ABC CONJECTURE AND ITS IMPLICATIONS

by

Tuğçe Koç

B.S., Mathematics, Boğaziçi University, 2016

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Mathematics

Boğaziçi University

2019

ACKNOWLEDGEMENTS

First of all, I would like to thank my advisor, Ekin Özman, for the continuous support throughout my studies, for her patience, immense knowledge, and advice . Her guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor.

I, especially, want to thank my best friend Ezgi Kan for her endless energy, encouragement, and for the days she had been dragged to many places for me to study, and, of course, for the tech tips. I would also like to thank my great friend Merve Taşan for her positivity and moral support during the hard times. I could not have done it if you two were not in my life.

I am deeply thankful for my lovely family, Hilal, Kadir and Burak, for always being there for me. This thesis stands as a testament to your unconditional love and encouragement.

My great friend Ayçin İplikçi deserves special thanks for her help and colleagueship.

I thankfully acknowledge the financial support given by TÜBİTAK PIA Bosphorus project 117F274 “Sonlu Cisimler Üzerinde Tanımlı Eğriler, Jakobyen Varyeteleri Ve Abelyen Varyeteler” (PI Alp Bassa). Finally, I would like to thank İlhan İkeda and Ayberk Zeytin for serving on my thesis defense committee.

ABSTRACT

ABC CONJECTURE AND ITS IMPLICATIONS

In this thesis, our aim is to state the importance of *abc conjecture* and prove the strong results we obtain with the help of *abc conjecture*. First, we give necessary notions and tools which are used throughout the thesis. Then we introduce *Hall conjecture*, *Fermat's last theorem* and *Mordell conjecture*, and their relations with *abc conjecture*. In particular, we give the effective proof of *Mordell conjecture* using *abc conjecture*, given in the article of Noam Elkies, [1], and also get another height bound by combining with a different theorem. Finally, we give three examples where we use both of the height bounds. This thesis was supported by TÜBİTAK Project 117F274.

ÖZET

ABC VARSAYIMI VE ONUN ÇIKARIMLARI

Bu tezde, amacımız abc varsayımının önemini anlatmak ve onu kullanarak birçok güçlü ifadeyi kanıtlayabileceğimizi göstermektir. İlk olarak, tez boyunca kullanacağımız bazı kavramlar ve araçlardan bahsediyoruz. Daha sonra Hall varsayımı, Fermat'ın son teoremi, Mordell varsayımı ifadelerini tanıtıp abc varsayımından çıkan ispatlarını veriyoruz. Özellikle, Noam Elkies'in makalesinde kanıtladığı [1], abc varsayımından çıkan Mordell varsayımının efektif kanıtı üzerinde duruyoruz. Aynı zamanda, bu kanıtı başka bir teoremle birlikte kullanarak farklı bir yükseklik sınırı elde ediyoruz. Sonunda da bu iki farklı yükseklik sınırını kullanarak örnekler veriyoruz. Bu tez TÜBİTAK 117F274 projesi tarafından desteklenmiştir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SYMBOLS

$\mathbb{A}^n$	affine n space
$\mathbb{A}^n(K)$	the set of K rational points of $\mathbb{A}^n$
$b_f(x)$	branch number of f in $x \in C$
C	algebraic curve
C_f	$= \overline{K}[x,y]/\langle f \rangle$, ring of continuous functions on the curve $Z_f(\overline{K})$
$C(K)$	the set of K rational points of C
$C(\mathbb{Q})$	the points on C with rational coordinates
$C(Z_f(\overline{K}), \overline{K})$	the set of continuous functions from $Z_f(\overline{K})$ to $\overline{K}$
D	discriminant of a polynomial
$\|D\|$	complete linear system of the divisor D
$D_1 \geq D_2$	indication that $D_1 - D_2$ is a positive divisor
$deg(D)$	degree of a divisor
$deg(\phi)$	degree of the map ϕ
$disc_{\mathbb{Z}}(\mathcal{O}_K)$	discriminant of the ring of integers
$Div(C)$	divisor group of a curve
$Div^0(C)$	group of divisors of degree 0 on a curve
$Div_K(C)$	group of divisors defined over K
$div(f)$	divisor of the function f
$div_0(f)$	divisor of zeros of f
$div_\infty(f)$	divisor of poles of f
$div(w)$	divisor of the differential form w
$e_\phi(P)$	ramification index of the map ϕ
$\hat{E}$	reduction of the elliptic curve E modulo p
$E(K)$	group of K-rational points on the elliptic curve E/K
E_{tors}	torsion subgroup of the elliptic curve E
g	genus
G_K	the absolute Galois group $Gal(\overline{K}/K)$
$G_{\mathbb{Q}}$	Galois group on $\mathbb{P}^n(\overline{\mathbb{Q}})$

$Gal(\overline{K}/K)$	the Galois group of $\overline{K}/K$
h	absolute logarithmic height
H	absolute multiplicative height
h_f	height relative to a morphism f
$H(P)$	height of a point in $\mathbb{P}^n(\mathbb{Q})$
$H_K(P)$	relative height of a point in $\mathbb{P}^n(\mathbb{Q})$
$h_{V,D}$	height on the variety V relative to the divisor D
$H(\alpha)$	abbreviation for $H((x,1))$
K	number field
K^*	the unit group of K
$\overline{K}$	an algebraic closure of K
K_C	canonical divisor on a curve C
$K(C)$	function field of C/K
$\overline{K}[C]_P$	local ring of C at P
$K(P)$	minimal field of definition of the point P
$\overline{K}[X]$	the polynomial ring $\overline{K}[X_1,\ldots,X_n]$
$\overline{K}(Z_f)$	the field of fractions of the ring C_f
$\mathcal{L}(D)$	space of functions satisfying $div(f) \geq -D$
$l(D)$	dimension of the space of functions $\mathcal{L}(D)$
$L(E,s)$	L-function of E
$L_p(T)$	local factor at p of the L-series
M_K	a complete set of inequivalent absolute values on K
${M_K}^\infty$	the archimedean absolute values in K
${M_K}^0$	the nonarchimedean absolute values in K
M_P	ideal associated to the point P
$Max(C_f)$	the set of maximal ideals of C_f
$\mathbb{N}$	norm
$n_0(f)$	number of distinct roots of f
N_p	number of points in the reduction of the curve modulo p
O	the identity element on an elliptic curve
$O(1)$	a bounded function

$\mathcal{O}_K$	the ring of integers of K
ord_p	the p-adic order of a rational number
$ord_{\mathcal{P}}$	valuation attached to the prime ideal
ord_P	normalized valuation on $\overline{K}[C]_P$ and on $\overline{K}(C)$
p	rational prime
$\mathcal{P}$	prime ideal
$\mathcal{P}_P$	prime rational divisor
P^σ	the action of $\sigma \in Gal(\overline{K}/K)$ on the point P
$\mathbb{P}^n$	projective n space
$\mathbb{P}^n(K)$	set of K rational points of $\mathbb{P}^n$
$Pic(C)$	the Picard group of C
$Pic^0(C)$	Picard group of degree-zero divisor classes
R_E	rank of an elliptic curve E
$Rad(n)$	product of distinct prime factors of n
$\text{rank}_{an}(E/\mathbb{Q})$	analytic rank of $E/\mathbb{Q}$
$Spec(A)$	the set of prime ideals of a ring A
$Supp(D)$	support of the divisor D
V	projective variety
$v(x)$	$= -\log \lvert x\rvert_v$, for an absolute value $v \in M_K$
w	invariant differential on an elliptic curve
$\lvert x\rvert_p$	p-adic absolute value
$\lvert x\rvert_{\mathcal{P}}$	$\mathcal{P}$-adic absolute value associated to a prime ideal
$\lvert x\rvert_\phi$	absolute value associated to an embedding in $\mathbb{C}$
$\lvert x\rvert_\infty$	archimedean absolute values
$Z_f(K)$	the set of points of f with coordinates in K
Δ_E	discriminant of an elliptic curve E
Δ'_E	minimal discriminant of an elliptic curve E
ϕ^*	map of function fields induced by rational map of curves
ϕ_*	map of function fields induced by rational map of curves
ϕ^*	map induced on divisors by the rational map ϕ

ϕ_D	rational map associated to the divisor D
$\phi_{D,F}$	rational map associated to the divisor D and a basis F of $\mathcal{L}(D)$
Ω_C	the space of meromorphic differential forms on the curve C
$\oplus$	group law on an elliptic curve
$\sim$	linear equivalence of divisors

1. INTRODUCTION

In any sum of three relatively prime integers, it is not possible for all three terms to be divisible by many high powers of primes. That is the basic intuition behind the famous *abc conjecture* in number theory. It seems like a simple conjecture about the sum of integers, yet nobody has been able to prove it. Goldfeld described *abc conjecture* as "the most important unsolved problem in Diophantine analysis." [2] It is first proposed by Joseph Osterlee and David Masser in the 1980s.

The importance of *abc conjecture* mainly stems from two reasons. First, the conjecture investigates the relation between two arithmetic operations, addition and multiplication. Divisibility and primes are multiplicative notions. *Abc conjecture* sets a subtle limit on how these notions can interact with addition. Since number theory is mostly built over the relation between addition and multiplication, it is useful to understand such constraints. Secondly, *abc conjecture* has really important consequences such as *"eventual" Fermat's last theorem, Catalan's conjecture for any large parameters, Roth's theorem, Mordell conjecture* and *Hall conjecture.*

As it is not provable till the current day, it is also not disprovable. The exponent $1 + \epsilon$ makes the conjecture very powerful, and thus hard to find a counterexample. *Abc conjecture* got some media attention in 2012 when Professor Shinichi Mochizuki published a possible proof for the conjecture, which is 500 pages long based on perhaps another 1000 pages of previous technical papers. After 7 years of struggle, almost no one could completely understand his proof let alone approve, and its reason is that Mochizuki developed a whole new type of mathematics called Inter-universal Teichmüller theory. So mathematicians who want to read his proof first need to understand his new theory. Workshops and conferences were held on Mochizuki's work on *abc conjecture.* Brian Conrad, who is a math professor at Stanford and was one of the participants at the Oxford workshop, summarizes the issue on his notes from the workshop as: "The method as currently formulated by Mochizuki does not yield explicit constants, so it cannot be used to establish an effective proof of the Mordell

Conjecture. But if correct it would nonetheless be a tremendous achievement, setting many difficult open problems, and would yield a new proof of the Mordell conjecture (as shown long ago by Noam Elkies)." [3]

In this thesis, we first give two equivalent statements of *abc conjecture* and a few examples. Then, we state and give the proof of *Mason-Stothers theorem*, which is the inspiration for *abc conjecture*, and an analogy of it in polynomials, hence a reason for us to believe the truth of *abc conjecture*. *Mason-Stothers theorem* puts a bound on the degrees of polynomials using the addition relation between them. The idea of changing "polynomial" with "integer" led Masser and Osterle to *abc conjecture.* Indeed this is just the tip of the iceberg, they developed the conjecture with deep considerations of algebraic geometry and the theory of modular functions.

Later, as another important result of *abc conjecture*, we introduce *Hall conjecture* and give the proof of *Weak Hall conjecture* using *abc conjecture*. *Hall conjecture* puts a bound on the size of the integral points on a Mordell curve, in addition this bound simply depends on a coefficient of the defining equation of the curve.

Finally, we give the proofs for the most famous applications of *abc conjecture*, which are *Fermat's last theorem* for large exponents and *Mordell conjecture*. Even though *Fermat's last theorem* was proved by Wiles, proving it from *abc conjecture* provides a much simpler proof. The same is also true for *Mordell conjecture*, now called *Faltings' theorem.* Furthermore, the proof of Faltings is not effective, in other words it proves that there are finitely many rational points on a curve of genus > 1, but does not give a method for finding them or give an upper bound on the size of the points. On the other hand, the proof using *abc conjecture*, given by Noam Elkies [1], provides an effective upper bound on the size of the points, hence limits the search criteria for finding the points. Moreover, the proof could be modified to get *Siegel's theorem* which states that an algebraic affine curve of genus ≥ 1, or of genus zero with 3 points at infinity and defined over a number field K has only finitely many integral points. But it is not exciting as *Mordell's*, since effective methods for this case are already available.

2. PRELIMINARIES

We have a long Preliminaries section, including many tools and theorems, which will be helpful for us to understand the proofs and examples in the following sections.

2.1. Absolute Values

In this section, we define the notion of an absolute value (or multiplicative valuation) and give various examples. The majority of the content is taken from Algebraic Number Theory Notes by Milne [4], in which one can find more details and the proofs.

Definition 2.1. *An absolute value or multiplicative valuation on a field K is a real-valued function $K \to \mathbb{R}$ with $x \mapsto |x|_v$ such that*

(i) $|x|_v \geq 0$ *with equality if and only if* $x = 0$
(ii) $|xy|_v = |x|_v|y|_v$
(iii) *there is a* $C_v \in \mathbb{R}$ *such that* $|x+y|_v \leq C_v max(|x|_v, |y|_v)$ *for all* $x, y \in K$.

Note that $C_v \geq 1$, which can be easily seen by taking $y = 0$ in (iii). If we can take $C_v = 1$, then $|\,|_v$ is called *nonarchimedean*; otherwise *archimedean.*

By the first two conditions, we see that $|\,|_v : K^* \to \mathbb{R}_{>0}$ is a homomorphism where both of them are multiplicative groups. The valuation homomorphism $|\,|_v$ sends all roots of unity in K^* to 1, and $|-x|_v = |x|_v$ for all $x \in K^*$. We say $|\,|_v$ is *discrete* when $|K^*|_v$ is a discrete subgroup of $\mathbb{R}_{>0}$.

A discrete (additive) valuation *ord* on K determines an absolute value by $|x|_v = e^{-ord(x)}$, any $e > 1$. When we feel like additive valuations, we can pass from multiplicative by taking logs gives $\log_e |x|_v = -ord(x)$, or $ord(x) = -\log_e |x|_v$.

Definition 2.2. *We define the trivial absolute value as* $|a|_v = 1$ *for all* $a \in K^*$.

Proposition 2.3. *Let $|\,|_v$ be a nontrivial nonarchimedean absolute value, and put $v(x) = -\log|x|_v$, $x \neq 0$ (log to base e for any real $e > 1$). Then $v : K^* \mapsto \mathbb{R}$ satisfies the following conditions:*

- $v(xy) = v(x) + v(y)$,
- $v(x+y) \geq min\{v(x), v(y)\}$.

If $v(K^*)$ is discrete in $\mathbb{R}$, then v is a multiple of a discrete valuation $ord : K^* \hookrightarrow \mathbb{Z} \subset \mathbb{R}$.

Definition 2.4. *Let K be a field. A discrete valuation on K is a nonzero homomorphism $v : K^* \mapsto \mathbb{Z}$ such that $v(a+b) \geq min(v(a), v(b))$. The image of v is of the form $m\mathbb{Z}$ for some $m \in \mathbb{Z}$. If $m = 1$, then $v : K^* \mapsto \mathbb{Z}$ is surjective, and v is said to be normalized.*

Example 2.5. *Let K be an arbitrary number field, then any (real or complex) embedding gives rise to an archimedean absolute value with complex-conjugate embeddings yielding the same absolute value because $|a+bi| = |a-bi|$ in $\mathbb{C}$. Moreover, the only way for two archimedean embeddings to define the same absolute value is when they come from a pair of complex-conjugate embeddings. Let r_1 denote the number of real embeddings of K, and r_2 denote the number of pairs of complex-conjugate embeddings of K. Then K has $r_1 + 2r_2$ many embeddings and $r_1 + r_2$ many archimedean absolute values up to equivalence.*

For a real embedding $\phi : K \hookrightarrow \mathbb{R}$, $|a|_\phi := |\phi(a)|_\infty$, with $|\,|_\infty$ usual real absolute value.

For a pair of complex-conjugate embeddings $(\phi_1, \phi_2) : K \hookrightarrow \mathbb{C}$, $|a|_{(\phi_1,\phi_2)} = |\phi_1(a).\phi_2(a)|_\infty$.

A concrete example: $K = \mathbb{Q}(\sqrt{3})$ has two real embeddings $\phi_1 : a+b\sqrt{3} \mapsto a+b\sqrt{3}$ and $\phi_2 : a+b\sqrt{3} \mapsto a-b\sqrt{3}$, hence has two archimedean absolute values. So for $2+5\sqrt{3}$, we get $|2+5\sqrt{3}|_{\phi_1} = |\phi_1(2+5\sqrt{3})|_\infty = 2+5\sqrt{3}$ and $|2+5\sqrt{3}|_{\phi_2} = |\phi_2(2+5\sqrt{3})|_\infty = 5\sqrt{3}-2$.

Another concrete example: $K = \mathbb{Q}(i)$ has two complex embeddings, which are complex-conjugate, $\phi_1 : a + bi \mapsto a + bi$ and $\phi_2 : a + bi \mapsto a - bi$, hence K has only one archimedean absolute value. Let $\phi = (\phi_1, \phi_2)$. For $3 + 5i$, we get $|3 + 5i|_\phi = (3 + 5i)(3 - 5i) = 34$.

Example 2.6. *With an additive discrete valuation* $ord : K^* \to \mathbb{Z}$ *and a real number* $e > 1$, *we define a nonarchimedean absolute value for nonzero* $a \in K : |a| = (1/e)^{ord(a)}$, *and set* $|0| = 0$. *On the field of rational numbers* $\mathbb{Q}$, *for every prime* p, *we get the p-adic absolute value* $|\ |_p : |a|_p = (1/e)^{ord_p(a)}$ *with* $e > 1$ *a real number. We take the base* $e = p$ *to normalize. Hence* $|a|_p = (1/p)^{ord_p(a)} = 1/p^r$ *where* $a = a_0.p^r$ *with* $ord_p(a_0) = 0$.

A concrete example for $\mathbb{Q}$: $\left|\frac{3}{5}\right|_3 = \left(\frac{1}{3}\right)^{ord_3(3/5)} = \frac{1}{3}$ and $\left|\frac{3}{5}\right|_5 = \left(\frac{1}{5}\right)^{ord_5(3/5)} = 5$.

Let $\mathcal{P}$ be a nonzero prime ideal in $\mathcal{O}_K$. If $\alpha \in K^*$ and $ord_\mathcal{P}(\alpha) \geq 0$, then $\alpha = x/y$ with $x, y \in \mathcal{O}_K$ and $ord_\mathcal{P}(y) = 0$. Similarly, on a number field K, we get a *normalized p-adic absolute value* for every nonzero prime ideal $\mathcal{P}$ by setting the base as the norm of the prime ideal. We denote the norm by $\mathbb{N}$. Then we have:

For $\alpha \in K^*$ and for a nonzero prime $\mathcal{P}$, $|a|_\mathcal{P} = (1/\mathbb{N}(\mathcal{P}))^{ord_\mathcal{P}(\alpha)}$ with $(\alpha) = \alpha\mathcal{O}_K = (P)^{ord_\mathcal{P}(\alpha)}.\frac{\mathcal{I}}{\mathcal{J}}$ and $\mathcal{P} \nmid \mathcal{I}\mathcal{J}$.

Fact 2.7. *K number field,* $a \in K^*$. *The norm of a principal fractional ideal* $a\mathcal{O}_K$ *is given by* $\mathbb{N}(a\mathcal{O}_K) = |\mathbb{N}_{K/\mathbb{Q}}(a)|$.

A concrete example for K : Let $K = \mathbb{Q}(\sqrt{2})$. The ideal (2) factorizes into prime ideals in the ring of integers $\mathbb{Z}[\sqrt{2}]$ as $2\mathbb{Z}[\sqrt{2}] = (2, \sqrt{2})^2$. Let $\mathcal{P} = (2, \sqrt{2})$. Then $|2|_\mathcal{P} = (1/\mathbb{N}_\mathcal{P})^2 = \left(\frac{1}{2}\right)^2 = \frac{1}{4}$.

From an absolute value, one can construct a metric on K with $d(a, b) = |a - b|$, and hence a topology on K.

Proposition 2.8 ([4], Chapter 7, Proposition 7.8)**.** *Let* $|\,|_v$, $|\,|_w$ *be nontrivial absolute values on* K*. Then the followings are equivalent:*

(i) $|\ |_v$ *and* $|\ |_w$ *define the same topology on* K*;*
(ii) If $|\ |_v < 1$*, then* $|\ |_w < 1$*;*
(iii) $|\ |_w = (|\ |_v)^a$ *for some* $a > 0$*.*

In this case $|\,|_v$ *and* $|\,|_w$ *are said to be equivalent.*

There is a classification theorem on valuations on the field of rational numbers:

Theorem 2.9 (Ostrowski Theorem, [5])**.** *Let* $|\,|_v$ *be a nontrivial absolute value on* $\mathbb{Q}$*.*

- *If* $|\,|_v$ *is archimedean, then* $|\,|_v$ *is equivalent to* $|\,|_\infty$*.*
- *If* $|\,|_v$ *is nonarchimedean, then* $|\,|_v$ *is equivalent to* $|\,|_p$ *for exactly one prime* p*.*

Note that $|\,|_\infty$ is the usual real absolute value, and we say that $|\,|_\infty$ is *normalized.*

Theorem 2.10 (Product Formula for $\mathbb{Q}$)**.** *For* $p = 2, 3, 5, 7, \ldots, \infty$*, let* $|\,|_p$ *be the corresponding normalized absolute value on* $\mathbb{Q}$*. For any nonzero rational number* a*,*

$$\prod_{p \in M_{\mathbb{Q}}} |a|_p = 1.$$

This is a simple reflection of the fact that $\mathbb{Z}$ has unique factorization.

There is also a classification theorem of valuations for number fields:

Theorem 2.11 (Ostrowski Theorem for K, [5])**.** *Every nontrivial absolute value on* K *is equivalent to a* p*-adic absolute value for a unique prime ideal* $\mathcal{P}$ *in* $\mathcal{O}_K$ *or is equivalent to an archimedean absolute value coming from a real or complex-conjugate pair of embeddings.*

Theorem 2.12 (Product Formula for K)**.** *Let $a \in K^*$. Then $\prod |a|_v = 1$, where v ranges over all normalized valuations of K.*

Example 2.13. *Let $K = \mathbb{Q}(\sqrt{5})$ and $a = 3$. Then K has two real embeddings $\phi_1 : a + b\sqrt{5} \mapsto a + b\sqrt{5}$ and $\phi_2 : a + b\sqrt{5} \mapsto a - b\sqrt{5}$, hence has two archimedean absolute values. For $a = 3$, we have $|3|_{\phi_1} = |3|_{\phi_2} = 3$. Moreover, the element $3 \in K$ forms the principal ideal $3\mathcal{O}_K$ where it has the factorization $3\mathcal{O}_K = \mathcal{P}$ for some prime ideal $\mathcal{P}$ with norm 9. Then $|3|_{\mathcal{P}} = (1/9)$. Therefore multiplying over all absolute values for the element 3, we obtain 1.*

2.2. Prime Ideals and Factorization

In the following four subsections, we introduce some key notions from algebraic number theory such as plane curves, factorization of ideals, and ramified primes. It is a summary of three chapters from An Invitation to Arithmetic Geometry by Dino Lorenzini. For more details and proofs, see [6].

2.2.1. Introduction

Let K be a field and $\overline{K}$ be an algebraic closure of K. In the first subsection of the sequence, we explain that a nonsingular plane curve, defined by an irreducible polynomial $f(x, y)$ in $K[x, y]$, monic in y, corresponds to the integral closure of $K[x]$ in the function field of the given curve.

Let $f(x, y) \in K[x, y]$ be a polynomial in two variables with coefficients in K and let d be the degree of f. Let

$$Z_f(K) := \{(a, b) \in K \times K | f(a, b) = 0\}.$$

Definition 2.14. *The set $Z_f(\overline{K})$ is called an affine plane curve. The set $Z_f(K)$ is the set of points with coordinates in K of the affine plane curve of degree d defined by $f(x, y)$.*

Lemma 2.15. *Let $f(x,y) \in K[x,y]$ be a polynomial of degree $d > 0$. Then $Z_f(\overline{K})$ is an infinite set. In particular, $Z_f(\overline{K})$ is not empty.*

We assume that $Z_f(\overline{K})$ and $\overline{K}$ are endowed with Zariski topology, which will allow us to view the ring $C_f := \overline{K}[x,y]/\langle f\rangle$ as a ring of continuous functions on the curve $Z_f(\overline{K})$.

Let $C(Z_f(\overline{K}), \overline{K})$ denote the set of continuous functions from $Z_f(\overline{K})$ to $\overline{K}$. The map of sets $i_f : C_f \to C(Z_f(\overline{K}), \overline{K})$ is injective. We call C_f *the ring of algebraic functions on* $Z_f(\overline{K})$, or simply *the ring of functions of* $Z_f(\overline{K})$.

Definition 2.16. *Let $f \in \overline{K}[x,y]$ be an irreducible polynomial, so that the ring $C_f := \overline{K}[x,y]/\langle f\rangle$ is an integral domain. We denote by $\overline{K}(Z_f)$ the field of fractions of the ring C_f and call it the field of rational functions of the affine curve defined by f. The elements of $\overline{K}(Z_f)$ are called the rational functions of $Z_f(\overline{K})$.*

Definition 2.17. *Let $f(x,y) \in K[x,y]$. The degree of f in y, $deg_y(f)$, is the degree of the polynomial f viewed as a polynomial in the variable y with coefficients in $K[x]$.*

Let $f(x,y) = a_n(x)y^n + \cdots + a_0(x)$ be an irreducible polynomial in $\overline{K}[x,y]$. The natural map

$$\phi : \overline{K}[x] \to C_f := \overline{K}[x,y]/\langle f\rangle$$

is injective if $f(x,y) \neq cx + d$, for all $c, d \in \overline{K}$. In this case, it can be extended to an injection $\overline{K}(x) \to \overline{K}(Z_f)$ with $g(x)/h(x) \mapsto \phi(g(x))/\phi(h(x))$ of the field of fractions $\overline{K}(x)$ of $\overline{K}[x]$ into the field of fractions $\overline{K}(Z_f)$ of C_f. The field $\overline{K}(Z_f)$ is isomorphic to $\overline{K}(x)[y]/(f)$ and, therefore, the extension $\overline{K}(Z_f)/\overline{K}(x)$ is of finite degree $deg_y(f)$. We have obtained in this way a triple of $\overline{K}[x], \overline{K}(x), \overline{K}(Z_f)$. It is not always the case that C_f is the integral closure of $\overline{K}[x]$ in $\overline{K}(Z_f)$.

Fact 2.18. *Let A be a domain of dimension* 1. *Let B be a domain containing A and such that each element of B is integral over A. Then B has dimension* 1.

Fact 2.19. *Let $A \subseteq B$ be two rings. If A is noetherian and B is a finitely generated A-module, then B is noetherian.*

When $a_n(x) = 1$, every element of C_f is integral over $\overline{K}[x]$, and hence C_f is an integral extension of $\overline{K}[x]$. When $a_n(x) = 1$, the ring C_f has dimension 1 by Fact 2.18. Furthermore, since C_f is generated as a ring over $\overline{K}[x]$ by the classes of the elements $1, y, \ldots, y^{n-1}$, we conclude from Fact 2.19 that C_f is noetherian. Therefore, the domain C_f is integrally closed if and only if it is a Dedekind domain.

2.2.2. Points and Maximal Ideals

Here, we show the association between the set of points of the curve $Z_f(\overline{K})$ and the set of maximal ideals $Max(C_f)$ of the ring $C_f := \overline{K}[x, y]/(f)$.

Proposition 2.20. *Let M be a maximal ideal of $\overline{K}[x, y]$. Then there exists a point $(a, b) \in \mathbb{A}^2(\overline{K})$ such that M is generated by $(x - a)$ and $(y - b)$.*

Corollary 2.21. *Let $f \in \overline{K}[x, y]$ be an irreducible polynomial. An ideal M of the ring $C_f := \overline{K}[x, y]/(f)$ is maximal if and only if it can be generated by the images in C_f of two elements of the form $x - a$ and $y - b$ in $\overline{K}[x, y]$, with $f(a, b) = 0$. Let $I_f(a, b)$ denote the ideal of C_f generated by the images of $x - a$ and $y - b$ in C_f. The map $I_f : (a, b) \mapsto I_f(a, b)$ is a bijection from $Z_f(\overline{K})$ to $Max(C_f)$.*

2.2.3. Morphisms of Curves

We introduce the notion of morphism of curves and the relation between morphism of curves and the homomorphism of rings of functions.

Let $\phi : Z_f(\overline{K}) \to Z_g(\overline{K})$ be any map between two curves. This map ϕ uniquely determines two maps $\phi_1, \phi_2 : Z_f(\overline{K}) \to \overline{K}$ such that $\phi(a, b) := (\phi_1(a, b), \phi_2(a, b))$.

