

# POLYNOMIALS, GALOIS GROUPS, AND DEEP LEARNING

ELIRA SHASKA

*Department of Computer Science  
College of Computer Science and Engineering,  
Oakland University,  
Rochester, MI, 48309.*

TONY SHASKA

*Department of Mathematics and Statistics,  
College of Arts and Sciences  
Oakland University,  
Rochester, MI, 48309*

ABSTRACT. This paper introduces a novel approach to understanding Galois theory, one of the foundational areas of algebra, through the lens of machine learning. By analyzing polynomial equations with machine learning techniques, we aim to streamline the process of determining solvability by radicals and explore broader applications within Galois theory. This summary encapsulates the background, methodology, potential applications, and challenges of using data science in Galois theory.

## CONTENTS

1. Introduction	1
2. Preliminaries	3
3. Equivalences of polynomials	5
4. Heights of polynomials	8
5. Binary forms	9
6. Galois groups of a polynomials	14
7. Reduction modulo $p$	21
8. Transitive groups	22
9. Databases	23
10. Neuro-symbolic networks	26
11. Concluding remarks	26
References	27

## 1. INTRODUCTION

Galois theory, a cornerstone of modern algebra, provides profound insights into the solvability of polynomial equations. Since its inception by Évariste Galois, it has explained why there are no general formulas for polynomials of degree five or higher by radicals, unlike the well-known quadratic, cubic, and quartic

formulas. This theory links the algebraic structure of field extensions to the symmetry of polynomial roots encapsulated by their Galois groups. While traditional methods allow us to determine solvability for lower-degree polynomials through invariants like discriminants, the complexity escalates dramatically for higher degrees, where the Galois group might not be solvable, leading to no radical solution.

This project embarks on an innovative journey to merge the abstract realm of Galois theory with the practical capabilities of machine learning (ML). Our goal is to harness ML's pattern recognition and prediction abilities to address some of the most challenging aspects of Galois theory, potentially revolutionizing how we understand and approach polynomial solvability and other related problems. At the heart of Galois theory is the connection between a polynomial's roots and its Galois group, which describes how these roots can be permuted while preserving the field operations. A polynomial is solvable by radicals if its Galois group is solvable; this means there exists a chain of normal subgroups where each quotient is cyclic, allowing for the roots to be constructed by sequential additions, multiplications, and extractions of roots. However, for degrees five and above, generic polynomials often have non-solvable groups like  $S_n$  (symmetric group), rendering them unsolvable by radicals.

We propose an approach where we compile or generate datasets of polynomials with known Galois groups. Key to our approach will be identifying or creating features from polynomials that are indicative of Galois group properties or solvability. These might include traditional invariants like discriminants or novel features derived from root distributions or algebraic properties. Using supervised learning, we aim to predict the Galois group or solvability of polynomials, potentially using neural networks for their ability to handle complex patterns or decision trees for interpretability. Unsupervised methods could explore clustering of polynomials, perhaps revealing new mathematical insights. By learning from simpler polynomials, we hope to generalize these insights to more complex polynomials, possibly using techniques like transfer learning where models adapt knowledge from one task to another.

This integration could lead to automated solvability prediction, offering mathematicians tools to quickly assess if a polynomial can be solved by radicals, and might uncover patterns or invariants not yet recognized by traditional mathematics. The methodology could extend to other areas like field theory or algebraic geometry. However, several challenges loom, including the computational cost of handling high-degree polynomials, ensuring interpretability of ML models to enhance theoretical understanding, and balancing between providing practical tools and contributing to the theoretical body of Galois theory.

This project stands at the intersection of pure mathematics and cutting-edge computational science. By leveraging machine learning, we aim not only to solve practical problems within Galois theory but also to catalyze new theoretical advancements. This exploration could redefine how we approach some of the oldest and most fundamental questions in algebra, potentially opening new avenues for research in both mathematics and computer science.

A neuro-symbolic network is a type of artificial intelligence system that combines the strengths of neural networks (good at pattern recognition) with symbolic reasoning (based on logic and rules) to create models that can both learn from data and reason through complex situations, essentially mimicking human-like cognitive abilities by understanding and manipulating symbols to make decisions; this approach aims to overcome limitations of either method alone, providing better explainability and adaptability in AI systems. In this paper we experiment with such models to study some classical questions of Galois theory.

## 2. PRELIMINARIES

In this section we will go over some preliminary results on polynomials. Even though we will start with the general setup of polynomials defined over number fields and their rings of integers, later in the paper we will mostly focus on  $\mathbb{Q}$  and its ring of integers  $\mathbb{Z}$ . For any field  $k$ ,  $\mathbb{A}_k^n$  and  $\mathbb{P}_k^n$  denote the affine and projective spaces of dimension  $n$  over  $k$ , respectively.

**2.1. Polynomials.** Let  $R$  is a commutative ring with identity. An expression of the form

$$(1) \quad f(x) = \sum_{i=0}^n a_i x^i = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n,$$

where  $a_i \in R$  and  $a_n \neq 0$ , is called a **polynomial over  $R$**  with **variable  $x$** . The elements  $a_0, a_1, \dots, a_n$  are called **coefficients** of  $f(x)$ . The coefficient  $a_n$  is called the **leading coefficient**. A polynomial is called **monic** if its leading coefficient is 1.

If  $n$  is the largest non negative integer for which  $a_n \neq 0$ , then we say that the **degree** of  $f(x)$  is  $n$  and write  $\deg f(x) = n$ . The set of all polynomials, with coefficient in a ring  $R$  is denoted by  $R[x]$ . It is also a commutative ring with identity. Two **polynomials are equal** if their corresponding coefficients are equal, so if we have

$$(2) \quad \begin{aligned} p(x) &= a_0 + a_1 x + \cdots + a_n x^n \\ q(x) &= b_0 + b_1 x + \cdots + b_m x^m, \end{aligned}$$

then  $p(x) = q(x)$  if and only if  $a_i = b_i$  for every  $i = 0, \dots, \max\{m, n\}$ .

Let  $p(x)$  and  $q(x)$  be polynomials in  $R[x]$ , where  $R$  is a integral ring. Then,

$$\deg(p \cdot q) = \deg p + \deg q.$$

Moreover,  $R[x]$  is a integral ring. If  $\mathbb{F}$  is a field, then  $\mathbb{F}[x]$  is a Euclidean domain with norm  $N : \mathbb{F}[x] \rightarrow \mathbb{Z}^{\geq 0}$ , such that  $N(p(x)) = \deg(p(x))$ .

**Lemma 1** (Division Algorithm). *Let  $f(x)$  and  $g(x)$  be two nonzero polynomials in  $\mathbb{F}[x]$ , where  $\mathbb{F}$  is a field and  $g(x)$  is a non-constant polynomial. Then, there exist unique polynomials  $q(x), r(x) \in \mathbb{F}[x]$  such that*

$$f(x) = g(x)q(x) + r(x),$$

where  $\deg r(x) < \deg g(x)$  and  $r(x)$  is a nonzero polynomial.

Let  $p(x)$  be a polynomial in  $\mathbb{F}[x]$  and  $\alpha \in \mathbb{F}$ . We say that  $\alpha$  is a **zero** or **root** of  $p(x)$ , if  $p(x)$  is in the kernel of the homomorphism  $\phi_\alpha$  or we say  $\alpha$  is a zero of  $p(x)$  if  $p(\alpha) = 0$ .

**Corollary 1.** *Let  $\mathbb{F}$  be a field. An element  $\alpha \in \mathbb{F}$  is a zero of  $p(x) \in \mathbb{F}[x]$ , if and only if  $(x - \alpha)$  is a factor of  $p(x)$  in  $\mathbb{F}[x]$ . A nonzero polynomial  $p(x)$  with degree  $n$  in  $\mathbb{F}[x]$  has at most  $n$  distinct zeroes in  $\mathbb{F}$ .*

A monic polynomial  $d(x)$  is called **greatest common divisor** of polynomials  $p(x), q(x) \in \mathbb{F}[x]$  if  $d(x)$  divides  $p(x)$  and  $q(x)$ ; and if for every other polynomial  $d'(x)$  that divides  $p(x)$  and  $q(x)$ ,  $d'(x) \mid d(x)$ . We write

$$d(x) = \gcd(p(x), q(x)).$$

Two polynomials  $p(x)$  and  $q(x)$  are **relatively prime** if  $\gcd(p(x), q(x)) = 1$ . Similarly as for the greatest common divisor of integers, we have the following:

**Lemma 2.** *Let  $\mathbb{F}$  be a field and assume that  $d(x)$  is the greatest common divisor of two polynomials  $p(x)$  and  $q(x)$  in  $\mathbb{F}[x]$ . Then, there exist polynomials  $r(x)$  and  $s(x)$  such that*

$$d(x) = r(x) \cdot p(x) + s(x) \cdot q(x).$$

Moreover, the greatest common divisor of two polynomials is unique.

A polynomial  $f(x) \in \mathbb{F}[x]$  is called **irreducible** if it has degree  $\geq 1$  and can not be written as

$$f(x) = g(x) \cdot h(x)$$

for some  $g, h \in \mathbb{F}[x]$  and both  $g, h \notin \mathbb{F}$ . Elements of  $\mathbb{F}$  are called **constant polynomials**.

Let  $A$  be a UFD and  $k$  its field of fractions. We take  $a \in k$  such that  $a = \frac{r}{s}$ , where  $(r, s) = 1$ . For any prime element  $p \in A$ , we can write

$$a = p^m a'$$

where  $m$  is a integer and  $a' \in k$  such that  $p$  does not divide numerator or denominator of  $a'$ . The **order of  $a$  in  $p$**  is defined as  $m$ , say  $\text{ord}_p(a) = m$ . For  $f(x) \in \mathbb{F}[x]$  given as in Eq. (1) we define

$$\text{ord}_p(f) = \min \{ \text{ord}_p(a_i) \mid a_i \neq 0 \}.$$

The **content** of  $f(x)$ , which is denoted  $\text{cont}(f)$ , is defined as the product (up to multiplication to a unit in  $A$ )

$$(3) \quad \text{cont}(f) := \prod p^{\text{ord}_p(f)},$$

taking all  $p$  such that  $\text{ord}_p(f) \neq 0$ . If  $\text{cont}(f) = 1$ , then  $f(x)$  is called a **primitive polynomial**. Thus, every polynomial  $f(x) \in \mathbb{F}[x]$  can be written as

$$f(x) = \text{cont}(f) \cdot f_1(x),$$

where  $f_1(x)$  is primitive and  $f_1(x) \in A[x]$ . Notice that if  $f \in A[x]$  then  $\text{cont}(f)$  is simply

$$\text{cont}(f) = \gcd(a_0, \dots, a_n).$$

The **height** of  $f(x)$  is defined as

$$\mathfrak{h}(f) := \max \{ \text{ord}_p(a_i) \mid a_i \neq 0 \}$$

The following result is known as Gauss' lemma.

**Lemma 3** (Gauss Lemma). *Let  $A$  be a UFD,  $k$  its field of fractions and  $f, g \in \mathbb{F}[x]$ . Then,*

$$\text{cont}(fg) = \text{cont}(f) \cdot \text{cont}(g)$$

*Moreover, for  $f, g \in A[x]$ ,  $fg$  is primitive if and only if  $f$  and  $g$  are both primitive.*

**2.2. Several variables.** A polynomial with  $n$  variables is denoted by

$$f(x_1, \dots, x_n) = \sum_{i=(i_1, \dots, i_n) \in I} a_i x_1^{i_1} \cdots x_n^{i_n}$$

where all  $a_i \in K$ ,  $I \subset \mathbb{Z}^{\geq 0}$ , and  $I$  is finite. We use lexicographic ordering to order the terms in a given polynomial, and let

$$x_1 > x_2 > \cdots > x_n.$$

While the primary goal of this paper are polynomials with one variable, we will use polynomials with several variables when we discuss invariants of binary forms.

**2.3. Weighted polynomials.** Given any integer  $n \geq 1$ , let  $\mathbf{w} = (q_0, \dots, q_n)$  be a vector of positive integers. Consider the polynomial ring  $R = k_{\mathbf{w}}[x_0, \dots, x_n]$  where  $x_i$  has weight  $q_i$  for  $i = 0, 1, \dots, n$ .

Every polynomial is a sum of monomials  $x^d = \prod x_i^{d_i}$  with weight  $\sum_{i=1}^n q_i d_i$ . For every  $\lambda \in k^*$  and any weighted homogeneous polynomial  $f$  of degree  $d$ , we have

$$f(\lambda^{q_0} x_0, \lambda^{q_1} x_1, \dots, \lambda^{q_n} x_n) = \lambda^d f(x_0, \dots, x_n).$$

A degree  $d$  binary weighted form, where  $w = (q_0, q_1)$  be respectively the weights of  $x_0$  and  $x_1$ , is given by

$$f(x_0, x_1) = \sum_{d_0, d_1} a_{d_0, d_1} x_0^{d_0} x_1^{d_1}, \quad \text{such that } d_0 q_0 + d_1 q_1 = d$$

and in decreasing powers of  $x_0$  we have

$$f(x_0, x_1) = a_{d/q_0, 0} x_0^{d/q_0} + \cdots + a_{d_0, d_1} x_0^{d_0} x_1^{d_1} + \cdots + a_{0, d/q_1} x_1^{d/q_1}$$

By dividing with  $x_1^{d/q_1}$  and making a change of coordinates  $X = x_0^{q_1}/x_1^{q_0}$  we get

$$(4) \quad f(x_0, x_1) = a_{d/q_0, 0} X^{d/q_0 q_1} + \cdots + a_{d_0, d_1} X^{d_0/q_1} + \cdots + a_{0, d/q_1} = f(X)$$

Notice that the condition  $f(P) = 0$  is well defined on  $\mathbb{P}_{\mathbf{w}, k}^n$ .

### 3. EQUIVALENCES OF POLYNOMIALS

Two polynomial  $f(x)$  and  $g(x)$  are called **equivalent** if there is a nonzero scalar  $\lambda$  such that  $f(x) = \lambda g(x)$ . Hence,  $f(x)$  (up to multiplication by a scalar) can be conveniently thought as a point  $[a_0 : \cdots : a_n]$  in the projective space  $\mathbb{P}^n$ .

Since we want to identify polynomials up to multiplication by a non-zero constant it is convenient sometimes to think of them in their projective form.

**3.1. Binary forms.** Let  $k[x, y]$  be the polynomial ring in two variables and  $V_d$  denote the  $(d+1)$ -dimensional subspace of  $k[x, y]$  consisting of homogeneous polynomials

$$(5) \quad f(x, y) = a_d x^d + a_{d-1} x^{d-1} y + \cdots + a_0 y^d$$

of degree  $d$ . Elements of  $V_d$  are called **binary forms** of degree  $d$ .

