Parametric Solutions to the Generalised Fermat Equation

Jane Shi

August 30, 2023

Contents

1 Introduction

In this essay, we will present some remarkable work about finding integer solutions to the following diophantine equation (the generalised Fermat equation):

$$AX^p + BY^q + CZ^r = 0, \gcd(X, Y, Z) = 1, XYZ \neq 0 \tag{1}$$

where $p, q, r, A, B, C \in \mathbb{Z}, p, q, r \geq 2$ are fixed and $X, Y, Z \in \mathbb{Z}$ are the unknowns.
Based on the exponents p, q, r, there are three cases:

- $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$: The Euclidean case. The only possibilities of p, q, r, up to permutation, are $\{(3, 3, 3), (2, 3, 6), (2, 4, 4)\}$. This case is well-studied and relates to the problem of finding rational points on elliptic curves.
- $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} > 1$: The spherical case.
  The only possibilities of (p, q, r), up to permutation, are $\{(2, 2, n), (2, 3, 3), (2, 3, 4), (2, 3, 5)\}$ where $n \geq 2$ is an integer. The rest of the essay focuses on the spherical case. One property that makes the spherical case special is that parameterised families of solutions exist in the spherical case. Soon in the Subsection 1.2, we will talk about what parametrised solutions are.
- $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} < 1$: The hyperbolic case.
  Fermat's Last Theorem falls under this case. It was proven by Darmon and Granville in [4] that in this case, (1) has finitely many solutions.

  It is also notable that when setting $A = 1, B = 1, C = -1$, the equation obtained with the constraint that $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} < 1$:
  $$X^p + Y^q - Z^r = 0$$
  only has ten solutions (that have have been found up to now), if we count $1^k + 2^3 = 3^2$ for all k as one single solution. Some of the larger ones include $43^8 + 96222^3 = 30042907^2$ and $1414^3 + 2213459^2 = 65^7$.

In this essay, we are only interested in solving the equation for the spherical case.

1.1 Why is this problem important?

- The famous Fermat's Last Theorem, proven by Sir Andrew Wiles in 1995, is a famous special case of the hyperbolic case of the generalised Fermat Equations. Lots of new mathematics were generated in order to study of Fermat's Last Theorem.
- The studies of this problem under different settings give unexpected connection to other mathematical theory. Later we will see how it relates to the group actions that permute vertices of the platonic solids. Another example is that work by Hellegouarch, Frey, Serre and Ribet shows connections between the Fermat equation to elliptic curves, modular forms, and Galois representation. This is mentioned in [2].
- For number theorists with computational interests, there is a computational aspect to produce the solutions to the generalised Fermat equation. More specifically, later in this essay, we will present an algorithm to compute parametrised solutions in the spherical case. It will be the focus of Section 5.

  Another example is that, of the ten known solutions to the hyperbolic case, the larger ones
  - $33^8 + 1,549,034^2 = 15,613^3$
  - $43^8 + 96,222^3 = 30,042,907^2$
  - $9,262^3 + 15,312,283^2 = 113^7$

  are resulted from Beukers's and Zagier's work using computational methods, as mentioned in [3].

1.2 Main goals and contributions for this essay

The main goal of the paper is to explore how the group of rotations on a platonic solid can be used to generate parametrised solutions to (1) in the spherical case. The main contribution is an implementation of the algorithm in [6]'s work in Python. In addition to the implementation, we will also include technical details of the implementation and a demonstration of how to obtain solutions to the diophantine equation from the outputs.

1.3 Motivations and Background of this Essay

Beukers's theorem states that there exists a finite set of parametrised solutions, such that every solution to (1) is a specialisation of one of the parametrised solutions. We will present and prove Beukers's theorem in a special case.
Since one of the main idea of the essay revolves around the finiteness of families of parametrised solutions, we first introduce the concept parametrised solutions.

Definition 1.1 (Parameterised Solution): A parametrised solution to (1) is a triple of polynomials $P_X, P_Y, P_Z \in \mathbb{Z}[x_1, x_2]$ such that $\gcd(P_X, P_Y, P_Z) = 1$ and it satisfies

$$AP_X(x_1, x_2)^p + BP_Y(x_1, x_2)^q + CP_Z(x_1, x_2)^r = 0 \tag{2}$$

The equality here means the equality of polynomials.
By families of parametrised solutions, we identify two parametrised solutions as the same if they integer specialises to the same set of solutions to the diophantine equation. Later in Section 2, we will see that we can relate two elements in the family of parametrised solution by $\mathrm{SL}_2(\mathbb{Z})$ action.

Definition 1.2 (Integer Specialisations of Parametrised Solution): Given a parametrised solution $P_X, P_Y, P_Z \in \mathbb{Z}[x_1, x_2]$, specifying the value of $x_1, x_2 \in \mathbb{Z}$ results in a solution $(X, Y, Z), X, Y, Z \in \mathbb{Z}$, of (1). This is called an integer specialisation of the parametrised solution.

As shown in [3], an example of parametrised solution can look like this:
The equation is

$$X^2 + 27Y^2 = 4Z^3 \tag{3}$$

A parametrised solution is given by

- $P_X(x_1, x_2) = (x_1 - x_2)(2x_1^2 + 5x_1x_2 + 2x_2^2)$
- $P_Y(x_1, x_2) = x_1x_2(x_1 + x_2)$
- $P_Z(x_1, x_2) = x_1^2 + x_1x_2 + x_2^2$

We can specialize (x_1, x_2) to any pairs of integers. For example, specializing (x_1, x_2) to $(2, 1)$ yields $(X, Y, Z) = (20, 6, 7)$ which is indeed a solution to (3). Setting (x_1, x_2) to any integers would yield infinitely many solutions to (3).
One might ask, how many parametrised solutions exist for an equation with fixed, A, B, C, p, q, r in a spherical case? And how to generate them? This essay will answer these questions.
Both Beukers and Edwards studied the solutions of (1) in the spherical case in the forms of parametrised solutions, yet their approaches are slightly different.

1.3.1 Beukers's work

In Beukers's work [3], he proved the following theorem

Theorem 1.3 (Beukesr's Theorem): Given (1) in spherical case,

- There exists a finite number of families of parameterised solutions S. Furthermore, if (x, y, z) is a solution to (1), then we must be able to find a parameterised solution in S such that (x, y, z) is an inter specialisation of it.
- If we can find one solution to (1), then we can find infinitely many.

Beukers's work involves mathematics that could require more technical background to understand. For example, it includes the Hilbert's 90 theorem and complex reflection polynomials. In this essay, we will not talk about Beukers's theorem in detail, but we will present some ideas briefly and how it relates to Edwards's work in Section 6.

1.3.2 Edwards's work

As a student of Beukers's, Edwards also worked on solving the generalized Fermat equation in the spherical case following Beukers's work in 1998. In the work that Edwards published 2004, Edwards presented a complete solution to a subset of equations of the form (1). In his work, he studied the equations of the form

$$X^2 + Y^3 + dZ^r = 0, r \in \{3, 4, 5\} \tag{4}$$

with $d \neq 0$ an integer. He used a different approach compared to Beukers: his approach is algorithmic. Using invariant theory and Hermite reduction theory, he presented a bound on the coefficients of binary forms such that all the parametrised families of solutions can result from looping through all possibilities within that bound. Since the bound is finite, there are finitely many families of parametrised solutions to (4). Following the proof of the correctness of the bound, he presented an algorithm to explicitly list all the parametrised families of solutions with proof. We will study Edwards's approach in detail.

1.3.3 Comparison of Beukers's work to Edwards's work

- Beukers's approach works for the generalised Fermat equations whereas Edwards's algorithm only proves the Beukers's theorem for a subset of the spherical cases, but not all.
- Edwards's work presents an algorithm to compute all the parametrised solutions, whereas Beukers's work does not present an explicit method to generate solutions.

1.4 Roadmap of the essay

We will first present the background on invariant theory in Section 2, then the background of the geometry of numbers in Section 3. Next, we will present Edwards's approach in Section 4 and how Edwards uses his approach in Section 4 to produce the algorithm in Section 5. Finally, we will briefly talk about Beukers's theorem in 6. [1]

[1] All the main ideas in this essay are understood from the references, with some examples and calculations I produced. I will mention it explicitly when I present examples and calculations that I came up with.

2 Basic Invariant Theory

In this section, we will provide necessary background for understanding how the platonic solids relate to parametrisations of the Fermat equation. In Subsection 2.1, we will give a geometric perspective on the action of $\mathrm{PGL}_2(\mathbb{C})$ on the Riemann Sphere. In Subsection 2.2, we will introduce terminologies associated to binary forms. In Subsection 2.3, we will see the connection.

2.1 The Projective Linear Group, a famous theorem by Klein, and the Platonic Solids

In this subsection, we will introduce the Riemann sphere, the Möbius transformations on it, and a famous result by Klein about rotations on the Riemann sphere.
We first recall Riemann Sphere and its automorphism group. The following definitions and proposition are retrieved from article by [8] and notes by [1].

Definition 2.1 (The Riemann Sphere): The Riemann Sphere $\mathbb{C}_\infty$ is the extended complex plane $\mathbb{C} \cup \{\infty\}$.

There is a bijection between $\mathbb{C}_\infty = \mathbb{C} \cup \{\infty\}$ to the unit sphere of radius 1 in $\mathbb{R}^3$ centered in the origin by the stereographic projection.
The bijection is defined as follows: We embed $\mathbb{C}$ in $\mathbb{R}^3$, by identifying it with the $xy-$plane, mapping $a + bi \in \mathbb{C}, a, b \in \mathbb{R}$ to $(a, b, 0)$ in $\mathbb{R}^3$. We also identify $(0, 0, 1) \in \mathbb{R}^3$ as the north pole N. Given a point P on the unit sphere, if it is N, then it is mapped to ∞ on the Riemann Sphere. Otherwise, it is mapped to the point where the line (NP) intersects with the xy-plane.

Definition 2.2 (Möbius Transformations of the Riemann Sphere): A Möbius Transformation is a map of the form $f : \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\}$, $z \mapsto \frac{az+b}{cz+d}$, with $a, b, c, d \in \mathbb{C}$ and $ad - bc \neq 0$.

It is worth noting that $f(\infty) = \frac{a}{c}, f\left(-\frac{d}{c}\right) = \infty$ and that when $ad - bc = 0$, f maps $\mathbb{C} \cup \{\infty\}$ to one single point a/c, making it not an invertible map.
A rotation on the Riemann Sphere can be viewed as a Möbius transformation. I.e. identify points on $\mathbb{C}_\infty$ with the unit sphere in $\mathbb{R}^3$ by reverse stereographic projection, rotate the sphere, then apply stereographic projection again.
We next introduce the projective linear group $\mathrm{PGL}_2(\mathbb{C})$ and its construction.

Definition 2.3 ($\mathrm{PGL}_2(\mathbb{C})$): $\mathrm{PGL}_2(\mathbb{C})$ is the group $\mathrm{GL}_2(\mathbb{C})/\sim$, where $M_1 \sim M_2, M_1, M_2 \in \mathrm{GL}_2(\mathbb{C})$ if $M_1 = cI_2M_2$ for some $c \in \mathbb{C}$.

This group is important as it acts on the Riemann Sphere and can help us to generate invariant polynomials that will be useful later in our constructions.
Given a Möbius transformation $f : \mathbb{C}_\infty \to \mathbb{C}_\infty, z \mapsto \frac{az+b}{cz+d}$, we notice that $\frac{az+b}{cz+d} = \frac{(\lambda a)z+(\lambda b)}{(\lambda c)z+(\lambda d)}$ for any $\lambda \neq 0, \lambda \in \mathbb{C}$.
We have the following statements:

Proposition 2.4:

- The Möbius functions $f : \mathbb{C}_\infty \to \mathbb{C}_\infty$ form a group under function compositions.
- The map $\phi : \mathrm{PGL}_2(\mathbb{C}) \to \{\text{the group of Möbius transformations}\}$ given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \left(z \mapsto \frac{az+b}{cz+d} \right)$$

  is a group homomorphism, where $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is any representative of an element in $\mathrm{PGL}_2(\mathbb{C})$.

Therefore, we can view $\mathrm{PGL}_2(\mathbb{C})$ as a group of Möbius transformations on the Riemann Sphere.
We quote a famous theorem by Klein in [7]:

Theorem 2.5 (Klein): All finite subgroups of $\mathrm{PGL}_2(\mathbb{C})$ are isomorphic to one of the following:

- C_n, a cyclic group of order n
- D_n, a dihedral group of order $2n$ where $n \geq 2$
- A_4, the tetrahedral group of order 12. This is the group of rotational symmetry (orientation preserving symmetry) of a regular tetrahedron.
- S_4, the octahedral group of order 24. This is the group of rotational symmetry of a regular octahedron.
- A_5, the icosahedral group of order 60. This is the group of rotational symmetry of a regular icosahedron.

We will focus on how the rotations of three platonic solids (the tetrahedron, the octahedron, and the icosahedron) relate to invariant binary forms (which will be defined later), and how these forms will help us generate parametrised solutions.

2.2 Binary forms, Actions on Polynomials and Covariants

In this subsection, we will introduce binary forms, actions on binary forms, and covariants of forms. The ideas are important for the next subsection, where we introduce a special kind of binary forms.

Definition 2.6 (Binary form): A binary form of degree n in variables x_1, x_2, is a homogenous polynomial of degree n in x_1, x_2, written as

$$\sum_{k=0}^{n} \binom{n}{k} a_k x_1^k x_2^{n-k}$$

in the context of this section, $a_k \in \mathbb{C}$.

Definition 2.7 (G-Invariant polynomials): Let $G \subset \mathrm{GL}_2(\mathbb{C})$ be a subgroup.

Let $g = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in G$, then g acts on $\mathbb{C}^2$ by sending (x_1, x_2) to $(a_{11}x_1 + a_{12}x_2, a_{21}x_1 + a_{22}x_2)$.
g also acts on $f(x_1, x_2) \in \mathbb{C}[x_1, x_2]$ by sending $f(x_1, x_2)$ to $g \cdot f = f(g^{-1}(x_1, x_2))$.
Then $\mathbb{C}[x_1, x_2]^G$, the set of $G-$invariant polynomials is set to be the subring of polynomials that are invariant under the action of G.

We associate the symmetry groups acting on the Riemann Sphere with the elements of $G = \mathrm{PGL}_2(\mathbb{C})$ acting on $\mathbb{C}_\infty$ via the stereographic projection.
When G acts on binary forms, we have the following definition:

Definition 2.8 (Covariants): Given $f = \sum_{k=0}^{n} \binom{n}{k} a_k x_1^k x_2^{n-k} \in \mathbb{C}[x_1, x_2]$, a binary form $C(f) \in \mathbb{C}[a_0, \ldots, a_k, x_1, x_2]$, depending on f, is called a covariant binary form if it is a homogenous polynomial and there exists a positive integer p, called the weight of the covariant, such that for all $g \in \mathrm{GL}_2(\mathbb{C})$, we have

$$g \cdot C(f) = \det(g)^p C(g \cdot f)$$

This is an equivalence of polynomials where the left hand side is an element in $\mathrm{GL}_2(\mathbb{C})$ acting on the binary form $C(f)$. The right hand side is a binary form that depends on the form $g \cdot f$, multiplied by a constant.