Definition 2.22. *A map $\phi : Z_f(\overline{K}) \to Z_g(\overline{K})$ between two plane curves is a morphism of affine plane curves if there exist two polynomials $\alpha(x, y)$ and $\beta(x, y)$ in $\overline{K}[x, y]$ such that $\phi_1(a, b) = \alpha_1(a, b)$ and $\phi_2(a, b) = \beta(a, b)$ for all $(a, b) \in Z_f(\overline{K})$.*

Lemma 2.23. *Let $f, g \in \overline{K}[x, y]$ be irreducible polynomials. Endow $Z_f(\overline{K})$ and $Z_g(\overline{K})$ with the Zariski topology. Let $\phi : Z_f(\overline{K}) \to Z_g(\overline{K})$ be a morphism of curves. Then ϕ is a continuous map.*

Any such continuous map $\phi : Z_f(\overline{K}) \to Z_g(\overline{K})$ defines a map:

$$\phi_C{}^* : C(Z_g(\overline{K}), \overline{K}) \to C(Z_f(\overline{K}), \overline{K})$$

with $h \mapsto h \circ \phi$. Consider the following diagram:

$$\begin{array}{ccc} C_g & & C_f \\ \downarrow{\scriptstyle i_g} & & \downarrow{\scriptstyle i_f} \\ C(Z_g(\overline{K}), \overline{K}) & \xrightarrow{\phi_C{}^*} & C(Z_f(\overline{K}), \overline{K}) \end{array} .$$

Lemma 2.24. *Let $\phi : Z_f(\overline{K}) \to Z_g(\overline{K})$ be a continuous map between two algebraic curves endowed with the Zariski topology. The map ϕ is a morphism of plane curves if and only if $(\phi_C{}^* \circ i_g)(C_g) \subseteq i_f(C_f)$.*

If ϕ is a morphism of curves, then the map $\phi^*_{C|_{i_g(C_g)}}$ can be used to define a homomorphism of $\overline{K}$-algebras between C_g and C_f, denoted by ϕ^*, in such a way that the following diagram is commutative:

$$\begin{array}{ccc} C_g & \xrightarrow{\phi^*} & C_f \\ \downarrow{\scriptstyle i_g} & & \downarrow{\scriptstyle i_f} \\ C(Z_g(\overline{K}), \overline{K}) & \xrightarrow{\phi_C{}^*} & C(Z_f(\overline{K}), \overline{K}) \end{array} .$$

Let $\phi^* : C_g \to C_f$ be any homomorphism of $\overline{K}-$algebras. The homomorphism ϕ^* defines in a natural way a morphism of curves $\phi : Z_f(\overline{K}) \to Z_g(\overline{K})$. Let $\phi_u(x, y) \in \overline{K}[x, y]$ be a polynomial such that the class of $\phi_u(x, y)$ in C_f is ϕ^*(class of u).

Similarly, let $\phi_v(x,y) \in \overline{K}[x,y]$ be a polynomial such that the class of $\phi_v(x,y)$ in C_f is ϕ^*(class of u). We set

$$\phi : Z_f(\overline{K}) \to Z_g(\overline{K})$$

with $(a,b) \mapsto (\phi_u(a,b), \phi_v(a,b))$. The map ϕ is well-defined and does not depend on the choices of $\phi_u(x,y)$ and $\phi_v(x,y)$.

Lemma 2.25. *Let $\phi^* : C_g \to C_f$ be any homomorphism of $\overline{K}-$algebras. Let $\phi : Z_f(\overline{K}) \to Z_g(\overline{K})$ be the induced morphism of curves described above. Let*

$$\phi_C{}^* : C(Z_g(\overline{K}), \overline{K}) \to C(Z_f(\overline{K}), \overline{K})$$

with $h \mapsto h \circ \phi$. The following diagram is commutative:

$$\begin{array}{ccc} C_g & \xrightarrow{\phi^*} & C_f \\ \downarrow{\scriptstyle i_g} & & \downarrow{\scriptstyle i_f} \\ C(Z_g(\overline{K}), \overline{K}) & \xrightarrow{\phi_C{}^*} & C(Z_f(\overline{K}), \overline{K}) \end{array} .$$

Definition 2.26. *The set of prime ideals of a ring A is called the spectrum of A and is denoted by $Spec(A)$. Given any ring homomorphism $\phi : A \to B$, we let $Spec(\phi)$ denote the natural map:*

$$Spec(\phi) : Spec(B) \to Spec(A)$$

with $P \mapsto \phi^{-1}(P)$.

Let $Z_f(\overline{K})$ be any plane curve. Recall that the map $I_f : Z_f(\overline{K}) \to Max(C_f)$, with $(a,b) \mapsto \langle x-a, y-b \rangle$, is bijective.

By following the set-up in "An Invitation to Arithmetic Geometry" by Dino Lorenzini, we denote the quotient as $\langle x-a, y-b \rangle$ by abuse of notation.

Lemma 2.27. *Let C_f and C_g be the rings of functions of two affine plane curves given by irreducible polynomials $f(x,y)$ and $g(u,v)$, respectively. Let $\phi^* : C_g \to C_f$ be a homomorphism of $\overline{K}-$algebras. Then the map $Spec(\phi^*)$ restricts to a map $\phi' : Max(C_f) \to Max(C_g)$, with $M \mapsto (\phi^*)^{-1}(M)$. Moreover, the following diagram is commutative:*

$$\begin{array}{ccc} Z_f(\overline{K}) & \xrightarrow{I_f} & Max(C_f) \\ \downarrow{\scriptstyle \phi} & & \downarrow{\scriptstyle \phi'} \\ Z_g(\overline{K}) & \xrightarrow{I_g} & Max(C_g) \end{array} .$$

Theorem 2.28. *Let $f(x,y) \in \overline{K}[x,y]$ be an irreducible polynomial. The ring $C_f := \overline{K}[x,y]/(f)$ is integrally closed if and only if the curve $Z_f(\overline{K})$ is nonsingular.*

Theorem 2.29. *Let $f \in \overline{K}[x,y]$ be an irreducible polynomial. Every nonzero prime ideal of $\overline{K}[x,y]/(f)$ is a maximal ideal generated by the images of $(x-a)$ and $(y-b)$ for some $(a,b) \in \overline{K} \times \overline{K}$ such that $f(a,b) = 0$.*

In conclusion of the section, for an irreducible polynomial $f \in \overline{K}[x,y]$, the set $Z_f(\overline{K})$ is in bijection with $Max(C_f)$.

2.2.4. Factorization of Ideals

In the final subsection of the sequence, we describe the factorization of ideals and ramification, then give several important field extensions such as Kummer extensions and Eisenstein extensions.

Theorem 2.30. *In a Dedekind domain A, any nonzero proper ideal has a unique factorization as a product of nonzero prime ideals.*

Lemma 2.31. *Let A be a Dedekind domain and let K denote its field of fractions. Let B be the integral closure of A in a finite extension L of K. Assume that B is a finitely generated $A-$module and let $\mathcal{P}$ be a maximal ideal of A. Then $\mathcal{P}B \neq B$.*

Then in our set-up above, by Lemma 2.31, it follows that the ideal $\mathcal{P}B$ factors in B into a product of prime ideals of B, since B is a Dedekind domain. Let

$$\mathcal{P}B = M_1^{e_1} \dots M_s^{e_s},$$

for some positive integers $e_1, \dots, e_s$. The integer $e_{M_i/\mathcal{P}} := e_i$ is called the *ramification index* of M_i over $\mathcal{P}$.

Since $M_i \cap A = \mathcal{P}$, the inclusion $A \subseteq B$ induces injections

$$A/\mathcal{P} \to B/M_i,$$

for $i = 1, \dots, s$. Since B is a finitely generated $A-$module, the field B/M_i is a finite extension of the field $A/\mathcal{P}$. Let $f_{M_i/\mathcal{P}} := [B/M_i : A/\mathcal{P}]$ be the dimension of B/M_i as an $(A/\mathcal{P})-$vector space. The field $A/\mathcal{P}$ is called the *residue field* of A at $\mathcal{P}$. The integer $f_{M_i/\mathcal{P}}$ is called the *residual degree* of M_i over $\mathcal{P}$.

Remark 2.32. *For a maximal ideal M of B, the prime ideal $\mathcal{P} := M \cap A$ is maximal.*

Theorem 2.33. *Let A be a Dedekind domain with field of fractions K. Let L/K be a finite extension. Let B denote the integral closure of A in L. If B is a finitely generated $A-$module, then*

$$[L : K] = \sum_{M|\mathcal{P}B} e_{M/\mathcal{P}} f_{M/\mathcal{P}}.$$

Example 2.34. *Let $A = \mathbb{Z}$. Let $K = \mathbb{Q}(\sqrt{d})$ where d is a square-free integer. Then the ring of integers $\mathcal{O}_K$ is equal to $\mathbb{Z}[\sqrt{d}]$ when $d \equiv 2, 3 \mod 4$, and $\mathbb{Z}[(1 + \sqrt{d})/2]$ when $d \equiv 1 \mod 4$.*

Let $p \in \mathbb{Z}$ be a prime number. The factorization of the primes of $\mathbb{Z}$ above in $\mathcal{O}_K$ is known for every case:

- *If* $p \mid d$, *then* $p\mathcal{O}_K = \langle p, \sqrt{d}\rangle^2$.

- *If* $p = 2$ *and* $2 \nmid d$, *then*

$$2\mathcal{O}_K = \begin{cases} \text{prime} & \text{if } d \equiv 5 \mod 8 \\ \left\langle 2, \frac{1+\sqrt{d}}{2} \right\rangle \left\langle 2, \frac{1-\sqrt{d}}{2} \right\rangle & \text{if } d \equiv 1 \mod 8 \\ \langle 2, 1+\sqrt{d}\rangle^2 & \text{if } d \equiv 3 \mod 4 \end{cases} \tag{2.1}$$

- *If* p *is odd and* $p \nmid d$, *then*

$$p\mathcal{O}_K = \begin{cases} \text{prime} & \text{if } d \text{ is not a square} \mod p \\ \langle p, \sqrt{d}+n\rangle\langle p, \sqrt{d}-n\rangle & \text{if } n^2 \equiv d \mod p \end{cases} \tag{2.2}$$

Proposition 2.35. *Let* A *be any Dedekind domain. Let* $f(y)$ *be an irreducible monic polynomial in* $A[y]$. *Let* α *denote a root of* $f(y)$ *in a fixed algebraic closure of* K, *the field of fractions of* A. *Let* $M \in Max(A[\alpha])$. *Let* $P := M \cap A$. *Factor the reduction* $\overline{f}(y)$ *of* $f(y)$ *in* $(A/P)[y]$ *as* $\overline{f}(y) = \prod_{i=1}^{s} \overline{g_i}(y)^{e_i}$, *with* $\overline{g_i}(y)$, $i = 1, \ldots, s$, *distinct monic irreducible polynomials. Write* $f(y) = h(y) + \prod_{i=1}^{s} g_i(y)^{e_i}$, *with* $h(y) \in PA[y]$, *and such that the reduction of* $g_i(y)$ *in* $(A/P)[y]$ *is equal to* $\overline{g_i}(y)$, *for all* $i = 1, \ldots, s$. *Then the ideal* M *is generated by the elements of* P *and by* $g_{i_0}(\alpha)$, *for some unique* $i_0 \in \{1, \ldots, s\}$. *Moreover, the following statements are equivalent:*

- $f'(\alpha) \notin M$.
- $e_{i_0} = 1$ *and the extension* $A[\alpha]/M$ *is separable over* A/P.

Definition 2.36. *Let* A *be a subring of a ring* B. *The ring extension* B/A *is called simple or monogenic if there exists an element* $\alpha \in B$ *such that* $B = A[\alpha]$.

Example 2.37. *Let $K = \mathbb{Q}(\sqrt{2}, \sqrt{3})$, we find its ring of integers.*

Let $\alpha \in \mathbb{Z}_K$. Then $\alpha = a + b\sqrt{2} + ci + d\sqrt{2}i$ for some $a, b, c, d \in \mathbb{Q}$. Since α is an algebraic integer, then indeed all of its conjugates are so: $\alpha_2 = a - b\sqrt{2} + ci - d\sqrt{2}i$, $\alpha_3 = a + b\sqrt{2} - ci - d\sqrt{2}i$, $\alpha_4 = a - b\sqrt{2} - ci + d\sqrt{2}i$. By the closedness under addition in $\mathbb{Z}_K$, these are also algebraic integers: $\alpha + \alpha_2 = 2a + 2ci$, $\alpha + \alpha_3 = 2a + 2b\sqrt{2}$, $\alpha + \alpha_4 = 2a + 2d\sqrt{2}i$ and integral if $2a, 2b, 2c, 2d \in \mathbb{Z}$. Rewrite $\alpha = \dfrac{A + B\sqrt{2} + Ci + D\sqrt{2}i}{2}$ for integers $A = 2a, B = 2b, C = 2c$, and $D = 2d$. By the closedness under multiplication, $\alpha\alpha_2 = a^2 - 2b^2 - c^2 + 2d^2 + 2aci - 4bdi = \dfrac{A^2 - 2B^2 - C^2 + 2D^2}{4} + \dfrac{AC - 2BD}{2}i$ is also integral. Therefore $4 | A^2 - 2B^2 - C^2 + 2D^2$ and $2 | AC - 2BD$, then $2 | AC$. If only A or C were even, then the expression $A^2 - 2B^2 - C^2 + 2D^2$ would be odd, and this would give a contradiction with divisibility to 4. Hence both of them are even. Using this in the first expression, we get $4 | -2B^2 + 2D^2$ or $2 | D^2 - B^2$. So either both of them are even or both of them are odd. We find that integers are all of the form $\alpha = a + b\sqrt{2} + ci + d\sqrt{2}i$ for some $a, c \in \mathbb{Z}$ and b and d both integral or both halves of odd integers. Such elements are integer linear combinations of $1, \sqrt{2}, i$, and $\dfrac{1+i}{\sqrt{2}}$, thus elements of this form are all integers.

We have showed what kind of elements $\mathbb{Z}_K$ has. Now our claim is $\mathbb{Z}_K = \mathbb{Z}[\gamma]$, where $\gamma = \dfrac{1+i}{\sqrt{2}}$:

$\mathbb{Z}[\gamma] \subset \mathbb{Z}_K$ is obvious, since $\gamma \in \mathbb{Z}_K$.

Using the relations $\gamma^2 = i$, $\gamma - \gamma^3 = \sqrt{2}$, each element in the integral basis can be written as an element in $\mathbb{Z}[\gamma]$, so $\mathbb{Z}_K \subset \mathbb{Z}[\gamma]$ and we are done.

Proposition 2.38. *Suppose that K is a number field, and that $\mathbb{Z}_K = \mathbb{Z}[\gamma]$. Write $g(X) \in \mathbb{Z}[X]$ for the minimal polynomial of γ over $\mathbb{Z}$. Let p be a prime in $\mathbb{Z}$, and let*

$$\overline{g}(X) = \overline{g_1}(X)^{e_1} \cdots \overline{g_r}(X)^{e_r}$$

be the factorization of the minimal polynomial g modulo p of γ into irreducibles. Then $p\mathbb{Z}_K = \mathcal{P}_1^{e_1} \cdots \mathcal{P}_r^{e_r}$, for certain distinct ideals $\mathcal{P}_i$ of $\mathbb{Z}_K$; the inertia degree of $\mathcal{P}_i$ is simply given by the degree of $\overline{g_i}(X)$.

Kummer extensions:

Let $f(y) = y^n - a$ be an irreducible polynomial in $\mathbb{Z}[y]$. Let α be a root of $f(y)$ in $\mathbb{C}$. The extension $K = \mathbb{Q}(\alpha)$ is called a *Kummer extension* of $\mathbb{Q}$.

Proposition 2.39. *Let p be a prime. Let $a = q_1 \cdots q_s$ be a product of distinct primes. Let $\alpha := \sqrt[p]{a}$ be a root of $y^p - a$. If $p \mid a$, then $\mathbb{Z}[\alpha]$ is equal to the ring of integers of $\mathbb{Q}(\alpha)$. If $p \nmid a$, then $\mathbb{Z}[\alpha]$ is equal to the ring of integers of $\mathbb{Q}(\alpha)$ if and only if $p^2 \nmid a^p - a$. Moreover, when $\mathbb{Z}[\alpha] = \mathcal{O}_{\mathbb{Q}(\alpha)}$, the primes p and q_i, $i = 1, \ldots, s$, are the only primes of $\mathbb{Z}$ that ramify in $\mathbb{Z}[\alpha]$.*

In addition, a factorization in $\mathbb{Z}[\alpha]$ of a prime q such that $q \mid a$, can be explicitly determined: $\langle q \rangle \mathbb{Z}[\alpha] = \langle \alpha, q \rangle^p$.

Example 2.40. *Let K be any field and $f(y) = y^n - a$ be an irreducible polynomial in $\overline{K}[x, y]$. So $a := a(x)$ is a polynomial in $\overline{K}[x]$. Then Theorem 2.28 implies that $C_f := \overline{K}[x, y]/\langle y^n - a(x)\rangle$ is integrally closed if and only if the curve $Z_f(\overline{K})$ has no singular points. Note that for the polynomial $f(x, y) = y^n - a(x)$, we have $f_x = a'(x)$ and $f_y(x) = ny^{n-1}$. If the characteristic of K is prime to n, then $Z_f(\overline{K})$ is nonsingular if and only if $a(x)$ is square-free. Assume that this is the case, i.e. C_f is a Dedekind domain. We write $a(x) = c\prod_{i=1}^{n}(x - a_i)$ with $a_i \in \overline{K}$ and $a_i \neq a_j$ if $i \neq j$. Consider the map $\overline{K}[x] \mapsto C_f$. The maximal ideal $\langle x - d\rangle \in \overline{K}[x]$ factors as the product of n distinct maximal ideals of C_f if and only if $d \neq a_i$ for all $i = 1, \ldots, n$. We find that $\langle x - a_i\rangle C_f = \langle y, x - a_i\rangle^n$. In particular, the only ideals in $\overline{K}[x]$ that ramify in C_f are the ideals $\langle x - a_i\rangle$, $i = 1, \ldots, r$.*

Example 2.41. *Let $f(x, z) = z^n + x^n + 1$ and $C_f := \mathbb{C}[x, z]/\langle f\rangle$. Consider the map $\pi : Z_f(\mathbb{C}) \to \mathbb{A}^1(\mathbb{C})$, with $[x : 1 : z] \mapsto [x : 1]$, corresponding to the inclusion $\mathbb{C}[x] \to C_f$. The degree of the map $\deg \pi = [C_f : \pi^*(\mathbb{C}[x])]$ is equal to n, since z^n+x^n+1 is irreducible and monic in x. All partial derivatives of the homogeneous polynomial of f vanish at $[0 : 0 : 0]$, but there is no such point in the projective plane. Hence the curve is smooth. Since the curve $Z_f(\mathbb{C})$ is nonsingular, the ring C_f is the integral closure of the ring $\mathbb{C}[x]$ in $\mathbb{C}(Z_f)$. By Example 2.40, $a(x) = x^n + 1 = \prod_{i=1}^{n}(x - \xi_i)^n$ where ξ_i, $i = 1, \ldots, n$ are the n-th roots of -1 and $\langle x - \xi_i\rangle C_f$ are the only ramified ideals in C_f, which are explicitly factorized as $\langle x - \xi_i\rangle C_f = \langle x - \xi_i, z\rangle^n$. Therefore, from the bijection between the nonsingular points of the curve and the maximal ideals of C_f, we say that there are exactly n ramification points with ramification degree n.*

Eisenstein extensions:

Definition 2.42. *Let A be any domain. Let $P \in Max(A)$. Let $f(y) = a_n y^n + a_{n-1}y^{n-1} + \cdots + a_0 \in A[y]$ be a polynomial of positive degree such that $a_n \notin P$, $a_i \in P$, for all $i = 0, \ldots, n-1$, and $a_0 \notin P^2$. We call $f(y)$ an Eisenstein polynomial, or a P-Eisenstein polynomial, when the dependence on P must be specified.*

Lemma 2.43. *Let p be a maximal ideal of $\mathbb{Z}$. Let $f(y) = y^n + \sum_{i=0}^{n-1} a_i y^i$ be a p-Eisenstein polynomial, then $f(y)$ is irreducible in $K[y]$. Let α be a root of $f(y)$ in $\mathbb{C}$. Then $A[\alpha]$ is a local principal ideal domain in $K(\alpha)$ and, in particular, $A[\alpha]$ is the integral closure of A in $K(\alpha)$. Moreover, $pA[\alpha] = \langle\alpha\rangle^n$.*

Cyclotomic extensions:

Let p be a prime number and ξ_p denote a root of $f(y) := 1 + y + \cdots + y^{p-1}$.

Proposition 2.44. *The ring $\mathbb{Z}[\xi_p]$ is the ring of integers of $\mathbb{Q}(\xi_p)$. Let $\mathfrak{P}$ denote the ideal of $\mathbb{Z}[\xi_p]$ generated by $\xi_p - 1$. Then $\mathfrak{P}$ is a prime ideal and $\langle p\rangle\mathbb{Z}[\xi_p] = \mathfrak{P}^{p-1}$. The prime p is the only prime of $\mathbb{Z}$ that ramifies in $\mathbb{Z}[\xi_p]$.*

Proposition 2.45. *Let $q \neq p$ be a prime of $\mathbb{Z}$. Let f_q denote the order of the class of q in $(\mathbb{Z}/p\mathbb{Z})^*$. Then there exist exactly $s := (p-1)/f_q$ distinct prime ideals $\mathcal{L}_1, \ldots, \mathcal{L}_s$ of $\mathbb{Z}[\xi_p]$ over $\langle q\rangle$, and $\langle q\rangle\mathbb{Z}[\xi_p] = \mathcal{L}_1 \cdots \mathcal{L}_s$. Moreover, the residual index $f_{\mathcal{L}_i/q}$ of each prime $\mathcal{L}_i$ is equal to f_q.*

Galois extensions:

Let L/K be a finite Galois extension with Galois group $G = Gal(L/K)$.

Proposition 2.46. *Let p be a maximal ideal of $\mathbb{Z}$. Let M_i and M_j be two distinct maximal ideals in $\pi^{-1}(p)$. Then there exists $\sigma \in G$ such that $\sigma(M_i) = M_j$. Moreover, $e_{M_1/p} = \cdots = e_{M_s/p} = e$, and $f_{M_1/p} = \cdots = f_{M_s/p} = f$. In particular,*

$$pB = \langle M_1 \cdots M_s\rangle^e, \text{ with } efs = [K : \mathbb{Q}].$$

Theorem 2.47 ([7], Theorem 1.3)**.** *For a number field K, the primes which ramify are those dividing the integer $disc_{\mathbb{Z}}(\mathcal{O}_K)$.*

2.3. Algebraic Curves Basics

In this section, we introduce some notions about algebraic curves such as divisors and rational maps. We also give famous *Riemann-Roch theorem* and *Riemann Hurwitz formula*. We mostly benefit from The Arithmetic of Elliptic Curves by Silverman and Generalization of the ABC-conjecture by Nils Bruin. For more details, see [8] and [9].

2.3.1. Divisors

Throughout this section, we take K as a perfect field unless something different is specified.

Definition 2.48. *Let C be a curve defined over K with $\overline{K} = K$. A divisor of C is a formal sum $D := \sum_{P \in C} n_P P$ with $n_P \in \mathbb{Z}$ and all but finitely many $n_P = 0$. The set of points P for which $n_P \neq 0$ is called the support of D, denoted by Supp(D).*

The divisors of C form a free abelian group under addition, the divisor group, $Div(C)$.

Now we generalize to the case where K does not need to be algebraically closed. Let $G_K = Gal(\overline{K}/K)$ be the absolute Galois group of K, i.e., the group of all automorphisms of $\overline{K}$ that fix K.

Definition 2.49. *The prime divisors of C over K are defined by $\mathcal{P}_P := \sum_{\sigma \in G_K} P^\sigma$ where $P \in C(\overline{K})$.*

Definition 2.50. *A divisor $D = \sum_{P \in C} n_P P \in Div(C)$ is defined over K if for all $\sigma \in G_K$, we have $D^\sigma = D$, where $D^\sigma = \left(\sum_{P \in C} n_P P \right)^\sigma = \sum_{P \in C} n_P P^\sigma$. The subset of $Div(C)$, comprising of those divisors that are defined over K, forms the subgroup of K-rational divisors.*

Note that $D = D^\sigma$ does not necessarily imply that $P = P^\sigma$ for all points in the support of D, i.e., it does not have to be fixed pointwise, but rather setwise. But if $D = D^\sigma$ for all $\sigma \in G_K$, then $n_{P^\sigma} = n_P$ for all $\sigma \in G_K$. Hence we can group the terms of a K-rational divisor into G_K-orbits with a single coefficient n_P applied to all the points in the orbit. Equivalently, we can view a K-rational divisor as a sum over G_K-orbits of points $P \in C(\overline{K})$.

Remark 2.51. *The orbit of any algebraic number is the set of all of its conjugates, hence finite.*

Definition 2.52. *Let C be a curve defined over K. Then G_K-orbits of $C(\overline{K})$ are called closed points, which we also denote by P. A rational divisor of C/K is a formal sum $D := \sum n_P P$, where P ranges over the closed points of C/K, with $n_P \in \mathbb{Z}$ and all but finitely many $n_P = 0$. The group of rational divisors on C is defined as the subgroup generated by the prime rational divisors, and denoted by $Div_K(C)$.*

The points in $C(K)$ correspond to a proper subset of the set of closed points, the trivial G_K-orbits consisting of a single element. But every point in $C(\overline{K})$ is contained in a closed point of C/K.

Let C be a curve defined over a perfect field K, we denote by C/K. Let $\overline{K}(C)$ be the function field of C over $\overline{K}$, $\overline{K}[C]_P$ be the local ring of C at P, and M_P be the maximal ideal of $\overline{K}[C]_P$.

Definition 2.53. *Let $P \in C$ a smooth point. The normalized valuation on $\overline{K}[C]_P$ is given by*

$$ord_P : \overline{K}[C]_P \to \{0, 1, 2, \dots\} \cup \{\infty\},$$

$$ord_P(f) = sup\{d \in \mathbb{Z} : f \in {M_P}^d\}.$$

Using $ord_P(f/g) = ord_P(f) - ord_P(g)$, *we can extend the valuation to* $\overline{K}(C)$,

$$ord_P : \overline{K}(C) \to \mathbb{Z} \cup \infty.$$

Let C be a smooth curve, and let $f \in \overline{K}(C)^*$, then we can associate f to a divisor $div(f) = \sum_{P \in C} ord_P(f)P$. Since each ord_P is a valuation, the map

$$div : \overline{K}(C)^* \to Div(C)$$

is a homomorphism of abelian groups.

Definition 2.54. *Let* $f \in \overline{K}(C)^*$. *The divisor of* f *is,* $divf := \sum_{P \in C_F} ord_P(f)P$, *called a principal divisor.*

Note that for any closed point, a function $f \in K(C)$ vanishes at a point $P \in C(\overline{K})$ if and only if it vanishes on the whole G_K-orbit of P.

For a principal divisor $divf = \sum_{n_P} P$, the divisors $div_0 f := \sum_{n_P > 0} n_P P$ and $div_\infty f := \sum_{n_P < 0} -n_P P$ are called *the divisor of zeros* and *the divisor of poles*, respectively. We have the equality $divf = div_0 f - div_\infty f$. Note that $div_0 f$ is sometimes called the zero divisor of f, e.g. in the paper, which we study later, "Abc Implies Mordell" by Noam Elkies.

The Picard group of C, denoted by $Pic(C)$, is the quotient of $Div(C)$ by its subgroup of principal divisors.

Definition 2.55. *Let* C/K *be a curve. If* P *is a closed point of* C/K, *we define the degree of* P *to be* $degP = [K(P) : K]$, *where* $K(P)$ *is the minimal field extension* K' *of* K *in* $\overline{K}$ *such that* P *is* K'*-rational. Equivalently,* $degP$ *is the cardinality of the closed point* P *as a* G_K*-orbit of points in* $C(\overline{K})$.

The degree of a divisor $D = \sum n_P P$ in the group of K-rational divisors is $degD = \sum n_P degP$. When K is algebraically closed, we have $degP = 1$ for all points P, so $degD = \sum n_P$.

The divisors of degree 0 form a subgroup of $Div(C)$, which we denote by $Div^0(C) = \{D \in Div(C) : degD = 0\}$.

Example 2.56. *Let F_5 be a curve over $\mathbb{Q}$, where F_5 is the zero locus of $x^5+y^5+z^5 = 0$. Let $f : F_5(\overline{\mathbb{Q}}) \to \mathbb{P}^1(\mathbb{C})$ with $(x, y, z) \mapsto (-x^5, y^5)$. Let D be the zero divisor of f. Let $\xi_5 \neq -1$ be a 5th root of -1. Then $D = 5((0, 1, \xi_5)+(0, 1, {\xi_5}^2)+(0, 1, {\xi_5}^3)+(0, 1, {\xi_5}^4))+ 5(0, 1, -1)$, where $P_1 := (0, 1, \xi_5)+(0, 1, {\xi_5}^2)+(0, 1, {\xi_5}^3)+(0, 1, {\xi_5}^4)$ and $P_2 := (0, 1, -1)$ are prime rational divisors of degrees $deg(P_1) = 4$ and $deg(P_2) = 1$. Then $deg(D) = 25$.*

Now we partially order divisors by defining the relation $\leq$ on $Div_K C$ by $D_1 \leq D_2$ if and only if $n_P(D_1) \leq n_P(D_2)$ for all P where P ranges over all closed points of C/K. Then we have some consequences:

- If $D_1 \leq D_2$, then $D_1 + E \leq D_2 + E$ for any divisor E
- If $D_1 \leq D_2$ and $E_1 \leq E_2$, then $D_1 + E_1 \leq D_2 + E_2$.

Note that $\leq$ is not a total ordering on $Div_K C$ since most pairs of divisors are incomparable.

Definition 2.57. *A divisor $D = \sum n_P P$ is called effective or positive if $n_P \geq 0$ for all P. We denote by $D \geq 0$.*

Like principal divisors $divf = div_0 f - div_\infty f$, every divisor can be written uniquely as a difference of two effective divisors, $D = D_0 - D_\infty$, where $D_0 := \sum_{n_P > 0} n_P P$ and $D_\infty := \sum_{n_P < 0} -n_P P$.

Definition 2.58. *Two divisors are linearly equivalent if their difference is the divisor of a function.*

Proposition 2.59. *Let C be a smooth curve and let $f \in \overline{K}(C)^*$. Then $deg(div(f)) = 0$.*

Definition 2.60. *Let V_1 and $V_2 \subset \mathbb{P}^n$ be projective varieties. A rational map from V_1 to V_2 is a map of the form $f : V_1 \to V_2$ with $f = [f_0, \ldots, f_n]$, where the functions $f_0, \ldots, f_n \in \overline{K}(V_1)$ have the property that for every point $P \in V_1$ at which $f_0, \ldots, f_n$ are all defined, $f(P) = [f_0(P), \ldots, f_n(P)] \in V_2$. Multiplying with an appropriate function g, we can assume that all gf_i are defined in $x \in V_1$ without changing the value of f in x. We say f is regular at x if there exists a representative $(gf_0, \ldots, gf_n)$ that is defined in x. If f is regular for all $x \in C$, then we say that f is a morphism.*

Let $\phi : C_1 \to C_2$ be a nonconstant map of smooth curves, then ϕ induces maps on the function fields of C_1 and C_2,

$$\phi^* : \overline{K}(C_2) \to \overline{K}(C_1) \text{ and } \phi_* : \overline{K}(C_1) \to \overline{K}(C_2).$$

We similarly define maps of divisor groups as follows:

$$\phi^* : Div(C_2) \to Div(C_1) \text{ given by } (Q) \mapsto \sum_{P \in \phi^{-1}(Q)} e_\phi(P)(P),$$

$$\phi_* : Div(C_1) \to Div(C_2) \text{ given by } (P) \mapsto (\phi P)$$

and extend $\mathbb{Z}-$linearly to arbitrary divisors.