To every polynomial  $f(x)$  we associate a binary form  $f(x, y) = y^n f\left(\frac{x}{y}\right)$  as above, which is called the *homogenization* of  $f(x)$ . Conversely, every binary form  $f(x, y)$  can be associated to a polynomial  $f(x, 1)$ , called the *dehomogenization* of  $f(x, y)$ .

Notice that any polynomial  $f \in \mathbb{Q}[x]$  can be written as  $f = \lambda g(x)$  for some  $g \in \mathbb{Z}[x]$ . Since  $f(x)$  and  $g(x) = \lambda f(x)$  have the same Galois group over  $\mathbb{Q}$ , it is enough to consider only polynomials in  $\mathbb{Z}[x]$ .

Let  $\text{GL}_2(\mathbb{Z})$  be the subgroup of  $\text{GL}_2(\mathbb{Q})$  such that matrices have integer entries. Hence every matrix  $M \in \text{GL}_2(\mathbb{Z})$  has determinant  $\det M = \pm 1$  and entries in  $\mathbb{Z}$ .

Two polynomials  $f, g \in \mathbb{Z}[x]$  of degree  $n$  are called  **$\mathbb{Z}$ -equivalent** if  $f(x) = a^n g(ax+b)$  for some  $a = \pm 1$  and  $b \in \mathbb{Z}$ .

Two degree  $n$  binary forms  $f, g \in \mathbb{Z}[x, y]$  are called  **$\text{GL}_2(\mathbb{Z})$ -equivariant** if  $g(x, y) = \pm f(ax+by, cx+dy)$  for some  $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \text{GL}_2(\mathbb{Z})$ . Two degree  $n$  polynomials  $f, g \in \mathbb{Z}[x]$  are called  **$\text{GL}_2(\mathbb{Z})$ -equivalent** if their homogenizations are  $\text{GL}_2(\mathbb{Z})$ -equivalent, in other words if

$$g(x) = \pm (cx+d)^n f\left(\frac{ax+d}{cx+d}\right), \quad \text{for some} \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \text{GL}_2(\mathbb{Z}).$$

$f, g \in \mathbb{Q}[x]$  are called  **$\mathbb{Q}$ -equivalent** if  $f(x) = g\left(\frac{ax+b}{cx+d}\right)$  for  $a, b, c, d \in \mathbb{Q}$ .

**Lemma 4.** *Let  $f, g \in \mathbb{Z}[x]$ . If  $f, g$  are  $\mathbb{Z}$ -equivalent, then they are  $\text{GL}_2(\mathbb{Z})$ -equivalent and their homogenizations are  $\text{GL}_2(\mathbb{Q})$ -equivalent.*

Hence the  $\text{GL}_2(\mathbb{Q})$  orbit, is partitioned into  $\text{GL}_2(\mathbb{Z})$ -orbits and each  $\text{GL}_2(\mathbb{Z})$ -orbit into  $\mathbb{Z}$ -orbits.

**3.2. Tschirnhaus-equivalent.**  $f$  and  $g$  (monic separable and irreducible of the same degree) are Tschirnhaus-equivalent iff they have the same splitting field  $E$  and moreover, if we let  $P$  and  $Q$  be the subgroups of  $G := \text{Gal}(E/k)$  fixing a root of  $f$  and  $g$  respectively, then  $P$  and  $Q$  are conjugate in  $G$ .

**3.3. Hermite equivalence.** Let  $f(x) \in \mathbb{Z}[x]$  given as in Eq. (1) and  $\alpha_1, \dots, \alpha_d \in \mathbb{C}$  its roots. Hence

$$f(x) = \sum_{i=0}^d a_i x^i = a_d \prod_{i=0}^d (x - \alpha_i)$$

To every root  $\alpha_i$  we associate a linear form in new variables  $x_1, \dots, x_d$  via

$$\alpha_i \rightarrow \alpha_i^{d-1} x_1 + \alpha_i^{d-2} x_2 + \dots + \alpha_i x_{d-1} + x_d$$

Then we associate to  $f$  the  $d$ -ary form

$$f \longrightarrow a_d^{d-1} \prod_{i=1}^d (\alpha_i^{d-1} x_1 + \alpha_i^{d-2} x_2 + \dots + \alpha_i x_{d-1} + x_d) =: [f]$$

The  $d$ -ary form  $[f]$  is called the **Hermite form associated to  $f$** . It is easy to show that the Hermite form is given by the resultant with respect to  $x$  of  $f(x)$  and  $g(x) = x_1 x^{d-1} + x_2 x^{d-2} + \dots + x_{d-1} x + x_d$ , namely

$$[f] = \text{Res}(f, g, x)$$

Hence,  $[f](x_1, \dots, x_d)$  is a  $d$ -ary form with integer coefficients. Moreover

$$\text{cont}([f]) = (\text{cont}(f))^{d-1}$$

Two polynomials  $f, g \in \mathbb{Z}[x]$  of degree  $n$  are called **Hermite equivalent** if their corresponding Hermite forms are  $\text{GL}_n(\mathbb{Z})$ -equivalent.

The discriminant of a decomposable  $d$ -ary form

$$F(x_1, \dots, x_d) = \prod_{i=1}^d (\alpha_{i,1} x_1 + \dots + \alpha_{i,d} x_d)$$

is defined as

$$\Delta(F) = \left( \det (\alpha_{i,j})_{i,j=1,\dots,d} \right)^2$$

Here are some properties of Hermitian forms. For proofs one can check [4].

**Lemma 5.** *The following are true:*

- (i) *The discriminant of any polynomials is the same as the discriminant of its Hermite form. In other words,  $\Delta([f]) = \Delta_f$ .*
- (ii) *Two polynomials which are Hermite equivalent have the same discriminants.*
- (iii) *Let  $f, g \in \mathbb{Z}[x]$  be  $\text{GL}_2(\mathbb{Z})$ -equivalent polynomials. Then  $f$  and  $g$  are Hermite equivalent. Moreover, if  $f$  and  $g$  are monic and  $\mathbb{Z}$ -equivalent then they are Hermite equivalent.*
- (iv) *(Hermite) There are finitely many Hermite equivalence classes of polynomials in  $\mathbb{Z}[x]$  of a given degree and given discriminant  $\Delta \neq 0$ .*

**3.4. Julia equivalence.** Hermite defined the equivalence class of polynomials to develop a reduction theory for degree  $d > 2$  polynomials. A reduction theory that was developed further by Julia; see [17] and [36] for more recent treatments.

Let  $f(x, y) \in \mathbb{Z}[x, y]$  be a degree  $n$  binary form

$$f(x, y) = a_0 x^n + a_1 x^{n-1} y + \dots + a_n y^n$$

and suppose that  $a_0 \neq 0$ . Let the real roots of  $f(x, y)$  be  $\alpha_i$ , for  $1 \leq i \leq r$  and the pair of complex roots  $\beta_j, \bar{\beta}_j$  for  $1 \leq j \leq s$ , where  $r + 2s = n$ . The form can be factored as

$$(6) \quad f(x, 1) = \prod_{i=1}^r (x - \alpha_i) \cdot \prod_{i=1}^s (x - \beta_i)(x - \bar{\beta}_i).$$

The ordered pair  $(r, s)$  of numbers  $r$  and  $s$  is called the **signature** of the form  $f$ .

We associate to  $f$  the two quadratics  $T_r(x, 1)$  and  $S_s(x, 1)$  of degree  $r$  and  $s$  respectively given by the formulas

$$(7) \quad T_r(x, 1) = \sum_{i=1}^r t_i^2 (x - \alpha_i)^2, \quad \text{and} \quad S_s(x, 1) = \sum_{j=1}^s 2u_j^2 (x - \beta_j)(x - \bar{\beta}_j),$$

where  $t_i, u_j$  are to be determined. For a binary form  $f$  of signature  $(r, s)$  the quadratic  $Q_f$  is defined as

$$(8) \quad \boxed{Q_f(x, 1) = T_r(x, 1) + S_s(x, 1)}.$$

Let  $\beta_i = a_i + b_i \cdot I$ , for  $i = 1, \dots, s$ .

The discriminant of  $Q_f$  is a degree 4 homogenous polynomial in  $t_1, \dots, t_r, u_1, \dots, u_s$ . We pick values for  $t_1, \dots, t_r, u_1, \dots, u_s$  such that this discriminant is square free and minimal. Then we can use the reduction theory of quadratics (with square free, minimal discriminant) to determine the reduced form for  $Q_f$ . Define

$$(9) \quad \theta_T = \frac{a_0^2 \cdot \Delta_T}{t_1^2 \cdots t_r^2}, \quad \theta_S = \frac{a_0^2 \cdot \Delta_S}{u_1^4 \cdots u_s^4}$$

**Proposition 1.** *Let  $f \in V_{n, \mathbb{Q}}$  with signature  $(r, s)$  and equation as in Eq. (6). Then  $Q_f$  is a positive definite quadratic form with discriminant  $\mathfrak{D}_f$  given by the formula*

$$(10) \quad \mathfrak{D}_f = \Delta(T_r) + \Delta(S_s) - 8 \sum_{i,j} t_i^2 u_j^2 ((\alpha_i - a_j)^2 + b_j^2).$$

From the above formula it can be seen that  $\mathfrak{D}_f$  is expressed in terms of the root differences. Hence,  $\mathfrak{D}_f$  is fixed by all the transpositions of the roots. However, it is not an invariant of the binary form. In order to get an invariant we need to fix it by all symmetries of the roots, hence by an element of order  $n$ . Indeed  $\mathfrak{D}_f^n$  is an invariant of the binary form  $f$  as we will see later. We define the  $\theta_0$  of a binary form as follows

$$(11) \quad \theta_0(f) = \frac{a_0^2 \cdot |\mathfrak{D}_f|^{n/2}}{\prod_{i=1}^r t_i^2 \prod_{j=1}^s u_j^4}.$$

Notice that in order for  $f$  to be in somewhat "simpler" or "minimal" form we would like the discriminant  $\mathfrak{D}_f$  to be minimal. Hence, we would like  $\theta_0(f)$  to be minimal. Consider  $\theta_0(t_1, \dots, t_r, u_1, \dots, u_s)$  as a multivariable function in the variables  $t_1, \dots, t_r, u_1, \dots, u_s$ . We would like to pick these variables such that  $Q_f$  is a reduced quadratic, hence  $\mathfrak{D}_f$  is minimal. This is equivalent to  $\theta_0(t_1, \dots, t_r, u_1, \dots, u_s)$  obtaining a minimal value.

**Proposition 2.** *The function  $\theta_0 : \mathbb{R}^{r+s} \rightarrow \mathbb{R}$  obtains a minimum at a unique point  $(\bar{t}_1, \dots, \bar{t}_r, \bar{u}_1, \dots, \bar{u}_s)$ .*

Choosing  $(\bar{t}_1, \dots, \bar{t}_r, \bar{u}_1, \dots, \bar{u}_s)$  that make  $\theta_0$  minimal gives a unique positive definite quadratic  $Q_f(x, z)$ . We call this unique quadratic  $Q_f(x, z)$  for such a choice of  $(\bar{t}_1, \dots, \bar{t}_r, \bar{u}_1, \dots, \bar{u}_s)$  the **Julia quadratic** of  $f(x, z)$ , denote it by  $\mathcal{J}_f(x, z)$ , and the quantity  $\theta_f := \theta_0(\bar{t}_1, \dots, \bar{t}_r, \bar{u}_1, \dots, \bar{u}_s)$  the **Julia invariant**.

**Lemma 6.** *Consider  $\text{SL}_2(\mathbb{Q})$  acting on  $V_{n, \mathbb{Q}}$ . Then  $\theta$  is an  $\text{SL}_2(\mathbb{Q})$ -invariant and  $\mathcal{J}$  is an  $\text{SL}_2(\mathbb{Q})$ -covariant of order 2.*

Performing Julia reduction symbolically is very difficult, but a machine learning approach is used in [20] to perform Julia reduction to higher degree polynomials. Hence, our database will have irreducible polynomials  $f(x) \in \mathbb{Q}[x]$  (up to the above equivalence) which are represented as polynomials in  $\mathbb{Z}[x]$ . There are two main issues here:

- i) identifying  $\mathbb{Q}$ -equivalence classes of polynomials,
- ii) determining a method of listing and ordering such polynomials.

The first issue can be addressed via the classical invariant theory of binary forms, which motivates the material for the rest of this section. The second issue can be addressed via heights of polynomials which is the focus of next section.

## 4. HEIGHTS OF POLYNOMIALS

Let  $K$  be a number field,  $\mathcal{O}_K$  its ring of integers, and  $M_K$  the set of absolute values of  $K$ . The **(affine) multiplicative height of a polynomial**  $f(x)$  is defined as

$$H_K^{\mathbb{A}}(f) = \prod_{v \in M_K} \max \left\{ 1, |f|_v^{n_v} \right\},$$

where

$$|f|_v := \max_j \left\{ |a_j|_v \right\}$$

is the **Gauss norm** for any  $v \in M_K$ . The **(affine) logarithmic height of  $f$**  is defined to be

$$h_K^{\mathbb{A}}(f) = h_K([1, \dots, a_j, \dots]_{j \in I}).$$

Hence, the affine height of a polynomial is defined to be the height of its coefficients taken as affine coordinates. The affine height sometimes is called the **naive height**.

The **(projective) multiplicative height** is

$$(12) \quad H_K(f) = \prod_{v \in M_K} |f|_v^{n_v}$$

where  $n_v$  is the completion of  $K_v$ ; see [6] among other sources. The **(projective) absolute multiplicative height** is defined as

$$H : \mathbb{P}^n(\mathbb{Q}) \rightarrow [1, \infty) \\ H(f) = H_K(f)^{1/[K:\mathbb{Q}]},$$

**Example 1.** Let  $f(x) \in \mathbb{Z}[x]$  and assume that  $f(x)$  is primitive. Then the projective height of  $f(x)$  is simply the maximum of the absolute values of its coefficients.

It is a consequence of Northcott's theorem that:

**Lemma 7.** *There are only finitely many polynomials of bonded height. In particular, for any polynomial  $f(x_1, \dots, x_n) \in K[x_1, \dots, x_n]$  there are only finitely many polynomials  $g(x_1, \dots, x_n) \in K[x_1, \dots, x_n]$  such that  $H_K(g) \leq H_K(f)$ .*

Let  $f(x_0, \dots, x_n)$  and  $g(y_0, \dots, y_n)$  be polynomials in different variables. Then, the projective height has the following property

$$H(f \cdot g) = H(f) \cdot H(g)$$

Before considering the height of polynomials in the same variables, we will consider  $|f \cdot g|_v$ . The following lemma is true for the product of a finite number of polynomials.

