Mordell's Conjecture
Canonical Heights and Vojta's Inequality

Bowen Yang

April 4, 2018

Submitted to the
Department of Mathematics and Statistics
of Amherst College
in partial fulfillment of the requirements
for the degree of
Bachelor of Arts with honors

Faculty Advisor: Professor Gregory S. Call

Abstract

The paper is part of an effort to understand some key results in an area called *Diophantine Geometry*. Among them are the Mordell-Weil theorem, Roth's theorem, Siegel's theorem, and Mordell's Conjecture (Faltings' theorem).

The particular approach to the Mordell's Conjecture I studied (which differs from Faltings' original proof) relies on the former three theorems. The theory of canonical height functions on abelian varieties plays a paramount role to the translation of the Mordell's Conjecture to a finite-dimensional linear algebra problem. Then an important inequality due to Vojta implies Faltings' theorem.

Throughout the paper, I strive to keep the account elementary by defining all the basic algebro-geometric objects in ways I found most helpful when I first learnt them. The first section of each chapter is dedicated to some informal introduction using metaphors closest to my heart. My account is by no means complete and should be viewed as a story about height functions and, to a lesser extent, Vojta's inequality.

Recommended Books

The main reference for the paper is Hindry and Silverman [5], however, the mathematical background needed to understand the proof of Mordell's Conjecture (Faltings' Theorem) might be beyond what is usually taught to undergraduate math majors. As an effort to make this paper more accessible to my fellow students as well as an excuse to record some books I truly enjoyed, I dedicate this page to books I have personally enjoyed reading during my undergraduate years. It should not be mistaken as the bibliography which is at the end of the paper.

- For Galois theory, I enjoyed the section on Galois theory in Dummit and Foote [2]. Another good read is [14] by David Cox.
- For algebraic number theory, you can never go wrong with the comprehensive book by Neukirch [10]. It is one of the rare 'bricks' that are suitable for reading from front to back.
- For introductory algebraic geometry, [3] was the first book I read. I then read a part of [15] by Eisenbud and Harris. I enjoyed both books as introductions, but there are many more books on algebraic geometry.
- For Riemann Surfaces, I absolutely love the book by Donaldson [7] and think anyone with an interest in geometry would agree.

Acknowledgements

I am most grateful for my advisor Gregory Call whose helpful guidance, kind words, and unwavering support carried me far in my four years at Amherst.

I owe thanks to my Mom, Dad and my girlfriend Yueying for being in my life.

Contents

1 The Basics **1**

1.1 The Canvas 1

1.2 Algebraic Varieties 1

1.3 Algebraic Curves 4

2 Height Functions **7**

2.1 Height Means Height 7

2.2 Absolute Values 7

2.3 Height on Projective Space 9

2.4 Divisors 15

2.5 Height on Varieties 18

2.6 Counting Rational Points on Curves 20

3 Abelian Varieties **22**

3.1 Bigger Picture 22

3.2 Abelian Varieties and Jacobian Varieties 22

3.3 Mordell-Weil Theorem 24

3.4 Canonical Height 24

4 Mordell's Conjecture **27**

4.1 Introduction 27

4.2 Vojta's Inequality 27

4.3 Proof of Vojta's Inequality 28

4.4 An Upper Bound 30

Bibliography **33**

List of Notation

k	a field.
$\bar{k}$	the algebraic closure of the field k.
G_k	the Galois group $\mathrm{Gal}(\bar{k}/k)$.
K	a number field i.e. a finite extension of $\mathbb{Q}$.
$\mathbb{A}^n$	affine n-space.
$\mathbb{A}^n(k)$	the set of k-rational points of $\mathbb{A}^n$.
$\mathbb{P}^n$	projective n-space.
$\mathbb{P}^n(k)$	the set of k-rational points of $\mathbb{P}^n$.
$k(P)$	field of definition of the point P.
C/K	a smooth projective curve of genus $g \geq 2$ defined over K.
J/K	the Jacobian variety of C.
$\mathrm{Div}(X)$	group of Weil divisors on X.
$D \sim D'$	linear equivalence of the divisors.
(f)	the principal divisor of the function f.
$L(D)$	space of rational functions with $(f) + D \geq 0$.
Θ	the theta divisor on J.
$H_k(P)$	relative height of a point in $\mathbb{P}^n(k)$.
$h_{X,D}$	height on the variety X relative to the divisor D.
$h_{X,\phi,D}$	canonical height on X relative to ϕ and D.

Chapter 1

The Basics

1.1 The Canvas

I think of Diophantine Geometry as paintings of number theory on algebraic geometry. Geometric spaces such as varieties are like the canvas, whereas the points of the varieties are like the painting on the canvas. In this chapter, we set up our canvas and then state Mordell's Conjecture (Faltings' theorem) which asserts that in fact many paintings have only finitely many points.

1.2 Algebraic Varieties

Let k be a field and let $\bar{k}$ be the algebraic closure of k.

Definition 1.2.1. An *affine n-space (over k)* is the set

$$\mathbb{A}^n = \{(x_1, \ldots, x_n) \mid x_i \in \bar{k}\}.$$

The set of *k-rational points* of $\mathbb{A}^n$ is the set

$$\mathbb{A}^n(k) = \{(x_1, \ldots, x_n) \in \mathbb{A}^n \mid x_i \in k\}.$$

An *affine algebraic set* X is a set of the form

$$X = V(I) = \{x \in \mathbb{A}^n \mid P(x) = 0 \text{ for all } P \in I\},$$

where I is an ideal in the polynomial ring $\bar{k}[X_1, \ldots, X_n] = \bar{k}[X]$.
For any subset S of $\mathbb{A}^n$, we define the ideal

$$I(S) = \{P \in \bar{k}[X] \mid P(x) = 0 \text{ for all } x \in S\},$$

and

$$\bar{k}[S] = \bar{k}[x_1, \ldots, x_n]/I(S).$$

The set S is said to be *defined over* k if the ideal $I(S)$ is generated by polynomials in $k[X]$.

Remark 1.2.2. Using the definitions above, we can define the *Zariski topology* on $\mathbb{A}^n$ by declaring its closed sets to be all algebraic sets. The Zariski topology on algebraic sets is simply the subspace topology (see [1] Chapter 2). Hence, there is the notion of irreducible algebraic set coming from point-set topology. In fact, we call an irreducible algebraic set of some $\mathbb{A}^n$ an affine variety.

Definition 1.2.3. A *projective n-space* $\mathbb{P}^n$ is the set of lines through the origin in $\mathbb{A}^{n+1}$. More precisely,

$$\mathbb{P}^n = \{(x_0, \ldots, x_n) \in \mathbb{A}^{n+1} \backslash \{0\}\} / \sim,$$

where the equivalence relation $\sim$ is defined by

$$(x_0, \ldots, x_n) \sim (y_0, \ldots, y_n) \iff (x_0, \ldots, x_n) = \lambda(y_0, \ldots, y_n) \text{ for some } \lambda \in \bar{k}^*.$$

The Galois group $G_k = \mathrm{Gal}(\bar{k}/k)$ acts on $\mathbb{P}^n$ by acting on the coordinates,

$$\sigma(P) = (\sigma(x_0), \ldots, \sigma(x_n)) \qquad \text{for } P = (x_0, \ldots, x_n) \in \mathbb{P}^n,\ \sigma \in G_k.$$

Then, we define the *k-rational points* in $\mathbb{P}^n$ as

$$\mathbb{P}^n(k) = \{P \in \mathbb{P}^n \mid \sigma(P) = P \text{ for all } \sigma \in G_k\}.$$

The *field of definition* of a point $P = (x_0, \ldots, x_n) \in \mathbb{P}^n$ is the smallest extension of k over which P is rational, namely,

$$k(P) = k(x_0/x_j, x_1/x_j, \ldots, x_n/x_j) \qquad \text{for any } j \text{ with } x_j \neq 0.$$

In order to define projective algebraic sets, we need to consider homogeneous polynomials in $\bar{k}[x_0, \ldots, x_n]$. They are the polynomials $f \in \bar{k}[x_0, \ldots, x_n]$ with the property that

$f(\lambda x_0, \ldots, \lambda x_n) = \lambda^n f(x_0, \ldots, x_n)$ for all $\lambda \in \bar{k}^*$ where n is the degree of the polynomial f. Notice when f is homogeneous,

$$f(x_0, \ldots, x_n) = 0 \iff f(\lambda x_0, \ldots, \lambda x_n) = 0 \quad \text{for all } \lambda \in \bar{k}^*.$$

An ideal of $\bar{k}[x_0, \ldots, x_n]$ is homogeneous if it is generated by homogeneous polynomials.