Definition 2.61. *The degree of a morphism of curves $\phi : C_1 \to C_2$ is the degree of the corresponding extension of function fields $deg(\phi) = [K(C_1) : \phi^*(K(C_2))]$.*

Proposition 2.62. *Let $\phi : C_1 \to C_2$ be a nonconstant map of smooth curves.*

- *$deg(\phi^* D) = (deg(\phi))(degD)$ for all $D \in Div(C_2)$.*
- *$\phi^*(div(f)) = div(\phi^* f)$ for all $f \in \overline{K}(C_2)^*$.*
- *$deg(\phi_* D) = degD$ for all $D \in Div(C_1)$.*
- *$\phi_*(div(f)) = div(\phi_* f)$ for all $f \in \overline{K}(C_1)^*$.*
- *$\phi_* \circ \phi^*$ acts as multiplication by $deg(\phi)$ on $Div(C_2)$.*

2.3.2. Differentials

Definition 2.63. *Let C be a curve. The space of differential forms on C, denoted by Ω_C, is the $\overline{K}-$vector space generated by symbols of the form dx for $x \in \overline{K}(C)$, subject to the usual relations:*

- $d(x+y) = dx + dy$ *for all* $x, y \in \overline{K}(C)$.
- $d(xy) = xdy + ydx$ *for all* $x, y \in \overline{K}(C)$.
- $da = 0$ *for all* $a \in \overline{K}$.

Let $\phi : C_1 \to C_2$ be a nonconstant map of curves, then associated function field map $\phi^* : \overline{K}(C_2) \to \overline{K}(C_1)$ induces a map on differentials,

$$\phi^* : \Omega_{C_2} \to \Omega_{C_1}, \text{ given by } \phi^*\left(\sum f_i dx_i\right) = \sum (\phi^* f_i) d(\phi^* x_i).$$

Proposition 2.64.
- Ω_C *is a* 1*-dimensional* $\overline{K}(C)$*-vector space.*
- *Let* $x \in \overline{K}(C)$. *Then* dx *is a* $\overline{K}(C)$*-basis for* Ω_C *if and only if* $\overline{K}(C)/\overline{K}(x)$ *is a finite separable extension.*
- ϕ *is separable if and only if* $\phi^* : \Omega_{C_2} \to \Omega_{C_1}$ *is injective.*

Definition 2.65. *Let $w \in \Omega_C$. The divisor associated to w is*

$$div(w) = \sum_{P \in C} ord_P(w) P \in Div(C).$$

The differential $w \in \Omega_C$ is holomorphic if $ord_P(w) \geq 0$ for all $P \in C$, and it is nonvanishing if $ord_p(w) \leq 0$ for all $P \in C$.

For more details, please see Chapter 2 in Arithmetic of Elliptic Curves by Silverman [8].

If $w_1, w_2 \in \Omega_C$ are nonzero differentials, then there is a function $f \in \overline{K}(C)^*$ such that $w_1 = f w_2$. Thus $div(w_1) = div(f) + div(w_2)$.

Definition 2.66. *The canonical divisor class on C is the image in $Pic(C)$ of $div(w)$ for any nonzero differential $w \in \Omega_C$. Any divisor in this class is called a canonical divisor.*

Proposition 2.67. *Let E be an elliptic curve. Then the invariant differential w associated to a Weierstrass equation for E is holomorphic and nonvanishing, i.e., $div(w) = 0$.*

2.3.3. Riemann-Roch Theorem

The partial order we put on divisors is useful for describing zeros and poles of functions.

Definition 2.68. *Let $D \in Div(C)$. We associate to D the set of functions*

$$\mathcal{L}(D) = \{f \in \overline{K}(C)^* : div(f) \geq -D\} \cup \{0\},$$

which is known as Riemann-Roch space associated to the divisor D. It is sometimes called the linear system associated with D. Its dimension is denoted by $l(D) = dim_{\overline{K}}\mathcal{L}(D)$.

Proposition 2.69. *Let $D \in Div(C)$.*

- *If $degD < 0$, then $\mathcal{L}(D) = \{0\}$ and $l(D) = 0$.*
- *$\mathcal{L}(D)$ is a finite-dimensional $\overline{K}$-vector space.*
- *If $D' \in Div(C)$ is linearly equivalent to D, then*

$$\mathcal{L}(D) \cong \mathcal{L}(D'), \text{ and so } l(D) = l(D').$$

Theorem 2.70 (Riemann-Roch Theorem)**.** *Let C be a smooth curve and let K_C be a canonical divisor on C. There is an integer $g \geq 0$, called the genus of C, such that for every divisor $D \in Div(C)$,*

$$l(D) - l(K_C - D) = degD - g + 1.$$

Corollary 2.71. *(i)* $l(K_C) = g$.

(ii) $degK_C = 2g - 2$.

(iii) If $degD > 2g - 2$, *then* $l(D) = degD - g + 1$.

Proof. (i) Using Theorem 2.70 with $D = 0$, we get $l(K_C) = l(0) + g - 1$. If $D = 0$, then $\mathcal{L}(D)$ is the set of functions that have no poles at all, i.e., $div(f) \geq 0$. Then the rational map defined by f, $f : C \to \mathbb{P}^1$, is a constant map, hence $f \in \overline{K}$ and $\mathcal{L}(0) \cong \overline{K}$. So we have $l(0) = 1$ and $l(K_C) = g$.

(ii) Using Theorem 2.70 with $D = K_C$ and Part (i), we get $l(K_C) = 2g - 2$.

(iii) If $degD > 2g-2$, then by Part (ii), $deg(K_C - D) < 0$. Take an arbitrary nonzero $f \in \mathcal{L}(K_C - D)$, then $0 = deg(div(f)) \geq -deg(K_C - D)$, so $deg(K_C - D) \geq 0$. Hence $\mathcal{L}(K_C - D) = 0$, and $l(K_C - D) = 0$. Now applying Theorem 2.70, we have $l(D) = degD - g + 1$.

□

Proposition 2.72. *Let* C/K *be a smooth curve and let* $D \in Div_K(C)$. *Then* $\mathcal{L}(D)$ *has a basis consisting of functions in* $K(C)$.

Proposition 2.73. *If an elliptic curve* E *given by a Weierstrass equation is singular, then there exists a rational map* $\phi : E \to \mathbb{P}^1$ *of degree one.*

Lastly, we look at the behavior of a map in the neighborhood of a point. Let C be a curve over some number field K and $f \in \overline{\mathbb{Q}}(C)$ be a rational function. We define the *branch number* of f in $x \in C$ by

$$b_f(x) = ord_x(f - f(x)) - 1 \text{ if } f(x) \text{ is finite, and}$$

$$b_f(x) = ord_x\left(\frac{1}{f}\right) - 1 \text{ if } f \text{ has a pole at } x.$$

For $\alpha \in K \cup \{\infty\}$ we define the *ramification* of f above α by

$$b_f(\alpha) = \sum_{x:f(x)=\alpha} b_f(x).$$

We have $deg(f) = |f^{-1}(\{\alpha\})| + b_f(\alpha)$.

We give a theorem on how the genera of curves linked by a nonconstant map are related.

Theorem 2.74 (Hurwitz, [8])**.** *Let $\phi : C_1 \to C_2$ be a nonconstant separable map of smooth curves of genera g_1 and g_2, respectively. Then*

$$2g_1 - 2 \geq (deg(\phi))(2g_2 - 2) + \sum_{P \in C_1} (e_\phi(P) - 1).$$

Further, equality holds if and only if one of the following two conditions is true:

(i) $char(K) = 0$.

(ii) $char(K) = p > 0$, *where p is prime, and p does not divide $e_\phi(P)$ for all $P \in C_1$.*

It turns out that the total ramification of a rational function on a curve is closely connected with the genus of that curve.

Remark 2.75. *Let C be a curve of genus g over an algebraically closed field K and let f be a nonconstant rational function on that curve. Then*

$$2deg(f) = \sum_{\alpha \in K} b_f(\alpha) + 2 - 2g.$$

If C is a curve over a number field K, then this need not be true. However, since the branch number $b_f(\alpha)$ is nonzero for only finitely many $\alpha \in \overline{\mathbb{Q}}$, the theorem is true for some finite extension of K, dependent on f.

Definition 2.76. *The Fermat curve F_n is the algebraic curve in the complex projective plane defined in homogeneous coordinates (x, y, z) by the Fermat equation $x^n+y^n+z^n = 0$.*

Example 2.77. *Let $f(x, z) = z^n + x^n + 1$ be an irreducible polynomial with $deg_z(f) = n > 0$ and monic in z. We consider the map $\pi : Z_f(\mathbb{C}) \to \mathbb{A}'(\mathbb{C})$. In Example* 2.41, *we showed that there are exactly n ramification points with ramification index n, and the degree of the map π is n. Note that the projective line $\mathbb{P}^1(\mathbb{C})$ has no handles and hence its genus is* 0. *Using Hurwitz Theorem 2.74, we see that the genus of the Fermat curve is*

$$g = \frac{(n-1)(n-2)}{2}.$$

Hence for $n \geq 4$, the curve has genus ≥ 2, hence has finitely many rational points by Mordell conjecture 5.1.

2.3.4. Rational Maps

Let C be a curve over K and let $f_0, \ldots, f_n \in K(C)$. Then $f : C \to \mathbb{P}^n$ with $x \mapsto (f_0(x), \ldots, f_n(x))$ is a rational map from C into $\mathbb{P}^n$, which is defined for $x \in C$ if not all $f_i(x) = 0$ and no f_i has a pole in x. Since C is a curve, we can construct for each $x \in C$ a function g with a pole or a zero at x of arbitrary order. Therefore, if f is regular at some point, i.e., $f \neq (0, \ldots, 0)$, then f is a morphism.

Definition 2.78. *Let C be a smooth projective curve and $D \in Div(C)$. Let $F = \{f_0, \ldots, f_n\}$ be a basis for $\mathcal{L}(D)$, then $\phi_{D,F} : C \to \mathbb{P}^n$ with $x \mapsto (f_0(x), \ldots, f_n(x))$ is a rational map. If this map is a nonconstant morphism, then $\mathcal{L}(D)$ is said to be without base point.*

Remark 2.79. *For curves, $\mathcal{L}(D)$ is without base point once $l(D) \geq 2$ because the associated map ϕ_D must be nonconstant and regular somewhere.*

Definition 2.80. *If ϕ_D is injective and regular, then D is called very ample. If some multiple of D is very ample, then D is called ample.*

There is no canonical basis for $\mathcal{L}(D)$. Let $G = \{g_0, \ldots, g_n\}$ be another basis for $\mathcal{L}(D)$, then $g_i = \sum_{j=0}^{n} a_{ij} f_j$ for $i = 0, \ldots, n$, and $\phi_{D,G} = M \circ \phi_{D,F}$ where M is the linear transformation of $\mathbb{P}^n$ determined by the coefficients a_{ij}. Therefore, we can associate a rational map ϕ_D with a divisor D defined up to linear transformation.

For any linearly equivalent divisors $D_1 \sim D_2$, there exists a function $g \in \overline{K}(C)$ such that $D_1 = D_2 + div(g)$. Then we have $\mathcal{L}(D_1) = \{f \in \overline{K}(C) : div(f) + D_2 + div(g) \geq 0\} = \{f \in \overline{K}(C) : div(fg) + D_2 \geq 0\}$ and $\mathcal{L}(D_2) = g\mathcal{L}(D_1)$. So if $\{f_0, \ldots, f_n\}$ is a basis for $\mathcal{L}(D_1)$, then $\{gf_0, \ldots, gf_n\}$ is a basis for $\mathcal{L}(D_2)$. Thus, we have $\phi_{D_1} = \phi_{D_2}$ since we map into projective space.

Therefore, if $\overline{D} \in Pic(C)$, we can define $\phi_{\overline{D}} : C \to \mathbb{P}^n$ up to linear transformation.

Fact 2.81. *If $\mathcal{L}(D)$ and $\mathcal{L}(D')$ are two spaces without base points with bases $\{f_0, \ldots, f_n\}$ and $\{g_0, \ldots, f_m\}$, then $\mathcal{L}(D + D')$ is spanned by $\{f_i g_j\}$, where $i = 1, \ldots, n$ and $j = 1, \ldots, m$.*

Definition 2.82. *A complete linear system on a variety is defined as the set of all effective divisors linearly equivalent to some given divisor D. It is denoted by $|D|$.*

Definition 2.83. *Suppose that $|D|$ is a complete linear system of divisors on some variety V. The set of base points of $|D|$ is the intersection of supports of all divisors in $\mathcal{L}(D)$.*

The rational map ϕ_D associated to divisor D is defined outside of the base points of $|D|$.

Definition 2.84. *A divisor is called base point free if $|D|$ has no base points.*

Fact 2.85 ([10], Fact 9.15)**.** *Let D be any divisor and H be an ample divisor. Then $D + mH$ is base point free for m big enough. Furthermore, if D is base point free, then $D + H$ is very ample.*

Corollary 2.86. *Every divisor D can be written as a difference of two base point free (very) ample divisors.*

Proof. Choose a very ample divisor, say H. By Fact 2.85, for m big enough, $D + mH$ and mH are base point free very ample divisors. Let $D = (D + mH) - mH$ and we are done. □

Fact 2.87. *Every divisor on a curve of sufficiently high degree is very ample by Corollary 2.71 (iii).*

2.3.5. Primes of Good Reduction

Let K be a number field. Let C be a curve consisting of points $(x_0, \ldots, x_n) \in \mathbb{P}^n$ satisfying the homogeneous equations $h_1(x_0, \ldots, x_n) = 0, \ldots, h_k(x_0, \ldots, x_n) = 0$ with $k \geq n-1$ and h_i irreducible. Let C be defined over K, i.e., the coefficients of $h_1, \ldots, h_k$ lie in K. Also let $f : C \to \mathbb{P}^1$ be a map defined over K, where $f = (f_0, f_1)$ and f_i homogeneous polynomials of the same degree with coefficients in K. Multiply all h_i's and f_i's by a common denominator so that all polynomials have integral coefficients. Then in the reduction of a finite prime $\mathcal{P}$, we take each coefficient modulo $\mathcal{P}$. We set an algebraic number α to 0 if $ord_{\mathcal{P}}(\alpha) > 0$, and to ∞ if $ord_{\mathcal{P}} < 0$.

In the reduction of some primes, some h_i may become indeterminate or the degree of f may decrease or f may map a point to $(0, 0)$. Also, given a divisor $D \in Div_K(C)$, some points in D may coalesce or may become indeterminate. We call such primes as *bad primes*.

We give a procedure for eliminating these finitely many bad primes, which is taken from *Abc Implies Roth's Theorem and Mordell Conjecture* by *Machiel van Frankenhuysen*, [11]: Take an arbitrary point $a \in \mathbb{P}^1(K)$ such that f is not ramified over a and take another point $b \in \mathbb{P}^1(K)$ different from a. If we have $\overline{a} = 0$ or $\overline{b} = 0$ or $\overline{a} = \overline{b}$ in the reduction, we exclude such primes. Also if the points in the divisors $\sum_{P \in f^{-1}(a)} e_P(f)(P)$ and $\sum_{P \in f^{-1}(b)} e_P(f)(P)$ coalesce or become indeterminate, we exclude such primes, too. Now for the remaining primes, we have $\overline{a} \neq \overline{b}$, so $\overline{f}$ is not constant, and $deg(\overline{f}) = deg(\overline{f}^*(\overline{a})) = deg(f)$. Still $\overline{f}$ may map a point of C to an indeterminate point, but this happens for only finitely many primes.

2.4. Elliptic Curves

In this section, we introduce elliptic curves and then give some important properties which are necessary for the Section 3, *Hall conjecture.* For more details and proofs, see [8] and [12].

Definition 2.88. *Let K be a field. An elliptic curve over K is a smooth cubic projective curve defined over K that has genus 1 and a distinguished rational point O, namely the point at infinity.*

Proposition 2.89 ([8], Chapter 3, Proposition 3.1)**.** *Let K be a perfect field and E be an elliptic curve defined over K with the point at infinity O.*

(i) There exist functions $x, y \in K(E)$ such that the map $\phi : E \to \mathbb{P}^2$ with $\phi = [x, y, 1]$ gives an isomorphism of E/K onto a curve given by a Weierstrass equation

$$C : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

with coefficients $a_1, \ldots, a_6 \in K$; and such that $\phi(O) = [0, 1, 0]$.

(ii) Conversely, every smooth cubic curve C given by a Weierstrass equation as above is an elliptic curve defined over K with origin $O = [0, 1, 0]$.

Proof. (i) First we look at the vector spaces $\mathcal{L}(n(O))$ for $n \in \mathbb{Z}^+$. E has genus 1, and $deg(n(O)) > 2g-2 = 0$, then by Corollary 2.71 (iii), $l(n(O)) = deg(n(O)) = n$ for all $n \geq 1$. Thus by Proposition 2.72, $\mathcal{L}(n(O))$ has a basis consisting of functions in $K(E)$. The set $\{1, x\}$ is a basis for $\mathcal{L}(2(O))$ and $\{1, x, y\}$ is a basis for $\mathcal{L}(3(O))$. The Riemann Roch space $\mathcal{L}(2(O))$ is the vector space of functions such that $div(f) \geq -2(O)$, so functions in $\mathcal{L}(2(O))$ have poles at the point of infinity of order at most 2. Note that x must have a pole of exact order 2 at O, otherwise it would be in $\mathcal{L}(1(O))$. By similar reasoning, y must have a pole of exact order 3 at O. Observe that the seven functions $1, x, y, x^2, xy, x^3, y^2$ are all in $\mathcal{L}(6(O))$, since they satisfy $div(f) \geq -6O$. However, $l(6(O)) = 6$, so there must be a K-linear relation between them.

Say $A_1 + A_2x + A_3y + A_4x^2 + A_5xy + A_6y^2 + A_7x^3 = 0$ with $A_i \in K$. Notice that first five terms in the sequence $1, x, y, x^2, xy, x^3, y^2$ have poles of order $0, 2, 3, 4, 5$, respectively, but x^3 and y^2 both have poles of order 6. Thus leaving out one of them would give a basis for $\mathcal{L}(6(O))$, so the coefficients of x^3 and y^2 in the dependence relation must be nonzero, i.e., $A_6.A_7 \neq 0$.

Replacing x and y by $-A_6A_7x$ and $A_6{A_7}^2y$, then dividing by ${A_6}^3{A_7}^4$, we get

$$C : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \text{ with coefficients } a_1, \ldots, a_6 \in K.$$

This gives a map $\phi : E \to C \subset \mathbb{P}^2$ with $P \mapsto [x(P), y(P), 1]$, which is a morphism since E is smooth and surjective since the map is not constant. Moreover, ϕ maps O to $[0, 1, 0]$ because y has a higher order pole at point at infinity than x.

Now we want to show that ϕ has degree one, or equivalently $K(E) \cong K(x, y)$. Consider the map $[x, 1] : E \to \mathbb{P}^1$. Since x has a double pole at O, using the definition of the degree of a map by ramification, this map has degree 2. Hence $[K(E) : K(x)] = 2$. In a similar way, the map $[y, 1] : E \to \mathbb{P}^1$ has degree 3, and so $[K(E) : K(y)] = 3$. By tower rule, $[K(E) : K(x, y)]$ must divide both 2 and 3, hence $[K(E) : K(x, y)] = 1$, and $K(E) \cong K(x, y)$.

In addition, we want to show that C is smooth. So assume for a contradiction C is singular. Then from Proposition 2.73, there is a rational map $\varphi : C \to \mathbb{P}^1$ of degree 1. Taking the composition of two maps, we get that $\varphi \circ \phi : E \to \mathbb{P}^1$ is a map of degree 1, but then $E \cong \mathbb{P}^1$. We get a contradiction since E has genus 1 and $\mathbb{P}^1$ has genus 0. In the end, we have a map $\phi : E \to C$ of degree 1 between two smooth curves. Hence $E \cong C$.

(ii) Let E be given by a nonsingular Weierstrass equation. Using Proposition 2.67, we know that the differential $w = \dfrac{dx}{2y + a_1x + a^3} \in \Omega_E$ does not have any zeros or poles, so $div(w) = 0$. Then, by Theorem 2.70, we get $2g - 2 = deg(div(w)) = 0$, where g is the genus of E, and E has genus 1. Also taking $[0, 1, 0]$ as the base point makes E into an elliptic curve.

□

Proposition 2.89 above provides us the general *Weierstrass form*. Then we use some linear change of variables to obtain *short Weierstrass form*.

Consider the equation $C : y^2+axy+by = x^3+cx^2+dx+e$. We use the substitution $y \mapsto y - \frac{a}{2}x - \frac{b}{2}$ to eliminate xy and y terms.

$$\left(y - \frac{a}{2}x - \frac{b}{2}\right)^2 + ax\left(y - \frac{a}{2}x - \frac{b}{2}\right) + b\left(y - \frac{a}{2}x - \frac{b}{2}\right) = y^2 - \frac{a^2}{4}x^2 - \frac{ab}{2}x - \frac{b^2}{4}$$

Combining with the right hand side, we get $y^2 = x^3 + \left(c + \frac{a^2}{4}\right)x^2 + \left(d + \frac{ab}{2}\right)x + \left(e + \frac{b^2}{4}\right)$. Let $a' := c + \frac{a^2}{4}, b' := d + \frac{ab}{2}$, and $c' := e + \frac{b^2}{4}$, then $y^2 = x^3 + a'x^2 + b'x + c'$. Finally, use the substitution $x \mapsto x - \frac{a'}{3}$ to eliminate x^2.

$$y^2 = x^3 - a'x^2 + \frac{(a')^2}{3}x - \frac{(a')^3}{27} + a'x^2 - \frac{2(a')^2}{3}x + \frac{(a')^3}{9} + b'x - \frac{a'b'}{3} + c'$$

So we get $y^2 = x^3 + Ax + B$ with $A = b' - \frac{(a')^2}{3}, B = c' + \frac{2(a')^3}{27} - \frac{a'b'}{3}$.

Elliptic curves have a special group structure in their geometric nature. Let E be an elliptic curve defined over $\mathbb{Q}$. Then $(E(\mathbb{Q}), +)$ is a group with the point at infinity O as the identity element. The addition is defined as follows: Let P and Q be two rational points in $E(\mathbb{Q})$ and let $\mathbb{L}$ be the line joining P and Q. (If $P = Q$, then $\mathbb{L}$ is the tangent line to E at P.) Since E is cubic, the line $\mathbb{L}$ and E intersect at a third point, say R. Next we join O and R by a line and take the third intersection point with E to be $P + Q$. The addition law is commutative, hence $(E(\mathbb{Q}), +)$ is an abelian group.

Theorem 2.90 (Mordell-Weil Theorem, [12])**.** *$E(\mathbb{Q})$ is a finitely generated abelian group. In other words, there are points $P_1, \ldots, P_n$ such that any other point Q in $E(\mathbb{Q})$ can be expressed as a linear combination*

$$Q = a_1P_1 + a_2P_2 + \cdots + a_nP_n \quad \textit{for some } a_i \in \mathbb{Z}.$$

Thanks to *Mordell-Weil theorem*, we know the general structure of the group completely:

$$E(\mathbb{Q}) \cong E(\mathbb{Q})_{tors} \oplus \mathbb{Z}^{R_E}$$

where $E(\mathbb{Q})_{tors} = \{P \in E(\mathbb{Q}) : \exists n \in \mathbb{N} \text{ such that } nP = O\}$ and R_E is the rank of E. The first part of the decomposition is a finite group formed by the torsion points of E, i.e., points of finite order. The second part is called the free part, and generated by points of infinite order. The minimum number of generators gives the rank.

There is a simple algorithm to determine $E(\mathbb{Q})_{tors}$:

Theorem 2.91 (Nagell-Lutz)**.** *Let $E/\mathbb{Q}$ be an elliptic curve with Weierstrass equation $y^2 = x^3 + Ax + B$, $A, B \in \mathbb{Z}$. Then, every torsion point $P \neq O$ of E satisfies:*

(i) The coordinates of P are integers, i.e. $x(P), y(P) \in \mathbb{Z}$.
(ii) If P is a point of order $n \geq 3$, then $4A^3 + 27B^2$ is divisible by $y(P)^2$.
(iii) If P is of order 2, then $y(P) = 0$ and $x(P)^3 + Ax(P) + B = 0$.

Let $E/\mathbb{Q}$ be an elliptic curve given by a Weierstrass equation $y^2 = x^3 + Ax + B$ with $A, B \in \mathbb{Z}$. Let p be a prime. If we reduce A, B modulo p, then we obtain the equation of a curve $\hat{E}$ given by a cubic curve and defined over the field $\mathbb{F}_p$. Even though E is smooth as a curve over $\mathbb{Q}$, the curve $\hat{E}$ may be singular over $\mathbb{F}_p$.

Definition 2.92. *Let E be an elliptic curve given by $y^2 = x^3 + Ax + B$, with $A, B \in \mathbb{Q}$.*

(i) We define Δ_E, the discriminant of E, by $\Delta_E = -16(4A^3 + 27B^2)$.
(ii) Let S be the set of all elliptic curves E' that are isomorphic to E over $\mathbb{Q}$ and such that the discriminant of E' is an integer. The minimal discriminant of E is the integer Δ'_E that attains the minimum of the set $\{|\Delta'_E| : E' \in S\}$. If E' is the minimal model for E with minimal discriminant, we say that E' is a minimal model for E.

Let $\hat{E}$ be a cubic curve over a field K with Weierstrass equation $f(x,y) = 0$, where $f(x,y) = y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x - a^6$, and suppose that $\hat{E}$ has a singular point $P = (x_0, y_0)$, i.e., $f_x(P) = f_y(P) = 0$. Thus, one can write the Taylor expansion of $f(x,y)$ around (x_0, y_0) :

$$f(x,y) - f(x_0,y_0) = ((y-y_0) - \alpha.(x-x_0)).((y-y_0) - \beta.(x-x_0)) - (x-x_0)^3$$

for some $\alpha, \beta \in \overline{K}$.

Definition 2.93. *The singular point $P \in \hat{E}$ is a node if $\alpha \neq \beta$. In this case there are two different tangent lines to E at P, namely*

$$y - y_0 = \alpha.(x - x_0), \quad y - y_0 = \beta.(x - x_0).$$

If $\alpha = \beta$, then we say that P is a cusp, and there is a unique tangent line at P.

Definition 2.94. *Let $E/\mathbb{Q}$ be an elliptic curve given by a minimal model, let $p \geq 2$ be a prime and let $\hat{E}$ be the reduction curve of E modulo p. We say that $E/\mathbb{Q}$ has good reduction modulo p if $\hat{E}$ is smooth and hence is an elliptic curve over $\mathbb{F}_p$. If $\hat{E}$ is singular at a point $P \in E(\mathbb{F}_p)$, then we say that $E/\mathbb{Q}$ has bad reduction at p and we distinguish two cases:*

(i) If $\hat{E}$ has a cusp at P, then we say that E has additive reduction.

(ii) If $\hat{E}$ has a node at P, then we say that E has multiplicative reduction. If the slopes of the tangent lines, α and β, are in $\mathbb{F}_p$, then the reduction is said to be split multiplicative; and non-split otherwise.

Proposition 2.95. *Let K be a field and let E/K be a cubic curve given by $y^2 = f(x)$, where $f(x)$ is a monic cubic polynomial in $K[x]$. Suppose that $f(x) = (x-\alpha)(x-\beta)(x-\gamma)$ with $\alpha, \beta, \gamma \in \overline{K}$ and put*

$$D = (\alpha - \beta)^2(\alpha - \gamma)^2(\beta - \gamma)^2.$$

Then E is nonsingular if and only if $D \neq 0$.

The quantity D in Proposition 2.95 is the discriminant of the polynomial $f(x)$. The discriminant of $E/\mathbb{Q}$, Δ_E as in Definition 2.92, is equal to $16D$. This fact with Proposition 2.95 yields:

Corollary 2.96. *Let $E/\mathbb{Q}$ be an elliptic curve with coefficients in $\mathbb{Z}$. Let $p \geq 2$ be a prime. If E has bad reduction at p, then $p|\Delta_E$. In fact, if E is given by a minimal model, then $p|\Delta_E$ if and only if E has bad reduction at p.*

Let E be an elliptic curve over $\mathbb{Q}$ given by a minimal model

$$y^2 + a_1xy + a_3y = x^3 + a^2x^2 + a^4x + a_6$$

with integer coefficients. For a prime p of good reduction for $E/\mathbb{Q}$, we define N_p as the number of points in the reduction of the curve modulo p, i.e., the number of points in $E(\mathbb{F}_p)$. Also, let $a_p = p + 1 - N_p$. We define the local factor at p of the L-series to be $L_p(T) =:$

$$\begin{cases} 1 - a_pT + pT^2, \text{ if } E \text{ has good reduction at } p, \\ 1 - T, \text{ if } E \text{ has split multiplicative reduction at } p, \\ 1 + T, \text{ if } E \text{ has non-split multiplicative reduction at } p, \\ 1, \text{ if } E \text{ has additive reduction at } p. \end{cases} \tag{2.3}$$

Definition 2.97. *The L-function of the elliptic curve E is defined to be*

$$L(E, s) = \prod_{p \geq 2} \frac{1}{L_p(p^{-s})},$$

where the product is over all primes $p \geq 2$ and $L_p(T)$ is the local factor defined above.

