Minkowski Theory and Rational Points on Hyperelliptic Curves

by

Shanzhao Wang

B.Sc., University of Toronto, 2015

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
Department of Mathematics
Faculty of Science

SIMON FRASER UNIVERSITY
Fall 2018

Approval

Name:	**Shanzhao Wang**
Degree:	**Master of Science (Mathematics)**
Title:	**Minkowski Theory and Rational Points on Hyperelliptic Curves**
Examining Committee:	**Chair:** Ladislav Stacho Associate Professor
	Imin Chen Senior Supervisor Associate Professor
	Stephen Choi Supervisor Professor
	Nils Bruin Internal Examiner Professor
Date Defended:	**September 18, 2018**

Abstract

A remarkable recent result of M. Bhargava shows in a certain precise sense that 'most' hyperelliptic curves over $\mathbb{Q}$ have no rational points. An object central to his proof is a certain representation $\mathbb{Z}^2 \otimes \mathrm{Sym}_2 \mathbb{Z}^2$ of $\mathrm{GL}_n(\mathbb{Z})$. Elements in the set $\mathbb{Z}^2 \otimes \mathrm{Sym}_2 \mathbb{Z}^2$ can be viewed as a pair of $n \times n$ symmetric matrices with entires in $\mathbb{Z}$ up to a $\mathrm{GL}_n(\mathbb{Z})$ equivalence.

Alternatively, $\mathbb{Z}^2 \otimes \mathrm{Sym}_2 \mathbb{Z}^2$ has an algebraic number theoretic description. By taking advantage of this property, in this thesis, we investigate $\mathbb{Z}^2 \otimes \mathrm{Sym}_2 \mathbb{Z}^2$ from the point of view of Minkowski theory. In particular, by assuming the hyperelliptic curve C is given by $z^2 = f(x, y)$, where $f(x, y)$ is irreducible over $\mathbb{Q}$, we gave a direct 'Minkowski' style proof that a certain part of the set $\mathbb{Z}^2 \otimes \mathrm{Sym}_2 \mathbb{Z}^2$, which contains the elements arising from the rational points on C, is finite.

Although, our main principle of proof mirrors the classical proof of finiteness for the class number of a number field, we develop new arguments when there exist some notable differences, and we strive to give self-contained proofs of some of the components of Bhargava's paper which we utilize.

Keywords: Rational points, hyperelliptic curves, Bhargava theory, Minkowski theory.

Acknowledgements

I am thankful Lord for everything that You allow to cross my path. Thankful for the decisions that You allow me to make and the lessons that come from these decisions, especially for the three years I spent at SFU. Also, I would like to express my deepest appreciation to all those who provided me the possibility to complete this thesis. A special gratitude I give to my senior supervisor: Dr. Imin Chen whose contribution in stimulating suggestions and encouragement, helped me in writing this thesis. Last but not least, many thanks go to my family, my friends and Tiantian.

Table of Contents

Chapter 1

Introduction

A hyperelliptic curve over $\mathbb{Q}$ is a smooth, geometrically irreducible, complete curve C over $\mathbb{Q}$ equipped with a fixed map of degree 2 to $\mathbb{P}^1$. Explicitly, we view any hyperelliptic curve over $\mathbb{Q}$ of genus g as an equation of the form

$$C : z^2 = f(x, y) = f_0 x^n + f_1 x^{n-1} y + \ldots + f_{n-1} x y^{n-1} + f_n y^n \tag{1.1}$$

where $n = 2g + 2$, the coefficients f_i lie in $\mathbb{Z}$, and f factors into distinct linear factors over $\bar{\mathbb{Q}}$.

A $\mathbb{Q}$-*rational point* on C is a triple $(x_0, y_0, z_0) \neq (0, 0, 0) \in \mathbb{Q}^3$ such that $z_0^2 = f(x_0, y_0)$. Define the height $H(C)$ of C by

$$H(C) := H(F) := \max\left\{|f_i|\right\}. \tag{1.2}$$

In [3], Bhargava proved the following remarkable result:

Theorem 1.1. *As $g \to \infty$, a density approaching 100% of hyperelliptic curves over $\mathbb{Q}$ of genus g possess no $\mathbb{Q}$-rational points.*

An object which is central to the proof of Theorem 1.1 is the representation $\mathbb{Z}^2 \otimes \mathrm{Sym}_2 \mathbb{Z}^n$ of $\mathrm{GL}_n(\mathbb{Z})$. An element v of $\mathbb{Z}^2 \otimes \mathrm{Sym}_2 \mathbb{Z}^n$ can be viewed as a pair (A, B) of symmetric $n \times n$ matrices with entries in $\mathbb{Z}$. Then an element $g \in \mathrm{GL}_n(\mathbb{Z})$ acts on (A, B) by the formula $g \cdot (A, B) = (gAg^t, gBg^t)$. To such a pair $v = (A, B)$, we may associate a binary form f_v of degree n, given by

$$f_v(x, y) = (-1)^{\frac{n}{2}} \det(Ax - By). \tag{1.3}$$

The coefficients of f_v in fact generate the ring of polynomial invariants for the action of $\mathrm{GL}_n(\mathbb{Z})$ on $\mathbb{Z}^2 \otimes \mathrm{Sym}_2 \mathbb{Z}^n$ (see, e.g., the work of Schwarz [14]), and f_v is called the invariant binary n-ic form associated to $v \in \mathbb{Z}^2 \otimes \mathrm{Sym}_2 \mathbb{Z}^n$.

The orbits of $\mathrm{GL}_n(\mathbb{Z})$ on $\mathbb{Z}^2 \otimes \mathrm{Sym}_2 \mathbb{Z}^n$ were first considered in the case $n = 2$ by Hardy and Williams [8] and more generally by Morales [10, 11]. A classification of the orbits in the cases $n = 2$ and $n = 3$, in terms of ideal classes in quadratic and cubic rings, was given in

[1] and [2], while a complete classification for general n in terms of module classes of rings of rank n was given by Wood [16].

The key algebraic construction used in the proof of Theorem 1.1 and Corollary 1.2 is the observation that a $\mathbb{Q}$-rational point on C gives rise to an element $v \in \mathbb{Z}^2 \otimes \mathrm{Sym}_2 \mathbb{Z}^n$ such that $f_v = f$. Bhargava then shows by intricate geometry of numbers counting arguments that, for 'most' integral binary n-ic forms f (in the sense of Theorem 1.1), there do not exist any such integral orbits with invariant binary form equal to f.

The key algebraic construction above also extends to showing that an element of the fake 2-Selmer set of C gives rise to a $\mathrm{GL}_n(\mathbb{Z})$ orbit of $\mathbb{Z}^2 \otimes \mathrm{Sym}_2 \mathbb{Z}^n$ having invariant form f. Thus, Theorem 1.1 implies

Corollary 1.2. *As $g \to \infty$, a density approaching 100% of hyperelliptic curves over $\mathbb{Q}$ of genus g possess an empty fake 2-Selmer set.*

The fake 2-Selmer set of C can be used to as a criterion to determine if C has no $\mathbb{Q}$-rational points [4]. Hence, the above corollary implies for 'most' hyperelliptic curves C over $\mathbb{Q}$, the method in [4] succeeds in establishing C has no $\mathbb{Q}$-rational points.

In this thesis, we investigate the representation $\mathbb{Z}^2 \otimes \mathrm{Sym}_2 \mathbb{Z}^2$ from the point of view of Minkowski theory. This is possible because $\mathbb{Z}^2 \otimes \mathrm{Sym}_2 \mathbb{Z}^2$ has an algebraic number theoretic description (see paragraph after Definition 4.8).

In particular, we show how to carry a direct 'Minkowski' style proof that a certain part of the set $\mathbb{Z}^2 \otimes \mathrm{Sym}_2 \mathbb{Z}^2$ (which contains the elements arising from $\mathbb{Q}$-rational points on C) is finite (Theorem 5.8). For simplicity, we only treat the case that $f \in \mathbb{Z}[x, y]$ is irreducible in $\mathbb{Q}[x, y]$.

While the principle of proof mirrors the classical proof of finiteness for the class number of a number field, some notable differences arise which require new arguments. Additionally, we strive to give self-contained proofs of some of the components from [3] which we require to prove our main theorems.

1.1 Terminology and Notation

1. $\mathbb{R}^+$ is the set of positive real numbers.
2. $\mathbb{Z}^+ = \mathbb{N}$ is the set of positive integers.
3. Let N be a $\mathbb{Z}$-submodule of M. Then $(M : N)$ is the index of $\mathbb{Z}$-module N in M.
4. Let K be a field and K^* be the multiplicative subgroup of all nonzero elements of K.
5. $\mathcal{O}$ is an order in the number field K.
6. f denotes an homogeneous polynomial in $\mathbb{Z}[x, y]$ of even degree n which is irreducible in $\mathbb{Q}[x, y]$
7. K_f is the number field generated by $\mathbb{Q}[x]/(f)$ and R_f is the order associated with f.
8. $N_f(I) := N_{R_f}(I)$ is the norm of I with respect to the order R_f.
9. $\langle \alpha_1, \ldots, \alpha_m \rangle$ denotes $\mathbb{Z}$-module generated by $\alpha_1, \ldots, \alpha_m$ lying a $\mathbb{Z}$-module M.
10. $\mathrm{GL}_n(R)$ denotes the general linear group of degree n over a ring R.

Chapter 2

$\mathbb{Z}$-modules

In this chapter, we review some basic results on modules over $\mathbb{Z}$, or more generally, a principal ideal domain R.

Lemma 2.1. *Let R be a principal ideal domain and M be a finitely generated R-module. Then any submodule of M is also finitely generated. Moreover, if M is generated by n elements, then any submodule of M can also be generated by n elements.*

Proof. See [5, Corollary 2, §10.6]. □

Theorem 2.2. *Let R be a principal ideal domain, then any submodule of R^n is a direct sum of cyclic modules. More precisely, if M is a finitely generated R-module, then*

$$M \cong R^s \oplus R/a_1R \oplus R/a_2R \oplus \ldots \oplus R/a_rR, \tag{2.1}$$

where the a_i's are non-zero non-units and

$$a_i | a_{i+1}, i = 1, 2, ..., r-1; \tag{2.2}$$

the decomposition (2.1)*, subject to* (2.2) *is unique, up to isomorphism.*

Proof. We follow the proof in [5, Theorem 2, §10.6] by using the presentation (see Lemma 2.1 above):

$$0 \to R^m \to R^n \to M \to 0,$$

where $m \leq n$, and the mapping $f : R^m \to R^n$ can be represented by an $m \times n$ matrix $A = (a_{ij})$. Thus by a suitable choice of bases we can take A in the diagonal form and it leads to the form (2.1) for M, where $s = n - m$ if we omit terms corresponding to units. The a_i are unique as the invariant factors of A. □

Lemma 2.3. *Let $M' \subseteq M$ both be free $\mathbb{Z}$-modules of rank n with bases $\{y_1, y_2, \ldots, y_n\}$ and $\{x_1, x_2, \ldots, x_n\}$ respectively. If we write $y_i = \sum_{i=1}^{n} c_{ij} x_i$ with $c_{ij} \in \mathbb{Z}$, then the index $(M : M')$ equals $|\det(c_{ij})|$.*

Proof. See [7, Theorem §4.15]. □

Definition 2.4. ***Bound for a quotient module***

Let N be a $\mathbb{Z}$-submodule of a free $\mathbb{Z}$-module M. If there exists a positive integer δ, such that $\delta M \subseteq N$, we say that M/N is bounded by δ.

Lemma 2.5. *Let $N \subseteq M$ be free $\mathbb{Z}$-modules both of rank n. Then M/N is bounded by δ, where $\delta = (M : N)$.*

Proof. Recall the proof of Theorem 2.2 proceeds by picking isomorphisms $M \cong \mathbb{Z}^n$ and $N \cong \mathbb{Z}^n$. Then the inclusion $N \subseteq M$ corresponds to a $\mathbb{Z}$-module homomorphsim $\phi : \mathbb{Z}^n \to \mathbb{Z}^n$, which in turn, can be represented by a $n \times n$ matrix A with integer entries.

By the elementary divisor theorem [5, Theorem 2, §10.6], by changing bases for the source and target copies of $\mathbb{Z}^n$, the matrix A can be put in diagonal form with entries $a_1, \ldots, a_n$ such that $a_i | a_{i+1}$ for $i = 1, \ldots, n-1$. Thus, M/N has the structure

$$M/N \cong \mathbb{Z}/a_1\mathbb{Z} \oplus \mathbb{Z}/a_2\mathbb{Z} \ldots \oplus \mathbb{Z}/a_n\mathbb{Z} \tag{2.3}$$

where $a_i | a_{i+1}$ for $i = 1, ..., n-1$.

If we multiply $\delta = a_1 \cdot a_2 \cdot \ldots \cdot a_n = (M : N)$ to (2.3), we will have $\delta M/N = 0$, which implies M/N is bounded by δ. □

Corollary 2.6. *Let $N \subseteq M$ be free $\mathbb{Z}$-modules, both of rank n. If there exists a positive integer δ' such that $\delta' M \subseteq N$, then we have $(M : N) | \delta'^n$.*

Proof. M/N has the same structure as in (2.3), and if $\delta' M \subseteq N$, we have that $a_i | \delta'$ for $i = 1, \ldots, n$. Hence, $(M : N) = a_1 \cdot a_2 \cdot \ldots \cdot a_n \Big| (\delta')^n$. □

Chapter 3

Minkowski Theory for Orders

Definition 3.1. ***Norm*** *and* ***Trace***

Let $L|K$ be a field extension of number field K. The ***trace*** *and* ***norm*** *of an element $x \in L^*$ are defined to be the trace and determinant of the endomorphism*

$$T_x : L \longrightarrow L, \quad T_x(\alpha) = x\alpha,$$

of the K-vector space L:

$$Tr_{L|K}(x) = \mathrm{T}r(T_x),$$

$$N_{L|K}(x) = \det(T_x).$$

Definition 3.2. ***Integrality***, ***Integral Bases*** *and* ***Integral Closure***

Let $f : A \longrightarrow B$ be a ring homomorphism. An element $b \in B$ is called ***integral*** *over A, if it satisfies a monic polynomial*

$$x^n + a_1 x^{n-1} + \ldots + a_n = 0$$

with coefficients $a_i \in A$ and degree $n \geq 1$.

A system of integral elements $\omega_1, \omega_2, ..., \omega_n$ in B is called an ***integral basis*** *of B over A if and only if for any $b \in B$,*

$$b = a_1\omega_1 + ... + a_n\omega_n$$

for unique determined coefficients $a_i \in A$.

If all elements in B are integral over A and f is injective, we call B is an ***integral closure*** *$\bar{A}$ of A. Notice that B can also be called an integral extension of A and denote it by B/A.*

Utilizing Definition 3.2, we can have another way to compute **trace** and **norm** of the element x in Definition 3.1.

Proposition 3.3. *If $L|K$ is a separable extension and $\sigma : L \longrightarrow \bar{K}$ varies over different K-embeddings over L into algebraic closure $\bar{K}$ of K, then we have*

$$Tr_{L|K}(x) = \sum_{\sigma} \sigma x$$
$$N_{L|K}(x) = \prod_{\sigma} \sigma x$$

Proof. See [13]. □

Definition 3.4. ***Ring of Integers*** $\mathcal{O}_K$

The ring of integers $\mathcal{O}_K \subseteq K$ is defined to be the integral closure of $\mathbb{Z} \subset \mathbb{Q}$ in an algebraic number field.

Definition 3.5. ***Order*** $\mathcal{O}$

*Let K be a number field of degree n. An **order** of K is a subring $\mathcal{O}$ of $\mathcal{O}_K$ which has an integral basis over $\mathbb{Z}$ of length n. The ring $\mathcal{O}_K$ is called the maximal order of K. In concrete terms, orders are obtained as ring of the form*

$$\mathcal{O} = \mathbb{Z}[\alpha_1, \ldots, \alpha_n],$$

where $\alpha_1, \ldots, \alpha_n$ are all algebraic integers such that $K = \mathbb{Q}[\alpha_1, \ldots, \alpha_n]$.

Definition 3.6. ***Discriminant*** $d(\alpha_1, \ldots, \alpha_n)$

*The **discriminant** of a basis $\alpha_1, ..., \alpha_n$ of a separable extension $L|K$ is defined by*

$$d(\alpha_1, ..., \alpha_n) = \det\left((\sigma_i \alpha_j)\right)^2,$$

where $\sigma_i, i = 1, 2, ..., n$ varies over the K-embeddings of $L \hookrightarrow \bar{K}$. Because of the relation

$$Tr_{L|K}\left(\alpha_i \alpha_j\right) = \sum_{k} \left(\sigma_k \alpha_i\right)\left(\sigma_k \alpha_j\right),$$

the matrix $\left(Tr_{L|K}(\alpha_i \alpha_j)\right)$ is the product of the matrix $(\sigma_k \alpha_i)^T$ and $(\sigma_k \alpha_j)$. Thus one may also write

$$d(\alpha_1, ..., \alpha_n) = \det\left(Tr_{L|K}(\alpha_i \alpha_j)\right).$$

We now recall some basic theory involving the ring of integers $\mathcal{O}_K$, which we at the same time generalize to a general order $\mathcal{O}$, where possible.

Lemma 3.7. *Suppose $\{\alpha_1, \ldots, \alpha_n\}$ is a basis of K over $\mathbb{Q}$, which is contained in an order $\mathcal{O}$ in K and let $d = d(\alpha_1, \ldots, \alpha_n)$ be the discriminant. Then one has*

$$d\mathcal{O} \subseteq \mathbb{Z}\alpha_1 + \ldots + \mathbb{Z}\alpha_n.$$

Proof. We know from Lemma 2.9 in [13] that if $\{\alpha_1, \ldots, \alpha_n\}$ is a basis of K over $\mathbb{Q}$, which is contained in $\mathcal{O}_K$, and has discriminant $d = d(\alpha_1, \ldots, \alpha_n)$, then,

$$d\mathcal{O}_K \subseteq \mathbb{Z}\alpha_1 + \ldots + \mathbb{Z}\alpha_n.$$

As $d\mathcal{O} \subseteq d\mathcal{O}_K$, we have the result of Lemma 3.7. □

Lemma 3.8. *Let $\mathcal{O}$ be an order in K. For any $k \in K$, k can be written as $\frac{a}{z}$, where $a \in \mathcal{O}$ and $z \in \mathbb{Z}, z \neq 0$.*

Proof. Since $K|\mathbb{Q}$ is an algebraic extension, $k \in K$ satisfies

$$a_n k^n \ + a_{n-1}k^{n-1} + a_{n-2}k^{n-2} \ + \ldots \ + a_0 = 0, \tag{3.1}$$

where $a_0, a_1, \ldots, a_n \in \mathbb{Z},\ a_n \neq 0$.

Multiplying a_n^{n-1} on both sides of (3.1), we have

$$(a_n k)^n + a_{n-1}(a_n k)^{n-1} + \ldots + a_0 a_n^{n-1} = 0,$$

which implies $a_n k \in \mathcal{O}_K$.

From lemma 2.5, we know the nonzero $\delta = (\mathcal{O}_K : \mathcal{O}) \in \mathbb{Z}^+$ makes $\delta a_n k \in \mathcal{O}$. So we obtain k as $\frac{a}{z}$, for an $a \neq 0 \in \mathcal{O}$ and $z = \delta a_n \in \mathbb{Z}$. □

Theorem 3.9. *Let M be a nonzero finitely generated $\mathcal{O}$-module in K. Then M is a free $\mathbb{Z}$-module of rank $(K : \mathbb{Q})$.*

Proof. Since M is a nonzero finitely generated $\mathcal{O}$-module in K, it has a generating set $S = \{\mu_1, \ldots, \mu_r\}$.