Note that a covariant depends on the form f, but it has not been explicitly stated in what way.
One example of a covariant is $C(f) = f$ with weight 0. Covariants are important because they are the building blocks to the parametrised families of solutions to (4). We will soon see other examples of covariants.

2.3 Binary Forms Introduced by Klein

In this subsection, we will link the platonic solids (introduced in Subsection 2.1) and covariants (introduced in Subsection 2.2) together. These ideas are introduced by Klein in [7]. We will introduce the binary forms that are used to represent vertices on the platonic solids and some examples of covariants. We will also start to see how to generate parametrised solutions to (1) based on these binary forms and covariants.
We only consider three platonic solids in this essay: the tetrahedron, the octahedron, and the icosahedron. For each platonic solid, we consider it inscribed within $\mathbb{C}_\infty$, the Riemann Sphere. That is, the platonic solids intersects the sphere only at its vertices.
To start, we find a homogeneous polynomial $f \in \mathbb{C}[x_1, x_2]$ such that its roots corresponds of the point of intersections of the platonic solid and C_∞.
The polynomials f are presented by Klein:

Platonic Solid	f
Tetrahedron	$x_2^4 - 2\sqrt{3}x_1^2x_2^2 - x_1^4$
Octahedron	$x_1x_2(x_1^4 - x_2^4)$
Icosahedron	$x_1x_2(x_1^{10} + 11x_1^5x_2^5 - x_2^{10})$

Table 1: Platonic Solids and the Polynomial whose Roots are their Vertices

Here is a brief remark about how these polynomials are obtained. The octahedron case is the easiest to see: after the steregraphic projection, its six vertices corresponds to $0, \infty, \pm 1, \pm i \in \mathbb{C}_\infty$. Consider the polynomial $x_1(x_1+1)(x_1-1)(x_1+i)(x_1-i)$, whose roots are precisely these vertices other than ∞. We obtain

the root ∞ after homogenizing this polynomial, hence we arrive at $x_1x_2(x_1^4 - x_2^4)$. For the tetrahedron case, the equation is established via observing the intersection of the coordinate axes and the sphere, then multiplying out the linear factors and homogenizing. See [7] Part I. Chap II. Section 2 for details. For the Icosahedron case, the derivation is similar except it requires more details in [7] Part I. Chap II. Section 6. Note that the set of vertices are invariant under the corresponding rotation groups specified in Subsection 2.1.

These polynomials are invariant polynomials under the actions. Each invariant polynomial above represents a set of points on the platonic solid that remains the same set after rotating with respect to the group actions. For example, the six vertices on the Octahedron remains the same with elements under the group action of S_4. However, the set of points under invariant polynomials need not to be the vertices of the solids, they can also be the mid-edge points or the points in the center of the faces.

Klein also presented the concept of special ground forms, which are three invariants forms associated to each of the three platonic solids. Not only do the roots of these three invariants forms form a group of points on the platonic solid, if we raise each special ground form to a suitable power, then their linear combination is 0.

For example, the Icosahedron's three associated special ground forms are as follows:

$$\begin{cases} I_{12} : x_1x_2(x_1^{10} + 11x_1^5x_2^5 - x_2^{10}) \\ I_{20} : -(x_1^{20} + x_2^{20}) + 228(x_1^{15}x_2^5 - x_1^5x_2^{15}) - 494x_1^{10}x_2^{10} \\ I_{30} : (x_1^{30} + x_2^{30}) + 522(x_1^{25}x_2^5 - x_1^5x_2^{25}) - 10005(x_1^{20}x_2^{10} + x_1^{10}x_2^{20}) \end{cases}$$

- The first polynomial comes from Table 1. It is the first special ground form corresponding to the vertices of the icosahedron (I_{12}).
- The roots of the second polynomial corresponds to the twenty center of each faces, denoted by I_{20}. It is worthy to point out that the icosahedron and the dodecahedron are duals to each other, hence I_{20} also corresponds to the vertices of the dodecahedron.
- The roots of the third polynomial corresponds to the thirty mid-points of the edegs of the icosahedron, usually denoted by I_{30}.

The above is the special ground forms for the icosahedron. One may ask, we know where I_{12} came from, but how about the polynomials corresponding to I_{20} and I_{30}? Now we will present the special ground forms more generally.

Definition 2.9 (Hessian and the functional determinant): Given f as Table 1, and k the degree of f, the Hessian of f, denoted $H(f)$ is given by

$$H(f) = \left(\frac{1}{k(k-1)}\right) \begin{vmatrix} f_{x_1x_1} & f_{x_1x_2} \\ f_{x_2x_1} & f_{x_2x_2} \end{vmatrix}$$

and the functional determinant, $t(f)$ is given by

$$t(f) = \left(\frac{1}{k(k-2)}\right) \begin{vmatrix} f_{x_1} & f_{x_2} \\ H_{x_1} & H_{x_2} \end{vmatrix}$$

Theorem 2.10: $H(f)$ is a covariant of weight 2 and $t(f)$ is a covariant of weight 3.

Proof : For the proof, please see Section 2.2 in [6]. □

For each of the three platonic solids, f, the Hessian of f, and the functional determinant of f are the three special ground forms. Each of them is also an invariant form. In Classical Invariant theory([7]), it was shown that for the tetrahedron case, the Hessian and the functional determinant are the only two independent covariants.
Here is what $f, H(f), t(f)$ are for each platonic solid:
Tetrahedron[2]:

$$\begin{cases} f = x_1^4 - 2\sqrt{3}x_1^2x_2^2 - x_2^4 \\ H(f)\cdot 3 = -\sqrt{3}x_1^4 - 6x_1^2x_2^2 + \sqrt{3}x_2^4 \\ t(f) = -4(x_1^5x_2 + x_1x_2^5) \end{cases}$$

Octahedron[3]:

$$\begin{cases} f = x_1x_2(x_1^4 - x_2^4) \\ H(f)\cdot 36 = -x_1^8 - 14x_1^4x_2^4 - x_2^8 \\ t(f)\cdot 108 = x_1^{12} - 33(x_1^8x_2^4 + x_1^4x_2^8) + x_2^{12} \end{cases}$$

Icosahedron:

$$\begin{cases} f = x_1x_2(x_1^{10} - 11x_1^5x_2^5 - x_2^{10}) \\ H(f)\cdot 144 = -(x_1^{20} + x_2^{20}) - 228(x_1^{15}x_2^5 - x_1^5x_2^{15}) - 494x_1^{10}x_2^{10} \\ t(f)\cdot 864 = (x_1^{30} + x_2^{30}) - 522(x_1^{25}x_2^5 - x_1^5x_2^{25}) - 10005(x_1^{20}x_2^{10} + x_1^{10}x_2^{20}) \end{cases}$$

Note that the linear combination of invariant forms remains invariant. Klein's relation is a linear combination of some powers of these forms equating zero.
Klein's relation is as follows:

Platonic Solid	Klein Relation[4]
Tetrahedron	$\frac{1}{3\sqrt{3}}f^3 + H(f)^3 + \left(\frac{1}{2}t(f)\right)^2 = 0$
Octahedron	$\frac{1}{432}f^4 + H(f)^3 + \left(\frac{1}{2}t(f)\right)^2 = 0$
Icosahedron	$\frac{1}{1728}f^5 + H(f)^3 + \left(\frac{1}{2}t(f)\right)^2 = 0$

Table 2: Platonic Solids and Klein Relations

The three forms f for each of the three platonic solids are examples of Klein forms, which we will formally define later in section 4.
Indeed, this is starting to look like the equations in (4). Consider the equation $X^2 + Y^3 + dZ^r = 0$, each platonic solid corresponds to an exponent of Z. More specifically, the Tetrahedron, Octahedron and Icosahedron corresponds to $r = 3, r = 4$ and $r = 5$ respectively. However, one notices that the above Klein relations provide parametrisations for only one value of d. We will explore how to extend the Klein relations and the covariants to make the equation work for more general values of d in Section 4.

3 Introduction to the Geometry of Numbers

The main subject of interest in this section is reduction theory, which studies binary forms under the $\mathrm{SL}_2(\mathbb{Z})$ action. These ideas are relevant to our main interest to generate parametrized solutions to (4) as they inspire a method to list all the parametrized solutions up to $\mathrm{GL}_2(\mathbb{Z})$ equivalence.

[2]I calculated this using the `sympy` library in Python
[3]I calculated this using the `sympy` library in Python
[4]I verified this table with the `sympy` in Python

In this section, we will first introduce definitions and results in the reduction theory of binary quadratic forms ($ax_1^2 + bx_1x_2 + cx_2^2$) in Subsection 3.1. Then, we will show how Hermite reduction theory generalizes these ideas to higher-order binary forms (homogenous of order $k \geq 2$) in Subsection 3.2, present some useful results in Subsection 3.3 and comment on how they can be used to generate parametrisations of solutions.
The following definitions, theorem and proposition are retrieved from lecture notes on [9].

3.1 Reduction Theory in Binary Quadratic Forms

In the theory of quadratic binary forms, we can associate each positive definite real binary form with a point - its unique root in $\mathbb{H}$ (the upper half plane of $\mathbb{C}$). Reduced forms are exactly the forms whose associated point lies in the fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$, and it has been shown that every form is $\mathrm{SL}_2(\mathbb{Z})$ equivalent to a reduced form. The reduced forms have a useful property, which is that their coefficients are bounded by a number depending on its discriminant.

Definition 3.1 (Binary quadratic form): A binary quadratic form is a function of the form $ax_1^2 + bx_1x_2 + cx_2^2$ where $a, b, c \in \mathbb{R}$. Equivalently, it is a binary form of degree 2.

Definition 3.2 (Discriminant): The discriminant of a quadratic binary form $f = ax_1^2 + bx_1x_2 + cx_2^2$ is $\mathrm{disc}(f) = b^2 - 4ac$.

Note that since elements in $\mathrm{SL}_2(\mathbb{Z})$ have determinant 1, discriminants stay the same under the $\mathrm{SL}_2(\mathbb{Z})$ actions, defined similarly as the $\mathrm{GL}_2(\mathbb{Z})$ action in Section 2.

Definition 3.3 (Positive definitive quadratic binary forms): A quadratic binary form f is positive definitive if $f(x_1, x_2) > 0$ for all $(x_1, x_2) \in \mathbb{R}^2 \setminus \{(0, 0)\}$.

Note that being positive definitive is equivalent to the condition that $\mathrm{disc}(f) < 0$ and $a > 0$. This is because $\mathrm{disc}(f) > 0$ and $a > 0$ implies $a > 0$ and $c > 0$. This also implies that the quadratic equation we obtain by dehomogenizing (in either x_1 or x_2) has no roots and obtain only positive values for points that are not the origin.

Definition 3.4 (Reduced): A positive definitive quadratic binary form is reduced if $|b| \leq a \leq c$ and that whenever one of the inequalities is an equality, then $b \geq 0$.

Theorem 3.5 (An equivalent condition of being reduced): Consider a form $f = ax_1^2 + bx_1x_2 + cx_2^2$. Let τ be its root in $\mathbb{H}$. Then, f is reduced if and only if τ lies in the fundamental domain of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{H}$, shown below:[5]

$$\{z = x + iy \in \mathbb{H}, x, y \in \mathbb{R} \mid x \in [-1/2, 1/2), |z| > 1, \text{ or } |z| = 1 \text{ and } x \leq 0\}$$

[5]The fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$ defined in [9] differs slightly compared to the definition in [6]. For example, $\frac{1}{2} + i$ belongs to the fundamental region in [6] but not [9].

Proof : See [9]. □

Proposition 3.6: Given a reduced binary quadratic form $f = ax_1^2 + bx_1x_2 + cx_2^2$ with discriminant d then $b^2 \leq |ac| \leq |d|/3$ when $ac \neq 0$.

Proof : From $|b| \leq a \leq c$, we can obtain $4b^2 \leq ac = b^2 - D$. The proposition follows. □

The above proposition provides a bound on the coefficients for reduced binary forms. Although the above definitions and theorems will not be used later in this essay, it provides some motivation for how terms in Hermite reduction theory are defined.

3.2 Hermite Reduction Theory

Hermite reduction theory draws parallels to the ideas above to higher-order binary forms. We will introduce Hermite determinant for higher-order forms, which plays the role of the discriminant. We will also introduce the representative point that is associated with each form, which plays the role of the unique root in $\mathbb{H}$. Reduced forms are defined similarly. Definitions and propositions in this section references [6]. Throughout these next two sections, we let $f \in \mathbb{R}[x_1, x_2]$ be a form of order k (a binary form of order k). Denote f's roots by $(\mu_i, \nu_i) \in \mathbb{P}_1(\mathbb{C})$. We can always rewrite f in the following way:

$$f = A\prod_{i=1}^{k}(\nu_i x_1 - \mu_i x_2)$$

where $A \in \mathbb{C}^*$.
Below is a way to define a real quadratic form based on f:

Definition 3.7 ($\varphi(f, \vec{t})$): Let $\vec{t} \in (\mathbb{R}^*)^k$ be a vector representing some weights, then

$$\varphi(f, \vec{t}) = \sum_{i=1}^{k} t_i^2(\nu_i x_1 - \mu_i x_2)(\overline{\nu_i} x_1 - \overline{\mu_i} x_2)$$

Note that $\varphi(f, \vec{t})$ is a quadratic form. Furthermore, by expanding each term we verify that its coefficients are real. It is also a positive definitive binary quadratic form, as $(x_1, x_2) \neq (0, 0)$ always give a positive value of $\varphi(f, \vec{t})$.

Definition 3.8 ($\Phi(f, \vec{t})$): Fix a representative of roots $(\mu_i, \nu_i)_{1 \leq i \leq k}$ of f in $\mathbb{P}(\mathbb{C})$, denonte

$$\Phi(f, \vec{t}) = \frac{|A|^2(-\operatorname{disc}(f)/4)^{k/2}}{(\prod_{i=1}^{k} t_i)^2}$$

Definition 3.9 (Hermite Covariant): Let $f \in \mathbb{R}[x_1, x_2]$ be a form of degree k as usual and let $z \in \mathbb{C}$ be arbitrary. Define the Hermite covariant as follows:

$$\Theta(f, z) = \min \Phi(f, \vec{t}), \text{ over all } \vec{t} \in (\mathbb{R}^*)^k \text{ such that } \varphi(f, \vec{t})(z, 1) = 0$$

If such $\Phi(f, \vec{t})(z, 1) = 0$ for all $\vec{t}$ then we set $\Theta(f, z)$ to ∞.

There are two points to notice. First, using min here makes the Hermite covariant not dependent on the representatives of the roots (μ_i, ν_i). Second, since the form is real and positive definitive, replacing z with $\overline{z}$ still yields 0. So we can assume $z \in \mathbb{H}$.

