi_1 =& -\frac{1}{6}(12c_3c_4 + 3c_2c_0 - 3c_2c_5 - c_1^2 + 2c_1c_6 - 3c_0c_7 + 3c_7c_5 - c_6^2)\\
\xi_2 =& -\frac{1}{72}(72c_3c_1c_4 + 27c_3c_0^2 - 54c_3c_0c_5 - 72c_3c_6c_4 + 27c_3c_5^2 - 27c_2^2c_4 - 9c_2c_1c_0 + 9c_2c_1c_5\\
& + 9c_2c_0c_6 + 54c_2c_7c_4 - 9c_2c_6c_5 + 2c_1^3 - 6c_1^2c_6 + 9c_1c_0c_7 - 9c_1c_7c_5 + 6c_1c_6^2 - 9c_0c_7c_6\\
& - 27c_7^2c_4 + 9c_7c_6c_5 - 2c_6^3)\\
\xi_3 =& \frac{1}{3}(27c_3c_0^2 + 18c_3c_0c_5 + 3c_3c_5^2 - 3c_2^2c_4 - 9c_2c_1c_0 + c_2c_1c_5 - 15c_2c_0c_6 + 9c_7c_6c_5 - 2c_6^3\\
& - 18c_2c_7c_4 - c_2c_6c_5 + 2c_1^3 + 2c_1^2c_6 + 9c_1c_0c_7 + 15c_1c_7c_5 - 2c_1c_6^2 - 9c_0c_7c_6 - 27c_7^2c_4)\\
\xi_4 =& \frac{1}{9} - (18c_3c_2c_0c_4 + 6c_3c_2c_5c_4 + 12c_3c_1^2c_4 - 36c_3c_1c_0c_5 + 24c_3c_1c_6c_4 - 12c_3c_1c_5^2 + 54c_3c_0^2c_6\\
& + 54c_3c_0c_7c_4 + 18c_3c_7c_5c_4 + 12c_3c_6^2c_4 - 6c_3c_6c_5^2 - 6c_2^2c_1c_4 + 9c_2^2c_0^2 + 6c_2^2c_0c_5 - 12c_2^2c_6c_4 + 3c_2^2c_5^2\\
& - 6c_2c_1^2c_0 + 2c_2c_1^2c_5 - 6c_2c_1c_0c_6 + 2c_2c_1c_6c_5 - 18c_2c_0^2c_7 - 24c_2c_0c_7c_5 + 6c_2c_0c_6^2 - 36c_2c_7c_6c_4\\
& + 6c_2c_7c_5^2 + 2c_2c_6^2c_5 + c_1^4 + 6c_1^2c_0c_7 + 6c_1^2c_7c_5 - 2c_1^2c_6^2 - 6c_1c_0c_7c_6 + 54c_1c_7^2c_4 - 6c_1c_7c_6c_5\\
& + 27c_0^2c_7^2 - 18c_0c_7^2c_5 + 6c_0c_7c_6^2 + 9c_7^2c_5^2 - 6c_7c_6^2c_5 + c_6^4)\\
\xi_5 =& -\frac{1}{4}(36c_3^2c_1c_0^2c_4 + 24c_3^2c_1c_0c_5c_4 + 4c_3^2c_1c_5^2c_4 + 27c_3^2c_0^4 + 36c_3^2c_0^2c_6c_4 - 18c_3^2c_0^2c_5^2 - 18c_7^3c_6c_5c_4\\
& + 24c_3^2c_0c_6c_5c_4 - 8c_3^2c_0c_5^3 + 4c_3^2c_6c_5^2c_4 - c_3^2c_5^4 + 4c_3c_2^2c_1c_4^2 - 6c_3c_2^2c_0^2c_4 - 8c_3c_2^2c_0c_5c_4\\
& + 4c_3c_2^2c_6c_4^2 - 2c_3c_2^2c_5^2c_4 - 8c_3c_2c_1^2c_0c_4 - 8c_3c_2c_1^2c_5c_4 - 18c_3c_2c_1c_0^3 - 18c_3c_2c_1c_0^2c_5\\