By Lemma 3.8, there exists a nonzero integer z such that $z\mu_i \in \mathcal{O}, \forall \mu_i \in S$. Therefore $zM \subseteq \mathcal{O}$. Then from Lemma 3.7,

$$dzM \subseteq d\mathcal{O} \subseteq \mathbb{Z}\alpha_1 + \ldots + \mathbb{Z}\alpha_n,$$

where $d = d(\alpha_1, \ldots, \alpha_n)$ is the discriminant and $\{\alpha_1, \ldots, \alpha_n\}$ is a basis of $K|\mathbb{Q}$.

Let $M_0 := \mathbb{Z}\alpha_1 + \ldots + \mathbb{Z}\alpha_n$. By Theorem 2.2, $dzM \subseteq M_0$ is a free $\mathbb{Z}$-module, and hence also M, as the map

$$M \longrightarrow dzM$$
$$m \longmapsto dzm$$

is an isomorphism between $\mathbb{Z}$-modules. Thus,

$$rank_{\mathbb{Z}}(M) = rank_{\mathbb{Z}}(dzM) \leq rank_{\mathbb{Z}}(M_0) = (K : \mathbb{Q}).$$

So finally,

$$n = (K : \mathbb{Q}) = rank_{\mathbb{Z}}(\mathcal{O}) \leq rank_{\mathbb{Z}}(M) = rank_{\mathbb{Z}}(dzM) \leq rank_{\mathbb{Z}}(M_0) = (K : \mathbb{Q}) = n.$$

□

Definition 3.10. *The discriminant of an order $\mathcal{O}$ is defined as $\Delta(\mathcal{O}) := d(\alpha_1, \ldots, \alpha_n)$, where $\{\alpha_1, \ldots, \alpha_n\}$ is a $\mathbb{Z}$-basis for $\mathcal{O}$.*

Lemma 3.11. *If $\mathcal{O}$ is an order, $\Delta(\mathcal{O}) \in \mathbb{Z}$, where $\Delta(\mathcal{O})$ is the discriminant of $\mathcal{O}$.*

Proof. Let $\mathcal{O}$ be an order with a $\mathbb{Z}$-basis $\{\alpha_1, \ldots, \alpha_n\}$. From Definition 3.6, we can write $\Delta(\mathcal{O}) = \det\left(Tr_{L|K}(\alpha_i\alpha_j)\right)$.

Since each $a = \alpha_i\alpha_j \in \mathcal{O} \subseteq \mathcal{O}_K$ and $\forall a \in \mathcal{O}_K$, $Tr_{L|K}(\alpha) \in \mathbb{Z}$. Therefore, each $Tr_{L|K}(\alpha_i\alpha_j) \in \mathbb{Z}$ and $\Delta(\mathcal{O}) = \det\left(Tr_{L|K}(\alpha_i\alpha_j)\right) \in \mathbb{Z}$. □

Lemma 3.12. *Let $\mathfrak{a} \subseteq \mathfrak{a}'$ be two nonzero finitely generated $\mathcal{O}$-modules in K. From Theorem 3.9, we know $\mathfrak{a} \subseteq \mathfrak{a}'$ are both free $\mathbb{Z}$-modules of rank n. Then, the index $(\mathfrak{a}' : \mathfrak{a})$ is finite and satisfies*

$$d(\mathfrak{a}) = (\mathfrak{a}' : \mathfrak{a})^2 \cdot d(\mathfrak{a}').$$

Proof. As defined in Definition 3.6:

$$d(\mathfrak{a}) = d(\alpha_1, \ldots, \alpha_n) = \det\left((\sigma_i\alpha_j)\right)^2,$$

where $\{\alpha_1, \ldots, \alpha_n\}$ is a $\mathbb{Z}$-basis of $\mathfrak{a}$ and $d(\mathfrak{a})$ is independent of the choice of $\mathbb{Z}$-basis.

Let $\{\beta_1, \beta_2, \ldots, \beta_n\}$ be a $\mathbb{Z}$-basis of $\mathfrak{a}'$, then

$$d(\mathfrak{a}) = d(\alpha_1, \ldots, \alpha_n) = \det(T)^2 \cdot d(\beta_1, \ldots, \beta_n),$$

where $T = (b_{ij})$ is the base change matrix from $\{\beta_1, \ldots, \beta_n\}$ to $\{\alpha_1, \ldots, \alpha_n\}$ with $\alpha_i = \sum b_{ij}\beta_j$.

By Definition 3.6 again,

$$\det(T)^2 \cdot d(\beta_1, \ldots, \beta_n) = \det(T)^2 \cdot d(\mathfrak{a}').$$

Therefore, we have

$$d(\mathfrak{a}) = \det(T)^2 \cdot d(\mathfrak{a}').$$

Notice by Lemma 2.3 the index $(\mathfrak{a}' : \mathfrak{a})$ equals the absolute value of $\det(T)$. Thus, we have $d(\mathfrak{a}) = (\mathfrak{a}' : \mathfrak{a})^2 \cdot d(\mathfrak{a}')$ as required. □

Definition 3.13. ***Lattice***

Let V be an n-dimensional $\mathbb{R}$-vector space. A ***lattice*** *Γ in V is a subgroup of the form*

$$\Gamma = \mathbb{Z}v_1 + \ldots + \mathbb{Z}v_m$$

with $\mathbb{R}$-linearly independent vectors $v_1, \ldots, v_m$ of V. The m-tuple $(v_1, \ldots, v_m)$ is called a basis and the set

$$\Phi = \{x_1v_1 + \ldots + x_mv_m \mid x_i \in \mathbb{R}, 0 \leq x_i < 1\}$$

a ***fundamental mesh*** *of the lattice. The lattice is called complete if $m = n$.*

A subset X of V is called centrally symmetric if, given any point $x \in X$, then $-x \in X$; and it is called convex if, given any two points $x, y \in X$, the whole line segment $\{ty + (1 - t)x \mid 0 \leq t \leq 1\}$ joining x with y is contained in X.

Definition 3.14. ***Volume of a lattice***

Let V be an euclidean vector space, i.e., an $\mathbb{R}$-vector space of finite dimension n equipped with a symmetric, positive definite bilinear form

$$\langle\ ,\ \rangle : V \times V \longrightarrow \mathbb{R}.$$

The parallelopiped spanned by n linearly independent vectors $\{v_1, v_2, \ldots, v_n\}$,

$$\Phi = \{x_1v_1 + \cdots + x_nv_n \mid x_i \in \mathbb{R},\ 0 \leq x_i < 1\}$$

has ***volume***

$$\mathrm{vol}(\Phi) = |\det A|$$

where $A = (a_{ij})$ is the matrix if the base change form $\{e_1, \ldots, e_n\}$ to $\{v_1, v_2, \ldots, v_n\}$ and so that $v_i = \sum_k a_{ik}e_k$. Here, $\{e_1, \ldots, e_n\}$ is an orthonormal basis spans a cube with volume 1.

Since

$$(\langle\ v_i,\ v_j\rangle) = \left(\sum_{k,l} a_{ik}a_{jl}\langle e_k, e_l\rangle\right) = \left(\sum_{k} a_{ik}a_{jk}\right) = AA^T,$$

we have the invariant notation

$$\mathrm{vol}(\Phi) = \left|\det\left(\langle v_i, v_j\rangle\right)\right|^{\frac{1}{2}}.$$

Let Γ be a lattice spanned by $\{v_1, v_2, \ldots, v_n\}$. Then Φ is a fundamental mesh of Γ, and we write short

$$\mathrm{vol}(\Gamma) = \mathrm{vol}(\Phi).$$

With the above definitions, we now state the well-known Minkowski's Lattice Point Theorem .

Theorem 3.15. ***Minkowski's Lattice Point Theorem***

Let Γ be a complete lattice in the Euclidean vector space V and X a centrally symmetric and convex body of V. Suppose that

$$\mathrm{vol}(X) > 2^n \mathrm{vol}(\Gamma).$$

Then X contains at least one non-zero lattice point $\gamma \in \Gamma$.

Proof. See Theorem 4.4 in [13]. □

We consider the canonical mapping

$$j : K \longrightarrow K_{\mathbb{C}} := \prod_{\tau} \mathbb{C}, \quad a \longmapsto ja = (\tau a), \tag{3.2}$$

which results from the n complex embeddings $\tau : K \to \mathbb{C}$ (see [13, Chapter 4]). The $\mathbb{C}$-vector space $K_{\mathbb{C}}$ is equipped with the *Hermitian inner product*

$$\langle x, y\rangle = \sum_{\tau} x_\tau \bar{y_\tau}. \tag{3.3}$$

We recall that a Hermitian inner product is given by a form $H(x, y)$ which is linear in the first variable and satisfies $\overline{H(x, y)} = H(y, x)$ as well as $H(x, y) > 0$ for $x \neq 0$ and we can view $K_{\mathbb{C}}$ as a hermitian space with respect to (3.3).

Next, we will introduce the definition of *Minkowski space*. First, notice that every embedding $\tau : K \to \mathbb{C}$ is either real or complex, and if the embeddings $\tau : K \to \mathbb{C}$ are real, their images already landed in $\mathbb{R}$. Second, all the other complex embeddings come as in pairs. Also, they can be thought of as embeddings into $\mathbb{R}^2$ by splitting into real and imaginary parts.

Definition 3.16. ***Minkowski space*** $K_{\mathbb{R}}$

Let

$$\rho_1, \ldots, \rho_r \ : \ K \longrightarrow \mathbb{R}$$

be the real embeddings. The complex embeddings come in pairs

$$\sigma_1, \bar{\sigma_1}, \ldots, \sigma_s, \bar{\sigma_s} \ : \ K \longrightarrow \mathbb{C}$$

We choose from each pair some fixed complex embeddings, and let ρ vary over the family of real embeddings and σ over the family of chosen complex embeddings. So we define the Minkowski space $K_{\mathbb{R}}$ of K as

$$K_{\mathbb{R}} = \{(z_\tau) \in \prod_\tau \mathbb{C} \mid z_\rho \in \mathbb{R}, z_{\bar{\sigma}} = \bar{z}_\sigma\}. \tag{3.4}$$

From Proposition §5.1 in Neukirch[13], we know there is an isomorphism

$$f : K_{\mathbb{R}} \longrightarrow \prod_\tau \mathbb{R} = \mathbb{R}^{r+2s}$$

given by the rule $(z_\tau) \longmapsto (x_\tau)$ where

$$x_\rho = z_\rho, \ x_\sigma = Re(z_\sigma), \ x_{\bar{\sigma}} = Im(z_\sigma).$$

Now consider V again from Definition 3.14. This isomorphism transforms the canonical metric $\langle \ , \ \rangle$ into the inner product

$$\langle x, y \rangle = \sum_\tau \alpha_\tau x_\tau y_\tau$$

where $\alpha_\tau = 1$ (resp. $\alpha_\tau = 2$), if τ is real (resp. τ is complex).

Lemma 3.17. *Modification of Proposition 5.2 in [13]*

Let $\mathfrak{a}$ be a nonzero ideal of $\mathcal{O}$, then $\Gamma = j\mathfrak{a}$ is a complete lattice in $K_{\mathbb{R}}$ and its fundamental mesh has volume

$$\mathrm{vol}(\Gamma) = \sqrt{|d_{\mathcal{O}}|}(\mathcal{O} : \mathfrak{a}),$$

where j is the canonical mapping defined in (3.2).

Proof. Let $\{\alpha_1, \ldots, \alpha_n\}$ be a $\mathbb{Z}$-basis of $\mathfrak{a}$, then

$$\Gamma = \mathbb{Z} \ j\alpha_1 + \ldots + \mathbb{Z} \ j\alpha_n.$$

We choose a numbering of the embeddings $\tau : K \longrightarrow \mathbb{C},\ \tau_1, ..., \tau_n$, and form the matrix $A = (\tau_l \alpha_i)$. Then according to Lemma 3.12, we have

$$d(\mathfrak{a}) = d(\alpha_1, ..., \alpha_n) = (\det A)^2 = (\mathcal{O} : \mathfrak{a})^2 d(\mathcal{O})$$

and on the other hand

$$(\langle j\alpha_i, j\alpha_k \rangle) = \left(\sum_{l=1}^{n} \tau_l \alpha_i \bar{\tau}_l \alpha_k \right) = A\bar{A}^T.$$

By Definition 3.14, this yields

$$\mathrm{vol}(\Gamma) = \left| \det \left(\langle j\alpha_i, j\alpha_k \rangle \right) \right|^{\frac{1}{2}} = |\det(A)| = \sqrt{|d_{\mathcal{O}}|}(\mathcal{O} : \mathfrak{a}). \tag{3.5}$$

□

Theorem 3.18. *Modification of Theorem 5.3 in [13]*

Let $\mathfrak{a}$ be a nonzero integral ideal of K, and let $c_\tau > 0$, for $\tau \in Hom(K, \mathbb{C})$, be real numbers such that $c_\tau = c_{\bar{\tau}}$ and

$$\prod_{\tau} c_\tau > A(\mathcal{O} : \mathfrak{a}),$$

where $A = \left(\frac{2}{\pi}\right)^2 \sqrt{|d_{\mathcal{O}}|}$. Then there exists $a \in \mathfrak{a}, a \neq 0$, such that

$$|\tau a| < c_\tau,\ \forall \tau \in Hom(K, \mathbb{C}).$$

Proof. The set $X = \{(z_\tau) \in K_{\mathbb{R}} \ \Big|\ |z_\tau| < c_\tau\}$ is centrally symmetric and convex. Its volume $\mathrm{vol}(X)$ can be computed via the map

$$f : K_{\mathbb{R}} \xrightarrow{\sim} \prod_{\tau} \mathbb{R}, \quad (z_\tau) \mapsto (x_\tau),$$

given by $x_\rho = z_\rho, x_\sigma = \mathrm{Re}(z_\sigma), x_{\bar{\sigma}} = \mathrm{Im}(z_\sigma)$.

It comes out to be 2^s times the Lebesgue-volume of the image

$$f(X) = \left\{ (x_\tau) \in \prod_{\tau} \mathbb{R} \Big| \quad |x_\rho| < c_\rho, x_\rho^2 + x_{\bar{\rho}}^2 < c_\rho^2 \right\}$$

This gives

$$\mathrm{vol}(X) = 2^s \mathrm{vol}_{Lebesgue}\left(f(X)\right) = 2^s \prod_{\rho} (2c_\rho) \prod_{\sigma} \left(\pi c_\rho^2\right) = 2^{r+s} \pi^s \prod_{\tau} c_\tau.$$

Now using Lemma 3.17, we obtain

$$\mathrm{vol}(X) > 2^{r+s}\pi^s \left(\frac{2}{\pi}\right)^s \sqrt{|d_{\mathcal{O}}|}(\mathcal{O} : \mathfrak{a}) = 2^n \mathrm{vol}(\Gamma).$$

Thus, the hypothesis of Minkowski's lattice point theorem is satisfied. So there exists a lattice point $ja \in X, a \neq 0, a \in \mathfrak{a}$. In other words, $|\tau a| < c_\tau$. □

Theorem 3.19. *Modification of Lemma 6.2 in [13]*

In every nonzero ideal $\mathfrak{a}$ of an arbitrary order $\mathcal{O} \subseteq \mathcal{O}_K$, there exists an $a \in \mathfrak{a}, a \neq 0$ such that

$$|N_{K/\mathbb{Q}}(a)| \leqslant \left(\frac{2}{\pi}\right)^s \sqrt{|d_{\mathcal{O}}|} N(\mathfrak{a}),$$

where $N(\mathfrak{a}) = (\mathcal{O} : \mathfrak{a})$.

Proof. Given $\varepsilon > 0$, we choose positive real numbers c_τ, for $\tau \in Hom(K, \mathbb{C})$, such that $c_\tau = c_{\bar{\tau}}$, and

$$\prod_\tau c_\tau = \left(\frac{2}{\pi}\right)^s \sqrt{|d_{\mathcal{O}}|} N(\mathfrak{a}) + \varepsilon.$$

Then by previous theorem we find an element $a \in \mathfrak{a}, a \neq 0$, satisfying $|\tau a| < c_\tau$. Thus

$$|N_{K/\mathbb{Q}}(a)| = \prod_\tau |a_\tau| < \left(\frac{2}{\pi}\right)^s \sqrt{|d_{\mathcal{O}}|} N(\mathfrak{a}) + \varepsilon.$$

This being true for all $\varepsilon > 0$ and since $|N_{K/\mathbb{Q}}(a)|$ is always a positive integer, there has to exist an $a \in \mathfrak{a}, a \neq 0$, such that

$$|N_{K/\mathbb{Q}}(a)| \leq \left(\frac{2}{\pi}\right)^s \sqrt{|d_{\mathcal{O}}|} N(\mathfrak{a}).$$

□

Definition 3.20. ***Inverse of an $\mathcal{O}$-module***

Let $\mathfrak{a}$ be a nonzero $\mathcal{O}$-module in K. We define the inverse of $\mathfrak{a}$ in a similar manner as when $\mathfrak{a}$ is an $\mathcal{O}_K$-module, that is $\mathfrak{a}^{-1} = \{x \in K \mid x\mathfrak{a} \subseteq \mathcal{O}\}$.

One can check by the module definition that $\mathfrak{a}^{-1}$ is indeed an $\mathcal{O}$-module in K.

Lemma 3.21. *Let $\mathfrak{a}$ be a nonzero finitely generated $\mathcal{O}$-module in K, then $\mathfrak{a}^{-1}$ is a nonzero finitely generated $\mathcal{O}$-module as well.*

Proof. Given $\mathfrak{a}$ is a nonzero finitely generated $\mathcal{O}$-module, we know from Theorem 3.9 that $\mathfrak{a}$ is a free $\mathbb{Z}$-module of rank n, that is

$$\mathfrak{a} = \mathbb{Z}\beta_1 + \ldots + \mathbb{Z}\beta_n \tag{3.6}$$

where $\{\beta_1, \ldots, \beta_n\}$ is a $\mathbb{Z}$-basis of $\mathfrak{a}$.

Take an element β_1 from the $\mathbb{Z}$-basis $\{\beta_1, \ldots, \beta_n\}$ of $\mathfrak{a}$. We know from Definition 3.20 that

$$x\beta_1 \in \mathcal{O}, \quad \forall x \in \mathfrak{a}^{-1}. \tag{3.7}$$

By Definition 3.5, $\mathcal{O}$ is a free $\mathbb{Z}$-module of rank n. So (3.7) implies

$$x\beta_1 \in \mathcal{O} = \mathbb{Z}\alpha_1 + \ldots + \mathbb{Z}\alpha_n, \quad \forall x \in \mathfrak{a}^{-1} \tag{3.8}$$

where $\{\alpha_1, \ldots, \alpha_n\}$ is a $\mathbb{Z}$-basis of $\mathcal{O}$.

Let $\frac{\alpha_i}{\beta_1} = r_i$, $1 \leq i \leq n$. We can write (3.8) as

$$x \in \mathfrak{a}^{-1} \subseteq \mathbb{Z}r_1 + \ldots + \mathbb{Z}r_n := B, \quad \forall x \in \mathfrak{a}^{-1}. \tag{3.9}$$

Then (3.9) implies the $\mathcal{O}$-module $\mathfrak{a}^{-1}$ is a submodule of a finitely generated $\mathbb{Z}$-module B. Therefore, by Lemma 2.1, $\mathfrak{a}^{-1}$ is a finitely generated $\mathbb{Z}$-module as well. □

Definition 3.22. ***Product of two $\mathcal{O}$-modules***

Let $\mathfrak{a}$ and $\mathfrak{b}$ be two nonzero $\mathcal{O}$-modules, where $\mathcal{O}$ is an order from Definition 3.5.