Proposition 2.98. *Let $E/\mathbb{Q}$ be an elliptic curve, and let $L(E,s)$ be its L-function. Define Fourier coefficients a_n for all $n \geq 1$ as follows. Let $a_1 = 1$. If $p \geq 2$ is prime, we define $a_p =$:*

$$\begin{cases} p+1-N_p, \text{ if } E \text{ has good reduction at } p, \\ 1, \text{ if } E \text{ has split multiplicative reduction at } p, \\ -1, \text{ if } E \text{ has non-split multiplicative reduction at } p, \\ 0, \text{ if } E \text{ has additive reduction at } p. \end{cases} \tag{2.4}$$

If $n = p^r$ for some $r \geq 1$, we define a_{p^r} recursively using the relation

$$a_p.a_{p^r} = a_{p^{r+1}} + p.a_{p^{r-1}}$$

if $E/\mathbb{Q}$ has good reduction at p and $a_{p^r} = (a_p)^r$ if $E/\mathbb{Q}$ has bad reduction at p. Finally, if $(m,n) = 1$, then we define $a_{mn} = a_m.a_n$. Then the L-function of E can be written as the series

$$L(E,s) = \sum_{n \geq 1} \frac{a_n}{n^s}.$$

Conjecture 2.99 (First part of Birch and Swinnerton-Dyer conjecture)**.** *Let E be an elliptic curve over $\mathbb{Q}$, and let $L(E,s)$ be the L-function attached to E. Then $L(E,s)$ has a zero at $s = 1$ of order equal to the rank R_E of $E(\mathbb{Q})$. In other words, the Taylor expansion of $L(E,s)$ at $s = 1$ is of the form*

$$L(E,s) = c_0.(s-1)^{R_E} + c_1.(s-1)^{R_E+1} + \dots$$

where c_0 is a non-zero constant.

Theorem 2.100 (Gross-Zagier, Kolyvagin, [12], Chapter 5, Theorem 5.2.8)**.** *Let $E/\mathbb{Q}$ be an elliptic curve of algebraic rank R_E. Suppose that the analytic rank of $E/\mathbb{Q}$ is ≤ 1, i.e., $ord_{s=1}L(E,s) \leq 1$. Then the first part of Birch and Swinnerton-Dyer conjecture holds for $E/\mathbb{Q}$, i.e.,*

$$R_E = rank(E(\mathbb{Q})) = rank_{an}(E/\mathbb{Q}) = ord_{s=1}L(E,s).$$

Example 2.101. *Let $E : y^2 = x^3 - 13$. Then $E(\mathbb{Q})$ does not have any torsion point of order 2 by Theorem 2.91, (ii). If (x,y) is a torsion point of E, then by Theorem 2.91, y is an integer and y divides $4A^3+27B^2 = 27.13^2$. Then $y(P) = \pm 1, \pm 3, \pm 13, \pm 39$, which implies $x(P)^3 = 14, 22, 182, 1534$, respectively. Since $x(P)$ is an integer, $E(\mathbb{Q})_{tors}$ is trivial. Furthermore, the discriminant of E is $\Delta_E = -73008 = 2^4.3^3.13^2$ and E is a minimal model since powers of the primes in the discriminant do not exceed 12. Then by Corollary 2.96, E has bad reduction at primes $2, 3, 13$.*

(i) E has additive reduction at $p = 2$, since modulo 2 we can write the equation as $((y-1) - 0.(x-0))^2 - x^3$ and the unique slope is 0.

(ii) E has additive reduction at $p = 3$, since modulo 3 we can write the equation as $((y-0) - 0.(x-1))^2 - (x-1)^3$ and the unique slope is 0.

(iii) E has additive reduction at $p = 13$, since modulo 13 we can write the equation as $((y-0) - 0.(x-0))^2 - x^3$ and the unique slope is 0.

Then $$L(E,s) = \prod_{p \neq 2,3,13} \frac{1}{1 - a_p p^{-s} + p^{1-2s}} = \sum_{n \geq 1} \frac{a_n}{n^s} = 1 + \frac{a_5}{5^s} + \frac{a_7}{7^s} + \frac{a_{11}}{11^s} + \frac{a_{17}}{17^s} + \dots$$

Computing a few terms of the Taylor expansion of L-function about $s = 1$ and evaluating $L(E,1)$ and derivatives at $s = 1$ in Magma, we see that $L(E,1) = 0$ and the first derivative is non-zero. This suggests that $L(E,s)$ has a zero of order 1 at $s = 1$. Then by Conjecture 2.99 and Theorem 2.100, the rank of E is 1.

2.5. Height Function

For the study of rational and integral points on an algebraic variety, it is required to have a good size function to measure the "size" of a point. By good, we actually mean two essential properties for these functions. First, the function should have a finiteness property, in other words there should exist only finitely many points of bounded size. Second, it should reflect the underlying geometry of the variety and the arithmetic nature of the point. We introduce height functions here as an example of a good size function. We mostly benefit from Generalization of the ABC-conjecture by Nils Bruin and Diophantine Geometry: An Introduction by Silverman and Hindry. For more details and proofs, see [9] and [13].

We start with defining height functions for points in a projective space.

2.5.1. Heights on Projective Space

Let $P \in \mathbb{P}^n(\mathbb{Q})$. Then we can write $P = (x_0, x_1, \ldots, x_n)$ with $x_0, x_1, \ldots, x_n \in \mathbb{Z}$ and $gcd(x_0, x_1, \ldots, x_n) = 1$. We define the *height* of P as H(P)=max$\{|x_0|, |x_1|, \ldots, |x_n|\}$.

For any M, the set $\{P \in \mathbb{P}^n(\mathbb{Q}) | H(P) < M\}$ is finite, since there are finitely many integers $x \in \mathbb{Z}$ satisfying $|x| \leq M$.

We fix K as a number field for the rest and generalize in the following way:

Definition 2.102. *Let $P = (x_0, x_1, \ldots, x_n) \in \mathbb{P}^n(K)$. The height of P, relative to K, is*

$$H_K(P) = \prod_{v \in M_K} max\{|x_0|_v, |x_1|_v, \ldots, |x_n|_v\},$$

where v ranges over all normalized valuations of K, denoted as M_K.

We will denote $H_K(P)$ as $H(P)$ since we fixed K.

We also define *logarithmic height* $h(P) := \log H(P)$, also called as *additive* and the height H above is also called as the *multiplicative height*.

Lemma 2.103. *Let* $P = (x_0, x_1, \ldots, x_n) \in \mathbb{P}^n(K)$.

(i) The height $H(P)$ *is well-defined, independent of the choice of homogeneous coordinates for* P.

(ii) $H(P) \geq 1$ *for all* $P \in \mathbb{P}^n(K)$.

Proof. (i) Take another choice of coordinates for P, which is in the form of $(cx_0, cx_1, \ldots, cx_n)$ for some $c \in K^*$. Using the Product Formula 2.12 and the multiplicativity of absolute values, we get $\prod_{v \in M_K} \max\{|cx_0|_v, |cx_1|_v, \ldots, |cx_n|_v\} =$

$$\left(\prod_{v \in M_K} |c|_v \right) \left(\prod_{v \in M_K} \max\{|x_0|_v, |x_1|_v, \ldots, |x_n|_v\} \right)$$

$$= \prod_{v \in M_K} \max\{|x_0|_v, |x_1|_v, \ldots, |x_n|_v\}.$$

(ii) Choose homogeneous coordinates for P so that some coordinate equals 1. Then $H(P) \geq 1$ by definition.

□

Definition 2.104. *The height of an element* $\alpha \in K$ *is defined as the height of the corresponding projective point* $(\alpha, 1) \in \mathbb{P}^1(K)$:

$$H(\alpha) = \prod_{v \in M_K} max\{|\alpha|_v, 1\}.$$

Remark 2.105. *For all* $r \in K^*$, *we have* $H(r) = H(1/r) = H(r-1) + O(1)$. *The first equation follows from Lemma 2.103* (i) *and the second follows from Definition 2.1* (iii).

Remark 2.106. *Let $P \in \mathbb{P}^n(\overline{\mathbb{Q}})$ and let $\sigma \in G_{\mathbb{Q}}$, which is the Galois group on $\mathbb{P}^n(\overline{\mathbb{Q}})$. Then $H_{\sigma(K)}(\sigma(P)) = H_K(P)$, since the absolute value $\sigma(v) \in M_{\sigma(K)}$ is defined by $|\sigma(x)|_{\sigma(v)} = |x|_v$.*

The automorphism σ of $\overline{\mathbb{Q}}$ defines an isomorphism between K and $\sigma(K)$, and it likewise identifies the sets of absolute values on K and $\sigma(K)$.

Recall that the field of definition of a point $P = (x_0, x_1, \ldots, x_n) \in \mathbb{P}^n(\overline{\mathbb{Q}})$ is the field $\mathbb{Q}(P) = \mathbb{Q}(x_0/x_j, x_1/x_j, \ldots, x_n/x_j)$ for any j with $x_j \neq 0$.

Theorem 2.107. *For any number $M \leq 0$ and for any fixed number field K, the set $\{P \in \mathbb{P}^n(K) | H(P) \leq M\}$ is finite.*

Proof. Choose homogeneous coordinates for $P = (x_0, x_1, \ldots, x_n)$ so that some coordinate equals 1. Then for any absolute value v, we have

$$\max\{|x_0|_v, |x_1|_v, \ldots, |x_n|_v\} \geq \max\{|x_i|_v, 1\}$$

for all $0 \leq i \leq n$. Multiplying over all $v \in M_K$, we get $H(P) \geq H(x_i)$ for all $0 \leq i \leq n$. Hence it is enough to show that the set $\{x \in \overline{\mathbb{Q}} | H(x) \leq M\}$ is finite.

Let $x \in \overline{\mathbb{Q}}$ have degree d and let $K = \mathbb{Q}(x)$. Denote the conjugates of x over $\mathbb{Q}$ as $x_1, x_2, \ldots, x_d$. Let

$$f_x(t) = \prod_{j=1}^{d} (t - x_j)$$

be the minimal polynomial of x over $\mathbb{Q}$. We can write this minimal polynomial $f_x(t)$ so that its coefficients are composed of symmetric polynomials, say $s_r(x)$:

$$\sum_{r=0}^{d} (-1)^r s_r(x) t^{d-r}.$$

Then for any absolute value $v \in M_K$, using Definition 2.1 (iii), we have

$$|s_r(x)|_v = \left| \sum_{1 \le i_1 < \cdots < i_r \le d} x_{i_1} \dots x_{i_r} \right|_v$$

$$\le C_v \max_{1 \le i_1 < \cdots < i_r \le d} |x_{i_1} \dots x_{i_r}|_v \le C_v \max(|x_{i_1}|_v, \dots, |x_{i_r}|_v)^r,$$

where $C_v = 1$ when v is nonarchimedean and $C_v = \binom{d}{r}$ when v is archimedean. Then

$$\max\{|s_0(x)|_v, \dots, |s_d(x)|_v\} \le C \prod_{i=1}^{d} \max\{|x_i|_v, 1\}^d,$$

where we can take $C = 2^d \ge \sum\limits_{1 \le r \le d} \binom{d}{r}$. Multiplying over all absolute values $v \in M_K$, we estimate

$$H(s_0(x), \dots, s_d(x)) \le 2^d \prod_{i=1}^{d} H(x_i)^d.$$

Moreover, we can write $H(s_0(x), \dots, s_d(x)) \le 2^d H(x)^{d^2}$ by the Remark 2.106.

Finally, let $x \in \{x \in \overline{\mathbb{Q}} | H(x) \le M\}$. Then x is the root of a polynomial $f_x(t) \in \mathbb{Q}[t]$ whose coefficients $s_0, \dots, s_d$ satisfy $H_K(s_0(x), \dots, s_d(x)) \le 2^d M^{d^2}$. Since there are only finitely many points of bounded height in $\mathbb{P}^n(\mathbb{Q})$, there are also finitely many possibilities for the polynomial $f_x(t)$, which gives us only finitely many possibilities for x. Hence the set $\{x \in \overline{\mathbb{Q}} | H(x) \le M\}$ is finite. □

This property of height functions is an essential key for the proofs of many fundamental finiteness theorems in Diophantine geometry such as *Faltings' theorem*, *Siegel's theorem* and *Mordell-Weil theorem.*

Now we define a product. Let $x = (x_0, \ldots, x_n) \in \mathbb{P}^n(K)$ and $y = (y_0, \ldots, y_m) \in \mathbb{P}^m(K)$. We define

$$x \otimes y := (x_0y_0, \ldots, x_0y_m, \ldots, x_ny_0, \ldots, x_ny_m) \in \mathbb{P}^{(n+1)(m+1)-1}(K),$$

$x^r := ({x_0}^r, \ldots, {x_n}^r)$ and $x^{(r)} := x \otimes \cdots \otimes x$.

Proposition 2.108. *Let $x = (x_0, \ldots, x_n) \in \mathbb{P}^n(K)$ and $y = (y_0, \ldots, y_m) \in \mathbb{P}^m(K)$ be projective points. Then we have the following properties:*

(i) $H(x \otimes y) = H(x)H(y)$

(ii) $H(x^r) = H(x^{(r)}) = H(x)^r$

(iii) *Let $f : \mathbb{P}^n(K) \to \mathbb{P}^m(K)$ be a rational map of degree d. Then there exists an effectively computable constant $C = C_f$ such that for all $x \in \mathbb{P}^n(K)$ where f is regular, we have $H(f(x)) \leq CH(x)^d$.*

Proof. (i) $H(x \otimes y) = \prod_{v \in M_K} \max_{i,j}|x_iy_j|_v = \prod_{v \in M_K} \max_{i,j}|x_i|_v|y_j|_v$

$$= \prod_{v \in M_K} \max_i|x_i|_v \prod_{v \in M_K} \max_j|y_j|_v = H(x)H(y).$$

(ii) Let $|x_{i_0}|_v = \max_i|x_i|_v$, then ${|x_{i_0}|_v}^r = \max_{i_1,\ldots,i_r}|x_{i_1} \ldots x_{i_r}|_v$. Hence $H(x^r) = H(x)^r$. The second equality $H(x^{(r)}) = H(x)^r$ follows from (i).

(iii) Let $f = (f_0, \ldots, f_m)$ be a representation of f where f_i are homogeneous polynomials of degree d. Then for all $x \in \mathbb{P}^n(K)$ where f is regular, we can find a f_i so that $f_i(x) \neq 0$. We can write f_i as

$$f_i(x) = (c_{i,1}, \ldots, c_{i,t}).x^{(d)},$$

where t is the number of entries in the vector $x^{(d)}$, and the product . is the standard inner product of vectors. (We give an example below to understand this representation better, please see Example 2.110.) Let $c_i := (c_{i,1}, \ldots, c_{i,t})$. Then by Definition 2.1 (iii), since there exist t terms in the sum of monomials,

there exists a constant C_v such that

$$H(f(x)) = \prod_{v \in M_K} \max_i |c_i.x^{(d)}|_v \leq \prod_{v \in M_K} {C_v}^{t-1} \max_i (\max_j |c_{i,j}|_v |(x^{(d)})_j|_v),$$

where $(x^{(d)})_j$ denotes the $j-$th place. Then

$$\prod_{v \in M_K} {C_v}^{t-1} \max_i (\max_j |c_{i,j}|_v |(x^{(d)})_j|_v) \leq \left(\prod_{v \in M_K} {C_v}^{t-1} \right) H(c_0, \ldots, c_m) H(x)^d.$$

Here we get $C_v > 1$ for only finitely many $v \in M_K$, namely for archimedean absolute values, hence it is a finite product and by Product Formula 2.12, we have $\prod_{v \in M_K} {C_v}^{t-1} = 1$. Letting $C := H(c_0, \ldots, c_m)$, we are done.

□

Remark 2.109. *Linear transformations of $\mathbb{P}^n(K)$ change the logarithmic height by only a bounded function, i.e., $h(f(x)) \leq h(x) + O(1)$. In particular, logarithmic height is only dependent on the choice of the coordinates by a bounded function.*

Example 2.110. *Let $f : \mathbb{P}^2(K) \to \mathbb{P}^1(K)$ and let $f = (f_0, f_1)$ with $f_0(x_0, x_1, x_2) = {x_1}^2 + x_0 x_2$ and $f_1(x_0, x_1, x_2) = {x_2}^2$. One can write f_0 as*

$$f_0(x) = (c_{i,1}, \ldots, c_{i,9}).x^{(2)} =$$

$$(c_{i,1}, \ldots, c_{i,9}).(x_0x_0, x_0x_1, x_0x_2, x_1x_0, x_1x_1, x_1x_2, x_2x_0, x_2x_1, x_2x_2),$$

where $c_{03} = c_{05} = 1$ and the other coefficients are 0.

Let $x \in \mathbb{P}^2(K)$ where f is regular, say $f_0(x) \neq 0$. Then we have $H(f(x)) \leq CH(x)^2$ for some constant C.

Proposition 2.111. *Let F be a homogeneous polynomial over K of degree d in variables $X_0, \ldots, X_n$ with the coefficient of X_n is nonzero, i.e., the point $(0, \ldots, 0, 1)$ does not satisfy the equation of F. Let π be the projection from $(0, \ldots, 0, 1)$ defined by $\pi : \mathbb{P}^n \to \mathbb{P}^{n-1}$ with $(x_0, \ldots, x_n) \mapsto (x_0, \ldots, x_{n-1})$. Then there exists an effectively computable constant C dependent on F such that for any $x \in \mathbb{P}^n(K)$ on the hypersurface determined by F, we have*

$$h(x) - C \leq h(\pi(x)) \leq h(x).$$

Proof. For any $x \in \mathbb{P}^n(K)$ on the hypersurface determined by F, we have $F(x) = 0$. Then ${x_n}^d$ can be written as a linear combination of other monomials. The point $x^{(d-1)} \otimes \pi(x)$ contains all those monomials. Indeed, it contains all monomials of degree d in $n+1$ variables but with the degree of x_n strictly less than d. Since we can write ${x_n}^d$ using monomials in $x^{(d-1)} \otimes \pi(x)$, there exists some linear map determined by F sending $x^{(d-1)} \otimes \pi(x)$ to $x^{(d)}$. Then by Proposition 2.108 (iii), there exists a constant C such that

$$h(x^{(d)}) \leq h(x^{(d-1)} \otimes \pi(x)) + C.$$

Finally, using Proposition 2.108 (ii), we get $h(x) \leq h(\pi(x)) + C$. The second inequality $h(\pi(x)) \leq h(x)$ follows directly from the definition. □

We can generalize Proposition 2.111 for any projection from a point not on the hypersurface with a slight adjustment $h(x) - C_1 \leq h(\pi(x)) \leq h(x) + C_2$, since the logarithmic height depends on the choice of coordinates by a bounded function by Remark 2.109.

Theorem 2.112. *Let $f : \mathbb{P}^n(K) \to \mathbb{P}^m(K)$ be a morphism of degree d. Then there exist effectively computable constants C_1, C_2 such that for $x \in \mathbb{P}^n(K)$, we have*

$$h(f(x)) - C_1 \leq dh(x) \leq h(f(x)) + C_2.$$

Proof. We only need to show $dh(x) \leq h(f(x))+C_2$, since we showed the first inequality in Proposition 2.108. Let $f = (f_0, \ldots, f_m)$ be a representation of f where f_i are homogeneous polynomials of degree d. Let $p : \mathbb{P}^n(K) \to \mathbb{P}^t(K)$ with $x \mapsto x^{(d)}$. Then $f_i(x)$ are linear combinations of the coordinates of $p(x)$. Since f_i's have no common zero, we can write the morphism as $f = \pi \circ p$ where π is a combination of projections from points outside $f(\mathbb{P}^n)$ and linear transformations. Notice that we take projections from points outside $f(\mathbb{P}^n)$ in order to not lose any points, since in the end we project into $f(\mathbb{P}^n)$. Then by Proposition 2.111, we have $h(\pi(p(x))) \geq h(p(x)) - C_2$ for $x \in \mathbb{P}^n(K)$. Lastly, using Proposition 2.108 (ii) with $h(p(x)) = h(x^{(d)}) = dh(x)$, we get $h(f(x)) = h(\pi(p(x))) + C_2 \geq dh(x)$. □

2.5.2. Heights on Curves

Let V be a smooth projective variety defined over K. In order to satisfy the second essential property of height functions, we begin by defining a height function for each projective embedding of V. Then using the relationship between projective embeddings and divisors, we get an equivalence class of height functions for each divisor class on V. Therefore we get information about the rational points on V by taking from the structure of its divisor class group. This construction is called *Weil's Height Machine*.

We restrict ourselves to curves for the construction. Let C be a curve over K. For every morphism $f : C \to \mathbb{P}^n$, we have a *height* on C defined by $h_f := h \circ f$ where h is the logarithmic height on $\mathbb{P}^n(K)$.

For every very ample divisor, we have such a morphism, defined up to a linear transformation, since ϕ_D is regular and injective. Then by Theorem 2.112, we get a height function for every very ample divisor, defined up to a bounded function, $h_{C,D} = h \circ \phi_D + O(1)$. This determines the height function for very ample divisors.

Now since linearly equivalent divisors have the same associated morphism, they have the same associated height: if $D \sim D'$, then $h_D = h_{D'}$.

Moreover, by Fact 2.81, we have $L(D_1+D_2) = L(D_1)\otimes L(D_2)$ for divisors without base points, particularly for very ample divisors. Then for the associated morphisms, we have $\phi_{D_1+D_2} = \pi\circ(\phi_{D_1}\otimes\phi_{D_2})$ for some linear map π, which is regular on the image of C. Using Proposition 2.108 (i) and Proposition 2.111, we get $h_{D_1+D_2} = h_{D_1}+h_{D_2}+O(1)$, the additivity property.

We associate equivalence classes of height functions, i.e., height functions modulo bounded functions, first with all very ample divisors and then with all divisors by enforcing additivity, since any divisor can be written as the difference of very ample divisors by Fact 2.86.

Let $D \in Div(C)$, then using Fact 2.86, we write $D = D_1 - D_2$ where D_1, D_2 are very ample divisors. Then we define $h_{C,D}(P) = h_{C,D_1}(P) - h_{C,D_2}(P)$ for all $P \in C(\overline{K})$. This gives us a height function $h_{C,D}$ for every divisor D on every curve C.

Next we check the additivity property, which we already know for base point free divisors. Let D, E be arbitrary divisors and we write them as differences $D = D_1 - D_2$ and $E = E_1 - E_2$ of base point free divisors using Fact 2.86. Then $D_1 + E_1$ and $D_2 + E_2$ are base point free, so $h_{C,D+E} = h_{C,D_1+E_1} - h_{C,D_2+E_2} + O(1) = h_{C,D_1} + h_{C,E_1} - h_{C,D_2} - h_{C,E_2} + O(1) = h_{C,D} + h_{C,E} + O(1)$.

We are now ready to give Weil's construction that associates a height function to every divisor.

Theorem 2.113 (Weil's Height Machine, [13])**.** *Let K be a number field. For every smooth projective variety V/K there exists a map*

$$h_V : Div(V) \to \{\textit{functions } V(\overline{K}) \to \mathbb{R}\}$$

with the following properties:

- *(Normalization) Let $H \subset \mathbb{P}^n$ be a hyperplane, and let $h(P)$ be the absolute logarithmic height on $\mathbb{P}^n$. Then*

$$h_{\mathbb{P}^n,H}(P) = h(P) + O(1) \text{ for all } P \in \mathbb{P}^n(\overline{K}).$$

- *(Functoriality) Let $\phi : V \to W$ be a morphism and let $D \in Div(W)$. Then*

$$h_{V,\phi^* D}(P) = h_{W,D}(\phi(P)) + O(1) \text{ for all } P \in V(\overline{K}).$$

- *(Additivity) Let $D, E \in Div(V)$. Then*

$$h_{V,D+E}(P) = h_{V,D}(P) + h_{V,E}(P) + O(1) \text{ for all } P \in V(\overline{K}).$$

- *(Linear Equivalence) Let $D, E \in Div(V)$ with D linearly equivalent to E. Then*

$$h_{V,D}(P) = h_{V,E}(P) + O(1) \text{ for all } P \in V(\overline{K}).$$

- *(Positivity) Let $D \in Div(V)$ be an effective divisor, and let B be the base locus of the linear system $|D|$. Then*

$$h_{V,D}(P) \geq O(1) \text{ for all } P \in (V - B)(\overline{K}).$$

- *(Finiteness) Let $D \in Div(V)$ be ample. Then for every finite extension K'/K and every constant M, the set*

$$\{P \in V(K') | h_{V,D}(P) \leq M\}$$

  is finite.
- *(Uniqueness) The height functions $h_{V,D}$ are determined, up to $O(1)$, by normalization, functoriality just for embeddings $\phi : V \hookrightarrow \mathbb{P}^n$, and additivity.*

Remark 2.114. *If the variety V is not smooth, Weil's Height Machine is still valid, provided that we work with Cartier divisors instead of Weil divisors. For the theory of Cartier divisors, see [13].*

The height associated to the zero divisor of a rational function is essentially the same as the height function induces itself.

Proposition 2.115. *Let C be a curve over K and let f be a rational function on that curve. Let $(f)_0$ be the zero divisor of f. Then*

$$h(f(x)) = h\left(\frac{1}{f}(x)\right) = h_{(f)_0}(x) + O(1).$$

Proof. Since $(f)_0$ is positive, using Fact 2.112, we choose an m such that $m(f)_0$ is very ample. We have $\frac{1}{f^m} \in L(m(f)_0)$, since $\left(\frac{1}{f}\right)_\infty = (f)_0$. Moreover, $1 \in L(m(f)_0)$, since $(f)_0$ is positive. Now we take a basis $\left\{1, \frac{1}{f^m}, g_1, \ldots, g_s\right\}$ for $L(m(f)_0)$. None of g_i's has a pole of higher order than $\frac{1}{f^m}$, then $(0, 0, x_1, \ldots, x_s) \notin V$ for any x_i if V is identified with its embedding. By Proposition 2.111, we know that the projection onto the first two coordinates changes the height by a bounded function. Therefore,

$$h(f(x)) = h\left(1, \frac{1}{f}(x)\right) = h\left(1, \frac{1}{f}(x), g_1(x), \ldots, g_s(x)\right) + O(1).$$

Since the height associated to a divisor is defined via the projective embedding, we are done. □

Theorem 2.116 ([14], Page 45)**.** *Let V be a nonsingular projective variety over K, $c \in Pic^0(V)$, and c' an element of $Pic(V)$ which is ample. Then on $V(\overline{K})$,*

$$h_c \leq O(\sqrt{h_{c'}}) + O(1).$$

Theorem 2.117 ([15], Chapter 4, Corollary 3.5)**.** *Let C be a curve over K and let $D_1, D_2 \in Div_K(C)$. Then for every $\epsilon > 0$, there exist constants C_1, C_2 such that for all $P \in C$, it holds that*

$$(1-\epsilon)deg(D_2)h_{D_1}(P) + C_1 \leq deg(D_1)h_{D_2}(P) \leq (1+\epsilon)deg(D_2)h_{D_1}(P) + C_2.$$

2.6. Belyi's theorem

For curves over number fields, we have a theorem called *Belyi* showing that we can control where a function is ramified. Moreover, the proof of *Belyi* is effective, in other words, this function can be constructed explicitly.

Definition 2.118. *A smooth algebraic curve C is defined over a subfield K of $\mathbb{C}$ if it is isomorphic to the set of zeros, in some affine or projective space over $\mathbb{C}$, of a finite set of polynomials with coefficients in K. We call K a field of definition of C.*

Example 2.119. *The Fermat curve F_n, which is the zero locus of $x^n + y^n + z^n = 0$, is defined over $\mathbb{Q}$ for all $n > 0$.*

If a curve C is defined over $\overline{\mathbb{Q}}$, then finitely many coefficients of the defining polynomials of C all lie in some finite extension of $\mathbb{Q}$, i.e., in a number field. This gives us the lemma below.

Lemma 2.120. *An algebraic curve is defined over a number field if and only if it is defined over $\overline{\mathbb{Q}}$.*

Theorem 2.121 (Belyi, [16], Theorem 2.6.1)**.** *Let C be an algebraic curve. Then C is definable over $\overline{\mathbb{Q}}$ if and only if there exists a morphism $f : C \to \mathbb{P}^1$ unramified outside $\{0, 1, \infty\}$.*

For the proof of *Mordell conjecture* using *abc conjecture*, we merely need the "only if" part of *Belyi's theorem*, which results from a surprisingly simple construction.

Before proving *Belyi*, we prove another theorem, which is in fact *Belyi* restricted to the case when $C = \mathbb{P}^1$, and then *Belyi* easily follows.

Theorem 2.122. *If $S \subseteq \mathbb{P}^1$ is a finite set, then there is a map $\phi : \mathbb{P}^1 \to \mathbb{P}^1$, defined over $\overline{\mathbb{Q}}$, and ramified only over $0, 1, \infty$ such that $\phi(S) \subseteq \{0, 1, \infty\}$.*

Proof. Let $S \subseteq \mathbb{P}^1$ be a finite set and $\phi : \mathbb{P}^1 \to \mathbb{P}^1$ be a nonconstant morphism. Denote the set $S_\phi := \phi(S) \cup \{x \in \mathbb{P}^1 : \phi \text{ is ramified over } x\}$. Let $\alpha \in S$ be a nonrational point in $\overline{\mathbb{Q}}$ and let $\alpha^n + a_1\alpha^{n-1} + \cdots + a_n = 0$ be its minimal equation over $\mathbb{Q}$. Assume that n is the maximal degree over $\mathbb{Q}$ of the elements of S and the number of such elements with degree n over $\mathbb{Q}$ is p. We take $\phi(z) = z^n + a_1 z^{n-1} + \cdots + a_n$, which is defined over $\mathbb{Q}$, and consider the set

$$S_\phi = \phi(S) \cup \{x \in \mathbb{P}^1 : \phi \text{ is ramified over } x\} = \tag{2.5}$$

$$\phi(S) \cup \{\infty\} \cup \{\phi(z) : \phi'(z) = 0\} = \phi(S) \cup \{\infty\} \cup \phi(S') \tag{2.6}$$

where S' consists of $z \in \overline{\mathbb{Q}}$ such that $\phi'(z) = 0$, hence S' contains elements of degree $\leq n-1$. The points of $\phi(S)$ have degree $\leq n$ and the number of these elements of degree n is $\leq p-1$ since $\phi(\alpha) = 0$. Therefore, S_ϕ contains less elements of degree n than S. By repeating this step, S will contain only rational points eventually. We may assume that $\{0, 1, \infty\} \subseteq S$, because if $a \in S$, then $z \mapsto \dfrac{1}{z-a}$ is unramified and it maps a to ∞, in other words, it does not affect the critical values, just translate the value. Similarly, if $\{a, \infty\} \subseteq S$, the map $z \mapsto z - a$ is unramified and it maps a, ∞ to $0, \infty$. Finally, if $\{a, 0, \infty\} \subseteq S$, then $z \mapsto z/a$ is unramified and it maps $a, 0, \infty$ to $1, 0, \infty$.