Definition 3.10 (Hermite Determinant): For a form $f \in \mathbb{R}[x_1, x_2]$, we define its Hermite determinant to be

$$\Theta(f) = \min_{z \in \mathbb{H}} \Theta(f, z)$$

Definition 3.11 (Signature of a form): The signature of a form f is (r, s) where r is the number of real roots and s the number of pairs of complex roots.

Definition 3.12 (Representative point): For a form $f \in \mathbb{R}[x_1, x_2]$, its representative point is any point $z \in \mathbb{H}$ such that $\Theta(f) = \Theta(f, z)$. In other words, it is the point in $\mathbb{H}$ that achieves the minimum value of $\Theta(f, z)$.

Note that the representative points are usually unique for a given form f. Referencing Proposition 4.2.2. of Edwards's work in [6], if f's signature is (r, s) , it has distinct roots, and either $k > 2$ or $s > 0$, then its representative point is unique.

Definition 3.13 (Reduced): A form $f \in \mathbb{R}[x_1, x_2]$ is Hermite reduced (or simply, reduced) if it has a representative point in the fundamental domain for $SL_2(\mathbb{Z})$, as follows:

$$\mathcal{D} = \{z = x + iy \in \mathbb{C}, x, y \in \mathbb{R} \mid |z| \geq 1, -1/2 \leq x \leq 1/2\}$$

Note that the definition of the fundamental domain differs slightly from Subsection 3.1.

Remark 3.14: In [10], Stoll and Cremona provided a method to find a reduced form that is $\mathrm{SL}_2(\mathbb{Z})$ equivalent to a given binary form of degree $n \geq 3$.

3.3 Useful results from Hermite reduction theory

Upon defining the terms for Hermite Reduction theory in the previous subsection 3.2, we will also introduce some results in Hermite reduction theory. We divide the results into two parts.
The first part consists of several results that will help us determine $\Theta(f)$ for a given form f.
Assuming that we have the value of $\Theta(f)$, the second part will help us to determine the bounds on the

coefficients of the reduced forms, using $\Theta(f)$. At the end of this section, we will comment on why these results are important.
The following theorems, definitions, and propositions are from [6].

3.3.1 Results that help compute $\Theta(f)$

Theorem 3.15 (Covariance): $f \in \mathbb{R}[x_1, x_2]$ a binary form of order k. Let $g \in \mathrm{GL}_2(\mathbb{R})$, then

$$\Theta(f \circ g, z) = |\det(f)|^k \Theta(f, gz)$$

If we set z to be the point where $f \circ g$ obtains its minimum, we obtain another identity that looks like how covariance is defined in Section 2.

$$\Theta(f \circ g) = |\det(g)|^k \Theta(f)$$

Proof : See Theorem 4.2.1 in [6]. □

The identity above becomes helpful when we work with the parametrisations of Klein Forms for different coefficients in the general Fermat equation.

Proposition 3.16: Let $f = A\prod_{i=1}^{k}(\nu_i x_1 - \mu_i x_2)$ be a real form of order k, where $k \geq 3$. Suppose that f's roots are distinct. If $z = x + iy \in \mathbb{H}$ is f's representative point, then $\Theta(f)$ is given by

$$\Theta(f) = \left(\frac{k}{2y}\right)^k |A|^2 \prod_{j=1}^{k}(|\nu_j x - \mu_j|^2 + |\nu_j y|)^2$$

Proof : See Proposition 4.2.3 in [6]. □

The proposition above allows us to write $\Theta(f)$ using its representative point. This will help us to compute $\Theta(f)$ directly.

Proposition 3.17: Let f be a real 4-form whose roots are all finite. Then the following table lists the weights at which f attains its Hermite determinant. There are three possibilities for f's signatures.

signature	roots	weights
$(4,0)$	$\alpha_1, \alpha_2, \alpha_3, \alpha_4$	$t(\alpha_i) = 1/(f'(\alpha_i, 1))$ where f' is the derivative with respect to the first variable.
$(2,1)$	$\alpha_1, \alpha_2, \beta, \overline{\beta}$	$t(\alpha_i)^2 = \lvert\beta - \overline{\beta}\rvert\lvert\alpha_{3-i} - \beta\rvert^2$ $t(\beta)^2 = t(\overline{\beta})^2 = \lvert\alpha_1 - \alpha_2\rvert\lvert\alpha_1 - \beta\rvert\lvert\alpha_2 - \beta\rvert$
$(0,2)$	$\beta_1, \overline{\beta_1}, \beta_2, \overline{\beta_2}$	$t(\beta_i)^2 = t(\overline{\beta_i})^2 = \lvert\beta_i - \overline{\beta}_i\rvert$

Table 3: The weights that yield the Hermite Determinant for 4 forms

Proof : See Proposition 4.2.4 of [6]. □

Lemma 3.18:
Let f be a real form of order $k \geq 3$ and let its signature be (r, s). If f has distinct roots and $f(x_2, -x_1) = \pm f(x_1, x_2)$, then the representative point of f is i.
Additionally, we factor $f = f_1 f_2$, where f_1 only has real roots and f_2 only has complex roots, then:

- If $r > 2$, then i is also the representative point of f_1
- If $s > 0$ then i is also the representative point of f_2. If f_2 has exactly one pair of complex roots, then i is the unique root of f_2 in $\mathbb{H}$.

Proof : Consider the map $z \mapsto -1/z$. Note that if $z = (x_2 : -x_1)$ is a root of f, then so is $-1/z = (x_1 : x_2)$. Therefore, this map permutes the roots of f_1 and f_2. Additionally, this map maps $\mathbb{H}$ to itself, with fixed points $\pm i$. By Theorem 4.2.2. of [6], the representative point of f is unique in $\mathbb{H}$. By Theorem 3.15, the Hermite determinant does not change under action by $\mathrm{SL}_2(\mathbb{Z})$. Hence the unique representative point must be i. Therefore, i is the representative point. □

3.3.2 Results that compute bounds

Theorem 3.19: Write $f \in \mathbb{R}[x_1, x_2]$ as below

$$f = \sum_{i=1}^{k} \binom{k}{i} a_i x_1^{k-i} x_2^i$$

If f is Hermite reduced,, then

$$|a_i a_j| \leq \left(\frac{4}{3k^2}\right)^{\frac{k}{2}} \Theta(f), \text{ for all } i, j \text{ such that } i + j \leq k$$

Proof : See Theorem 4.2.5 in [6]. □

The theorem above is extremely useful to us. Having the Hermite determinant, we can use this bound to generate the list of all possible reduced forms by exhaustively checking all possibilities of the coefficients, which is finite. This will produce a list of reduced forms. Yet we know that all binary forms are $\mathrm{SL}_2(\mathbb{C})$ equivalent to some reduced forms, and a later result will show that all $\mathrm{SL}_2(\mathbb{C})$-equivalent forms yield the same parametrisation, this helps us to determine a complete list of parametrisations.

4 A Computational Approach for the Diophantine Equation

After building some knowledge about invariant theory and the geometry of numbers, we have the preliminary knowledge to understand Edwards's approach in [6] to generate all parameterised solutions to the equations of the form $(X^2 + Y^3 + dZ^r = 0)$ as in (4).
This section and the next section are tightly connected, in the sense that this section introduces three key

theoretical ingredients that serves as the core of Edwards's Algorithm, which is what next section focuses on.
Before diving deeper into the technical details, it is important to get a sense of why these three ingredients are important. Below is a quick preview on how each ingredient is used in Edwards's algorithm.

- **Subsection 4.1, Klein forms and their connection to parametrisations**: Building on Section 3, this section shows us which binary forms can be used to construct parameterised solutions. Furthermore, this section provides us an explicit way to generate solutions to (4) given any Klein form (to be defined later, though we already saw examples in Subsection 2). In the algorithm, this helps to produce concrete numeric solutions in order to check whether the generated parameterisations yield coprime solutions.
- **Subsection 4.2, Applying Hermite Reduction Theory to Klein forms**: Building on Hermite Reduction Theory in Section 3, this section helps us in two ways. First, it gives us a bound for the coefficients for the Klein forms, so that we can enumerate all possibilities of Klein forms explicitly in the algorithm. Second, it shows us how to compute representative points of forms, so that we can throw away all the forms that are not reduced, which are precisely those whose representative points are not in the fundamental region.
- **Subsection 4.3, Lifting solutions to Klein forms**: This section helps us to determine the coefficients of a Klein form given a single solution $(X, Y, Z) \in \mathbb{Z}^3$. In the algorithm, we will first generate all possible solutions of the form $(X, Y, Z) \in \mathbb{Z}^3$ (within a justified bound). Then for each solution (X, Y, Z), we will use this theory to 'lift' it to a Klein form f such that (X, Y, Z) is one of the integer specialisations of f.

Now, we should shift our attention back to this section and focus on understanding the technical details behind it.

4.1 Klein forms and their connection to parametrisations

Continuing on Section 2, in this subsection, we will first introduce Klein form and a series of ideas that links Klein forms and the parametrised solutions of $X^2 + Y^3 = dZ^r$ together.
Following Edwards's work in [6], we first define a set of constants for each of the three platonic solids:

Definition 4.1 (Constants $\tilde{f}_r, k, N, \beta_r$):

Platonic Solid	r	$\tilde{f}_r$	$k = \deg(f)$	N	β_r	$\deg(H(f))$	$\deg(t(f))$
Tetrahedron	3	$\tilde{f}_3 = x_2^4 - 2\sqrt{3}x_1^2x_2^2 - x_1^4$	4	12	$3\sqrt{3}$	4	6
Octahedron	4	$\tilde{f}_4 = x_1x_2(x_1^4 - x_2^4)$	6	24	432	8	12
		$\tilde{f}_4{}^* = x_1x_2(x_1^4 + x_2^4)$					
Icosahedron	5	$\tilde{f}_5 = x_1x_2(x_1^{10} + 11x_1^5x_2^5 - x_2^{10})$	12	60	1728	20	30

Table 4: Constants for each Platonic Solid

The following is the meaning of each constant in the above table:

- r denotes the exponent of Z in the equation $X^2 + Y^3 = dZ^r$, where the parametrisation of this equation corresponds to the platonic solid as demonstrated in Klein's Relation (Table 2).
- $\tilde{f}_r$ is the equation whose roots correpsonds to the vertices of the platonic solid. Note that $\tilde{f}_4{}^*$ is another equation whose roots correspond to the vertices of the Octahedron.

- k denotes the number of vertices in the platonic solid, and it also corresponds to the degree of the binary form $\tilde{f}_r$.
- N denotes the order of the group of rotational symmetries of the platonic solid.
- β_r denotes the reciprocals of the constants in front of f in Table 2.
- $t(f), H(f)$ denotes the Hessian and the functional determinant of f, respectively, as defined in Definition 2.9. Their degrees are $2k-4, 3k-6$, respectively.

Now we can rewrite Klein's relation with our new notation, as in [6]:

$$\left(\frac{1}{2}t(\tilde{f}_r)\right)^2 + H(\tilde{f}_r)^3 + \frac{1}{\beta_r}\tilde{f}_r{}^r = 0 \tag{5}$$

and for $\tilde{f}_4{}^*$, we get

$$\left(\frac{1}{2}t(\tilde{f}_4^*)\right)^2 + H(\tilde{f}_4^*)^3 - \frac{1}{\beta_4}\tilde{f}_4^{*4} = 0 \tag{6}$$

Although (5) looks very similar to (4), we are only limited to produce parameterised solutions to the equation whose coefficient in front of Z^r is $\frac{1}{\beta_r}$. This gives parameterised solutions to only one value of d. To address this issue, we introduce a series of definitions. This allows d, the coefficient in front of $\tilde{f}_r$, to be arbitrary.

Definition 4.2 ($\mathscr{C}(r), \mathscr{C}(r,d)$):
Let $r \in \{3,4,5\}$ and $d \in \mathbb{C}^*$, we define

$$\mathscr{C}(r) = \{\tilde{f}_r \circ g \mid g \in \mathrm{GL}_2(\mathbb{C})\}$$

$$\mathscr{C}(r,d) = \left\{ f \in \mathscr{C}(r) \middle| \left(\frac{1}{2}t(f)\right)^2 + H(f)^3 + df^r = 0 \right\}$$

Note that both $\mathscr{C}(r)$ and $\mathscr{C}(r,d)$ consists of binary forms. In particular, each binary forms in $\mathscr{C}(r,d)$ presents a parametrised solution to $X^2 + Y^3 = dZ^r$.

This notation allows us to work with different values of d for (4), not just one d.

Definition 4.3 (Parametrisation): Given $f \in \mathscr{C}(r,d)$, a binary form, we write

$$\Phi(f) = \left(\frac{1}{2}t(f), H(f), f\right)$$

which is a parametrised solution to $X^2 + Y^3 = dZ^r$.

Next, we will present two results that allows us to write $\mathscr{C}(r,d)$ in terms of $\tilde{f}_r$, referencing [6].

Lemma 4.4:
First, $\tilde{f}_r \in \mathscr{C}(r, \beta_r^{-1})$.
Second, if $f \in \mathscr{C}(r,d)$, i.e. $\left(\frac{1}{2}t(f)\right)^2 + H(f)^3 + df^r = 0$, then

1. If $g \in \mathrm{GL}_2(\mathbb{C})$, then $f \circ g \in \mathscr{C}(r, \det(g)^6 d)$.
2. If $\lambda \in \mathbb{C}^*$, then $\lambda f \in \mathscr{C}(r, \lambda^{6-r} d)$.

Proof : $\tilde{f}_r \in \mathscr{C}(r, \beta_r^{-1})$ follows from the definition of $\mathscr{C}(r, \beta_r^{-1})$ and (5).
To show 1, since $H(f), t(f), f^r$ are covariants of weights $2, 3, 0$ respectively, so

$$g \circ t(f) = \det(g)^3 t(g \circ f)$$

$$g \circ H(f) = \det(g)^2 H(g \circ f)$$

$$g \circ (f^r) = g \circ f^r$$

From $\left(\frac{1}{2}t(f)\right)^2 + H(f)^3 + df^r = 0$ and from squaring the first equation, cubing the second, and multiplying the third by $\det(g)^6$, we obtain that $f \circ g \in \mathscr{C}(r, \det(g)^6 d)$.
To show 2, note that $H(f)$ and $t(f)$ can both be seen as homogeneous polynomials in the coefficients $a_0, \ldots a_k$. $H(f)$ is degree 2 in the a_is and $t(f)$ is degree 3 in the a_is, hence $H(f)^3$ and $t(f)^2$ are both of degree 6 in the $a_i s$. Hence, multiplying $\left(\frac{1}{2}t(f)\right)^2 + H(f)^3 + df^r = 0$ by λ^6 shows that the new constant in front of f^r is $d \cdot \lambda^{6-r}$. □

Proposition 4.5: We can write $\mathscr{C}(r, d)$ and $\mathscr{C}(r)$ as follows

$$\mathscr{C}(r, d) = \left\{ f = \tilde{f}_r \circ g \mid \det(g)^6 = \beta_r d \right\}$$

$$\mathscr{C}(r) = \bigcup_{d \in \mathbb{C}^*} \mathscr{C}(r, d)$$

Proof : To show the first identity, note that $\tilde{f}_r \in \mathscr{C}(r, \beta_r^{-1})$, so $\tilde{f}_r \circ g \in \mathscr{C}(r, \det(g)^6 \beta_r^{-1})$. So it follows.
The second identity follows from the first and the original definition of $\mathscr{C}(r)$. □

Definition 4.6 (Klein forms): We define Klein forms as the union $\mathscr{C}(3) \cup \mathscr{C}(4) \cup \mathscr{C}(5)$.