We define the product of $\mathfrak{a}$ and $\mathfrak{b}$ by

$$\mathfrak{ab} = \left\{ \sum_i a_i b_i \,\middle|\, a_i \in \mathfrak{a}, b_i \in \mathfrak{b} \right\}.$$

Definition 3.23. ***Extension of an $\mathcal{O}$-module to an $\mathcal{O}_K$-module***

Let $\mathfrak{a}$ be a nonzero finitely generated $\mathcal{O}$-module in K and suppose

$$\mathfrak{a} = \mathcal{O}\alpha_1 + \ldots + \mathcal{O}\alpha_n, \tag{3.10}$$

where $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is an $\mathcal{O}$-generating set of $\mathfrak{a}$.

We define

$$\mathcal{O}_K\mathfrak{a} := \mathcal{O}_K\alpha_1 + \ldots + \mathcal{O}_K\alpha_n.$$

One can check by module definition that $\mathcal{O}_K\mathfrak{a}$ is an $\mathcal{O}_K$-module that contains $\mathfrak{a}$. Also, the following lemma shows $\mathcal{O}_K\mathfrak{a}$ is independent of the $\mathcal{O}$-generating set of $\mathfrak{a}$, and therefore, it is well-defined.

Lemma 3.24. *$\mathcal{O}_K\mathfrak{a}$ is the smallest $\mathcal{O}_K$-module in K that contains $\mathfrak{a}$, which means*

$$\mathcal{O}_K\mathfrak{a} = \bigcap_{M \in \Omega} M$$

where Ω is the set of all $\mathcal{O}_K$-modules in K which contain $\mathfrak{a}$.

Proof. $\mathcal{O}_K\mathfrak{a} \subseteq \bigcap_{M\in\Omega} M$ follows from the construction of $\mathcal{O}_K\mathfrak{a}$ and the fact that each M is an $\mathcal{O}_K$-module in K containing $\mathfrak{a}$. On the other hand, we have from Definition 3.23 that $\mathcal{O}_K\mathfrak{a}$ is an $\mathcal{O}_K$-module in K containing $\mathfrak{a}$, so $\mathcal{O}_K\mathfrak{a}$ will be one of those M's in Ω, and therefore $\bigcap_{M\in\Omega} M \subseteq \mathcal{O}_K\mathfrak{a}$. □

Lemma 3.25. *Let $\mathcal{O}_K\mathfrak{a}$ be defined as in Definition 3.23. Then $(\mathcal{O}_K\mathfrak{a} : \mathfrak{a}) \mid \delta^n$, where $\delta = (\mathcal{O}_k : \mathcal{O})$.*

Proof. From Lemma 2.5, $\delta\mathcal{O}_K \subseteq \mathcal{O}$. So $\delta\mathcal{O}_K\mathfrak{a} \subseteq \mathcal{O}\mathfrak{a} = \mathfrak{a}$, and by Corollary 2.6, we have $(\mathcal{O}_K\mathfrak{a} : \mathfrak{a})\Big|\delta^n$. □

Definition 3.26. *Let $\mathcal{R}$ be a collection of ordered pairs (M, N) of nonzero finitely generated $\mathcal{O}$-modules in K. We write $M \lesssim_{\mathcal{R}} N$ if there is a positive integer c such that $cM \subseteq N$ for every $(M, N) \in \mathcal{R}$. We write $M \approx_{\mathcal{R}} N$ if and only if $M \lesssim_{\mathcal{R}} N$ and $N \lesssim_{\mathcal{R}} M$.*

Remark 3.27. *Our convention for the definition of $M \lesssim_{\mathcal{R}} N$ is that the constant c can always be in principle made explicit (possibly dependent on $\mathcal{O}$, $\mathcal{R}$).*

In what follows, we will often express the condition $M \lesssim_{\mathcal{R}} N$ in a more informal way. For example, the statement

"Let $\mathfrak{a}$ be any nonzero $\mathcal{O}$-module in K. Then $\mathcal{O} \approx \mathfrak{a} \cdot \mathfrak{a}^{-1}$."

translates more formally to

"$\mathcal{O} \approx_{\mathcal{R}} \mathfrak{a}\cdot\mathfrak{a}^{-1}$ where $\mathcal{R} = \left\{(\mathcal{O}, \mathfrak{a} \cdot \mathfrak{a}^{-1}) : \mathfrak{a} \text{ is a nonzero finitely generated } \mathcal{O}\text{-module in } K\right\}$."

As one can see, the formal statements can be cumbersome; however, the informal statement carries a slight abuse of quantification, since it could be interpreted that $\mathfrak{a}$ is fixed first (which is not the case).

Lemma 3.28. *Let M, N, T be any three nonzero finitely generated $\mathcal{O}$-modules in K. Then*

- $M \approx N, N \approx T \implies M \approx T$,
- $M \approx N \implies MT \approx NT$,
- $M \approx N \implies \mathcal{O} \approx M^{-1}N$.

Proof.

- By Definition 3.26, $M \approx N, N \approx T$ means $M \lesssim N, N \lesssim T$ as well as $N \lesssim M, T \lesssim N$. From $M \lesssim N, N \lesssim T$, we know there exist positive integers c_1 and c_2 for which

$$c_1 M \subseteq N, \quad c_2 N \subseteq T.$$

Then, we have

$$c_1 c_2 M \subseteq c_2 N \subseteq T$$

and it gives $M \lesssim T$.

We use the same argument for $T \lesssim N, N \lesssim M$, and it results $T \lesssim M$. Thus $M \approx T$.

- By Definition 3.26 again, $M \approx N$ gives $c_1 M \subseteq N$ and $c_2 N \subseteq M$, where c_1 and c_2 are positive integers.

  Since T is a nonzero $\mathcal{O}$-module. If $c_1 M \subseteq N$, then $c_1 MT \subseteq NT$. So

  $$MT \lesssim NT. \tag{3.11}$$

  Similarly, $c_2 N \subseteq M$ will give us $c_2 NT \subseteq MT$, which implies

  $$NT \lesssim MT. \tag{3.12}$$

  Therefore,

  $$MT \approx NT. \tag{3.13}$$

- Notice that T in (3.13) can be replaced by M^{-1}, which is proven to be a nonzero finitely generated $\mathcal{O}$-module by Lemma 3.21. So, (3.13) becomes

  $$MM^{-1} \approx NM^{-1}. \tag{3.14}$$

  From Lemma 3.30, we have $\mathcal{O} \approx MM^{-1}$, which implies

  $$\mathcal{O} \approx NM^{-1}. \tag{3.15}$$

□

Lemma 3.29. *Let $\mathfrak{a}$ be any nonzero finitely generated $\mathcal{O}$-module. Then $\mathfrak{a} \approx \mathcal{O}_K\mathfrak{a}$, where we regard $\mathcal{O}_K\mathfrak{a}$ as an $\mathcal{O}$-module.*

Proof. Take $\mathcal{O}_K\mathfrak{a}$ to be the one defined in Definition 3.22 and we know from Lemma 3.24 that $\mathfrak{a} \subseteq \mathcal{O}_K\mathfrak{a}$ as $\mathcal{O}$-modules. So $\mathfrak{a} \lesssim \mathcal{O}_K\mathfrak{a}$ is trivially true.

On the other hand, we know from Lemma 2.5 that $\delta\mathcal{O}_K\mathfrak{a} \subseteq \mathfrak{a}$ as $\mathcal{O}$-modules, where $\delta = (\mathcal{O}_K : \mathcal{O}) \in \mathbb{Z}^+$, and we have $\mathcal{O}_K\mathfrak{a} \lesssim \mathfrak{a}$.

□

Lemma 3.30. *Let $\mathfrak{a}$ be any nonzero finitely generated $\mathcal{O}$-module in K. For the positive integer $\delta = (\mathcal{O}_K : \mathcal{O})$, we have that $\delta^2\mathcal{O} \subseteq \mathfrak{a}\mathfrak{a}^{-1} \subseteq \mathcal{O}$, that is $\mathcal{O} \lesssim \mathfrak{a} \cdot \mathfrak{a}^{-1} \lesssim \mathcal{O}$. Hence, we see that $\mathcal{O} \approx \mathfrak{a} \cdot \mathfrak{a}^{-1}$.*

Proof. From Definition 3.23,

$$\mathfrak{a} \subseteq \mathcal{O}_K\mathfrak{a}, \quad \mathfrak{a}^{-1} \subseteq \mathcal{O}_K\mathfrak{a}^{-1}.$$

As $\mathcal{O}_K\mathfrak{a}^{-1} \cdot \mathcal{O}_K\mathfrak{a} \subseteq \mathcal{O}_K$, we know

$$\mathcal{O}_K\mathfrak{a}^{-1} \subseteq (\mathcal{O}_K\mathfrak{a})^{-1}.$$

Taken together, we have

$$\mathfrak{a} \subseteq \mathcal{O}_K\mathfrak{a} \text{ and } \mathfrak{a}^{-1} \subseteq \mathcal{O}_K\mathfrak{a}^{-1} \subseteq (\mathcal{O}_K\mathfrak{a})^{-1} \tag{3.16}$$

and it can be depicted by the following diagram:

$$\begin{array}{ccc} \mathcal{O}_K\mathfrak{a} & & (\mathcal{O}_K\mathfrak{a})^{-1} \\ & & \cup| \\ & & \mathcal{O}_K\mathfrak{a}^{-1} \\ \cup| & & \\ & & \cup| \\ \mathfrak{a} & & \mathfrak{a}^{-1} \end{array} \tag{3.17}$$

On the left-hand side of (3.17), let $\{\alpha_1, \ldots, \alpha_n\}$ be an $\mathcal{O}$-generating set of $\mathfrak{a}$ so that

$$\mathfrak{a} = \mathcal{O}\alpha_1 + ... + \mathcal{O}\alpha_n \tag{3.18}$$

and

$$\mathcal{O}_K\mathfrak{a} = \mathcal{O}_K\alpha_1 + ... + \mathcal{O}_K\alpha_n. \tag{3.19}$$

From Lemma 2.5, we know there exists a nonzero integer $\delta = (\mathcal{O}_K : \mathcal{O})$ such that $\delta\mathcal{O}_K \subseteq \mathcal{O}$. So we multiply by δ on both sides of (3.19) and it gives

$$\delta\mathcal{O}_K\mathfrak{a} = \delta\mathcal{O}_K\alpha_1 + \ldots + \delta\mathcal{O}_K\alpha_n \subseteq \mathcal{O}\alpha_1 + ... + \mathcal{O}\alpha_n = \mathfrak{a}. \tag{3.20}$$

Therefore, by Definition 3.26, $\mathcal{O}_K\mathfrak{a} \lesssim \mathfrak{a}$.

On the right-hand side of (3.17), for each generator $\alpha_i \in \{\alpha_1, \ldots, \alpha_n\} \subseteq \mathcal{O}_K\mathfrak{a}$ in (3.19), we have

$$x \cdot \alpha_i \in \mathcal{O}_K, \quad \forall i, 1 \leq i \leq n \tag{3.21}$$

where x is any element in $(\mathcal{O}_K\mathfrak{a})^{-1}$.

Multiply again by δ on both sides of (3.21), we have

$$\delta \cdot x \cdot \alpha_i \in \delta \cdot \mathcal{O}_K \subseteq \mathcal{O}, \quad \forall i, 1 \leq i \leq n$$

and it implies

$$(\delta \cdot x) \in \mathfrak{a}^{-1}, \quad \forall x \in (\mathcal{O}_K\mathfrak{a})^{-1}. \tag{3.22}$$

Therefore,

$$(\mathcal{O}_K\mathfrak{a})^{-1} \lesssim \mathfrak{a}^{-1}. \tag{3.23}$$

With the above two "approximate" inequalities in hand, our next task is to show the inequality holds for the product of $\mathcal{O}_K\mathfrak{a}$ and $(\mathcal{O}_K\mathfrak{a})^{-1}$, that is, $\mathcal{O}_K\mathfrak{a}\cdot(\mathcal{O}_K\mathfrak{a})^{-1} \lesssim \mathfrak{a}\cdot\mathfrak{a}^{-1}$.

Notice that $(\mathcal{O}_K\mathfrak{a})^{-1}$ is a finitely generated $\mathcal{O}_K$-module with the property:

$$\mathcal{O}_K\mathfrak{a}\cdot(\mathcal{O}_K\mathfrak{a})^{-1} = \mathcal{O}_K. \tag{3.24}$$

Multiplying by δ^2 on both sides of (3.24), we have

$$\delta^2\cdot\mathcal{O}_K\mathfrak{a}\cdot(\mathcal{O}_K\mathfrak{a})^{-1} = \delta^2\cdot\mathcal{O}_K,$$

which can be written in concrete terms:

$$\delta^2\cdot\mathcal{O}_K = \delta^2(\mathcal{O}_K\mathfrak{a})^{-1}(\mathcal{O}_K\mathfrak{a}) = \delta^2\sum_{i,j=1}^{n}\mathcal{O}_K\beta_j\alpha_i = \sum_{i,j=1}^{n}(\delta\mathcal{O}_K\alpha_i)\cdot(\delta\beta_j), \quad \forall i,j, 1\le i,j\le n, \tag{3.25}$$

where $\{\beta_1,\dots,\beta_n\}$ is an $\mathcal{O}_K$-generating set of $(\mathcal{O}_K\mathfrak{a})^{-1}$ and $\{\alpha_1,\dots,\alpha_n\}$ is from (3.18).

From (3.20) and (3.22), the right-hand side of (3.25) satisfies

$$\sum_{i,j=1}^{n}(\delta\mathcal{O}_K\alpha_i)\cdot(\delta\beta_j) \subseteq \mathfrak{a}\cdot\mathfrak{a}^{-1} \subseteq \mathcal{O}, \quad \forall i,j,\ 1\le i,j\le n. \tag{3.26}$$

Therefore, all together we have $\delta^2\cdot\mathcal{O} \subseteq \delta^2\cdot\mathcal{O}_K \subseteq \mathfrak{a}\cdot\mathfrak{a}^{-1} \subseteq \mathcal{O}$, and that shows the lemma. □

Corollary 3.31. *Let $\mathfrak{a}$ any nonzero finitely generated $\mathcal{O}$-module in K. Then $\mathfrak{a} \approx (\mathfrak{a}^{-1})^{-1}$.*

Proof. We have that

$$\begin{aligned}\mathfrak{a} = \mathfrak{a}\cdot\mathcal{O} &\approx \mathfrak{a}\cdot(\mathfrak{a}^{-1}\cdot(\mathfrak{a}^{-1})^{-1})\\ &\approx (\mathfrak{a}\cdot\mathfrak{a}^{-1})\cdot(\mathfrak{a}^{-1})^{-1}\\ &\approx \mathcal{O}\cdot(\mathfrak{a}^{-1})^{-1}.\end{aligned}$$

□

Lemma 3.32. *Let $\mathfrak{a}$ be any nonzero finitely generated $\mathcal{O}$-module in K. Then $\mathcal{O}_K\mathfrak{a}^{-1} \approx (\mathcal{O}_K\mathfrak{a})^{-1}$.*

Proof. Notice from Definition 3.26, (3.16) gives

$$\mathfrak{a}^{-1} \lesssim \mathcal{O}_K\mathfrak{a}^{-1} \lesssim (\mathcal{O}_K\mathfrak{a})^{-1} \tag{3.27}$$

and it directly combines with $(\mathcal{O}_K\mathfrak{a})^{-1} \lesssim \mathfrak{a}^{-1}$ (3.23) to give

$$(\mathcal{O}_K\mathfrak{a})^{-1} \approx \mathcal{O}_K\mathfrak{a}^{-1}.$$

□

Definition 3.33. *Let $\mathcal{O}$ be an order in K. A nonzero $\mathcal{O}$-module M is called a based $\mathcal{O}$-module if M is equipped with an ordered $\mathbb{Z}$-basis of rank n over $\mathbb{Z}$. Furthermore, if M is an ideal of $\mathcal{O}$, M is called a based ideal of $\mathcal{O}$.*

Let $\mathcal{O}$ be an order from Definition 3.5 and $\mathfrak{a}$ be a nonzero finitely generated $\mathcal{O}$-module in K. By Theorem 3.9, $\mathfrak{a}$ is a free $\mathbb{Z}$-module of rank n, thus a based $\mathcal{O}$-module.

Definition 3.34. ***Norm of a based $\mathcal{O}$-module***

Let $\mathfrak{a}$ be a based $\mathcal{O}$-module equipped with a $\mathbb{Z}$-basis $\{\beta_1, \ldots, \beta_n\}$. We define the ***Norm*** *$N_{\mathcal{O}}(\mathfrak{a})$ of $\mathfrak{a}$ as the determinant of the K-linear transformation mapping α_i to β_i for $i = 1, \ldots, n$, where $\{\alpha_1, \ldots, \alpha_n\}$ is a $\mathbb{Z}$-basis of the order $\mathcal{O}$.*

Remark 3.35. *Let M be the matrix corresponding to the K-linear transformation defined above. If different $\mathbb{Z}$-bases of $\mathfrak{a}$ and $\mathcal{O}$ are used, M will become PMQ, where $P, Q \in GL_n(\mathbb{Z})$. So $N_{\mathcal{O}}(\mathfrak{a})$ depends on $\mathbb{Z}$-bases of $\mathfrak{a}$ and $\mathcal{O}$, but $|N_{\mathcal{O}}(\mathfrak{a})|$ will be well-defined.*

Remark 3.36. *If $\mathfrak{a}$ is a nonzero ideal of $\mathcal{O}$, then the K-linear transformation in Definition 3.34 is indeed a $\mathbb{Z}$-linear transformation and $|N_{\mathcal{O}}(\mathfrak{a})| \in \mathbb{Z}$. Moreover, we know from Lemma 2.3 that $|N_{\mathcal{O}}(\mathfrak{a})| = (\mathcal{O} : \mathfrak{a})$.*

Remark 3.37. *$N_{\mathcal{O}_K}(\mathcal{O}_K\mathfrak{a})$ is defined in a similar manner as $N_{\mathcal{O}}(\mathfrak{a})$ since $\mathcal{O}_K$ is the maximal order of K and $\mathcal{O}_K\mathfrak{a}$ is an $\mathcal{O}_K$-module.*

Lemma 3.38. *Let $\mathfrak{a}$ be a principal ideal of order $\mathcal{O}$, which is generated by $a \in K^*$. Then, $|N_{\mathcal{O}}(\mathfrak{a})| = |N_{K/\mathbb{Q}}(a)|$, where $N_{K/\mathbb{Q}}(a)$ is from Definition 3.1.*

Proof. From Definition 3.5, $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is a $\mathbb{Z}$-basis of the order $\mathcal{O}$. Then $\{a\alpha_1, a\alpha_2, \ldots, a\alpha_n\}$ is a $\mathbb{Z}$-basis of $\mathfrak{a} = a \cdot \mathcal{O}$, and if $M = (m_{ij})$ denotes the matrix, where $a\alpha_i = \sum m_{ij}\alpha_j$. Then by Definition 3.1, $\det(M) = N_{K/\mathbb{Q}}(a)$. So we have $|N_{\mathcal{O}}(\mathfrak{a})| = |N_{K/\mathbb{Q}}(a)|$. □

Definition 3.39. *Let $\mathcal{R}$ be a collection of ordered pairs of elements in $\mathbb{R}^+$. We say $x \lesssim_{\mathcal{R}} y$ if and only if there is a constant $c \in \mathbb{R}^+$ such that $x \leq cy$ for every $(x, y) \in \mathcal{R}$. We say $x \approx_{\mathcal{R}} y$ if and only if $x \lesssim_{\mathcal{R}} y$ and $y \lesssim_{\mathcal{R}} x$.*

Remark 3.40. *Our convention for the definition of $x \lesssim_{\mathcal{R}} y$ is that the constant c can always be in principle made explicit (possibly dependent on $\mathcal{O}$, $\mathcal{R}$).*

As with Definition 3.26, we will express the condition $x \lesssim_{\mathcal{R}} y$ in a more informal way, with a slight abuse of quantification. Thus, when we make a statement such as

" Let $\mathfrak{a}, \mathfrak{b}$ be any nonzero finitely generated $\mathcal{O}$-modules in K. Then $N_{\mathcal{O}}(\mathfrak{a}\cdot\mathfrak{b}) \approx N_{\mathcal{O}}(\mathfrak{a})N_{\mathcal{O}}(\mathfrak{b})$.",
it should be read as

"$N_{\mathcal{O}}(\mathfrak{a}\cdot\mathfrak{b}) \approx_{\mathcal{R}} N_{\mathcal{O}}(\mathfrak{a})N_{\mathcal{O}}(\mathfrak{b})$ where
$\mathcal{R} = \{(N_{\mathcal{O}}(\mathfrak{a}\cdot\mathfrak{b}), N_{\mathcal{O}}(\mathfrak{a})N_{\mathcal{O}}(\mathfrak{b})) : \mathfrak{a}, \mathfrak{b}$ are nonzero finitely generated $\mathcal{O}$-modules in $K\}$."