Now let $\{0, 1, \infty, \beta\} \subseteq S$ and $|S| = n \geq 4$ with β be a rational number different from $0, 1, \infty$. We seek a map $\phi(z) = z^A(1-z)^B$ with $A, B \in \mathbb{Q}$, $A, B, A+B \neq 0$. Then $\phi(0), \phi(1), \phi(\infty) \in \{0, \infty\}$. Ramified points of the map are the points z such that $\phi'(z) = 0$, or $\dfrac{\phi'(z)}{\phi(z)} = A\dfrac{dz}{z} + B\dfrac{dz}{z-1} = 0$ since $x \neq 0, 1, \infty$, which gives $z = \dfrac{A}{A+B}$. Now choose A, B integers so that $\alpha = \dfrac{A}{A+B}$. Then

$$S_\phi = \phi(S) \cup \{z \in \mathbb{P}^1 : \phi \text{ is ramified over } z\} \subseteq \{0, \infty, \phi(\alpha)\} \cup \phi(S')$$

where $S' = S - \{0, 1, \infty, \alpha\}$. So $\phi(S')$ can have at most $n-4$ elements, and $|S_\phi| \leq n-1$. Therefore, S_ϕ contains less elements than S. By repeating this step, S will only contain $\{0, 1, \infty\}$ eventually. $\square$

Proof of Belyi 2.121. Let f be any nonconstant rational function $C \to \mathbb{P}^1$ defined over a number field $K \subset \overline{\mathbb{Q}}$. Let $S \subset \mathbb{P}^1(\overline{\mathbb{Q}})$ be the finite set of ramification points of f. Using Theorem 2.122, if we take $\pi = \phi \circ f$, we are done. □

Example 2.123. *Let $C : y^2 = x^7 - 1$ be defined over $\mathbb{Q}$. We need a meromorphic function on C, so we start with the projection*

$$C \to \mathbb{P}^1$$
$$(x, y) \mapsto x$$

which is ramified at points $({\xi_7}^k, 0)$, $k = 0, 1, \ldots, 6$ and (∞, ∞). Indeed, in the affine plane with $z = 1$, the map is only ramified where the polynomial of x vanishes, by Example 2.40 of Kummer extension, which are the 7th roots of unity. On the other hand, for the homogeneous polynomial $Y^2Z^5 = X^7 - Z^7$, when $Z = 0$, we have $X^7 = 0$. Hence, we get the ramified point $\infty = (0, 1, 0) \mapsto (1, 0) = \infty$.

We need to eliminate the critical values ${\xi_7}^k$, where $k = 1, \ldots, 6$. So we decompose this projection with the function $x \mapsto x^6 + x^5 + x^4 + x^3 + x^2 + x + 1$ which sends ${\xi_7}^k$ to the rational number 0 for $k = 1, \ldots, 6$, and $\infty \mapsto \infty$, $1 \mapsto 7$. For the affine plane with $w = 1$, we have the corresponding map $\mathbb{C}[x] \to \mathbb{C}[x]$. Let M be a maximal ideal of $\mathbb{C}[x]$. Using Proposition 2.35, M is ramified over $\mathbb{C}[x]$ if and only if $f'(\alpha) \in M$ where α is a root of f. Here since $f' = 1$, and $1 \notin M$ for any maximal ideal, the map is unramified. In the other affine plane with $x = 1$, we get $(1, w) \mapsto (w^6+w^5+w^4+w^3+w^2+w+1, w^6)$, or $w \mapsto \dfrac{w^6}{w^6 + w^5 + w^4 + w^3 + w^2 + w + 1}$. When $w = 0$, we have $0 \mapsto 0$ with $w^6 = 0$. Hence the map $\mathbb{P}^1 \to \mathbb{P}^1$ is only ramified over $(1, 0) = \infty$.

In order to deal with the resulting critical value 7, we now apply the function $x \mapsto x/7$, which is unramified, since the degree is 1, and $7, 0, \infty \mapsto 1, 0, \infty$.

The composition

$$f : (x, y) \mapsto x \mapsto x^6 + x^5 + x^4 + x^3 + x^2 + x + 1 \mapsto \frac{x^6 + x^5 + x^4 + x^3 + x^2 + x + 1}{7}$$

maps the first critical values $1, \infty, {\xi_7}^k$, *with* $k = 1, \ldots, 6$ *into* $\{0, 1, \infty\}$ *as* $1 \mapsto 1$, $\infty \mapsto \infty$, ${\xi_7}^k \mapsto 0$. *Hence we have a Belyi function* f.

Example 2.124. *Let* $C : (y - \sqrt{2})^2 = (x^5 - 1)$ *be defined over* $\mathbb{Q}(\sqrt{2})$. *We need a meromorphic function on* C, *so we start with the projection*

$$C \to \mathbb{P}^1$$
$$(x, y) \mapsto x$$

which is ramified at points $({\xi_5}^k, \sqrt{2})$, $k = 0, 1, \ldots, 4$ *and* (∞, ∞). *Indeed, in the affine plane with* $z = 1$, *the map is only ramified where the polynomial of* x *vanishes, by Example 2.40 of Kummer extension, which are the* 5*th roots of unity. On the other hand, for the homogeneous polynomial* $Y^2Z^3 - 2\sqrt{2}YZ^4 - X^5 + 3Z^5 = 0$, *when* $Z = 0$, *we have* $X^5 = 0$. *Hence, we get the ramified point* $\infty = (0, 1, 0) \mapsto (1, 0) = \infty$.

We need to eliminate the critical values ${\xi_5}^k$, *where* $k = 1, \ldots, 4$. *So we decompose this projection with the function* $x \mapsto x^4 + x^3 + x^2 + x + 1$ *which sends* ${\xi_5}^k$ *to the rational number* 0 *for* $k = 1, \ldots, 4$, *and* $\infty \mapsto \infty$, $1 \mapsto 5$. *For the affine plane with* $w = 1$, *we have the corresponding map* $\mathbb{C}[x] \to \mathbb{C}[x]$. *Let* M *be a maximal ideal of* $\mathbb{C}[x]$. *Using Proposition 2.35,* M *is ramified over* $\mathbb{C}[x]$ *if and only if* $f'(\alpha) \in M$ *where* α *is a root of* f. *Here since* $f' = 1$, *and* $1 \notin M$ *for any maximal ideal, the map is unramified. In the other affine plane with* $x = 1$, *we get* $(1, w) \mapsto (w^4 + w^3 + w^2 + w + 1, w^4)$, *or* $w \mapsto \dfrac{w^4}{w^4 + w^3 + w^2 + w + 1}$. *When* $w = 0$, *we have* $0 \mapsto 0$ *with* $w^4 = 0$. *Hence the map* $\mathbb{P}^1 \to \mathbb{P}^1$ *is only ramified over* $(1, 0) = \infty$.

In order to deal with the resulting critical value 5, *we now apply the function* $x \mapsto x/5$, *which is unramified, since the degree is* 1, *and* $5, 0, \infty \mapsto 1, 0, \infty$.

The composition

$$f : (x, y) \mapsto x \mapsto x^4 + x^3 + x^2 + x + 1 \mapsto \frac{x^4 + x^3 + x^2 + x + 1}{5}$$

maps the first critical values $1, \infty, {\xi_5}^k$, *with* $k = 1, \ldots, 4$ *into* $\{0, 1, \infty\}$ *as* $1 \mapsto 1$, $\infty \mapsto \infty$, ${\xi_5}^k \mapsto 0$. *Hence we have a Belyi function* f.

3. ABC CONJECTURE

In the first chapter into the main subject, we start by giving two equivalent statements of *abc conjecture* and a few examples of different cases.

Definition 3.1. *The radical of a positive integer n is the product of its distinct prime factors:*

$$Rad(n) = \prod_{p|n} p.$$

Conjecture 3.2 (abc conjecture, 1, [17], Page 27)**.** *For all $\epsilon > 0$, there exists a constant $\lambda_\epsilon > 0$ such that, if a, b and c in $\mathbb{Z}^+$ are relatively prime and satisfy $a + b = c$, then*

$$c = max(a, b, c) < \lambda_\epsilon Rad(abc)^{1+\epsilon}. \tag{3.1}$$

Conjecture 3.3 (abc conjecture, 2)**.** *For all $\epsilon > 0$, there exist only finitely many triples (a, b, c) of coprime positive integers with $a + b = c$ satisfying*

$$c = max(a, b, c) > Rad(abc)^{1+\epsilon}.$$

Both statements can be easily obtained from each other.

Conjecture 3.3 implies conjecture 3.2: For a given ϵ, if there are only finitely many (a, b, c) triples for which $Rad(abc)^{1+\epsilon} < c$, then there are only finitely many values of $c/Rad(abc)^{1+\epsilon}$, and we can choose the maximum such value, call it m. Letting $\lambda_\epsilon := 1 + m$, we are done.

Conjecture 3.2 implies conjecture 3.3: For a given ϵ, we have $c < \lambda_\epsilon Rad(abc)^{1+\epsilon}$. If $\lambda_\epsilon < 1$, it is trivial. If not, let $\epsilon' := \dfrac{\epsilon}{2}$. Then there exists a positive constant λ'_ϵ

such that $c < \lambda'_\epsilon Rad(abc)^{1+\epsilon'}$, or $\frac{c^{1/1+\epsilon'}}{(\lambda'_\epsilon)^{1/1+\epsilon'}} < Rad(abc)$. Note that $\frac{1}{1+\epsilon'} = \frac{1}{1+\epsilon} + \frac{\epsilon}{(1+\epsilon)(2+\epsilon)}$ and $\frac{\epsilon}{(1+\epsilon)(2+\epsilon)} = \left(\frac{1}{1+\epsilon'}\right)\left(\frac{\epsilon}{2+2\epsilon}\right)$ since $\epsilon' = \frac{\epsilon}{2}$. Thus for c large enough, i.e., $c^{\epsilon/2+2\epsilon} > \lambda'_\epsilon$, we have

$$Rad(abc) > \frac{c^{1/1+\epsilon'}}{(\lambda'_\epsilon)^{1/1+\epsilon'}} > \frac{c^{1/1+\epsilon'}}{\lambda'_\epsilon} = c^{1/1+\epsilon}\left(\frac{c^{\epsilon/2+2\epsilon}}{\lambda'_\epsilon}\right)^{1/1+\epsilon'} > c^{1/1+\epsilon},$$

hence $c < Rad(abc)^{1+\epsilon}$. Since λ'_ϵ is a finite number, there are only finitely many c for which $c^{\epsilon/2+2\epsilon} \leq \lambda'_\epsilon$. So there are only finitely many (a, b, c) triples satisfying $c > Rad(abc)^{1+\epsilon}$.

In the equation $a+b = c$, where a, b and c are relatively prime positive integers, the case $c \ll_\epsilon Rad(abc)^{1+\epsilon}$ happens most of the time, so we call it *usual*. Roughly speaking, if there are lots of primes on the left side of the equation, then we get only a few primes on the right side. Otherwise, we call it *unusual*. When we take the exponent of the radical of abc to be 1 in the conjecture, there are infinitely many exceptions, i.e., unusual cases, and we cannot find a constant that bounds c. However, when the exponent is > 1, there exist only finitely many exceptions, i.e., unusual cases. Hence one can list them.

We give a few examples of usual and unusual cases:

Example 3.4. *Take* $a = 256$, $b = 27$, *so* $c = 283$, *where* 283 *is prime. One can write as* $2^8 + 3^3 = 283$. *Then* $Rad(abc) = 2.3.283 = 1698$, *and we have* $283 < 1698$. *This is the usual case.*

Example 3.5. *Take* $a = 625 = 5^4$, $b = -1$, *so* $c = 624 = 2^4.39$. *Then* $Rad(abc) = 2.5.39 = 390$, *and we have* $625 > 390$. *This is unusual.*

Example 3.6. *Take* $a = 6561$, $b = 1$, $c = 6562$, *then* $3^8 + 1 = 2.17.193$. *Since there are prime numbers dividing the left hand side too many times, their presence is balanced out by larger primes dividing the right hand side only a few times.*

3.1. The Origin of Abc: Mason-Stothers Theorem

Mason-Stothers theorem is an analogy of *abc conjecture* with polynomials.

Theorem 3.7 (Mason-Stothers Theorem, [18])**.** *Let $f, g, h \in \mathbb{C}[t]$ be nonconstant relatively prime polynomials satisfying $f + g = h$. Then*

$$max(deg(f), deg(g), deg(h)) \leq n_0(fgh) - 1$$

where n_0 denotes the number of distinct roots.

Instead of the original proof of the theorem, we give the version of Synder's.

Lemma 3.8. *Let f be a nonzero polynomial in $\mathbb{C}[t]$, then $deg(f) = deg(f, f') + n_0(f)$.*

Proof. If f is constant, then the proof is trivial. So assume that f is not constant, and let β be a root of f. We write f as a sum of powers of $(t - \beta)$ since $(t - \beta)|f$.

$$f(t) = c_1(t - \beta)^{a_1} + \cdots + c_n(t - \beta)^{a_n}$$

where constants $c_i > 0$ and $a_i > 0$. Then order the multiplicities a_i in a decreasing manner so that a_n is the smallest.

Taking the derivative, we get $f'(t) = a_1c(t - \beta)^{a_1-1} + \cdots + a_nc(t - \beta)^{a_n-1}$, then $(t - \beta)^{a_n-1}$ is the largest power of $(t - \beta)$ dividing f'. Apply this to all of the roots of f.

Let $\beta_1, \ldots, \beta_r$ be distinct roots of f. Using the unique factorization in $\mathbb{C}[t]$, we write $f(t) = c(t - \beta_1)^{b_1} \ldots (t - \beta_r)^{b_r}$ with some $0 \neq c \in \mathbb{C}$. Since the only factors of degree 1 that could be shared by f and f' are of the form $(t - \beta_i)$ for some i, we write $(f, f') = c'(t - \beta_1)^{k_1} \ldots (t - \beta_r)^{k_r}$ for some $c' \in \mathbb{C}$ and $k_i \geq 0$, and in fact, $k_i = b_i - 1$ for each $1 \leq i \leq r$ by the reasoning above.

Then $deg(f) = b_1 + \cdots + b_r = (b_1 - 1) + \cdots + (b_r - 1) + r = deg(f, f') + n_0(f)$, and we are done. □

Proof of Mason-Stothers Theorem. We have $f + g = h$, by taking the derivative of each side we get $f' + g' = h'$. Then

$$f'g - fg' = f'(f + g) - f(f' + g') = f'h - fh'.$$

Now, we have three relations: $(f, f')|f'h - fh'$, $(g, g')|f'h - fh'$, and $(h, h')|f'g - fg'$. Since f, g, h are relatively prime, we get $(f, f').(g, g').(h, h')|f'g - fg'$. Then

$$deg(f, f') + deg(g, g') + deg(h, h') \leq deg(f'g - fg') \leq deg(f) + deg(g) - 1,$$

and adding $deg(h)$ to the leftmost and rightmost sides gives us

$$deg(f) - deg(f, f') + deg(g) - deg(g, g') + deg(h) - deg(h, h') - 1 \geq deg(h)$$

so $n_0(f) + n_0(g) + n_0(h) - 1 \geq deg(h)$. Finally, since they are relatively prime, we have $n_0(fgh) = n_0(f) + n_0(g) + n_0(h)$, so the result follows. □

One can also derive *Fermat's last theorem* for polynomials from *Mason-Stothers theorem* effortlessly.

4. HALL CONJECTURE

Consider the equation $x^3 - y^2 = k$, where x, y, and $k \in \mathbb{Z}$. This equation is known as *Mordell equation*. In the case of $k = 0$, we have infinitely many solutions which can be seen by parametrizing with (t^2, t^3), where t is an integer. However, when k is nonzero, we have only finitely many solutions by *Mordell's theorem* in 1922.

Theorem 4.1 (Mordell's theorem, [19])**.** *For each $k \in \mathbb{Z} - \{0\}$, the equation $y^2 = x^3 + k$ has finitely many integer solutions.*

Later in 1929, this finiteness was handled by a more general theorem about elliptic curves by Siegel. *Mordell curve* is, in fact, an elliptic curve, hence *Siegel's theorem* contains Mordell curves. For more details, see Section 2.4.

Theorem 4.2 (Siegel's Theorem)**.** *Every elliptic curve has only finitely many integer points.*

Before giving the statement of the main conjecture, *Hall conjecture*, on these solutions, we give a few examples on their computations.

Recall that for odd primes p, -1 is equivalent to a square mod p if and only if p is equivalent to 1 mod 4.

Example 4.3. *Let $E : y^2 = x^3 + 511$. Assume for a contradiction there exists an integral solution (x, y) on E. If x is even, then $y^2 \equiv 511 \equiv 3 \mod 4$, which is not possible. Hence x is odd, and y is even. Note that $x^3 \equiv -511 \equiv 1 \mod 4$, then $x \equiv 1 \mod 4$ since x is odd. Write the equation as:*

$$y^2 + 1 = x^3 + 512 = (x + 8)(x^2 - 8x + 64).$$

We have $x^2 - 8x + 64 \equiv 59 \equiv 3 \mod 4$, and $x^2 - 8x + 64 = (x - 4)^2 + 48 > 0$, then there exists a prime $p \equiv 3 \mod 4$ such that $p \mid x^2 - 8x + 64$, so $p \mid y^2 + 1$. It

follows that -1 *is a square* $\mod p$ *which can happen only when* p *is* $1 \mod 4$. *It is a contradiction. Hence the elliptic curve does not have any integral points other than the point at infinity* O.

Example 4.4. *The integral solutions of* $y^2 = x^3 + 4$ *are* $(0, \pm 2)$, *which can be seen easily by substituting* $x = 0$.

Example 4.5. *Let* $E : y^2 = x^3 - 13$. *If* x *is even, then* $y^2 \equiv -13 \equiv 3 \mod 4$, *which is a contradiction. So* x *is odd, and* y *is even. Then* $y = 2m$ *and* $x = 2n + 1$ *for some* $m, n \in \mathbb{Z}$. *Plug in, we get*

$$4m^2 = 8n^3 + 12n^2 + 6n - 12.$$

Looking at the equation mod 4, *we get* $n \equiv 0 \mod 4$, *hence* $n = 4k$ *for some* $k \in \mathbb{Z}$, *and* $x = 8k + 1$. *Again plug in, we get*

$$4m^2 = 512k^3 + 192k^2 + 24k - 12,$$

then divide the both sides by 4 :

$$m^2 = 128k^3 + 48k^2 + 6k - 3.$$

Looking at the equation mod 4, *we see that* $m^2 \equiv 2k + 1 \mod 4$. *The system is not solvable for* $m^2 \equiv 0 \mod 4$. *When* $m^2 \equiv 1 \mod 4$, *we have* $k \equiv 2 \mod 4$. *Trying the equations for* $k = 2$, *we get two integral points on our Mordell curve:* $(17, 70)$ *and* $(17, -70)$.

Moreover, these are the the only integral points on the curve. In Example 2.101, we find that the torsion group of E is an abelian group of order 1, and the rank of E is 1. So E has no torsion points other than the point at infinity O and $E(\mathbb{Q}) \cong \mathbb{Z}$. Then two rational points $(17, 70)$ and $(17, -70)$ are of infinite order. Moreover, $(17, 70)$ is a generator for the group $E(\mathbb{Q})$. That means one can find all of the rational points using the group structure with $(17, 70)$. Since it is too messy to do by hand, we use the

code *IntegralPoints(E)* on Magma, and this is exactly what Magma does to search for integral points; it adds the rational points and checks whether they are integral or not. However, Magma does not do this process forever, it uses some bounds on its search criteria to shorten the time.

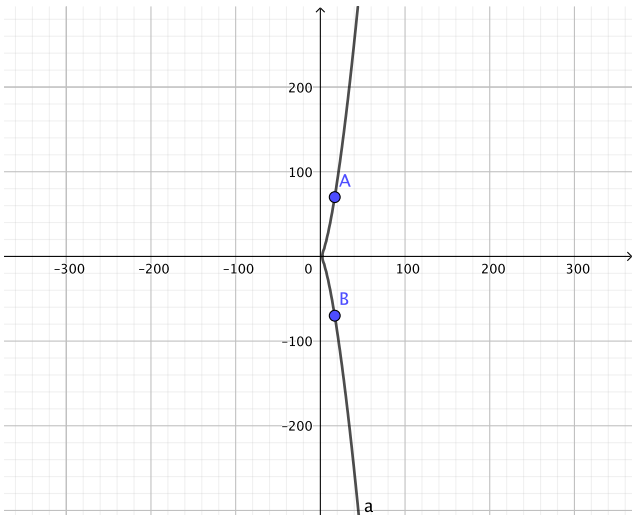

Figure 4.1. An example of Mordell curve

In Figure 4.1 above we see our example, the curve given by $y^2 = x^3 - 13$, with its two and only integral points.

There are many conjectures on integral solutions of Mordell equations or more generally elliptic curves, but the most famous one for Mordell equations is *Hall conjecture,* which was originally formulated by Marshall Hall, Jr. in 1970.

Conjecture 4.6 (Hall Conjecture, [19])**.** *There is a constant $C > 0$ such that if $y^2 = x^3 + k$ in $\mathbb{Z}$ with $k \neq 0$ then $|x| \leq C|k|^2$ and $|y| \leq C|k|^3$.*

Although it seems like a useful conjecture at first for making our job easier on listing the integral solutions on a Mordell curve, integral points can be quite large with respect to k. Very simple equations with small k values can have really large integral solutions.

Example 4.7. $447884928428402042307918^2 = 58538865167812223^3 - 1641843$, *found by Noam Elkies in* 1998.

4.1. Abc Implies Weaker Hall

We state a weaker version of *Hall conjecture* which can be derived from *abc conjecture*.

Conjecture 4.8 (Weak Hall Conjecture, [19])**.** *Pick $\epsilon > 0$. There is $C_\epsilon > 0$ such that for each $k \in \mathbb{Z} - \{0\}$, if (x, y) is a point on the curve $y^2 = x^3 + k$ in $\mathbb{Z}$ then $|x| \leq C_\epsilon |k|^{2(1+\epsilon)}$, $|y| \leq C_\epsilon |k|^{3(1+\epsilon)}$.*

Theorem 4.9. *Assume abc conjecture, then for all $\epsilon > 0$ there is $C_\epsilon > 0$ such that whenever $y^2 = x^3 + k$ in $\mathbb{Z}$ with $k \neq 0$,*

$$|x| \leq C_\epsilon |k|^{2(1+\epsilon)}, \;\; |y| \leq C_\epsilon |k|^{3(1+\epsilon)}.$$

Proof. Note that x and y cannot be both zero, since k is non-zero. If one of them is zero, then it is trivial. Hence assume $x, y \neq 0$. Let $n = gcd(x^3, y^2)$. Divide both sides of Mordell equation by n,

$$\frac{y^2}{n} = \frac{x^3}{n} + \frac{k}{n}.$$

Let $a = \dfrac{x^3}{n}$, $b = \dfrac{k}{n}$, and $c = \dfrac{y^2}{n}$. Now we have $c = a + b$, using *abc conjecture* we get two inequalities:

$$\frac{|x|^3}{n} \leq \lambda_\epsilon Rad(abc)^{1+\epsilon}, \quad \frac{|y|^2}{n} \leq \lambda_\epsilon Rad(abc)^{1+\epsilon}, \text{ so } |x|^3, |y|^2 \leq n\lambda_\epsilon Rad(abc)^{1+\epsilon}.$$

For the radical, we have $Rad(abc) \leq \prod_{p|ac} p . \prod_{p|b} p \leq |x||y|Rad(b) \leq |x||y|\dfrac{|k|}{n}$.

Plugging into above inequality, we get

$$|x|^3, |y|^2 \leq n\lambda_\epsilon \left(\frac{|x||y||k|}{n}\right)^{1+\epsilon} < \lambda_\epsilon (|x||y|)^{1+\epsilon} |k|^{1+\epsilon}.$$

Considering Mordell equation, we have either $|y|^2 \leq |x|^3$ or $|x|^3 \leq |y|^2$.

Case 1: If $|y|^2 \leq |x|^3$, then using $|y| \leq |x|^{3/2}$, we get $|x|^3 < \lambda_\epsilon(|x|)^{(5/2)(1+\epsilon)}|k|^{1+\epsilon}$,

$$\text{hence } |x|^{(1-5\epsilon)/2} < \lambda_\epsilon |k|^{1+\epsilon}.$$

Therefore for $0 < \epsilon < 1/5$, we have $|x| < {\lambda_\epsilon}^{2/(1-5\epsilon)}|k|^{2(1+\epsilon)/(1-5\epsilon)}$ and $|y| \leq |x|^{3/2} < {\lambda_\epsilon}^{3/(1-5\epsilon)}|k|^{3(1+\epsilon)/(1-5\epsilon)}$.

Case 2: If $|x|^3 \leq |y|^2$, then using $|x| \leq |y|^{2/3}$, we get $|y|^2 < \lambda_\epsilon(|y|)^{(5/3)(1+\epsilon)}|k|^{1+\epsilon}$,

$$\text{hence } |y|^{(1-5\epsilon)/3} < \lambda_\epsilon |k|^{1+\epsilon}.$$

Therefore for $0 < \epsilon < 1/5$, we have $|y| < {\lambda_\epsilon}^{3/(1-5\epsilon)}|k|^{3(1+\epsilon)/(1-5\epsilon)}$ and $|x| \leq |y|^{2/3} < {\lambda_\epsilon}^{2/(1-5\epsilon)}|k|^{2(1+\epsilon)/(1-5\epsilon)}$.

In conclusion we get the same bounds for both cases:

$$|x| < {\lambda_\epsilon}^{2/(1-5\epsilon)}|k|^{2(1+\epsilon)/(1-5\epsilon)},\ |y| < {\lambda_\epsilon}^{3/(1-5\epsilon)}|k|^{3(1+\epsilon)/(1-5\epsilon)}.$$

Let $\dfrac{1+\epsilon}{1-5\epsilon} = 1 + \epsilon'$. Then when $0 < \epsilon < 1/5$, we have $0 < \epsilon' < \infty$, and ϵ' is small if and only if ϵ is small. Finally setting $C_{\epsilon'} = max({\lambda_\epsilon}^{2/(1-5\epsilon)}, {\lambda_\epsilon}^{3/(1-5\epsilon)})$, we are done. $\square$

4.2. Abc Implies Fermat for Large Exponents

One can derive *Fermat's last theorem* for large exponents using *abc conjecture* easily.

Theorem 4.10 (Fermat's Last Theorem, [19])**.** *The equation* $x^n + y^n = z^n$*, where* x, y, z*, and* n *are integers, has no non-zero solutions for* $n > 2$*.*

Theorem 4.11. *Assume abc conjecture. Let* x, y, z, n *be positive integers such that* $gcd(x, y, z) = 1$ *and* $x^n + y^n = z^n$*. Then there is no solution for* $n > 6$*.*

Proof. Suppose $x^n + y^n = z^n$ with x, y, z relatively prime positive integers and $n \geq 3$. Let $\lambda_1 = 1$, then by *abc conjecture* we have, $z^n \leq Rad(x^n y^n z^n)^2$ where $Rad(x^n y^n z^n)^2 = Rad(xyz)^2$. Then $z^n \leq Rad(xyz)^2 \leq (xyz)^2 \leq z^6$, which shows that $n \leq 6$. Hence we have *Fermat's Last Theorem* for $n \geq 7$. □

5. MORDELL CONJECTURE

Let $f(x_1, x_2, \ldots, x_n)$ be a polynomial with coefficients in $\mathbb{Q}$. Finding rational solutions to $f = 0$ is equivalent to finding rational points on V, where V is the affine space given by f. The variety V will be $(n-1)$-dimensional, if f is not identically zero. When finitely many polynomials are taken into account, V is defined to be the common zero locus.

A smooth curve C as defined above may be thought of as a Riemann surface with finitely many punctures; after taking an appropriate compactification by filling these punctures, the resulting compact Riemann surface has a *topological genus*, which is essentially the number of holes. For simplicity, we will assume that our curve C is projective nonsingular curve and it is connected.

When the curve C has genus 0 or 1, it is possible for the curve to have infinitely many rational points. For the genus 0 case, it is easily determined by *Hasse Principle:*

Hasse Principle: [20] A genus 0 curve C has a rational point if and only if it has a solution over the real numbers $\mathbb{R}$ and the p-adic numbers $\mathbb{Q}_p$ for all primes p.

When the genus is 1, it is a big mystery. Genus 1 curves with a given rational point are known as elliptic curves. For more details, see Section 2.4.

On the other hand, curves with genus ≥ 2 cannot have infinitely many rational points.

Conjecture 5.1 (Mordell Conjecture, 1922)**.** *A curve of genus greater than* 1 *over the field* $\mathbb{Q}$ *of rational numbers has only finitely many rational points.*

Faltings proved *Mordell conjecture* and generalized it over number fields.

Theorem 5.2 (Faltings, 1983)**.** *If K is any number field and C is any curve of genus > 1 defined over K, then C has only a finite number of K-rational points.*

Example 5.3. *The curve defined by $x^n + y^n + 1$ has finitely many rational points when $n \geq 4$ by Example* 2.77.