**Lemma 8** (Gauss's lemma). *Let  $K$  be a number field and  $f, g \in K[x_1, \dots, x_n]$ . If  $v$  is not Archimedean, then  $|fg|_v = |f|_v |g|_v$ .*

The proof can be found in [6, pg. 22]. An analogous Archimedean estimate is given by the following lemma. Gauss's lemma and the following are used to give an estimate of  $H(f_1 f_2 \cdots f_r)$  in terms of  $H(f_i)$  for  $1 \leq i \leq r$  and  $f_1, f_2, \dots, f_r \in K[x_1, \dots, x_n]$ .

**Lemma 9.** *Let  $f_1, \dots, f_r \in \mathbb{C}[x_1, \dots, x_n]$ ,  $f = f_1 \cdots f_r$ , and  $d_i = \deg(f, x_i)$ . Then,*

$$(13) \quad \prod_{i=1}^r |f_i|_v \leq e^{(d_1 + \cdots + d_n)} |f|_v.$$



The proof of this can be found in [14, pg. 232] and uses the concept of Mahler measure which is defined as follows. Let  $f(x_1, \dots, x_n) \in \mathbb{C}[x_1, \dots, x_n]$ . The **Mahler measure** is

$$M(f) := \exp \left( \int_{\mathbb{T}^n} \log |f(e^{i\theta_1}, \dots, e^{i\theta_n})| d\mu_1 \cdots d\mu_n \right)$$

where  $\mathbb{T}$  is the unit circle  $\{e^{i\theta} | 0 \leq \theta \leq 2\pi\}$  equipped with the standard measure  $d\mu = \frac{1}{2\pi} d\theta$ . Then

$$M(fg) = M(f) \cdot M(g),$$

see [14, pg. 230] for proof.

**Lemma 10.** *Let  $K$  be a number field and  $f_1, \dots, f_r \in K[x_1, \dots, x_n]$ . Denote with  $\deg f_j$  the total degree of  $f_j$ . Then the following are true*

- (i)  $H^{\mathbb{A}}(f_1 f_2 \cdots f_r) \leq N \cdot \prod_{j=1}^r H^{\mathbb{A}}(f_j) \leq r \cdot \max_{1 \leq j \leq r} \{h(f_j) + (\deg f_j + m) \log 2\}$ .
- (ii)  $H^{\mathbb{A}}(f_1 + f_2 + \cdots + f_r) \leq r \cdot \prod_{j=1}^r H^{\mathbb{A}}(f_j)$ .
- (iii) *If  $f_1, \dots, f_r \in \mathcal{O}_K[x_1, \dots, x_n]$ , then*

$$H^{\mathbb{A}}(f_1 + f_2 + \cdots + f_r) \leq r \cdot \max_j \left\{ H^{\mathbb{A}}(f_j) \right\}^{[K:\mathbb{Q}]}$$

The converse of part (i) is known as Gelfand's inequality.

**Lemma 11** (Gelfand's inequality). *Let  $f_1, \dots, f_r \in \overline{\mathbb{Q}}[x_1, \dots, x_n]$ ,  $d_i = \deg f_i$  such that  $\deg(f_1 \cdots f_r, x_i) \leq d_i$  for each  $1 \leq i \leq r$ . Then*

$$\prod_{i=1}^r H(f_i) \leq e^{(d_i + \cdots + d_n)} \cdot H(f_1 \cdots f_r).$$

## 5. BINARY FORMS

$\mathrm{GL}_2(k)$  acts as a natural group of automorphisms on  $k[x, y]$ . Denote by  $f \rightarrow f^M$  this action. It is well known that  $\mathrm{SL}_2(k)$  leaves a bilinear form (unique up to scalar multiples) on  $V_d$  invariant. If  $k$  is algebraically closed then  $f(x, y)$  can be factored as

$$(14) \quad f(x, y) = (\beta_1 x - \alpha_1 y) \cdots (\beta_d x - \alpha_d y) = \prod_{1 \leq i \leq d} \det \begin{pmatrix} x & \alpha_i \\ y & \beta_i \end{pmatrix}$$

Points with homogeneous coordinates  $(\alpha_i, \beta_i) \in \mathbb{P}^1$  are called the **projective roots** of  $f$ . For  $M \in \mathrm{GL}_2(k)$  we have

$$f^M(x, y) = (\det M)^d (\beta'_1 x - \alpha'_1 y) \cdots (\beta'_d x - \alpha'_d y), \quad \text{where} \quad \begin{pmatrix} \alpha'_i \\ \beta'_i \end{pmatrix} = M^{-1} \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix}.$$

Consider  $a_0, a_1, \dots, a_d$  as transcendentals over  $k$  (coordinate functions on  $V_d$ ). Then the coordinate ring of  $V_d$  can be identified with  $k[a_0, \dots, a_d]$ . There is an action of  $\mathrm{GL}_2(k)$  on  $k[a_0, \dots, a_d]$  via

$$\begin{aligned} \mathrm{GL}_2(k) \times k[a_0, \dots, a_d] &\rightarrow k[a_0, \dots, a_d] \\ (M, F) &\rightarrow F^M := F(f^M), \quad \text{for all } f \in V_d. \end{aligned}$$

Thus for a polynomial  $F \in k[a_0, \dots, a_d]$  and  $M \in \mathrm{GL}_2(k)$ , define  $F^M \in k[a_0, \dots, a_d]$  as  $F^M(f) := F(f^M)$ , for all  $f \in V_d$ . Then  $F^{MN} = (F^M)^N$ . The homogeneous degree in  $a_0, \dots, a_d$  is called the **degree** of  $F$ , and the homogeneous degree in  $x, y$  is called the **order** of  $F$ . An **invariant** is usually referred to an  $\mathrm{SL}_2(k)$ -invariant on  $V_d$ . Hilbert's theorem says that the ring of invariants  $\mathcal{R}_d$  is finitely generated. Thus,  $\mathcal{R}_d$  is a finitely generated graded ring.

Let  $\xi_0, \dots, \xi_n$  be a minimal set of generators of  $\mathcal{R}_d$  and  $\deg \xi_i = q_i$ . The set of degrees  $(q_0, \dots, q_n)$  is often called the **set of weights**.

**Lemma 12.** *Let  $f, g \in V_d, M \in \mathrm{GL}_2(k), \lambda = (\det M)^{\frac{d}{2}}$ . Then  $f = g^M$  if and only if*

$$(\xi_0(f), \dots, \xi_i(f), \dots, \xi_n(f)) = (\lambda^{q_0} \xi_0(g), \dots, \lambda^{q_i} \xi_i(g), \dots, \lambda^{q_n} \xi_n(g)).$$

*If  $k = \mathbb{Q}$  we can choose  $\xi_0, \dots, \xi_n \in \mathbb{Z}[a_0, \dots, a_d]$  and primitive.*

The theory of binary forms is quite extensive and well understood. However, the main goal of this paper is to construct a database of irreducible polynomials  $f \in \mathbb{Q}[x]$  so we can study their Galois groups. Hence, we have to consider some other equivalences of polynomials.

**5.1. Proj  $\mathcal{R}_d$  as a weighted projective space.** Let  $\xi_0, \dots, \xi_n$  be the generators of  $\mathcal{R}_d$  with degrees  $q_0, \dots, q_n$  respectively. Since all  $\xi_0, \dots, \xi_i, \dots, \xi_n$  are homogenous polynomials then  $\mathcal{R}_d$  is a graded ring and Proj  $\mathcal{R}_d$  as a weighted projective space.

Let  $\mathbf{w} := (q_0, \dots, q_n) \in \mathbb{Z}^{n+1}$  be a fixed tuple of positive integers called **weights**. Consider the action of  $k^* = k \setminus \{0\}$  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  $\mathbb{WP}_{(q_0, \dots, q_n)}^n(k)$ . It is the projective variety  $\mathrm{Proj}(k[x_0, \dots, x_n])$  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 denote greatest common divisor of  $q_0, \dots, q_n$  by  $\gcd(q_0, \dots, q_n)$ . The space  $\mathbb{WP}_w^n$  is called **well-formed** if

$$\gcd(q_0, \dots, \hat{q}_i, \dots, q_n) = 1, \quad \text{for each } i = 0, \dots, n.$$

We denote a point  $\mathbf{p} \in \mathbb{WP}_w^n(k)$  by  $\mathbf{p} = [x_0 : x_1 : \dots : x_n]$ .

Let  $\xi_0, \xi_1, \dots, \xi_n$  be the generators of the ring of invariants  $\mathcal{R}_d$  of degree  $d$  binary forms. A  $k$ -isomorphism class of a binary form  $f$  is determined by the point

$$\xi(f) := [\xi_0(f), \xi_1(f), \dots, \xi_n(f)] \in \mathbb{WP}_{\mathbf{w}}^n(k).$$

Moreover, for any two forms  $f$ , and  $g$  we have that  $f = g^M$  for some  $M \in \mathrm{GL}_2(k)$  if and only if  $\xi(f) = \lambda \star \xi(g)$ , for  $\lambda = (\det A)^{\frac{d}{2}}$ .

**5.2. Generators of the ring of invariants.** Finding generators for the ring of invariants  $\mathcal{R}_d$  is a classical problem of the XIX-century. Such generators are obtained in terms of transvections or root differences. Below we list the generating set of  $\mathcal{R}_d$  for  $d \leq 10$ . From here on

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

For given binary invariants  $f, g \in V_d$  the  $r$ -th transvection of  $f$  and  $g$  is denoted by  $(f, g)_r$ .

While there is no method known to determine a generating set of invariants for any  $\mathcal{R}_d$ , we display a minimal generating set for all  $3 \leq d \leq 10$ . For the rest of this section  $f(x, y)$  is given as in Eq. (5) and a minimal set of invariants is always picked as in lemma 12.

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

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

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

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

**5.2.3. Quintics.** A generating set for  $\mathcal{R}_5$  is  $\xi = [\xi_0, \xi_1, \xi_2]$  with  $\mathbf{w} = (4, 8, 12)$ , where

$$\xi_0 = (c_1, c_1)_2, \quad \xi_1 = (c_4, c_1)_2, \quad \xi_2 = (c_4, c_4)_2,$$

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

5.2.4. *Sextics*. The case of sextics was studied in detail by XIX-century mathematicians (Bolza, Clebsch, et al.) when  $\text{char } k = 0$  and by Igusa for  $\text{char } k > 0$ . 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  $\mathbf{w} = (2, 4, 6, 10)$ , where

$$\xi_0 = (f, f)_6, \quad \xi_1 = (c_1, c_1)_4, \quad \xi_2 = (c_4, c_1)_4, \quad \xi_3 = (c_4, c_3^2)_4$$

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.

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

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

5.2.6. *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  $\mathbf{w} = (2, 3, 4, 5, 6, 7)$ . We define them as follows. Let

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

Then the invariants are:

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

5.2.7. *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  $\mathbf{w} = (4, 8, 10, 12, 12, 14, 16)$ . Let

$$\begin{aligned} c_1 &= (f, f)_8, & c_2 &= (f, f)_6, & c_4 &= (f, f)_2, & c_5 &= (f, c_1)_2, & c_6 &= (f, c_2)_6, \\ c_7 &= (c_2, c_2)_4, & c_9 &= (c_5, c_5)_4, & c_{21} &= (f, c_2)_2, & c_{25} &= (c_4, c_4)_{10}, & c_{27} &= (c_6^3, c_6)_3 \\ \xi_0 &= (c_1, c_1)_2, & \xi_1 &= (c_2, c_6^2)_6, & \xi_2 &= (((c_{25}, f)_6, c_{21})_5, c_2)_6, \\ \xi_3 &= ((c_7, c_7)_2, c_7)_4, & \xi_4 &= (c_9, c_1^3)_6, & \xi_5 &= ((c_2, c_{27})_3)_6, \\ \xi_6 &= ((c_5, c_5)_2, c_1^5)_{10}. \end{aligned}$$

5.2.8. *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  $\mathbf{w} = (2, 4, 6, 6, 8, 9, 10, 14, 14)$ . Let

$$\begin{aligned} c_1 &= (f, f)_8, & c_2 &= (f, f)_6, & c_5 &= (f, c_1)_4, & c_6 &= (f, c_2)_8, \\ c_7 &= (c_2, c_2)_6, & c_8 &= (c_5, c_5)_4, & c_9 &= (c_2, c_7)_4, & c_{10} &= (c_1, c_1)_2, \\ c_{16} &= (c_5, c_5)_2, & c_{19} &= (c_5, c_1)_1, & c_{25} &= (c_7, c_7)_2 \\ \xi_0 &= (f, f)_{10}, & \xi_1 &= (c_1, c_1)_4, & \xi_2 &= (c_5, c_5)_6, \\ \xi_3 &= (c_6, c_6)_2, & \xi_4 &= (c_1, c_8)_4, & \xi_5 &= (c_{19}, c_1^2)_8, \\ \xi_6 &= (c_{16}, c_1^2)_8, & \xi_7 &= (c_{25}, c_9)_4, & \xi_8 &= (c_{10}^2, c_{16})_8. \end{aligned}$$

**5.3. Root differences.** Invariants can also be expressed in terms of root differences. For example the discriminant is given by

$$\Delta(f) = \prod_{i \neq j} (\alpha_i - \alpha_j).$$

An excellent article on invariants including root differences is [22]. Multiplicities of the roots determine the stability of the binary forms via the Hilbert-Mumford criterion; see [11].

- (i) If  $f$  has a root of multiplicity  $r > \frac{d}{2}$  then  $\xi(f) = (\xi_0, \dots, \xi_n) = (0, \dots, 0)$ .
- (ii) If  $d$  is even, then all binary forms with a root of multiplicity  $\frac{d}{2}$  have the same invariants.

