Machine learning and moduli spaces

Tony Shaska Sr.[1, 2] Elira Curri[1]

[1] Department of Computer Science, Research Institute of Science and Technology (RISAT), Vlorë, Albania

[2] Department of Mathematics, Oakland University, Rochester, MI.

Polynomial Computer Algebra 2023
April 17-22, 2023
Euler International Mathematical Institute, St. Petersburg, RUSSIA

Abstract

Can we develop a machine learning model to get information on the arithmetic of the moduli space of curves $\mathcal{M}_g$? We propose new methods, to apply machine learning to various databases which have emerged in the study of the moduli spaces of algebraic curves.

Basic questions about the moduli space $\mathcal{M}_g$

Let $g > 1$ be a fixed integer, k a number field, and $\mathcal{M}_g$ the moduli space of smooth genus g curves defined over $\bar{k}$. It is a (quasi)projective variety of dimension $3g - 3$.

Consider the following questions:

- What is the singular locus of $\mathcal{M}_g$?
- What is the list of automorphisms which can occur as $\mathrm{Aut}(\mathfrak{p})$ for $\mathfrak{p} \in \mathcal{M}_g$?
- What is the stratification of $\mathcal{M}_g$ based on the automorphisms?
- For what $\mathfrak{p} \in \mathcal{M}_g$ the corresponding Jacobian has complex multiplication? What is the distribution of such points?
- How can the points $\mathfrak{p} \in \mathcal{M}_g$ be described explicitly?
- Given a k-rational point $\mathfrak{p} \in \mathcal{M}_g(k)$, is there a curve $\mathcal{X}$, defined over k, such that $\mathfrak{p} = [\mathcal{X}]$? For such points we say that field of moduli is a field of definition
- What is the distribution of points in $\mathcal{M}_g$ for which the field of moduli is a field of definition?
- Let $\mathfrak{p} \in \mathcal{M}_g(k)$ and $\mathcal{X}$ defined over k corresponding to $\mathfrak{p}$. Can we determine an equation for $\mathcal{X}$? Is this equation canonical in some sense? Reduction type A
- Is there a way to create some kind of a "database" of k-rational moduli points $\mathfrak{p} \in \mathcal{M}_g(k)$ such that the field of moduli is a field of definition? In other words, how to order points in $\mathcal{M}_g$ or define some kind of "size" for $\mathfrak{p} \in \mathcal{M}_g(k)$? Can we choose a coordinate in $\mathcal{M}_g$ such that points in $\mathcal{M}_g(k)$ are of "small size" Reduction Type B or moduli reduction

Learning models

- Supervised ML methods
- Unsupervised Machine Learning methods

Genus 2 as a case study

- Invariants for genus 2 curves
- Subloci of M_2

Automorphisms

- Definition of strata based on automorphisms; Hurwitz spaces
- Stratification of $\mathcal{M}_g$ based on automorphisms

From hyperelliptic to superelliptic

Moduli points and Geometric Invariant Theory (GIT)

- Weighted projective spaces
- Sorting points in the moduli space; weighted heights

Complex multiplication

SUPERVISED ML METHODS

Supervised learning is a machine learning approach that?s defined by its use of labeled datasets. These datasets are designed to train or *supervise* algorithms into classifying data or predicting outcomes accurately. Using labeled inputs and outputs, the model can measure its accuracy and learn over time.

1. Determine the type of training examples. Before doing anything else, the user should decide what kind of data is to be used as a training set.
2. Gather a training set. The training set needs to be representative of the real-world use of the function. Thus, a set of input objects is gathered and corresponding outputs are also gathered, either from human experts or from measurements.
3. Determine the input feature representation of the learned function. The accuracy of the learned function depends strongly on how the input object is represented. Typically, the input object is transformed into a feature vector, which contains a number of features that are descriptive of the object.
4. Determine the structure of the learned function and corresponding learning algorithm. For example, the engineer may choose to use support-vector machines or decision trees.
5. Complete the design. Run the learning algorithm on the gathered training set. Some supervised learning algorithms require the user to determine certain control parameters. These parameters may be adjusted by optimizing performance on a subset (called a validation set) of the training set, or via cross-validation.
6. Evaluate the accuracy of the learned function. After parameter adjustment and learning, the performance of the resulting function should be measured on a test set that is separate from the training set.

Unsupervised Machine Learning methods

Unsupervised learning uses machine learning algorithms to analyze and cluster unlabeled datasets. These algorithms discover hidden patterns or data groupings without the need for human intervention.

- Clustering
- Association Rules
- Dimensionality reduction

Challenges of unsupervised learning

While unsupervised learning has many benefits, some challenges can occur when it allows machine learning models to execute without any human intervention. Some of these challenges can include:

- Computational complexity due to a high volume of training data
- Longer training times
- Higher risk of inaccurate results
- Human intervention to validate output variables
- Lack of transparency into the basis on which data was clustered

COUNTING RATIONAL POINTS ON $\mathcal{M}_g$

Let g be an integer $g \geq 2$ and $\mathcal{M}_g$ the moduli space of smooth, irreducible curves of genus g. $\mathcal{M}_g$ is an algebraic variety of dimension $3g - 3$. Hence, $\mathcal{M}_g$ is embedded in $\mathbb{P}^{3g-2}$. Here are a few facts:

- If $\mathcal{X}$ is a curve defined over $\mathbb{Q}$, then the corresponding moduli point is also defined over $\mathbb{Q}$, say $\mathfrak{p} \in \mathcal{M}_g(\mathbb{Q})$.
- If $\mathfrak{p} \in \mathcal{M}_g(\mathbb{Q})$ it is not true that we can find a curve $\mathcal{X}$ defined over $\mathbb{Q}$ corresponding to $\mathfrak{p}$. In other words, $\mathcal{M}_g$ os a coarse moduli space.
- Let $\mathfrak{p} \in \mathcal{M}_g(\mathbb{Q})$ and F a minimal field of definition of $\mathfrak{p}$. Then F is a number field and $F/\mathbb{Q}$ is called the obstruction of $\mathfrak{p}$.