Definition 1.2.4. A *projective algebraic set* is the set of zeros in $\mathbb{P}^n$ of a homogeneous ideal in $\bar{k}[x_0, \ldots, x_n]$. The *homogeneous coordinate ring* of a projective variety $V \subset \mathbb{P}^n$ is the quotient $S(V) = \bar{k}[x_0, \ldots, x_n]/I(V)$.

Remark 1.2.5. Similarly, the *Zariski topology* on $\mathbb{P}^n$ is defined by taking the projective algebraic sets to be the closed sets. The topology on the algebraic sets is the subspace topology. A *projective variety* is an irreducible projective algebraic set.

Definition 1.2.6. The *standard affine open set* U_i of $\mathbb{P}^n$ is the complement of the hyperplane defined by $x_i = 0$.

Thus, we can view $\mathbb{P}^n$ as a 'global' space covered by n copies of 'local' spaces $\mathbb{A}^n$. This comes from the observation that if $P = (x_0, \ldots, x_n) \in U_i$, then $x_i \in \bar{k}^*$. Thus we can represent P by $(x_0/x_i, \ldots, x_{i-1}/x_i, 1, x_{i+1}/x_i, \ldots, x_n/x_i)$ which is determined by a unique n-tuple of $\bar{k}$, namely, a point $(x_0/x_i, \ldots, x_{i-1}/x_i, x_{i+1}/x_i, \ldots, x_n/x_i)$ in $\mathbb{A}^n$. The notion of affine open sets enable us to specify functions and morphisms.

Definition 1.2.7. Let X be a variety and x a point on X. A function $f : X \to \bar{k}$ is *regular* at x if x has an affine open neighborhood $U \subset X$, and polynomials $P, Q \in \bar{k}[x_1, \ldots, x_n]$ such that $U \subset \mathbb{A}^n$, $Q(x) \neq 0$ and $f(y) = P(y)/Q(y)$ for all $y \in U$. A function is regular on X if it is regular on every point of X. The set of all regular functions on a variety form a ring called the *ring of regular functions* on X which is denoted by $\mathcal{O}(X)$.

Definition 1.2.8. Let X be a variety and $Y \subset X$ a subvariety. The *local ring of X along Y*, denoted by $\mathcal{O}_{Y,X}$, is the set of pairs (U, f), where U is an open subset of X with $U \cap Y \neq \emptyset$ and $f \in \mathcal{O}(U)$, and where we identify two pairs $(U_1, f_1) = (U_2, f_2)$ if $f_1 = f_2$ on $U_1 \cap U_2$. The *function field* of X, denoted by $\bar{k}(X)$, is defined to be $\mathcal{O}_{X,X}$.

As $\bar{k} = \mathbb{A}^1$, functions on a variety X are also maps $X \to \mathbb{A}^1$. This motivates the general definitions of morphisms and rational maps between varieties.

Definition 1.2.9. A map $\phi : X \to Y$ between varieties is a *morphism* if it is continuous, and if for every open set $U \subset Y$ and every regular function f on U, the function $f \circ \phi$ is regular on $\phi^{-1}(U)$. An *isomorphism* is a morphism with an inverse which is also a morphism. A map is *regular* at a point x if it is a morphism on some open neighborhood of x.
A *rational map* from a variety X to a variety Y is a map that is a morphism on some nonempty open subset of X. A *birational map* is a rational map that has a rational inverse. Two varieties are said to be *birationally equivalent* if there is a birational map between them.

While the definitions may seem overwhelming, the main takeaway for the readers should be that the zeros of polynomial equations are intimately related to a vast class of geometric objects called varieties. Like other geometric objects we are more familiar with, varieties have a notion of dimension and it can be defined strictly through algebra.

Definition 1.2.10. The *dimension* of a variety X defined over $\bar{k}$ is the transcendence degree of its function field $\bar{k}(X)$ over $\bar{k}$. The dimension of an algebraic set is the maximum of the dimensions of its irreducible components.

Remark 1.2.11. A good definition of dimension should agree with our intuition. For example, the dimension of the affine space $\mathbb{A}^n$ naturally should be n. Indeed, the function field of the affine space is simply $\bar{k}(x_1, \ldots, x_n)$ which has transcendence degree (see page 645 of [2] for the definition) n over $\bar{k}$.

Just like real curves and surfaces, there is a notion of smoothness (or non-singularity) for algebraic varieties. The details of the definition do not play a vital role here and can be found in any introductory book on algebraic geometry such as [3] and [4]. For the rest of the paper, we only consider smooth varieties.

1.3 Algebraic Curves

The heroes of the paper are one-dimensional varieties known aptly as algebraic curves. According to our definition, the function field of an algebraic curve has transcendence

degree 1. It is natural to identify two curves with isomorphic function fields. In this section, we will state some results specific to curves, then we will see the classification of curves using *genus.*

Theorem 1.3.1. *A rational map from a smooth algebraic curve to a projective variety extends to a morphism defined on the whole curve.*

Proof. See [4] III.3, Theorem 3 or [9] IV.6.2.1. □

Corollary 1.3.2. *A birational map between two smooth projective curves is an isomorphism.*

Proof. It follows immediately from the previous theorem. □

Theorem 1.3.3. *Any algebraic curve is birational to a unique (up to isomorphism) smooth projective curve.*

Proof. See [6] I, Corollary 6.11. □

It follows that to study smooth algebraic curves, it is sufficient to consider smooth projective curves up to birational equivalence.

Genus is the last piece of the setup before we can state the main theorem of this expository paper. Note there are many equivalent definitions of genus, and the readers do not have to stick to the version given below.

We are only concerned with curves over a number field K, so it suffices to let $\bar{k} = \mathbb{C} \supset \bar{K}$. We then have the following theorems from the theory of Riemann surfaces.

Theorem 1.3.4. *A smooth projective algebraic curve over $\mathbb{C}$ is a compact, connected and oriented surface over $\mathbb{R}$.*

Proof. A smooth projective algebraic curve over $\mathbb{C}$ is compact because it is a closed subset of a compact space, namely, the projective space. Using the holomorphic implicit function theorem, there is a complex atlas on the curve. Hence the curve is a compact Riemann surface. Then the theorem follows from the properties of compact Riemann surfaces. Interested readers may refer to Chapter 3 of [7] (which is my favorite book on Riemann surfaces). □

Theorem 1.3.5. *A smooth compact, connected and oriented surface is homeomorphic to a sphere Σ_0 or some surface Σ_g, formally the connected sum of g tori where g is a positive integer.*

Proof. See [7] Chapter 2 for an excellently explained and illustrated proof using Morse theory. □

Remark 1.3.6. Homeomorphism is an equivalence relation of two topological spaces (refer to [1] Chapter 2). The precise definition is found in any introductory topology textbook. Intuitively Σ_g is the surface of a special doughnut that has g holes rather than just 1.

Definition 1.3.7. The previous theorems indicates that for any smooth algebraic curve, we can associate to it a unique nonnegative integer g related to its number of 'holes'. We call this integer g the *genus* of the curve.

Remark 1.3.8. Besides the number of 'holes' there is a more algebraic interpretation according to the degree-genus formula. Recall that a projective variety is the zero set of a set of homogeneous polynomials. If we simply look at a single homogeneous polynomial of degree d with three variables X, Y, Z, its set of zeros gives a projective algebraic curve. The genus of the curve is

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

In particular, if $d > 3$, then $g > 1$.

Now we are ready to behold Faltings' Theorem, previously known as Mordell's Conjecture!

Theorem 1.3.9 (Faltings)**.** *Let K be a number field and C be a smooth algebraic curve of genus g defined over k. If $g > 1$, the set of K-rational points $C(K)$ is finite.*

Remark 1.3.10. When $g = 0$, it can be shown that $C(K)$ is either empty or infinite. And when $C(K)$ is nonempty, then C is isomorphic over K to $\mathbb{P}^1$.
When $g = 1$, the points $C(K)$ possess a natural group structure. The curve together with the said group structure is called an *elliptic curve*. It is an important example of an *abelian variety*. Abelian varieties play an important part in the proof of Faltings' theorem.