Note that Klein forms are our main objects of interest. In the rest of this essay, our main goals are to determine which binary form is a Klein form and to generate Klein forms.
We are now ready to study a subset of all Klein forms, called the integral Klein forms, as well as their properties.

Definition 4.7 ($\mathfrak{A}_r$): For each $r \in \{3, 4, 5\}$, we consider forms of degree $k \in \{4, 6, 12\}$ respectively.
For f, written as $f = \sum_{i=0}^{k} \binom{k}{i} a_i x_1^{k-i} x_2^i$, we define

$$\mathfrak{A}_3(f) = \{a_0, \ldots, a_4\}$$

$$\mathfrak{A}_4(f) = \{a_0, \ldots, a_6\}$$

$$\mathfrak{A}_5(f) = \{a_0, \ldots, a_5, 7a_6, a_7, \ldots a_{12}\}$$

Note that essentially, $\mathfrak{A}_r$ of the form f is just to extract the coefficient for each summand and dividing it by the weight of the binomial coefficient.
One exception is that the seventh coefficient for $\mathfrak{A}_5(f)$ is multiplied by 7. Note that this implies that it is possible for a_6 to be of the form $\frac{z}{7}$, where $z \in \mathbb{Z}$. However, the form itself still has $\mathbb{Z}$ coefficient in front of $x_1^6 x_2^6$, since $\binom{12}{6}$ has a factor of 7.

Definition 4.8 (R-Integral Klein Forms): Let $R \subseteq \mathbb{C}$ be a ring, then the subset of R-integral forms is defined as

$$\mathscr{C}(r,d)(R) = \{f \in \mathscr{C}(r,d) \mid \mathfrak{A}_r(f) \subseteq R\}$$

Note that the ring R is usually $\mathbb{Z}$ or $\mathbb{R}$.

We now present two results about R-integral Klein forms.

Proposition 4.9 (The set of R-integral Klein Forms are closed under $\mathrm{GL}_2(\mathbb{Z})$): Fix (r,k) to be one of $(3,4),(4,6)$ or $(5,12)$.
Then $\mathscr{C}(r,d)(\mathbb{Z})$ is closed under the action of $\mathrm{GL}_2(\mathbb{Z})$.

Proof : Note that $\mathrm{GL}_2(\mathbb{Z})$ is generated by S, T, U, where

$$S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, U = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Therefore, it suffices to check $\mathscr{C}(r,d)(\mathbb{Z})$ is closed under the action by S, T, U.
Write $f = \sum_{i=0}^{k} \binom{k}{i} a_i x_1^{k-i} x_2^i$. Since the action by S is just to replace a_i with a_{k-i} and multiply each odd indexed coefficient by -1, the set is closed under action by S. Similarly, the action by U is to simply multiply each odd-indexed coefficient by -1. Therefore this set is also closed under the action by U.
We now show the set is closed under the action by T. Action by T is the same as evaluating $f(x_1+x_2, x_2)$. We have

$$f(x_1 + x_2, x_2) = \sum_{i=0}^{k} \binom{k}{i} a_i' x_1^{k-i} x_2^i$$

where

$$a_t' = a_t + \sum_{i=0}^{t-1} \binom{t}{i} a_i, t \in \{0, \ldots, k\} \tag{7}$$

For $r = 3, 4$ this shows that a_t' is still in $\mathbb{Z}$ for each t.
For $r = 5$, for $0 \leq t \leq 6$, each of the summands above is an integer, hence each a_t' is still an integer. For $7 \leq t \leq 12$, 7 divides $\binom{t}{6}$, this guarantees that $\binom{t}{6} a_6$ is an integer, so is $\sum_{i=0}^{t-1} \binom{t}{i} a_i$. Therefore, this set is also integrally closed under action by T. □

Remark 4.10: Remark that formula (7) will be helpful to us in the next section. We will encounter the case where it is difficult to compute the representative point of f directly, yet we can easily compute the representative point of $T \circ f$.

The proposition below will tell us about the properties of $\mathbb{Z}$ integral forms when taking covariants.

Proposition 4.11: Fix (r, k) to be $(3, 4), (4, 6)$ or $(5, 12)$. Let $C = \sum_{i=0}^{n} C_i x_1^{n-i} x_2^i$ be a covariant. Suppose that

- $C_0 \in \mathbb{Z}[a_0, \ldots, a_k]$, and
- If $r = 5$, then the covariant has weight ≤ 5
- $\mathfrak{A}_r(f) \subseteq \mathbb{Z}$

Then, $C(f) \in \mathbb{Z}[x_1, x_2]$. That is $C(f)$ is a binary form whose coefficients are integral.

Proof : See Proposition 2.4.2 in [6]. □

The above proposition tells us that if $f \in \mathscr{C}(r, d)$ is $\mathbb{Z}-$integral, then $C(f)$ has coefficients in $\mathbb{Z}$.
Next, we present a theorem that characterizes all Klein forms.

Definition 4.12 (The 4th and 6th covariants): Let f be a degree k form.
We define Ω as follows:

$$\Omega = \left(\frac{\delta^2}{\delta x \delta y'} - \frac{\delta^2}{\delta y \delta x'} \right)$$

where x, y, x', y' are variables.
We define the $4th$ and $6th$ covariants as follows:

$$\tau_4(f) = \frac{1}{2} \left(\frac{(k-4)!}{k!} \right)^2 \Omega^4 f(x, y) f(x', y') \mid_{x, x' = x_1, y, y' = x_2}$$

$$\tau_6(f) = \frac{1}{2} \left(\frac{(k-6)!}{k!} \right)^2 \Omega^6 f(x, y) f(x', y') \mid_{x, x' = x_1, y, y' = x_2}$$

Essentially, the 4th and the 6th covariants are obtained by computing the derivatives Ω for 4, 6 times respectively, multiplying a constant, and then substitute x, x' by x_1 and y, y' by x_2 and finally obtaining a binary form in x_1, x_2. Note that $\tau_4(f)$ is a $2k - 8$ form and $\tau_6(f)$ is a $2k - 12$ form, having weights $4, 6$, respectively.

Definition 4.13 (The catelecticant invariant j): Let f be a 4−form and we define $j(f)$ to be

$$j(f) = \begin{vmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \\ a_2 & a_3 & a_4 \end{vmatrix}$$

Note that it has weight 6.

Theorem 4.14 (Classification of Klein forms): Let $\mathbb{C}[x_1, x_2]_k$ denote k-degree binary forms whose coefficients are in $\mathbb{C}$.
Fix $d \in \mathbb{C}^*$ as the coefficient of Z^r, then

$$\mathscr{C}(3, d) = \{f \in \mathbb{C}[x_1, x_2]_4, \mid \tau_4(f) = 0, j(f) = 4d\}$$

$$\mathscr{C}(4, d) = \{f \in \mathbb{C}[x_1, x_2]_6, \mid \tau_4(f) = 0, \tau_6(f) = 72d\}$$

$$\mathscr{C}(5, d) = \{f \in \mathbb{C}[x_1, x_2]_{12}, \mid \tau_4(f) = 0, \tau_6(f) = \frac{360}{7} d \cdot f\}$$

Proof : See Theorem 2.5.1, Lemma 2.5.2, and Theorem 2.5.3 on [6]. □

This theorem presents the necessary and sufficient conditions for a form to be in $\mathscr{C}(r, d)$. In fact, for a form f, to test whether $f \in \mathscr{C}(r, d)$, we just need to compute two of $\tau_4(f), \tau_6(f)$ or $j(f)$. On the other hand, as we will see in Subsection 4.3, we can compute $\tau_4(f)$ and $\tau_6(f)$ in terms of the coefficients of f explicitly. This allows us to generate equations in the a_is that are equivalent to $\tau_4(f) = 0$, $\tau_6(f) = 72d$, or $\tau_6(f) = \frac{360}{7} d \cdot f$. Thus, to check whether $f \in \mathscr{C}(r, d)$, it suffices to check whether its coefficients satisfies the equations in terms of the a_is. These equations in a_is are called the defining equations, as we will see later.
This concludes the Subsection 4.1. In this subsection, we introduced Klein forms and showed how to use them to generate parametrised solutions.

4.2 Applying Hermite Reduction Theory to Klein forms

In the previous subsection, we studied Klein forms, how they produce parameterised solutions, and the equivalent conditions for a form to be a Klein form in Theorem 4.14. This is the second key ingredient in Edwards' algorithm.
In this subsection, we will show how to use Hermite Reduction Theory on Klein forms in order to obtain bounds on the coefficients of reduced Klein forms. Recall that in Section 3, we have grouped the results of Hermite Reduction Theory into two parts: how to compute $\Theta(f)$ and and how to compute the coefficients given $\Theta(f)$. In this subsection, we will combine them together to obtain explicit bounds for the coefficients and also provide a method to compute representative points of Klein forms.
Again, the results presented in this section references [6].
We first present a result about the roots of Klein forms.

Proposition 4.15: If $f \in \mathscr{C}(r, d)(\mathbb{R})$, then it has at least one real root.

Proof : See Proposition 4.3.1. of [6]. □

Theorem 4.16: Let $f \in \mathscr{C}(r,d)(\mathbb{R})$ and let $f' \in \mathscr{C}(r,d')(\mathbb{R})$ be Klein forms whose coefficients are all real, and $d, d' \neq 0$.
Denote $\mathrm{GL}_2(\mathbb{R})^+$ be the set of matrices in $\mathrm{GL}_2(\mathbb{R})$ with positive determinants. Then

- If d, d' have the same sign, then f, f' are $\mathrm{GL}_2(\mathbb{R})^+$ equivalent.
- If d, d' have different signs and r is 3 or 5, then f and $-f'$ are $\mathrm{GL}_2(\mathbb{R})^+$ equivalent.

Proof :
Consider Lemma 4.4. When d, d' have different signs, and r is 3 or 5, we can change f by applying $-I$, so that $-f \in \mathscr{C}(r,-d)$. Now it remains to prove the first claim.
We use Lemma 4.4 again, we can apply $g, g' \in \mathrm{GL}_2(\mathbb{R})$ with determinants $d^{-1/6}$, $d'^{-1/6}$ respectively, to f, f'. This ensures that $f \circ g, f' \circ g' \in \mathscr{C}(r,1) \cup \mathscr{C}(r,-1)$. Additionally, they either both belong to $\mathscr{C}(r,1)$ or $\mathscr{C}(r,-1)$.
So we can now assume $d = d' = \pm 1$. By Proposition 4.15, f, f' has at least one real root. By techniques analogous to the proof of Theorem 4.23, we may apply a $\mathrm{GL}_2(\mathbb{R})^+$ transformation to f such that its real root maps to ∞, and the first three coefficients of f are $(0,1,0)$. Similarly, we can change f' such that its real root also maps to ∞ with its first three coefficients $(0,1,0)$.
However, by the defining equations for $r = 3,4,5$, which is a relation about the coefficients of the Klein forms (See Remark 4.22), the rest of the coefficients are determined uniquely by the first three coefficients. We will see similar proof techniques in Remark 4.22 and Theorem 4.23 later in this section. Therefore, there exists a matrix in $\mathrm{GL}_2(\mathbb{R})^+$ that serves as a proof that f to f' are $\mathrm{GL}_2(\mathbb{R})^+$ equivalent. □

Now, we will show two theorems that are directly useful for computation: the first one will be helpful for computing Representative points of a Klein form, and the second one will be helpful to compute the bounds of the coefficients of a Klein form that is also Hermite Reduced.

Theorem 4.17: Suppose that $f \in \mathscr{C}(r,d)(R)$, then, the following is its Hermite determinant.

Class	A representative of this class	Signature	$\Theta(f)$
$\mathscr{C}(3,d)$	$\tilde{f}_3$	$(2,1)^6$	$2^6 3^3 \lvert d\rvert^{2/3}$
$\mathscr{C}(4,d), d>0$	$\tilde{f}_4$	$(4,1)$	$2^8 3^9 \lvert d\rvert$
$\mathscr{C}(4,d), d<0$	$\tilde{f}_4^*$	$(2,2)$	$2^8 3^9 \lvert d\rvert$
$\mathscr{C}(5,d)$	$\tilde{f}_5$	$(4,4)$	$2^{24} 3^{18} 5^5 \lvert d\rvert^2$

Table 5: Hermite Determinant for each class of Klein Forms

Furthermore, if we factor $f = f_1 f_2$, where all of f_1's roots are real and all of f_2's roots are complex, here is a fast way to find a representative point for each class of Klein forms:

Class	Method
$\mathscr{C}(3,d)$	Use Proposition 3.17 directly on f, since f is a 4 form
$\mathscr{C}(4,d), d>0$	Return the unique root of f_2 in $\mathbb{H}$
$\mathscr{C}(4,d), d<0$	Return the representative point of f_2 using Proposition 3.17
$\mathscr{C}(5,d)$	Return the representative point of f_1 using Proposition 3.17

Table 6: Find Representative Points

Remark 4.18: [7]
Suppose a 4-form has finite roots. Then, to compute its representative points using its roots r_i and the weights t_i^2 obtained from Table 3.17, we only need to compute the root of the quadratic binary form

$$\phi(z,1) = \sum_{i=1}^{k} t_i^2 (z - r_i)(z - \overline{r_i})$$

in $\mathbb{H}$. This is by definitions of the Hermite covariant and the Hermite determinant.

We will use the above theorem in the algorithm. It provides us a method to compute representative points of a form given its coefficients.

Proof : [8] We first verify that the signature is correct for each of the $\tilde{f}_r$. Indeed,

- $\tilde{f}_3$ has roots $[\pm 0.517638, \pm 1.931852i]$
- $\tilde{f}_4$ has roots $[\infty, 0, \pm 1, \pm i]$.
- $\tilde{f}_4^*$ has roots $[\infty, 0, \pm\sqrt{i}, \pm\sqrt{-i}]$
- $\tilde{f}_5$ has roots $[\infty, 0, \pm 0.618034, -0.5 \pm 1.538842i, -0.5 \pm 0.363271i, 1.309017 \pm 0.951057i, 0.190983 \pm 0.587785i]$

Therefore, the signatures are correct for $\tilde{f}_r$.
Next, we show that i is a representative point for each of $\tilde{f}_r$. Indeed,

- For $\tilde{f}_4, \tilde{f}_4^*$ and $\tilde{f}_5$, we can use Lemma 3.18. By verifying that $f(x_2, -x_1) = \pm f(x_1, x_2)$ we conclude that their representative points are i.
- For $\tilde{f}_3$, note that $f(x_2, -x_1) \notin \pm f(x_1, x_2)$, so we cannot use Lemma 3.18. To find its representative point, we use Table 3 to find the roots and the weights:
  - Roots are $[0.517638, -0.517638, 1.931852i, -1.931852i]$
  - Respective weights are $[15.454813, 15.454813, 4.141105, 4.141105]$

  Using Remark 4.18, indeed i is its representative point.