6. ABC IMPLIES MORDELL

6.1. Adaptation and Examples

The aim is to prove *Mordell conjecture* using *abc conjecture.* Before giving the proof, we construct the necessary set-up.

Theorem 6.1 (Mordell conjecture 1922, Faltings theorem)**.** *Any curve of genus at least* 2 *over a number field* K *has only finitely many* K*-rational points.*

First we reconcile our terminology with the paper of Elkies [1], hence restate *abc conjecture.*

Conjecture 6.2 (abc conjecture)**.** *For any relatively prime nonzero* A, B, $C \in \mathbb{Z}$ *such that* $A + B + C = 0$, *and for all* $\epsilon > 0$,

$$N(A,B,C) \gg_\epsilon H(A,B,C)^{1-\epsilon}, \tag{6.1}$$

with the constant implied in $\gg$ *depending on* ϵ *but not on* A, B, C, *where*

$$N(A,B,C) = \prod_{p|ABC} p \ , \ p \text{ prime} \tag{6.2}$$

and

$$H(A,B,C) = max(|A|,|B|,|C|). \tag{6.3}$$

Note that here H stands for the *height*, and N for the *conductor* where both of them are usually known with definitions on curves. This is the first hint of the connection we make between the theory of Diophantine approximation and the theory of points on curves of high genera.

Example 6.3. *Let* $A = 128$, $B = -125$, *and* $C = -3$. *Then* $N(A, B, C) = 30$, *and* $H(A, B, C) = 128$. *By abc conjecture, there is a constant depending on* ϵ, *say* λ_ϵ, *so that* $128 \leq \lambda_\epsilon .30^{1+\epsilon}$. *As seen clearly in this example, the conjecture is not true in the form of* $H(A, B, C) \leq N(A, B, C)$, *in other words when simply* $\epsilon = 1$, *and* $\lambda_\epsilon = 1$. *There are infinitely many exceptions in this case. However, once the exponent is greater than* 1, *there exist only finitely many exceptions.*

The statement of *abc conjecture* is for integers, whereas the statement of *Mordell conjecture* is for number fields. So we give a more general construction of *abc conjecture*.

First, we state the conjecture over $\mathbb{Q}$ and hence remove the condition of being relatively prime with the new definitions of the height and the conductor:

Conjecture 6.4 (abc conjecture over $\mathbb{Q}$)**.** *For any nonzero* A, B, $C \in \mathbb{Q}$ *such that* $A + B + C = 0$, *and for all* $\epsilon > 0$,

$$N(A, B, C) \gg_\epsilon H(A, B, C)^{1-\epsilon}, \tag{6.4}$$

with the constant implied in $\gg$ *depending on* ϵ *but not on* A, B, C, *where*

$$H(A, B, C) = \prod_{v \in M_{\mathbb{Q}}} max(|A|_v, |B|_v, |C|_v), \tag{6.5}$$

and

$$N(A, B, C) = \prod_{p \in I} p, \tag{6.6}$$

where $I = \{p \, prime : max(|A|_p, |B|_p, |C|_p) > min(|A|_p, |B|_p, |C|_p)\}$.

Notation: The set of normalized valuations on $\mathbb{Q}$ is the set $M_{\mathbb{Q}}$ consisting of the archimedean absolute value $|\,|_\infty$ and the $p-$adic absolute values $|\,|_p$ for every prime p.

Note that Formulas 6.5 and 6.6 are finite products: Let $A = a/b \in \mathbb{Q} - \{0\}$ with $(a, b) = 1$ and $a, b \in \mathbb{Z}$. Then $|A|_p = 1$ unless $p \mid a$ or $p \mid b$. So there are finitely many primes that contribute to both of the products.

Moreover, the values of $N(A, B, C)$ and $H(A, B, C)$ do not change when A, B, and C are multiplied by a scalar in $\mathbb{Q}^*$ and they agree with Formulas 6.2 and 6.3 when A, B, and C are relatively prime integers. We prove these in the Proposition 6.5 below.

Proposition 6.5. *(i) The height N and the conductor N are scaling invariant.*

(ii) When A, B, C are relatively prime integers, Formulas 6.5 and 6.6 agree with Formulas 6.2 and 6.3, respectively.

Proof. (i) Let $\lambda \in \mathbb{Q}^*$. Using the Product Formula 2.10, we find that

$$H(\lambda A, \lambda B, \lambda C) = \prod_{v \in M_{\mathbb{Q}}} max(|\lambda A|_v, |\lambda B|_v, |\lambda C|_v)$$

$$= \left(\prod_{v \in M_{\mathbb{Q}}} \lambda \right) \left(\prod_{v \in M_{\mathbb{Q}}} max(|A|_v, |B|_v, |C|_v) \right) = \prod_{v \in M_{\mathbb{Q}}} max(|A|_v, |B|_v, |C|_v) = H(A, B, C).$$

For the conductor, we have $N(\lambda A, \lambda B, \lambda C) = \prod_{p \in I} p$, where

$$I = \{p \,\text{prime} : max(|\lambda A|_p, |\lambda B|_p, |\lambda C|_p) > min(|\lambda A|_p, |\lambda B|_p, |\lambda C|_p)\}.$$

By the multiplicativity of valuation, we can also write the set I as

$$I = \{p \,\text{prime} : (|\lambda|_p).max(|A|_p, |B|_p, |C|_p) > (|\lambda|_p).min(|A|_p, |B|_p, |C|_p)\}.$$

Hence, the primes contributing to the set I stay invariant and $N(\lambda A, \lambda B, \lambda C) = N(A, B, C)$.

(ii) Let A, B, C be relatively prime nonzero integers. First, looking at the height formula

$$H(A, B, C) = \prod_{v \in M_{\mathbb{Q}}} max(|A|_v, |B|_v, |C|_v),$$

the p-adic absolute values in $M_{\mathbb{Q}}$ affects the formula if the primes appear in the factorization of A, B, C. However, A, B, C are relatively prime, so none of them shares the same prime factor. Moreover, they are all integers, so their p-adic values are either less than 1 or equal to 1. In the end, we get 1 as the maximum value from each finite prime. Therefore, the height formula becomes $H(A, B, C) = max(|A|, |B|, |C|)$.

For the conductor formula, the only primes in the set I are the primes in the factorization of A, B, C, in other words $p \in I$ if and only if $p|ABC$. So the conductor formula becomes $N(A, B, C) = \prod_{p|ABC} p$. □

We give an example to show the consistency of the old formulas with the new formulas when relatively prime integers are taken.

Example 6.6. *Take $A = -9$, $B = 5$, $C = 4$. Applying the old Formulas 6.2 and 6.3, we get $H = 9$, and $N = 2.3.5 = 30$. Applying Formulas 6.6 and 6.5, we get:*

- *For the prime $p = 2$, we have $|A|_2 = 1$, $|B|_2 = 1$, $|C|_2 = 1/4$.*
- *For $p = 3$: $|A|_3 = 1/9$, $|B|_3 = 1$, and $|C|_3 = 1$.*
- *For $p = 5$: $|A|_5 = 1$, $|B|_5 = 1/5$, $|C|_5 = 1$.*
- *For all primes other than 2, 3 or 5, we have $|A|_p = |B|_p = |C|_p = 1$.*
- *For the usual real absolute value, we have $|A|_\infty = 9$, $|B|_\infty = 5$, and $|C|_\infty = 4$.*

To evaluate the height, we multiply the maximum values from each valuation: $H = 9$; and to evaluate the conductor, we multiply the primes at which the maximum value is strictly larger than the minimum value: $N = 2.3.5 = 30$.

Example 6.7. *We also give an example to show that the products are scaling invariant:*

(i) If λ has a new prime factor:

Let $\lambda = 7$ and take the set-up in Example 6.6, then $A = -63$, $B = 35$, $C = 28$. It is enough to check for the new prime and the usual absolute value since there is no change in the computations related to old primes.

We get $|A|_7 = |B|_7 = |C|_7 = 1/7$ and $|A|_\infty = 63$, $|B|_\infty = 35$, and $|C|_\infty = 28$, then we have $H = 63.\frac{1}{7} = 9$ and $N = 2.3.5 = 30$. For the new prime, whether it is in the nominator or in the denominator, the value $|\,|_\lambda$ is the same for all of them because we multiply A, B, and C with the same multiplicity of λ. Also $|\,|_\infty$ increases or decreases by the same multiplicity, hence ends up canceling any change in the height. Since they are all multiplied with the same factor, the conductor stays the same.

(ii) If λ is a prime which already exists in the prime factorization of one of A, B, or C : Let $\lambda = 3$, then for $A = -27$, $B = 15$, and $C = 12$, we get $|A|_3 = 1/27$, $|B|_3 = 1/3$, $|C|_3 = 1/3$ and $|A|_\infty = 27$, $|B|_\infty = 15$, and $|C|_\infty = 12$. So $H = 27.\frac{1}{3} = 9$ and $N = 2.3.5 = 30$. The effect of the $p-$adic valuation of the new prime λ and the effect of the usual absolute value cancel each other out again.

Example 6.8. *Take $A = \frac{3}{5}$, $B = \frac{-1}{6}$, and $C = \frac{-13}{30}$.*

- *For $p = 2$: $|A|_2 = 1$, $|B|_2 = 2$, $|C|_2 = 2$.*
- *For $p = 3$: $|A|_3 = 1/3$, $|B|_3 = 3$, and $|C|_3 = 3$.*
- *For $p = 5$: $|A|_5 = 5$, $|B|_5 = 1$, $|C|_5 = 5$.*
- *For $p = 13$: $|A|_{13} = 1$, $|B|_{13} = 1$, $|C|_{13} = 1/13$.*
- *$|A|_\infty = 3/5$, $|B|_\infty = 1/6$, and $|C|_\infty = 13/30$.*

Then $H = 2.3.5.\frac{3}{5} = 18$, $N = 2.3.5.13 = 390$ and we have $N = 390 > 18$ as usual.

Conjecture 6.9 (abc conjecture over K)**.** *Let K be a number field. For all $\epsilon > 0$, and for all nonzero A, B, $C \in K$ with $A + B + C = 0$,*

$$N(A, B, C) \gg_\epsilon H(A, B, C)^{1-\epsilon} \tag{6.7}$$

with the constant implied in $\gg$ depending on K and ϵ, but not on A, B, C, where

$$H(A, B, C) = \prod_{v \in M_K} max(|A|_v, |B|_v, |C|_v), \tag{6.8}$$

and $N(A, B, C)$ is the product of the absolute norms of all the finite primes $\mathcal{P}$ of K at which $max(|A|_\mathcal{P}, |B|_\mathcal{P}, |C|_\mathcal{P})$ strictly exceeds $min(|A|_\mathcal{P}, |B|_\mathcal{P}, |C|_\mathcal{P})$.

Notation: The set of normalized valuations on a number field K is the set M_K consisting of all absolute values on K whose restriction to $\mathbb{Q}$ is one of the normalized valuations on $\mathbb{Q}$. We write ${M_K}^\infty$ for the set of archimedean absolute values in M_K, and ${M_K}^0$ for the set of nonarchimedean absolute values in M_K.

Note that archimedean absolute value is also denoted as *infinite absolute value*, and nonarchimedean as *finite*.

H and N are finite products: there are finitely many nonarchimedean valuations, since there are finitely many prime ideals in the factorization of an ideal in a Dedekind domain; and finitely many archimedean valuations, because K is a finite extension of $\mathbb{Q}$, so the number of embeddings of K is finite.

Example 6.10. *Set $K = \mathbb{Q}(\sqrt{2})$ with $\mathcal{O}_K = \mathbb{Z}[\sqrt{2}]$. Let $A = \sqrt{2}$, $B = \frac{1}{2}$, and $C = -\sqrt{2} - \frac{1}{2}$.*

Archimedean: K has two real embeddings, say $\phi_1 : a + b\sqrt{2} \mapsto a + b\sqrt{2}$ and $\phi_2 : a + b\sqrt{2} \mapsto a - b\sqrt{2}$, hence two archimedean absolute values. First, we look at the image under the embedding, then take the usual real absolute value of the outcome.

$$|A|_{\phi_1} = \sqrt{2},\ |B|_{\phi_1} = \frac{1}{2},\ |C|_{\phi_1} = \sqrt{2} + \frac{1}{2}$$

$$|A|_{\phi_2} = \sqrt{2},\ |B|_{\phi_2} = \frac{1}{2},\ |C|_{\phi_2} = \sqrt{2} - \frac{1}{2}$$

Nonarchimedean: *We decompose principal fractional ideals into prime ideals in* $\mathcal{O}_K$*, then consider valuations up to these resulting primes.*

Note that the discriminant of K *is* 8*, so* 2 *is the only prime that ramifies in* K*.*

- $A = \sqrt{2}$ *creates the principal fractional ideal* $\langle\sqrt{2}\rangle = \sqrt{2}\mathcal{O}_K$*. We find its prime factorization via norm:* $\mathbb{N}(\sqrt{2}\mathcal{O}_K) = |\mathbb{N}_{K/\mathbb{Q}}(\sqrt{2})| = 2$*. We look at how* 2 *decomposes. Using the method in Example 2.34, we find that* $2\mathbb{Z}[\sqrt{2}] = \langle 2, \sqrt{2}\rangle^2$*, where* $\mathbb{N}(2\mathbb{Z}[\sqrt{2}]) = 4 = \mathbb{N}(\langle 2, \sqrt{2}\rangle)^2$*, so* $\mathbb{N}(\langle 2, \sqrt{2}\rangle) = 2$*, and* $\sqrt{2}\mathcal{O}_K = \langle 2, \sqrt{2}\rangle$*. Let* $\mathcal{P} = \langle 2, \sqrt{2}\rangle$*.* $|A|_{\mathcal{P}} = \frac{1}{2}$ *and* $|A|_{\mathcal{Q}} = 1$ *for all other primes* $\mathcal{Q}$*.*

- $B = \frac{1}{2}$ *creates* $\left\langle \frac{1}{2} \right\rangle = \frac{1}{2}\mathcal{O}_K$*. We compute its norm:* $\mathbb{N}\left(\frac{1}{2}\mathcal{O}_K\right) = \left|\mathbb{N}_{K/\mathbb{Q}}\left(\frac{1}{2}\right)\right| = \frac{1}{4}$*. Then* $\frac{1}{2}\mathcal{O}_K = \langle 2, \sqrt{2}\rangle^{-2}$*, so* $|B|_{\mathcal{P}} = \left(\frac{1}{2}\right)^{-2} = 4$ *and* $|B|_{\mathcal{Q}} = 1$ *for all other primes* $\mathcal{Q}$*.*

- $C = -\sqrt{2} - \frac{1}{2}$ *creates* $\left\langle -\sqrt{2} - \frac{1}{2} \right\rangle = \left(-\sqrt{2} - \frac{1}{2}\right)\mathcal{O}_K$*. We compute its norm:* $\mathbb{N}\left(\left(-\sqrt{2} - \frac{1}{2}\right)\mathcal{O}_K\right) = \left|\mathbb{N}_{K/\mathbb{Q}}\left(-\sqrt{2} - \frac{1}{2}\right)\right| = \frac{7}{4}$*. Using the method in Example 2.34 with* $p = 7$*, we see that* $7\mathcal{O}_K$ *is prime. Let* $7\mathcal{O}_K = \mathcal{P}_1$*, then* $\mathbb{N}(\mathcal{P}_1) = 49$*. Then the ideal of norm* $\frac{7}{4}$ *is* $\left\langle -\sqrt{2} - \frac{1}{2} \right\rangle = \mathcal{P}_1.\mathcal{P}^{-2}$*. Hence,* $|C|_{\mathcal{P}} = 4$*,* $|C|_{\mathcal{P}_1} = \frac{1}{7}$*, and* $|C|_{\mathcal{Q}} = 1$ *otherwise.*

In the end, multiplying these maximum values, we obtain $H = 8 + 2\sqrt{2}$*; and choosing the valuations where maximum value strictly exceeds minimum value out of finite primes, which is all of them in this case, we get* $N = 2.7 = 14$*. We have* $N = 14 > 8 + 2\sqrt{2} + H$ *as usual.*

Example 6.11. *Set $K = \mathbb{Q}(i)$ with $\mathcal{O}_K = \mathbb{Z}[i]$. Let $A = 3+2i$, $B = -5$, and $C = 2-2i$.*

Archimedean: K has two embeddings which are complex-conjuagates, $\sigma_1 : a+bi \mapsto a+bi$ and $\sigma_2 : a+bi \mapsto a-bi$, hence one archimedean absolute value. Let $\sigma = (\sigma_1, \sigma_2)$.

$$|A|_\sigma = |\sigma_1(A).\sigma_2(A)|_\infty = |(3+2i)(3-2i)|_\infty = 13,\ |B|_\sigma = 25,\ |C|_\sigma = 8.$$

Nonarchimedean: Note that the discriminant of K is -4, so 2 is the only prime that ramifies in K.

- *$A = 3+2i$ generates the principal fractional ideal $(3+2i)\mathcal{O}_K$ with norm $\mathbb{N}((3+2i)\mathcal{O}_K) = |\mathbb{N}_{K/\mathbb{Q}}(3+2i)| = 13$. Using the method in Example 2.34 with $p = 13$, we get the factorization $13\mathcal{O}_K = \langle 13, i+8\rangle\langle 13, i-8\rangle$. Let $\mathcal{P}_1 = \langle 13, i+8\rangle$ and $\mathcal{P}_2 = \langle 13, i-8\rangle$, where $\mathbb{N}(\mathcal{P}_1) = \mathbb{N}(\mathcal{P}_2) = 13$. Then we have either $(3+2i)\mathcal{O}_K = \mathcal{P}_1$ or $(3+2i)\mathcal{O}_K = \mathcal{P}_2$. In both cases, the same result follows: $|A|_{\mathcal{P}_1} = |A|_{\mathcal{P}_2} = \frac{1}{13}$. So we can take $\mathcal{P}_1$ into account without loss of generality. Then $|A|_{\mathcal{P}_1} = \frac{1}{13}$ and $|A|_{\mathcal{L}} = 1$ for all other primes $\mathcal{L}$.*

- *$B = -5$ generates $(-5)\mathcal{O}_K$ with norm $\mathbb{N}(-5\mathcal{O}_K) = |\mathbb{N}_{K/\mathbb{Q}}(-5)| = 25 = 5^2$. Using the method in Example 2.34 with $p = 5$, we get the decomposition $5\mathcal{O}_K = (-5)\mathcal{O}_K = \langle 5, i+2\rangle\langle 5, i-2\rangle$. Let $\mathcal{Q}_1 = \langle 5, i+2\rangle$ and $\mathcal{Q}_2 = \langle 5, i-2\rangle$, where $\mathbb{N}(\mathcal{Q}_1) = \mathbb{N}(\mathcal{Q}_2) = 5$. We get $|B|_{\mathcal{Q}_1} = |B|_{\mathcal{Q}_2} = \frac{1}{5}$. So take $\mathcal{Q}_1$ into account without loss of generality, and $|B|_{\mathcal{L}} = 1$ for all other primes $\mathcal{L}$.*

- *$C = 2-2i$ generates $(2-2i)\mathcal{O}_K$ with norm $\mathbb{N}((2-2i)\mathcal{O}_K) = |\mathbb{N}_{K/\mathbb{Q}}(2-2i)| = 8 = 2^3$. The prime 2 is ramified in K, and its factorization is $2\mathcal{O}_K = \langle 2, 1+i\rangle^2$ with $\mathbb{N}(\langle 2, 1+i\rangle) = 2$. Then $(2-2i)\mathcal{O}_K = \langle 2, 1+i\rangle^3$. Let $\mathfrak{P}$ denote the prime ideal $\langle 2, 1+i\rangle$, so $|C|_{\mathfrak{P}} = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$, and $|C|_{\mathcal{L}} = 1$ for all other primes $\mathcal{L}$.*

In the end, the only contribution to H comes from the archimedean absolute value for this case: $H = 25$, and $N = 2.5.13 = 130$. So we have $N > H$ as usual.

Example 6.12. *Let* $K = \mathbb{Q}(\sqrt{2}, i)$. *Let* $A = \dfrac{\sqrt{2}+3i}{2}$, $B = \dfrac{-5i}{3}$, *and* $C = \dfrac{-3\sqrt{2}+i}{6}$.

Archimedean: *K has four complex embeddings, which are*

$$\begin{cases} \phi_1 : a + b\sqrt{2} + ci + d\sqrt{2}i \mapsto a + b\sqrt{2} + ci + d\sqrt{2}i, \\ \phi_2 : a + b\sqrt{2} + ci + d\sqrt{2}i \mapsto a + b\sqrt{2} - ci - d\sqrt{2}i, \\ \phi_3 : a + b\sqrt{2} + ci + d\sqrt{2}i \mapsto a - b\sqrt{2} + ci - d\sqrt{2}i, \\ \phi_4 : a + b\sqrt{2} + ci + d\sqrt{2}i \mapsto a - b\sqrt{2} - ci + d\sqrt{2}i. \end{cases} \tag{6.9}$$

As seen clearly there are two pairs of complex-conjugate embeddings, say $(\phi_1, \phi_2) := \sigma_1$ *and* $(\phi_3, \phi_4) := \sigma_2$, *hence two archimedean absolute values.*

$$|A|_{\sigma_1} = |\phi_1(A).\phi_2(A)|_\infty = \Big|\Big(\frac{\sqrt{2}+3i}{2}\Big).\Big(\frac{\sqrt{2}-3i}{2}\Big)\Big|_\infty = \frac{11}{4},$$

$$|B|_{\sigma_1} = |\phi_1(B).\phi_2(B)|_\infty = \Big|\Big(\frac{-5i}{3}\Big)\Big(\frac{5i}{3}\Big)\Big|_\infty = \frac{25}{9},$$

$$|C|_{\sigma_1} = |\phi_1(C).\phi_2(C)|_\infty = \Big|\Big(\frac{-3\sqrt{2}+i}{6}\Big)\Big(\frac{-3\sqrt{2}-i}{6}\Big)\Big|_\infty = \frac{19}{36}.$$

$$|A|_{\sigma_2} = |\phi_3(A).\phi_4(A)|_\infty = \Big|\Big(\frac{-\sqrt{2}+3i}{2}\Big).\Big(\frac{-\sqrt{2}-3i}{2}\Big)\Big|_\infty = \frac{11}{4},$$

$$|B|_{\sigma_2} = |\phi_3(B).\phi_4(B)|_\infty = \Big|\Big(\frac{-5i}{3}\Big)\Big(\frac{5i}{3}\Big)\Big|_\infty = \frac{25}{9},$$

$$|C|_{\sigma_2} = |\phi_3(C).\phi_4(C)|_\infty = \Big|\Big(\frac{3\sqrt{2}+i}{6}\Big)\Big(\frac{3\sqrt{2}-i}{6}\Big)\Big|_\infty = \frac{19}{36}.$$

Nonarchimedean: *Again we decompose principal fractional ideals into prime ideals in $\mathcal{O}_K$ using norm, then we consider valuations up to primes that appear in the factorization.*

The discriminant of K is equal to $147456 = 2^{14}.3^2$. *Therefore the only primes ramifying in K are* 2 *and* 3.

Note that we can write K *as* $Q(\sqrt{2}+i)$ *by Primitive Element Theorem, then the minimal polynomial* $m(x)$ *of* $\sqrt{2}+i$ *over* $\mathbb{Q}$ *is* $m(x)=x^4-2x^2+9$.

- $A=\dfrac{\sqrt{2}+3i}{2}$ *creates the principal fractional ideal* $\left\langle \dfrac{\sqrt{2}+3i}{2} \right\rangle = \Big(\dfrac{\sqrt{2}+3i}{2}\Big)\mathcal{O}_K$.
  We compute the norm as $\Big|\mathbb{N}_{K/\mathbb{Q}}\Big(\dfrac{\sqrt{2}+3i}{2}\Big)\Big| = \dfrac{11^2}{2^4}$.
  In order to find the factorization into prime ideals, one can look at the minimal polynomial to use Proposition 2.38, since $\mathcal{O}_K$ *is monogenic by Example 2.37. The minimal polynomial* $m(x)=x^4-2x^2+9$ *has no solution mod* 11, *hence* $11\mathcal{O}_K=\mathcal{P}$ *with residual degree* 4 *and* $\mathbb{N}(\mathcal{P})=11^2$. *On the other hand,* $m(x)\equiv(x-1)^4 \mod 2$, *hence* $2\mathcal{O}_K=\mathcal{Q}^4$ *with* $\mathbb{N}(\mathcal{Q})=2$. *Then* $\Big(\dfrac{\sqrt{2}+3i}{2}\Big)\mathcal{O}_K=\mathcal{P}.\mathcal{Q}^{-4}$. *So, we get* $|A|_{\mathcal{P}}=\dfrac{1}{121}$, $|A|_{\mathcal{Q}}=16$ *and* 1 *for all the other primes.*

- $B=\dfrac{-5i}{3}$ *creates* $\left\langle \dfrac{-5i}{3} \right\rangle = \Big(\dfrac{5i}{3}\Big)\mathcal{O}_K$ *with the norm* $\mathbb{N}\Big(\dfrac{5i}{3}\mathcal{O}_K\Big)=\dfrac{5^4}{3^4}$.
  The minimal polynomial $m(x)=x^4-2x^2+9$ *factorizes mod* 5 *as* $m(x)\equiv(x^2-x+2)(x^2+x+2)$, *hence* $5\mathcal{O}_K=\mathcal{I}\mathcal{J}$ *with* $\mathbb{N}(\mathcal{I})=5^2$ *and* $\mathbb{N}(\mathcal{J})=5^2$, *where* $\mathcal{I}=\langle x^2-x+2\rangle$ *and* $\mathcal{J}=\langle x^2+x+2\rangle$. *However,* $m(x)$ *stays irreducible mod* 3, *so* 3 *stays prime above:* $3\mathcal{O}_K=\mathcal{H}$ *with* $\mathbb{N}(\mathcal{H})=3^4$. *Then* $\Big(\dfrac{5i}{3}\Big)\mathcal{O}_K=\mathcal{I}.\mathcal{J}.\mathcal{H}^{-1}$, *and* $|B|_{\mathcal{I}}=\Big(\dfrac{1}{25}\Big)$, $|B|_{\mathcal{J}}=\Big(\dfrac{1}{25}\Big)$, $|B|_{\mathcal{H}}=\Big(\dfrac{1}{81}\Big)^{-1}=81$, *and for all the other primes we get* 1.

- $C=\dfrac{-3\sqrt{2}+i}{6}$ *creates* $\Big(\dfrac{-3\sqrt{2}+i}{6}\Big)\mathcal{O}_K$ *with the norm* $\mathbb{N}\Big(\Big(\dfrac{-3\sqrt{2}+i}{6}\Big)\mathcal{O}_K\Big)=\dfrac{19^2}{2^4.3^4}$.
  The minimal polynomial $m(x)=x^4-2x^2+9$ *has no solution mod* 19, *so* $19\mathcal{O}_K=\mathcal{L}$ *with residual degree* 4 *and* $\mathbb{N}(\mathcal{L})=19^2$. *Then* $\Big(\dfrac{-3\sqrt{2}+i}{6}\Big)\mathcal{O}_K=\mathcal{L}.\mathcal{Q}^{-4}.\mathcal{H}^{-1}$. *We obtain* $|C|_{\mathcal{L}}=\dfrac{1}{271}$, $|C|_{\mathcal{Q}}=16$, $|C|_{\mathcal{H}}=81$, *and* 1 *for all the other primes.*

In the end, we have $N>H$ *with* $H=3600$ *and* $N=121.2.25.25.81.271=3320088750$.

Example 6.13. *An example of a Kummer but not a Galois extension: we take* $K = \mathbb{Q}(\sqrt[3]{2})$ *with* $A = \dfrac{1-\sqrt[3]{2}}{2}$, $B = \dfrac{\sqrt[3]{2}}{3}$, *and* $C = \dfrac{\sqrt[3]{2}-3}{6}$.

Note that the minimal polynomial of $\sqrt[3]{2}$ *over* $\mathbb{Q}$ *is* x^3-2, *and the conjugates of* $\sqrt[3]{2}$ *are* $\sqrt[3]{2}w$ *and* $\sqrt[3]{2}w^2$ *with* $w = e^{2\pi i/3}$.

An element in K *can be expressed as* $a + b\sqrt[3]{2} + c\sqrt[3]{4}$ *where* $a, b, c \in \mathbb{Q}$.

Archimedean: K *has three embeddings, which are*

$$\begin{cases} \phi_1 : a + b\sqrt[3]{2} + c\sqrt[3]{4} \mapsto a + b\sqrt[3]{2} + c\sqrt[3]{4}, \\ \phi_2 : a + b\sqrt[3]{2} + c\sqrt[3]{4} \mapsto a + b\sqrt[3]{2}w + c\sqrt[3]{4}w^2, \\ \phi_3 : a + b\sqrt[3]{2} + c\sqrt[3]{4} \mapsto a + b\sqrt[3]{2}w^2 + c\sqrt[3]{4}w. \end{cases} \tag{6.10}$$

We get two archimedean absolute values, ϕ_1 *and* $(\phi_2, \phi_3) := \sigma_1$. *Then*

$$|A|_{\phi_1} = |\phi_1(A)|_\infty = \frac{\sqrt[3]{2}-1}{2},$$

$$|B|_{\phi_1} = |\phi_1(B)|_\infty = \frac{\sqrt[3]{2}}{3},$$

$$|C|_{\phi_1} = |\phi_1(C)|_\infty = \frac{3-\sqrt[3]{2}}{6}.$$

$$|A|_{\sigma_1} = |\phi_2(A).\phi_2(A)|_\infty = \left|\left(\frac{1-\sqrt[3]{2}w}{2}\right).\left(\frac{1-\sqrt[3]{2}w^2}{2}\right)\right|_\infty = \frac{1+\sqrt[3]{2}+\sqrt[3]{4}}{4},$$

$$|B|_{\sigma_1} = |\phi_2(B).\phi_2(B)|_\infty = \left|\left(\frac{\sqrt[3]{2}w}{3}\right).\left(\frac{\sqrt[3]{2}w^2}{3}\right)\right|_\infty = \frac{\sqrt[3]{4}}{9},$$

$$|C|_{\sigma_1} = |\phi_2(C).\phi_2(C)|_\infty = \left|\left(\frac{\sqrt[3]{2}w-3}{6}\right).\left(\frac{\sqrt[3]{2}w^2-3}{6}\right)\right|_\infty = \frac{9+3\sqrt[3]{2}+\sqrt[3]{4}}{36}.$$

Nonarchimedean: The discriminant of K *is* -108, *so only primes that ramify in* K *are* 2 *and* 3.