Chapter 2

Height Functions

2.1 Height Means Height

When one looks for a nearby restaurant on a map, one is assigning a number to each point on the map, namely the distance from where one stands. On the other hand, when one looks for a mountain to climb, another number, namely the altitude from sea level, gets assigned to each point. Nevertheless, they both demonstrate the convenience of such a real-valued function that captures relevant aspects of each point. In this chapter, we are going to study a systematic way of defining such functions on the points of our varieties. We call these functions *heights* because they are just like asking how high a point is above sea level. More directly, the height measures the arithmetic complexity of a point's coordinates.

2.2 Absolute Values

Definition 2.2.1. An *absolute value* on a field k is a real-valued function

$$|\cdot| : k \longrightarrow [0, \infty)$$

with the following properties:

(a) $|x| = 0$ if and only if $x = 0$.

(b) $|xy| = |x||y|$.

(c) $|x + y| \leq |x| + |y|$.

The absolute value is *nonarchimedean* if it satisfies

(c′) $|x + y| \leq \max\{|x|, |y|\}$.

For $\mathbb{Q}$ we have $|\cdot|_\infty$ inherited from the standard absolute value on $\mathbb{R}$ and the p-adic absolute values (defined below). They form the set $M_\mathbb{Q}$ of all standard absolute values on $\mathbb{Q}$.

Definition 2.2.2. Let p be a fixed prime. Any nonzero rational $r \in \mathbb{Q}$ can be written uniquely as

$$r = p^s u/v,\ s, u, v \in \mathbb{Z}, v > 0,\ p \nmid u, p \nmid v,\ \text{and } u, v \text{ relatively prime.}$$

We define the p-adic absolute value

$$|r|_p = p^{-s}.$$

The set of standard absolute values on a number field K is the set M_K of all absolute values on K whose restriction to $\mathbb{Q}$ is in $M_\mathbb{Q}$. We write M_K^∞ for the set of archimedean ones and M_K^0 for the nonarchimedean ones. We denote absolute value associated to $v \in M_k$ by $|\cdot|_v$. We define

$$v(x) = -\log |x|_v \text{ for all } x \in K^*,\ v(0) = \infty$$

to be the valuation associated to the absolute value $|\cdot|_v$.

Let K'/K be an extension of number fields and let $v \in M_K,\ w \in M_{K'}$, we say w *divides* v and write $w|v$ if the restriction of w to K is v.

Clearly,

$$\prod_{v \in M_\mathbb{Q}} |x|_v = 1 \ \text{ for all } x \in \mathbb{Q}^*.$$

For $v \in M_K$ we write K_v for the completion of K with respect to v. We have the following result.

Proposition 2.2.3 (Degree Formula)**.** *Let K'/K be an extension of number fields, and let $v \in M_K$. Then*

$$\sum_{w \in M_{K'}, w|v} [K'_w : K_v] = [K' : K].$$

Proof. See [10] II.8.4 □

Definition 2.2.4. Let $v \in M_K$. Then the *local degree* of v is the number $n_v = [K_v : \mathbb{Q}_v]$. The *normalized absolute value* associated to v is

$$\|x\|_v = |x|_v^{n_v}.$$

Proposition 2.2.5 (Product Formula)**.** *K a number field and $x \in K^*$. Then*

$$\prod_{v \in M_K} \|x\|_v = 1.$$

Proof. See [10] II.2.1. □

There is an alternative description of absolute values on a number field using specific real and complex embeddings as well as prime ideals. Interested readers can find it in most algebraic number theory textbooks such as [10].

2.3 Height on Projective Space

There is a natural way to measure the size of a rational point $P \in \mathbb{P}^n(\mathbb{Q})$. Write $P = (x_0, x_1, \ldots, x_n)$ with $x_0, \ldots, x_n \in \mathbb{Z}$ and relatively prime, define the height of P to be

$$H(P) = \max\{|x_0|, |x_1|, \ldots, |x_n|\}.$$

Clearly, for any $B \in \mathbb{R}$, the set

$$\{P \in \mathbb{P}^n | H(P) \leq B\}$$

is finite. This will be useful for a proof later.

Definition 2.3.1. Let K be a number field, and let $P = (x_0, x_1, \ldots, x_n) \in \mathbb{P}^n(K)$ be a point whose homogeneous coordinates are chosen in K. The *height* of P (relative to K) is

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

We also use the logarithmic height defined as $h_K(P) = \log H_K(P)$ when we are in an additive mood rather than a multiplicative one.

We have three properties of the height (see [5] B.2.1):

(a) It is independent of the choice of homogeneous coordinates.

(b) $H_K(P) \geq 1$ for all $P \in \mathbb{P}^n(K)$.

(c) K'/K finite extension. Then $H_{K'}(P) = H_K(P)^{[K':K]}$.

This motivates the following field invariant definition.

Definition 2.3.2. The *absolute height* on $\mathbb{P}^n$ is the function

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

where K is any field with $P \in \mathbb{P}^n(K)$. Similarly, we have $h(P) = \log H(P)$.

See [5] B.2 for the proof that H is well-defined and thus independent of the field K.

We also define the height of an element $\alpha \in K$ to be the height of $(\alpha, 1) \in \mathbb{P}^2(K)$.

Remark 2.3.3. We can see by this definition if $P = (x_0, x_1, \ldots, x_n) \in \mathbb{P}^n(\mathbb{Q})$ where $x_0, \ldots, x_n$ are relatively prime integers, then $H(P) = \prod_{v \in M_\mathbb{Q}} \max\{\|x_0\|_v, \|x_1\|_v \ldots, \|x_n\|_v\} = \max\{|x_0|, |x_1|, \ldots, |x_n|\}$. This follows from the definition of $H_\mathbb{Q}(P)$ and the facts that $\|x\|_v = |x|_v$ for all $v \in M_\mathbb{Q}$ and $\max\{\|x_0\|_v, \|x_1\|_v, \ldots, \|x_n\|_v\} = 1$ for all nonarchimedean $v \in M_\mathbb{Q}$. Thus, the absolute height is indeed a generalization of the natural height on $\mathbb{P}^n(\mathbb{Q})$, which justifies our using the same notation for both definitions.

Recall $\mathbb{Q}(P) := \mathbb{Q}(x_0/x_j, \ldots, x_n/x_j)$ for any $x_j \neq 0$. We now prove a theorem of fundamental importance.

Theorem 2.3.4. *For any numbers $B, D \geq 0$, the set*

$$\{P \in \mathbb{P}^n(\bar{\mathbb{Q}}) : H(P) \leq B \text{ and } [\mathbb{Q}(P) : \mathbb{Q}] \leq D\}$$

is finite. In particular, let K be a number field, the set

$$\{P \in \mathbb{P}^n(K) \leq B\}$$

is finite.

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

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

Multiplying over all v and taking an appropriate root, we see that

$$H(P) \geq H(x_i) \quad \text{for all } \ 0 \leq i \leq n.$$

Further, it is clear that $\mathbb{Q}(P) \supset \mathbb{Q}(x_i)$. Hence it suffices to prove that for each $1 \leq d \leq D$, the set

$$\{x \in \bar{\mathbb{Q}} \mid H(x) \leq B \ \text{ and } \ [\mathbb{Q}(x) : \mathbb{Q}] = d\} \qquad \text{is finite.}$$

Let $x \in \bar{\mathbb{Q}}$ have degree d and let $K = \mathbb{Q}(x)$. We write $x_1, \cdots, x_d$ for the conjugates of x over $\mathbb{Q}$, and we let

$$F_x(T) = \prod_{j=1}^{d}(T - x_j) = \sum_{r=0}^{d}(-1)^r s_r(x) T^{d-r}$$

be the minimal polynomial of x over $\mathbb{Q}$. More precisely, $s_r(x) = \sum_{1 \leq i_1 < \cdots < i_r \leq d} x_{i_1} \cdots x_{i_r}$. For any absolute value $v \in M_K$, we can estimate

$$\begin{aligned} |s_r(x)|_v &= \left| \sum_{1 \leq i_1 < \cdots < i_r \leq d} x_{i_1} \cdots x_{i_r} \right|_v \\ &\leq c(v, r, d) \max_{1 \leq i_1 < \cdots < i_r \leq d} |x_{i_1} \cdots x_{i_2}|_v \\ &\leq c(v, r, d) \max_{1 \leq i \leq d} |x_i|_v^r. \end{aligned}$$