& - 16c_3c_2c_1c_0c_6c_4 + 2c_3c_2c_1c_0c_5^2 + 24c_3c_2c_1c_7c_4^2 - 16c_3c_2c_1c_6c_5c_4 + 2c_3c_2c_1c_5^3 + 18c_3c_2c_0^3c_6\\
& + 6c_3c_2c_0^2c_6c_5 - 24c_3c_2c_0c_7c_5c_4 - 8c_3c_2c_0c_6^2c_4 + 6c_3c_2c_0c_6c_5^2 + 24c_3c_2c_7c_6c_4^2 - 8c_3c_2c_7c_5^2c_4\\
& - 8c_3c_2c_6^2c_5c_4 + 2c_3c_2c_6c_5^3 + 4c_3c_1^3c_0^2 + 8c_3c_1^3c_0c_5 + 4c_3c_1^3c_5^2 - 12c_3c_1^2c_0^2c_6 + 24c_3c_1^2c_0c_7c_4\\
& - 8c_3c_1^2c_0c_6c_5 - 8c_3c_1^2c_7c_5c_4 + 4c_3c_1^2c_6c_5^2 + 18c_3c_1c_0^3c_7 - 6c_3c_1c_0^2c_7c_5 + 48c_3c_1c_0c_7c_6c_4\\
& + 14c_3c_1c_0c_7c_5^2 - 16c_3c_1c_0c_6^2c_5 + 36c_3c_1c_7^2c_4^2 - 16c_3c_1c_7c_6c_5c_4 + 6c_3c_1c_7c_5^3 + 18c_3c_0^3c_7c_6\\
& + 54c_3c_0^2c_7^2c_4 - 42c_3c_0^2c_7c_6c_5 + 16c_3c_0^2c_6^3 + 24c_3c_0c_7c_6^2c_4 - 10c_3c_0c_7c_6c_5^2 + 36c_3c_7^2c_6c_4^2\\
& - 6c_3c_7^2c_5^2c_4 - 8c_3c_7c_6^2c_5c_4 + 2c_3c_7c_6c_5^3 - c_2^4c_4^2 + 2c_2^3c_1c_0c_4 + 2c_2^3c_1c_5c_4 + 4c_2^3c_0^3 + 4c_2^3c_0^2c_5\\
& + 6c_2^3c_0c_6c_4 - 8c_2^3c_7c_4^2 + 2c_2^3c_6c_5c_4 - c_2^2c_1^2c_0^2 - 2c_2^2c_1^2c_0c_5 - c_2^2c_1^2c_5^2 + 2c_2^2c_1c_0^2c_6 + 4c_7^3c_5^3 + 4c_7^2c_6^3c_4\\
& - 10c_2^2c_1c_0c_7c_4 + 6c_2^2c_1c_7c_5c_4 + 4c_2^2c_1c_6^2c_4 - 2c_2^2c_1c_6c_5^2 - 12c_2^2c_0^3c_7 - 4c_2^2c_0^2c_7c_5 + 3c_2^2c_0^2c_6^2\\
& + 14c_2^2c_0c_7c_6c_4 - 8c_2^2c_0c_7c_5^2 + 2c_2^2c_0c_6^2c_5 - 18c_2^2c_7^2c_4^2 + 2c_2^2c_7c_6c_5c_4 + 4c_2^2c_6^3c_4 - c_2^2c_6^2c_5^2\\
& + 2c_2c_1^2c_0^2c_7 + 4c_2c_1^2c_0c_7c_5 - 16c_2c_1^2c_7c_6c_4 + 2c_2c_1^2c_7c_5^2 - 8c_2c_1c_0^2c_7c_6 - 42c_2c_1c_0c_7^2c_4\\
& + 8c_2c_1c_0c_7c_6c_5 + 6c_2c_1c_7^2c_5c_4 - 8c_2c_1c_7c_6^2c_4 + 24c_2c_0^2c_7^2c_5 - 10c_2c_0^2c_7c_6^2 - 6c_2c_0c_7^2c_6c_4\\
& - 4c_2c_0c_7^2c_5^2 + 4c_2c_0c_7c_6^2c_5 - 18c_2c_7^2c_6c_5c_4 + 4c_2c_7^2c_5^3 + 8c_2c_7c_6^3c_4 - 2c_2c_7c_6^2c_5^2 + 16c_1^3c_7^2c_4\\