Lemma 3.41. *Let x_1, x_2, y_1, y_2 be any four numbers $\in \mathbb{R}^+$ such that $x_1 \approx x_2$, $y_1 \approx y_2$. Then $x_1y_1 \approx x_2y_2$.*

Proof. Given that $x_1 \approx x_2$ and $y_1 \approx y_2$, from Definition 3.39 we get

$$x_1 \leq c_1x_2, \quad y_1 \leq c_2y_2, \text{ where } c_1, c_2 \in \mathbb{R}^+.$$

Therefore, we can have

$$x_1y_1 \leq c_1c_2 \cdot x_2y_2,$$

which gives

$$x_1y_1 \lesssim x_2y_2.$$

Similarly, we have $x_2y_2 \lesssim x_1y_1$ and the lemma is proven. □

Lemma 3.42. *Let $\mathfrak{a}, \mathfrak{b}$ be two nonzero based $\mathcal{O}$-modules in K, where $\mathfrak{b} \subseteq \mathfrak{a}$. Then the index*

$$(\mathfrak{a} : \mathfrak{b}) = \frac{|N_{\mathcal{O}}(\mathfrak{b})|}{|N_{\mathcal{O}}(\mathfrak{a})|} = c \in \mathbb{Z}^+.$$

Proof. Let $\{\beta_1, \ldots, \beta_n\}$ and $\{\gamma_1, \ldots, \gamma_n\}$ be $\mathbb{Z}$-bases of $\mathfrak{a}$ and $\mathfrak{b}$ respectively.

Then we have

$$(\mathfrak{a} : \mathfrak{b}) = |\det(A)| = c \in \mathbb{Z}^+, \quad \text{by Lemma 2.3} \tag{3.28}$$

where A is the matrix such that

$$\begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix} = A \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \tag{3.29}$$

Now, consider $\{\alpha_1, \ldots, \alpha_n\}$ as a $\mathbb{Z}$-basis of the order $\mathcal{O}$. We can write down matrices B and C as follows:

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = B \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \tag{3.30}$$

$$\begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix} = C \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}. \tag{3.31}$$

From (3.29) and (3.30), we have

$$\begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix} = AB \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}. \tag{3.32}$$

Comparing (3.32) with (3.31), we obtain

$$|\det(A)| \cdot |\det(B)| = |\det(C)|, \tag{3.33}$$

and by Definition 3.34, (3.33) can be rewritten as

$$(\mathfrak{a} : \mathfrak{b}) \cdot |N_{\mathcal{O}}(\mathfrak{a})| = |N_{\mathcal{O}}(\mathfrak{b})|, \tag{3.34}$$

which gives the desired result. □

Corollary 3.43. *Let $\mathfrak{a}, \mathfrak{b}$ be two nonzero based $\mathcal{O}$-modules, where $\mathfrak{b} \subseteq \mathfrak{a} \subseteq \mathcal{O}$ and $|N_{\mathcal{O}}(\mathfrak{b})| = |N_{\mathcal{O}}(\mathfrak{a})|$. Then $\mathfrak{a} = \mathfrak{b}$.*

Proof. By Lemma 3.42,

$$(\mathfrak{a} : \mathfrak{b}) \cdot |N_{\mathcal{O}}(\mathfrak{a})| = \left|N_{\mathcal{O}}(\mathfrak{b})\right|.$$

Then we have $(\mathfrak{a} : \mathfrak{b}) = 1$ since $|N_{\mathcal{O}}(\mathfrak{a})| = |N_{\mathcal{O}}(\mathfrak{b})|$. Furthermore, $\mathfrak{b} \subseteq \mathfrak{a}$ implies $\mathfrak{a} = \mathfrak{b}$. □

There is an "approximate" version of the above corollary:

Lemma 3.44. *Let $\mathfrak{a}, \mathfrak{b}$ be any two nonzero based $\mathcal{O}$-modules where $\mathfrak{b} \subseteq \mathfrak{a}$ and $|N_{\mathcal{O}}(\mathfrak{b})| \approx |N_{\mathcal{O}}(\mathfrak{a})|$. Then $\mathfrak{a} \approx \mathfrak{b}$.*

Proof. Given $\mathfrak{b} \subseteq \mathfrak{a}$, we have

$$(\mathfrak{a} : \mathfrak{b}) = \frac{|N_{\mathcal{O}}(\mathfrak{b})|}{|N_{\mathcal{O}}(\mathfrak{a})|} = c \in \mathbb{Z}^+,$$

from Lemma 3.42. Furthermore, $|N_{\mathcal{O}}(\mathfrak{b})| \approx |N_{\mathcal{O}}(\mathfrak{a})|$ implies an existence of $c_1 \in \mathbb{R}^+$ such that

$$1 \leq \frac{|N_{\mathcal{O}}(\mathfrak{b})|}{|N_{\mathcal{O}}(\mathfrak{a})|} \leq c_1.$$

Let c_2 be the lcm of all integers from 1 to $\lfloor c_1 \rfloor$. Then Lemma 2.5 implies

$$c_2 \cdot \mathfrak{a} \subseteq \mathfrak{b}.$$

As $\mathfrak{b} \subseteq \mathfrak{a}$ is given, we have $\mathfrak{a} \approx \mathfrak{b}$. □

Lemma 3.45. *Suppose $\mathfrak{a}$ and $\mathfrak{b}$ are any two nonzero based $\mathcal{O}$-modules in K such that $\mathfrak{a} \approx \mathfrak{b}$. Then $|N_{\mathcal{O}}(\mathfrak{a})| \approx |N_{\mathcal{O}}(\mathfrak{b})|$.*

Proof. From $\mathfrak{a} \approx \mathfrak{b}$, we know

$$c_1\mathfrak{a} \subseteq \mathfrak{b}, \tag{3.35}$$

$$c_2\mathfrak{b} \subseteq \mathfrak{a}, \tag{3.36}$$

where c_1 and $c_2 \in \mathbb{R}^+$. By Lemma 3.42, (3.35) gives

$$(\mathfrak{b} : c_1\mathfrak{a}) = \frac{|N_{\mathcal{O}}(c_1\mathfrak{a})|}{|N_{\mathcal{O}}(\mathfrak{b})|} = c \in \mathbb{Z}^+. \tag{3.37}$$

Notice $c_1\mathfrak{a}$ is a free $\mathbb{Z}$-module whose $\mathbb{Z}$-basis can be obtained by multiplying c_1 to each element in a $\mathbb{Z}$-basis of $\mathfrak{a}$. Then from Definition 3.34, we obtain

$$|N_{\mathcal{O}}(c_1\mathfrak{a})| = c_1^n \cdot |N_{\mathcal{O}}(\mathfrak{a})|,$$

and therefore (3.37) becomes

$$(\mathfrak{b} : c_1\mathfrak{a}) = \frac{c_1^n \cdot |N_{\mathcal{O}}(\mathfrak{a})|}{|N_{\mathcal{O}}(\mathfrak{b})|} = c \in \mathbb{Z}^+. \tag{3.38}$$

Notice from (3.35) and (3.36) we have

$$c_1c_2\mathfrak{b} \subseteq c_1\mathfrak{a} \subseteq \mathfrak{b}. \tag{3.39}$$

Thus, the index $c = (\mathfrak{b} : c_1\mathfrak{a})$ is less than or equal to the index of $(\mathfrak{b} : c_1c_2\mathfrak{b}) = (c_1c_2)^n$ and we have

$$\frac{c_1^n \cdot |N_{\mathcal{O}}(\mathfrak{a})|}{|N_{\mathcal{O}}(\mathfrak{b})|} = c \leq (c_1c_2)^n,$$

which implies

$$\frac{1}{c_2^n} \cdot |N_{\mathcal{O}}(\mathfrak{a})| \leq |N_{\mathcal{O}}(\mathfrak{b})|. \tag{3.40}$$

Therefore,

$$|N_{\mathcal{O}}(\mathfrak{a})| \lesssim |N_{\mathcal{O}}(\mathfrak{b})|.$$

Instead of (3.39), we can reverse the roles of $\mathfrak{a}$ and $\mathfrak{b}$ to get another inclusion

$$c_1 c_2 \mathfrak{a} \subseteq c_2 \mathfrak{b} \subseteq \mathfrak{a}. \tag{3.41}$$

Following a similar argument, (3.41) gives $|N_{\mathcal{O}}(\mathfrak{b})| \lesssim |N_{\mathcal{O}}(\mathfrak{a})|$. Therefore, we have $|N_{\mathcal{O}}(\mathfrak{b})| \approx |N_{\mathcal{O}}(\mathfrak{a})|$ from $\mathfrak{a} \approx \mathfrak{b}$. □

Lemma 3.46. *Let $\mathfrak{a}$ be any nonzero based $\mathcal{O}$-module, and $\mathcal{O}_K\mathfrak{a}$ be the corresponding $\mathcal{O}_K$-module defined in Definition 3.23. Then $|N_{\mathcal{O}_K}(\mathcal{O}_K\mathfrak{a})| \approx |N_{\mathcal{O}}(\mathfrak{a})|$.*

Proof. Given $\mathfrak{a}$ is a nonzero $\mathcal{O}$-module, $\mathfrak{a}$ is a free $\mathbb{Z}$-module equipped with a $\mathbb{Z}$-basis: $\{\beta_1, \ldots, \beta_n\}$. Let A denote the transition matrix from $\{\beta_1, \ldots, \beta_n\}$ to $\{\alpha_1, \ldots, \alpha_n\}$, where $\{\alpha_1, \ldots, \alpha_n\}$ is a $\mathbb{Z}$-basis of $\mathcal{O}$. Next, we consider B, the transition matrix from a $\mathbb{Z}$-basis of $\mathcal{O}$ to a $\mathbb{Z}$-basis of $\mathcal{O}_K$ to, and by Lemma 3.24, $\mathcal{O}_K\mathfrak{a}$ is an $\mathcal{O}_K$-module with a $\mathbb{Z}$-basis as well, so we let C be the transition matrix from a $\mathbb{Z}$-basis of $\mathcal{O}_K\mathfrak{a}$ to a $\mathbb{Z}$-basis of $\mathcal{O}_K$. Lastly, consider D, which is the transition matrix from the $\mathbb{Z}$-basis of $\mathfrak{a}$ to the $\mathbb{Z}$-basis of $\mathcal{O}_K\mathfrak{a}$.

The above can be depicted by the following diagram:

$$\begin{array}{ccc} \mathcal{O}_K & \xleftarrow{\quad C \quad} & \mathcal{O}_K\mathfrak{a} \\ \mathrm{B}\uparrow & & \uparrow\mathrm{D} \\ \mathcal{O} & \xleftarrow{\quad A \quad} & \mathfrak{a} \end{array}$$

As $\{\alpha_1, \ldots, \alpha_n\}$ and $\{\beta_1, \ldots, \beta_n\}$ are fixed,

$$A = B^{-1} \cdot C \cdot D,$$

which gives

$$\det(A) = \frac{1}{\det(B)} \cdot \det(C) \cdot \det(D) \tag{3.42}$$

From Definition 3.34, we have

$$|N_{\mathcal{O}_K}(\mathcal{O}_K\mathfrak{a})| = |\det(C)|, \quad |N_{\mathcal{O}}(\mathfrak{a})| = |\det(A)|.$$

Furthermore, we know by Lemma 2.5,

$$(\mathcal{O}_K : \mathcal{O}) = \delta = |\det(B)|,$$

so that (3.42) becomes

$$|N_{\mathcal{O}}(\mathfrak{a})| = \frac{1}{\delta} \cdot |N_{\mathcal{O}_K}(\mathcal{O}_K\mathfrak{a})| \cdot |\det(D)|,$$

which is

$$\delta \cdot |N_{\mathcal{O}}(\mathfrak{a})| = |\det(D)| \cdot |N_{\mathcal{O}_K}(\mathcal{O}_K\mathfrak{a})|. \tag{3.43}$$

From Lemma 3.25 and Lemma 2.3, $|\det(D)| = (\mathcal{O}_K\mathfrak{a} : \mathfrak{a})\Big|\delta^n$, and $|\det(D)| \in \mathbb{N}$. Hence,

$$1 \leq |\det(D)| \leq \delta^n$$

and from (3.43) we have

$$1 \cdot |N_{\mathcal{O}_K}(\mathcal{O}_K\mathfrak{a})| \leq \delta \cdot |N_{\mathcal{O}}(\mathfrak{a})| = |\det(D)| \cdot |N_{\mathcal{O}_K}(\mathcal{O}_K\mathfrak{a})| \leq \delta^n \cdot |N_{\mathcal{O}_K}(\mathcal{O}_K\mathfrak{a})|,$$

which implies

$$\frac{1}{\delta} \cdot |N_{\mathcal{O}_K}(\mathcal{O}_K\mathfrak{a})| \leq |N_{\mathcal{O}}(\mathfrak{a})| \tag{3.44}$$

as well as

$$|N_{\mathcal{O}}(\mathfrak{a})| \leq \delta^{n-1} \cdot |N_{\mathcal{O}_K}(\mathcal{O}_K\mathfrak{a})|. \tag{3.45}$$

Therefore, from (3.44) and (3.45)

$$|N_{\mathcal{O}_K}(\mathcal{O}_K\mathfrak{a})| \approx |N_{\mathcal{O}}(\mathfrak{a})|.$$

□

For a nonzero $\mathcal{O}_K$-module $\mathfrak{a}$ in $\mathcal{O}_K$, $|N_{\mathcal{O}_K}(\mathfrak{a})| = (\mathcal{O}_K : \mathfrak{a})$, which coincides with the usual definition of the norm $N(\mathfrak{a})$ of $\mathfrak{a}$ from algebraic number theory. Recall this satisfies $N(\mathfrak{ab}) = N(\mathfrak{a})N(\mathfrak{b})$ for any nonzero $\mathcal{O}_K$-modules $\mathfrak{a}, \mathfrak{b}$ in $\mathcal{O}_K$ using the fact that $\mathcal{O}_K$ is a Dedekind domain. This property can be extended to $|N_{\mathcal{O}_K}(\cdot)|$ for all nonzero $\mathcal{O}_K$-modules.

Lemma 3.47. *Let $\mathfrak{a}, \mathfrak{b}$ be nonzero $\mathcal{O}_K$-modules in K. Then*

$$|N_{\mathcal{O}_K}(\mathfrak{ab})| = |N_{\mathcal{O}_K}(\mathfrak{a})||N_{\mathcal{O}_K}(\mathfrak{b})|.$$

Proof. Let $c \in \mathbb{N}$ be such that $c\mathfrak{a}$ and $c\mathfrak{b}$ are contained in $\mathcal{O}_K$. Then

$$|N_{\mathcal{O}_K}(c\mathfrak{a} \cdot c\mathfrak{b})| = N(c\mathfrak{a} \cdot c\mathfrak{b}) = N(c\mathfrak{a})N(c\mathfrak{b}) = |N_{\mathcal{O}_K}(c\mathfrak{a})||N_{\mathcal{O}_K}(c\mathfrak{b})|.$$

By Proposition 5.1,

$$\begin{aligned}
|N_{\mathcal{O}_K}(c\mathfrak{a} \cdot c\mathfrak{b})| &= |N_{K/\mathbb{Q}}(c^2)||N_{\mathcal{O}_K}(\mathfrak{ab})| \\
|N_{\mathcal{O}_K}(c\mathfrak{a})| &= |N_{K/\mathbb{Q}}(c)||N_{\mathcal{O}_K}(\mathfrak{ab})| \\
|N_{\mathcal{O}_K}(c\mathfrak{b})| &= |N_{K/\mathbb{Q}}(c)||N_{\mathcal{O}_K}(\mathfrak{b})|.
\end{aligned}$$

Hence, we obtain the result upon cancelling $N_{K/\mathbb{Q}}(c)^2 \neq 0$ from both sides. □

Lemma 3.48. *Let $\mathfrak{a}$ and $\mathfrak{b}$ be any nonzero based $\mathcal{O}$-modules in K. Then $|N_{\mathcal{O}}(\mathfrak{ab})| \approx |N_{\mathcal{O}}(\mathfrak{a})| \cdot |N_{\mathcal{O}}(\mathfrak{b})|$.*

Proof. We know from Definition 3.22 that $\mathfrak{ab}$, the product of two $\mathcal{O}$-modules, is also an $\mathcal{O}$-module in K. Then from Lemma 3.25, we have that $\mathcal{O}_K\mathfrak{ab}$ is a nonzero $\mathcal{O}_K$-module, and Lemma 3.46 gives

$$|N_{\mathcal{O}}(\mathfrak{ab})| \approx |N_{\mathcal{O}_K}(\mathcal{O}_K\mathfrak{ab})|. \tag{3.46}$$

Since $\mathcal{O}_K$ is a Dedekind domain, $|N_{\mathcal{O}_K}(\mathcal{O}_K\mathfrak{ab})| = |N_{\mathcal{O}_K}(\mathcal{O}_K\mathfrak{a})| \cdot |N_{\mathcal{O}_K}(\mathcal{O}_K\mathfrak{b})|$, by Lemma 3.41 (3.46) becomes

$$|N_{\mathcal{O}}(\mathfrak{ab})| \approx |N_{\mathcal{O}_K}(\mathcal{O}_K\mathfrak{a} \cdot \mathcal{O}_K\mathfrak{b})| = |N_{\mathcal{O}_K}(\mathcal{O}_K\mathfrak{a})| \cdot |N_{\mathcal{O}_K}(\mathcal{O}_K\mathfrak{b})| \approx |N_{\mathcal{O}}(\mathfrak{a})| \cdot |N_{\mathcal{O}}(\mathfrak{b})|, \tag{3.47}$$

which gives the desired result. □

Theorem 3.49. *Let K be a number field and $\mathcal{O}$ be an order in K. If $\mathfrak{a}$ is any nonzero finitely generated $\mathcal{O}$-module in K, then there exists an ideal $\mathfrak{a}_1$ of $\mathcal{O}$ such that*

$$|N_{\mathcal{O}}(\mathfrak{a}_1)| \lesssim \left(\frac{2}{\pi}\right)^s \sqrt{|d_{\mathcal{O}}|},$$

where $\mathfrak{a}_1 = \alpha\mathfrak{a}$ for some $\alpha \in K^*$.