PROBLEM

Can we somehow count the rational points in $\mathcal{M}_g$? Moreover, can we count how many of them have non-trivial obstruction.

Let $\mathfrak{p} \in \mathcal{M}_g$. We call the **moduli height** $\mathfrak{h}(\mathfrak{p})$ the usual height $H(P)$ in the projective space $\mathbb{P}^{3g-2}$. Obviously, when we fix some coordinate in $\mathcal{M}_g$, $\mathfrak{h}(\mathfrak{p})$ is an invariant of the curve.

LEMMA

For any constant $c \geq 1$, degree $d \geq 1$, and genus $g \geq 2$ there are finitely many curves $\mathcal{X}_g$ defined over the ring of integers $\mathcal{O}_K$ of an algebraic number field K such that $[K : \mathbb{Q}] \leq d$ and $\mathfrak{h}(\mathcal{X}_g) \leq c$.

Genus 2 curves

Every genus 2 curve has equation

$$Y^2Z^4 = F(X,Z) = a_6X^6 + a_5X^5Z + \cdots + a_1XZ^5 + a_0Z^6$$

Bolza determined invariants of binary sextics (Bolza, 1887) in char $k \neq 2$ and Igusa extended it for char $k = 2$. Hence, in the literature such invariants are mistakenly known as Igusa invariants.

$$
\begin{aligned}
J_2 :=& -240a_0a_6 + 40a_1a_5 - 16a_2a_4 + 6a_3^2\\
J_4 :=& 48a_0a_4^3 + 48a_2^3a_6 + 4a_2^2a_4^2 + 1620a_0^2a_6^2 + 36a_1a_3^2a_5 - 12a_1a_3a_4^2 - 12a_2^2a_3a_5 + 300a_1^2a_4a_6 + 300a_0a_5^2a_2\\
&+ 324a_0a_6a_3^2 - 504a_0a_4a_2a_6 - 180a_0a_4a_3a_5 - 180a_1a_3a_2a_6 + 4a_1a_4a_2a_5 - 540a_0a_5a_1a_6 - 80a_1^2a_5^2\\
J_6 :=& -a_5^2a_4^2a_2^2 + 1600a_1^3a_5a_4a_6 + 1600a_1a_5^3a_0a_2 - 2240a_1^2a_5^2a_0a_6 + 20664a_0^2a_4a_6^2a_2 - 640a_0a_4a_2^2a_5^2 - 18600a_0a_4a_1^2a_6^2 + 76a_1a_3a_2a_4^3 - 198a_1a_3^3a_2a_6\\
&+ 26a_1a_3a_2^2a_5^2 + 616a_2^3a_5a_1a_6 + 28a_1a_4^2a_2^2a_5 - 640a_1^2a_4^2a_2a_6 + 26a_1^2a_4^2a_3a_5 + 616a_1a_4^3a_0a_5 + 59940a_0^2a_5a_6^2a_1 + 330a_0a_5^2a_3^2a_2 + 8a_2^2a_3^2a_4^2 - 24a_2^2a_3^2a_5\\
&+ 60a_2^3a_3^2a_6 + 60a_0a_4^3a_3^2 - 192a_2^3a_0a_6^2 - 320a_2^4a_4a_6 + 176a_1^2a_5^2a_3^2 + 2250a_1^3a_3a_6^2 - 900a_2^2a_1^2a_6^2 - 900a_0^2a_5^2a_4^2 - 10044a_0^2a_6^2a_3^2 + 162a_0a_6a_3^4 - 36a_2^4a_5^2\\
&- 36a_1^2a_4^4 + 76a_2^3a_2 - 320a_1^3a_5^3 + 484a + 492a_0a_4^2a_2a_3a_5 + 492a_0a_4^2a_2a_3a_5 + 3060a_0^2a_4a_6a_3a_5 - 468a_0a_4a_3^2a_2a_6 + 3472a_0a_4a_2a_5a_1a_6 + 492a_1a_3a_2^2a_4a_6\\
&- 238a_1a_3^2a_2a_4a_5 + 1818a_1a_3^2a_0a_6a_5 - 876a_2^2a_0a_6a_3a_5 - 3a_5 - 198a_0a_4a_3^3a_5 + 330a_1^2a_3^2a_6a_4 + 72a_1a_3^4a_5 - 24a_1a_3^3a_4^2 + 2250a_0^2a_5^3a_3 - 1860a_1a_4a_0a_5^2a_3\\
&+ 3060a_1a_3a_0a_6^2a_2 - 876a_0a_4^2a_1a_6a_3 - 1860a_1^2a_3a_2a_5a_6 - 18600a_0^2a_5^2a_6a_2 - 24a_2^3a_4^3 - 119880a_0^3a_6^3\\
J_{10} :=& a_6^{-1} Res_X \left(f, \frac{\partial f}{\partial X} \right)
\end{aligned}
$$

Two genus 2 curves $\mathcal{X}$ and $\mathcal{X}'$ are isomorphic over k if and only if exists $\lambda \in k^\star$ such that $J_{2i}(\mathcal{X}) = \lambda^{2i} J_{2i}(\mathcal{X}')$.