& + 3c_1^2c_0^2c_7^2 - 10c_1^2c_0c_7^2c_5 + 3c_1^2c_7^2c_5^2 + 6c_1c_0^2c_7^2c_6 + 18c_1c_0c_7^3c_4 - 8c_1c_0c_7^2c_6c_5 + 18c_1c_7^3c_5c_4\\
& - 12c_1c_7^2c_6^2c_4 + 2c_1c_7^2c_6c_5^2 + 3c_0^2c_7^2c_6^2 + 18c_0c_7^3c_6c_4 - 12c_0c_7^3c_5^2 + 2c_0c_7^2c_6^2c_5 + 27c_7^4c_4^2 - c_7^2c_6^2c_5^2)
\end{aligned}
$$

$$
\begin{aligned}
J_6 = {} & 81c_3^4c_0^2 - 54c_3^3c_2c_1c_0 + 54c_3^3c_2c_0c_4 + 12c_3^3c_1^3 - 36c_3^3c_1^2c_4 + 54c_3^3c_1c_0c_5 \\
& - 108c_3^3c_1c_4^2 + 108c_3^3c_0^2c_6 + 378c_3^3c_0c_5c_4 + 324c_3^3c_4^3 + 12c_3^2c_2^3c_0 - 3c_3^2c_2^2c_1^2 \\
& + 6c_3^2c_2^2c_1c_4 - 36c_3^2c_2^2c_0c_5 + 45c_3^2c_2^2c_4^2 + 6c_3^2c_2c_1^2c_5 - 126c_3^2c_2c_1c_0c_6 \\
& - 60c_3^2c_2c_1c_5c_4 - 162c_3^2c_2c_0c_6c_4 - 108c_3^2c_2c_0c_5^2 - 234c_3^2c_2c_5c_4^2 + 28c_3^2c_1^3c_6 \\
& - 18c_3^2c_1^2c_0c_7 + 12c_3^2c_1^2c_6c_4 + 45c_3^2c_1^2c_5^2 - 108c_3^2c_1c_0c_7c_4 - 18c_3^2c_1c_0c_6c_5 \\
& - 252c_3^2c_1c_6c_4^2 + 150c_3^2c_1c_5^2c_4 + 54c_3^2c_0^2c_6^2 - 162c_3^2c_0c_7c_4^2 + 162c_3^2c_0c_6c_5c_4 \\
& - 60c_3^2c_0c_5^3 - 108c_3^2c_6c_4^3 + 45c_3^2c_5^2c_4^2 + 40c_3c_2^3c_0c_6 - 10c_3c_2^2c_1^2c_6 + 48c_3c_2^2c_1c_0c_7 \\
& + 20c_3c_2^2c_1c_6c_4 + 144c_3c_2^2c_0c_7c_4 + 72c_3c_2^2c_0c_6c_5 + 150c_3c_2^2c_6c_4^2 - 10c_3c_2c_1^3c_7 \\
& - 18c_3c_2c_1^2c_7c_4 - 28c_3c_2c_1^2c_6c_5 + 96c_3c_2c_1c_0c_7c_5 - 66c_3c_2c_1c_0c_6^2 + 162c_3c_2c_1c_7c_4^2 \\
& - 104c_3c_2c_1c_6c_5c_4 + 288c_3c_2c_0c_7c_5c_4 - 126c_3c_2c_0c_6^2c_4 + 24c_3c_2c_0c_6c_5^2 + 378c_3c_2c_7c_4^3 \\
& - 60c_3c_2c_6c_5c_4^2 - 22c_3c_1^3c_7c_5 + 20c_3c_1^3c_6^2 - 12c_3c_1^2c_0c_7c_6 - 126c_3c_1^2c_7c_5c_4 + 68c_3c_1^2c_6^2c_4 \\
& + 6c_3c_1^2c_6c_5^2 - 72c_3c_1c_0c_7c_6c_4 + 48c_3c_1c_0c_7c_5^2 - 30c_3c_1c_0c_6^2c_5 - 162c_3c_1c_7c_5c_4^2 \\