Here $c(v, r, d) = \binom{d}{r} \leq 2^d$ if v is archimedean, and $c(v, r, d) = 1$ if v is nonachimedean. It follows that

$$\max\{|s_0(x)|_v, \ldots, |s_d(x)|_v\} \leq c(v, d) \prod_{i=1}^{d} \max\{|x_i|_v, 1\}^d$$

where $c(v, d) = 2^d$ if v is archimedean, and $c(v, d) = 1$ otherwise. Now we multiply this inequality over all $v \in M_K$ and take the $[K : \mathbb{Q}]^{\text{th}}$ root (note the number of archimedean v is exactly $[K : \mathbb{Q}]$):

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

From the definition of height, it is not hard to show that it is invariant under the action of the Galois group (see [5] B.2.2 for proof). Thus $H(x_i) = H(x)$ for any i. Hence,

$$H(s_0(x), \ldots, s_d(x)) \leq 2^d H(x)^{d^2}.$$

Now suppose that $x \in \{x \in \bar{\mathbb{Q}} \mid H(x) \leq B, [\mathbb{Q}(x) : \mathbb{Q}] = d\}$. Then x is the root of a polynomial $F_x(T) \in \mathbb{Q}[T]$ whose coefficients $s_0, \ldots, s_d$ satisfy $H(s_0, \ldots, s_d) \leq 2^d B^{d^2}$. However, earlier we noted the $\mathbb{P}^d(\mathbb{Q})$ has only finitely many points of bounded height, so there are only finitely many possibilities for the polynomial $F_x(T)$ and hence finitely many possibilities for x.

□

Corollary 2.3.5 (Kronecker's theorem)**.** *Let K be a number field, and let $P = (x_0, x_1, \ldots, x_n) \in \mathbb{P}^n(K)$. Fix any i with $x_i \neq 0$. Then $H(P) = 1$ if and only if the ratio x_j/x_i is a root of unity or zero for every j.*

Proof. Without loss of generality, we may divide the coordinates of P by x_i and reorder them to get that $P = (1, x_1, x_2, \ldots, x_n)$. If every x_j is a root of unity, then $|x_j|_v = 1$ for every absolute value on K, and hence $H(P) = 1$.

Conversely, suppose that $H(P) = 1$. For each $r = 1, 2, \ldots$, let $P^r = (x_0^r, \ldots, x_n^r)$. It is clear that $H(P^r) = H(P)^r = 1$, for any $r \geq 1$. So the theorem implies that $P, P^2, P^3, \ldots$ contains only finitely many distinct points. Then, for some $s > r \geq 1$, $x_j^s = x_j^r$ for each j, $0 \leq j \leq n$. Therefore each x_j is a root of unity or zero. □

The following proposition shows the interplay between geometry and arithmetic.

Theorem 2.3.6. *Let $\phi : \mathbb{P}^n \to \mathbb{P}^m$ be a rational map of degree d defined over $\bar{\mathbb{Q}}$, so ϕ is given by an $(m+1)$-tuple $\phi = (f_0, \ldots, f_m)$ of homogeneous polynomials of degree d. Let $Z \subset \mathbb{P}^n$ be the subset of common zeros of the f_i's. Notice that ϕ is defined on $\mathbb{P}^n \backslash Z$.*
(a) We have

$$h(\phi(P)) \leq dh(P) + O(1)$$

for all $P \in \mathbb{P}^n(\bar{\mathbb{Q}}) \backslash Z$.
(b) Let X be a closed subvariety of $\mathbb{P}^n$ with the property that $X \cap Z = \emptyset$. Then

$$h(\phi(P)) = dh(P) + O(1)$$

for all $P \in X(\bar{\mathbb{Q}})$.

Proof. The following proof is found in [5] B.2.5. Nevertheless, it is included to give the readers a glimpse of the style of proofs in height theory.

Fix a field of definition k for ϕ, write

$$\phi = (f_0, f_1, \ldots, f_m) \quad \text{with } f_0, \ldots, f_m \in k[X_0, \ldots, X_n]_d.$$

(That is, every f_i is homogeneous polynomial of degree d.) We write f_i explicitly as

$$f_i(X) = \sum_{|e|=d} a_{i,e} X^e,$$

where $e = (e_0, \ldots, e_n)$ is a multi-index, $|e| = e_0 + \cdots + e_n$, and $X^e = X_0^{e_0} \cdots X_n^{e_n}$.

For any point $P = (x_0, \ldots, x_n)$ with $x_j \in k$ and any absolute value v, we write $|P|_v = \max\{|x_j|_v\}$. Similarly, for any polynomial $f = \sum a_e X^e \in k[X]$ we let $|f|_v = \max\{|a_e|_v\}$. We also set the notation $\epsilon_v(r)$ to be r if v is archimedean and 1 otherwise. With this notation, the triangle inequality can be written uniformly as

$$|a_1 + a_2 + \cdots + a_r|_v \leq \epsilon_v(r) \max\{|a_1|_v, \ldots, |a_r|_v\}.$$

Now consider any point $P \in \mathbb{P}^n(\bar{\mathbb{Q}})$. Extending k if necessary, we may assume that $P \in \mathbb{P}^n(k)$ and write $P = (x_0, \ldots, x_n)$ with $x_i \in k$. Then for any $v \in M_k$ and any i we have

$$\begin{aligned} |f_i(P)|_v &= |\sum_{|e|=d} a_{i,e} x^e|_v \\ &\leq \epsilon_v \binom{n+d}{n} (\max_e |a_{i,e}|_v)(\max_e |x_0^{e_0} \cdots x_n^{e_n}|_v) \\ &\leq \epsilon_v \binom{n+d}{n} |f_i|_v |P|_v^d. \end{aligned}$$

Now take the maximum over $0 \leq i \leq m$, raise to the $n_v/[k:\mathbb{Q}]$ power, and multiply over all $v \in M_k$. This gives

$$H(\phi(P)) \leq \binom{n+d}{n} H(\phi) H(P)^d,$$

where we are writing $H(\phi)$ for

$$H(\phi) = \prod_{v \in M_k} \max\{|f_0|_v, \ldots, |f_m|_v\}^{n_v/[k:\mathbb{Q}]}.$$

Note that we have made use of the identity

$$\prod_{v\in M_k} \epsilon_v(r)^{n_v} = \prod_{v\in M_k^\infty} r^{n_v} = r^{[k:\mathbb{Q}]},$$

which follows from the degree formula. Taking the logarithms gives

$$h(\phi(P)) \le dh(P) + h(\phi) + \log \binom{n+d}{n},$$

which completes the proof of (a).

In order to get a complementary inequality, we need to use the fact that we are choosing points $P \in X$ and that ϕ is a morphism on X. Let $p_1, \dots, p_r$ be homogenous polynomials generating the ideal of X. Then we know that $p_1, \dots, p_r, f_0, \dots, f_m$ have no common zeros in $\mathbb{P}^n$. The Nullstellensatz tells us that the ideal they generate has a radical equal to the ideal generated by $X_0, X_1, \dots, X_n$. This means that we can find polynomials g_{ij}, q_{ij} (assumed to be homogeneous) and an exponent $t \ge d$ such that

$$g_{0j}f_0 + \cdots + g_{mj}f_m + q_{1j}p_1 + \cdots + q_{rj}p_r = X_j^t$$

for $0 \le j \le n$. Extending the fields if necessary, we may assume that the g_{ij}'s and q_{ij}'s have coefficients in k. Now let $P = (x_0, \dots, x_n) \in X(k)$. Then $p_i(P) = 0$ for all i, so when we evaluate the above formula at P we obtain

$$g_{0j}(P)f_0(P) + \cdots + g_{mj}(P)f_m(P) = x_j^t, \text{ for } 0 \le j \le n.$$

Hence

$$\begin{aligned}
|P|_v^t &= \max_j |x_j^t|_v \\
&= \max_j |g_{0j}(P)f_0(P) + \cdots + g_{mj}(P)f_m(P)|_v \\
&\le \epsilon_v(m+1)(\max_{i,j} |g_{ij}(P)|_v)(\max_i |f_i(P)|_v) \\
&\le \epsilon_v(m+1)\big[\epsilon_v \binom{t-d+n}{n} (\max_{i,j} |g_{i,j}|_v)|P|_v^{t-d}\big](\max_i |f_i(P)|_v).
\end{aligned}$$

Now raise to the $n_v/[k:\mathbb{Q}]$ power and multiply over all $v \in M_k$. This yields

$$H(P)^t \le cH(P)^{t-d}H(\phi(P)),$$

where c is a certain constant depending on f_i's, g_{ij}'s and t, but independent of P. Taking the logarithms gives the desired inequality

$$dh(P) \leq h(\phi(P)) + O(1).$$

This completes the proof of (b).