GENUS 2 CURVES

Let the space of all tuples (J_2, J_4, J_6, J_{10}) be S. Define the following relation in S as follows. Two tuples

$$(J_2, J_4, J_6, J_{10}) \sim (J_2', J_4', J_6', J_{10}') \iff \exists \lambda \in k^\star,\ (J_2, J_4, J_6, J_{10}) = \left(\lambda^2 J_2', \lambda^4 J_4', \lambda^6 J_6', \lambda^{10} J_{10}'\right)$$

Set of equiv. classes is called a weighted projective space denoted by $\mathbb{WP}_{(2,4,6,10)}(k)$. It is embedded into $\mathbb{P}^3$:

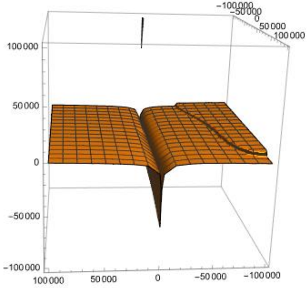
$$\text{Veronese embedding:} \quad \mathbb{WP}_{(2,4,6,10),k} \to \mathbb{P}^3_k, \quad \text{via} \quad [J_2 : J_4 : J_6 : J_{10}] \to \left[J_2^{30} : J_4^{15} : J_6^{10} : J_{10}^6\right].$$

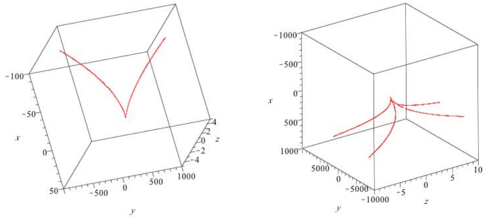
Since $J_{10} \neq 0$, then $[J_2^{30} : J_4^{15} : J_6^{10} : J_{10}^6] \equiv \left[\frac{J_2^{30}}{J_{10}^6} : \frac{J_4^{15}}{J_{10}^6} : \frac{J_6^{10}}{J_{10}^6} : 1\right]$. Thus, two curves are isomorphic iff they have the same absolute invariants $j_1 := \frac{J_2^{30}}{J_{10}^6}$, $j_2 := \frac{J_4^{15}}{J_{10}^6}$, $j_3 := \frac{J_6^{10}}{J_{10}^6}$. To avoid high degrees sometimes different invariants have been used, where $i_1 = \frac{J_4}{J_2^2}$, $i_2 = \frac{J_2J_4 - J_6}{J_2^3}$, $i_3 = \frac{J_{10}}{J_2^5}$, but they are not defined everywhere in $\mathcal{M}_2$.

Wouldn't it make more sense to keep track of only tuples (J_2, J_4, J_6, J_{10}) instead?

In (Shaska et al., 2020) we introduce normalized points in $\mathbb{WP}_{2,4,6,10}(\mathbb{Q})$ which uniquely determine the minimal representative of $[J_2 : J_4 : J_6 : J_{10}]$ and and sort them according to their **weighted heights**.

INCLUSION AMONG THE LOCI





$\mathcal{M}_2$ as a projective space is the set of (affine) points (i_1, i_2, i_3) or projective points $[J_2^{30} : J_4^{15} : J_6^{10} : J_{10}^{6}]$.

GENUS 2: A CASE STUDY

Input: a sextic polynomial $f(t)$

`J30`	J_{30}: the V_4-locus
`Igusa`	Igusa invariants $[J_2, J_4, J_6, J_{10}]$
`RatMod`	Rational model of the curve over when such model exists.
`RatModMe`	Rational model over $\mathbb{Q}$, when such model exists, as in Mestre (Mestre, 1991)
`height`	Height of the sextic
`EquivBin`	Checks if sextics are equivalent
`RatModTable`	Rational Model from the Table of minimal models
`MinField`	Minimal field of definition
`Info`	Displays information about the curve $y^2 = f(t)$
`RatForm`	Rational Model from Malmendier/Shaska (Malmendier and Shaska, 2019)

Input: the moduli point (J_2, J_4, J_6, J_{10})

`J30_j`	J_{30}, V_4-locus
`L_D4`	Locus of curves with group D_4
`L_D6`	Locus of curves with group D_6
`AutGroup`	Automorphism group of the curve
`ModHeight`	Modular height

Moduli Space

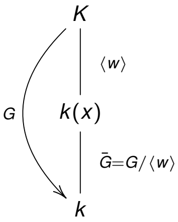
`curves_moduli`	Computes the number of rational points of height $\mathfrak{h}$ in the moduli space and how many of those have a rational model
`NumbCurvMod`	number of rational points of moduli height $\mathfrak{h}$, how many of them have a rational model over $\mathbb{Q}$, how many of them have automorphisms
`moduli_points`	Computes the number of rational points of height $\mathfrak{h}$ in the moduli space
`MoPtsCurvAut`	Moduli points with automorphisms

Creating the databases

`Curves(h, L)`	Creates the dictionary $\mathcal{L}_1$ of curves with height h
`CurvesAut(h, L)`	Creates the dictionary $\mathcal{L}_2$ of curves with automorphisms
`CurvHe`	Number of curves with height h
`CurvHeW(h, w)`	Number of curves with height h and w
`NCWT(h, w)`	Number of curves with height h and twists w
`CurvesTabOverQ(h, w)`	Counts the number of curves over $\mathbb{Q}$, including twist, for given height.