**5.4. Heights and moduli heights.** Next we focus on heights of binary forms and their invariants. Let  $K$  be a number field,  $f \in K[x_0, \dots, x_n]$  a homogenous polynomial of degree  $d$ . We define

$$|c(d, n)|_v := \begin{cases} \binom{n+d}{n} & \text{if } v \text{ is Archimedean} \\ 1 & \text{if } v \text{ is non-Archimedean} \end{cases}$$

**Lemma 13.** *Let  $K$  be a number field,  $f \in K[x_0, \dots, x_n]$  a homogenous polynomial of degree  $d$ , and  $\alpha = (\alpha_0, \dots, \alpha_n) \in \overline{K}^{n+1}$ . Then,*

$$|f(\alpha)|_v \leq |c(d, n)|_v \cdot \max_j \left\{ |\alpha_j|_v \right\}^d \cdot |f|_v,$$

where  $|c(d, n)|_v$  is  $\binom{n+d}{n}$  if  $v$  is non-Archimedean and 1 otherwise. Moreover,

$$H(f(\alpha)) \leq c_0 \cdot H(\alpha)^d \cdot H(f).$$

Lemma 13 can be used to determine the height of invariants of binary forms.

**Corollary 2.** *Let  $f \in K[x, y]$  as in Eq. (5) and  $\alpha = (\alpha_0, \alpha_1) \in \overline{K}^2$ . Then,*

$$H(f(\alpha)) \leq \min \{d+1, 2^{d+1}\} \cdot H(\alpha)^d \cdot H(f).$$

**5.5. Minimal and moduli heights of forms.** Let  $f(x, y)$  be a binary form and  $Orb(f)$  its  $GL_2(K)$ -orbit in  $V_d$ . As a consequence of Northcott's theorem, there are only finitely many  $f' \in Orb(f)$  such that  $H(f') \leq H(f)$ . Define the height of the binary form  $f(x, y)$  as follows

$$\tilde{H}(f) := \min \left\{ H(f') \mid f' \in Orb(f), H(f') \leq H(f) \right\}$$

we want to consider the following problem. For every  $f$  let  $f'$  be the binary form such that  $f' \in Orb(f)$  and  $\tilde{H}(f) = H(f')$ . Determine a matrix  $M \in GL_2(K)$  such that  $f' = f^M$ .

Let  $\mathcal{B}_d$  be the moduli space of degree  $d$  binary forms defined over an algebraically closed field  $k$ . Then  $\mathcal{B}_d$  is a quasi-projective variety with dimension  $d-3$ . We denote the equivalence class of  $f$  by  $\mathfrak{f} \in \mathcal{B}_d$ . The **moduli height** of  $f(x, z)$  is defined as

$$\mathfrak{H}(f) = H(\mathfrak{f})$$

where  $\mathfrak{f}$  is considered as a point in the projective space  $\mathbb{P}^{d-3}$ . A natural question would be to investigate if the minimal height  $\tilde{H}(f)$  has any relation to the moduli height  $\mathfrak{H}(f)$ .

Let  $\{I_{i,j}\}_{j=1}^{j=s}$  be a basis of  $\mathcal{R}_d$ . Here the subscript  $i$  denotes the degree of the homogenous polynomial  $I_{i,j}$ . The fixed field of invariants is the space  $V_d^{GL_2(K)}$  and is generated by rational functions  $t_1, \dots, t_r$  where each of them is a ratio of polynomials in  $I_{i,j}$  such that the combined degree of the numerator is the same as that of the denominator.

**Theorem 5.1** ([39]). *Let  $f$  be a binary form. Then, For any  $SL_2(k)$ -invariant  $I_i$  of degree  $i$  we have that*

$$H(I_i(f)) \leq c \cdot H(f)^d \cdot H(I_i)$$

Moreover,  $\mathfrak{H}(f) \leq c \cdot \tilde{H}(f)$ , for some constant  $c$ .

For a given degree  $d$  the constant  $c$  of the theorem can be explicitly computed. For binary sextics this constant is  $c = 2^{28} \cdot 3^9 \cdot 5^5 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \cdot 43$ ; see [39].

**5.6. Weighted moduli height.** For any point  $\mathbf{p} = [x_0 : \cdots : x_n] \in \mathbb{P}_{\mathbf{w},k}^n$  we can assume, without loss of generality, that  $\mathbf{p} = [x_0 : \cdots : x_n] \in \mathbb{P}_{\mathbf{w},k}^n(\mathcal{O}_k)$ . Let  $\mathbf{w} = (q_0, \dots, q_n)$  be a set of weights and  $\mathbb{P}_{\mathbf{w},k}^n$  the weighted projective space over a number field  $k$ . Let  $\mathbf{p} \in \mathbb{P}_{\mathbf{w},k}^n$  a point such that  $\mathbf{p} = [x_0, \dots, x_n]$ . We define the **weighted multiplicative height** of  $\mathbf{p}$  as

$$(15) \quad \mathfrak{H}_k(\mathbf{p}) := \prod_{v \in M_k} \max \left\{ |x_0|_v^{\frac{n_v}{q_0}}, \dots, |x_n|_v^{\frac{n_v}{q_n}} \right\}.$$

The **absolute weighted height** of  $\mathbf{p} \in \mathbb{P}_{\mathbf{w},k}^n$  is the function  $\mathfrak{H} : \mathbb{P}_{\mathbf{w},\overline{\mathbb{Q}}}^n \rightarrow [1, \infty)$ ,

$$(16) \quad \mathfrak{H}(\mathbf{p}) = \mathfrak{H}_k(\mathbf{p})^{1/[k:\mathbb{Q}]},$$

where  $\mathbf{p} \in \mathbb{P}_{\mathbf{w},k}^n$ , for any  $k$  which contains  $\mathbb{Q}(\overline{\text{wgcd}}(\mathbf{p}))$ . The **absolute (logarithmic) weighted height** on  $\mathbb{P}_{\mathbf{w},\overline{\mathbb{Q}}}^n$  is the function  $\mathfrak{s} : \mathbb{P}_{\mathbf{w},\overline{\mathbb{Q}}}^n \rightarrow [0, \infty)$

$$\mathfrak{s}(\mathbf{p}) = \log \mathfrak{H}(\mathbf{p}) = \frac{1}{[k:\mathbb{Q}]} \log \mathfrak{H}_k(\mathbf{p}).$$

where again  $\mathbf{p} \in \mathbb{P}_{\mathbf{w},k}^n$ , for any  $k$  which contains  $\mathbb{Q}(\overline{\text{wgcd}}(\mathbf{p}))$ .

Let  $\mathbb{P}_{\mathbf{w},k}$  be a well-formed weighted projective space and  $\mathbf{x} = [x_0 : \cdots : x_n] \in \mathbb{P}_{\mathbf{w},k}(k)$ . Assume  $\mathbf{x}$  normalized (i.e.  $\text{wgcd}_k(\mathbf{x}) = 1$ ). Clearly  $\text{wgcd}_k(\mathbf{x}) \mid \text{gcd}(x_0, \dots, x_n)$  and therefore  $\text{wgcd}(\mathbf{x}) \leq \text{gcd}(x_0, \dots, x_n)$ . Let  $\mathbf{x}$  be absolutely normalized. Then  $\text{gcd}(x_0, \dots, x_n) = 1$ . If  $\mathbf{x} = [x_0 : \dots, x_n]$  is a normalized point then by definition of the height

$$\mathfrak{H}_k(\mathbf{x}) = \max_{i=0}^n \{ |x_i|^{\frac{1}{q_i}} \}$$

## 6. GALOIS GROUPS OF A POLYNOMIALS

Let  $\mathbb{F}$  be a perfect field. For simplicity we only consider the case when  $\text{char}\mathbb{F} = 0$ . Let  $f(x)$  be a degree  $n = \deg f$  irreducible polynomial in  $\mathbb{F}[x]$  which is factored as follows:

$$(17) \quad f(x) = (x - \alpha_1) \dots (x - \alpha_n)$$

in a splitting field  $E_f$ . Then,  $E_f/\mathbb{F}$  is Galois because it is a normal extension and separable. The group  $\text{Gal}(E_f/\mathbb{F})$  is called **the Galois group** of  $f(x)$  over  $\mathbb{F}$  and denoted by  $\text{Gal}_{\mathbb{F}}(f)$ . The elements of  $\text{Gal}_{\mathbb{F}}(f)$  permute roots of  $f(x)$ . Thus, the Galois group of polynomial has an isomorphic copy embedded in  $S_n$ , determined up to conjugacy by  $f$ . The main goal of this section is to determine  $\text{Gal}_{\mathbb{F}}(f)$ .

**Proposition 3.** *The following are true:*

- (i)  $\deg f \mid |G|$
- (ii) Let  $G = \text{Gal}_{\mathbb{F}}(f)$  and  $H = G \cap A_n$ . Then  $H = \text{Gal}(E_f/\mathbb{F}(\sqrt{\Delta_f}))$ . In particular,  $G$  is contained in the alternating group  $A_n$  if and only if the discriminant  $\Delta_f$  is a square in  $\mathbb{F}$ .
- (iii) The irreducible factors of  $f$  in  $\mathbb{F}[x]$  correspond to the orbits of  $G$ . In particular,  $G$  is a transitive subgroup of  $S_n$  if and only if  $f$  is irreducible.

*Proof.* The first part is a basic property of the splitting field  $E_f$ . (ii) We have  $\Delta_f = d_f^2$ , where  $d_f = \prod_{i>j}(\alpha_i - \alpha_j)$ . For  $g \in G$  we have  $g(d_f) = \text{sgn}(g)d_f$ . Thus  $H = G \cap A_n$  is the stabilizer of  $d_f$  in  $G$ . But this stabilizer equals  $\text{Gal}(E_f/\mathbb{F}(d_f))$ . Hence the claim.

(iii)  $G$  acts transitively on the roots of each irreducible factor of  $f$ . □

**Lemma 14.** *The following are true:*

- (1) If  $\sigma \in \text{Gal}(E_f/\mathbb{F})$  is a transposition then  $\sigma(\Delta_f) = -\Delta_f$ .
- (2) If  $\sigma \in \text{Gal}(E_f/\mathbb{F})$  is an even permutation then  $\sigma(\Delta_f) = \Delta_f$ .
- (3)  $\text{Gal}(E_f/\mathbb{F})$  is isomorphic to a subgroup of  $A_n$  if and only if  $\Delta_f \in \mathbb{F}$ .

When  $n = 2$  then  $f(x) = a_2x^2 + a_1x + a_0$ . Thus,  $\Delta_f = a_1^2 - 4a_0a_2$ . Hence  $\text{Gal}(f) \cong A_2 = \{1\}$  if and only if  $\Delta_f$  is a square.

**Lemma 15.** *Let  $f(x) \in \mathbb{F}[x]$  be an irreducible polynomial of degree  $\deg f = n$ . Then  $\text{Gal}_{\mathbb{F}}(f)$  is an affine invariant of  $f(x)$ . In other words,  $\text{Gal}(f) \cong \text{Gal}(g)$  for any  $g(x) = f(ax + b)$ , for  $a, b \in \mathbb{F}$  and  $a \neq 0$ .*

Let  $f(x, y) \in \mathbb{F}[x, y]$  be a binary form of degree  $\deg f = n$ . Let  $g(x) = f(x, 1)$ . Can  $\text{Gal}(g)$  be characterized in terms of invariants of the binary form  $f(x, y)$ ? From section 5.2 we know that invariants of binary forms do not change under linear substitutions. Also from lemma 15 is invariant under such substitutions. Hence, we must be able to determine  $\text{Gal}(g)$  in terms of invariants of  $f(x, y)$ . For the rest of this section we will see how this can be done explicitly for cubics, quartics, and quintics.

**6.1. Cubics.** Let  $f(x)$  be an irreducible cubic polynomial in  $\mathbb{F}[x]$ . From ?? we know that  $[E_f : \mathbb{F}] = 3$  or 6. Hence, the Galois group  $\text{Gal}_{\mathbb{F}}(f)$  is a subgroup of  $S_3$  with order 3 or 6. Thus,  $\text{Gal}_{\mathbb{F}}(f) \cong A_3$  if and only if  $\Delta_f$  is a square in  $\mathbb{F}$ , otherwise  $\text{Gal}_{\mathbb{F}}(f) \cong S_3$ .

**Lemma 16.** *Let  $f(x) \in \mathbb{F}[x]$  be an irreducible cubic. Then  $G = A_3$  if and only if  $\xi_0(f) = \Delta_f$  is a square in  $\mathbb{F}$ . Moreover, the following hold:*

- (i)  $\Delta_f > 0$  if and only if  $f$  has three distinct real roots.
- (ii)  $\Delta_f < 0$  iff  $f$  has one real root and two non-real complex conjugate roots.

Since both  $A_3$  and  $S_3$  are solvable, we should be able to determine formulas to give the roots of  $f(x)$  in terms of radicals. These formulas are known as Cardano's formulas and we will skip them here.

**Remark 1.** *What we notice from the cubics is that we can determine the Galois group simply by condition on invariants. We will see next if that can be done for higher degree polynomials.*

**6.2. Quartics.** Let  $f(x) \in \mathbb{F}[x]$  be an irreducible polynomial of degree 4. Then  $G := \text{Gal}(f)$  is a transitive subgroup of  $S_4$ . Further  $4 \mid |G|$ , see Prop. 3. So the order of  $G$  is 4, 8, 12, or 24. It can be easily checked that transitive subgroups of  $S_4$  of order 4, 8, 12, or 24 are isomorphic to one of the following groups

$$C_4, D_4, V_4, A_4, S_4.$$

Consider the normalized polynomial

$$(18) \quad f(x) = x^4 + ax^2 + bx + c = (x - \alpha_1) \dots (x - \alpha_4)$$

with  $a, b, c \in \mathbb{F}$ . Let  $E_f = \mathbb{F}(\alpha_1, \dots, \alpha_4)$  be the splitting field of  $f$  over  $\mathbb{F}$ . Since  $f$  has no  $x^3$ -term, we have  $\alpha_1 + \dots + \alpha_4 = 0$ . We assume  $\Delta_f \neq 0$ , so  $\alpha_1, \dots, \alpha_4$  are distinct. Let  $G = \text{Gal}_{\mathbb{F}}(f)$ , viewed as a subgroup of  $S_4$  via permuting  $\alpha_1, \dots, \alpha_4$ .