We then verify that $\Theta(f)$ is correct for each of $\tilde{f}_r$. This is done by using Proposition 3.16 with d equal to the reciprocal of β_r (as shown in (5)). We can find the value of β_r in Table 4.

[6] In [6], this should be $(2,1)$ instead of $(2,2)$.
[7] I thought of on my own.

Upon verifying that the formula $\Theta(f)$ is correct when f is one of the $\tilde{f}_3, \tilde{f}_4, \tilde{f}_4^*$ or $\tilde{f}_5$, we now verify the formula for all $f \in \mathscr{C}(r,d)$.
Let $f \in \mathscr{C}(r,d)$ be arbitrary, then

- By Theorem 4.16, $\pm f = \tilde{f}_r \circ g$ for some $g \in \mathrm{GL}_2(\mathbb{R})^+$.
- By Theorem 3.15, $\Theta(f) = |\det(g)|^k \Theta(\tilde{f})$
- By Lemma 4.4, $\det(g)^6 = \beta_r d$

Combing the above, for $r = 3$, we have

$$\Theta(f) = |\det(g)|^k \Theta(\tilde{f}_3) = |\det(g)|^k 2^6 3^3 \left(\frac{1}{\beta_r}\right)^{2/3} = 2^6 3^3 |d|^{2/3}$$

The case for $r = 4, 5$ is verified similarly. This verifies the table for all $f \in \mathscr{C}(r,d)$.
The correctness of methods listed in Table 6 follows from the two points mentioned in Lemma 3.18. Therefore, finding the representative point of a form in $\mathscr{C}(4,d)$, where signature is $(4,1)$, is equivalent to finding its unique root in $\mathbb{H}$. Finding the representative point of a form in $\mathscr{C}(4,d)$ with signature $(2,2)$ is equivalent to finding the representative point of its complex factor. Similarly, finding the representative point of a form in $\mathscr{C}(5,d)$ is equivalent to finding its representative point of its real factor.

□

Theorem 4.19: Let $f \in \mathscr{C}(r,d)$ be a Hermite reduced Klein Form. Let k denote its degree. Then, we have the following bound on its coefficients:

$$\text{For all } i,j, \text{ such that } i+j \leq k, |a_i a_j| \leq B^2$$

where B is given by the following table:

Class	B
$\mathscr{C}(3,d)$	$2 \cdot 3^{1/2} \lvert d \rvert^{1/3}$
$\mathscr{C}(4,d)$	$16 \lvert d \rvert^{1/2}$
$\mathscr{C}(5,d)$	$1600 \cdot 5^{1/2} \lvert d \rvert$

Table 7: Bound on the Coefficients of Klein Forms

Proof : We use the bounds presented in Theorem 3.19. We simply plug in the value of $\Theta(f)$ specified in Table 5 of Theorem 4.17 to obtain B.

□

In Edwards' Algorithm, the above theorem is very important. It is useful for generating a list of candidates to be Hermite reduced Klein forms. This will provide us a with a list of forms as a starting point. This concludes Subsection 4.2. In this subsection, we presented methods to compute representative points

[8] I computed the values in this proof using the `numpy` library in Python.

and the bounds on Hermite reduced Klein forms.

4.3 Lifting solutions to Klein forms

In this subsection, we will present two important theorems by Edwards in [6] that is used to 'lift' solutions of the form $(X, Y, Z) \in \mathbb{Z}^3$ to a Klein form. In his work, Edwards describes this as the 'Arithmetic heart' of this paper. This is the third key ingredient in Edwards' algorithm.
This subsection consists of two important theorems by Edwards, the lifting theorem and the uniqueness theorem.
Essentially, if we are given a triple $(X, Y, Z) \in \mathbb{Z}^3$ such that $\gcd(X, Y, Z) = 1$, and it is a solution to

$$X^2 + Y^3 + dZ^r = 0$$

Then, the existence theorem would tell us that there exists a form $f \in \mathscr{C}(r, d)$ such that (X, Y, Z) is an integer specialisation of f. Furthermore, the uniqueness theorem will tell us that if there exists $f' \in \mathscr{C}(r, d)$ such that (X, Y, Z) is also an integer specialisation of f', then f' is $\mathrm{SL}_2(\mathbb{C})$ equivalent to f. That is, all of the Klein forms in $\mathbb{C}(r, d)$ that integer specializes to a particular (X, Y, Z) are all equivalent under the action of $\mathrm{SL}_2(\mathbb{C})$.
For a specific (X, Y, Z) that solves $X^2 + Y^3 + dZ^r = 0$, once you find one form that integer specializes to it, you can find all by $\mathrm{SL}_2(\mathbb{C})$ actions.
Before stating and proving the theorems, it is also worth pointing out that, as a result of the proof the lifting theorem, we also obtain a recipe for computing the form f that specializes to (X, Y, Z), given the values of (X, Y, Z).
Now we are ready to state the two theorems.

Theorem 4.20 (The lifting theorem): Fix $d \in \mathbb{Z}, d \neq 0$ and $r \in \{3, 4, 5\}$. Let $(X, Y, Z) \in \mathbb{Z}^3$ be a solution to the spherical diophantine equation

$$X^2 + Y^3 + dZ^r = 0$$

such that $\gcd(X, Y, Z) = 1$, then there exists a binary form $f \in \mathscr{C}(r, d)(\mathbb{Z})$ where (X, Y, Z) is f's integer specialisation. That is, there exists $(c_1, c_2) \in \mathbb{Z}^2$ such that $\Phi(f)(c_1, c_2) = (X, Y, Z)$.

Before we prove the theorem, we first introduce a lemma and state the defining equations for $\mathscr{C}(r, d)$.

Lemma 4.21:
Fix $f \in \mathscr{C}(r, d)$ a form. Let $(X, Y, Z) \in \mathbb{C}$ be such that $X^2 + Y^3 = dZ^r$. Then, f (written more explicitly, $f(x_1, x_2)$) specializes to (X, Y, Z) with complex arguments. That is, there exists $c_1, c_2 \in \mathbb{C}$ such that $\Phi(f)(c_1, c_2) = (X, Y, Z)$.

Proof (of Lemma 4.21): Consider the system of equations in the unknowns s_1, s_2:

$$\begin{cases} H(f)(s_1, s_2) = Y \\ f(s_1, s_2) = Z \end{cases}$$

They are binary forms of degree $2k - 4$ and k respectively. Using methods from elimination theory (this is a system of equations in two variables), we can obtain solutions $(c_1, c_2) \in \mathbb{C}^2$.
Now, since $\left(\frac{1}{2}t(f), H(f), f\right)$ satisfies $X^2 + Y^3 + dZ^r = 0$ and $H(f)(c_1, c_2) = Y$, $f(c_1, c_2) = Z$, we must have

$\frac{1}{2}t(f)(c_1, c_2) = \pm X$.
Here are two cases:

- If $\frac{1}{2}t(f)(c_1, c_2) = X$, then $\Phi(f)(c_1, c_2) = (X, Y, Z)$
- Otherwise, $\frac{1}{2}t(f)(c_1, c_2) = -X$. In this case, we change c_1, c_2 by multiplying it with a matrix so that the new vector yields a parametrisation for (X, Y, Z).

  Since $f = \tilde{f}_r \circ h$ for some $h \in \mathrm{GL}_2(\mathbb{C})$, (for $r = 4$, set $\tilde{f}_r$ to $\tilde{f}_r^4$.)

  For the cases $r = 3, 4, 5$, set $\tilde{m}$ to be $\begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}$, $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}$ respectively.

  Observe that in each case, $\tilde{m}$ is a matrix of determinant -1 that fixes $\tilde{f}$. Therefore, $h^{-1}\tilde{m}h := m$ is a matrix of determinant -1 that also fixes f. We replace $\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$ by $\begin{bmatrix} c_1' \\ c_2' \end{bmatrix} = m \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$.

  Now, since $f, H(f), t(f)$ are covariants of weighs $0, 2, 3$ respectively, we have that $f(c_1', c_2') = f(c_1, c_2) = Z$ and $H(f)(c_1', c_2') = f(c_1, c_2) = Y$. Yet $t(f)(c_1', c_2')$ has its sign flipped so $t(f)(c_1', c_2') = -(-X) = X$.

  Hence $\Phi(f)(c_1', c_2') = (X, Y, Z)$ as desired.

□

Remark 4.22:
From Theorem 4.14, we are able to obtain a set of defining equations for each $r = 3, 4, 5$. That is, a form $f = [a_0, \ldots, a_k]$ is in $\mathscr{C}(r, d)$ if and only if it satisfies the defining equations. We will use the defining equations in the proof of Lifiting theorem. This is also useful to us when determining $a_3, \ldots, a_k$ from a_0, a_1, a_2. Below are the defining equations obtained from Appendix A of [5]. [9]
The defining equations for $\mathscr{C}(3, d)$ are given by:

$$0 = a_0a_4 - 4a_1a_3 + 3a_2^2$$

$$-4d = a_0a_2a_4 + 2a_1a_2a_3 - a_2^3 - a_0a_3^2 - a_1^2a_4$$

The defining equations for $\mathscr{C}(4, d)$ are given by:

$$0 = a_0a_4 - 4a_1a_3 + 3a_2^2$$

$$0 = a_0a_5 - 3a_1a_4 + 2a_2a_3$$

$$0 = a_0a_6 - 9a_2a_4 + 8a_3^2$$

$$0 = a_1a_6 - 3a_2a_5 + 2a_3a_4$$

$$0 = a_2a_6 - 4a_3a_5 + 3a_4^2$$

$$-72d = a_0a_6 - 6a_1a_5 + 15a_2a_4 - 10a_3^2$$

To make this essay short, I omitted the defining equations for $\mathscr{C}(5, d)$. It is available in Appendix A of [5]. To obtain the defining equations for $\mathscr{C}(5, d)$, one would need to write $f \in \mathbb{C}[x_1, x_2]_{12} = \sum_{i=0}^{12} a_i x_1^{12-i} x_2^i$,

[9] When implementing Edwards's algorithm, I used the defining equations with the signs of d flipped. Otherwise the `sympy` library will not find a solution.

expand $\tau_4(f) = 0$ and $\tau_6(f) = \frac{360}{7} d \cdot f$. Then we end up with equations in the a_i and d, which are precisely the defining equations for $\mathscr{C}(5, d)$.
In my implementation of Edwards's algorithm, I only computed the tetrahedral and the octaheral case. Hence I only used the defining equations for $\mathscr{C}(3, d)$ and $\mathscr{C}(4, d)$.
We can finally prove the theorem.

Proof (of Theorem 4.20):
This proof consists of four steps

- Step 1. find some $f \in \mathscr{C}(r, d), c_1, c_2 \in \mathbb{C}$ where f is not necessarily in $\mathscr{C}(r, d)(\mathbb{Z})$.
- Step 2. apply transitions on a_0, a_1, a_2, a_3, such that X, Y, Z can be directly expressed using a_0, a_1, a_2, a_3.
- Step 3. apply another transition so that $a_0, a_1, a_2 \in \mathbb{Z}$
- Step 4. finally, show that $\mathfrak{A}_r(f) \subseteq \mathbb{Z}$. i.e. the rest of the coefficients $a_4, \ldots, a_k \in \mathbb{Z}$ (or $7a_6 \in \mathbb{Z}$ when $r = 5$.)

Step 1.
We start with any $f \in \mathscr{C}(r, d)$ and use Lemma 4.21 to find $c_1, c_2 \in \mathbb{C}$ such that $\Phi(f)(c_1, c_2) = (X, Y, Z)$.
Step 2.
Since $\mathscr{C}(r, d)$ is closed under transformations in $\mathrm{SL}_2(\mathbb{C})$, we can apply a transformation m to f such that $\Phi(f \circ m)(1, 0) = (X, Y, Z)$, where m is a matrix of determinant 1 with $m \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Now, for convenience, we denote f for the form after transformation.
Recall that we defined $a_0, \ldots, a_k$ be the coefficients of f after diving by a factor of binomial coefficients. Expanding $\Phi(f \circ m)(1, 0) = (X, Y, Z)$, we obtain

$$f(1, 0) = Z = a_0 \tag{8}$$
$$H(1, 0) = Y = a_0 a_2 - a_1^2 \tag{9}$$
$$t(1, 0) = 2X = a_0^2 a_3 - 3a_0 a_1 a_2 + 2a_1^3 \tag{10}$$

The right-hand equality is obtained by taking the coefficient of $x_1^{\deg(\text{covariant})} x_2^0$ for each covariant as a result of evaluating at $(x_1, x_2) = (1, 0)$. The exact expressions are obtained from expanding $H(f), t(f)$ given f's coefficients.
Step 3.
Now we make the second transformation. We will act on f by the matrix $\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$, or equivalently, replacing $f(x_1, x_2)$ by $f(x_1 + \lambda x_2, x_2)$ in order to ensure that $a_0, \ldots, a_3$ are integral. Note that the matrix maps $(1, 0)$ to $(1, 0)$ so that the above equations are still true. Furthermore, the matrix has determinant 1, so it still lies in $\mathscr{C}(r, d)$.
λ is defined as follows:

- Case when $Z = 0$.

  In this case, $a_0 = 0$, then a_1 will not be 0 as Klein forms cannot have multiple roots. We can pick λ such that $a_2 = 0$.

  Since $Z = 0$ and $\gcd(X, Y, Z) = 1$, we have $\gcd(X, Y) = 1$. Combined with the fact that $X^2 + Y^3 = dZ^r = 0$, we obtain that $X = \pm 1, Y = -1, Z = 0$. Plugging this back into the equation, we have $a_0 = 0, a_1 = \pm 1, a_2 = 0$.

- Case when $Z \neq 0$.

  If $Z \neq 0$, then, for any desired value of a_1, we can pick λ such that after transformation by $\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$, a_1 is equal to that desired value. Since $\gcd(Y, Z) = 1$, Y is a unit in $\mathbb{Z}/(Z^r\mathbb{Z})$, we can pick λ such that $a_1 \in \mathbb{Z}$ and that $a_1 = -X/Y \pmod{Z^r}$.