AUTOMORPHISMS OF CURVES

Let $\mathcal{X}_g$ denote an algebraic curve of genus $g \geq 2$, defined over $\bar{k} = k$, and $K = k(\mathcal{X}_g)$. The automorphism group $\mathrm{Aut}(\mathcal{X}_g)$ of $\mathcal{X}_g$ is the group of automorphisms of K defined over k. From Riemann-Hurwitz formula we derive what is now known as the **Hurwitz bound**. $|\mathrm{Aut}(\mathcal{X}_g)| \leq 84\,(g-1)$

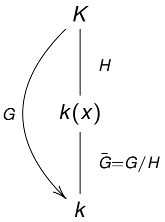


Let $\mathcal{X}_g$ be hyperelliptic. Then, $\mathcal{X}_g : \; y^2 = f(x)$, where $\deg f = 2g+2$. Let $G = \mathrm{Aut}(\mathcal{X}_g)$ and $w : (x, y) \to (-x, y)$ be the hyperelliptic involution. Then, w is central in G.

The group $\bar{G} := G/\langle w \rangle$ is called the **reduced automorphism group** of $\mathcal{X}_g$. Hence, $\bar{G}$ is finite and

$$\bar{G} \hookrightarrow \mathrm{Aut}(k(x)/k) \cong PGL(2, k)$$

Hence, G is a degree 2 central extensions of $\bar{G}$ and $\bar{G} \cong C_n, D_n, A_4, S_4, A_5$.



Let $\mathcal{X}_g$ be a curve and $H := \langle \tau \rangle$ be a **normal cyclic subgroup of order** n of $G = \mathrm{Aut}(\mathcal{X}_g)$ which fixes a genus 0 space $\mathcal{X}_g/H$. The group $\bar{G} = G/H$ is called the **reduced automorphism group** of $\mathcal{X}_g$.

We call such curves **superelliptic curves**. They have affine equation $y^n = f(x)$, for some polynomial $f(x)$. Then $\tau : (x, y) \to (x, \zeta y)$, where $\zeta^n = 1$.

STRATIFICATION OF $\mathcal{M}_g$ BASED ON AUTOMORPHISMS

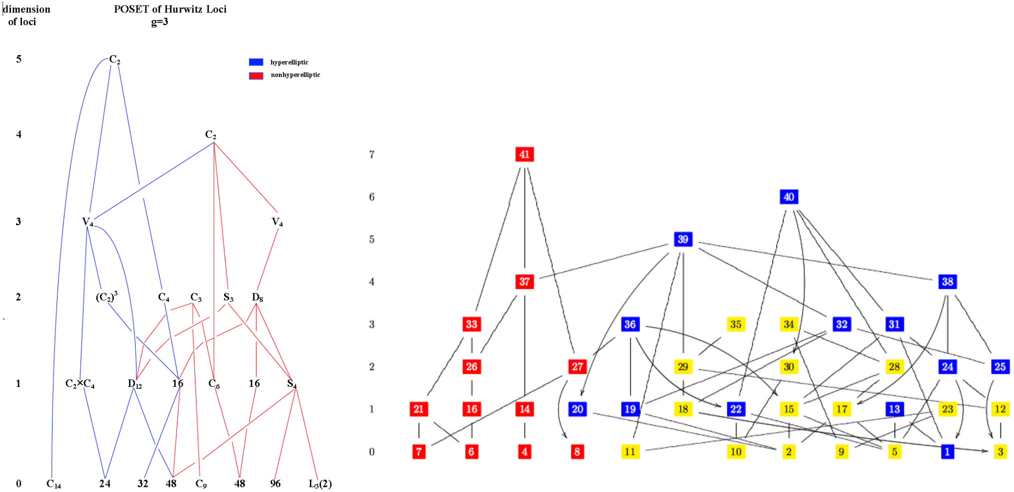


FIGURE: Inclusion of the loci in $\mathcal{M}_g$ for genus 3 and 4; see (Magaard et al., 2002)

FROM HYPERELLIPTIC TO SUPERELLIPTIC

In (Malmendier and Shaska, 2019) we make the case that superelliptic loci are the building blocks of understanding the general theory of $\mathcal{M}_g$ (no surprise here!) Think of Galois theory and cyclic extensions.