Using the basis $\{1, \sqrt[3]{2}, \sqrt[3]{4}\}$, *the matrix representation for multiplication by* $\alpha = a + b\sqrt[3]{2} + c\sqrt[3]{4}$ *where* $a, b, c \in \mathbb{Q}$ *is obtained by multiplying* α *by* $1, \sqrt[3]{2}$, *and* $\sqrt[3]{4}$: $\alpha.1 = a + b\sqrt[3]{2} + c\sqrt[3]{4}$, $\alpha.\sqrt[3]{2} = 2c + a\sqrt[3]{2} + b\sqrt[3]{4}$, *and* $\alpha.\sqrt[3]{4} = 2b + 2c\sqrt[3]{2} + a\sqrt[3]{4}$, *then using these calculations we get* $[m_\alpha] =$

$$\begin{bmatrix} a & 2c & 2b \\ b & a & 2c \\ c & b & a \end{bmatrix}$$

Therefore the norm of an element is $\mathbb{N}_{K/\mathbb{Q}}(a + b\sqrt[3]{2} + c\sqrt[3]{4}) = a^3 + 2b^3 + 4c^3 - 6abc$.

- $A = \dfrac{1 - \sqrt[3]{2}}{2}$ *creates the principal deal* $\Big(\dfrac{1 - \sqrt[3]{2}}{2}\Big)\mathcal{O}_K$ *with the norm* $\dfrac{1}{8}$. *Since* $\mathcal{O}_K = \mathbb{Z}[\sqrt[3]{2}]$, *we can factorize into prime ideals by looking at the factorization of the minimal polynomial mod p. As* $x^3 - 2 \equiv x^3 \mod 2$, 2 *is totally ramified in* K, *and we have* $\Big(\dfrac{1 - \sqrt[3]{2}}{2}\Big)\mathcal{O}_K = \mathcal{P}^{-3}$ *with* $\mathbb{N}(\mathcal{P}) = 2$. *So,* $|A|_\mathcal{P} = 8$ *and for all the other primes we get* 1.

- $B = \dfrac{\sqrt[3]{2}}{3}$ *creates* $\Big(\dfrac{\sqrt[3]{2}}{3}\Big)\mathcal{O}_K$ *with the norm* $\dfrac{2}{3^3}$. *Since* $x^3 - 2 = (x-2)(x^2+2x+1) = (x+1)(x+1)^2 = (x+1)^3 \in F_3[x]$, *we have* $\Big(\dfrac{\sqrt[3]{2}}{3}\Big)\mathcal{O}_K = \mathcal{P}\mathcal{Q}^{-3}$ *with* $\mathbb{N}(\mathcal{Q}) = 3$. *So,* $|B|_\mathcal{P} = \dfrac{1}{2}$, $|B|_\mathcal{Q} = 27$, *and for all the other primes we get* 1.

- $C = \dfrac{\sqrt[3]{2} - 3}{6}$ *creates* $\Big(\dfrac{\sqrt[3]{2} - 3}{6}\Big)\mathcal{O}_K$ *with the norm* $\dfrac{5^2}{2^3.3^3}$. *Since* 5 *does not ramify, we have* $\Big(\dfrac{\sqrt[3]{2} - 3}{6}\Big)\mathcal{O}_K = \mathcal{P}^{-3}\mathcal{Q}^{-3}\mathcal{R}$ *with* $\mathbb{N}(\mathcal{R}) = 25$. *So,* $|C|_\mathcal{P} = 8$, $|C|_\mathcal{Q} = 27$, $|C|_\mathcal{R} = \dfrac{1}{25}$, *and for all the other primes we get* 1.

In the end, we have $N > H$ *with* $H = 27$ *and* $N = 30$.

Example 6.14. *An example of a Kummer, Galois extension: we take* $K = \mathbb{Q}(\alpha)$ *with* $\alpha := \sqrt[7]{6}$. *Let* $A = \dfrac{1-2\alpha}{3}$, $B = \dfrac{2\alpha}{3}$, *and* $C = \dfrac{-1}{3}$.

A basis for K *is* $\{1, \alpha, \alpha^2, \alpha^3, \alpha^4, \alpha^5, \alpha^6\}$. *Hence an element in* K *can be shown as* $a + b\alpha + c\alpha^2 + d\alpha^3 + e\alpha^4 + f\alpha^5 + g\alpha^6$ *with all coefficients in* $\mathbb{Q}$. *Moreover the minimal polynomial of* α *is* $x^7 - 6$, *so its conjugates are* $\alpha, \alpha w, \alpha w^2, \alpha w^3, \alpha w^4, \alpha w^5$, *and* αw^6 *where* $w = e^{2\pi i/7} = cos\dfrac{2\pi}{7} + i.sin\dfrac{2\pi}{7}$.

Archimedean: K *has seven embeddings:*

$$\begin{aligned}
\phi_1 &: a + b\alpha + c\alpha^2 + d\alpha^3 + e\alpha^4 + f\alpha^5 + g\alpha^6 \mapsto a + b\alpha + c\alpha^2 + d\alpha^3 + e\alpha^4 + f\alpha^5 + g\alpha^6,\\
\phi_2 &: a + b\alpha + c\alpha^2 + d\alpha^3 + e\alpha^4 + f\alpha^5 + g\alpha^6 \mapsto a + b\alpha w + c\alpha^2 w^2 + d\alpha^3 w^3 + e\alpha^4 w^4\\
&\qquad + f\alpha^5 w^5 + g\alpha^6 w^6,\\
\phi_3 &: a + b\alpha + c\alpha^2 + d\alpha^3 + e\alpha^4 + f\alpha^5 + g\alpha^6 \mapsto a + b\alpha w^2 + c\alpha^2 w^4 + d\alpha^3 w^6 + e\alpha^4 w\\
&\qquad + f\alpha^5 w^3 + g\alpha^6 w^5,\\
\phi_4 &: a + b\alpha + c\alpha^2 + d\alpha^3 + e\alpha^4 + f\alpha^5 + g\alpha^6 \mapsto a + b\alpha w^3 + c\alpha^2 w^6 + d\alpha^3 w^2 + e\alpha^4 w^5\\
&\qquad + f\alpha^5 w + g\alpha^6 w^4,\\
\phi_5 &: a + b\alpha + c\alpha^2 + d\alpha^3 + e\alpha^4 + f\alpha^5 + g\alpha^6 \mapsto a + b\alpha w^4 + c\alpha^2 w + d\alpha^3 w^5 + e\alpha^4 w^2\\
&\qquad + f\alpha^5 w^6 + g\alpha^6 w^3,\\
\phi_6 &: a + b\alpha + c\alpha^2 + d\alpha^3 + e\alpha^4 + f\alpha^5 + g\alpha^6 \mapsto a + b\alpha w^5 + c\alpha^2 w^3 + d\alpha^3 w + e\alpha^4 w^6\\
&\qquad + f\alpha^5 w^4 + g\alpha^6 w^2,\\
\phi_7 &: a + b\alpha + c\alpha^2 + d\alpha^3 + e\alpha^4 + f\alpha^5 + g\alpha^6 \mapsto a + b\alpha w^6 + c\alpha^2 w^5 + d\alpha^3 w^4 + e\alpha^4 w^3\\
&\qquad + f\alpha^5 w^2 + g\alpha^6 w,
\end{aligned} \tag{6.11}$$

We get four archimedean absolute values coming from one real embedding ϕ_1 *and three complex-conjugate pairs, which we denote as* $(\phi_2, \phi_7) := \sigma_1$, $(\phi_3, \phi_6) := \sigma_2$, $(\phi_4, \phi_5) := \sigma_3$. *Then*

$$|A|_{\phi_1} = |\phi_1(A)|_\infty = \frac{2\alpha - 1}{3}, \quad |B|_{\phi_1} = |\phi_1(B)|_\infty = \frac{2\alpha}{3}, \quad |C|_{\phi_1} = |\phi_1(C)|_\infty = \frac{1}{3}.$$

The maximum value for ϕ_1 *is* B.

$$|A|_{\sigma_1} = |\phi_2(A).\phi_7(A)|_\infty = \Big|\Big(\frac{1-2\alpha w}{3}\Big).\Big(\frac{1-2\alpha w^6}{3}\Big)\Big|_\infty = \frac{1 - 4\alpha\cos(2\pi/7) + 4\alpha^2}{9},$$

$$|B|_{\sigma_1} = |\phi_2(B).\phi_7(B)|_\infty = \Big|\Big(\frac{2\alpha w}{3}\Big).\Big(\frac{2\alpha w^6}{3}\Big)\Big|_\infty = \frac{4\alpha^2}{9},$$

$$|C|_{\sigma_1} = |\phi_2(C).\phi_7(C)|_\infty = \frac{1}{9}.$$

Since $\cos(2\pi/7)$ *is approximately equal to* 0.64, *the maximum value for* σ_1 *is* B.

$$|A|_{\sigma_2} = \Big|\Big(\frac{1-2\alpha w^2}{3}\Big).\Big(\frac{1-2\alpha w^5}{3}\Big)\Big|_\infty = \frac{1 - 4\alpha\cos(4\pi/7) + 4\alpha^2}{9},$$

$$|B|_{\sigma_2} = \frac{4\alpha^2}{9}, \quad |C|_{\sigma_2} = \frac{1}{9}.$$

Since $\cos(4\pi/7)$ *gives us a negative value, the maximum value for* σ_2 *is* A.

$$|A|_{\sigma_3} = \Big|\Big(\frac{1-2\alpha w^3}{3}\Big).\Big(\frac{1-2\alpha w^4}{3}\Big)\Big|_\infty = \frac{1 - 4\alpha\cos(6\pi/7) + 4\alpha^2}{9},$$

$$|B|_{\sigma_3} = \frac{4\alpha^2}{9}, \quad |C|_{\sigma_3} = \frac{1}{9}.$$

Since $\cos(6\pi/7)$ *gives us a negative value, the maximum value for* σ_2 *is* A.

<u>*Nonarchimedean:*</u> *Using the tools from Proposition 2.39, we find that* $\mathcal{O}_K = \mathbb{Z}[\alpha] = \mathbb{Z}[\sqrt[7]{6}]$. *Moreover, the only primes of* $\mathbb{Z}$ *that ramify in* $\mathbb{Z}[\sqrt[7]{6}]$ *are* $2, 3$ *and* 7. *Explicitly,* $2\mathbb{Z}[\sqrt[7]{6}] = \langle \alpha, 2\rangle^7$, *and* $3\mathbb{Z}[\sqrt[7]{6}] = \langle \alpha, 3\rangle^7$. *Let* $\mathcal{P} := \langle \alpha, 2\rangle$ *and* $\mathcal{Q} := \langle \alpha, 3\rangle$.

- $A = \dfrac{1-2\alpha}{3}$ *creates the principal fractional ideal* $\Big(\dfrac{1-2\alpha}{3}\Big)\mathcal{O}_K$ *with the norm* $\dfrac{-13.59}{3^7}$. *Then* $\Big(\dfrac{1-2\alpha}{3}\Big)\mathcal{O}_K = \mathcal{I}\mathcal{J}\mathcal{Q}^{-7}$ *with* $\mathbb{N}(\mathcal{I}) = 13$, $\mathbb{N}(\mathcal{J}) = 59$, *and* $\mathbb{N}(\mathcal{Q}) = 3$. *So,* $|A|_{\mathcal{I}} = \dfrac{1}{13}$, $|A|_{\mathcal{J}} = \dfrac{1}{59}$, $|A|_{\mathcal{Q}} = 3^7 = 2187$, *and for other primes we get* 1.

- $B = \frac{2\alpha}{3}$ *creates* $\left(\frac{2\alpha}{3}\right)\mathcal{O}_K$ *with the norm* $\frac{2^8}{3^6}$. *Then* $\left(\frac{2\alpha}{3}\right)\mathcal{O}_K = \mathcal{P}^8\mathcal{Q}^{-6}$ *with* $\mathbb{N}(\mathcal{P}) = 2$. *So,* $|B|_{\mathcal{P}} = \frac{1}{256}$, $|B|_{\mathcal{Q}} = 729$, *and for all the other primes we get* 1.

- $C = \frac{-1}{3}$ *creates* $\left(\frac{-1}{3}\right)\mathcal{O}_K$ *with norm* $\frac{-1}{3^7}$. *Then* $\left(\frac{-1}{3}\right)\mathcal{O}_K = \mathcal{Q}^{-7}$. *So,* $|C|_{\mathcal{Q}} = 2187$, *and for all the other primes we get* 1.

In the end, we have $N > H$ with $H = 8\alpha^3(1-4\alpha\cos(4\pi/7)+4\alpha^2)(1-4\alpha\cos(6\pi/7)+4\alpha^2)$ and $N = 13.59.2.3 = 4602$.

Remark: In fact, both of the formulas, H and N, depend on a ratio:

Divide the equation $A + B + C = 0$ by $-B$, then the values of H and N do not change, since H and N are scaling invariant. We get:

$$-A/B - 1 + (A + B)/B = 0$$

where $r := (-A/B)$, with $r \in \mathbb{P}^1K - \{0, 1, \infty\}$, since A, C, and B cannot be zero, respectively. Hence,

$$H(A, B, C) = H(-A/B, -1, A/B + 1) = H(r, -1, 1 - r),$$

$$N(A, B, C) = N(r, -1, 1 - r).$$

We simplify our notation:

Proposition 6.15. *Let r_1 be the number of real embeddings of K and r_2 be the number of the complex embeddings of K. Then*

$$H(A,B,C) = \prod_{v \in M_K} max(1, |r|_v) \times \kappa, \tag{6.12}$$

where $1 \leq \kappa \leq 2^{r_1+r_2}$.

Proof. Since $|-1|_v = |1|_v = 1$, we have

$$H(A,B,C) = H(r,-1,1-r) = \prod_{v \in M_K} max(1, |r|_v, |1-r|_v).$$

Now we divide the product into two parts: over all nonarchimedean absolute values and over all archimedean absolute values.

$$\prod_{v \in M_K} max(1, |r|_v, |1-r|_v) = \prod_{v \in {M_K}^0} max(1, |r|_v, |1-r|_v) \times \prod_{v \in {M_K}^\infty} max(1, |r|_v, |1-r|_v).$$

By definition, $|1-r|_v \leq max(1, |r|_v)$ for all nonarchimedean valuations, then we have $\prod_{v \in {M_K}^0} max(1, |r|_v, |1-r|_v) = \prod_{v \in {M_K}^0} max(1, |r|_v)$.

For the archimedean part of the product:

Case 1: If $|1-r|_v, |r|_v \leq 1$, then we take 1, and in this case $\kappa = 1$.

Case 2: If $|1-r|_v \leq |r|_v$, then we take $|r|_v$, and in this case $\kappa = 1$.

Case 3: If $1 < |r|_v < |1-r|_v$, then we multiply $\prod_{v \in {M_K}^\infty} max(1, |r|_v, |1-r|_v)$ by $\dfrac{|1-r|_v}{|r|_v}$ to make it equal with $\prod_{v \in {M_K}^\infty} max(1, |r|_v)$, hence $\kappa = \left(\dfrac{|1-r|_v}{|r|_v}\right)^m$ where m is the number of the archimedean absolute values satisfying the inequality.

Moreover, $\frac{|1-r|_v}{|r|_v} \leq \frac{1+|r|_v}{|r|_v} = 1+\frac{1}{|r|_v} < 2$. Since there are r_1+r_2 many archimedean absolute values in total, we have $2^0 = 1 \leq \kappa \leq 2^{r_1+r_2}$.

Case 4: If $|r|_v \leq 1 < |1-r|_v$, then $\kappa = (|1-r|_v)^m$ where m is the number of the archimedean absolute values satisfying the inequality. Since $|1-r|_v \leq 1+|r|_v \leq 2$, we are done. □

Further setting $M \geq (r_1+r_2).ln2$ with $|O(1)| \leq M$, we shortly write

$$H(A,B,C) = H(r) \times exp(O(1)), \tag{6.13}$$

where $H(r)$ denotes the *naive height* $\prod_{v \in M_K} max(1, |r|_v)$ by abuse notation.

Proposition 6.16. *$N(A,B,C)$ is the product of the absolute norms of all the finite primes of K at which r, $1/r$ or $r-1$ has a positive valuation.*

Proof. By definition $N(r,-1,1-r)$ is the product of the absolute norms of all the finite primes of K at which $max(1,|r|_{\mathcal{P}},|1-r|_{\mathcal{P}})$ strictly exceeds $min(1,|r|_{\mathcal{P}},|1-r|_{\mathcal{P}})\}$.

First, we check if the finite primes in the product N are the primes such that r, $1/r$ or $r-1$ has a positive valuation at, i.e., $ord_{\mathcal{P}} > 0$.

Since finite primes are nonarchimedean absolute values, we have $|1-r|_{\mathcal{P}} \leq max(1,|r|_{\mathcal{P}})$ for all primes $\mathcal{P}$, which gives us the cases below to consider:

(i) When $|1-r|_{\mathcal{P}} < max(1,|r|_{\mathcal{P}})$: those primes contribute to the product because of the strict inequality.

- If $max(1,|r|_{\mathcal{P}}) = 1$, then $|1-r|_{\mathcal{P}} < 1$ and $|r|_{\mathcal{P}} < 1$. So both r and $1-r$ have positive valuation at $\mathcal{P}$.
- If $max(1,|r|_{\mathcal{P}}) = |r|_{\mathcal{P}}$ and $|1-r|_{\mathcal{P}} < 1 < |r|_{\mathcal{P}}$, then $|r|_{\mathcal{P}} > 1$. So both $1/r$ and $1-r$ have positive valuation at $\mathcal{P}$.

- If $max(1, |r|_{\mathcal{P}}) = |r|_{\mathcal{P}}$ and $1 < |1 - r|_{\mathcal{P}} < |r|_{\mathcal{P}}$, then $1/r$ has a positive valuation at $\mathcal{P}$.

(ii) When $|1 - r|_{\mathcal{P}} = max(1, |r|_{\mathcal{P}})$:

- If $|1 - r|_{\mathcal{P}} = 1 > |r|_{\mathcal{P}}$, then $\mathcal{P} \in I$, and r has a positive valuation at $\mathcal{P}$.
- If $|1 - r|_{\mathcal{P}} = 1 = |r|_{\mathcal{P}}$, then $\mathcal{P} \notin I$.
- If $|1 - r|_{\mathcal{P}} = |r|_{\mathcal{P}} < 1$, then $\mathcal{P} \in I$, and $1/r$ has a positive valuation at $\mathcal{P}$.

For the other way around:

- If r has a positive valuation at $\mathcal{P}$, then $|r|_{\mathcal{P}} < 1$, hence $\mathcal{P} \in I$.
- If $r - 1$ has a positive valuation at $\mathcal{P}$, then $|r - 1|_{\mathcal{P}} = |1 - r|_{\mathcal{P}} < 1$, hence $\mathcal{P} \in I$.
- If $1/r$ has a positive valuation at $\mathcal{P}$, then $|1/r|_{\mathcal{P}} < 1$ or equivalently $|r|_{\mathcal{P}} > 1$, hence $\mathcal{P} \in I$.

□

We shortly denote as $N(r) := N(r, -1, 1-r)$, and factor $N(r) = N_0(r).N_1(r).N_\infty(r)$ as the product of absolute norms of the prime ideals containing r, $r-1$, $1/r$ respectively.

6.2. Motivation for the Proof

The motivation for the proof of Elkies is the Fermat bound obtained by *abc conjecture*. As in the set-up of *abc conjecture*, we start with integers.

Proposition 6.17. *Assume abc conjecture holds, then there are only finitely many integer solutions to* $x^n + y^n + z^n = 0$ *with* $gcd(x, y, z) = 1$ *and* $n \geq 3$.

Proof. Let $n \geq 3$ and x, y, z be relatively prime, nonzero integers such that $x^n + y^n + z^n = 0$. Taking $(A, B, C) = (x^n, y^n, z^n)$, we have

$$N(A, B, C) = N(x^n, y^n, z^n) = N(x, y, z) \leq |xyz| < max\{|x|, |y|, |z|\}^3 = H(x^n, y^n, z^n)^{3/n}, \tag{6.14}$$

or $N(A, B, C) < H(A, B, C)^{3/n}$, hence we get a contradiction with *abc conjecture* for all $\epsilon < 1 - (3/n)$ once $H(A, B, C)$ is large enough. Therefore there are only finitely many (A, B, C) satisfying the equation. As another way to see this, we plug the inequality we obtanied $N(A, B, C) < H(A, B, C)^{3/n}$ into *abc conjecture*, then we get $H(A, B, C)^{1-\epsilon} < C_\epsilon H(A, B, C)^{3/n}$ or $H(A, B, C) < {C_\epsilon}^{1-\epsilon-3/n}$, for some constant C_ϵ, which gives us an effective height bound, limiting the possibilities to a finite number. □

Next, we carry this set-up to rationals by considering the ratios of variables:

Proposition 6.18. *Assume abc conjecture holds, then there are only finitely many rational solutions to* $x^n + y^n + z^n = 0$ *with* $n \geq 3$.

Proof. Let $(x, y, z) \in \mathbb{P}^2(\mathbb{Q})$, then we can take x, y, z to be relatively prime integers. We divide the equation $x^n + y^n + z^n = 0$ by $-y^n$, then we get $-(x/y)^n - 1 + (x^n + y^n/y^n) = 0$. Let $r := -(x/y)^n \in \mathbb{Q}$, then $1 - r = 1 + (x/y)^n = -(z/y)^n$ and $1/r = -(y/x)^n$. Note that all fractions are in lowest terms because x, y, z are relatively prime. Then we are looking for points $(r, -1, 1 - r)$ with $r \in \mathbb{Q}$.

Since N is scaling invariant by Proposition 6.5, $N(A,B,C) = N(r,-1,1-r) = N_0(r).N_1(r).N_\infty(r)$ using Proposition 6.16. Explicitly,

$$N_0(r): \text{ product of primes } p \text{ such that } r \equiv 0 \mod p \text{ or } -(x/y)^n \equiv 0 \mod p \tag{6.15}$$

$$N_1(r): \text{ product of primes } p \text{ such that } 1-r \equiv 0 \mod p \text{ or } r \equiv 1 \mod p \tag{6.16}$$

$$N_\infty(r): \text{ product of primes } p \text{ such that } 1/r \equiv 0 \mod p \text{ or ``} r \equiv \infty \mod p\text{''}. \tag{6.17}$$

Hence $N(r,-1,1-r)$ becomes the product of primes p such that $r \mod p$ is one of $0,1,\infty$, in other words, primes dividing x,y,z, respectively for $N_0(r), N_1(r), N_\infty(r)$.

Without loss of generality, assume x is positive and $max\{|x|_\infty, |y|_\infty, |z|_\infty\} = x$. For the height, we have

$$H(r,-1,1-r) = H(-x^n/y^n, -1, -y^n/x^n) = \prod_{v \in M_{\mathbb{Q}}} max||r||_v, 1, ||1-r||_v.$$

We get $|-x^n/y^n|_\infty$ for archimedean valuation from our assumption. For the finite primes, we consider the primes dividing denominators; otherwise we choose 1 as maximum. Hence multiplying for all nonarchimedean valuations, we get $x^n y^n$. Then $H(r,-1,1-r) = x^{2n}$.

We obtain the inequality $N(r,-1,1-r) = N_0(r).N_1(r).N_\infty(r) \leq |xyz| < x^3 = H(r,-1,1-r)^{3/2n}$, where $3/2n < 1$ since $n \geq 3$. Then $N < H^{3/2n}$, so $(r,-1,1-r)$ gives a counter example to abc conjecture over $\mathbb{Q}$ for $\epsilon < 1-(3/2n)$ once H is large enough, i.e., for all but finitely many $(r,-1,1-r)$.

□

In the final step of set-up, we take the problem into consideration in a geometric manner. We take (x, y, z) as a rational point on the Fermat curve and we build the rational function $f = (-x/y)^n$ on the Fermat curve, then one can interpret as that N reflects the ramification of the map above $0, 1, \infty$ and H represents the degree of the map. Hence, we first compute the degree of the map f and the number of preimages of f corresponding the points in the set $\{0, 1, \infty\}$, which will be useful in the proof.

Proposition 6.19. *Let F_n be the zero locus of $x^n + y^n + z^n = 0$ and let $f : F_n(\overline{\mathbb{Q}}) \to \mathbb{P}^1(\mathbb{C})$ with $(x, y, z) \mapsto (-x^n, y^n)$. Then the degree of f is n^2.*

Proof. On the affine part of the curve F_n given by $y = 1$ and $x^n + z^n + 1 = 0$, we have $f : (x, 1, z) \mapsto (-x^n, 1)$, which induces the map

$$f^* : \mathbb{C}[x] \to C_{F_n} := \mathbb{C}[x, z]/\langle z^n + x^n + 1\rangle \text{ with } g \mapsto g \circ f + \langle z^n + x^n + 1\rangle.$$

Then $deg\, f = [C_{F_n} : f^*(\mathbb{C}[x])] = [C_{F_n} : \mathbb{C}[x]].[\mathbb{C}[x] : f^*(\mathbb{C}[x])]$.

Since $z^n + x^n + 1$ is irreducible and monic in z,

$$[C_{F_n} : \mathbb{C}[x]] = [\mathbb{C}[x, z]/\langle z^n + x^n + 1\rangle : \mathbb{C}[x]] = n.$$

Moreover, as a $\mathbb{C}[x^n]-$vector space, $\mathbb{C}[x]$ has dimension n since $f^* : x \mapsto x^n$, and $f^* : a \mapsto a$ for all $a \in \mathbb{C}$. So

$$[\mathbb{C}[x] : f^*(\mathbb{C}[x])] = [\mathbb{C}[x] : \mathbb{C}[x^n]] = n.$$

Therefore, we get $deg\, f = n^2$. □

Proposition 6.20. *Let F_n be the zero locus of $x^n + y^n + z^n = 0$ and let $f : F_n(\overline{\mathbb{Q}}) \to \mathbb{P}^1(\mathbb{C})$ with $(x, y, z) \mapsto (-x^n, y^n)$. Then f is ramified over $0, 1, \infty$ and $|P \in C(\overline{\mathbb{Q}}) : f(P) \in \{0, 1, \infty\}| = 3n$.*

Proof.
- The value $0 \in \mathbb{C}$ is represented by a pair $(-x^n, y^n)$ where $-x^n = 0$ with $y \neq 0$. Looking at the affine plane with $y = 1$, we have $x = 0$ and since $x^n + z^n = -1$, there are n distinct solutions for z, namely nth roots of -1. Hence there are n preimages for the value 0.
- The value $1 \in \mathbb{C}$ is represented by a pair $(-x^n, y^n)$ where $-x^n = y^n$ with $y \neq 0$. In the affine plane with $y = 1$, when $-(x/y)^n = 1$, we have $z = 0$ and hence there are n distinct solutions for x, which are nth roots of -1. Hence there are n preimages for the value 1.
- The value $\infty \in \mathbb{P}^1(\mathbb{C})$ is represented by a pair $(-x^n, y^n)$ where $y = 0$, while $-x^n \neq 0$. Then looking at the affine plane with $x = 1$, we have $z^n = -1$, hence there are n distinct solutions for z, namely nth roots of -1. So there are n preimages for the value ∞.

We see that f is ramified above $0, 1, \infty$ since the number of the preimages for these points is less than the degree of the map. □

Definition 6.21 ([11], Section 3.1)**.** *The field of maps $f : C(\mathbb{C}) \to \mathbb{P}^1(\mathbb{C})$, has the valuations $v_x(f) = -ord_x(f)$ for each point of $x \in C(\mathbb{C})$. For a nonconstant map $f : C(\mathbb{C}) \to \mathbb{P}'(\mathbb{C})$, we define the height and the radical of $P = (f, 1-f, 1) \in \mathbb{P}^2(C(\mathbb{C}))$ by $h(P) = deg\, f$, and $\log N(P) = |f^{-1}(\{0, 1, \infty\})|$.*

Proposition 6.22. *Let F_n be the zero locus of $x^n + y^n + z^n = 0$ and let $f : F_n(\overline{\mathbb{Q}}) \to \mathbb{P}^1(\mathbb{C})$ with $(x, y, z) \mapsto (-x^n, y^n)$. Then F_n has finitely many rational points.*

Proof. Using Proposition 6.19 and Proposition 6.20, we have

$$3n = |P \in C(\overline{\mathbb{Q}}) : f(P) \in \{0, 1, \infty\}| < deg\, f = n^2$$

since $n > 3$. So *abc conjecture* is regarded as a bound on the "ramification" of the rational number r above $0, 1$, and ∞. Since we have $\log N < \log H$, we conclude that the Fermat curve has only finitely many rational points. □

In general, it is rare for there to be a rational function f on a curve C whose degree exceeds the size of $f^{-1}(\{0,1,\infty\})$ by a large factor. As a result, we expect that any curve C over an arbitrary number field K, provided there is a rational function $f \in K(C)$ such that

$$|P \in C(\overline{\mathbb{Q}}) : f(P) \in \{0,1,\infty\}| < deg\, f,$$

has only finitely many $K-$rational points.