Proof. Given $\mathfrak{a}$ is a nonzero finitely generated $\mathcal{O}$-module in K, we consider $\mathfrak{a}^{-1}$ first. Lemma 3.21 shows that $\mathfrak{a}^{-1}$ is also a nonzero finitely generated $\mathcal{O}$-module, thus a free $\mathbb{Z}$-module of rank n, and we write:

$$\mathfrak{a}^{-1} = \mathbb{Z}\beta_1 + \ldots + \mathbb{Z}\beta_n, \tag{3.48}$$

where $\{\beta_1, \ldots, \beta_n\}$ is a $\mathbb{Z}$-basis of $\mathfrak{a}^{-1}$.

From Lemma 3.8, we know each β_i in the above $\mathbb{Z}$-basis has the property such that $\beta_i = \frac{\mathcal{O}}{\mathbb{Z}}$. Therefore, we can choose a nonzero $r \in \mathbb{Z}$ to clean all the denominators from β_1 to β_n. Then we have

$$r\beta_i \in \mathcal{O},\ 1 \leq i \leq n, i \in \mathbb{Z} \Longleftrightarrow r\mathfrak{a}^{-1} \subseteq \mathcal{O} \tag{3.49}$$

Let $\mathfrak{b} := r\mathfrak{a}^{-1}$ and we can tell $\mathfrak{b}$ is a nonzero finitely generated $\mathcal{O}$-module, which is contained in $\mathcal{O}$. Thus $\mathfrak{b}$ is an ideal of $\mathcal{O}$ and from Theorem 3.19, there exists an $a \in \mathfrak{b}, a \neq 0$ such that

$$|N_{K/\mathbb{Q}}(a)| \leqslant \left(\frac{2}{\pi}\right)^s \sqrt{|d_{\mathcal{O}}|} \cdot |N(\mathfrak{b})|, \tag{3.50}$$

where $|N(\mathfrak{b})| = (\mathcal{O} : \mathfrak{b})$.

Notice that (3.50) can be rewritten as

$$|N_{\mathcal{O}_K}(\mathcal{O}_K(a))| \leqslant \left(\frac{2}{\pi}\right)^s \sqrt{|d_{\mathcal{O}}|} \cdot |N_{\mathcal{O}}(\mathfrak{b})|, \tag{3.51}$$

since by using Lemma 3.38 and Lemma 2.3, we have that

$$|N_{K/\mathbb{Q}}(a)| = |N_{\mathcal{O}_K}(\mathcal{O}_K(a))|, \quad |N(\mathfrak{b})| = |N_{\mathcal{O}}(\mathfrak{b})|.$$

Then (3.51) implies

$$\frac{|N_{\mathcal{O}_K}(\mathcal{O}_K(a))|}{|N_{\mathcal{O}}(\mathfrak{b})|} \leqslant \left(\frac{2}{\pi}\right)^s \sqrt{|d_{\mathcal{O}}|}. \tag{3.52}$$

Apply Lemma 3.46 to $|N_{\mathcal{O}}\mathfrak{b}|$ on the left-hand side of (3.52), we get

$$\begin{aligned}
\frac{|N_{\mathcal{O}_K}(\mathcal{O}_K(a))|}{|N_{\mathcal{O}}(\mathfrak{b})|} &\approx \frac{|N_{\mathcal{O}_K}(\mathcal{O}_K(a))|}{|N_{\mathcal{O}_K}(\mathcal{O}_K\mathfrak{b})|} \\
&= |N_{\mathcal{O}_K}(\mathcal{O}_K(a)\cdot(\mathcal{O}_K\mathfrak{b})^{-1})| \quad \text{by Lemma 3.47} \\
&\approx |N_{\mathcal{O}_K}(\mathcal{O}_K a\mathfrak{b}^{-1})| \quad \text{by Lemmas 3.32, 3.28, and 3.29} \\
&\approx |N_{\mathcal{O}}(a\mathfrak{b}^{-1})| \quad \text{by Lemma 3.46 again.}
\end{aligned}$$

Therefore,

$$|N_{\mathcal{O}}(a\mathfrak{b}^{-1})| \lesssim \left(\frac{2}{\pi}\right)^s \sqrt{|d_{\mathcal{O}}|}. \tag{3.53}$$

Recall that $\mathfrak{b} := r\mathfrak{a}^{-1}$ is an ideal of $\mathcal{O}$. Then $\mathfrak{b}^{-1}$ is a finitely generated $\mathcal{O}$-module and so is $a\mathfrak{b}^{-1}$. Moreover, $a \in \mathfrak{b}$, so $a\mathfrak{b}^{-1} \subseteq \mathcal{O}$ and $a\mathfrak{b}^{-1}$ is indeed an ideal of $\mathcal{O}$.

Now $\{r\beta_1, \ldots, r\beta_n\}$ is an $\mathbb{Z}$-basis of $\mathfrak{b}$ and consider any element $x \in \mathfrak{b}^{-1}$ so that,

$$xr\beta_i \in \mathcal{O}, \quad \forall i \in \mathbb{Z}, 1 \leq i \leq n. \tag{3.54}$$

Recall from (3.48) that $\{\beta_1, \ldots, \beta_n\}$ is a $\mathbb{Z}$-basis of $\mathfrak{a}^{-1}$, and we deduce from (3.54) that $xr \in (\mathfrak{a}^{-1})^{-1}$, so $x \in \frac{(\mathfrak{a}^{-1})^{-1}}{r}$. Therefore,

$$\mathfrak{b}^{-1} \subseteq \frac{\mathfrak{a}}{r} \Rightarrow a\mathfrak{b}^{-1} \subseteq \frac{a}{r}(\mathfrak{a}^{-1})^{-1}, \tag{3.55}$$

that is, $a\mathfrak{b}^{-1} \subseteq \alpha(\mathfrak{a}^{-1})^{-1}$, where $\alpha = \frac{a}{r} \in K^*$.

On the other hand, we know that $\frac{\mathfrak{a}}{r}\cdot\mathfrak{b} = \frac{\mathfrak{a}}{r}\cdot r\mathfrak{a}^{-1} \subseteq \mathcal{O}$ so $\frac{\mathfrak{a}}{r} \subseteq \mathfrak{b}^{-1}$. Hence, $\alpha\mathfrak{a} = \frac{a}{r}\mathfrak{a} \subseteq a\mathfrak{b}^{-1}$. Therefore, we have

$$\alpha\mathfrak{a} \subseteq a\mathfrak{b}^{-1} \subseteq \alpha(\mathfrak{a}^{-1})^{-1} \approx \alpha\mathfrak{a}$$

by Corollary 3.31. Hence, $\alpha\mathfrak{a} \approx a\mathfrak{b}^{-1}$ and finally we have

$$|N_{\mathcal{O}}(\alpha\mathfrak{a})| \approx |N_{\mathcal{O}}\left(a\mathfrak{b}^{-1}\right)| \lesssim \left(\frac{2}{\pi}\right)^s \sqrt{|d_{\mathcal{O}}|}.$$

□

Chapter 4

Symmetric Tensors

Let C be a hyperelliptic curve over $\mathbb{Q}$ given in the form as in (1.1)

$$C : z^2 = f(x, y) = f_0 x^n + f_1 x^{n-1} y + \ldots + f_n y^n, \tag{4.1}$$

where $f \in \mathbb{Z}[x, y]$ is a homogeneous polynomial of even degree $n = 2g + 2$, $g \geq 0, g \in \mathbb{Z}$.

Lemma 4.1. *Let $f(x, y) = f_0 x^n + \ldots + f_n y^n \in \mathbb{Q}[x, y]$ be a homogeneous polynomial. If $f(x, y)$ is reducible in $\mathbb{Q}[x, y]$, then all factors of $f(x, y)$ are homogeneous polynomials as well.*

Proof. Suppose

$$f(x, y) = f_1(x, y) \cdot \ldots \cdot f_k(x, y),$$

with total degree of each $f_i(x, y)$ being denoted by $d_i \in [1, n)$.

Notice that the highest degree part of each $f_i(x, y)$ has degree d_i, and their product is a monomial of $f(x, y)$ with degree n. On the other hand, the non-highest degree parts of each $f_i(x, y)$ all have degree $\leq d_i$, but their products are monomials of $f(x, y)$ with degree $< n$, which contradicts the fact that $f(x, y)$ is a homogeneous polynomial. Thus, the $f_i(x, y)$ are homogeneous polynomials. □

Lemma 4.2. *Let $f(x, y) = f_0 x^n + \ldots + f_n y^n \in \mathbb{Q}[x, y]$ be a homogeneous polynomial with $f_0 \neq 0$.*

Then $f(x, y) \in \mathbb{Q}[x, y]$ is irreducible if and only if $f(x, 1) \in \mathbb{Q}[x]$ is irreducible.

Proof. Suppose $f(x, y) \in \mathbb{Q}[x, y]$ is a reducible. Then by Lemma 4.1, f can be written as

$$f(x, y) = f_1(x, y) \cdot \ldots \cdot f_k(x, y), \tag{4.2}$$

where each $f_i(x, y) \in \mathbb{Z}[x, y]$ is homogeneous polynomial of degree ≥ 1. Since $f_0 \neq 0$, the degree of $f_i(x, y)$ is equal to the degree of $f_i(x, 1)$. Replacing y with 1 in (4.2) will give a

corresponding factorization of $f(x,1) \in \mathbb{Z}[x]$,

$$f(x,1) = f_1(x,1) \cdot \ldots \cdot f_k(x,1),$$

where $f_i(x,1) \in \mathbb{Q}[x]$ has degree ≥ 1. Thus, $f(x,1)$ is reducible as well.

Suppose $f(x,1) \in \mathbb{Q}[x]$ is reducible. Then it can be written as

$$g(x) = f(x,1) = f_1(x) \cdot \ldots \cdot f_k(x),$$

where each $f_i(x) \in \mathbb{Q}[x]$ is a polynomial of degree ≥ 1. Then

$$f(x,y) = y^n g(x/y) = y^{n_1} f_1(x/y) \cdot \ldots \cdot y^{n_k} f_k(x/y),$$

where n_i is the degree of $f_i(x)$. Thus, $f(x,y)$ is reducible. □

Definition 4.3. *Let $f(x,y) \in \mathbb{Z}[x,y]$ and $f'(x',y') \in \mathbb{Z}[x',y']$ be homogeneous polynomials.*

We say $f(x,y)$ is equivalent to $f'(x',y')$ via γ, denoted $f(x,y) \sim f'(x',y')$, if and only if there is a $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that

$$f(x,y) = f'(x',y'),$$

where $(x,y) = (x',y') \cdot \gamma$.

We make the following assumptions on f which will remain in force throughout the thesis:

1. Up to equivalence by an element in $\mathrm{SL}_2(\mathbb{Z})$, we may assume without loss of generality that $f_0 \neq 0$ throughout.
2. $f(x,y) \in \mathbb{Z}[x,y]$ is irreducible in $\mathbb{Q}[x,y]$, which implies $f(x,1) \in \mathbb{Z}[x]$ is irreducible in $\mathbb{Q}[x]$.

Lemma 4.4. *Suppose that $(x_0,y_0,z_0) \in \mathbb{Q}^3$ is a $\mathbb{Q}$-rational point on $z^2 = f(x,y)$. Then without loss of generality, by scaling (x,y) and z, we can assume $x_0,y_0,z_0 \in \mathbb{Z}$ and $(x_0,y_0) = 1$.*

Proof. Given a solution $(x_0,y_0,z_0) \in \mathbb{Q}^3$ to $z^2 = f(x,y)$, let λ be a nonzero integer such that $\lambda(x_0,y_0) \in \mathbb{Z}^2$. Write $\lambda(x_0,y_0) = d(x_1,y_1)$, where $d \in \mathbb{Z}$ is nonzero and the gcd of

(x_1, y_1) is 1. Then

$$\begin{aligned} f(x_1, y_1) &= f\left(\frac{\lambda x_0}{d}, \frac{\lambda y_0}{d}\right) \\ &= \left(\frac{\lambda}{d}\right)^n \cdot f(x_0, y_0) \\ &= \left(\left(\frac{\lambda}{d}\right)^{\frac{n}{2}} \cdot z_0\right)^2 \end{aligned}$$

By Gauss' Lemma, $\left(\frac{\lambda}{d}\right)^{\frac{n}{2}} \cdot z_0$ is an integer. □

Consider $K_f := \mathbb{Q}[x]/(f(x,1)) = \mathbb{Q}[\theta]$, where θ denotes the image of x in the $\mathbb{Q}$-algebra K_f. Since we assume $f(x,1)$ is irreducible in $\mathbb{Q}[x]$, K_f is indeed a number field rather than a general $\mathbb{Q}$-algebra.

The following is proven in [12, Prop. 1.1].

Theorem 4.5. *Let*

$$R_f = \langle 1, \zeta_1, \ldots, \zeta_{n-1} \rangle \tag{4.3}$$

be the $\mathbb{Z}$-module generated by the elements $1, \zeta_1, \ldots, \zeta_{n-1}$, where

$$\zeta_k = f_0\theta^k + f_1\theta^{k-1} + \ldots + f_{k-1}\theta, \quad 1 \leq k \leq n-1. \tag{4.4}$$

Then

1. *$\{1, \zeta_1, \ldots, \zeta_{n-1}\}$ is a $\mathbb{Z}$-basis for the $\mathbb{Z}$-module R_f,*
2. *R_f is an order in the number field K_f.*
3. *The multiplication in R_f is explicitly given by*

$$\zeta_i\zeta_j = \sum_{k=j+1}^{\min(i+j,n)} f_{i+j-k}\zeta_k - \sum_{k=\max(i+j-n,1)}^{i} f_{i+j-k}\zeta_k. \tag{4.5}$$

Theorem 4.6. *Let*

$$I_f(k) = \langle 1, \theta, \theta^2, \ldots, \theta^k, \zeta_{k+1}, \ldots, \zeta_{n-1} \rangle \tag{4.6}$$

be $\mathbb{Z}$-modules generated by $1, \theta, \ldots, \theta^k, \zeta_{k+1}, \ldots, \zeta_{n-1}$, where $0 \leq k \leq n-1$.

Then

1. *$\{1, \theta, \ldots, \theta^k, \zeta_{k+1}, \ldots, \zeta_{n-1}\}$ forms a $\mathbb{Z}$-basis for the $\mathbb{Z}$-module $I_f(k)$. Furthermore, $I_f(k) = I_f^k$ for $1 \leq k \leq n-1$, where $I_f = I_f(1)$.*
2. *$I_f(k)$ are ideals of the ring R_f.*

Proof. 1. See [3].

2. See [15, Prop. A.3].

□

Remark 4.7. *Let $I \subseteq K_f$ be a nonzero based R_f-module of the order $R_f \subseteq K_f$. Definition 3.34 gives the norm $N_{R_f}(I)$ of I. Since R_f in (4.3) depends only on the polynomial $f(x,y)$ in (4.1), and we will study this dependence in this chapter, it's more convenient to adopt Bhargava's notation, $N_f(I)$ instead of $N_{R_f}(I)$. This convention will remain in force throughout this chapter.*

Definition 4.8. ***Symmetric Tensors***
Consider a pair (I, α) such that

$$I^2 \subseteq \alpha \cdot I_f^{n-3}, \tag{4.7}$$

$$N_f(I)^2 = N_f(\alpha \cdot R_f) \cdot N_f(I_f^{n-3}), \tag{4.8}$$

where $N_f(I)$ is from Remark 4.7, $I \subseteq K_f$ is a nonzero based R_f-module, $\alpha \in K_f$, and $I^2 \subseteq \alpha I_f^{n-3}$.

Two pairs (I, α) and (J, β) are equivalent if and only if there exists $\kappa \in K_f^\times$ such that $J = \kappa I$ and $\beta = \kappa^2 \alpha$, which we denote by $(I, \alpha) \sim (J, \beta)$.

Let S_f denote the set of all equivalence classes of pairs (I, α) as above.

The set S_f defined above is stated in [3, Theorem 6] to be in bijection with elements in $\mathbb{Z}^2 \otimes \mathrm{Sym}_2\mathbb{Z}^n$ which have an invariant the binary n-ic form f with nonzero discriminant, where $\mathbb{Z}^2 \otimes \mathrm{Sym}_2\mathbb{Z}^n$ denotes the set of pairs (A, B) of symmetric $n \times n$ matrices with entries in $\mathbb{Z}$. Geometrically, such pairs (A, B) correspond to a pencil of quadrics which have invariant $(-1)^{n/2} \det(Ax - By) = f$. The general form of the bijection is proven in [16, Theorem 3.1 and 5.5].

Recall from [3, p.5]: If $f(x,y) \sim f'(x',y')$, via $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$, then the root θ' of $f'(x', 1)$ is given by

$$\theta' = \frac{d\theta - c}{-b\theta + a}, \tag{4.9}$$

where θ is the root of $f(x, 1)$, or equivalently,

$$\theta = \frac{a\theta' + c}{b\theta' + d}. \tag{4.10}$$

We state and give a self-contained proof of [3, (7)] below.

Theorem 4.9. *Suppose $f(x,y) \sim f'(x',y')$ via $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$.*

Then there is a $\mathbb{Z}$-module isomorphism $\phi : I_f(k) \cong I_{f'}(k)$ given by

$$\phi : \ \delta \longmapsto (-b\theta + a)^{-k} \cdot \delta, \quad \forall \delta \in I_f(k). \tag{4.11}$$

Proof. As $(-b\theta + a)^{-k}$ is a nonzero element of the number field K_f, ϕ is clearly a $\mathbb{Z}$-module isomorphism between $I_f(k)$ and $I_f(k)(-b\theta + a)^{-k}$, and it is left to show

$$I_f(k)(-b\theta + a)^{-k} = I_{f'}(k). \tag{4.12}$$

To show (4.12), it suffices to show the containment $I_f(k)(-b\theta + a)^{-k} \subseteq I_{f'}(k)$, as switching the roles of f and f', and replacing γ with its inverse γ^{-1} will obtain the reverse containment. The required containment (4.12) will be shown below.

□

Theorem 4.10. *$SL_2(\mathbb{Z})$ is generated by two elements $\gamma_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $\gamma_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.*

Proof. See [9, Theorem 4.1]. □

Lemma 4.11. *In addition to γ_1 and γ_2, $SL_2(\mathbb{Z})$ is also generated by γ_1, γ_2', where $\gamma_2' = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$.*

Proof. Let $\gamma_2' = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$, we have $\gamma_2' = \gamma_1 \cdot \gamma_2 \cdot \gamma_1^{-1}$, and it proves the lemma. □

It suffices to prove (4.12) for any set of generators for $SL_2(\mathbb{Z})$, which will be done in the next two lemmas.

Lemma 4.12. (4.12) *is true when $\gamma = \gamma_1$.*

Proof. Since $\gamma = \gamma_1$, we have that

$$(x, y) = (x', y') \cdot \gamma_1, \tag{4.13}$$

from Definition 4.3.

By (4.13), we have that $f'(x', y') = f(y', -x')$, so we obtain

$$\begin{aligned} f'(x', y') &= f_0' x'^n + f_1' x'^{n-1} y' + \ldots + f_{n-1}' x' y'^{n-1} + f_n' y'^n \\ &= f(y', -x') \\ &= f_0 y'^n + f_1(-x') y'^{n-1} + \ldots + f_{n-1} y' - x'^{n-1} + f_n - x'^n \\ &= f_0' y^n + (-f_1) x' y'^{n-1} + \ldots + (-f_{n-1}) y' x'^{n-1} + f_n x'^n. \end{aligned}$$

Equating coefficients of the corresponding monomials, we obtain the following relations between coefficients:

$$\begin{aligned} f_n &= f'_0 \\ -f'_{n-1} &= f'_1 \\ &\vdots \\ f_0 &= f'_n. \end{aligned}$$

Noticing on the right-hand side of (4.12), $I_{f'}(k)$ is a free $\mathbb{Z}$-module given by

$$I_{f'}(k) = \langle 1, \theta', \theta'^2, \ldots, \theta'^k, \zeta'_{k+1}, \ldots, \zeta'_{n-1} \rangle \tag{4.14}$$

$$= \langle 1, \theta^{-1}, \theta^{-2}, \ldots, \theta^{-k}, \zeta'_{k+1}, \ldots, \zeta'_{n-1} \rangle \tag{4.15}$$

by (4.13).