There are 3 partitions of  $\{1, \dots, 4\}$  into two pairs.  $S_4$  permutes these 3 partitions, with kernel

$$(19) \quad V_4 = \{(12)(34), (13)(24), (14)(23), id\}.$$

Thus  $S_4/V_4 \cong S_3$ , the full symmetric group on these 3 partitions. Associate with these partitions the elements

$$(20) \quad \beta_1 = \alpha_1\alpha_2 + \alpha_3\alpha_4, \quad \beta_2 = \alpha_1\alpha_3 + \alpha_2\alpha_4, \quad \beta_3 = \alpha_1\alpha_4 + \alpha_2\alpha_3$$

of  $E_f$ . If  $\beta_1 = \beta_2$  then  $\alpha_1(\alpha_2 - \alpha_3) = \alpha_4(\alpha_2 - \alpha_3)$ , a contradiction. Similarly,  $\beta_1, \beta_2, \beta_3$  are 3 distinct elements. Then  $G$  acts as a subgroup of  $S_4$  on  $\alpha_1, \dots, \alpha_4$ , and as the corresponding subgroup of  $S_3 \cong S_4/V_4$  on  $\beta_1, \dots, \beta_3$ . Thus the subgroup of  $G$  fixing all  $\beta_i$  is  $G \cap V_4$ . This proves the following:

$$\begin{array}{c} E_f := \mathbb{F}(\alpha_1, \alpha_2, \alpha_3, \alpha_4) \\ \left| \bar{G}=G \cap V_4 \right. \\ E := \mathbb{F}(\beta_1, \beta_2, \beta_3) \\ \left| d \right. \\ \mathbb{F} \end{array}$$

**Lemma 17.** *The subgroup  $G \cap V_4 \leq G$  corresponds to the subfield  $\mathbb{F}(\beta_1, \beta_2, \beta_3)$ , which is the splitting field over  $\mathbb{F}$  of the cubic polynomial (cubic resolvent)*

$$(21) \quad g(x) = (x - \beta_1)(x - \beta_2)(x - \beta_3) = x^3 - ax^2 - 4cx + -b^2 + 4ac$$

The roots  $\beta_i$  of the cubic resolvent can be found by Cardano's formulas. The extension

$$\mathbb{F}(\alpha_1, \dots, \alpha_4)/k(\beta_1, \beta_2, \beta_3)$$

has Galois group  $\leq V_4$ , hence is obtained by adjoining at most two square roots to  $\mathbb{F}(\beta_1, \beta_2, \beta_3)$ . Moreover,  $\Delta(f, x) = \Delta(g, x)$ .

In general, for an irreducible quartic

$$f(x) = x^4 + ax^3 + bx^2 + cx + d$$

we can first eliminate the coefficient of  $x^3$  by the substituting  $x$  with  $x - \frac{a}{4}$ . In terms of the binary forms this corresponds to the transformation

$$(x, y) \rightarrow \left(x - \frac{a}{4}y, y\right)$$

and the new quartic is  $f^M$  for  $M = \begin{bmatrix} 1 & -a/4 \\ 0 & 1 \end{bmatrix}$ . Since  $M \in \text{SL}_2(\mathbb{Q})$  then  $\det M = 1$  and the invariants of  $f^M$  are the same as those of  $f$ , namely

$$\begin{aligned}\xi_0(f) &= 2a_0a_4 - \frac{a_1a_3}{2} + \frac{a_2^2}{6} \\ \xi_1(f) &= a_0a_2a_4 - \frac{3a_0a_3^2}{8} - \frac{3a_1^2a_4}{8} + \frac{a_1a_2a_3}{8} - \frac{a_2^3}{36}\end{aligned}$$

Moreover  $g(x)$  is

$$(22) \quad g(x) := x^3 - bx^2 + (ac - 4d)x - a^2d + 4bd - c^2.$$

The discriminant of  $f(x)$  is the same as the discriminant of  $g(x)$  and is given below:

$$(23) \quad \begin{aligned}\Delta_f &= -27a^4d^2 + 18a^3bcd - 4a^3c^3 - 4a^2b^3d + a^2b^2c^2 + 144a^2bd^2 - 6a^2c^2d - 80ab^2cd \\ &\quad + 18abc^3 + 16b^4d - 4b^3c^2 - 192acd^2 - 128b^2d^2 + 144bc^2d - 27c^4 + 256d^3\end{aligned}$$

We denote by  $d := [\mathbb{F}(\beta_1, \beta_2, \beta_3) : \mathbb{F}]$ . Then we have the following:

**Lemma 18.** *The Galois group of  $f(x)$  is one of the following:*

- (i)  $d = 1 \iff G \cong V_4$ .
- (ii)  $d = 3 \iff G \cong A_4$ .
- (iii)  $d = 6 \iff G \cong S_4$ .
- (iv) *If  $d = 2$  then we have*
  - a)  $f(x)$  is irreducible over  $F \iff G \cong D_4$
  - b)  $f(x)$  is reducible over  $F \iff G \cong C_4$

6.2.1. *Solving quartics.* The element  $(\alpha_1 + \alpha_2)(\alpha_3 + \alpha_4)$  is fixed by  $G \cap V_4$ , hence lies in  $K(\beta_1, \beta_2, \beta_3)$ . We find

$$(24) \quad -(\alpha_1 + \alpha_2)^2 = (\alpha_1 + \alpha_2)(\alpha_3 + \alpha_4) = \beta_2 + \beta_3$$

By this and symmetry we get **Ferrari's formulas**

$$(25) \quad \begin{aligned}\alpha_1 + \alpha_2 &= \sqrt{-\beta_2 - \beta_3} \\ \alpha_1 + \alpha_3 &= \sqrt{-\beta_1 - \beta_3} \\ \alpha_1 + \alpha_4 &= \sqrt{-\beta_1 - \beta_2}\end{aligned}$$

or

$$(26) \quad \begin{aligned}\alpha_1 &= \frac{\sqrt{-\beta_1 - \beta_2} + \sqrt{-\beta_1 - \beta_3} + \sqrt{-\beta_2 - \beta_3}}{2} \\ \alpha_2 &= \frac{-\sqrt{-\beta_1 - \beta_2} - \sqrt{-\beta_1 - \beta_3} + \sqrt{-\beta_2 - \beta_3}}{2} \\ \alpha_3 &= \frac{-\sqrt{-\beta_1 - \beta_2} + \sqrt{-\beta_1 - \beta_3} - \sqrt{-\beta_2 - \beta_3}}{2} \\ \alpha_4 &= \frac{\sqrt{-\beta_1 - \beta_2} - \sqrt{-\beta_1 - \beta_3} - \sqrt{-\beta_2 - \beta_3}}{2}\end{aligned}$$

This completes the case for the quartics.



**6.3. Quintics.** Now we are ready to handle quintics which has such a special case in the history of Galois theory.

**Lemma 19.** *Let  $f(x) \in \mathbb{F}[x]$  be an irreducible quintic. Then its Galois group is one of the following  $C_5$ ,  $D_5$ ,  $F_5 = AGL(1, 5)$ ,  $A_5$ ,  $S_5$ .*

*Proof.*  $G$  is transitive, hence its 5-Sylow subgroup is isomorphic to  $C_5$  (generated by a 5-cycle). If  $C_5$  is not normal, then  $G$  has at least 6 of 5-Sylow subgroups; then  $|G| \geq 6 \cdot 5 = 30$ , hence  $[S_5 : G] \leq 4$  which implies  $G = S_5, A_5$ . If  $C_5$  is normal in  $G$  then  $G$  is conjugate either  $C_5$ ,  $D_5$  (dihedral group of order 10) or  $F_5 = AGL(1, 5)$ , the full normalizer of  $C_5$  in  $S_5$ , of order 20 (called also the Frobenius group of order 20).  $\square$

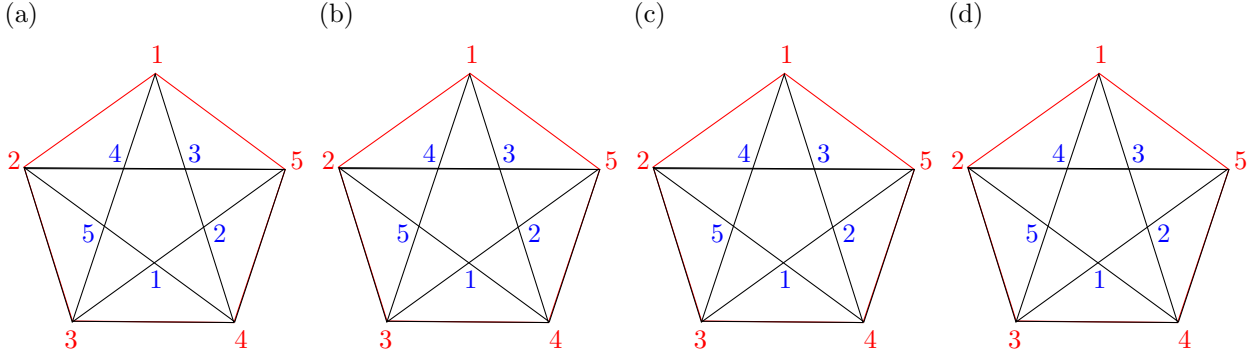
**Remark 2.** *If the discriminant of the quintic is a square in  $\mathbb{F}$  then  $\text{Gal}(f)$  is contained in  $A_5$ . Hence, it is  $C_5, D_5$ , or  $A_5$ .*

**6.3.1. Solvable quintics.** If  $G = S_5, A_5$  then the equation  $f(x) = 0$  is not solvable by radicals. We want to investigate here the case  $G$  is not isomorphic to  $S_5$  or  $A_5$ . Let  $f(x)$  be an irreducible quintic in  $\mathbb{F}[x]$  given by

$$(27) \quad f(x) = x^5 + c_4x^4 + \cdots + c_0 = (x - \alpha_1) \cdots (x - \alpha_5)$$

Let  $G = \text{Gal}(f)$ , viewed as a (transitive) subgroup of  $S_5$  via permuting the (distinct) roots  $\alpha_1, \dots, \alpha_5$ . As before  $E_f = \mathbb{F}(\alpha_1, \dots, \alpha_5)$  denotes the splitting field.

A 5-cycle in  $S_5 = \text{Sym}(\{1, \dots, 5\})$  corresponds to an oriented pentagon with vertices  $1, \dots, 5$ . A 5-cycle and its inverse correspond to a (non-oriented) pentagon, and the full  $C_5$  corresponds to a pentagon together with its "opposite".



Thus  $F_5$ , the normalizer of  $C_5$  in  $S_5$ , is the subgroup permuting the pentagon and its opposite.  $D_5$  is the subgroup of  $F_5$  fixing the pentagon (symmetry group of the pentagon), and  $C_5$  is the subgroup of rotations. For example,  $F_5$  is generated by

$$(28) \quad F_5 = \langle \sigma, \tau \mid \sigma^5 = \tau^4 = (\sigma\tau)^4 = \sigma\sigma\tau\sigma^{-1}\tau^{-1} \rangle,$$

where  $\sigma = (12345)$  and  $\tau = (2453)$ . Thus if  $G \leq F_5$  then  $G$  fixes

$$(29) \quad \begin{aligned} \delta_1 = & (\alpha_1 - \alpha_2)^2(\alpha_2 - \alpha_3)^2(\alpha_3 - \alpha_4)^2(\alpha_4 - \alpha_5)^2(\alpha_5 - \alpha_1)^2 \\ & - (\alpha_1 - \alpha_3)^2(\alpha_3 - \alpha_5)^2(\alpha_5 - \alpha_2)^2(\alpha_2 - \alpha_4)^2(\alpha_4 - \alpha_1)^2 \end{aligned}$$

where the first (resp., second) term corresponds to the edges of the pentagon (resp., its opposite). There are six 5-Sylow subgroups of  $S_5$  given by

$$\begin{aligned} H_1 &= \langle (1, 2, 3, 4, 5) \rangle = \{(), (1, 2, 3, 4, 5), (1, 3, 5, 2, 4), (1, 4, 2, 5, 3), (1, 5, 4, 3, 2)\} \\ H_2 &= \langle (1, 2, 3, 5, 4) \rangle = \{(), (1, 2, 3, 5, 4), (1, 3, 4, 2, 5), (1, 5, 2, 4, 3), (1, 4, 5, 3, 2)\} \\ H_3 &= \langle (1, 2, 4, 5, 3) \rangle = \{(), (1, 2, 4, 5, 3), (1, 4, 3, 2, 5), (1, 5, 2, 3, 4), (1, 3, 5, 4, 2)\} \\ H_4 &= \langle (1, 2, 4, 3, 5) \rangle = \{(), (1, 2, 4, 3, 5), (1, 4, 5, 2, 3), (1, 3, 2, 5, 4), (1, 5, 3, 4, 2)\} \\ H_5 &= \langle (1, 2, 5, 3, 4) \rangle = \{(), (1, 2, 5, 3, 4), (1, 5, 4, 2, 3), (1, 3, 2, 4, 5), (1, 4, 3, 5, 2)\} \\ H_6 &= \langle (1, 3, 4, 5, 2) \rangle = \{(), (1, 3, 4, 5, 2), (1, 4, 2, 3, 5), (1, 5, 2, 4, 3), (1, 5, 3, 2, 4)\} \end{aligned}$$

To see the full invariance properties, we need to "projectivize" and use the invariants of binary forms; see section 5.2. Let  $y = 1 = \beta_i$ . The generalized version of the  $\delta_1$ 's is  $\tilde{\delta}_1$ , formed by replacing  $\alpha_i - \alpha_j$  by

$$D_{ij} = \det \begin{bmatrix} \gamma_i & \beta_i \\ \gamma_j & \beta_j \end{bmatrix} \text{ in the formulas defining the } \delta_i \text{'s. In particular,}$$

$$(30) \quad \tilde{\delta}_1 = D_{12}^2 D_{23}^2 D_{34}^2 D_{45}^2 D_{51}^2 - D_{13}^2 D_{35}^2 D_{52}^2 D_{24}^2 D_{41}^2$$

Since  $S_5$  has six 5-Sylow subgroups let  $\delta_1, \dots, \delta_6$  be the elements associated in this way to the six 5-Sylow's of  $S_5$ , i.e., to the six pentagon-opposite pentagon pairs on five given letters. We can write them all explicitly as

$$\begin{aligned} \tilde{\delta}_2 &= D_{12}^2 D_{23}^2 D_{35}^2 D_{54}^2 D_{41}^2 - D_{13}^2 D_{34}^2 D_{42}^2 D_{25}^2 D_{51}^2 \\ \tilde{\delta}_3 &= D_{12}^2 D_{24}^2 D_{45}^2 D_{53}^2 D_{31}^2 - D_{14}^2 D_{43}^2 D_{32}^2 D_{25}^2 D_{51}^2 \\ \tilde{\delta}_4 &= D_{12}^2 D_{24}^2 D_{43}^2 D_{35}^2 D_{51}^2 - D_{14}^2 D_{45}^2 D_{52}^2 D_{23}^2 D_{31}^2 \\ \tilde{\delta}_5 &= D_{12}^2 D_{25}^2 D_{53}^2 D_{34}^2 D_{41}^2 - D_{15}^2 D_{54}^2 D_{42}^2 D_{23}^2 D_{31}^2 \\ \tilde{\delta}_6 &= D_{13}^2 D_{34}^2 D_{45}^2 D_{52}^2 D_{21}^2 - D_{14}^2 D_{42}^2 D_{23}^2 D_{35}^2 D_{51}^2 \end{aligned} \quad (31)$$