  From the equations in Step 2 ((8), (9), (10)), we can deduce that

$$\begin{aligned} a_0 a_2 = Y + a_1^2 &\equiv Y + \left(\frac{X}{Y}\right)^2 && \mod Z^r \\ &\equiv \frac{-dZ^r}{Y^2} && \mod Z^r \end{aligned}$$

  and that

$$\begin{aligned} a_0^2 a_3 &= 2X + 3a_0 a_1 a_2 - 2a_1^3 \\ &= 2X + 3(a_0 a_1 a_2 - a_1^3) + a_1^3 \\ &= 2X + 3a_1 Y + 3a_1^3 \\ &\equiv -X + \frac{X^3}{Y^3} && \mod Z^r \\ &\equiv \frac{-XdZ^r}{Y^3} && \mod Z^r \end{aligned}$$

  This implies that $a_0, a_1, a_2, a_3 \in \mathbb{Z}$.

  Now, let v_p be the $p-$adic valuation on $\mathbb{Q}$. For any p such that $p \mid Z$, we have $v_p(a_0) = v_p(Z)$, and $v_p(a_1) = 0$.

  From the equations above, we have $v_p(a_2) + a_p(a_0) \geq v_p(Z^r) = rv_p(Z)$, so that $\frac{v_p(a_2)}{v_p(Z)} \geq r - 1$.

  Similarly, $\frac{v_p(a_3)}{v_p(Z)} \geq r - 3$.

Step 4.

Now, our goal is to show $\mathfrak{A}_r(f) \subseteq \mathbb{Z}$. In the previous step, in the case $Z = 0$, we obtained $a_0, \ldots, a_2$ and in the case $Z \neq 0$, we obtained $a_0, \ldots, a_3$. We now show that the rest of the coefficients are also in $\mathbb{Z}$. (or $7a_6 \in \mathbb{Z}$ in the case $r = 5$.)

We will prove the claim for the octahedral case.[10] Note that the results for the tetrahedral case and the icosahedral case can be found in [6].

Recall the defining equations for $r = 4$ in Remark 4.22. Again we split into two cases depending on whether $Z = 0$.

- Case when $Z = 0$. From Step 3, we can assume $(a_0, a_1, a_2) = (0, \pm 1, 0)$. Using the defining equations, we can deduce that $a_3 = 0$, $a_4 = 0$ and $a_6 = 0$. Then, we can deduce that $a_5 = 12d/a_1$. Therefore, all the coefficients are integral.

- Case when $Z \neq 0$. From step 3, we obtained that for any prime $p \mid Z$,

$$\frac{v_p(a_0)}{v_p(Z)} = 1, \frac{v_p(a_1)}{v_p(Z)} = 0, \frac{v_p(a_2)}{v_p(Z)} \geq 3, \frac{v_p(a_3)}{v_p(Z)} \geq 2$$

  comibing this with the defining equations, we obtain that

$$\frac{v_p(a_4)}{v_p(Z)} \geq 1, \frac{v_p(a_5)}{v_p(Z)} \geq 0, \frac{v_p(a_6)}{v_p(Z)} \geq 3$$

Also, from the set of defining equations, we can deduce that if any of the a_is have a denominator that is not 1, then any primes dividing its denominator would be an integer dividing $a_0 = Z$. (i.e. from $0 = a_0a_4 - 4a_1a_3 + 3a_2^2$, since a_1, a_2, a_3 are integers, a_0a_4 must also be integers. Any primes dividing a_4's denominator must divide a_0, other cases are similar). However, the above showed that the valuation of the a_is as rationals is at least as big as the valuation of Z. Therefore, all of the a_is are integers.

The cases where $r = 3, 5$ are shown similarly. □

Theorem 4.23 (The Uniqueness Theorem): Let $f, f' \in \mathscr{C}(r, d)$. Suppose there exists $(c_1, c_2) \in \mathbb{Z}^2$ such that

$$\Phi(f)(c_1, c_2) = (X, Y, Z)$$

Then,

- If there exists $(c_1', c_2') \in \mathbb{Z}^2$ such that $\Phi(f')(c_1', c_2') = (X, Y, Z)$, then f, f' are $\mathrm{SL}_2(\mathbb{Z})$ equivalent.
- If there exists $(c_1', c_2') \in \mathbb{Z}^2$ such that $\Phi(f')(c_1', c_2') = (-X, Y, Z)$, then f, f' are $\mathrm{GL}_2(\mathbb{Z})$ equivalent.

Proof : See Theorem 3.2.1 of [6].

□

This concludes Subsection 4.3. In this subsection, we presented a method to lift integers solutions of the form (X, Y, Z) to a Klein form that specializes to it. We also showed uniqueness of the form up to $\mathrm{SL}_2(\mathbb{C})$ transformation.
Now, we are ready to put the three key sections to practice - to enumerate all Hermite reduced Klein forms up to $\mathrm{GL}_2(\mathbb{Z})$ equivalences in the next section.

5 Edwards's algorithm and implementation

In this section, we will first present Edwards's algorithm in Subsection 5.1, which explicitly lists all Hermite reduced Klein forms (for integral d values) up to $\mathrm{GL}_2(\mathbb{C})$ equivalences. Then, we will see the walk-through of the algorithm on an example of (r, d) in Subsection 5.2. Then, we will see some results for my implementation in Subsection 5.3 and how to use them in Subsection 5.4.
This section is important because it gives us a practical way to explicitly generate all the possible coprime integer solutions to

$$X^2 + Y^3 + dZ^r = 0$$

for $r \in \{3, 4, 5\}, d \in \mathbb{Z}, d \neq 0$. It is also significant as it puts the three core ingredients in Section 4 to practice.

5.1 Edwards' Algorithm

In this subsecion, we will present Edwards' algorithm to explicitly compute all the Hermite reduced Klein form in $\mathscr{C}(r, d)$ for $r \in \{3, 4, 5\}$ and $d \in \mathbb{Z}, d \neq 0$, up to $\mathrm{GL}_2(\mathbb{C})$ equivalences. The algorithm is divided into

[10] I proved a different case compared to [6]. However, I did verify my proof with [5].

four steps:

- Step 1: Produce all the Klein forms that satisfy the correct bound for reduced forms.
- Step 2: Given Klein forms in step 1, eliminate those whose representative points are not in the fundamental region. Note that step 1 and 2 together are retrieved from section 5.1 of [6].
- Step 3: Given the result of Step 2, keep only the $\mathrm{GL}_2(\mathbb{Z})$ reduced form. In addition, keep only one form in each $\mathrm{GL}_2(\mathbb{Z})$ orbit. This corresponds to section 5.2 of [6].
- Step 4: Given the result of step 3, only keep the ones that yield relatively prime specialisations. This corresponds to section 5.4 of [6].

For each step, I will present the pseudocode, discuss its proof of correctness, and some technical details.

5.1.1 Step 1. Produce Klein Forms

This algorithm first enumerates all the possibilities of a_0, a_1, a_2 exhaustively within the bounds. Then, for each case, it determines X, Y, Z and the rest of the coefficients. If possible, it lifts the X, Y, Z to a Klein form. It eventually returns the Klein forms with integral coefficients.
Pseudocode

```
Algorithm_1(r,d):
    B = compute_B(r,d) # Based on the table in previous section
    resulting_forms = []
    for a0, a1, a2 in [-B...B]x[-B...B]x[-B...B] :
        Z = a0
        Y = a0*a2-a1^2
        x_abs = sqrt(-Y^3-d*(Z^r))
        if x_abs is not integer:
            'continue with the next loop iteration'
        for X in [-x_abs,x_abs]:
            if a0 == 0 and a1 == 0: # a0, a1 cannot be simultaneously zero
                'continue with the next loop iteration'
            # computes a3 based on a0,a1,a2,X
            a3 = compute_a3(a0,a1,a2,X)
            #check if a3 is an integer
            if a3 is not integer:
                'continue with the next loop iteration'
            coefficients = compute_coefficients(a0,a1,a2,a3,d) # computes the rest of the
                coefficients based on the defining equations
            if not all of the coefficients are integer: #in the r=5 case, check 7*a6
                'continue with the next loop iteration'
            if not check_bounds_are_correct(coefficients):# Check that all coefficients
                matches the bound on the table from previous section
                'continue with the next loop iteration'
            resulting_forms = resulting_forms + coefficients
    return resulting_forms
```

Proof of correctness
We claim that this algorithm produces all the Klein forms in $\mathscr{C}(r,d)$ up to $\mathrm{GL}_2(\mathbb{Z})$ equivalence such that its coefficients satisfy the bounds of Table 7.

By Theorem 4.19, the coefficients satisfy the bounds. By Theorem 4.20, these coefficients do indeed represent a Klein form. Additionally, we have also searched exhaustively within the bounds, and each (X, Y, Z) lifts to exactly one Klein form up to $\mathrm{GL}_2(\mathbb{Z})$ equivalence by Theorem 4.20. Therefore, the claim is proven.
Technical details:

- Computing a_3:

  The method to compute a_3 differs when $a_0 = 0$ and when $a_0 \neq 0$.

  - When $a_0 \neq 0$, it is determined from $a_0^2 a_3 - 3a_0 a_1 a_2 + 3a_1^3 = 2X$.
  - When $a_0 = 0$, then determine it using $a_3 = 3a_2^2/(4a_1)$.[11]

- Computing the rest of the a_is:

  I used the `sympy` library to compute the rest of the a_is using the defining equations.

5.1.2 Step 2. Keep only Hermite Reduced Forms

Given all the forms produced in Step 1, it throws away all the forms that is not Hermite reduced.
Pseudocode

```
Algorithm2(r, Step_1_results):
    resulting_forms = []
    for each form in Step_1_results:
        case r is 3:
            if a0 is 0:
                rep_point = compute_representative_point(form)
                #this case is slightly different, see technical details
            else assume signature is (4,0) or (2,1): # since (0,3) cannot be Klein form
                rep_point = compute_representative_point(form)
        case r is 4:
            assume signature is (4,1) or (2,2): # since (0,3) cannot be Klein form
                rep_point = compute_representative_point(form)
        case r is 5:
            assume signature is (4,4):
                rep_point = compute_representative_point(form)
        if rep_point is not in fundamental_region:
            'continue with next iteration'
        resulting_forms = resulting_forms + form
    return resulting_forms
```

Proof of correctness
The correctness follows from Theorem 4.17 and Definition 3.2.
Technical Details [12]

- Computing the representative point: When $a_0 \neq 0$, we just follow Table 6 and Remark 4.18 to find the representative point.

  However, when $a_0 = 0$, computing the representative points is quite tricky. This happens when one of the roots is ∞. In the octahedral cases, the representation point only depends on the complex

[11] I used the defining equations in Remark 4.22 for $\mathscr{C}(3, d)$ and $\mathscr{C}(4, d)$ to derive this.

part of the form (that is, if $f = f_1 f_2$, as Table 6, then its representative point only depends on f_2). Therefore, as ∞ is not a complex root, so we can still proceed as usual.

When $a_0 = 0$ in the tetrahedral and icosahedral case, we apply the following two tricks to obtain the representative point:

1. When the last coefficient $a_k \neq 0$, we reverse the coefficients (or equivalently, transform by $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$), compute its representative point, then take the reciprocal. Since the reciprocal of a value in $\mathbb{H}$ is not in $\mathbb{H}$, we take the conjugate.

   The reason why this trick works is because x_1, x_2 are interchangeable in the definitions of Hermite Covariant and the representative points. The only exception is that in the definition of Hermite Covariant, $\phi(z, 1) = 0$ is not symmetric in x_1 and x_2. Therefore, we take the reciprocal, then take conjugate.

2. It is possible that $a_0 = a_k = 0$. Hence, the trick of reversing its coefficients do not work.

   In this case, we shall apply a transformation by the matrix $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ to the form.

   The formula 7 for performing a T transformation matrix gives us an explicit way to implement it. Once we perform the T transformation and compute the representative point $\tilde{z}$, it satisfies $\phi(\tilde{z}, 1) = 0$. To recover the original representative point, z, we have $\phi(\tilde{z}, 1) = \phi(z + 1, 1) = 0$. So we subtract 1 from $\tilde{z}$.

- Floating point precision: I used `np.roots` to compute the roots and to rounded them to the nearest 6th decimal places. This is important when deciding whether a point is in the fundamental region, the real part of a point might be 0.5, but it is computed as $0.5 - 10^{-16}$ which is less than 0.5. Another reason to round it is so that it looks cleaner in the outputs.

- Remark: When considering a form as a homogenized polynomial, make sure to multiply each term by its respective binomial coefficients.

5.1.3 Step 3. Keep only $\mathrm{GL}_2(\mathbb{Z})$ reduced forms and only one per orbit

In this step, we keep only the $\mathrm{GL}_2(\mathbb{Z})$ reduced forms. Furthermore, we only keep one representative per equivalence class of $\mathrm{GL}_2(\mathbb{Z})$ equivalent forms.
Pseudocode

```
Algorithm3(Step_2_results):
    GL2Z_reduced_forms = []
    for each form in Step_2_results:
        if representative_point(form) is not in D-:
        #D- denotes the fundamental domain for GL2(Z)
            'continue with next iteration'
        GL2Z_reduced_forms = GL2Z_reduced_forms + form
    Keep only one reduced form in each GL2Z orbits
    return one form in each group
```

In here, $\mathscr{D}^- = \{z = x + iy \mid |z| \geq 1, -1/2 \leq x \leq 0\}$ denotes the fundamental domain for $\mathrm{GL}_2(\mathbb{Z})$, and a form f is $\mathrm{GL}_2(\mathbb{Z})$ reduced if its representative point is in $\mathscr{D}^-$.

[12] I thought of the three following bullet points on my own.

Now, two reduced forms f_1, f_2 are $\mathrm{GL}_2(\mathbb{Z})$ equivalent if they have the same representative point z and one is the transformation of another under $\mathrm{Stab}(z, \mathrm{GL}_2(\mathbb{Z}))/\pm I$, where I is the identity matrix.
Recall that

$$S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, U = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

As shown in [6], below table enumerates the elements of $\mathrm{Stab}(z) = \mathrm{Stab}(z, \mathrm{GL}_2(\mathbb{Z}))/\pm I$ depending on the value of z:

Case	z	Stab(z)	|Stab(z)|
z in the interior of $\mathscr{D}^-$	Any	$\langle I \rangle$	1
z on the boundary of $\mathscr{D}^-$, $z = i$ or $z = \omega := -\frac{1}{2} + \frac{\sqrt{3}}{2}$	i ω	$\langle S, U \rangle$ $\langle ST, US \rangle$	4 6
z on the boundary of $\mathscr{D}^-$, $z \neq i$ and $z \neq \omega$	$x = 0$ $\|z\| = 1$ $x = -\frac{1}{2}$	$\langle U \rangle$ $\langle US \rangle$ $\langle U \rangle$	2 2 2

Table 8: Stab(z) for each z

Proof of correctness
See lemma 5.2.1. in [6].
Technical Details

- Transformations: The implementation of T-transformation is discussed in step 2 of the algorithm. S tranformation is achieved by reversing the indices and negating coefficients of odd indices. U transformation is achieved by negating coefficients of odd indices.
- Generating the orbits: To group the forms into their $\mathrm{GL}_2(\mathbb{Z})$ orbits with Stab(z), I generated the orbits of the form under Stab(z) and threw away all the other forms in that orbit. Note that since k is even, applying $-I$ to the form does not change the value at all. However, in Edwards' outputs, we identify forms $\pm f(x_1, x_2)$ together, and we also identify $f(x_1, \pm x_2)$ and $f(x_2, x_1)$ as the same. Therefore, there could be up to $8 \cdot 6$ equivalent forms in the case where the representative point is ω. We need to make sure we generate all the possibilities.