& + 12c_3c_1c_6^2c_4^2 + 20c_3c_1c_6c_5^2c_4 + 12c_3c_0^2c_6^3 - 108c_3c_0c_7c_6c_4^2 + 144c_3c_0c_7c_5^2c_4 - 18c_3c_0c_6^2c_5c_4 \\
& - 8c_3c_0c_6c_5^3 + 54c_3c_7c_5c_4^3 - 36c_3c_6^2c_4^3 + 6c_3c_6c_5^2c_4^2 - 16c_2^4c_0c_7 + 4c_2^3c_1^2c_7 - 8c_2^3c_1c_7c_4 \\
& - 64c_2^3c_0c_7c_5 + 12c_2^3c_0c_6^2 - 60c_2^3c_7c_4^2 + 20c_2^2c_1^2c_7c_5 - 3c_2^2c_1^2c_6^2 + 16c_2^2c_1c_0c_7c_6 \\
& + 24c_2^2c_1c_7c_5c_4 + 6c_2^2c_1c_6^2c_4 + 48c_2^2c_0c_7c_6c_4 - 96c_2^2c_0c_7c_5^2 + 28c_2^2c_0c_6^2c_5 - 108c_2^2c_7c_5c_4^2 \\
& + 45c_2^2c_6^2c_4^2 - 6c_2c_1^3c_7c_6 - 30c_2c_1^2c_7c_6c_4 + 28c_2c_1^2c_7c_5^2 - 10c_2c_1^2c_6^2c_5 + 32c_2c_1c_0c_7c_6c_5 \\
& - 10c_2c_1c_0c_6^3 - 18c_2c_1c_7c_6c_4^2 + 72c_2c_1c_7c_5^2c_4 - 28c_2c_1c_6^2c_5c_4 + 96c_2c_0c_7c_6c_5c_4 - 64c_2c_0c_7c_5^3 \\
& - 22c_2c_0c_6^3c_4 + 20c_2c_0c_6^2c_5^2 + 54c_2c_7c_6c_4^3 - 36c_2c_7c_5^2c_4^2 + 6c_2c_6^2c_5c_4^2 + c_1^4c_7^2 + 12c_1^3c_7^2c_4 \\
& - 10c_1^3c_7c_6c_5 + 4c_1^3c_6^3 - 2c_1^2c_0c_7c_6^2 + 54c_1^2c_7^2c_4^2 - 66c_1^2c_7c_6c_5c_4 + 12c_1^2c_7c_5^3 + 20c_1^2c_6^3c_4 \\
& - 3c_1^2c_6^2c_5^2 - 12c_1c_0c_7c_6^2c_4 + 16c_1c_0c_7c_6c_5^2 - 6c_1c_0c_6^3c_5 + 108c_1c_7^2c_4^3 - 126c_1c_7c_6c_5c_4^2 \\
& + 40c_1c_7c_5^3c_4 + 28c_1c_6^3c_4^2 - 10c_1c_6^2c_5^2c_4 + c_0^2c_6^4 - 18c_0c_7c_6^2c_4^2 + 48c_0c_7c_6c_5^2c_4 - 16c_0c_7c_5^4 \\
& - 10c_0c_6^3c_5c_4 + 4c_0c_6^2c_5^3 + 81c_7^2c_4^4 - 54c_7c_6c_5c_4^3 + 12c_7c_5^3c_4^2 + 12c_6^3c_4^3 - 3c_6^2c_5^2c_4^2
\end{aligned}
$$

Department of Mathematics, American University of Cairo, Egypt
Email address: eslammath@aucegypt.edu

Department of Computer Science, Oakland University, Rochester, MI, 48309
Email address: elirashaska@oakland.edu

Department of Mathematics and Statistics, Oakland University, Rochester, MI, 48309
Email address: tanush@umich.edu