By γ_1, $I_f(k)(b\theta + a)^{-k}$ on the left-hand side of (4.12) becomes $I_f(k)(-\theta)^{-k}$, which is given by

$$I_f(k)(-\theta)^{-k} = (-\theta)^{-k} \cdot \langle 1, \theta, \theta^2, \ldots, \theta^k, \zeta_{k+1}, \ldots, \zeta_{n-1} \rangle \tag{4.16}$$

$$= \langle \theta^{-k}, \theta^{1-k}, \theta^{2-k}, \ldots, 1, \frac{\zeta_{k+1}}{\theta^k}, \ldots, \frac{\zeta_{n-1}}{\theta^k} \rangle. \tag{4.17}$$

A comparison between (4.17) and (4.15) shows an overlap of $\{1, \theta^{-1}, \theta^{-2}, \ldots, \theta^{-k}\}$, so it is left to show $\frac{\zeta_{k+1}}{\theta^k}, \ldots, \frac{\zeta_{n-1}}{\theta^k}$ from (4.17) are contained in (4.15) as a $\mathbb{Z}$-combination.

Now, let's consider $\frac{\zeta_{k+l}}{\theta^k}$, for $1 \leq l \leq n - (k+1)$, $l \in \mathbb{Z}$,

$$\begin{aligned} \frac{\zeta_{k+l}}{\theta^k} &= \frac{1}{\theta^k}\Big(f_0\theta^{k+l} + f_1\theta^{k+l-1} + \ldots + f_l\theta^k + \ldots + f_{k+l-1}\theta \Big) \quad \text{by (4.4)} \\ &= f_0\theta^l + f_1\theta^{l-1} + \ldots + f_{l-1}\theta + f_l + \frac{f_{l+1}}{\theta} + \ldots + \frac{f_{k+l-1}}{\theta^{k-1}} \\ &= \frac{f'_n}{(-\theta')^l} + \frac{-f'_{n-1}}{(-\theta')^{l-1}} + \ldots + \frac{(-1)^{l+1} \cdot f'_{n-l+1}}{(-\theta')} \\ &+ (-1)^l \cdot f'_{n-l} + (-1)^{l+1} \cdot f'_{n-l-1}(-\theta') + \ldots + (-1)^{k+l-1} \cdot f'_{n-k-l+1}(-\theta')^{k-1} \end{aligned}$$

by (4.13) and coefficients relation.

To simplify the above expression, we introduce the equivalence relation

$$a \equiv b \mod \Big(I_{f'}(k) \Big),$$

to mean

$$a \equiv b \bmod \Big(I_{f'}(k) \Big) \text{ if and only if } a - b \in I_{f'}(k),$$

where $a, b \in K_{f'}$. So according to this equivalence relation, $\frac{\zeta_{k+l}}{\theta^k} \equiv \frac{f'_n}{(-\theta')^l} + \frac{-f'_{n-1}}{(-\theta')^{l-1}} + \ldots + \frac{(-1)^{l+1} \cdot f'_{n-l+1}}{(-\theta')} \mod \Big(I_{f'}(k)\Big)$, that is,

$$\frac{\zeta_{k+l}}{\theta^k} \equiv \frac{f'_n + f'_{n-1}\theta' + \ldots + f'_{n-l+1}\theta'^{l-1}}{(\theta')^l} \mod \Big(I_{f'}(k)\Big). \tag{4.18}$$

Recall $(\theta', 1)$ is a root of $f'(x', 1)$, so that

$$f'_n + f'_{n-1}\theta' + \ldots + f'_{n-l+1}\theta'^{l-1} + f'_{n-l}\theta'^{l} + \ldots + f'_1\theta'^{n-1} + f'_0\theta'^{n} = 0.$$

Hence, we have that

$$f'_n + f'_{n-1}\theta' + \ldots + f'_{n-l+1}\theta'^{l-1} = -\Big(f'_{n-l}\theta'^{l} + \ldots + f'_1\theta'^{n-1} + f'_0\theta'^{n}\Big),$$

which can be subsituted into the top of (4.18). Thus, (4.18) becomes

$$\frac{\zeta_{k+l}}{\theta^k} \equiv f'_{n-l} + \ldots + f'_0\theta'^{n-l} \equiv \zeta'_{n-l} \mod (I_{f'}(k)), \quad 1 \leq l \leq n-(k+1), l \in \mathbb{Z}. \tag{4.19}$$

Then we know from (4.19) that $\frac{\zeta_{k+l}}{\theta^k} \equiv 0 \mod \Big(I_{f'}(k)\Big)$, for $1 \leq l \leq n-(k+1)$, which implies Lemma 4.12 is true when we use the generator $\gamma_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in SL_2(\mathbb{Z})$. □

Lemma 4.13. (4.12) *is true if* $\gamma = \gamma'_2$.

Proof. Since $\gamma = \gamma_2$, we have that

$$(x, y) = (x', y') \cdot \gamma'_2 \tag{4.20}$$

from Definition 4.3.

Notice from (4.20), we have $(x', y') = (x + y, y)$, so

$$\begin{aligned} f(x, y) &= f_0x^n + f_1x^{n-1}y + \ldots + f_{n-1}xy^{n-1} + f_ny^n \\ &= f'(x + y, y) \\ &= f'_0(x + y)^n + f'_1(x + y)^{n-1}y + \ldots + f'_{n-1}(x + y)y^{n-1} + f'_ny^n. \end{aligned}$$

From binomial expansion theorem, for $0 \leq k \leq n$, $k \in \mathbb{Z}$, we have

$$f_k = \sum_{l=0}^{k} f'_l \binom{n-l}{n-k}. \tag{4.21}$$

Now, let's look at both $I_f(k)(b\theta+a)^{-k}$ and $I_{f'}(k)$ in this case,

$$I_{f'}(k) = \langle 1, \theta', \theta'^2, \ldots, \theta'^k, \zeta'_{k+1}, \ldots, \zeta'_{n-1} \rangle. \tag{4.22}$$

$I_f(k)(b\theta+a)^{-k}$ becomes $I_f(k) = \langle 1, \theta, \theta^2, \ldots, \theta^k, \zeta_{k+1}, \ldots, \zeta_{n-1} \rangle$, and by $\theta' = \theta + 1$

$$I_f(k) = \langle 1, \theta'-1, (\theta'-1)^2, \ldots, (\theta'-1)^k, \zeta_{k+1}, \ldots, \zeta_{n-1} \rangle. \tag{4.23}$$

We can see from (4.23) and (4.22) that $1, \theta'-1, (\theta'-1)^2, \ldots, (\theta'-1)^k \in I_{f'}(k)$. So it suffices to show $\zeta_{k+1}, \ldots, \zeta_{n-1}, 0 \leq k \leq n-2$ in (4.23) can be written as $\mathbb{Z}$-linear combinations of $\zeta'_{k+1}, \ldots, \zeta'_{n-1}$ from (4.22).

In order to complete the argument, we require the next assertion.

Assertion 4.14. *For ζ_k in the basis $\{1, \zeta_1, \ldots, \zeta_k, \ldots, \zeta_{n-1}\}$ of R_f, we have that*

$$\zeta_k = a_k^{(k)}\zeta'_k + a_{k-1}^{(k)}\zeta'_{k-1} + \ldots + a_1^{(k)}\zeta'_1 - a_k^{(k)}\left(\zeta'_{k-1} + f'_{k-1}\right) - a_{k-1}^{(k)}\left(\zeta'_{k-2} + f'_{k-2}\right) - \ldots - a_1^{(k)} f'_0, \tag{4.24}$$

where

$$a_m^{(k)} = \binom{n-m}{n-k}, \quad \textit{for } 1 \leq m \leq k-1. \tag{4.25}$$

Proof. We prove it by induction on k.

The base case is when $k = 1$. By (4.4), $\zeta_1 = f_0\theta$, and from (4.20) and (4.21) we can write ζ_1 as $\zeta'_1 - f'_0$, which satisfies Assertion 4.14.

Next, we assume (4.24) is true for ζ_k, $1 < k < n-1$. By (4.4) again,

$$
\begin{aligned}
\zeta_{k+1} &= \theta\left(\zeta_k + f_k\right) \\
&= \left(\theta' - 1\right)\left(\zeta_k + f_k\right) \quad \text{by (4.20)} \\
&= \left(\theta' - 1\right)\left(\zeta_k + f_k' + \binom{n-k+1}{n-k} f_{k-1}' + \ldots + \binom{n}{n-k} f_0'\right) \quad \text{by (4.21)} \\
&= \left(\theta' - 1\right)\Bigg(\zeta_k' + a_{k-1}^{(k)}\zeta_{k-1}' + \ldots + a_1^{(k)}\zeta_1' - \left(\zeta_{k-1}' + f_{k-1}'\right) - a_{k-1}^{(k)}\left(\zeta_{k-2}' + f_{k-2}'\right) - \ldots - a_1^{(k)} f_0' \\
&\quad + f_k' + \binom{n-k+1}{n-k} f_{k-1}' + \ldots + \binom{n}{n-k} f_0'\Bigg) \quad \text{by induction hypothesis} \\
&= \left(\theta' - 1\right)\Bigg(\zeta_k' + f_k' + (a_{k-1}^{(k)} - 1)\zeta_{k-1}' + \left(\binom{n-k+1}{n-k} - 1\right) f_{k-1}' \\
&\quad + \left(a_{k-2}^{(k)} - a_{k-1}^{(k)}\right)\zeta_{k-2}' + \left(\binom{n-k+2}{n-k} - \binom{n-k+1}{n-k}\right) f_{k-2}' + \ldots \\
&\quad + (a_{k-l}^{(k)} - a_{k-l+1}^{(k)})\zeta_{k-l}' + \left(\binom{n-k+l}{n-k} - \binom{n-k+l-1}{n-k}\right) f_{k-l}' + \ldots \\
&\quad + (a_1^{(k)} - a_2^{(k)})\zeta_1' + \left(\binom{n-1}{n-k} - \binom{n-2}{n-k}\right) f_1' \\
&\quad + \left(\binom{n}{n-k} - a_1^{(k)}\right) f_0'\Bigg)
\end{aligned}
$$

By (4.25), we can write the above as

$$\begin{aligned}\zeta_{k+1} = \Big(\theta' - 1\Big)\Bigg(&\zeta'_k + f'_k \\ &+ \left(\binom{n-k+1}{n-k} - 1\right)\cdot(f'_{k-1} + \zeta'_{k-1}) \\ &+ \left(\binom{n-k+2}{n-k} - \binom{n-k+1}{n-k}\right)\cdot(f'_{k-2} + \zeta'_{k-2}) + \dots \\ &+ \left(\binom{n-k+l}{n-k} - \binom{n-k+l-1}{n-k}\right)\cdot(f'_{k-l} + \zeta'_{k-l}) + \dots \\ &+ \left(\binom{n-1}{n-k} - \binom{n-2}{n-k}\right)\cdot(f'_1 + \zeta'_1) \\ &+ \left(\binom{n}{n-k} - \binom{n-1}{n-k}\right)\cdot f'_0\Bigg).\end{aligned}$$

Notice that for $1 \le l \le k$, we have

$$\begin{aligned}\binom{n-k+l}{n-k} - \binom{n-k+l-1}{n-k} &= \frac{(n-k+l)!}{(n-k)!\ l!} - \frac{(n-k+l-1)!}{(n-k)!\ (l-1)!} \\ &= \frac{(n-k+l-1)!}{(n-k-1)!\ l!} \\ &= \binom{n-(k-l+1)}{n-(k+1)},\end{aligned}$$

which equals $a^{(k+1)}_{k-l+1}$ by (4.25).

Therefore we can rewrite ζ_{k+1} as

$$\begin{aligned}\zeta_{k+1} = \Big(\theta' - 1\Big)\Bigg(&\zeta'_k + f'_k \\ &+ a^{(k+1)}_k\cdot(f'_{k-1} + \zeta'_{k-1}) + a^{(k+1)}_{k-1}\cdot(f'_{k-2} + \zeta'_{k-2}) + \dots \\ &+ a^{(k+1)}_{k-l+1}\cdot(f'_{k-l} + \zeta'_{k-l}) + \dots + a^{(k+1)}_2\cdot(f'_1 + \zeta'_1) + a^{(k+1)}_1\cdot f'_0\Bigg) \\ = a^{(k+1)}_{k+1}&\zeta_{k'+1} + a^{(k+1)}_k\zeta'_k + a^{(k+1)}_{k-1}\zeta'_{k-1} + \dots + a^{(k+1)}_1\zeta'_1 \\ &- a^{(k+1)}_{k+1}\Big(\zeta'_k + f'_k\Big) - a^{(k+1)}_k\Big(\zeta'_{k-1} + f'_{k-1}\Big) - a^{(k+1)}_{k-1}\Big(\zeta'_{k-2} + f'_{k-2}\Big) - \dots - a^{(k+1)}_1 f'_0\end{aligned}$$

and it proves the desired result by using Assertion 4.14. □

An immediate result which follows from (4.12) is when $k = 0$,

$$I_f(0) = I_{f'}(0)$$

and it gives

$$R_f = R_{f'} \tag{4.26}$$

by (4.3) and (4.6). □

Corollary 4.15. *Suppose* $f(x,y) \sim f'(x',y')$ *via* $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$. *Let* S_f *and* $S_{f'}$ *denote sets of all equivalence classes of pairs* (I,α) *and* (I',α') *respectively as in Definition 4.8. Then there is a bijection* ψ *between* S_f *and* $S_{f'}$ *given by*

$$\psi: \ (I,\alpha) \longmapsto \left(I, (-b\theta + a)^{n-3} \cdot \alpha\right).$$

Proof. First, we check the image $\left(I, (-b\theta + a)^{n-3} \cdot \alpha\right) = \psi\left((I,\alpha)\right)$ is indeed an element in $S_{f'}$.

From Definition 4.8, $(I,\alpha) \in S_f$ satisfies

$$I^2 \subseteq \alpha \cdot I_f^{n-3}, \tag{4.27}$$

$$N_f(I)^2 = N_f(\alpha \cdot R_f) \cdot N_f(I_f^{n-3}). \tag{4.28}$$

By (4.12), (4.27) becomes

$$I^2 \subseteq (-b\theta + a)^{n-3}\alpha \cdot I_{f'}^{n-3}. \tag{4.29}$$

As for (4.28) we have

$$\begin{aligned} N_f(I)^2 &= N_f\left(\alpha \cdot (-b\theta + a)^{n-3} \cdot (-b\theta + a)^{-(n-3)} \cdot R_f\right) \cdot N_f(I_f^{n-3}) \\ &= N_f\left(\alpha \cdot (-b\theta + a)^{n-3} \cdot R_f\right) \cdot N_f\left((-b\theta + a)^{-(n-3)} \cdot I_f^{n-3}\right) \end{aligned}$$

by Theorem 5.1.

By (4.12) again, the above equation gives

$$N_f(I)^2 = N_f\left(\alpha \cdot (-b\theta + a)^{n-3} \cdot R_f\right) \cdot N_f(I_{f'}^{n-3}). \tag{4.30}$$

Moreover, (4.30) implies

$$N_{f'}(I)^2 = N_{f'}\left(\alpha \cdot (-b\theta + a)^{n-3} \cdot R_{f'}\right) \cdot N_{f'}(I_{f'}^{n-3}) \tag{4.31}$$

by (4.26) and Remark 3.35, Therefore, from (4.29) and (4.31) we can verify $\left(I, (-b\theta + a)^{n-3} \cdot \alpha\right) = \psi\left((I, \alpha)\right)$ is an element in $S_{f'}$. Note also that the map ψ is well-defined on equivalence classes.

Next, we switch the roles of $f(x, y)$ and $f'(x', y')$, and replacing γ with its inverse $\gamma^{-1} = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \in SL_2(\mathbb{Z})$. Then consider a map ψ' from $S_{f'}$ to S_f given by

$$\psi' : \left(I', \alpha'\right) \longmapsto \left(I', (-b'\theta' + a')^{n-3} \cdot \alpha'\right), \tag{4.32}$$

where θ' is the root of $f'(x', 1)$ satisfies (4.9).

Following the same argument as before, we can show the image $\left(I', (-b'\theta' + a')^{n-3} \cdot \alpha'\right) = \psi'\left(I', \alpha'\right)$ is an element in S_f.

Furthermore, we can verify that

$$\begin{aligned}
\psi\left(\psi'\left(I', \alpha'\right)\right) &= \psi\left(I', (-b'\theta' + a')^{n-3} \cdot \alpha'\right) \\
&= \left(I', (-b\theta + a)^{n-3} \cdot (-b'\theta' + a')^{n-3} \cdot \alpha'\right) \\
&= \left(I', \left(-b \cdot \left(\frac{a\theta' + c}{b\theta' + d}\right) + a\right)^{n-3} \cdot (-b'\theta' + a')^{n-3} \cdot \alpha'\right) \quad \text{by (4.10)} \\
&= \left(I', (-b'\theta' + a')^{-(n-3)} \cdot (-b'\theta' + a')^{n-3} \cdot \alpha'\right) \\
&= \left(I', \alpha'\right).
\end{aligned}$$

Also,

$$\begin{aligned}
\psi'\left(\psi\left(I, \alpha\right)\right) &= \psi'\left(I, (-b\theta + a)^{n-3} \cdot \alpha'\right) \\
&= \left(I, (-b'\theta' + a')^{n-3} \cdot (-b\theta + a)^{n-3} \cdot \alpha\right) \\
&= \left(I, \left(-b' \cdot \left(\frac{d\theta - c}{-b\theta + a}\right) + a'\right)^{n-3} \cdot (-b\theta + a)^{n-3} \cdot \alpha\right) \quad \text{by (4.9)} \\
&= \left(I, (-b\theta + a)^{-(n-3)} \cdot (-b\theta + a)^{n-3} \cdot \alpha\right) \\
&= \left(I, \alpha\right).
\end{aligned}$$

We can see from the above that $\psi \circ \psi' = 1_{S_{f'}}$ and $\psi' \circ \psi = 1_{S_f}$. So ψ is a bijection from S_f to $S_{f'}$. □

Theorem 4.16. *Suppose (x_0, y_0, z_0) is an integer solution to $z^2 = f(x, y)$ with $(x_0, y_0) = 1$. Then this solution gives rise to an element $(I, \alpha) \in S_f$.*

Proof. Applying a $\gamma \in SL_2(\mathbb{Z})$ to $f(x,y)$, we may assume $(x_0,y_0)\cdot\gamma = (0,1)$. Notice, in doing this operation, we produce a new $f'(x',y')$ with $z_0'^2 = f_n'$, and apply γ^{-1} will yield the original $f(x,y)$.

Next, set $\alpha' = \theta'$ and note that

$$\theta' I_{f'}^{n-3} = \langle z_0'^2, \theta', \theta'^2, \ldots, \theta'^{n-2}, f_0'\theta'^{n-1}\rangle. \tag{4.33}$$

Let

$$I' = \langle z_0', \theta' I_{f'}^{(n-4)/2}\rangle = \langle z_0', \theta', \theta'^2, \ldots, \theta'^{(n-2)/2}, \zeta_{n/2}', \ldots, \zeta_{n-1}'\rangle. \tag{4.34}$$

In order to show I' is a $R_{f'}$-module, recall $R_{f'} = \langle 1, \zeta_1', \ldots, \zeta_{n-1}'\rangle$. We need to check that for every for $i = 1, \ldots, n-1$, ζ_i' times each of the elements

$$z_0', \theta', \theta'^2, \ldots, \theta'^{(n-2)/2}, \zeta_{n/2}', \ldots, \zeta_{n-1}' \tag{4.35}$$

is a $\mathbb{Z}$-linear combination of the same elements above.