5.1.4 Step 4. Keep only relatively prime specialisations

The last step throws away forms that do not generate relatively prime solutions under any integer specialisations.
Pseudocode

```
Algorithm_4(Step_3_results,r,d):
    relatively_prime_forms = []
    for f in Step_3_results:
        compute H(f), t(f)
        N = N(r) # Recall N is the number of rotational symmetries
        N0 = 'product of all primes dividing N*d'
        for s1,s2 in [-N0... N0] x [-N0,... N0]:
            if gcd(Phi(f)(s1,s1)) == 1:
                relatively_prime_forms = relatively_prime_forms + form
                continue to the next form
    return relatively_prime_forms
```

This algorithm performs the following: For a form f, among all possible values of (s_1, s_2), $-N_0 \leq s_1, s_2 \leq N_0$, if pair of (s_1, s_2) yields a relatively prime solution $\Phi(f)(s_1, s_2)$, then it will be kept.
Proof of correctness
It suffices to show that for a fixed reduced Klein form f, if there exists $(s_1, s_2) \in \mathbb{Z}^2$ such that $\Phi(f)(s_1, s_2)$, then there exists (s_1', s_2'), $-N_0 \leq s_1, s_2 \leq N_0$ such that $\Phi(f)(s_1', s_2')$ are relatively prime. For the proof, see Edwards Section 5.4 in [6].
Technical Details

- To compute $H(f), t(f)$ and to compute $\Phi(f)(s_1, s_2)$ given (s_1, s_2), I used the `sympy` library.

Combining the four steps above, we obtained an algorithm that generates a list of $\mathrm{GL}_2(\mathbb{Z})$ equivalent Hermite reduced Klein forms that yields relatively prime solutions.

Remark 5.1: [13] One may point out, given the bounds, we are technically able to enumerate all possible integral forms, $a_0, \ldots, a_k$, check whether $\Phi(f)$ yields an answer, and then proceed with step 2, 3 and 4. That is, we can produce a list of forms without using the defining equations to lift the integer solutions. Then, why do we still bother to lift the solutions? The reason is to reduce runtime. If we try all the values of $a_3, \ldots a_k$ exhaustively, it would take longer.
Running the algorithm on $(r, d) = (4, 1)$ took 9.16779 seconds. If we wish to try all the possibilities of a_3, a_4, a_5, a_6, then runtime would be exponential in B, since each $|a_i a_j| \leq B^2$ for all $i + k \leq j$. Under some assumptions, the algorithm would take about 16^4 times longer to run, which is approximately 167 hours. Therefore, another importance of root lifting is to keep our program computationally feasible.

5.2 A walkthrough of the algorithm

In this subsection, we will present a walk-through of the algorithm with the parameter $(r, d) = (4, 2)$ to see how the algorithm produces $\mathrm{GL}_2(\mathbb{Z})$ reduced Klein forms. As a reminder, $[a_0, \ldots, a_6]$ represents the binary form $\sum_{i=0}^{6} \binom{6}{i} a_i x_1^{6-i} x_2^i$.
To obtain the following results, I traced through my own implementation of the algorithm with respect to the parameter $(4, 2)$.

- Step 1: In this step, the algorithm produces all the Klein forms that satisfy the correct bounds. We obtained 114 distinct forms of the form $[a_0, \ldots, a_6]$.
- Step 2: In this step, the algorithm removes all the Klein forms whose representative point is not in the fundamental region. We start with the 114 forms obtained from Step 1. In this step, 90 forms were removed and there were only 24 forms left.

  The following are some forms whose representative points are not in the fundamental region:

  - The form $[-13, 8, -5, 6, -9, 12, -9]$ with representative point $0.620 + 0.851i$ is eliminated. The representative point is not in the fundamental region as its real part is not in $[-1/2, 1/2]$.
  - The form $[1, 2, 5, 16, 53, 158, 337]$ with representative point $-2.125 + 1.961i$ is also eliminated. Its real part is also not in $[-1/2, 1/2]$
  - The form $[-18, -3, 0, 3, 2, 1, 4]$ with representative point $-0.195 + 0.762i$ is also eliminated. Although its real part is in $[-1/2, 1/2]$, its modulus is equal to 0.786, which is less than 1.

- Step 3: We start with the 24 forms obtained in Step 2.

  We first eliminate 8 forms that are not $\mathrm{GL}_2(\mathbb{Z})$ reduced. These are precisely the points whose real part is positive. Here are some examples:

[13] I thought of this remark on my own.

– The form $[-4, 1, -2, 3, 0, -3, 18]$ with representative point $0.316 + 1.231i$ is eliminated.

– The form $[-1, 0, -1, 2, 3, -4, 59]$ with representative point $0.125 + 1.961i$ is eliminated.

– The form $[8, -4, 4, -2, -2, 5, -13]$ with representative point $0.438 + 0.980i$ is eliminated.

We are left with 16 forms after removing the forms that are not $\mathrm{GL}_2(\mathbb{Z})$ reduced.

We next group the forms together if they are $\mathrm{GL}_2(\mathbb{Z})$ equivalent or identified by $\pm f(x_1, \pm x_2), \pm f(x_2, \pm x_1)$. For example $[-8, -4, -4, -2, 2, 5, 13]$ and $[8, 4, 4, 2, -2, -5, -13]$ are identified as $\pm f(x_1, x_2)$ are identified together. This gives us 8 forms.

(*Aside:* A not so simple example is in the case when $(r, d) = (4, -1)$, the two forms $(-8, -4, -4, -4, -2, 1, 7)$, $(-7, -6, -3, -2, -3, -6, -7)$ are identified up to $\mathrm{GL}_2(\mathbb{Z})$ equivalence and transformation by U and some sign changes).

- Finally, in step 4, we remove the forms that do not yield coprime parametrisations. Some examples of the forms that are eliminated include $[-4, -2, 0, -2, -4, -6, 8]$ and $[0, -2, 0, 0, 0, -12, 0]$. This leaves us with 5 forms in total.

5.3 Results of the implementation

I implemented all four steps of the algorithm for the tetrahedral case and the octahedral case in Python with Jupyter notebook. Please see [14] for the code and for a pdf print-out of the code and its results.
For the cases $(r, d) \in \{(3, \pm 1), (4, \pm 1)\}$ the outcome of my program matches exactly what was presented in Appendix B of [5].
The results of my calculations are presented in the form $[a_0, \ldots, a_k]$. As a quick reminder, we recover our binary form f by letting $f(x_1, x_2) = \sum_{i=0}^{k} \binom{k}{i} a_i x_1^{k-i} x_2^i$, where k is $4, 6, 12$ when r is $3, 4, 5$, respectively. Later in Subsection 5.4, we will see an example of recovering $f, H(f), t(f)$ from the form $[a_0, \ldots, a_k]$.
The following are the results that I obtained for the parameters (r, d) to be (3,1), (3,-1), (3,2), (3,-2), (3,3), (3,-3), (4,1), (4,-1), (4,2), (4,-2), (4,3) and (4,-3).

$(3, 1)$	Representative Point	$(3, -1)$	Representative Point
$[-2, -1, 0, -1, -2]$	$-0.268 + 0.963i$	$[2, 1, 0, 1, 2]$	$-0.268 + 0.963i$
$[-1, 0, -1, 0, 3]$	$1.316i$	$[1, 0, 1, 0, -3]$	$1.316i$
$[-1, 0, 0, -2, 0]$	$1.414i$	$[1, 0, 0, -2, 0]$	$1.414i$
$[-1, 1, 1, 1, -1]$	$-0.268 + 0.963i$	$[1, -1, -1, -1, 1]$	$-0.268 + 0.963i$
$[0, -1, 0, 0, -4]$	$1.414i$	$[0, -1, 0, 0, 4]$	$1.414i$
$[1, 0, -1, 0, -3]$	$1.316i$	$[1, 0, 1, 0, -3]$	$1.316i$

Table 9: Results for $r = 3, d \in \{\pm 1\}$.

The above table matches exactly with results in [6] up to sign changes from actions by U and $-I$. The actions are precisely flipping the sign of all odd-indexed coefficients, or flipping the sign of all coefficients, respectively.
The following two tables present the forms when $r = 3$ and $d \in \{\pm 2, \pm 3\}$. These are not presented in [6].

$(3, 2)$	Representative Point	$(3, -2)$	Representative Point
$[-1, 0, -1, -2, 3]$	$-0.158 + 1.509i$	$[1, 0, 1, 2, -3]$	$-0.158 + 1.509i$
$[0, -1, 0, 0, -8]$	$1.782i$	$[0, -1, 0, 0, 8]$	$1.782i$
$[1, 2, 1, 0, -3]$	$-0.436 + 1.011$	$[-1, -2, -1, 0, 3]$	$-0.436 + 1.011i$

Table 10: Results for $r = 3, d \in \{\pm 2\}$

[14] Please see below for the code and results. For the results, navigate to the end of the pdf document: `https://drive.google.com/drive/folders/1Epnd6sQfXXh7vCH-MNW6o8keXRszf75x?usp=sharing`

$(3,3)$	Representative Point	$(3,-3)$	Representative Point
$[-2,-1,0,-2,-4]$	$-0.318+1.180i$	$[2,1,0,2,4]$	$-0.318+1.180i$
$[-1,-1,0,-2,-8]$	$-0.426+1.580i$	$[1,1,0,2,8]$	$-0.426+1.580i$
$[0,-2,0,0,-3]$	$1.020i$	$[0,-2,0,0,3]$	$1.020i$
$[0,-1,0,0,-12]$	$2.040i$	$[0,-1,0,0,12]$	$2.040i$

Table 11: Results for $r=3, d\in\{\pm 3\}$

Note that when $r=3$, $f\in\mathscr{C}(r,d) \iff -f\in\mathscr{C}(r,-d)$, as t,H,f are covariants of weigh $2,3,0$ respectively, so flipping the sign of d and $z=f$ simultaneously yields the same parametrisation. Therefore, in the above tables for $r=3$, I placed equivalent forms in the same row.
The following tables list the forms for $r=4$.

$(4,1)$	Representative Point	$(4,-1)$	Representative Point
$[-3,-4,-1,0,1,4,3]$	$-0.268+0.963i$	$[-8,-4,-4,-4,-2,1,7]$	$-0.5+0.866i$
$[-1,0,1,0,3,0,-27]$	$1.732i$	$[-5,-1,1,3,3,3,9]$	$-0.436+1.011i$
$[0,-3,0,0,0,4,0]$	$1.075i$	$[-1,0,-1,0,3,0,27]$	$1.732i$
$[0,-1,0,0,0,12,0]$	$1.861i$	$[-1,0,0,-2,0,0,32]$	$1.782i$
		$[-1,1,1,1,-1,5,17]$	$-0.158+1.509i$
		$[0,-3,0,0,0,-4,0]$	$1.075i$
		$[0,-1,0,0,0,-12,0]$	$1.861i$

Table 12: Results for $r=4, d\in\{\pm 1\}$

Again, the above table matches exactly with results in [6] up to sign changes from action by U, $-I$ and US. These are actions that flip the sign of odd-indexed coefficients, flip the sign of every coefficient, and reverse the whole list (swapping x_1, x_2), respectively.
The following two tables present the forms when $r=4, d\in\{\pm 2,\pm 3\}$. These results weres not presented in [6].

$(4,2)$	Representative Point	$(4,-2)$	Representative Point
$[-5,-1,2,2,4,4,-8]$	$-0.310+1.133i$	$[-8,-4,-4,-2,2,5,13]$	$-0.438+0.98i$
$[-1,1,2,2,4,-4,-40]$	$-0.449+1.642i$	$[-4,-1,-2,-3,0,3,18]$	$-0.316+1.231i$
$[0,-3,0,0,0,8,0]$	$1.278i$	$[-1,0,-1,-2,3,4,59]$	$-0.125+1.961i$
$[0,-1,0,0,0,24,0]$	$2.213i$	$[0,-3,0,0,0,-8,0]$	$1.278i$
		$[0,-1,0,0,0,-24,0]$	$2.213i$

Table 13: Results for $r=4, d\in\{\pm 2\}$

$(4,3)$	Representative Point	$(4,-3)$	Representative Point
$[-3,-1,2,0,4,4,-24]$	$1.414i$	$[-15,-8,-5,0,5,8,15]$	$-0.5+0.866i$
$[0,-4,0,0,0,9,0]$	$1.225i$	$[-8,-3,-4,-4,0,4,16]$	$-0.385+1.04i$
$[0,-1,0,0,0,36,0]$	$2.449i$	$[-2,-1,0,-3,-6,-9,36]$	$-0.374+1.691i$
		$[-1,0,1,4,3,8,101]$	$-0.478+2.077i$
		$[0,-4,0,0,0,-9,0]$	$1.225i$
		$[0,-1,0,0,0,-36,0]$	$2.449i$

Table 14: Results for $r=4, d\in\{\pm 3\}$

5.4 How to use the outputs

In this subsection, we will give two example of how to turn the output of the algorithm to the binary form $f(x_1, x_2)$, and then how to use $f(x_1, x_2)$ to obtain the triple (X, Y, Z) to solve our equation.
We will work with two examples.

Example 5.2:
Consider the form $[a_1, \ldots, a_k] = [1, 2, 1, 0, -3] \in \mathscr{C}(3, 2)$, which can be obtained from the table $(r, d) = (3, 2)$ above. This is in the tetrahedral case. So $r = 3$, and $k = 4$.
First, we recover f by computing

$$f(x_1, x_2) = \sum_{i=0}^{4} \binom{4}{i} a_i x_1^{4-i} x_2^i = x_1^4 + 8x_1^3 x_2 + 6x_1^2 x_2^2 - 3x_2^4$$

Then, we compute $t(f)$ and $H(f)$ by using the formulas of the Hessian and the functional determinant in Definition 2.9. In my case, I used the Sympy library in Python.
They turned out to be

$$H(f) = -3x_1^4 - 4x_1^3 x_2 - 6x_1^2 x_2^2 - 12x_1 x_2^3 - 3x_2^4$$
$$t(f) = 10x_1^6 + 12x_1^5 x_2 - 30x_1^4 x_2^2 - 120x_1^3 x_2^3 - 90x_1^2 x_2^4 - 36x_1 x_2^5 - 18x_2^6$$

As a quick check, we must have

$$H(f)^3 + \left(\frac{t(f)}{2}\right)^2 + df^3 = 0$$

Upon checking, this is indeed the zero polynomial. One can also perform this check using the Sympy library in Python.
Now, we can obtain some values of (X, Y, Z) by computing integer specialisation of $t(f)/2, H(f), f$ by s_1, s_2:

$$X = \frac{t(f)(s_1, s_2)}{2}, Y = H(f)(s_1, s_2), Z = f(s_1, s_2)$$

The following are some examples of integer specialisations that yields solutions to $X^2+Y^3+2Z^3 = 0$ where $\gcd(X, Y, Z) = 1$.

s_1	s_2	X	Y	Z
1	0	5	-3	1
2	1	-433	-131	101
4	3	$-18,1705$	$-3,939$	$2,413$
8	5	$-5,472,865$	$-46,003$	$32,301$

Table 15: Some integer solutions to $X^2 + Y^3 + 2Z^3 = 0$

Example 5.3:
Now, we show an example in the octahedral case.
Consider the form $[-1, 0, 1, 4, 3, 8, 101]$ in $\mathscr{C}(4, -3)$. This form can be found in the table above for $(r, d) = (4, -3)$.