□

2.4 Divisors

Before we can talk about height functions on varieties in general, we need one more tool from algebraic geometry, namely divisor theory.

Definition 2.4.1. Let X be a smooth algebraic variety defined over field k. The *group of Weil divisors* on X is the free abelian group generated by the closed subvarieties of codimension one (that is dimension one less than the dimension of X) on X. It is denoted by $\mathrm{Div}(X)$. In other words, a divisor $D \in \mathrm{Div}(X)$ is a finite formal sum of the form $D = \sum n_Y Y$, where n_Y's are integers and Y's are codimension-one subvarieties of X.
The *support of the divisor* $D = \sum n_Y Y$, denoted by $\mathrm{supp}(D)$, is the union of Y's with $n_Y \neq 0$. The divisor is said to be *effective* or *positive* if every $n_Y \geq 0$. We write $D \geq D'$ if $D - D'$ is effective. The *degree* of a divisor $D = \sum n_Y Y$ is the integer $\sum n_Y$.

As we will only work with smooth varieties, we have an equivalent way of defining divisors, namely the *Cartier divisors*.

Definition 2.4.2. A Cartier divisor on a variety X is an (equivalence class of) collections of pairs $(U_i, f_i)_{i \in I}$ satisfying the following conditions:
(i) Then U_i's are open sets that cover X.
(ii) The f_i's are nonzero rational functions.
(iii) $f_i f_j^{-1} \in \mathcal{O}(U_i \cap U_j)^*$ for all i, j.
The *support* of a Cartier divisor is the set of zeros or poles of the f_i's.

The following definition is for Cartier divisors. For readers familiar with sheaves (see [6] II.1), it is immediate that Cartier divisors are sheaves on a variety. Sheaves are extremely

important to modern algebraic geometry as they are so-called functorial. Furthermore, results from an area named homological algebra are readily applied to sheaves. For interested readers, [13] is a great book that covers both the homological algebra foundation and its application to the derived category of coherent sheaves.

Regarding the equivalence between Weil divisors and Cartier divisors, the reader may refer to [5] A.3 which also gives a glimpse of the sheaf-theoretic language.

Recall that if Y is an irreducible divisor on X, then $\mathcal{O}_{Y,X}$ is the local ring of functions regular in a neighborhood of some point of Y. For $f \in \mathcal{O}_{Y,X}\backslash\{0\}$, then the *order of vanishing* of f along Y, written $\mathrm{ord}_Y(f)$, is the length of the module (see [2] Chapter 10 for an introduction to modules) $\mathcal{O}_{Y,X}/(f)$. We extend ord_Y to the function field $k(X)^*$ by defining $\mathrm{ord}_Y(g/h) = \mathrm{ord}_Y(g) - \mathrm{ord}_Y(h)$ for any $g, h \in \mathcal{O}_{Y,X}\backslash\{0\}$. Here is a summary of some main properties.

Lemma 2.4.3. *The order function* $\mathrm{ord}_Y : k(X)^* \to \mathbb{Z}$ *has the following properties:*
(a) $\mathrm{ord}_Y(fg) = \mathrm{ord}_Y(f) + \mathrm{ord}_Y(g)$ *for all* $f, g \in k(X)^*$.
(b) *Fix* $f \in k(X)^*$, *there are only finitely many* Y*'s with* $\mathrm{ord}_Y(f) \neq 0$.
(c) *Let* $f \in k(X)^*$, *then* $\mathrm{ord}_Y(f) \geq 0$ *if and only if* $f \in \mathcal{O}_{Y,X}$.

Proof. See [5] A.2.1.2. □

Definition 2.4.4. Let X be a variety, and let $f \in k(X)^*$ be a rational function on X. The *divisor* of f is the divisor

$$(f) = \mathrm{div}(f) = \sum_Y \mathrm{ord}_Y(f) Y \in \mathrm{Div}(X).$$

A divisor is said to be *principal* if it is the divisor of a function. Two divisors D and D' are said to be *linearly equivalent*, denoted by $D \sim D'$, if their difference is a principal divisor.

Sometimes it is convenient to write (f) as the difference of two effective divisors.

$$(f) = (f)_0 - (f)_\infty,$$

where

$$(f)_0 = \sum_{\mathrm{ord}_Y(f)>0} \mathrm{ord}_Y(f) Y, \quad (f)_\infty = \sum_{\mathrm{ord}_Y(f)<0} -\mathrm{ord}_Y(f) Y$$

Definition 2.4.5. The *divisor class group* of X is the group of divisor classes modulo linear equivalence. It is denoted by $\mathrm{Cl}(X)$. Since we are working with smooth varieties only, the Weil divisor class group is naturally isomorphic to its Cartier divisor class group (see [5] A.2.2.1) $\mathrm{Pic}(X)$. We will use the two notations interchangeably.

We can check the set $\mathrm{Pic}^0(X)$ of divisors on X of degree zero modulo linear equivalence is a subgroup of $\mathrm{Pic}(X)$. It will become important later.

Definition 2.4.6. Let $g : X \to Y$ be a morphism of varieties. Let $D \in \mathrm{Div}(Y)$ be a Cartier divisor defined by $\{(U_i, f_i)|i \in I\}$, and assume that $g(X)$ is not contained in the support of D. Then the Cartier divisor $g^*(D)$ on X is defined by

$$g^*(D) = \{(g^{-1}(U_i), f_i \circ g)|i \in I\}.$$

Remark 2.4.7. It is worth noting that if we work up to linearly equivalence, it is always possible to move the divisor D in the above definition so that g^* is defined. In fact, a morphism $f : X \to Y$ induces a group homomorphism $f^* : \mathrm{Pic}(Y) \to \mathrm{Pic}(X)$. If the reader is familiar with the categorical language, * describes a contravariant functor from the category of smooth varieties to the category of abelian groups.

Definition 2.4.8. For each divisor D on a variety X we associate the vector space

$$L(D) = \{f \in k(X)|D + \mathrm{div}(f) \geq 0\} \cup \{0\},$$

whose dimension is denoted by $l(D)$.

The set of effective divisors linearly equivalent to D is naturally parametrized by the projective space

$$\mathbb{P}(L(D)) \cong \mathbb{P}^{l(D)-1}.$$

Specifically, f mod k^* corresponds to D+div(f).

Definition 2.4.9. A *linear system* on a variety X is a set of effective divisors all linearly equivalent to a fixed divisor D and parametrized by a linear subvariety (a variety defined by linear equations) of $\mathbb{P}(L(D))$. The dimension of the linear system is the dimension of the linear subvariety. The set of effective divisors linearly equivalent to D is called the *complete linear system* of D and is denoted by $|D|$.

Definition 2.4.10. Let L be a linear system of dimension n parametrized by a linear projective variety $\mathbb{P}(V) \subset \mathbb{P}(L(D))$ where V is a vector subspace of $L(D)$. Select a basis $f_0, \ldots, f_n$ of $V \subset L(D)$. The *rational map associated to* L, denoted by ϕ_L, is the map

$$\begin{aligned} \phi_L : & X \longrightarrow \mathbb{P}^n \\ & x \mapsto (f_0(x), \ldots, f_n(x)). \end{aligned}$$

Definition 2.4.11. A linear system L on a projective variety X is *very ample* if the associated rational map $\phi_L : X \to \mathbb{P}^n$ is an embedding, that is, it is a morphism that maps X isomorphically onto its image. A divisor D is said to be *very ample* if the complete linear system $|D|$ is very ample. A divisor is said to be *ample* if some positive multiple of it is very ample.

At the end of this section, we see a theorem that is vital to the construction of Weil's height machine in the next section.

Theorem 2.4.12. *Every divisor can be written as the difference of two very ample divisors.*

Proof. See [5] A.3.2.3. □

2.5 Height on Varieties

Now we have what we need to define height functions on varieties in general.