**Lemma 20.**  $\delta_i^\sigma = \delta_i$  dhe  $\delta_i^\tau = \delta_i$  për  $i = 1, \dots, 6$ .

Clearly,  $G$  permutes  $\delta_1, \dots, \delta_6$ . If  $G$  is conjugate to a subgroup of  $F_5$ , it fixes one of  $\delta_1, \dots, \delta_6$ ; this fixed  $\delta_i$  must then lie in  $\mathbb{F}$ . Let  $\delta_1$  as in Eq. (29) and  $\delta_2, \dots, \delta_6$  as follows:

$$\begin{aligned} \delta_2 &= (\alpha_1 - \alpha_2)^2 (\alpha_2 - \alpha_3)^2 (\alpha_3 - \alpha_5)^2 (\alpha_5 - \alpha_4)^2 (\alpha_4 - \alpha_1)^2 \\ &\quad - (\alpha_1 - \alpha_3)^2 (\alpha_3 - \alpha_4)^2 (\alpha_4 - \alpha_2)^2 (\alpha_2 - \alpha_5)^2 (\alpha_5 - \alpha_1)^2 \\ \delta_3 &= (\alpha_1 - \alpha_2)^2 (\alpha_2 - \alpha_4)^2 (\alpha_4 - \alpha_5)^2 (\alpha_5 - \alpha_3)^2 (\alpha_3 - \alpha_1)^2 \\ &\quad - (\alpha_1 - \alpha_4)^2 (\alpha_4 - \alpha_3)^2 (\alpha_3 - \alpha_2)^2 (\alpha_2 - \alpha_5)^2 (\alpha_5 - \alpha_1)^2 \\ \delta_4 &= (\alpha_1 - \alpha_2)^2 (\alpha_2 - \alpha_4)^2 (\alpha_4 - \alpha_3)^2 (\alpha_3 - \alpha_5)^2 (\alpha_5 - \alpha_1)^2 \\ &\quad - (\alpha_1 - \alpha_4)^2 (\alpha_4 - \alpha_5)^2 (\alpha_5 - \alpha_2)^2 (\alpha_2 - \alpha_3)^2 (\alpha_3 - \alpha_1)^2 \\ \delta_5 &= (\alpha_1 - \alpha_2)^2 (\alpha_2 - \alpha_5)^2 (\alpha_5 - \alpha_3)^2 (\alpha_3 - \alpha_4)^2 (\alpha_4 - \alpha_1)^2 \\ &\quad - (\alpha_1 - \alpha_5)^2 (\alpha_5 - \alpha_4)^2 (\alpha_4 - \alpha_2)^2 (\alpha_2 - \alpha_3)^2 (\alpha_3 - \alpha_1)^2 \\ \delta_6 &= (\alpha_1 - \alpha_3)^2 (\alpha_3 - \alpha_4)^2 (\alpha_4 - \alpha_5)^2 (\alpha_5 - \alpha_2)^2 (\alpha_2 - \alpha_1)^2 \\ &\quad - (\alpha_1 - \alpha_4)^2 (\alpha_4 - \alpha_2)^2 (\alpha_2 - \alpha_3)^2 (\alpha_3 - \alpha_5)^2 (\alpha_5 - \alpha_1)^2 \end{aligned} \quad (32)$$

Thus, a necessary condition for the (irreducible) polynomial  $f(x)$  to be solvable by radicals is that one  $\delta_i$  lies in  $\mathbb{F}$ , i.e., that the polynomial

$$(33) \quad g(x) = (x - \delta_1) \cdots (x - \delta_6) \in \mathbb{F}[x]$$

has a root in  $\mathbb{F}$ . It is also sufficient:

**Lemma 21.** *If  $G$  fixes one  $\delta_i$  then  $G$  is conjugate to a subgroup of  $F_5$ , provided that  $\delta_1, \dots, \delta_6$  are all distinct.*

*Proof.* To check this it is enough to show that  $\delta_1, \dots, \delta_6$  are mutually distinct (under the hypothesis  $\Delta_f \neq 0$ ). hence, we have to show that  $\Delta_f \neq 0 \implies \Delta_g \neq 0$ . Using computational algebra we find  $\Delta_g$  and verify that

$$\Delta_g = ((\alpha_1 - \alpha_2)(\alpha_3 - \alpha_4)(\alpha_4 - \alpha_5)(\alpha_3 - \alpha_5))^4 \cdot \Delta_f \cdot I_2^2 \cdot I_3 \cdot I_4^2 \cdot I_6^2$$

where  $I_2, I_3, I_4$ , and  $I_6$  are given in [11]. Obviously  $\Delta_f \neq 0$  implies that  $\alpha_i - \alpha_j \neq 0$  for each  $i \neq j$ . This completes the proof.  $\square$

The coefficients of  $g(x)$  are symmetric functions in  $\alpha_1, \dots, \alpha_5$ , hence are polynomial expressions in  $c_0, \dots, c_4$ . The goal is to find these expressions explicitly. This gives an explicit criterion to check whether  $f(x) = 0$  is solvable by radicals.

**Lemma 22.** *Let  $s_r(x_1, \dots, x_6)$ ,  $r = 1, \dots, 6$ , be the elementary symmetric polynomials*

$$(34) \quad s_r = \sum_{i_1 < i_2 < \dots < i_r} x_{i_1} x_{i_2} \dots x_{i_r}.$$

*Then  $d_r := s_r(\tilde{\delta}_1, \dots, \tilde{\delta}_6)$  is a homogeneous polynomial expression in  $b_0, \dots, b_5$  of degree  $4r$ . These polynomials are invariant under the action of  $\text{SL}_2(\mathbb{F})$  on binary quintics: For any  $M \in \text{SL}_2(\mathbb{F})$  the quintic  $f^M$  has the same associated  $d_r$ 's.*

*Proof.* For  $\alpha_j := \gamma_j/\beta_j$  we have  $\tilde{\delta}_i = (\beta_1 \dots \beta_5)^4 \delta_i = b_5^4 \delta_i$ . Thus  $d_r = b_5^{4r} s_r(\delta_1, \dots, \delta_6)$ . But the  $s_r(\delta_1, \dots, \delta_6)$  are polynomial expressions in the  $c_j = b_j/b_5$ , for  $j = 0, \dots, 4$ . Thus  $d_r$  is a rational function in  $b_0, \dots, b_5$ , where the denominator is a power of  $b_5$ . Switching the roles of  $x$  and  $y$  yields that the denominator is also a power of  $b_0$ . Thus it is constant, i.e.,  $d_r$  is a polynomial in  $b_0, \dots, b_5$ . If we replace each  $\beta_j$  by  $c\beta_j$  for a scalar  $\lambda$  then each  $\tilde{\delta}_i$  gets multiplied by  $\lambda^4$ , so  $d_r$  gets multiplied by  $\lambda^{4r}$ . Thus  $d_r$  is homogeneous of degree  $4r$ . The rest of the claim is clear.  $\square$

$$(35) \quad \begin{array}{c} \mathbb{F}(\alpha_1, \dots, \alpha_5) \\ \swarrow \quad \downarrow \quad \searrow \quad \swarrow \quad \downarrow \quad \searrow \\ \mathbb{F}(\delta_1) \quad \mathbb{F}(\delta_2) \quad \mathbb{F}(\delta_3) \quad \mathbb{F}(\delta_4) \quad \mathbb{F}(\delta_5) \quad \mathbb{F}(\delta_6) \\ \swarrow \quad \downarrow \quad \searrow \quad \swarrow \quad \downarrow \quad \searrow \\ \mathbb{F}(\delta_1, \dots, \delta_6) \\ \downarrow \\ \mathbb{F}(s_1, \dots, s_6) \\ \downarrow \\ \mathbb{F} \end{array}$$

There are four basic invariants of quintics, denoted by  $J_4, J_8, J_{12}, J_{18}$ , of degrees 4, 8, 12 and 18, such that every  $\text{SL}(2, \mathbb{F})$ -invariant polynomial in  $b_0, \dots, b_5$  is a polynomial in  $J_4, J_8, J_{12}, J_{18}$ ; see e.g. I. Schur, Vorlesungen ueber Invariantentheorie, Springer 1968).

To define  $J_4, J_8, J_{12}$ , we need auxiliary quantities

$$A = \frac{1}{100} (20b_4 - 8b_1b_3 + 3b_2^2), \quad B = \frac{1}{100} (100b_5 - 12b_1b_4 + 2b_2b_3), \quad C = \frac{1}{100} (20b_1b_5 - 8b_2b_4 + 3b_3^2)$$

and  $D, E, F, G$  defined by

$$\begin{vmatrix} 10u + 2b_1v & 2b_1u + b_2v & b_2u + b_3v \\ 2b_1u + b_2v & b_2u + b_3v & b_3u + 2b_4v \\ b_2u + b_3v & b_3u + 2b_4v & 2b_4u + 10b_5v \end{vmatrix} = 10^3(Du^3 + Eu^2v + Fuv^2 + Gv^3)$$

Then  $J_2, J_8$ , and  $J_{12}$  are given by

$$\begin{aligned} J_4 &= 5^3(B^2 - 4AC) \\ (36) \quad J_8 &= 2^5 \cdot 5^6 [2A(3EG - F^2) - B(9DG - EF) + 2C(3FD - E^2)] \\ J_{12} &= -2^{10} \cdot 5^9 \cdot 3^{-1} [4(3EG - F^2)(3FD - E^2) - (9DG - EF)^2] \end{aligned}$$

By using special quintics one gets linear equations for the coefficients expressing the  $d_r$ 's in terms of  $J_4, J_8, J_{12}$ . The result is due to Berwick; see [19].

$$\begin{aligned} d_1 &= -10J_4 \\ d_2 &= 35J_4^2 + 10J_8 \\ d_3 &= -60J_4^3 - 30J_4J_8 - 10J_{12} \\ d_4 &= 55J_4^4 + 30J_4^2J_8 + 25J_8^2 + 50J_4J_{12} \\ d_5 &= -26J_4^5 - 10J_4^3J_8 - 44J_4J_8^2 - 59J_4^2J_{12} - 14J_8J_{12} \\ d_6 &= 5J_4^6 + 20J_4^2J_8^2 + 20J_4^3J_{12} + 20J_4J_8J_{12} + 25J_{12}^2 \end{aligned}$$

**Lemma 23.** *Let  $f(x)$  be a irreducible quintic over  $\mathbb{F}$  and  $d_1, \dots, d_6$  defined in terms of the coefficients of  $f(x)$  as above. Then  $f(x)$  is solvable by radicals if and only if  $g(x) = x^6 + d_1x^5 + \dots + d_5x + d_6$  has a root in  $\mathbb{F}$ .*

Extending the method of invariants becomes harder for higher degree equations. For degree six equations see [1] and [13]. We are not aware of explicit computations for degree  $d \geq 6$ .

$$\begin{aligned} \xi_0(f) &= -\frac{2}{625} \cdot (625a_0^2a_5^2 - 250a_0a_1a_4a_5 + 25a_0a_2a_3a_5 + 40a_0a_2a_4^2 - 15a_0a_3^2a_4 + 40a_1^2a_3a_5 + 9a_1^2a_4^2 \\ &\quad - 15a_1a_2^2a_5 - 19a_1a_2a_3a_4 + 6a_1a_3^3 + 6a_2^3a_4 - 2a_2^2a_3^2) \\ \xi_1(f) &= \frac{1}{1562500} (125000a_0^3a_2a_3a_5^3 - 50000a_0^3a_2a_4^2a_5^2 - 75000a_0^3a_3^2a_4a_5^2 + 50000a_0^3a_3a_4^3a_5 - 8000a_0^3a_4^5 \\ &\quad - 50000a_0^2a_1^2a_3a_5^3 + 20000a_0^2a_1^2a_4^2a_5^2 - 75000a_0^2a_1a_2^2a_5^3 + 55000a_0^2a_1a_2a_3a_4a_5^2 - 10000a_0^2a_1a_2a_4^3a_5 \\ &\quad + 30000a_0^2a_1a_3^3a_5^2 - 30000a_0^2a_1a_3^2a_4^2a_5 + 6000a_0^2a_1a_3a_4^4 + 30000a_0^2a_2^3a_4a_5^2 - 26250a_0^2a_2^2a_3^2a_5^2 \\ &\quad - 17000a_0^2a_2^2a_3a_4^2a_5 + 4400a_0^2a_2^2a_4^4 + 19500a_0^2a_2a_3^3a_4a_5 - 4800a_0^2a_2a_3^2a_4^3 - 3375a_0^2a_3^5a_5 + 900a_0^2a_3^4a_4^2 \\ &\quad + 50000a_0a_1^3a_2a_5^3 - 10000a_0a_1^3a_3a_4a_5^2 - 30000a_0a_1^2a_2^2a_4a_5^2 - 17000a_0a_1^2a_2a_3^2a_5^2 + 26600a_0a_1^2a_2a_3a_4^2a_5 \\ &\quad - 4320a_0a_1^2a_2a_4^4 - 1800a_0a_1^2a_3^3a_4a_5 + 120a_0a_1^2a_3^2a_4^3 + 19500a_0a_1a_3^3a_3a_5^2 - 1800a_0a_1a_3^2a_4^2a_5 \\ &\quad - 11000a_0a_1a_2^2a_3^2a_4a_5 + 2120a_0a_1a_2^2a_3a_4^3 + 2325a_0a_1a_2a_3^4a_5 - 300a_0a_1a_2a_3^3a_4^2 - 45a_0a_1a_3^5a_4 \\ &\quad - 3375a_0a_2^5a_5^2 + 2325a_0a_2^4a_3a_4a_5 - 380a_0a_2^4a_4^3 - 525a_0a_2^3a_3^3a_5 + 40a_0a_2^3a_3^2a_4^2 + 15a_0a_2^2a_3^4a_4 \\ &\quad - 8000a_1^5a_5^3 + 6000a_1^4a_2a_4a_5^2 + 4400a_1^4a_3^2a_5^2 - 4320a_1^4a_3a_4^2a_5 + 864a_1^4a_4^4 - 4800a_1^3a_2^2a_3a_5^2 + 120a_1^3a_2^2a_4^2a_5 \\ &\quad + 2120a_1^3a_2a_3^2a_4a_5 - 648a_1^3a_2a_3a_4^3 - 380a_1^3a_3^4a_5 + 152a_1^3a_3^3a_4^2 + 900a_1^2a_2^4a_5^2 - 300a_1^2a_2^3a_3a_4a_5 \\ &\quad + 152a_1^2a_2^3a_4^3 + 40a_1^2a_2^2a_3^3a_5 + 50a_1^2a_2^2a_3^2a_4^2 - 57a_1^2a_2a_3^4a_4 + 9a_1^2a_3^6 - 45a_1a_2^5a_4a_5 + 15a_1a_2^4a_3^2a_5 \\ &\quad - 57a_1a_2^4a_3a_4^2 + 37a_1a_2^3a_3^3a_4 - 6a_1a_2^2a_3^5 + 9a_2^6a_4^2 - 6a_2^5a_3^2a_4 + a_2^4a_3^4) \end{aligned}$$

We display  $\xi_2(f)$  in the appendix; see ??.