We recover f by computing the following:

$$f(x_1, x_2) = \sum_{i=0}^{6} \binom{6}{i} a_i x_1^{6-i} x_2^i = -x_1^6 + 15x_1^4x_2^2 + 80x_1^3x_2^3 + 45x_1^2x_2^4 + 48x_1x_2^5 + 101x_2^6$$

By using the formulas for the Hessian and the functional determinant in Definition 2.9, we obtain the following:

$$H(f) = -x_1^8 - 16x_1^7x_2 - 28x_1^6x_2^2 - 112x_1^5x_2^3 - 406x_1^4x_2^4 - 112x_1^3x_2^5 + 644x_1^2x_2^6 + 1520x_1x_2^7 + 239x_2^8$$

$$\begin{aligned} t(f) = {} & 4x_1^{12} + 24x_1^{11}x_2 + 264x_1^{10}x_2^2 + 1496x_1^9x_2^3 + 1980x_1^8x_2^4 + 1584x_1^7x_2^5 + 3696x_1^6x_2^6 + 20592x_1^5x_2^7 \\ & + 78012x_1^4x_2^8 + 144056x_1^3x_2^9 + 50952x_1^2x_2^{10} - 19272x_1x_2^{11} - 34556x_2^{12} \end{aligned}$$

Again, we check that $-3f^4 + H(f)^3 + \left(\frac{t(f)}{2}\right)^2$ is the zero polynomial.
The values of (X, Y, Z) are obtained by computing the integer specialisation of $t(f)/2, H(f), f$ by s_1, s_2 as follows:

$$X = \frac{t(f)(s_1, s_2)}{2}, Y = H(f)(s_1, s_2), Z = f(s_1, s_2)$$

The following are some integer specialisations that gives solution to $X^2 + Y^3 - 3Z^4 = 0$, where $\gcd(X, Y, Z) = 1$.

s_1	s_2	X	Y	Z
1	0	2	-1	-1
1	2	$-16,070,038$	$285,839$	$9,419$
2	3	$9,721,202,906$	$9,231,263$	$130,913$
5	4	$5,058,523,244,882$	$92,040,479$	$1,721,831$

Table 16: Some integer solutions to $X^2 + Y^3 - 3Z^4 = 0$

I did not implement the algorithm in the icosahedral case. However, the process of obtaining solutions to $X^2 + Y^3 + dZ^5 = 0$ from an output of the algorithm would be analogues to the previous two examples. This concludes Section 5. In this section, we presented Edwards's algorithm, our implementation, some of its outputs, as well as how to turn the output of the algorithm into a solution that solves (4).

6 A Geometric Approach for the Diophantine Equation

This section discusses Beukers's work in [3]. In Subsection 6.1, we will introduce some ideas in Beukers's approach and compare them to Edwards's approach. In subsection 6.2 we will point how the algorithm introduced Section 5 serves as a proof of a specialised version of Beuker's Theorem.

6.1 Beukers's Theorem

In this subsection, we will present Beukers's theorem and discuss some ideas of the proof. Considering the length of the essay, I decided to skip most of the details. The purpose of this section is to comment on Beukers's work and point out similarities between Beukers's approach and Edwards's approach.

Theorem 6.1 (Beukers's Theorem):
Recall the generalised Fermat equation in the spherical case:

$$AX^p + BY^q + CZ^r = 0, \gcd(X, Y, Z) = 1, XYZ \neq 0 \tag{11}$$

where $p, q, r, A, B, C \in \mathbb{Z}, p, q, r \geq 2$ are fixed and $X, Y, Z \in \mathbb{Z}$ are the unknowns.
Then

- There exists a finite number of parametrised solutions to the above equation and every solution of the above equation is a specialisation of one of the parametrised solutions.
- If the above equation has one solution $(X, Y, Z) \in \mathbb{Z}^3$, then, it has infinitely many solutions in $\mathbb{Z}$.

The setup of Beukers's work is as follows.
Let $G \subset \mathrm{GL}_n(\mathbb{C})$ be a finite subgroup. We consider the action of G on the polynomial ring $\mathbb{C}[X_1, \ldots, X_n]$ by letting $G \cdot f(X_1, \ldots, X_n) = f(X_1', \ldots, X_n')$, where $G[X_1, \ldots, X_n]^T = [X_1', \ldots, X_n']$. Now, we consider $\mathbb{C}[X_1, \ldots, X_n]^G$, which is the ring of polynomials that are G-invariant. Noetherian Normalisation Lemma says that this ring is finitely generated.
Now, let $I_1, \ldots, I_r$ be a set of homogenous polynomials that generates the ring $\mathbb{C}[X_1, \ldots, X_n]^G$, and let φ be defined as follows:

$$\varphi : \mathbb{C}^n \to \mathbb{C}^r$$
$$\vec{X} \to (I_1(\vec{X}), \ldots, I_r(\vec{X})), \text{ where } \vec{X} = (X_1, \ldots, X_n)$$

Below is an example from [3]:

Example 6.2: Consider the cyclic group $G \subset \mathrm{GL}_2(\mathbb{C})$ generated by $\begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$[15] Then, $\mathbb{C}[x_1, x_2]^G$ is generated by

- $I_1 = x_1x_2(x_1 + x_2)$
- $I_2 = x_1^2 + x_1x_2 + x_2^2$
- $I_3 = (x_1 - x_2)(2x_1^2 + 5x_1x_2 + 2x_2^2)$

In Beukers's work, he studied the image of φ as an affine variety V, and he used more theory in algebraic geometry to place some assumptions about V as a variety. We also assume that V can be defined over $\mathbb{Q}$. Along with some other restrictions, the following is our main subject of interest:
Suppose our variety V is given by a set of equations:

$$f_1(\vec{Y}) = f_2(\vec{Y}) = \ldots = f_k(\vec{Y}) = 0, \vec{Y} = (Y_1, \ldots, Y_r)$$

and that $f_i \in \mathbb{Z}[Y]$ with

$$f_i(I_1(\vec{X}), I_2(\vec{X}), \ldots I_r(\vec{X})) = 0$$

then we wish to solve the following set of equations with some additional restrictions about the variety V that we won't mention here:

$$f_i(\vec{y}) = 0, i = 1, \ldots, k \tag{12}$$

[15]Note that in [3], the matrix is $\begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}^T$ instead. The polynomials I_1, I_2, I_3 are invariant under the matrix we presented in the essay.

In other words, our main goal is to study the solutions to the diophantine equations $f_1, \ldots, f_k$. In this case, $\varphi : \mathbb{C}^n \to \mathbb{C}^r$, component wise reads $I_1, \ldots, I_r$, plays the role of parameterizations that are specialized by vector $\vec{X}$: given a vector $\vec{X}$, it turns $\vec{X}$ into a vector $\vec{Y}$ that solves our desired system of diophantine equations f_i.
In this case, φ plays the role of a parametrised of solution to f_i. For example, the map

$$(x_1, x_2) \mapsto (I_1(x_1, x_2), I_2(x_1, x_2), I_3(x_1, x_2)) = (y_1, y_2, y_3)$$

serves as a family of parameterised solution to $y_1^2 + 27y_2^2 - 4y_3 = 0$.
We may ask ourselves: as introduced in 3, binary forms can be $\mathrm{SL}_2(\mathbb{Z})$ equivalent to another binary form, and every binary form is $\mathrm{SL}_2(\mathbb{Z})$ equivalent to a reduced binary form. In this case, although the $I_1, \ldots, I_r$ are not necessarily binary forms, do we have anything similar with respect to the parametrisation φ?
Indeed, in Beukers's work, there is a similar idea. In his work, $\mathbb{Q}$-matrices are defined as the matrices in $\mathrm{GL}_n(\bar{\mathbb{Q}})$ such that $\varphi \cdot m$ is defined over $\mathbb{Q}$. Equivalence of φ and $\varphi \cdot m$ is analogues to the idea of $\mathrm{SL}_2(\mathbb{Z})$ equivalence in Edwards's approach.
We next present two key propositions in Beukers's work that leads to the proof of Beukers's theorems:

Proposition 6.3: Let $\vec{y}$ be a solution to (12), then there exists a $\mathbb{Q}-$matrix m and $\vec{s} \in \mathbb{Q}^n$, such that $y = (\varphi \cdot m)(\vec{s})$. [16]

This proposition plays the role of the lifting theorem (Theorem 4.20) in Section 4. They both start with a solution to the diophantine equation, then 'lifts' it to a parametrisation. Except that in Beuker's approach, the techniques are from algebraic number theory, whereas in Edwards's approach, the techniques are algebraic.

Proposition 6.4: There exists a set of $\mathbb{Q}-$matrices, denoted by M, such that whenever $\vec{y}$ is a solution to (12), then there exist a matrix $m \in M$, and vector $\vec{s} \in \mathbb{Q}^n$ such that $\vec{y} = (\varphi \cdot m)(\vec{s})$.[17]

The proposition above states that there are most finitely many $\mathbb{Q}-$matrices, such that every solutions to the diophantine equation can be written as $(\varphi \cdot m)(\vec{s})$. This proposition can be compared to Theorem 4.19 in Edwards's approach, which states that there are finite number of reduced binary form that are also Klein forms.
As we can see, using different techniques and working in different settings, both Edwards's and Beukers's approaches consists of 'lifting' a solution of the diophantine equation to a family, then showing that the number of families is finite.

6.2 Relation with Edwards's Algorithm

The correctness of the algorithm introduced in Section 5 serves as a proof to a specialised version of Beukers's theorem:

Theorem 6.5 (Specialised version of Beukers's Theorem): Consider the diophantine equations below in the spherical case with unknowns $X, Y, Z \in \mathbb{Z}$

$$X^2 + Y^3 + CZ^r = 0, \gcd(X, Y, Z) = 1, XYZ \neq 0$$

where $C \in \mathbb{Z}$ and $r \in \mathbb{Z}, r \geq 2$.

[16] Some details of this proposition are intentionally left out.
[17] Some details of this proposition are intentionally left out.

Then there exists a finite number of parametrised solutions, $X = P_X(x_1, x_2), Y = P_Y(x_1, x_2), Z = P_Z(x_1, x_2)$, to the above diophantine equation.

Proof : The proof follows from the correctness of Edwards's algorithm. □

7 Conclusion

The following is a brief summary of this essay. In Section 2, we introduced the Riemann sphere, the Möbius transformation on it, and how its groups of rotation relates to the platonic solids via Klein's theorem. In Section 3, we introduced quadratic binary forms and some results, then extended it to Hermite reduction theory, which generalizes the results to forms of degree greater than 2. In Section 6, we presented Beukers's theorem and introduced some of its proof ideas. In Section 4, we saw how Edwards studied a special case of equations in Beukers's theorem as well as applied invariant theory and Hermite reduction theory to them. In Section 5, we studied and implemented Edwards's algorithm and produced a complete set of parametrisations of reduced Klein forms under $\mathrm{GL}_2(\mathbb{Z})$ equivalences.
The following are three directions that an interested reader can explore:
Direction 1. One can compute and plot a diagram of how the number of families of parametrised solutions vary with respect to d. One may be interested to explore whether there is a pattern between the number of families and the value d.
Direction 2. In [6], Edwards mentioned that it took 6 hours to produce the list of 27 forms for the equation $X^2 + Y^3 + Z^5 = 0$. One may be interested in finding methods to reduce the runtime of this algorithm.
Direction 3. As mentioned in the Introduction, there are many interesting studies in the hyperbolic case of the generalised Fermat equations. We refer the interested readers to [2].
Overall, we studied some theory and methods that allowed us to generates the families of parametrised solutions for the generalised Fermat equations in the spherical case. We learned about the rotation groups of platonic solids, binary forms, Hermite reduction theory, and we saw them coming together in an algorithm.

References

[1] URL: http://www.warwickmaths.com/wp-content/uploads/2020/07/80_-M%C3%B6bius-Transformations.pdf (cit. on p. 6).

[2] Michael Bennett, Preda Mihailescu, and Samir Siksek. "The generalized Fermat equation". In: *Open Problems in Mathematics* (2016), pp. 173–205. DOI: 10.1007/978-3-319-32162-2_3 (cit. on pp. 3, 41).

[3] Frits Beukers. "The diophantine equation axp+byq=czr". In: *Duke Mathematical Journal* 91.1 (1998), pp. 61–88. DOI: 10.1215/s0012-7094-98-09105-0 (cit. on pp. 3–5, 38, 39).

[4] Henri Darmon and Andrew Granville. "On the equations Zm = F(x,y) and Axp+Byq = Czr". In: *Bulletin of the London Mathematical Society* 27.6 (1995), pp. 513–543. DOI: 10.1112/blms/27.6.513 (cit. on p. 3).

[5] Edward Jonathan Edwards. "Platonic Solids and Solutions to X2 + Y3 = dZr". PhD thesis. 2005 (cit. on pp. 26, 29, 35).

[6] Johnny Edwards. "A complete solution to X2 + y3 + z5 = 0". In: *Journal für die reine und angewandte Mathematik (Crelles Journal)* 2004.571 (2004), pp. 213–236. DOI: 10.1515/crll.2004.043 (cit. on pp. 4, 10–17, 20–23, 25, 28–30, 33–36, 41).

[7] Felix Klein. *Lectures on the icosahedron and the solution of equations of the fifth degree*. Dover, 1956 (cit. on pp. 7–10).

[8] Rob Siliciano. "Constructing Mobius Transformations with Spheres". In: *Rose-Hulman Undergraduate Mathematics Journal* 13.2 (2012), pp. 116–124. URL: https://scholar.rose-hulman.edu/cgi/viewcontent.cgi?article=1218&context=rhumj (cit. on p. 6).

[9] William Stein. *Lecture 23: Quadratic Forms III Reduction Theory*. URL: https://wstein.org/edu/Fall2001/124/lectures/lecture23/lecture23.pdf (cit. on pp. 11, 12).

[10] Michael Stoll and John E. Cremona. "On the reduction theory of binary forms". In: *Journal für die reine und angewandte Mathematik (Crelles Journal)* 2003.565 (2003). DOI: 10.1515/crll.2003.106 (cit. on p. 13).