Let K be a number field. For every smooth projective variety X/K, we build a map

$$h_X : \mathrm{Div}(X) \to \{\text{functions } X(\bar{K}) \to \mathbb{R}\}.$$

First for every very ample divisor $D \in \mathrm{Div}(X)$, the embedding $\phi_D : X \to \mathbb{P}^n$ defines a function $X(\bar{K}) \to \mathbb{R}$ by

$$h_{X,D}(P) = h(\phi_D(P)),$$

for any $P \in X(\bar{K})$. Next for any other divisor D, write $D = D_1 - D_2$ with D_1, D_2 very ample. Then we simply define

$$h_{X,D} = h_{X,D_1} - h_{X,D_2}.$$

There are many nice properties about this map called *Weil's Height Machine*.

Theorem 2.5.1. *The map h defined above has the following properties:*

- *Let $H \subset \mathbb{P}^n$ be a hyperplane, and let $h(P)$ be the absolute logarithmic height on $\mathbb{P}^n$ defined earlier. Then for all $P \in \mathbb{P}^n(\bar{K})$.*

$$h_{\mathbb{P}^n,H}(P) = h(P) + O(1)$$

- *Let $\phi : X \to Y$ be a morphism and let $D \in Div(Y)$. Then for all $P \in \mathbb{P}^n(\bar{K})$*

$$h_{X,\phi^* D}(P) = h_{Y,D}(\phi(P)) + O(1),$$

  where $\phi^*(D) = \phi^{-1}(D) \in Div(Y)$.

- *Let $D, E \in Div(X)$, then for all $P \in \mathbb{P}^n(\bar{K})$*

$$h_{X,D+E}(P) = h_{X,D}(P) + h_{X,E}(P) + O(1).$$

- *Let $D, E \in Div(X)$ with $D \sim E$. Then for all $P \in \mathbb{P}^n(\bar{K})$*

$$h_{X,D}(P) = h_{X,E}(P) + O(1).$$

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

$$\{P \in X(K') \mid h_{X,D}(P) \leq B\}$$

  is finite.

Proof. We begin by verifying up to $O(1)$, the height function $h_{X,D}$ assiciated to a divisor D is independent of the morphism ϕ_D. Let $\psi_D : X \to \mathbb{P}^m$ be another morphism associated to D. This means that $\phi_D^* H \sim D \sim \psi_D^* H'$, where H is a hyperplane in $\mathbb{P}^n$ and H' is a hyperplane in $\mathbb{P}^m$. Now it can be shown by definition that

$$h(\phi_D(P)) = h(\psi_D(P)) + O(1)$$

for all P.

After checking that the height machine is well defined up to $O(1)$, the rest of the properties follows from computations using the definition. Details can be found in [5], pages 186 to 190. □

In some cases, it is possible to find a particular height function within its $O(1)$ equivalence class that has nicer properties. For abelian varieties, this theory was developed by Néron and Tate [5] B.5.1. The construction below is due to Tate and was extended to this context by Call and Silverman in [11].

Theorem 2.5.2. *Let X/K be a smooth variety defined over a number field K, let $D \in Div(X)$ and $\phi : X \to X$ be a morphism. In addition, suppose that*

$$\phi^* D \sim \alpha D$$

for some number $\alpha > 1$. Then there is a unique function, called the canonical height on X relative to ϕ and D.

Proof. The canonical height can be computed as the limit

$$\hat{h}_{X,\phi,D}(P) = \lim_{n\to\infty} \frac{1}{\alpha^n} h_{V,D}(\phi^n(P)).$$

The proof follows from the observation that $\{\frac{1}{\alpha^n} h_{V,D}(\phi^n(P))\}$ is a Cauchy sequence. □

2.6 Counting Rational Points on Curves

Definition 2.6.1. Let C/K be a smooth curve over a number field K. Fix a multiplicative height function H on C relative to some ample divisor D. The *counting function* of $C(K)$ is

$$N(C(K), T) = \#\{P \in C(K) \mid H(P) \leq T\}.$$

Now we are ready to state a general theorem which implies Faltings' theorem and more.

Theorem 2.6.2. *Let K be a number field, let C/K be a smooth curve of genus g, and assume $C(K)$ is not empty. Then there are constants a and b, which depend on C/K and the height used in the counting function, such that*

$$N(C(K),T) \sim \begin{cases} aT^b & \text{if } g = 0 \ (a, b > 0), \\ a(\log T)^b & \text{if } g = 1 \ (a > 0, b \geq 0), \\ a & \text{if } g \geq 2. \end{cases}$$

Proof. See [5] B.6 for the proofs of genus 0 and 1 cases. The higher genus case is equivalent to Mordell's Conjecture. □

Chapter 3

Abelian Varieties

3.1 Bigger Picture

So far, our 'pictures' of interest are higher genus curves, but it is often beneficial to include them into a 'bigger picture' with more intrinsic structures. In this chapter, we will place a smooth algebraic curve into a variety called its Jacobian. A Jacobian of a curve has a group structure as it is a special type of variety called an abelian variety. Also, as a group, the rational points of abelian varieties are always finitely generated. Last, but not least, the canonical height construction in the previous chapter will show up naturally on abelian varieties.

3.2 Abelian Varieties and Jacobian Varieties

Definition 3.2.1. An *algebraic group* is a group that is also an algebraic variety, such that the group multiplication and inversion operations are given by regular maps on the variety.

Definition 3.2.2. An *abelian variety* is a projective variety that is also an algebraic group.

Two important classes of maps from an abelian variety A to itself are the translation maps, for each $a \in A$, $t_a : x \mapsto a + x$ for all $x \in A$ and, for each integer n, the multiplication maps $[n]$ which map an element in A to the sum of n copies of itself. When n is negative, we take the sign as taking the inverse. We will soon see that any abelian variety is an abelian group, which will justify the use of summation notations.

Proposition 3.2.3. *Let $\phi : A \to B$ be a morphism between two abelian varieties. Then ϕ is the composition of a translation and a homomorphism.*

Proof. See [5] A.7.1.2 for proof. □

Let A be an abelian varity. The inversion map $i : A \to A$ of the group law on A fixes the group identity. According to the proposition i has to be a homomorphism since nontrivial translation does not fix the identity. In particular,

$$xy = i(i(xy)) = i(i(x)i(y)) = i(i(y))i(i(x)) = yx.$$

From now on, we will only use additive notations for abelian varieties.

Remark 3.2.4. An important family of abelian varieties is *elliptic curves* which are genus one curves with a distinguished point on it. Though a higher genus curve is not an abelian variety by itself, it can be embedded into its *Jacobian* which is an abelian variety.

Theorem 3.2.5. *Let C be a smooth projective curve of genus $g \geq 1$. There exists an abelian variety $\mathrm{Jac}(C)$, called the Jacobian of C, and an injection $j : C \to \mathrm{Jac}(C)$, called the Jacobian embedding of C, with the following properties:*
(i) *Extend j linearly to divisors on C. Then j induces a group isomorphism between $\mathrm{Pic}^0(C)$ and $\mathrm{Jac}(C)$.*
(ii) *For each $r \geq 0$, define a subvariety $W_r \subset \mathrm{Jac}(C)$ by*

$$W_r = [r]j(C),$$

where for any integer n, $[n]$ denotes summation of n copies. Then

$$\dim(W_r) = \min(r, g), W_g = \mathrm{Jac}(C).$$

In particular, $\dim(\mathrm{Jac(C)}) = g$.
(iii) *Let $\Theta = W_{g-1}$. Then Θ is an irreducible ample divisor on $\mathrm{Jac}(C)$.*

Proof. See [5] A.8.1 for the details of the construction of the Jacobian embedding. See A.8.2 of the same book for properties of Θ. □

For our purposes, we assume $C(K)$ is nonempty (otherwise $C(K)$ is trivially finite). The following proposition makes the embedding j concrete.

Proposition 3.2.6. *Let C be a curve of genus g, let $P_0 \in C(K)$, and let $J = \mathrm{Jac}(C) \cong \mathrm{Pic}^0(C)$ be the Jacobian variety of C. We have an embedding $j : C \to J$ such that a point P is sent to the divisor class of $(P) - (P_0)$.*

Proof. See [5] A.8.2. □

3.3 Mordell-Weil Theorem

As we hope to study the rational points on higher genus smooth curves and, by the previous proposition, these points embed in their Jacobian, it is natural to look at the rational points on the Jacobian. An important result for rational points of abelian varieties in general is the Mordell-Weil theorem.