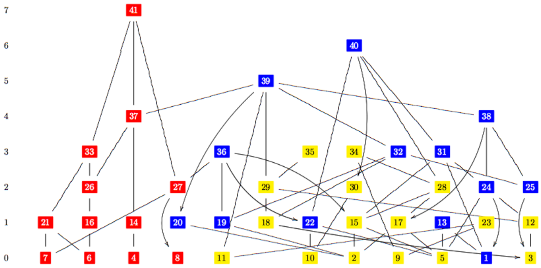


Table 2. Equations of genus 4 superelliptic curves

#	dim	aut	equation
23	1	(5,1)	$y^5 = x(x-1)(x-\lambda)$
9	0	(15,1)	$y^5 = x^3 - 1$
11	0	(10,2)	$y^5 = x(x^2-1)$
34	3	(3,1)	$y^3 = x(x-1)(x-\alpha_1)(x-\alpha_2)(x-\alpha_3)$
28	2	(6,2)	$y^3 = (x^2-1)(x^2-\alpha_1)(x^2-\alpha_2)$
15	1	(18,3)	$y^3 = x^6 + \lambda x^3 + 1$
17	1	(12,5)	$y^3 = (x^2-1)(x^4-\lambda x^2+1)$
2	0	(72,42)	$y^3 = x(x^4-1)$
5	0	(36,12)	$y^3 = x^6 - 1$
35	3	(3,1)	$y^3 = (x^2-2)(x^4+bx^2+cx+d)$
29	2	(6,1)	$y^3 - 1 = x\left(x^5 + (b-2)x^3 + x^3c - (2b+1/2)x - 2c\right)$
12	1	(36,10)	$y^3 - 1 = x^6 + \lambda x^3 + 1$
18	1	(12,4)	$y^3 - 1 = (x^2-1)(x^2-\alpha_1)(x^2-\alpha_2)$
3	0	(72,40)	$y^3 - 1 = x^6 - 1$
22	1	(6,2)	$y^6 = x(x-1)(x-\alpha)$
30	2	(4,1)	$y^4 = x^2(x-1)(x-\alpha_1)(x-\alpha_2)$
10	0	(12,2)	$y^4 = x^2(x^3-1)$
41	7	(2,1)	$y^2 = f(x),\ \deg f = 9, 10$
37	4	(4,2)	$y^2 = x^{10} + a_1x^8 + a_2x^6 + a_3x^4 + a_4x^2 + 1$
33	3	(4,1)	$y^2 = x(x^8 + a_1x^6 + a_2x^4 + a_3x^2 + 1)$
26	2	(8,3)	$y^2 = x(x^4 + \lambda_1 x^2 + 1)(x^4 + \lambda_2 x^2 + 1)$
27	2	(6,2)	$y^2 = x^9 + a_1x^6 + a_2x^3 + 1$
4	0	(40,8)	$y^2 = x^{10} - 1$
6	0	(32,19)	$y^2 = x(x^8-1)$
7	0	(24,3)	$y^2 = x(x^4-1)(x^4 + 2i\sqrt{3}x^2 + 1)$
8	0	(18,2)	$y^2 = x^9 + 1$
21	1	(8,4)	$y^2 = x(x^4-1)(x^4 + \lambda x^2 + 1)$
14	1	(20,4)	$y^2 = x^{10} + \lambda x^5 + 1$
16	1	(16,7)	$y^2 = x(x^8 + \lambda x^4 + 1)$

Genus $g = 4$. From 41 total cases, only 13 are non-superelliptic.

Geometric Invariant Theory (GIT)

$\mathcal{M}_g$ can be compactified by adding "boundary" points (semistable curves). The compactification is denoted by $\overline{\mathcal{M}}_g$. The dimension of $\mathcal{M}_g$ is $\dim(\mathcal{M}_g) = 3g - 3$. Indeed it is irreducible (Deligne/Mumford). The hyperelliptic sublocus $\mathcal{H}_g$ in $\mathcal{M}_g$ has dimension $2g - 1$ and when $g = 2$ we get $\mathcal{H}_g = \mathcal{M}_g$.

From GIT, for any curve $\mathcal{X}$ the corresponding moduli point is a set of invariants $(\xi_0, \dots, \xi_d)$. How can we explicitly describe points in $\mathcal{M}_g$?

- If $\mathcal{X}$ is a curve defined over $\mathbb{Q}$, then the corresponding moduli point is also defined over $\mathbb{Q}$, say $\mathfrak{p} \in \mathcal{M}_g(\mathbb{Q})$.
- If $\mathfrak{p} \in \mathcal{M}_g(\mathbb{Q})$ it is not true that we can find a curve $\mathcal{X}$ defined over $\mathbb{Q}$ corresponding to $\mathfrak{p}$. In other words, $\mathcal{M}_g$ is a coarse moduli space.
- Let $\mathfrak{p} \in \mathcal{M}_g(\mathbb{Q})$ and F a minimal field of definition of $\mathfrak{p}$. Then F is a number field and $F/\mathbb{Q}$ is called the obstruction of $\mathfrak{p}$.

Theorem (Shimura)

If $\mathfrak{p} \in \mathcal{M}_g$ and Aut$(\mathfrak{p}) = \{id\}$, then the field of moduli $k(\mathfrak{p})$ is a field of definition.

Hence, a generic point of $\mathcal{M}_g$ has no obstruction. However, the hyperelliptic locus is different. The generic curve $\mathfrak{p} \in \mathcal{H}_g$ ($|\text{Aut}(\mathfrak{p})| = 2$) is not necessarily defined over $k(\mathfrak{p})$, but it is defined over a quadratic extension of $k(\mathfrak{p})$.

Problem

Can we somehow count the rational points in $\mathcal{M}_g$? Moreover, can we count how many of them have non-trivial obstruction.

WEIGHTED PROJECTIVE SPACES

Let k be a field of characteristic zero and $\mathbf{w} = (q_0, \ldots, q_n) \in \mathbb{Z}^{n+1}$ 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

$$\lambda \star (x_0, \ldots, x_n) = \left(\lambda^{q_0} x_0, \ldots, \lambda^{q_n} x_n\right), \qquad \text{for } \lambda \in k^*.$$

The quotient of this action is called a weighted projective space and denoted by $\mathbb{WP}^n_{(q_0,\ldots,q_n),k}$. It is the projective variety $Proj\,(k[x_0, ..., x_n])$ associated to the graded ring $k[x_0, \ldots, x_n]$ where the variable x_i has degree q_i for $i = 0, \ldots, n$. Denote a point $\mathfrak{p} \in \mathbb{WP}^n_w(k)$ by $\mathfrak{p} = [x_0 : x_1 : \cdots : x_n]$.