7. REDUCTION MODULO  $p$ 

The reduction method uses the fact that once a every polynomial with rational coefficients can be transformed into a monic polynomial with integer coefficients without changing the splitting field.

Let  $f(x) \in \mathbb{Q}[x]$  be given by

$$(37) \quad f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$$

Let  $d$  be the common denominator of all coefficients  $a_0, \dots, a_{n-1}$ . Then  $g(x) := df(\frac{x}{d})$  is a monic polynomial with integer coefficients. Clearly the splitting field of  $f(x)$  is the same as the splitting field of  $g(x)$ . Thus, without loss of generality we can assume that  $f(x)$  is a monic polynomial with integer coefficients.

**Theorem 7.1. (Dedekind)** *Let  $f(x) \in \mathbb{Z}[x]$  be a monic polynomial such that  $\deg f = n$ ,  $\text{Gal}_{\mathbb{Q}}(f) = G$ , and  $p$  a prime such that  $p \nmid \Delta_f$ . If  $f_p := f(x) \bmod p$  factors in  $\mathbb{Z}_p[x]$  as a product of irreducible factors of degree  $n_1, n_2, n_3, \dots, n_k$ , then  $G$  contains a permutation of type  $(n_1)(n_2) \cdots (n_k)$*

The Dedekind theorem can be used to determine the Galois group in many cases since the *type* of permutation in  $S_n$  determines the conjugacy class in  $S_n$ . Consider for example polynomials of degree 5. The cycle types for all groups that occur as Galois groups of quintics are given below.

	(2)	(2) <sup>2</sup>	(3)	(4)	(3)(2)	(5)
$S_5$	10	15	20	30	20	24
$A_5$		15	20			24
$F_5$		5		10		4
$D_5$		5				4
$C_5$						4

TABLE 1. Cycle types for Galois groups of quintics

In table 2 we display the table for the type of elements in  $S_6$ . As it can be seen from the tables this method works well for degree 5 and 6. Unfortunately it does not work for degree  $d > 6$ .

	()	(2)	(2)(2)	(2)(2)(2)	(3)	(3)(2)	(3)(3)	(4)	(4)(2)	(5)	(6)	G
$S_6$	1	15	45	15	40	120	40	90	90	144	120	720
$A_6$	1	-	45	-	40	-	40	-	90	144	-	360
$S_5$	1	-	15	10	-	-	20	30	-	24	20	120
$(S_3 \times S_3) \rtimes C_2$	1	6	9	6	4	12	4	-	18	-	12	72
$A_5$	1	-	15	-	-	-	20	-	-	24	-	60
$C_2 \times S_4$	1	3	9	7	-	-	8	6	6	-	8	48
$(C_3 \times C_3) \rtimes C_4$	1	-	9	-	4	-	4	-	18	-	-	36
$S_3 \times S_3$	1	-	9	6	4	-	4	-	-	-	12	36
$S_4$	1	-	3	6	-	-	8	6	-	-	-	24
$S_4$	1	-	9	-	-	-	8	-	6	-	-	24
$C_2 \times A_4$	1	3	3	1	-	-	8	-	-	-	8	24
$C_3 \times S_3$	1	-	-	3	4	-	4	-	-	-	6	18
$A_4$	1	-	3	-	-	-	8	-	-	-	-	12
$D_{12}$	1	-	3	4	-	-	2	-	-	-	2	12
$S_3$	1	-	-	3	-	-	2	-	-	-	-	6
$C_6$	1	-	-	1	-	-	2	-	-	-	2	6

TABLE 2. Cycle types for Galois groups of sextics

## 8. TRANSITIVE GROUPS

Here is the number of transitive subgroups for  $n \leq 47$

$n$	# Subgroups	$n$	# Subgroups	$n$	# Subgroups	$n$	# Subgroups
5	5	6	16	7	7	8	50
9	34	10	45	11	8	12	301
13	9	14	63	15	104	16	1954
17	10	18	983	19	8	20	1117
21	164	22	59	23	7	24	25000
25	211	26	96	27	2392	28	1854
29	8	30	5712	31	12	33	162
34	115	35	407	36	121279	37	11
38	76	39	306	40	315842	41	10
42	9491	43	10	44	2113	45	10923

TABLE 3. Number of transitive subgroups of  $S_n$  for select values of  $n$

TABLE 4. Transitive Subgroups of  $S_n$  for  $n = 5, 6, 7, 11, 13, 17, 19$

$n$	Subgroups
5	$C(5) = 5, D(5) = 5 : 2, F(5) = 5 : 4, A_5, S_5$ ]
6	$C(6) = 6 = 3[x]^2, D_6(6) = [3]^2, D(6) = S(3)[x]^2, A_4(6) = [2^2]3, F_{18}(6) = [3^2]2 = 3 \wr 2,$ $2A_4(6) = [2^3]3 = 2 \wr 3, S_4(6d) = [2^2]S(3), S_4(6c) = \frac{1}{2}[2^3]S(3), F_{18}(6) : 2 = [\frac{1}{2}S(3)^2]2,$ $F_{36}(6) = \frac{1}{2}[S(3)^2]2, 2S_4(6) = [2^3]S(3) = 2 \wr S(3), L(6) = PSL(2, 5) = A_5(6),$ $F_{36}(6) : 2 = [S(3)^2]2 = S(3) \wr 2, L(6) : 2 = PGL(2, 5) = S_5(6), A_6, S_6$
7	$C(7) = 7, D(7) = 7 : 2, F_{21}(7) = 7 : 3, F_{42}(7) = 7 : 6, L(7) = L(3, 2), A_7, S_7$
11	$C(11) = 11, D(11) = 11 : 2, F_{55}(11) = 11 : 5, F_{110}(11) = 11 : 10, L(11) = PSL(2, 11)(11),$ $M(11), A_{11}, S_{11}$
13	$C(13) = 13, D(13) = 13 : 2, F_{39}(13) = 13 : 3, F_{52}(13) = 13 : 4, F_{78}(13) = 13 : 6,$ $F_{156}(13) = 13 : 12,$ $L(13) = PSL(3, 3), A_{13}, S_{13}$
17	$C(17) = 17, D(17) = 17 : 2, F_{68}(17) = 17 : 4, F_{136}(17) = 17 : 8, F_{272}(17) = 17 : 16,$ $L(17) = PSL(2, 16), L(17) : 2 = PZL(2, 16), L(17) : 4 = PYL(2, 16), A_{17}, S_{17}$
19	$C(19) = 19, D(19) = 19 : 2, F_{57}(19) = 19 : 3, F_{114}(19) = 19 : 6, F_{171}(19) = 19 : 9,$ $F_{342}(19) = 19 : 18, A_{19}, S_{19}$

## 9. DATABASES

**9.1. Datasets of irreducible polynomials.** In this section we want to create a database of irreducible polynomials  $f \in \mathbb{Z}[x]$  of degree  $\deg f = n$ . Data will be stored in a Python dictionary. A polynomial  $f(x) = \sum_{i=0}^n a_i x^i$  will be represented by its corresponding binary form  $f(x, y) = \sum_{i=0}^n a_i x^i y^{n-i}$ . Hence our points will be points in the projective space  $\mathbb{P}_{\mathbb{Q}}^n$ , i.e. points with integer coordinates

$$\mathbf{p} = [a_n : \cdots : a_0] \in \mathbb{P}_{\mathbb{Q}}^n,$$

such that  $\gcd(a_0, \dots, a_n)$ . Since  $f(x)$  is irreducible over  $\mathbb{Q}$  and of degree  $\deg f = n$ , then  $a_n \neq 0$  and  $a_0 \neq 0$ . Moreover,  $\Delta_f \neq 0$ .

**9.2. Datasets with bounded height.** Let us now trying to generate a dataset with a bounded height  $h$  as defined in Eq. (12). We will denote the set of such polynomials by  $\mathcal{P}_n^h$ . In other words

$$\mathcal{P}_n^h := \{[a_n : \cdots : a_0] \in \mathbb{P}_{\mathbb{Q}}^n \mid a_0 a_n \neq 0, \Delta_f \neq 0, H_{\mathbb{Q}}([a_n : \cdots : a_0]) \leq h\}$$

where  $H_{\mathbb{Q}}$  is defined as in Eq. (12). To ensure that the points in the database are not repeated we key the dictionary by the tuples  $\left(1, \frac{a_{n-1}}{a_n}, \dots, \frac{a_0}{a_n}\right)$ . This is safe since  $a_n \neq 0$ . A dictionary in Python does not allow key duplicates, which ensures that there are no duplicates in our data. For given  $h, n$  the cardinality of  $\mathcal{P}_n^h$  is bounded by

$$\#\mathcal{P}_n^h \leq 4h^2(2h+1)^{n-2}$$

The proof is a straightforward counting argument. There are more sophisticated methods to count algebraic points of bounded height on projective spaces; see for example [12] but we will work only over  $\mathbb{Q}$  and our heights will be relatively small which does not allow for much redundant data.

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

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

Below is the number of points for small  $d$  and  $h$ .

deg	h=1	h=2	h=3	h=4	h=5
3		136	668	1 940	4 936
4		694	4 823	18 528	569 912
5		3 616	34 860	174 120	639 476
6		18 602	249 498		

For every point  $\mathbf{p} = [a_n : \cdots : a_0]$  we will compute the following attributes

$$\left(1, \frac{a_{n-1}}{a_n}, \dots, \frac{a_0}{a_n}\right) : [\mathbf{p}, \xi_0, \dots, \xi_n, \Delta_f, H(f), \mathfrak{H}_k(\mathbf{p}), \mathfrak{H}(\mathbf{p}), \mathcal{T}_2, \mathcal{T}_3, \mathcal{T}_5, \mathcal{T}_7, \text{Gal}_{\mathbb{Q}}(f), \text{Relations},]$$

where

$\xi_0, \dots, \xi_n$	Invariants defined in <a href="#">section 5.2</a>
$\Delta_f$	Discriminant of $f(x)$
$H(f)$	Height of $f(x)$ defined in <a href="#">Eq. (12)</a>
$\mathfrak{H}_k(\mathbf{p})$	Weighted moduli height as in <a href="#">Eq. (15)</a>
$\mathfrak{H}(\mathbf{p})$	Absolute weighted moduli height as in <a href="#">Eq. (16)</a>
$\text{Gal}_{\mathbb{Q}}(f)$	Gap Identity of the Galois group of $f(x)$
Relations	Relations among invariants when possible determined by ....

**9.3. Cubics.** As a simple first exercise we start with irreducible cubics. We create a database of all rational points  $[c_0 : c_1 : c_2 : c_3]$  in  $\mathbb{P}^3$  with projective height  $h \leq 20$  such that

$$f(x) = c_0 + c_1x + c_2x^2 + c_3x^3$$

is an irreducible polynomial in  $\mathbb{Q}[x]$ . Since training a model for determining  $\text{Gal}(f)$  is trivial in this case we will focus mostly on comparing the naive height with the weighted moduli height and determining how the occurrence of  $A_3$  happens with the increase of  $h$ .

A slice of five random elements of our Python dictionary looks like:

Key	Value
(-1, -9, -20, 1)	[20, 98, 3.1463462836, 'A3']
(20, -9, -20, 1)	[20, 1458632, 34.752530588, 'A3']
(8, 12, -20, 1)	[20, 540800, 13.5590472788, 'A3']
(1, 17, -20, 1)	[20, 243602, 22.2162222997, 'A3']
(19, -9, -19, 1)	[19, 1204352, 16.5637384397, 'A3']

where the 'key' has the coefficients of the cubic and the entries in 'values' are respectively: naive height,  $J_4$  invariant, weighted height, and the Galois group.

**Lemma 24.** *The total number of rational points of heights in  $(0, 20]$  is 1 299 200. From those there are 1 178 856 irreducible polynomials and only 1328 of them have Galois group  $C_3$ . Moreover, the distribution of polynomials with Galois group  $C_3$  with respect to their naive height is given in fig. 1.*

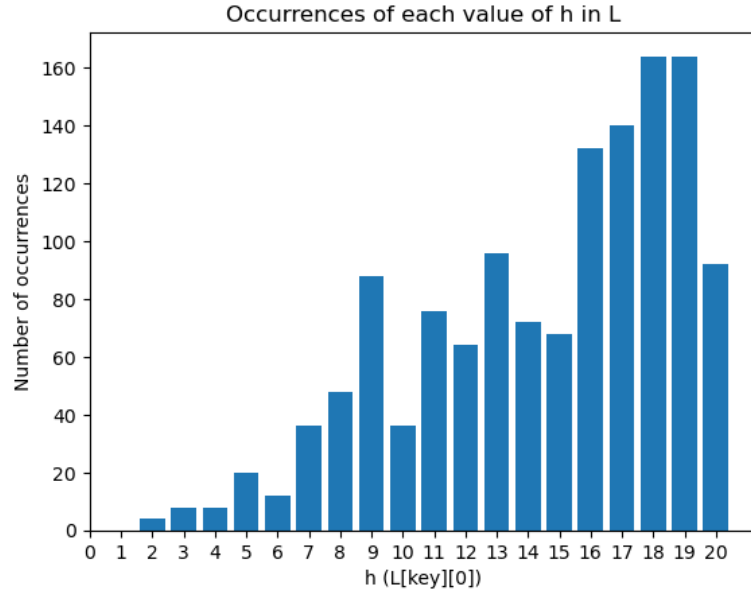


FIGURE 1. This distribution is only for cubics with Galois group  $C_3$ .

In [39] we give an estimate on the ratio of the moduli height over the naive height for binary sextics. Such bounds can be given for every degree  $d$  polynomial. In our case of cubics the minimum ratio is 0.074 for polynomial  $f(x) = 7x^3 - 5x^2 - 16x + 7$  and the maximum ratio is 2.008 for  $f(x) = 13x^3 - 19x^2 - 20x + 13$ .