Theorem 3.3.1 (Mordell-Weil)**.** *Let A be an abelian variety defined over a number field K. Then the group $A(K)$ of K-rational points of A is finitely generated.*

We can rephrase the theorem by saying that there are point $P_1, \ldots, P_r \in A(K)$ such that

$$A(K) = A(K)_{\mathrm{tors}} \oplus \mathbb{Z}P_1 \oplus \cdots \oplus \mathbb{Z}P_r,$$

where

$$A_{\mathrm{tors}} := \{P \in A(K) : [n_p]P = 0 \text{ for some } n_p \in \mathbb{N}\}.$$

The integer r is called the *rank* of the abelian variety A/K, and $A(K)$ is the *Mordell-Weil group* of A/K. Note that the torsion subgroup $A(k)_{\mathrm{tors}}$ is a finite abelian group.

Proof. See [5] Section C for a general proof. See [8] Chapter VIII for a less general but still elegant proof. Additionally, Cassels classic book [12] provides a proof of the Mordell-Weil theorem for elliptic curves over $\mathbb{Q}$. The latter book requires much less algebra than the former two and has the advantage of a less daunting thickness. □

3.4 Canonical Height

Theorem 3.4.1 (Mumford's formula)**.** *Let D be a divisor on an abelian variety A, and let $[n] : A \to A$ be the multiplication-by-n map. Then*

$$[n]^*(D) \sim \frac{n^2+n}{2}D + \frac{n^2-n}{2}[-1]^*(D).$$

In particular, $[n]^*(D) \sim n^2 D$ *if* D *is symmetric* ($[-1]^*D \sim D$) *and* $[n]^*(D) \sim nD$ *if* D *is antisymmetric* ($[-1]^*D \sim -D$).

Proof. See [5] A.7.2.5. □

Theorem 3.4.2 (Néron, Tate)**.** *Let* A/K *be an abelian variety defined over a number field, and let* $D \in Div(A)$ *be a divisor whose divisor class is symmetric. There is a height function*

$$\hat{h}_{A,D} : A(\bar{K}) \to \mathbb{R},$$

called the canonical height on A *relative to* D*, with the following properties:*

(a)

$$\hat{h}_{A,D}(P) = h_{A,D}(P) + O(1) \quad \text{for all } P \in A(\bar{K}).$$

(b) *For all integer* m,

$$\hat{h}_{A,D}([m]P) = m^2 \hat{h}_{A,D}(P) \quad \text{for all } P \in A(\bar{K}).$$

(c)

$$\hat{h}_{A,D}(P+Q) + \hat{h}_{A,D}(P-Q) = 2\hat{h}_{A,D}(P) + 2\hat{h}_{A,D}(Q) \quad \text{for all } P, Q \in A(\bar{K}).$$

(d) *The canonical height map is a quadratic form. The associated pairing* $\langle \cdot, \cdot \rangle_D : A(\bar{K}) \times A(\bar{K}) \to \mathbb{R}$ *defined by*

$$\langle P, Q \rangle_D = \frac{\hat{h}_{A,D}(P+Q) - \hat{h}_{A,D}(P) - \hat{h}_{A,D}(Q)}{2}$$

is bilinear and satisfies $\langle P, P \rangle = \hat{h}_{A,D}(P)$.

(e) *The canonical height depends only on the divisor class of the divisor* D*. It is uniquely determined by* (a) *and* (b) *for any integer* $m \geq 2$.

Proof. We take $\hat{h}_{A,D}$ to be the canonical height on A with respect to the doubling map $[2] : A \to A$. Note that $[2]^*D \sim 4D$, so we can apply the earlier theorem to obtain

$$\hat{h}_{A,D}(P) = \lim_{n \to \infty} \frac{1}{4^n} h_{A,D}([2^n]P).$$

The rest follows from the properties of the height machine.

For more details, please refer to [5] B.5. □

Chapter 4

Mordell's Conjecture

4.1 Introduction

In this chapter, we will see the key inequality by Vojta that implies Faltings' theorem. While the proof of the Vojta's inequality is involved, Faltings' theorem follows from it essentially by linear algebra. We will give the proof of Faltings' theorem from Vojta's inequality.

4.2 Vojta's Inequality

Theorem 4.2.1 (Vojta)**.** *Let C/K is a smooth projective curve of genus $g \geq 2$ defined over a number field K. Let J be the Jacobian variety of C and Θ the theta divisor on J. Let $|\cdot|$ and $\langle \cdot, \cdot \rangle$ be the norm and bilinear form on $J(\bar{K})$ associated to the canonical height relative to Θ. We choose a rational point in $C(K)$ and use it to fix an embedding $C \to J$. There exist constants $k_1 = k_1(C)$ and $k_2 = k_2(g)$ such that if $z, w \in C(\bar{K})$ are two points satisfying*

$$|z| \geq k_1 \text{ and } |w| \geq k_2|z|,$$

then

$$\langle z, w \rangle \leq \frac{3}{4}|z||w|.$$

First and foremost, let's see a profound implication of this inequality.

Proposition 4.2.2. *Let C/K be a curve of genus at least 2 defined over a number field K. Then Vojta's inequality implies Faltings' theorem that $C(K)$ is finite.*

Proof. We observe the kernel of the map $J(K) \to J(K) \otimes \mathbb{R}$ is the torsion subgroup $J(K)_{\text{tors}}$, which is finite. To prove that $C(K)$ is finite, it suffices to show that the image of $C(K)$ in $J(K) \otimes \mathbb{R}$ is finite. As such we denote the image by $C(K)$ as well.

The bilinear form $\langle \cdot, \cdot \rangle$ makes $V = J(K) \otimes \mathbb{R}$ into a finite-dimensional inner product space (or Euclidean space). For $x, y \in V$ we define the 'angle' $\theta(x, y)$ between x and y,

$$\cos \theta(x, y) = \frac{\langle x, y \rangle}{|x||y|}, \quad 0 \leq \theta(x, y) \leq \pi.$$

For any point x_0 and any angle θ_0, we consider the cone with interior angle $2\theta_0$ whose central axis is the ray from 0 through x_0,

$$\Gamma_{x_0, \theta_0} = \{x \in V | \theta(x, x_0) < \theta_0\}.$$

Assume $\#(\Gamma_{x_0,\theta_0} \cap C(K)) = \infty$. Since $J(K)$, and hence $C(K)$, contains only finitely many points of bounded norm, we can find a $z \in \Gamma_{x_0,\theta_0} \cap C(K)$ with $|z| \geq k_1$, and then we find $w \in \Gamma_{x_0,\theta_0} \cap C(K)$ with $|w| \geq k_2|z|$. Vojta's inequality tells us that

$$\langle z, w \rangle \leq \frac{3}{4}|z||w|,$$

or equivalently,

$$\cos \theta(z, w) \leq \frac{3}{4}.$$

As $z, w \in \Gamma_{x_0,\theta_0} \cap C(K)$ so

$$2\theta_0 > \theta(z, w) \geq \cos^{-1} \frac{3}{4} > \frac{\pi}{6}.$$

In particular $\Gamma_{x_0,\pi/12} \cap C(K)$ is finite for every $x_0 \in V$. It is possible to cover V by finitely many cones of the form $\Gamma_{x_0,\pi/12}$. Therefore $C(K)$ is finite.

□

4.3 Proof of Vojta's Inequality

For the rest of the chapter, we will present a sketch of the proof. In order for the proof to be complete, we need Siegel's theorem and Roth's theorem. The statements of the theorems are included and the readers are welcome to read more about them in [5] Chapter D.

One of the main obstacles to our task is the fact that Weil heights are defined only up to $O(1)$. Recall that if D is a divisor and we want to choose a particular height function $h_{X,D}$, we need to make the following choices:

(1) Choose D_1, D_2 very ample divisors such that $D = D_1 - D_2$.

(2) Choose embeddings $\phi_1 : X \to \mathbb{P}^n$ and $\phi_2 : X \to \mathbb{P}^m$ corresponding to D_1 and D_2 respecitvely.

(3) Set $h_{X,D}(P) = h(\phi_1(P)) - h(\phi_2(P))$.

To prove Vojta's inequality we need to choose height functions in a 'uniform' manner so that we can keep track of how the $O(1)$'s depend on certain parameters.