Note that in the newly produced polynomial $f'(x',y')$, we have

$$z_0'^2 = f_n'. \tag{4.36}$$

Also, by (4.4), we can verify that

$$\zeta_n' = -f_n' = z_0'^2 \in I' \tag{4.37}$$

$$\zeta_1', \ldots, \zeta_n' \in I'. \tag{4.38}$$

First, we can see that (4.38) guarantees each element in $\{1, \zeta_1', \ldots, \zeta_{n-1}'\}$ times z_0' is a $\mathbb{Z}$-linear combination of elements in (4.35).

Next, let us consider $\zeta_i \cdot (\theta')^k$, for $1 \le i \le n-1$ and $1 \le k \le \frac{n-2}{2}$. Notice that there are two possibilities of this product:

- $k + i \le n$.

  When we have this case,

$$\begin{aligned}\zeta_i' \cdot (\theta')^k &= \left(f_0'\theta'^i + \ldots + f_{i-1}'\theta'\right)\cdot\left(\theta'\right)^k \\ &= f_0'\theta'^{i+k} + \ldots + f_{i-1}'\theta'^{k+1} \\ &= \zeta_{i+k}' - f_i'\theta'^k - f_{i+1}'\theta'^{k-1} - \ldots - f_{i+k-1}'\theta'.\end{aligned}$$

  It's clear that the above product is a $\mathbb{Z}$-linear combination of elements in (4.35).

- $k + i > n$.

Again, we consider the product

$$\begin{aligned}
\zeta'_i \cdot (\theta')^k &= \left(f'_0\theta'^i + \ldots + f'_{i-1}\theta'\right) \cdot \left(\theta'\right)^k \\
&= \left(f'_0\theta'^i + \ldots + f'_{i-1}\theta'\right) \cdot \left(\theta'\right)^{n-i} \cdot \left(\theta'\right)^{k-n+i} \\
&= \left(f'_0\theta'^n + \ldots + f'_{i-1}\theta'^{n-i+1}\right) \cdot \left(\theta'\right)^{k-n+i} \\
&= \left(\zeta'_n - f'_i\theta'^{n-i} - \ldots - f'_{n-1}\theta'\right) \cdot \left(\theta'\right)^{k-n+i} \\
&= \zeta'_n \cdot \theta'^{k-n+i} - f'_i\theta'^k - f'_{i+1}\theta'^{k-1} - \ldots - f'_{n-1}\theta'^{k-n+i-1} \\
&= -f'_n \cdot \theta'^{k-n+i} - f'_i\theta'^k - f'_{i+1}\theta'^{k-1} - \ldots - f'_{n-1}\theta'^{k-n+i-1}
\end{aligned}$$

by (4.37).

From $k+i > n$, $1 \leq i \leq n-1$ and $1 \leq k \leq \frac{n-2}{2}$, we have $1 \leq n-i < k \leq \frac{n-2}{2}$ and that gives lower limits for exponents $k-n+i$ and $k-n+i-1$. $k-n+i = k-(n-i) < k \leq \frac{n-2}{2}$ and $k-n+i-1 = k-(n-i-1) \leq k \leq \frac{n-2}{2}$.

Therefore, we can see that $\zeta'_i \cdot (\theta')^k$ in this situation is also a $\mathbb{Z}$-linear combination of elements of (4.35).

Finally, for $i = 1, \ldots, n-1$ and $j = n/2, \ldots, n-1$, we have (4.5)

$$\zeta'_i\zeta'_j = \sum_{k=j+1}^{\min(i+j,n)} f_{i+j}\zeta'_k - \sum_{k=\max(i+j-n,1)}^{i} f_{i+j-k}\zeta'_k$$

which expresses this product as a $\mathbb{Z}$-linear combination of $\zeta'_1, \ldots, \zeta'_n$. By (4.38), we can conclude that $\zeta'_i\zeta'_j$ is a $\mathbb{Z}$-linear combination of elements of (4.35). This concludes the verification that I' is a $R_{f'}$-module.

By (4.4), we can verify that the elements

$$\theta', \theta'^2, \ldots, \theta'^{(n-2)/2}, \zeta'_{n/2}, \ldots, \zeta'_{n-1} \tag{4.39}$$

in the $\mathbb{Z}$-basis of I' can each be obtained as a $\mathbb{Z}$-combination of $\theta', \theta'^2, \ldots, \theta'^{n-2}, f'_0\theta'^{n-1}$ and hence lie in $\theta' \cdot I'^{n-3}_{f'}$ by (4.33). Thus, z_0 times any element in (4.39) can be written as as a $\mathbb{Z}$-combination of $\theta', \theta'^2, \ldots, \theta'^{n-2}, f'_0\theta'^{n-1}$ and hence lies in $\theta' \cdot I'^{n-3}_{f'}$. Also, from (4.34), it's clear that $(\theta')^i \cdot (\theta')^j, 1 \leq i, j \leq \frac{n-2}{2}$ lie in $\theta' \cdot I'^{n-3}_{f'}$. Lastly, we have verified $(\zeta_i)' \cdot (\theta')^k$ as well as $\zeta'_i\zeta'_j$ are contained in $\theta' \cdot I'^{n-3}_{f'}$. Therefore, we have that

$$I'^2 \subseteq \theta' \cdot I'^{n-3}_{f'}.$$

Furthermore, with (4.4), (4.33) and (4.34), we can write down transition matrices T_1 from I' to $R_{f'}$, and T_2 from $\theta' I'^{n-3}_{f'}$ to $R_{f'}$ as follows:

$$T_1 = \begin{pmatrix} z_0' & * & \dots & \dots & \dots & \dots & \dots & \dots & * \\ 0 & \frac{1}{f_0'} & * & \dots & \dots & \dots & \dots & \dots & * \\ \vdots & 0 & \frac{1}{f_0'} & * & \dots & \dots & \dots & \dots & * \\ \vdots & \vdots & \ddots & \ddots & * & \dots & \dots & \dots & * \\ \vdots & 0 & \dots & 0 & \frac{1}{f_0'} & * & \dots & \dots & * \\ \vdots & 0 & \dots & \dots & 0 & 1 & * & \dots & * \\ \vdots & \vdots & & & & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & & & & & \ddots & \ddots & * \\ 0 & 0 & \dots & 0 & \dots & \dots & \dots & 0 & 1 \end{pmatrix}$$

Noticing from Definition 3.34, $N_{f'}(I') = |\mathrm{Det}(T_1)|$ and for the sake of calculating $\mathrm{Det}(T_1)$, it's not necessary for us to figure out all $*$'s.

$$N_{f'}(I') = |\mathrm{Det}(T_1)| = |z_0'/f_0'^{n-2/2}|. \tag{4.40}$$

Similarly, for T_2, where

$$T_2 = \begin{pmatrix} z_0'^2 & * & \dots & \dots & \dots & \dots & \dots & \dots & * \\ 0 & \frac{1}{f_0'} & * & \dots & \dots & \dots & \dots & \dots & * \\ \vdots & 0 & \frac{1}{f_0'} & * & \dots & \dots & \dots & \dots & * \\ 0 & \vdots & \ddots & \ddots & * & \dots & \dots & \dots & * \\ \vdots & 0 & \dots & 0 & \frac{1}{f_0'} & * & \dots & \dots & * \\ \vdots & 0 & \dots & \dots & 0 & \frac{1}{f_0'} & * & \dots & * \\ \vdots & \vdots & & & & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & & & & 0 & \frac{1}{f_0'} & * \\ 0 & 0 & \dots & 0 & \dots & \dots & \dots & 0 & 1 \end{pmatrix}$$

and we can have

$$N_{f'}(\theta' I_{f'}^{n-3}) = |\mathrm{Det}(T_2)| = |z_0'^2/f_0'^{n-2}|. \tag{4.41}$$

Thus, by (4.40) and (4.41), we get

$$N_{f'}(I')^2 = \left(z_0'/f_0'^{(n-2)/2}\right)^2 = N_{f'}(\theta' I_{f'}^{n-3}),$$

which equals

$$N_{f'}(I')^2 = \left(z_0'/f_0'^{(n-2)/2}\right)^2 = N_{f'}(\theta') \cdot N_{f'}(I_{f'}^{n-3})$$

by Theorem 5.1. Therefore, (I', θ') lies in $S_{f'}$. By Corollary 4.15, $(I', \theta') \in S_{f'}$ corresponds to a pair $(I, \alpha) \in S_f$. It is verified in [3, p. 8] that the pair (I, α) in S_f does not depend on the choice of γ. □

Proposition 4.17. *Let $(x_0, y_0, z_0) \in \mathbb{Z}^3$ be a solution to $z^2 = f(x, y)$ with $(x_0, y_0) = 1$, and $(I, \alpha) \in S_f$ be the element associated to (x_0, y_0, z_0) from Theorem 4.16. Then $(I, \alpha) \sim (I_1, \alpha_1) \in S_f$ where*

$$c_0 \cdot \alpha_1 \in R_f,$$

for some positive integer c_0 which only depends on f.

Proof. We apply a $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$ such that $(x_0, y_0) \cdot \gamma = (0, 1)$ so that $(0, 1, z_0')$ is an integer solution to $z'^2 = f'(x', y')$, where $f(x, y) \sim f'(x', y')$ via γ.

By Theorem 4.16, the solution $(0, 1, z_0')$ to $z'^2 = f'(x', y')$ is associated with a pair (I', θ'). Moreover, $(I', \theta') \in S_{f'}$ corresponds to $\left(I', \theta' \cdot (-b\theta + a)^{n-3}\right) \in S_f$ from Corollary 4.15.

Using (4.9), we know

$$\left(I', \theta' \cdot (-b\theta + a)^{n-3}\right) = \left(I', \frac{d\theta - c}{(-b\theta + a)^{n-2}}\right),$$

which is equivalent to $\left((-b\theta + a)^{\frac{n-2}{2}} \cdot I', d\theta - c\right) \in S_f$ by Definition 4.8.

Recall from (4.3) and (4.4), $\zeta_1 = f_0\theta$ is an element in a $\mathbb{Z}$-basis of R_f. Thus, if we multiply $d\theta - c$ by $f_0 \in \mathbb{Z}$, we have both $d\theta \cdot f_0 \in R_f$ and $c \cdot f_0 \in \mathbb{Z} \subseteq R_f$. □

Chapter 5

Main Theorems

In this chapter, we prove the main theorems of this thesis and end with some concluding remarks.

Recall from Theorem 4.16 in Chapter 4 that a solution $(x, y, z) \in \mathbb{Z}^3$ to a hyperelliptic curve $z^2 = f(x, y)$ in (4.1) gives rise to an element $(I, \alpha) \in S_f$, where we recall that

$$I^2 \subseteq \alpha I_f^{n-3}, \tag{5.1}$$

$$N_f(I)^2 = N_{K_f/\mathbb{Q}}(\alpha) \cdot N_f(I_f^{n-3}). \tag{5.2}$$

From Theorem 3.49, there exists an element $\beta \in K_f^*$ such that $J = I\beta$ is an ideal of the order R_f with

$$N_f(J) \lesssim \left(\frac{2}{\pi}\right)^s \cdot \sqrt{|d_{R_f}|}.$$

The following proposition shows that the norm of a general nonzero order, i.e., $N_{\mathcal{O}}(\cdot)$ is multiplicative when one of the $\mathcal{O}$-modules is principal. Thus, this result can be applied to $N_f(\cdot)$.

Proposition 5.1. *Let I be a nonzero finitely generated $\mathcal{O}$-module and β be an element in K^*. Then $N_{\mathcal{O}}(I\beta) = N_{\mathcal{O}}(I \cdot \beta\mathcal{O}) = N_{\mathcal{O}}(I) \cdot N_{\mathcal{O}}(\beta\mathcal{O})$.*

Proof. We know from Theorem 3.9 that I is a free $\mathbb{Z}$-module of rank n, so we write

$$I = \mathbb{Z}\gamma_1 + \ldots + \mathbb{Z}\gamma_n, \tag{5.3}$$

where $\{\gamma_1, \ldots, \gamma_n\}$ is a $\mathbb{Z}$-basis of I.

We also know that $\mathcal{O}$ has a $\mathbb{Z}$-basis $\{\alpha_1, \ldots, \alpha_n\}$. Thus, $\{\beta\alpha_1, \ldots, \beta\alpha_n\}$ will be a $\mathbb{Z}$-basis of $\beta\mathcal{O}$ and we write

$$\beta\mathcal{O} = \mathbb{Z}\beta\alpha_1 + \ldots + \mathbb{Z}\beta\alpha_n. \tag{5.4}$$

Notice that (5.3) and (5.4) allows us to specifically write down transition matrices $A = (a_{ij})$ and $B = (b_{ij})$, from a $\mathbb{Z}$-basis of I to a $\mathbb{Z}$-basis of $\mathcal{O}$ and from a $\mathbb{Z}$-basis of $\beta\mathcal{O}$ to a

$\mathbb{Z}$-basis of $\mathcal{O}$ respectively, that are,

$$\gamma_i = \sum_{j=1}^{n} a_{ij}\alpha_j \quad \text{and} \quad \beta\alpha_i = \sum_{j=1}^{n} b_{ij}\alpha_j, \qquad \text{where } a_{ij}, b_{ij} \in K^*. \tag{5.5}$$

Next, we consider the product of two $\mathcal{O}$-modules I and $\beta\mathcal{O}$. From (5.3), we can see $\{\beta\gamma_1, \ldots, \beta\gamma_n\}$ is a $\mathbb{Z}$-basis for $I\beta = I \cdot \beta\mathcal{O}$, so we have

$$I \cdot \beta\mathcal{O} = \mathbb{Z}\beta\gamma_1 + \ldots + \mathbb{Z}\beta\gamma_n. \tag{5.6}$$

Let $C = (c_{ij})$ be the transition matrix from a $\mathbb{Z}$-basis of $I \cdot \beta\mathcal{O}$ to a $\mathbb{Z}$-basis of $\mathcal{O}$,

$$\sum_{j=1}^{n} c_{ik}\alpha_k = \beta\gamma_i \tag{5.7}$$

and because of (5.5) the right-hand side of (5.7) can be written as

$$\begin{aligned}\beta\gamma_i = \sum_{j=1}^{n} a_{ij}\beta\alpha_j &= \sum_{j=1}^{n} a_{ij} \sum_{k=1}^{n} b_{jk}\alpha_k, \\ &= \sum_{k=1}^{n} \left(\sum_{j=1}^{n} a_{ij}b_{jk} \right) \alpha_k.\end{aligned}$$

Therefore, we have $C = A \cdot B$, and by Definition 3.34, we obtain

$$N_{\mathcal{O}}(I\beta) = N_{\mathcal{O}}(I) \cdot N_{\mathcal{O}}(\beta\mathcal{O}).$$

□

We consider the pair (I, α) from Theorem 4.16 again, which satisfies (5.1) and (5.2), and let $\beta \in K^*$ be chosen from Theorem 3.49 such that $N_f(I\beta) \lesssim (\frac{2}{\pi})^s \cdot \sqrt{|d_{R_f}|}$.

By Theorem 5.1, if we multiply the pair (I, α) by β^2, norm equation (5.2) becomes

$$N_f(I\beta)^2 = N_{K_f/\mathbb{Q}}(\alpha\beta^2) \cdot N_f(I_f^{n-3}). \tag{5.8}$$

Next, we will give an an important property of $N_f(I_f^k)$ on the right-hand side of (5.8).

Proposition 5.2. *We have that* $N_f(I_f^k) = \frac{1}{f_0^k}$.

Proof. Recall from (4.6) in Theorem 4.6 that

$$I_f^k = I_f(k) = \langle 1, \theta, \theta^2, \ldots, \theta^k, \zeta_{k+1}, \ldots, \zeta_{n-1} \rangle,$$

where

$$\zeta_k = f_0\theta^k + f_1\theta^{k-1} + \ldots + f_{k-1}\theta, \quad \forall f_i's \in \mathbb{Z}.$$

From (4.3) we have that

$$R_f = \langle 1, \zeta_1, \zeta_2, \zeta_3, \zeta_4, \ldots, \zeta_{n-1} \rangle \subseteq I_f^k.$$

Then a transition matrix from a basis of I_f^k to a basis of the order R_f can be written down as follows:

$$T = \begin{pmatrix}
1 & * & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & * \\
0 & \frac{1}{f_0} & * & \ldots & \ldots & \ldots & \ldots & \ldots & * \\
\vdots & 0 & \frac{1}{f_0} & \ldots & \ldots & \ldots & \ldots & \ldots & * \\
\vdots & \vdots & \ddots & \ddots & \ldots & \ldots & \ldots & \ldots & * \\
\vdots & 0 & \ldots & 0 & \frac{1}{f_0} & * & \ldots & \ldots & * \\
\vdots & 0 & \ldots & \ldots & 0 & 1 & * & \ldots & * \\
\vdots & \vdots & & & & \ddots & \ddots & \ddots & \vdots \\
\vdots & \vdots & & & & & \ddots & \ddots & * \\
0 & 0 & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & 1
\end{pmatrix}$$

It's not necessary for us to figure out all $*$'s in T for the purpose of calculating $\det(T)$. Since there are k number of $\frac{1}{f_0}'s$ along the diagonal of T, by Definition 3.34, we have $N_f\left(I_f(k)\right) = N_f(I_f^k) = \frac{1}{f_0^k}$. □

By the preceding proposition, we know that $N_f(I_f^k) = \frac{1}{f_0^k}$. So $N_f(I_f^{n-3}) = \frac{1}{f_0^{n-3}}$ in (5.2) is fixed when the hyperelliptic curve $z^2 = f(x, y)$ is fixed.

Proposition 5.3. *Consider I_f^{n-3} from (5.1). There exists a nonzero constant c_1 dependent only on R_f such that $c_1 \cdot I_f^{n-3} \subseteq R_f$.*

Proof. Using the same idea as in Proposition 5.2, we write down a transition matrix T' from a $\mathbb{Z}$-basis of R_f to a $\mathbb{Z}$-basis of I_f^{n-3} as follows:

$$T' = \begin{pmatrix}
1 & 0 & 0 & \ldots & 0 & \ldots & \ldots & \ldots & 0 \\
0 & f_0 & f_1 & \vdots & f_{n-3} & \vdots & \vdots & \vdots & 0 \\
0 & 0 & f_0 & \vdots & \vdots & \ldots & \ldots & \ldots & 0 \\
0 & 0 & 0 & \ddots & f_1 & \ldots & \ldots & \ldots & 0 \\
\vdots & \vdots & \vdots & \vdots & f_0 & \ldots & \ldots & \ldots & 0 \\
0 & 0 & \vdots & \vdots & 0 & 1 & \ldots & \ldots & 0 \\
\vdots & 0 & \vdots & \vdots & 0 & 0 & 1 & \ldots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & \ldots & \ldots & \ldots & \ldots & 0 & 0 & 1
\end{pmatrix}$$

with determinant $\det(T') = f_0^{n-3} \neq 0 \in \mathbb{Z}$.