For an ordered tuple of integers $\mathbf{x} = (x_0, \ldots, x_n) \in \mathbb{Z}^{n+1}$, whose coordinates are not all zero, the weighted greatest common divisor with respect to the set of weights $\mathbf{w}$ is the largest integer d such that

$$d^{q_i} \mid x_i, \quad \text{for all } i = 0, \ldots, n, \qquad \text{denoted by} \qquad wgcd(x_0, \ldots, x_n) = wgcd(\mathbf{x})$$

An integral point $\mathbf{x} = [x_0 : \cdots : x_n]$ such that $wgcd(x_0, \ldots, x_n) = 1$ is called **normalized**.

The **absolute weighted gcd** of $\mathbf{x} = (x_0, \ldots, x_n)$ with respect to $\mathbf{w}$ is the largest **real number** d such that

$$d^{q_i} \in \mathbb{Z} \quad and \quad d^{q_i} \mid x_i, \quad \text{for all } i = 0, \ldots n.$$

We denote it by $\overline{wgcd}(\mathbf{x})$. An integer tuple $\mathbf{x}$ with $\overline{wgcd}(\mathbf{x}) = 1$ is called **absolutely normalized**. For $\mathfrak{p} \in \mathbb{WP}^n_{\mathbf{w}}(k)$, points

$$\mathbf{y} = \frac{1}{wgcd(\mathfrak{p})} \star \mathfrak{p}, \quad \text{and} \quad \bar{\mathbf{y}} = \frac{1}{\overline{wgcd}(\mathfrak{p})} \star \mathfrak{p},$$

are integral and normalized (resp. integral and absolutely normalized).

Veronese embedding

A weighted space $\mathbb{P}^n_{\mathbf{w},k}$ is called **reduced** if $\gcd(q_0, \cdots, q_n) = 1$. It is called **normalized** or **well-formed** if

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

Proposition

Given any tuple of weights $\mathbf{w} = (q_0, \ldots, q_n)$, the following hold:

(i) *Any weighted projective space $\mathbb{P}^n_{\mathbf{w},k}$ is isomorphic to $\mathbb{P}^n_{\mathbf{w}',k}$, where $\mathbf{w}'$ is a reduced tuple of weights.*

(ii) *If $\mathbb{P}^n_{\mathbf{w},k}$ is reduced and $d_i = \gcd(q_0, \cdots, \hat{q}_i, \cdots, q_n)$ for $0 \leq i \leq n$, then $\mathbb{P}^n_{\mathbf{w},k} \cong \mathbb{P}^n_{\mathbf{w}',k}$ with $\mathbf{w}' = \left(\frac{q_0}{d_i}, \ldots, \frac{q_{i-1}}{d_i}, q_i, \frac{q_{i+1}}{d_i}, \ldots, \frac{q_n}{d_i}\right)$.*

(iii) *Any weighted projective space is isomorphic to a reduced and well-formed one.*

(iv) *If $\mathbf{w}$ is reduced and all of m/q_i are coprime, where*

$$m = \operatorname{lcm}(q_0, \cdots, q_i),$$

then $\mathbb{P}^n_{\mathbf{w},k}$ is isomorphic to $\mathbb{P}^n_k$ by the following isomorphism:

$$\begin{aligned} \phi_m : \mathbb{P}^n_{\mathbf{w},k} &\longrightarrow \mathbb{P}^n_k, \\ \phi_m([x_0, \ldots, x_n]) &= [x_0^{m/q_0}, x_1^{m/q_1}, \ldots, x_n^{m/q_n}]. \end{aligned} \tag{1}$$

Let $\mathbf{w} = (q_0, \dots, q_n)$ be a set of weights and $\mathfrak{p} \in \mathbb{WP}^n(\bar{k})$ a point such that $\mathfrak{p} = [x_0, \dots, x_n]$.

Can we introduce a height function on $\mathbb{WP}^n_{\mathbf{w}}(k)$?

In (Shaska et al., 2020) we define the **weighted multiplicative height** of $\mathfrak{p}$ and the **logarithmic weighted height** are

$$\mathcal{S}(\mathfrak{p}) := \prod_{v \in M_k} \max \left\{ |x_0|_v^{\frac{n_v}{q_0}}, \dots, |x_n|_v^{\frac{n_v}{q_n}} \right\} \quad \text{and} \quad \log \mathcal{S}(\mathfrak{p}) := \log \mathcal{S}_k(\mathfrak{p}) = \sum_{v \in M_k} \max_{0 \le j \le n} \left\{ \frac{n_v}{q_j} \cdot \log |x_j|_v \right\}.$$

- $\mathcal{S}_k(\mathfrak{p})$ does not depend on the choice of coordinates of $\mathfrak{p}$.
- $\mathcal{S}_k(\mathfrak{p}) \geq 1$.
- If $\mathfrak{p}$ is normalized in $K = \mathbb{Q}(\overline{wgcd}(\mathfrak{p}))$, then $\mathcal{S}_K(\mathfrak{p}) = \mathcal{S}_\infty(\mathfrak{p}) = \max_{0 \le i \le n} \left\{ |x_i|_\infty^{\frac{n_\nu}{q_i}} \right\}$
- If L/K is a finite extension, then $\mathcal{S}_L(\mathfrak{p}) = \mathcal{S}_K(\mathfrak{p})^{[L:K]}$.
- Let $m = \text{lcm}(q_0, q_1 \cdots, q_n)$. Then $\mathcal{S}_k(\mathbf{x}) = H_k(\phi_m(\mathbf{x}))^{\frac{1}{m}}$ and $\mathfrak{s}_k(\mathbf{x}) = \frac{1}{m} \cdot h_k(\phi_m(\mathbf{x}))$,, for all $\mathbf{x} \in \mathbb{P}^n_{\mathbf{w}}(k)$, where ϕ_m is the Veronese map.