Remark 6.23. *When the curve has genus* 0, *by Hurwitz Formula 2.74, we have*

$$-2 = -2deg\, f + \sum_{Q} (deg\, f - |f^{-1}(Q)|).$$

By separating the ramification points, we get

$$-2 + 2deg\, f = \sum_{Q \in \{0,1,\infty\}} (deg\, f - |f^{-1}(Q)|) + \sum_{Q \notin \{0,1,\infty\}} (deg\, f - |f^{-1}(Q)|)$$

$$= 3deg\, f - |f^{-1}(\{0,1,\infty\})| + \sum_{Q \notin \{0,1,\infty\}} (deg\, f - |f^{-1}(Q)|),$$

since there may be ramification points other than $0, 1,$ *and* ∞. *Then,*

$$|f^{-1}(\{0,1,\infty\})| = deg\, f + 2 + \sum_{Q \notin \{0,1,\infty\}} (deg\, f - |f^{-1}(Q)|) > deg\, f.$$

Again by Hurwitz Formula 2.74, when the curve has genus 1, *we have*

$$0 = -2deg\, f + \sum_{Q \in \{0,1,\infty\}} (deg\, f - |f^{-1}(Q)|) + \sum_{Q \notin \{0,1,\infty\}} (deg\, f - |f^{-1}(Q)|),$$

then

$$|f^{-1}(\{0,1,\infty\})| = deg\, f + \sum_{Q \notin \{0,1,\infty\}} (deg\, f - |f^{-1}(Q)|) \geq deg\, f.$$

Therefore, we cannot find such $f \in K(C)$ when C has genus < 2.

Proposition 6.24. *Once the curve has genus $g \geq 2$, we can find such a rational function f satisfying $deg\, f > |f^{-1}(\{0,1,\infty\})|$.*

Proof. The proof of Belyi's theorem 2.121 gives an effective procedure for obtaining a rational function $f \in K(C)$ ramified only above $0, 1, \infty$. For a map that is only ramified over $0, 1$ and ∞, using Hurwitz Formula 2.74, we have

$$2g - 2 = -2deg\, f + \sum_{Q \in \{0,1,\infty\}} (deg\, f - |f^{-1}(Q)|) = deg\, f - |f^{-1}(\{0,1,\infty\})|,$$

hence $|f^{-1}(\{0,1,\infty\})| < deg\, f$, since $g \geq 2$. □

6.3. The Proof: Abc Implies Mordell

We prove *Mordell conjecture* using *abc conjecture* as in the paper of Elkies [1], which gives an effective height bound for the points on the curve.

Theorem 6.25. *Assume abc conjecture, then any curve of genus at least* 2 *over a number field* K *has only finitely many* K*-rational points.*

First, we prove a technical proposition which is one of the key points of the proof.

Notation: We write $f(x) = O(g(x))$ for functions $f(x)$ and $g(x)$ if and only if there exist constants C and k such that $|f(x)| \leq C|g(x)|$ for all $x > k$.

Proposition 6.26 ([1])**.** *Let* C *be any curve over* K *and* $f \in K(C)$ *be a rational function of degree* d*. Then for any* K*-rational point* $P \in C(K) - f^{-1}(0)$ *we have*

$$\log N_0(f(P)) < \left(1 - \frac{b_f(0)}{d}\right) \log H_P + O(\sqrt{\log H_P} + 1)$$

with the implied constant effective and depending on K, C, f *but not on* P *and* $H_P := H(f(P))$.

Proof. We use logarithmic height functions relative to divisors. Let $D = \sum_k m_k D_k$ be the zero divisor of f, where D_k are prime rational divisors of degrees d_k. Then we have $deg(D) = \sum_k m_k d_k = degf = d$ and $b_f(0) = d - \sum_k d_k$. Let $D' = \sum_k D_k$, then $deg(D') = \sum_k d_k$ and we can express the ramification above 0 as $b_f(0) = deg(D) - deg(D')$. By Proposition 2.115 and by additivity of Theorem 2.113, we have

$$\log H_P = \log H(f(P)) = h_D(P) + O(1) = \sum_k m_k h_{D_k}(P) + O(1).$$

Eliminating the finitely many bad primes as in the Subsection 2.3.5, a prime v contributes to $\log N_0(f(P))$ only if $f(P) \equiv 0 \mod v$, i.e., $ord_v(f(P)) > 0$, which holds also for the reduction $\overline{f}(\overline{P}) \equiv \overline{0} \mod v$. In this case, $\overline{P}$ is in the support of $\overline{D}$. Since the decomposition of D remain the same when reducing mod v, i.e., no coalescence or no vanishing, $\overline{P}$ is in the support of some $\overline{D_k}$. Moreover, P is in the support of some D_k, hence the prime contributes to $h_{D_k}(P)$ for some k. By the same reasoning applied backwards, we see that a prime contributes to $logN_0(f(P))$ if and only if it contributes to $h_{D_k}(P)$ for some k.

Now since the contribution of any prime, finite or infinite, to the height associated to an effective divisor is bounded below, we have $h_{D_k}(P) \geq O(1)$, by positivity of Theorem 2.113. Moreover by definition, $\log |f(P)|_v$ is a multiple of $\log(\mathbb{N}(v))$, i.e., $|h_{D_k}(P)| \geq \log(\mathbb{N}(v))$. Then summing over all primes, we obtain

$$\log N_0(f(P)) < \sum_k h_{D_k}(P) + O(1) = h_{D'}(P) + O(1),$$

with the second equation following from additivity of Theorem 2.113 and $<$ accounting for finite primes at which P and D' may meet nontransversally and infinite places at which P may come close the support of D'.

Lastly, we construct a degree zero divisor to make use of the bound in the Theorem 2.116. So let $E = (deg(D))D' - (deg(D'))D$. Then using the additivity of Theorem 2.113,

$$h_E(P) = (deg(D))h_{D'}(P) - (deg(D'))h_D(P) + O(1)$$

and by Proposition 2.115 and Theorem 2.116,

$$= (deg(D))h_{D'}(P) - (deg(D'))h(f(P)) + O(1) \leq O(\sqrt{\log H_P} + 1),$$

so

$$h_{D'}(P) \leq \frac{deg(D')}{deg(D)} h(f(P)) + O(\sqrt{\log H_P} + 1).$$

Combining with the above inequality, we get

$$\log N_0(f(P)) < \left(1 - \frac{b_f(0)}{d}\right) \log H_P + O(\sqrt{\log H_P} + 1)$$

since $b_f(0) = deg(D) - deg(D')$. □

Proof of Theorem 6.25. Let C be an arbitrary curve of genus $g \geq 2$. We fix a function $f \in K(C)$, satisfying $|P \in C(\overline{\mathbb{Q}}) : f(P) \in \{0, 1, \infty\}| < deg\, f$, which exists by Belyi's theorem 2.121 and the inequality is satisfied once $g \geq 2$ by Proposition 6.24. Let $m := |P \in C(\overline{\mathbb{Q}}) : f(P) \in \{0, 1, \infty\}|$ and $d := deg\, f$. Using Proposition 6.26, for any K-rational point $P \in C(K) - f^{-1}(0)$, we have

$$\log N_0(f(P)) < \left(1 - \frac{b_f(0)}{d}\right) \log H_P + O(\sqrt{\log H_P} + 1).$$

Replacing f by $f - 1$ in the Proposition 6.26, we get that for all $P \in C(K) - (f-1)^{-1}(0)$, i.e., $P \in C(K) - f^{-1}(1)$,

$$\log N_1(f(P)) < \left(1 - \frac{b_f(1)}{d}\right) \log H_P + O(\sqrt{\log H_P} + 1),$$

since $N_0((f-1)(P)) = N_0(f(P) - 1) = N_1(f(P))$ by definition and $\log H(f(P)) = \log H(f(P) - 1) + O(1)$ by Remark 2.105.

Similarly, replacing f by $1/f$ in the Proposition 6.26, we get that for all $P \in C(K) - (1/f)^{-1}(0)$, i.e., $P \in C(K) - f^{-1}(\infty)$,

$$\log N_\infty(f(P)) < \left(1 - \frac{b_f(\infty)}{d}\right) \log H_P + O(\sqrt{\log H_P} + 1),$$

since $N_0((1/f)(P)) = N_0(1/f(P)) = N_\infty(f(P))$ by definition and $\log H(f(P)) = \log H(1/f(P))$ due to H being scaling invariant.

Then for any K-rational point $P \in C(K)$ not in the finite set $f^{-1}(\{0,1,\infty\})$, we have:

$$\log N_P < \frac{m}{d} \log H_P + O(\sqrt{\log H_P} + 1),$$

with the implied O-constant effective and depending on K, C, f but not on P, since $\log N_0(f(P)) + \log N_1(f(P)) + \log N_\infty(f(P)) = \log N(f(P))$ and

$$\left(1 - \frac{b_f(0)}{d}\right) + \left(1 - \frac{b_f(1)}{d}\right) + \left(1 - \frac{b_f(\infty)}{d}\right) = \frac{\left(degf - \sum_{f(P)=0}(e_f(P) - 1)\right)}{d} +$$

$$\frac{\left(degf - \sum_{f(P)=1}(e_f(P) - 1)\right)}{d} + \frac{\left(degf - \sum_{f(P)=\infty}(e_f(P) - 1)\right)}{d} =$$

$$\frac{|f^{-1}(0)| + |f^{-1}(1)| + |f^{-1}(\infty)|}{d} = \frac{|f^{-1}(\{0,1,\infty\})|}{d} = \frac{m}{d}.$$

Taking logarithms of both sides in *abc conjecture,* it becomes $(1-\epsilon)\log H \leq \log N + C_\epsilon$ for some constant C_ϵ. Then $f(P)$ gives a counter example to *abc conjecture* over K for $\epsilon < 1 - (m/d)$ once H_P is large enough, i.e., for all but finitely many P. In particular, combining with the inequality, we get the effective height bound

$$\log H_P < \frac{O(\sqrt{\log H_P} + 1)}{1 - \epsilon - m/d}$$

for the points $P \in C(K)$ not in the finite set $f^{-1}(\{0,1,\infty\})$. Since $f^{-1}(\{0,1,\infty\})$ is finite, by Theorem 2.107, we are done. □

Remark 6.27. *Using Theorem 2.117 instead of Theorem 2.116 in the last paragraph above, we can get a different height bound on the points.*

Proposition 6.28. *Let C be any curve over K and $f \in K(C)$ be a rational function of degree d. Then for all $\epsilon > 0$, there is a constant C_ϵ such that for any K-rational point $P \in C(K) - f^{-1}(0)$ we have*

$$\log N_0(f(P)) < \left(1 - \frac{b_f(0)}{d}\right)(1+\epsilon)\log H_P + C_\epsilon.$$

Proof. With the same set-up constructed above, instead of defining a degree zero divisor, we use the inequality from Theorem 2.117 for the divisors D and D'. Then for all $\epsilon > 0$, there is a constant C_ϵ such that for all $P \in C(K) - f^{-1}(0)$, we have $h_{D'}(P) \leq (1+\epsilon)\frac{degD'}{degD}h_D(P) + C_\epsilon$. Then using Theorem 2.115, we get

$$\log N_0(f(P)) < \left(1 - \frac{b_f(0)}{d}\right)(1+\epsilon)\log H_P + C_\epsilon,$$

since $b_f(0) = degD - degD'$. □

Proof of Theorem 6.25. Replacing f by $f-1$ and $1/f$ in Proposition 6.28 and then adding all inequalities we obtained, we get for all points $P \in C(K)$ not in the finite set $f^{-1}(\{0,1,\infty\})$,

$$\log N_P < (1+\epsilon)\frac{m}{d}\log H_P + 3C_\epsilon.$$

Combining with *abc conjecture*, which says that $(1-\epsilon')\log H_P \leq \log N_P + C_{\epsilon'}$, we obtain

$$\left(1 - \epsilon' - \frac{(1+\epsilon)m}{d}\right)\log H_P < C_\epsilon + C_{\epsilon'}.$$

Choose $\epsilon, \epsilon' > 0$ so that $1 - \epsilon' - \frac{(1+\epsilon)m}{d} > 0$. Then we have the height bound $\log H_P < C$ with constant C depending on ϵ, ϵ'. By Theorem 2.107, we are done. □

6.4. Examples

We give various examples using two different height bounds we obtained in the last section.

Example 6.29. *Let F_5 be the Fermat curve of genus 6, computed by Example 2.77, defined over $\mathbb{Q}$ by the equation $x^5+y^5+z^5=0$. Let $f : F_5(\overline{\mathbb{Q}}) \to \mathbb{P}'(\mathbb{C})$ with $(x,y,z) \mapsto (-x^5, y^5)$ be a rational function. By Proposition 6.19 and Proposition 6.20, we know that $\deg f = 25$ and f is ramified over $\{0,1,\infty\}$ with ramification degrees 5. Using Hurwitz Formula 2.74, we see that f is only ramified over $\{0,1,\infty\}$ and nowhere else. Moreover, using $deg(f) = \sum_P e_f(P)$, we have $|\{P \in F_5(\overline{\mathbb{Q}}) : f(P) \in \{0,1,\infty\}\}| = 15$. Then f is a Belyi function satisfying $|\{P \in F_5(\overline{\mathbb{Q}}) : f(P) \in \{0,1,\infty\}\}| < \deg f$. Hence letting $\epsilon = 1/5$, for all $P \in F_5(\mathbb{Q})$ such that $f(P) \neq 0,1,\infty$, we get $\log H_P < 5O(\sqrt{\log H_P}+1)$ or $\log H_P < O(\sqrt{\log H_P}+1)$, since Big-Oh function is a set.*

Now let x,y,z be relatively prime integers, then $f(P) = (-x^5, y^5)$ or we could write as $f(P) = (-x^5/y^5 : 1)$ since $f(P) \neq \infty$. Then $\log H_P = \log H(f(P)) = \log max(|-x^5|_\infty, |y^5|_\infty)$ because for finite primes p, we have $|-x^5|_p \leq 1$, $|y^5|_p \leq 1$, but $gcd(x,y) = 1$ and we choose 1 for any finite prime. Therefore $\log H_P$ is equal to either $\log(\pm x^5)$ or $\log(\pm y^5)$. Without loss of generality, let $\log H_P = \log(x^5)$ where $x \in \mathbb{Z}_{>0}$ and $x > |y|_\infty$. Plugging in the height bound, we get $\log(x^5) < O(\sqrt{\log(x^5)}+1)$.

For any candidates of constants $C > 0$ and k, we can take $x > k$ and we would have to satisfy $|\log(x^5)| < C|\sqrt{\log(x^5)}+1|$, which gives us the interval $1 < x < 2.71828^{h(C)}$, where $h(C) = 0.1\sqrt{C^3(C+4)} + 0.1C^2 + 0.2C$, for a fixed C. Since there exist finitely many integers in an interval, we have finitely many options for x. There exist also finitely many options for y, since $x > |y|_\infty$. Then there are finitely many points $P \in F_5(\mathbb{Q})$ such that $f(P) \neq 0,1,\infty$, since z is dependent on x and y with the equation $x^5+y^5+z^5=0$. In the end, the Fermat curve F_5 has only finitely many rational points.

Example 6.30. *Let $C : y^2 = x^7 - 1$ be defined over $\mathbb{Q}$, of genus 3. We fix the rational map*

$$f : C \to \mathbb{P}'$$
$$(x, y) \mapsto (x^6 + x^5 + x^4 + x^3 + x^2 + x + 1, 7)$$

which is a Belyi function by Example 2.123.

We are interested in points $P \in C(\mathbb{Q})$ with $f(P) \neq 0, 1, \infty$, so we are in the affine plane with $z = 1$. Choose x, y relatively prime integers, then $\log H_P = \log max(|x^6 + x^5 + x^4 + x^3 + x^2 + x + 1|_\infty, 7)$.

(i) When $|x^6 + x^5 + x^4 + x^3 + x^2 + x + 1|_\infty \leq 7$, we have $\log H_P = \log 7$ and using the height bound, we obtain the inequality $\log 7 < O(\sqrt{\log 7} + 1)$, which is a true statement. Therefore, the integers satisfying $|x^6 + x^5 + x^4 + x^3 + x^2 + x + 1|_\infty \leq 7$ are options for the variable x, which are $-1, 0$, and 1.

(ii) When $|x^6 + x^5 + x^4 + x^3 + x^2 + x + 1|_\infty \geq 7$, we have $\log H_P = \log(|x^6 + x^5 + x^4 + x^3 + x^2 + x + 1|_\infty)$. Without loss of generality, choose $x > 1$, then using the height bound, we obtain the inequality $\log(x^6 + x^5 + x^4 + x^3 + x^2 + x + 1) < O(\sqrt{\log(x^6 + x^5 + x^4 + x^3 + x^2 + x + 1)} + 1)$, which gives us an interval. Then there exist finitely many options for x, since there are finitely many integers in an interval.

Since y depends on x, we say that the curve C has finitely many rational points.

Note that the functions in the inequality of Example 6.30 grows more slowly with respect to the ones of Example 6.29. For example, when $C = 600$ in Example 6.30, we have $-8.59523 < x < 8.25943$ as an interval; whereas, when $C = 50$ in Example 6.29, we have $1 \leq x < 5.618 \times 10^{225}$ as an interval. Hence it is easier to find the rational solutions of $y^2 = x^7 - 1$ by checking out of the points satisfying the height bound.

Example 6.31. *Let $C : (y-\sqrt{2})^2 = x^5 - 1$ be defined over $\mathbb{Q}(\sqrt{2})$, of genus 2. We fix the rational map*

$$f : C \to \mathbb{P}'$$
$$(x, y) \mapsto (x^4 + x^3 + x^2 + x + 1, 5)$$

which is a Belyi function by Example 2.124.

Take arbitrary $P = (x, y) \in C(\mathbb{Q}(\sqrt{2}))$ with $f(P) \neq 0, 1, \infty$, then

$$f(P) = \frac{x^4 + x^3 + x^2 + x + 1}{5} \in \mathbb{Q}(\sqrt{2}), \text{ and}$$

$$H_P = \prod_{v \in M_{\mathbb{Q}(\sqrt{2})}} max\{|x^4 + x^3 + x^2 + x + 1|_v, |5|_v\},$$

where ${M_{\mathbb{Q}(\sqrt{2})}}^0 = |\,|_{\mathcal{P}}$ determined by finite primes $\mathcal{P}$ of $\mathbb{Q}(\sqrt{2})$ and ${M_{\mathbb{Q}(\sqrt{2})}}^\infty = |\,|_{\phi_i}$ determined by embeddings $\phi_1 : a + b\sqrt{2} \mapsto a + b\sqrt{2}$, $\phi_2 : a + b\sqrt{2} \mapsto a - b\sqrt{2}$ of $\mathbb{Q}(\sqrt{2})$ into $\mathbb{C}$.

Note that for 5, we have $|5|_{\phi_1} = |5|_{\phi_2} = |5|_\infty = 5$. Moreover, $5\mathbb{Z}[\sqrt{2}] = \mathcal{P}$ with norm $\mathbb{N}(\mathcal{P}) = 25$ by Example 2.34. Then we have $|5|_{\mathcal{P}} = 1/25$ and $|5|_{\mathcal{Q}} = 1$ for all the other primes $\mathcal{Q}$.

We separate into cases:

(i) Let $x^4 + x^3 + x^2 + x + 1 =: a \in \mathbb{Z}$,

- *If $(a, 5) = 1$, then $\log H_P = \log\, max\{|a|_{\phi_1}, 5\} + \log\, max\{|a|_{\phi_2}, 5\} = 2 \log\, max\{|a|_\infty, 5\}$. Therefore, when $|a|_\infty \leq 5$, we have $\log H_P = 2 \log 5 = \log 25$. Using the height bound, we get $\log 25 < O(\sqrt{\log 25}+1)$, which is a true statement. So the integers satisfying $|x^4 + x^3 + x^2 + x + 1|_\infty \leq 5$ are options for x. On the other hand, when $|a|_\infty > 5$, we have $\log H_P = 2 \log |a|_\infty$. Using*

the height bound, we get $\log a^2 < O(\sqrt{\log a^2}+1)$, *which gives us an interval. Since there are finitely many integers in an interval, there are finitely many options for* x.

- *If* $(a,5) > 1$, *then* $\dfrac{x^4+x^3+x^2+x+1}{5} =: b \in \mathbb{Z}$ *and* $H(a,5) = H(b,1)$. *We get* $\log H_P = \log\, max\{|b|_{\phi_1}, 1\} + \log\, max\{|b|_{\phi_2}, 1\} = 2\log\, max\{|b|_\infty, 1\}$, *but* $f(P) \neq 1$, *so we have one case, which is* $|x^4+x^3+x^2+x+1|_\infty > 5$. *In that case, we get* $\log H_P = 2\log|b|_\infty$. *Using the height bound, we obtain* $\log b^2 < O(\sqrt{\log b^2}+1)$, *which gives us an interval. Since there are finitely many integers in an interval, there are finitely many options for* x.

(ii) Let $x^4+x^3+x^2+x+1 =: a+b\sqrt{2} \in \mathbb{Q}(\sqrt{2})$ *with* $b \neq 0$.

- *For archimedean valuations, we have* $|a+b\sqrt{2}|_{\phi_1} = |a+b\sqrt{2}|_\infty$ *and* $|a+b\sqrt{2}|_{\phi_2} = |a-b\sqrt{2}|_\infty$. *Then if* $|a+b\sqrt{2}|_\infty > 5$, *we choose* $|a+b\sqrt{2}|_\infty$ *as maximum in* H *for valuation* $|\,|_{\phi_1}$; *otherwise, we choose* 5. *For the other archimedean valuation, if* $|a-b\sqrt{2}|_\infty > 5$, *then we choose* $|a-b\sqrt{2}|_\infty$ *for* $|\,|_{\phi_2}$; *otherwise, we choose* 5.
- *We look into nonarchimedean valuations. The element* $a+b\sqrt{2}$ *forms the ideal* $(a+b\sqrt{2})\mathbb{Z}[\sqrt{2}]$ *with norm* $\mathbb{N}((a+b\sqrt{2})\mathbb{Z}[\sqrt{2}]) = a^2-2b^2$.
  - *If* $2^k \mid a^2-2b^2$ *for some* $k \in \mathbb{Z}$, *then* $|a+b\sqrt{2}|_{\langle 2,\sqrt{2}\rangle} = (1/2)^k$, *since* $2\mathbb{Z}[\sqrt{2}] = \langle 2,\sqrt{2}\rangle^2$ *with* $\mathbb{N}(\langle 2,\sqrt{2}\rangle) = 2$. *Moreover, we have* $|5|_{\langle 2,\sqrt{2}\rangle} = 1$. *Therefore, when* $k \geq 1$, *we choose* 1 *for* $|\,|_{\langle 2,\sqrt{2}\rangle}$; *otherwise we choose* $(1/2)^k$.
  - *If* $5^k \mid a^2-2b^2$ *for some* $k \in \mathbb{Z}$, *then* $|a+b\sqrt{2}|_{\mathcal{P}} = (1/25)^{k/2} = (1/5)^k$. *Moreover, we have* $|5|_{\mathcal{P}} = 1/25 = (1/5)^2$. *Therefore, when* $k \geq 2$, *we choose* $(1/5)^2$ *for* $|\,|_{\mathcal{P}}$; *otherwise we choose* $(1/5)^k$.
  - *When a prime* $q \equiv \pm 3 \mod 8$ *with* $q \neq 5$, *it stays prime above, i.e.,* $\mathcal{Q} \in q\mathbb{Z}[\sqrt{2}]$, *by Quadratic Reciprocity and Example 2.34. Say* $q^k \mid a^2-2b^2$ *for some* $k \in \mathbb{Z}$, *then* $|a+b\sqrt{2}|_{\mathcal{Q}} = (1/q)^k$. *Moreover, we have* $|5|_{\mathcal{Q}} = 1$. *Therefore, when* $k \geq 1$, *we choose* 1 *for* $|\,|_{\mathcal{Q}}$; *otherwise we choose* $(1/q)^k$.
  - *When a prime* $q \equiv \pm 1 \mod 8$, *it splits into two prime ideals* $\mathcal{Q}_1, \mathcal{Q}_2$ *above with* ${\mathcal{Q}_1}^{k_1}{\mathcal{Q}_2}^{k_2} \in q\mathbb{Z}[\sqrt{2}]$, *by Quadratic Reciprocity and Example 2.34. Say* $q^k \mid a^2-2b^2$ *for some* $k = k_1+k_2 \in \mathbb{Z}$, *then* $|a+b\sqrt{2}|_{\mathcal{Q}_1} =$

$(1/q)^{k_1}$ *and* $|a+b\sqrt{2}|_{\mathcal{Q}_2} = (1/q)^{k_2}$. *Moreover, we have* $|5|_{\mathcal{Q}_1} = |5|_{\mathcal{Q}_2} = 1$. *Therefore, if* $k_1 \geq 1$, *we choose* 1 *for* $|\,|_{\mathcal{Q}_1}$; *otherwise we choose* $(1/q)^{k_1}$. *Also, if* $k_2 \geq 1$, *we choose* 1 *for* $|\,|_{\mathcal{Q}_2}$; *otherwise we choose* $(1/q)^{k_2}$.

In the end, we get 2^7 *cases for the height* $\log H_P$. *Note that there is no possibility of having a value less than* 1 *inside* $\log$, *hence a negative value for* $\log H_P$. *Using Theorem 6.25 with the height bound in Proposition 2.117, for each case, we get* $\log H_P < C$ *for some constant* C. *Hence, by Theorem 2.107, there are only finitely many points* $P \in C(\mathbb{Q}(\sqrt{2}))$ *with* $f(P) \neq 0, 1, \infty$. *Since* $|f^{-1}(\{0,1,\infty\})|$ *is finite,* $C(\mathbb{Q}(\sqrt{2}))$ *is finite.*

7. CONCLUSION

In this thesis, we started by giving a long Preliminaries section providing the background necessary for further proofs and examples. We introduced many concepts and tools such as absolute values, height functions, Belyi maps and so on. Next, we worked on *abc conjecture*, which has been a hot topic in mathematics lately, and its relation between other conjectures, and theorems such as Mordell, Hall, and so on. Apart from its bridge role between the operations of addition and multiplication, *abc conjecture* provides such simple and elegant proofs for strong theorems. For example, we showed that one can prove *Fermat's last theorem* for large exponents using *abc conjecture* in one elementary paragraph, whereas Wiles had to prove Modularity conjecture for semi stable elliptic curves. We also gave the proof of *Mason-Stothers theorem*, which is the analogue of *abc conjecture* for polynomials.

Later, in Chapter 4, we introduced *Hall conjecture*, which puts a bound on the size of the integral points on a Mordell curve. We gave an example on how to put a bound on the size of those points using the properties of elliptic curves, by showing that a Mordell curve is in fact an elliptic curve. Then we gave the proof of *Weak Hall conjecture* using *abc conjecture*.

Finally, after giving the statement of *Mordell conjecture*, now *Faltings' theorem*, we moved on to the proof of *Mordell conjecture* using *abc conjecture*, which is a great result from the article "Abc Implies Mordell" by Noam Elkies, [1]. The proof represents a great connection between the theory of Diophantine approximation and the theory of points on curves of high genera. Before giving the proof, we gave the ideas which created the motivation for the proof. The effective proof of *Belyi* and the proof showing that the Fermat curve has no solutions, by using a rational map on the curve, were two of them. Then, we presented the proof using *abc conjecture*, given by Noam Elkies [1], which is of great importance because it is effective. The proof of Faltings is not effective; it proves that there are finitely many rational points on a curve of genus > 1, but does not give a method for finding them or give an upper bound on the size of the points. On

the other hand, the proof by Elkies provides an effective upper bound on the size of the points, hence limits the search criteria for finding the points. We also acquired another height bound by using a different theorem, bounding the height functions associated to divisors, with the technical proposition by Elkies [1]. In the end, we gave examples using both of the height bounds we obtained.

REFERENCES

1. Elkies, N. D., "ABC implies Mordell", *International Mathematics Research Notices*, Vol. 1991, No. 7, pp. 99–109, 1991.

2. Goldfeld, D., "Beyond the last theorem", *Math Horizons*, Vol. 4, No. 1, pp. 26–34, 1996.

3. O'Neil, C., "Notes on the Oxford IUT workshop by Brian Conrad", `https://mathbabe.org/2015/12/15/notes-on-the-oxford-iut-workshop-by-brian-conrad/`, accessed in February 2018.

4. Milne, J. S., "Algebraic Number Theory (v3.07)", `www.jmilne.org/math/`, accessed in May 2018.

5. Conrad, K., "Ostrowski for number fields", `https://kconrad.math.uconn.edu/blurbs/gradnumthy/ostrowskinumbfield.pdf`, accessed in February 2018.

6. Lorenzini, D., *An invitation to arithmetic geometry*, Vol. 9, American Mathematical Soc., 1996.

7. Conrad, K., "Discriminants and ramified primes", `https://kconrad.math.uconn.edu/blurbs/gradnumthy/disc.pdf`, accessed in March 2018.

8. Silverman, J. H., *The arithmetic of elliptic curves*, Vol. 106, Springer Science & Business Media, 2009.

9. Bruin, N., "Generalization of the ABC-conjecture", *Citeseer*, 1995.

10. Pazuki, F., "Heights", `https://www.math.u-bordeaux.fr/~abesheno/heights.pdf`, accessed in May 2018.

11. Van Frankenhuysen, M., "The ABC conjecture implies Roth's theorem and Mordell's conjecture", *Math. Contemp*, Vol. 16, pp. 45–72, 1999.

12. Lozano-Robledo, Á., *Elliptic curves, modular forms, and their L-functions*, American Mathematical Society Providence, RI, 2011.

13. Hindry, M. and J. H. Silverman, *Diophantine geometry: an introduction*, Vol. 201, Springer Science & Business Media, 2013.

14. Brown, M., J. P. Serre and M. Waldschmidt, *Lectures on the Mordell-Weil theorem*, Springer, 1989.

15. Lang, S., *Fundamentals of Diophantine geometry*, Springer Science & Business Media, 2013.

16. Ines, M., "Belyi's theorem", `https://alexjbest.github.io/buntes/sec-belyi-thm.html`, accessed in May 2018.

17. Waldschmidt, M., "On the abc Conjecture and some of its consequences", *Math Horizons*, February 2018.

18. Lang, S., *Math talks for undergraduates*, Springer Science & Business Media, 2012.

19. Conrad, K., "Examples of Mordell's equation", `https://kconrad.math.uconn.edu/blurbs/gradnumthy/mordelleqn1.pdf`, accessed in April 2018.

20. Ho, W., "How many rational points does a random curve have?", *Bulletin of the American Mathematical Society*, Vol. 51, No. 1, pp. 27–52, 2014.