**Lemma 25.** *There are only 40 cubics in the database with height  $\leq 5$  and Galois group of order 3. The discriminant  $\Delta_f$  of those forty polynomials has values  $\Delta_f = 7^2, 3^4, 13^2, 19^2, 31^2$ , and  $61^2$  as shown in the table 5*



TABLE 5. Irreducible degree 3 polynomials of height  $\leq 5$  and Galois group  $C_3$ 

#	$f$	$\Delta$	#	$f$	$\Delta$	#	$f$	$\Delta$
1	(1, 3, -4, 1)	$7^2$	15	(-1, -3, 0, 3)	$3^4$	29	(1, 2, -5, 1)	$19^2$
2	(-1, -4, -3, 1)	$7^2$	16	(1, -3, 0, 3)	$3^4$	30	(-1, -5, -2, 1)	$19^2$
3	(1, -1, -2, 1)	$7^2$	17	(5, 4, -5, 1)	$13^2$	31	(1, -5, 2, 1)	$19^2$
4	(1, -2, -1, 1)	$7^2$	18	(1, 1, -4, 1)	$13^2$	32	(-1, 2, 5, 1)	$19^2$
5	(-1, -2, 1, 1)	$7^2$	19	(5, -3, -2, 1)	$13^2$	33	(2, -1, -5, 2)	$31^2$
6	(-1, -1, 2, 1)	$7^2$	20	(-1, -4, -1, 1)	$13^2$	32	(2, -5, -1, 2)	$31^2$
7	(1, -4, 3, 1)	$7^2$	21	(1, -4, 1, 1)	$13^2$	35	(-2, -5, 1, 2)	$31^2$
8	(-1, 3, 4, 1)	$7^2$	22	(-5, -3, 2, 1)	$13^2$	36	(-2, -1, 5, 2)	$31^2$
9	(1, 0, -3, 1)	$3^4$	23	(-1, 1, 4, 1)	$13^2$	37	(3, -4, -5, 3)	$61^2$
10	(3, 0, -3, 1)	$3^4$	24	(-5, 4, 5, 1)	$13^2$	38	(3, -5, -4, 3)	$61^2$
11	(-1, -3, 0, 1)	$3^4$	25	(-1, -5, -4, 5)	$13^2$	39	(-3, -5, 4, 3)	$61^2$
12	(1, -3, 0, 1)	$3^4$	26	(1, -2, -3, 5)	$13^2$	40	(-3, -4, 5, 3)	$61^2$
13	(-3, 0, 3, 1)	$3^4$	27	(-1, -2, 3, 5)	$13^2$			
14	(-1, 0, 3, 1)	$3^4$	28	(1, -5, 4, 5)	$13^2$			

Below is the distribution of points in the database versus the invariant of cubics.

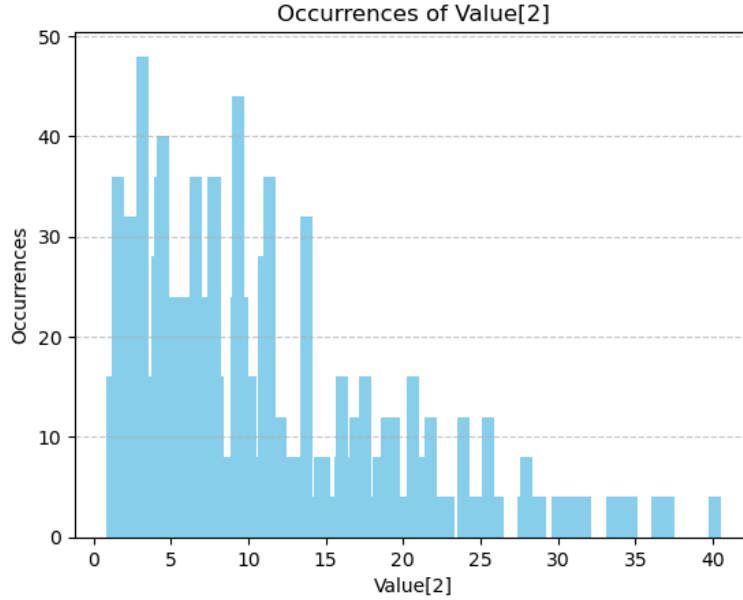


FIGURE 2. The number of occurrences versus the invariants

### 9.3.1. An unsupervised model on cubics.

9.4. **Quartics.** Let us now continue with the irreducible quartic polynomials. We refer to ?? for its invariants.

TABLE 6. Polynomials of degree  $4 \leq d \leq 5$  and height  $h \leq 5$  and the number of Galois groups in each case.

$n = 4$							
$h$	$\#\mathbb{P}_h^4(\mathbb{Q})$	$\# \text{ irred}$	$[4,1]$	$[4,2]$	$[8,3]$	$[12,3]$	$S_4$
1	121	34	2	2	10	-	20
2	1 441	694	2	10	114	4	564
3	8 161	4 823	4	25	422	32	4 340
4	27 841	18 528	24	90	1 318	52	17 044
5	78 721	56 500	42	142	2812	108	53 396

$n = 5$							
$h$	$\#\mathbb{P}_h^5(\mathbb{Q})$	$\# \text{ irred}$	$C_5$	$D_5$	$F(5)$	$A_5$	$S_5$
1	364	104	-	-	-	-	104
2	7 448	3 616	0	12	8	12	3 584
3	58 096	34 860	0	100	28	76	34 656
4	257 544	174 120	4	192	104	180	173 640
5	877 240	639 476	8	388	268	460	638 352

## 10. NEURO-SYMBOLIC NETWORKS

A neuro-symbolic network is a type of artificial intelligence system that combines the strengths of neural networks (good at pattern recognition) with symbolic reasoning (based on logic and rules) to create models that can both learn from data and reason through complex situations, essentially mimicking human-like cognitive abilities by understanding and manipulating symbols to make decisions; this approach aims to overcome limitations of either method alone, providing better explainability and adaptability in AI systems.

## 11. CONCLUDING REMARKS

This paper introduces an innovative approach to Galois theory by leveraging machine learning techniques to address challenges in understanding polynomial properties and their Galois groups. Combining classical algebraic structures with computational tools opens new avenues for exploring the connections between mathematics and data science.

We have demonstrated the potential of supervised learning to predict Galois groups and polynomial solvability, while unsupervised learning reveals latent structures in polynomial datasets. A comprehensive database of irreducible polynomials with known Galois groups has been compiled, and classical invariants such as discriminants, root differences, and moduli heights have been explored as features for machine learning models. Reduction theories, including Julia and Hermite equivalence, were employed to streamline classification, and the role of polynomial heights in minimal forms and equivalence classes was investigated. The geometric interpretation of polynomial transformations within weighted projective spaces further enhances this framework.

Future work could extend the polynomial database to higher degrees, incorporate multivariable polynomials, and develop novel invariants derived from machine learning. Advanced models, such as graph neural networks, could refine the analysis of root interactions and symmetries, while transfer learning may generalize insights to more complex cases. Automation of reduction methods and interactive visualization

tools could make these techniques accessible to a broader audience. Additionally, extending this framework to analyze field extensions and connections with algebraic geometry or physics could broaden its impact.

This work demonstrates the feasibility of integrating machine learning with classical mathematics, offering new tools for algebraists while uncovering deeper theoretical insights. By bridging abstract mathematics and computational science, this approach paves the way for a more interdisciplinary perspective in mathematical research.

## REFERENCES

- [1] W. E. H. Berwick, *On Soluble Sextic Equations*, Proc. London Math. Soc. (2) **29** (1928), no. 1, 1–28. MR1575303
- [2] L. Beshaj, J. Gutierrez, and T. Shaska, *Weighted greatest common divisors and weighted heights*, J. Number Theory **213** (2020), 319–346. MR4091944
- [3] L. Beshaj, R. Hidalgo, S. Kruk, A. Malmendier, S. Quispe, and T. Shaska, *Rational points in the moduli space of genus two*, Higher genus curves in mathematical physics and arithmetic geometry, 2018, pp. 83–115. MR3782461
- [4] Manjul Bhargava, Jan-Hendrik Evertse, Kálmán Györy, László Remete, and Ashvin A. Swaminathan, *Hermite equivalence of polynomials*, Acta Arith. **209** (2023), 17–58. MR4665252
- [5] A. Białostocki and T. Shaska, *Galois groups of prime degree polynomials with nonreal roots*, Computational aspects of algebraic curves, 2005, pp. 243–255. MR2182043
- [6] Enrico Bombieri and Walter Gubler, *Heights in Diophantine geometry*, New Mathematical Monographs, vol. 4, Cambridge University Press, Cambridge, 2006. MR2216774 (2007a:11092)
- [7] Nicolas Bourbaki, *Algebra I. Chapters 1–3*, Elements of Mathematics (Berlin), Springer-Verlag, Berlin, 1998. Translated from the French, Reprint of the 1989 English translation [MR0979982 (90d:00002)]. MR1727844
- [8] ———, *Algebra II. Chapters 4–7*, Elements of Mathematics (Berlin), Springer-Verlag, Berlin, 2003. Translated from the 1981 French edition by P. M. Cohn and J. Howie, Reprint of the 1990 English edition [Springer, Berlin; MR1080964 (91h:00003)]. MR1994218
- [9] A. Clebsch and P. Gordan, *Theorie der abelschen funktionen*, Teubner, 1866.
- [10] Elira Curri, *Zgjidhja e ekuacionit të gradës së 5-të* (2021March), 11 pp., available at <https://www.risat.org/pdf/Quintic.pdf>.
- [11] ———, *On the stability of binary forms and their weighted heights*, Albanian J. Math. **16** (2022), no. 1, 3–23. MR4448533
- [12] Quentin Guignard, *Counting algebraic points of bounded height on projective spaces*, J. Number Theory **170** (2017), 103–141. MR3541701
- [13] Thomas R. Hagedorn, *General formulas for solving solvable sextic equations*, J. Algebra **233** (2000), no. 2, 704–757. MR1793923
- [14] Marc Hindry and Joseph H. Silverman, *Diophantine geometry*, Graduate Texts in Mathematics, vol. 201, Springer-Verlag, New York, 2000. An introduction. MR1745599 (2001e:11058)
- [15] Jun-ichi Igusa, *Arithmetic variety of moduli for genus two*, Ann. of Math. (2) **72** (1960), 612–649. MR0114819
- [16] David Joyner and Tony Shaska, *Self-inversive polynomials, curves, and codes*, Higher genus curves in mathematical physics and arithmetic geometry, 2018, pp. 189–208. MR3782467
- [17] Gaston Julia, *Étude sur les formes binaires non quadratiques à indéterminées réelles, ou complexes, ou à indéterminées conjuguées* (1917), 293. MR3532882
- [18] George R. Kempf, *Instability in invariant theory*, Ann. of Math. (2) **108** (1978), no. 2, 299–316. MR506989
- [19] R. Bruce King, *Beyond the quartic equation*, Birkhäuser Boston, Inc., Boston, MA, 1996. MR1401346
- [20] Ilias Kotsireas and Tony Shaska, *A machine learning approach of Julia reduction*, RISAT preprints (202412), available at <https://www.risat.org/pdf/2024-06.pdf>.
- [21] Vishwanath Krishnamoorthy, Tanush Shaska, and Helmut Völklein, *Invariants of binary forms*, Progress in Galois theory, 2005, pp. 101–122. MR2148462
- [22] Joseph P. S. Kung and Gian-Carlo Rota, *The invariant theory of binary forms*, Bull. Amer. Math. Soc. (N.S.) **10** (1984), no. 1, 27–85. MR722856
- [23] Serge Lang, *Fundamentals of Diophantine geometry*, Springer-Verlag, New York, 1983. MR715605 (85j:11005)
- [24] A. Malmendier and T. Shaska, *The Satake sextic in F-theory*, J. Geom. Phys. **120** (2017), 290–305. MR3712162
- [25] ———, *A universal genus-two curve from Siegel modular forms*, SIGMA Symmetry Integrability Geom. Methods Appl. **13** (2017), Paper No. 089, 17. MR3731039
- [26] ———, *From hyperelliptic to superelliptic curves*, Albanian J. Math. **13** (2019), no. 1, 107–200. MR3978315
- [27] P. E. Newstead, *Geometric invariant theory*, Moduli spaces and vector bundles, 2009, pp. 99–127. MR2537067
- [28] Peter J. Olver, *Classical invariant theory*, London Mathematical Society Student Texts, vol. 44, Cambridge University Press, Cambridge, 1999. MR1694364

- [29] Sajad Salami and Tony Shaska, *Local and global heights on weighted projective varieties*, Houston J. Math. **49** (2023), no. 3, 603–636 (English).
- [30] George Salmon, *Modern higher algebra*, Cambridge University Press, Cambridge, 1876.
- [31] Bedri Shaska and Tanush Shaska, *Mësimdhënia e matematikës nëpërmjet problemeve klasike*, Albanian Journal of Mathematics **10** (2016), no. 1, 47–80.
- [32] Elira Shaska and Tony Shaska, *Irreducible sextics, invariants, and their galois groups*, RISAT preprints (202412), available at <https://www.risat.org/pdf/2024-07.pdf>.
- [33] ———, *Machine learning for moduli space of genus two curves and an application to isogeny based cryptography*, Journal of Algebraic Combinatorics (2024), available at <https://www.risat.org/pdf/2024-03.pdf>.
- [34] ———, *Polynomials, galois groups, and deep learning*, RISAT preprints (202412), available at <https://www.risat.org/pdf/2024-05.pdf>.
- [35] ———, *Rational functions on the projective line through machine learning techniques*, RISAT preprints (202412), available at <https://www.risat.org/pdf/2024-04.pdf>.
- [36] T. Shaska, *Reduction of superelliptic Riemann surfaces*, Automorphisms of Riemann surfaces, subgroups of mapping class groups and related topics, 2022, pp. 227–247. MR4375119
- [37] ———, *Artificial neural networks on graded vector spaces* (2024), available at [2407.19031](https://arxiv.org/abs/2407.19031).
- [38] T. Shaska and L. Beshaj, *The arithmetic of genus two curves*, Information security, coding theory and related combinatorics, 2011, pp. 59–98. MR2963126
- [39] ———, *Heights on algebraic curves*, Advances on superelliptic curves and their applications, 2015, pp. 137–175. MR3525576
- [40] Helmut Voelklein and Tanush Shaska (eds.), *Progress in Galois theory*, Developments in Mathematics, vol. 12, Springer, New York, 2005. MR2150438