We begin by fixing a divisor $A \in \mathrm{Div}(C)$ of degree 1. Take A such that

$$(2g-2)A \sim \mathcal{K}_C,$$

where $\mathcal{K}_C$ is the canonical divisor class on C. Of course A may not be defined over the field K, but it will be defined over a finite extension of K. For our proof we are free to replace K by a finite extension.

The divisor A defines an embedding

$$j_A : C \to J, \quad x \mapsto \mathrm{Cl}((x) - A).$$

Let $\Theta_A = [g-1]j_A(C)$ be a theta divisor on J. The choice of A implies the following results.

Lemma 4.3.1. *With the notations above and let $s_{12}, p_1, p_2 : J \times J \to J$ be the summation and projection maps and Δ be the diagonal divisor on $C \times C$, we have:*
(1) Θ_A is a symmetric divisor.
(2) $j_A^* \Theta_A \sim gA$.
(3) $(j_A \times j_A)^*(s_{12}^* \Theta_A - p_1^* \Theta_A - p_2^* \Theta_A) \sim -\Delta + p_1^* A + p_2^* A$.

Proof. See [5] E.2.1. □

Remark 4.3.2. We will often treat C as a subvariety of J. Moreover, the canonical height $\hat{h}_J$ and its corresponding norm and inner product on J will always be with respect to the divisor Θ_A. For example when we refer to the canonical height $\|x\|^2$ of a point $x \in C(\bar{K})$, we really mean the canonical height $\hat{h}_{J,\Theta_A}$ of the point $j_A(x)$.

To continue we first recall the following fact about smooth projective curves.

Proposition 4.3.3. *Let C be a smooth projective curve of genus g and let $D \in Div(C)$.* (i) *If $\deg(D) \geq 2g+1$, then D is very ample.* (ii) $\deg(D) > 1$ *if and only if D is ample.*

Proof. [5] A.4.2.4. □

In particular, A is ample. Choose $N \geq 2g+1$, so NA is very ample. Fix an embedding

$$\phi_{NA} : C \to \mathbb{P}^n.$$

We consider divisors on $C \times C$. We have slices $A \times C$ and $C \times A$ as well as the diagonal Δ. Now we are ready to define the divisors and height functions that will occupy most of our attention. For any given integer d_1, d_2 and d, we define

$$\Omega = \Omega(d_1, d_2, d) = (d_1 - d)(A \times C) + (d_2 - d)(C \times A) + d\Delta \in \mathrm{Div}(C \times C).$$

We say that Ω is a *Vojta divisor* if d_1, d_2 and d are positive integers satisfying the inequalities

$$gd^2 < d_1 d_2 < g^2 d^2.$$

The key height function we will study is $h_{C\times C,\Omega(d_1,d_2,d)}$.

4.4 An Upper Bound

Proposition 4.4.1. *There is a constant c_1 such that for all positive integers d_1, d_2, d all points $z, w \in C(\bar{K})$,*

$$h_{C\times C,\Omega(d_1,d_2,d)}(z,w) \leq \frac{d_1}{g}|z|^2 + \frac{d_2}{g}|w|^2 - 2d\langle z, w\rangle + c_1(d_1 + d_2 + d).$$

Proof. See [5] E.4.1. □

If we apply the proposition to Vojta divisors, we can 'prove' Vojta's inequality. In fact the condition $d_1 d_2 > g d^2$ implies that Ω is effective. The proof of this fact relies on the Riemann-Roch theorem (see [8] II.5) which is beyond the scope of this paper. As Ω is effective, the height is bounded below by some constant $-c_2$. The above proposition gives

$$-c_2 \leq \frac{d_1}{g}|z|^2 + \frac{d_2}{g}|w|^2 - 2d\langle z, w\rangle + c_1(d_1 + d_2 + d).$$

A little algebra gives us the estimate

$$\langle z, w\rangle \leq \frac{1}{2d}\left(\frac{d_1}{g}|z|^2 + \frac{d_2}{g}|w|^2\right) + O\left(\frac{d_1}{d} + \frac{d_2}{d} + 1\right).$$

Assuming $|z|, |w|$ are large, we choose the integers d_1, d_2, d to satisfy

$$d_1 \approx \sqrt{g}|w|^2, \quad d_2 \approx \sqrt{g}|z|^2, \quad d \approx |z||w|.$$

Then we have

$$\langle z, w\rangle \leq \frac{|z||w|}{\sqrt{g}} + O\left(\frac{|w|}{|z|} + \frac{|z|}{|w|}\right).$$

If $|z|, |w|$ are sufficiently large, the error term is small in comparison to the main term, we get

$$\langle z, w\rangle \leq (1 + \epsilon)\frac{|z||w|}{\sqrt{g}}.$$

As $g \geq 2$, we finally arrive at the desired bound

$$\langle z, w\rangle \leq \frac{3}{4}|z||w|.$$

Of course, this has to be too simple to be correct. The above proof has two gaps:

(1) The lower bound $-c_2$ depends on the choice of the Vojta divisor.

(2) The above estimates only works for points (z, w) not lying on the divisor Ω.

The two gaps are bridged by the use of Siegel's lemma and Roth's theorem respectively.

Lemma 4.4.2 (Siegel)**.** *Let K be a number field with $d = [K : \mathbb{Q}]$, let $a_{ij} \in K$ be elements not all zero, and let $A = H(\ldots, a_{ij}, \ldots)$ be the height of the vector formed by the a_{ij}'s. Given integers M, N such that $dM < N$, there exists a nonzero vector $x \in \mathbb{Z}^N$ such that*

$$\sum_{i=1}^{N} a_{ij}x_i = 0 \quad \textit{for all } 1 \leq j \leq M \quad \textit{and} \quad \max_{1\leq i\leq N} |x_i| \leq (NA)^{dM/(N-dM)}.$$

Proof. See [5] D.4.2. □

Theorem 4.4.3 (Roth)**.** *Let K be a number field, let $S \subset M_K$ be a finite set of absolute values on K, and assume that each absolute value in S has been extended in some way to $\bar{K}$. Let $\alpha \in \bar{K}$ and $\epsilon > 0$ be given. Then there are only finitely many $\beta \in K$ satisfying the inequality*

$$\prod_{v \in S} \min\{\|\beta - \alpha\|_v, 1\} \leq \frac{1}{H_K(\beta)^{2+\epsilon}}.$$

Proof. See [5] Chapter D for a self-contained proof. □

Interested readers are encouraged to look into [5] for the bridging of these gaps. Nevertheless, we hope this sketch captures the essence of the proof.

Bibliography

[1] J. Munkres. *Topology,* Pearson, 2000.

[2] D. Dummit and R. Foote. *Abstract Algebra,* Wiley, 2003.

[3] K. Smith, L. Kahanpää, P. Kekäläinen, and W. Traves. *An Invitation to Algebraic Geometry,* Universitext, Springer, 2000.

[4] I. Shafarevich. *Basic Algebraic Geometry 1: Varieties in Projective Space,* Springer, 2013.

[5] M. Hindry and J. Silverman. *Diophantine Geometry,* Graduate Texts in Mathematics 201, Springer, 2000.

[6] R. Hartshorne. *Algebraic Geometry,* Graduate Texts in Mathematics 52, Springer, 1977.

[7] S. Donaldson. *Riemann Surfaces,* Oxford Graduate Texts in Mathematics, Oxford, 2011.

[8] J. Silverman. *The Arithmetic of Elliptic Curves,* Graduate Texts in Mathematics 106, Springer, 1986.

[9] J. Silverman. *Advanced Topics in the Arithmetic of Elliptic Curves,* Graduate Texts in Mathematics 151, Springer, 1999.

[10] J. Neukirch. *Algebraic Number Theory,* Grundlehren der mathematischen Wissenschaften, Springer, 1999.

[11] G. Call and J. Silverman. *Canonical heights on varieties with morphisms,* Compos. Math. 89 (1993), 163-205.

[12] J. Cassels. *Lectures on Elliptic Curves,* London Mathematical Society Student Texts 24, Cambridge, 1991.

[13] D. Huybrechts. *Fourier-Mukai Transforms in Algebraic Geometry,* Oxford Mathematical Monographs, Oxford, 2006.

[14] D. Cox. *Galois Theory,* Wiley, 2012.

[15] D. Eisenbud and J. Harris. *The Geometry of Schemes,* Graduate Texts in Mathematics 197, Springer, 2001.