Then, its inverse $T'^{-1} = \frac{1}{\det(T')} \cdot \mathrm{Adj}(T')$ will be a transition matrix from a $\mathbb{Z}$-basis of I_f^{n-3} to a $\mathbb{Z}$-basis of R_f with entries $\in \mathbb{Q}$. Let $c_1 = \det(T') = f_0^{n-3}$, then $c_1 \cdot T'^{-1} = \mathrm{Adj}(T')$ is a transition matrix from a $\mathbb{Z}$-basis of $c_1 \cdot I_f^{n-3}$ to a $\mathbb{Z}$-basis of R_f and the definition of adjoint matrix guarantees all entries in $\mathrm{Adj}(T')$ to be integers. Therefore, $c_1 \cdot T'^{-1} = \mathrm{Adj}(T')$ is indeed a $\mathbb{Z}$-transition matrix from a $\mathbb{Z}$-basis of $c_1 \cdot I_f^{n-3}$ to a $\mathbb{Z}$-basis of R_f so that $c_1 \cdot I_f^{n-3} \subseteq R_f$ as $\mathbb{Z}$-modules. □

From Proposition 5.2 and (5.8), we obtain that

$$|N_{K_f/\mathbb{Q}}(\alpha\beta^2)| \lesssim \left(\frac{2}{\pi}\right)^s \sqrt{|d_{R_f}|}.$$

Note that this provides insufficient information to show there are finitely many possibilities for $\gamma = \alpha\beta^2$. However, the additional constraint which comes from the next proposition will be used later to show that there are only finitely many possibilities for γ.

Lemma 5.4. *Let $(I, \alpha) \in S_f$ arise from a $\mathbb{Q}$-rational point on the hyperelliptic curve as in Theorem 4.16. With Definition 3.39 and the above proposition, we can derive a new relationship between I^2 and αI_f^{n-3} in addition to* (5.1)*, that is,*

$$I^2 \approx \alpha I_f^{n-3}.$$

Proof. Let c_0, c_1 be positive integers result from Proposition 4.17 and 5.3.

Consider $c_3 = c_0 \cdot c_1$, then we have

$$c_3 \cdot \alpha I_f^{n-3} \subseteq R_f. \tag{5.9}$$

Recall from (5.1) that $I^2 \subseteq \alpha I_f^{n-3}$. Therefore, by (5.9), we have

$$c_3^2 I^2 \subseteq c_3^2 \alpha I_f^{n-3} \subseteq R_f. \tag{5.10}$$

Similar as in (5.8), (5.10) also gives a norm equation :

$$N_f(c_3 I)^2 = N_{K_f/\mathbb{Q}}(c_3^2\alpha) \cdot N_f(I_f^{n-3}) \tag{5.11}$$

and on whose right-hand side, we have

$$\begin{aligned} N_{K/\mathbb{Q}}(c_3^2\alpha) \cdot N_f(I_f^{n-3}) &= N_{\mathcal{O}_K}(c_3^2\alpha) \cdot N_f(I_f^{n-3}) \\ &\approx N_f(c_3^2\alpha) \cdot N_f(I_f^{n-3}) \quad \text{(by Lemma 3.46)} \\ &= N_f(c_3^2\alpha I_f^{n-3}) \quad \text{(by Theorem 5.1).} \end{aligned}$$

While, on the left-hand side of (5.11):

$$N_f(c_3 I)^2 = N_f(c_3 I) \cdot N_f(c_3 I) \approx N_f(c_3^2 I^2) \quad \text{(by Lemma 3.48).}$$

Therefore, we can derive an "approximate" version of (5.8) as follows:

$$N_f(c_3^2 I^2) \approx N_f(c_3 I)^2 = N_{K_f/\mathbb{Q}}(c_3^2 \alpha) N_f(I_f^{n-3}) \approx N_f(c_3^2 \alpha I_f^{n-3}). \tag{5.12}$$

Notice that (5.10) and (5.12) combines with Lemma 3.44 gives us $c_3^2 I^2 \approx c_3^2 \alpha I_f^{n-3}$, so we we obtain $I^2 \approx \alpha I_f^{n-3}$ as desired. □

Proposition 5.5. *Let (I, α) arise from a $\mathbb{Q}$-rational point on the hyperelliptic curve $z^2 = f(x, y)$. Let β be chosen as in Theorem 3.49 such that $I\beta \subseteq R_f$ and $|N_f(I\beta)| \lesssim M$, where $M = \left(\frac{2}{\pi}\right)^s \sqrt{|d_{R_f}|}$. Then there is a positive integer c only dependent on R_f such that $c\alpha\beta^2 \in R_f$.*

Proof. From Lemma 5.4, we have $I^2 \approx \alpha I_f^{n-3}$ and it gives $R_f \approx \alpha I^{-2} I_f^{n-3}$ by Lemma 3.28. Recall from the proof of Theorem 3.49, we choose a $r \in K^*$ so that

$$rI^{-1} := \mathfrak{b} \subseteq R_f \tag{5.13}$$

$$\implies r^2 I^{-2} = \mathfrak{b}^2 \subseteq R_f, \tag{5.14}$$

From Lemma 3.21, I^{-1} is an R_f-module, and we can see from (5.13) and (5.14) that both $\mathfrak{b}$ and $\mathfrak{b}^2$ are ideals of R_f.

Moreover, we know from Theorem 3.19 that there will exist an $a \in \mathfrak{b}$ such that $|N_{K/\mathbb{Q}}(a)| \leq (\frac{2}{\pi})^s \sqrt{|d_{R_f}|}$.

So the choice of β is given by $\beta = a/r$, and

$R_f \approx \alpha I^{-2} I_f^{n-3}$
$\implies R_f \approx \alpha r^{-2} \mathfrak{b}^2 I_f^{n-3}$ Since $r^2 I^{-2} = \mathfrak{b}^2$ by (5.14)
$\implies \alpha a^2 r^{-2} \mathfrak{b}^2 I_f^{n-3} \approx a^2 R_f \lesssim \mathfrak{b}^2 R_f$ Multiply by $a^2 \in \mathfrak{b}^2 \subseteq \mathfrak{b}$ on both sides
$\implies \alpha a^2 r^{-2} I_f^{n-3} \lesssim R_f$ Multiply by $\mathfrak{b}^{-2}$since it is also an R_f-module from Lemma 3.21
$\implies \alpha\beta^2 \lesssim R_f$ $\beta^2 = a^2 r^{-2}$ and $1 \in R_f \subseteq I_f^{n-3}$
$\implies$ There is a positive integer c dependent only on R_f such that $c\alpha\beta^2 \in R_f$.

□

Putting together Theorem 3.49 and Proposition 5.5, we obtain:

Theorem 5.6. *Let $f \in \mathbb{Z}[x, y]$ be a homogeneous polynomial of even degree $n \geq 2$ which is irreducible in $\mathbb{Q}[x, y]$. Let $(I, \alpha) \in S_f$ be any element which arises from a $\mathbb{Q}$-rational point on the hyperelliptic curve $z^2 = f(x, y)$ as in Theorem 4.16. Then $(I, \alpha) \sim (J, \gamma)$ where*

1. $J \subseteq R_f$ *and* $|N_f(J)| \lesssim (\frac{2}{\pi})^s \sqrt{|d_{R_f}|}$,

2. $|N_{K_f/\mathbb{Q}}(\gamma)| \lesssim (\frac{2}{\pi})^s \sqrt{|d_{R_f}|}$,

3. *there is a positive integer* c *only dependent on* f *such that* $c\gamma \in R_f$.

Lemma 5.7. *Let* K *be a number field and consider all embeddings* $\sigma : K \hookrightarrow \mathbb{C}$. *Let* $\mathcal{O}$ *be an order in* K. *For any principal* $\mathcal{O}$*-module* $\mathfrak{a}$ *in* K, *there exists an element* $b \in K^*$ *such that* $b\mathcal{O} = \mathfrak{a}$ *and* $|b_\sigma| \in O\left(|N_{\mathcal{O}}(\mathfrak{a})^{\frac{1}{n}}|\right)$ *for every embedding* σ, *where* b_σ *means* $\sigma(b)$.

Proof. Consider the Minkowski unit map given by

$$\begin{aligned} \tau : K^* &\longrightarrow \mathbb{C}^n \\ u &\longrightarrow \tau(u)_\sigma, \end{aligned}$$

where

$$\tau(u)_\sigma = \begin{cases} \left(\log|u_\sigma| - \frac{1}{n}\log|N_{K/\mathbb{Q}}(u)|\right)_\sigma, & \text{if } \sigma \text{ is a real embedding,} \\ \\ \left(\log|u_\sigma|^2 - \frac{1}{n}\log|N_{K/\mathbb{Q}}(u)|\right)_\sigma, & \text{if } \sigma \text{ is a complex embedding.} \end{cases}$$

Let $\sigma_1, \ldots, \sigma_r$ be all real embeddings and $\sigma_{r+1}, \ldots, \sigma_{r+s}$ be all complex embeddings. Then we apply the trace map to $\tau(u)_\sigma$ as follows:

$$\begin{aligned} Tr\left(\tau(u)_\sigma\right) &= \log\left(\frac{|u_{\sigma_1}|}{\left|N_{K/\mathbb{Q}}(u)^{\frac{1}{n}}\right|}\right) + \ldots + \log\left(\frac{|u_{\sigma_r}|}{\left|N_{K/\mathbb{Q}}(u)^{\frac{1}{n}}\right|}\right) \\ &+ \log\left(\frac{|u_{\sigma_{r+1}}|^2}{\left|N_{K/\mathbb{Q}}(u)^{\frac{1}{n}}\right|}\right) + \ldots + \log\left(\frac{|u_{\sigma_{r+s}}|^2}{\left|N_{K/\mathbb{Q}}(u)^{\frac{1}{n}}\right|}\right) \\ &= \log\left(\frac{|u_{\sigma_1}| \cdot \ldots \cdot |u_{\sigma_r}| \cdot |u_{\sigma_{r+1}}|^2 \cdot \ldots \cdot |u_{\sigma_{r+s}}|^2}{\left|N_{K/\mathbb{Q}}(u)\right|}\right) \\ &= \log\left(\frac{|u_{\sigma_1}| \cdot \ldots \cdot |u_{\sigma_r}| \cdot |u_{\sigma_{r+1}}| \cdot |u_{\bar{\sigma}_{r+1}}| \cdot \ldots \cdot |u_{\sigma_{r+s}}| \cdot |u_{\bar{\sigma}_{r+s}}|}{\left|N_{K/\mathbb{Q}}(u)\right|}\right) \\ &= \log\left(\frac{\left|N_{K/\mathbb{Q}}(u)\right|}{\left|N_{K/\mathbb{Q}}(u)\right|}\right) = 0 \end{aligned}$$

by Proposition 3.3.

As we can see from the above, due to the addition of $\frac{1}{n}\log|N_{K/\mathbb{Q}}(u)|$, $\tau(u)$ will be contained in the trace-zero hyperplane $H = \{x \in [\prod_\tau \mathbb{R}]^+ \Big| \ Tr(x) = 0\}$.

We can recall from Theorem (7.3) in [13]: there is a map $\lambda : \mathcal{O}_K^{\times} \longrightarrow H$ whose image forms a complete lattice in H. Moreover, Conrad proved in [6, p.7] that this property holds for any order in K, which means $\tau(\mathcal{O}^{\times})$ is a complete lattice in H as well.

Therefore, if a is a generator of an $\mathcal{O}$-module $\mathfrak{a}$, there will be a $u \in \mathcal{O}^{\times}$ such that $\tau(a) + \tau(u)$ lies in the fundamental parallelopiped of a complete lattice $\tau(\mathcal{O}^{\times})$.

Since $\tau(a) + \tau(u) = \tau(au)$, if we let $b = au$, then b will also be a generator of $\mathfrak{a}$ and $\tau(b)$ will lie in the fundamental parallelopiped of $\tau(\mathcal{O}^{\times})$ as well.

Let $\{u_1, u_2, \ldots, u_n\}$ be a basis of $\mathcal{O}^{\times}/\{\pm 1\}$, then for each embedding σ, we have

$$\begin{cases} \Big|\log|b_\sigma| - \frac{1}{n}\log\Big|N_{K/\mathbb{Q}}(b)\Big|\Big| \leq \frac{1}{2}\sum_{i=1}^{n-1}\Big|\log|u_{i,\sigma}|\Big|, & \text{if } \sigma \text{ is a real embedding} \\ \Big|\log|b_\sigma|^2 - \frac{1}{n}\log\Big|N_{K/\mathbb{Q}}(b)\Big|\Big| \leq \frac{1}{2}\sum_{i=1}^{n-1}\Big|\log|u_{i,\sigma}|\Big|, & \text{if } \sigma \text{ is a complex embedding.} \end{cases} \tag{5.15}$$

By the logarithmic properties, (5.15) becomes

$$\begin{cases} \left|\log\left(\frac{|b_\sigma|}{\left|N_{K|\mathbb{Q}}(b)^{\frac{1}{n}}\right|}\right)\right| \leq \Big|\log\Big(\prod_{i=1}^{n-1}|u_{i,\sigma}|\Big)^{\frac{1}{2}}\Big|, & \text{if } \sigma \text{ is a real embedding} \\ \left|\log\left(\frac{|b_\sigma|^2}{\left|N_{K|\mathbb{Q}}(b)^{\frac{1}{n}}\right|}\right)\right| \leq \Big|\log\Big(\prod_{i=1}^{n-1}|u_{i,\sigma}|\Big)^{\frac{1}{2}}\Big|, & \text{if } \sigma \text{ is a complex embedding.} \end{cases} \tag{5.16}$$

Take exponential function e^x on both sides of (5.16) and let $U_\sigma = \prod_{i=1}^{n-1} e^{|\log|u_{i,\sigma}||}/2$,

$$\begin{cases} \frac{1}{U_\sigma}\Big|N_{K/\mathbb{Q}}(b)^{\frac{1}{n}}\Big| \leq |b_\sigma| \leq U_\sigma\Big|N_{K/\mathbb{Q}}(b)^{\frac{1}{n}}\Big|, & \text{if } \sigma \text{ is a real embedding} \\ \sqrt{\frac{1}{U_\sigma}N_{K/\mathbb{Q}}(b)^{\frac{1}{n}}} \leq |b_\sigma| \leq \sqrt{U_\sigma\Big|N_{K/\mathbb{Q}}(b)^{\frac{1}{n}}\Big|}, & \text{if } \sigma \text{ is a complex embedding.} \end{cases}$$

Since b is a generator of an $\mathcal{O}$-module $\mathfrak{a}$, by Lemma 3.38, we have that $|N_{K|\mathbb{Q}}(b)| = |N_{\mathcal{O}}(\mathfrak{a})|$. So if $U = \max(U_\sigma)$, then the above becomes

$$\begin{cases} \frac{1}{U}|N_{\mathcal{O}}(\mathfrak{a})|^{\frac{1}{n}} \leq |b_\sigma| \leq U|N_{\mathcal{O}}(\mathfrak{a})|^{\frac{1}{n}}, & \text{if } \sigma \text{ is a real embedding} \\ \sqrt{\frac{1}{U}|N_{\mathcal{O}}(\mathfrak{a})|^{\frac{1}{n}}} \leq |b_\sigma| \leq \sqrt{U|N_{\mathcal{O}}(\mathfrak{a})|^{\frac{1}{n}}}, & \text{if } \sigma \text{ is a complex embedding.} \end{cases} \tag{5.17}$$

From (5.17), we can see $|b_\sigma| \in O(|N_{\mathcal{O}}(\mathfrak{a})|^{\frac{1}{n}})$ for all embeddings and if σ is a complex one, we have the better bound: $|b_\sigma| \in O(|N_{\mathcal{O}}(\mathfrak{a})|^{\frac{1}{2n}})$. □

Theorem 5.8. *Let $f \in \mathbb{Z}[x,y]$ be an homogeneous polynomial of even degree $n \geq 2$ which is irreducible in $\mathbb{Q}[x,y]$. Then there are only finitely many pairs $(J,\gamma) \in S_f$ satisfying conditions 1-3 of Theorem 5.6, up to multiplication by a unit of R_f.*

Proof. The quantity $|N_f(J)|$ is a positive integer and hence by condition 1, there are only finitely many possibilities for $|N_f(J)|$. Since $J \cong \mathbb{Z}^n$ is contained in $R_f \cong \mathbb{Z}^n$ and both are free $\mathbb{Z}$-modules of rank n, J is of the form $d_1\mathbb{Z} \oplus \ldots \oplus d_n\mathbb{Z} \subseteq \mathbb{Z}^n \cong R_f$ where $d_i | d_{i+1}$ and the d_i are positive integers. It follows that $J \supseteq d_n\mathbb{Z}^n$, where d_n divides $|N_f(J)| = d_1 \cdots d_n$. For each choice of d_n, there are only finitely many choices of $J \supseteq d_n\mathbb{Z}^n$.

Note if $(J,\gamma) \in S_f$ and satisfies conditions 1-3 of Theorem 5.6, then for any unit $u \in R_f^\times$, $(Ju, \gamma u)$ also lies in S_f and satisfies conditions 1-3 of Theorem 5.6. By Lemma 5.7, there exists a unit $u \in R_f^\times$ such that $(J',\gamma') = (Ju,\gamma u)$ and (J',γ') lies in S_f with conditions 1-3 of Theorem 5.6 being satisfied, and $|\gamma'_\sigma| \in O(M^{*\frac{1}{n}})$, where $M^* = \max(M,1)$.

Now, $j(\gamma')$ lies in the discrete set $\Gamma = j(\frac{1}{c}R_f)$ and the compact set

$$\Gamma = \left\{ x \in K_{\mathbb{R}} : |x| \leq O(M^{\frac{1}{n}}) \right\}.$$

Hence, there are finitely many choices for $j(\gamma')$, and thus for γ'. □

Bibliography

[1] Manjul Bhargava. Higher composition laws. I. A new view on Gauss composition, and quadratic generalizations. *Ann. of Math. (2)*, 159(1):217–250, 2004.

[2] Manjul Bhargava. Higher composition laws. II. On cubic analogues of Gauss composition. *Ann. of Math. (2)*, 159(2):865–886, 2004.

[3] Manjul Bhargava. Most hyperelliptic curves over $\mathbb{Q}$ have no rational points. *arXiv preprint arXiv:1308.0395*, 2013.

[4] Nils Bruin and Michael Stoll. Two-cover descent on hyperelliptic curves. *Math. Comp.*, 78(268):2347–2370, 2009.

[5] P. M. Cohn. *Algebra, Vol. 1.* John Wiley & Sons, London-New York-Sydney, 1974.

[6] Keith Conrad. Dirichlet's unit theorem. `http://www.math.uconn.edu/~kconrad/blurbs/gradnumthy/unittheorem.pdf`.

[7] Keith Conrad. Modules over a PID. `http://www.math.uconn.edu/~kconrad/blurbs/linmultialg/modulesoverPID.pdf`.

[8] Kenneth Hardy and Kenneth S. Williams. The class number of pairs of positive-definite binary quadratic forms. *Acta Arith.*, 52(2):103–117, 1989.

[9] Toshitsune Miyake. *Modular forms*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, english edition, 2006. Translated from the 1976 Japanese original by Yoshitaka Maeda.

[10] Jorge Morales. The classification of pairs of binary quadratic forms. *Acta Arith.*, 59(2):105–121, 1991.

[11] Jorge Morales. On some invariants for systems of quadratic forms over the integers. *J. Reine Angew. Math.*, 426:107–116, 1992.

[12] Jin Nakagawa. Binary forms and orders of algebraic number fields. *Invent. Math.*, 97(2):219–235, 1989.

[13] Jürgen Neukirch. *Algebraic number theory*, volume 322 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999. Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder.

[14] Gerald W. Schwarz. Representations of simple Lie groups with regular rings of invariants. *Invent. Math.*, 49(2):1–12, 1978.

[15] Melanie Matchett Wood. Rings and ideals parameterized by binary n-ic forms. *J. Lond. Math. Soc. (2)*, 83(1):208–231, 2011.

[16] Melanie Matchett Wood. Parametrization of ideal classes in rings associated to binary forms. *J. Reine Angew. Math.*, 689:169–199, 2014.