In $\mathbb{WP}^n(\overline{\mathbb{Q}})$ we define the **absolute (multiplicative) weighted height** $\tilde{\mathfrak{h}} : \mathbb{WP}^n(\bar{\mathbb{Q}}) \to [1, \infty)$

$$\tilde{\mathfrak{h}}(\mathfrak{p}) = \mathcal{S}_K(\mathfrak{p})^{1/[K:\mathbb{Q}]},$$

where $\mathfrak{p} \in \mathbb{WP}^n(K)$, for any K which contains $\mathbb{Q}(\overline{wgcd}(\mathfrak{p}))$.

- The height is invariant under Galois conjugation. For $\mathfrak{p} \in \mathbb{WP}^n(\overline{\mathbb{Q}})$ and $\sigma \in G_{\mathbb{Q}}$, $\mathcal{S}(\mathfrak{p}^\sigma) = \mathcal{S}(\mathfrak{p})$
- For any point $\mathfrak{p} \in \mathbb{WP}^n_{\mathbf{w}}(\overline{\mathbb{Q}})$, we have $[\mathbb{Q}(\mathfrak{p}) : \mathbb{Q}] \leq q \cdot [\mathbb{Q}(\phi(\mathfrak{p})) : \mathbb{Q}]$
- (Northcott) Let c_0 and d_0 be constants and $\mathbb{WP}^n_w(\overline{\mathbb{Q}})$ the weighted projective space with weights $\mathbf{w} = (q_0, \ldots, q_n)$. Then the set

  $$\{\mathfrak{p} \in \mathbb{WP}^n_w(\overline{\mathbb{Q}}) : \mathcal{S}_{\mathbb{Q}}(\mathfrak{p}) \leq c_0 \text{ and } [\mathbb{Q}(\mathfrak{p}) : \mathbb{Q}] \leq d_0\}$$

  contains only finitely many points.
- There are finitely many points $\mathfrak{p} \in \mathbb{WP}^n_{\mathbf{w}}(\overline{\mathbb{Q}})$ of bounded height, $\{\mathfrak{p} \in \mathbb{WP}^n_w(\overline{\mathbb{Q}}) : \mathcal{S}_{\bar{\mathbb{Q}}}(\mathfrak{p}) \leq c_0\}$ is finite.

How good are the weighted heights?

They solve the sorting problem in weighted moduli spaces.

CM-curves

Problem

Let $\mathcal{X}$ be a smooth, algebraic curve of genus $g \geq 1$, defined over a field k, and its Jacobian $\mathrm{Jac}_k(\mathcal{X})$. Determine when $\mathrm{Jac}\,\mathcal{X}$ has complex multiplication (CM).

They were first studied by M. Deuring (Deuring, 1941; 1949) for elliptic curves and generalized to Abelian varieties by (Shimura and Taniyama, 1961).

$\mathcal{X}$ is said to have complex multiplication when $\mathrm{Jac}_k(\mathcal{X})$ is of CM-type.

CM is a property of the Jacobian, so it is an invariant of $\mathcal{X}$.

Is there anything special about the points in $\mathcal{M}_g$ for which $\mathrm{Jac}_k(\mathcal{X})$ is of CM-type?

F. Oort asked if curves with *many automorphisms* are all of CM-type?

The answer to this question is negative.

Theorem ((Obus and Shaska, 2021))

If $\mathcal{X}$ is a superelliptic curve with many automorphisms, then $\mathcal{X}$ is one of the curves on the following Table. $\mathcal{X}$ is has CM if and only if it is one of the cases in Table **??**.

THINGS LEFT TO DO!

Will this work? Are we able to get reliable results?

Using the idea of weighted projective spaces and weighted heights we can do a lot of computation in $\mathcal{M}_g$ and arithmetic statistics.

The first case $\mathcal{M}_2$:

- Using the universal equation in (Malmendier and Shaska, 2017) and the database of genus 2 curves, study the distribution of rational points in $\mathcal{M}_2$ for which the field of moduli is a field of definition.
- Using (van Wamelen, 1999) study a distribution of CM-curves in $\mathcal{M}_2$.
- In Beshaj's thesis a constant was determined such that

$$\text{naive height } (C) \leq \lambda \cdot \text{ moduli height } (C)$$

  for some constant λ. Determine from the database a smaller universal constant λ.

The case $\mathcal{M}_3$:

The case of $\mathcal{M}_3$ is also well understood, but computations are rather more difficult. Some databases for hyperelliptic and non-hyperelliptic genus 3 curves do exist:

- Based on weighted heights sort all points in $\mathcal{M}_3$.
- Based on stratification of $\mathcal{M}_3$, determine parametric equations for each case (including non-superelliptic cases)
- Determine a distribution of points in $\mathcal{M}_3$ for which the field of moduli is not a field of definition.
- Determine a distribution of CM-points in $\mathcal{M}_3$

Things left to do!

Higher genus

- We can determine the stratification of the $\mathcal{M}_g$ for any $g \geq 2$
- We can determine all superelliptic loci, including parametric equations, dimension of each locus, inclusions among loci.
- Completing the above gives a precise understanding of about 80% of loci
- Superelliptic curves correspond to binary forms. So their isomorphism classes correspond to invariants of binary forms (another classical problem where ML can be used).
- From GIT we know that, in general, there are invariants describing points in $\mathcal{M}_g$. Very little is known about explicitly writing them down for higher genus.

Goal:

We hope to have explicit results by the end of the Summer for $g = 4$, building on databases from $g = 2, 3$.

REFERENCES,

Alessandretti, Laura, Andrea Baronchelli, and Yang-Hui He. 2019. *Machine learning meets number theory: The data science of birch-swinnerton-dyer*.

Beshaj, L., R. Hidalgo, S. Kruk, A. Malmendier, S. Quispe, and T. Shaska. 2018. *Rational points in the moduli space of genus two*, Higher genus curves in mathematical physics and arithmetic geometry, pp. 83–115. MR3782461

Bolza, Oskar. 1887. *On binary sextics with linear transformations into themselves*, Amer. J. Math. **10**, no. 1, 47–70. MR1505464

Bombieri, E. and W. Gubler. 2006. *Heights in Diophantine geometry*, New Mathematical Monographs, vol. 4, Cambridge University Press,.

Bost, Jean-Benoit. 1994. *Semi-stability and heights of cycles*, Invent. Math. **118**, no. 2, 223–253. MR1292112

Cardona, Gabriel and Jordi Quer. 2005. *Field of moduli and field of definition for curves of genus 2*, Computational aspects of algebraic curves, pp. 71–83. MR2181874

Cornalba, Maurizio and Joe Harris. 1988. *Divisor classes associated to families of stable varieties, with applications to the moduli space of curves*, Ann. Sci. École Norm. Sup. (4) **21**, no. 3, 455–475. MR974412

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

Deligne, P. and D. Mumford. 1969. *The irreducibility of the space of curves of given genus*, Inst. Hautes Études Sci. Publ. Math. **36**, 75–109. MR262240

Deuring, Max. 1941. *Die Typen der Multiplikatorenringe elliptischer Funktionenkörper*, Abh. Math. Sem. Hansischen Univ. **14**, 197–272. MR5125

______. 1949. *Algebraische Begründung der komplexen Multiplikation*, Abh. Math. Sem. Univ. Hamburg **16**, no. nos., nos. 1-2, 32–47. MR32601

Hindry, M. and J. Silverman. 2000. *Diophantine geometry*, Graduate Texts in Mathematics, vol. 201, Springer-Verlag, New York.

Magaard, K., T. Shaska, S. Shpectorov, and H. Völklein. 2002. *The locus of curves with prescribed automorphism group*, Sūrikaisekikenkyūsho Kōkyūroku **1267**, 112–141. Communications in arithmetic fundamental groups (Kyoto, 1999/2001). MR1954371

Malmendier, A. and T. Shaska. 2019. *From hyperelliptic to superelliptic curves*, Albanian J. Math. **13**, no. 1, 107–200. MR3978315

Malmendier, Andreas and Tony Shaska. 2017. *A universal genus-two curve from Siegel modular forms*, SIGMA Symmetry Integrability Geom. Methods Appl. **13**, Paper No. 089, 17. MR3731039

REFERENCES,

Mestre, Jean-François. 1991. *Construction de courbes de genre* 2 *à partir de leurs modules*, Effective methods in algebraic geometry (Castiglioncello, 1990), pp. 313–334. MR1106431

Mumford, David and John Fogarty. 1982. *Geometric invariant theory*, Second, Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas], vol. 34, Springer-Verlag, Berlin. MR719371

Obus, Andrew and Tanush Shaska. 2021. *Superelliptic curves with many automorphisms and CM Jacobians*, Math. Comp. **90**, no. 332, 2951–2975. MR4305376

Salami, Sajad and Tony Shaska. 2022. *Local and global heights on weighted projective varieties and vojta's conjecture*, arXiv.

Shaska, T. 2022. *Reduction of superelliptic riemann surfaces*, Contemporary Math.

Shaska, T., L. Beshaj, and J. Gutierrez. 2020. *Weighted greatest common divisors and weighted heights*, J. Number Theory.

Shimura, Goro and Yutaka Taniyama. 1961. *Complex multiplication of abelian varieties and its applications to number theory*, Publications of the Mathematical Society of Japan, vol. 6, The Mathematical Society of Japan, Tokyo. MR0125113

van Wamelen, Paul. 1999. *Examples of genus two CM curves defined over the rationals*, Math. Comp. **68**, no. 225, 307–320. MR1609658

Vojta, P. 2011. *Diophantine approximation and nevanlinna theory*, Arithmetic geometry., pp. 111–224 (English).

Vojta, Paul. 2012. *Multiplier ideal sheaves, nevanlinna theory, and diophantine approximation*, Number theory, analysis and geometry. in memory of serge lang, pp. 647–658 (English).

Zhang, Shouwu. 1996. *Heights and reductions of semi-stable varieties*, Compositio Math. **104**, no. 1, 77–105. MR1420712

Thank you for your attention!