UNIVERSITY OF
GOTHENBURG

Effective quasiparallelogram laws on elliptic curves over number fields

Master's thesis in Mathematics

DOUGLAS MOLIN

Department of Mathematical Sciences
UNIVERSITY OF GOTHENBURG
Gothenburg, Sweden 2021

Master's thesis 2021

Effective quasiparallelogram laws on elliptic curves over number fields

DOUGLAS MOLIN

UNIVERSITY OF
GOTHENBURG

Department of Mathematical Sciences
Division of Algebra and Geometry
University of Gothenburg
Gothenburg, Sweden 2021

Effective quasiparallelogram laws on elliptic curves over number fields

DOUGLAS MOLIN

Supervisor: Per Salberger, Department of Mathematical Sciences
Examiner: Jan Stevens, Department of Mathematical Sciences

Master's Thesis 2021
Department of Mathematical Sciences
Division of Algebra and Geometry
University of Gothenburg
SE-412 96 Gothenburg

Effective quasiparallelogram laws on elliptic curves over number fields

DOUGLAS MOLIN
Department of Mathematical Sciences
University of Gothenburg

Abstract

We introduce the classical theory of heights on projective space and prove explicit quasiparallelogram laws for the ordinary height and the naive height on elliptic curves over number fields with short Weierstrass equations. As corollaries, we obtain bounds for the differences between the classical heights and the canonical height, similar to the well-known Silverman bounds. The results are analyzed through a number of examples.

Keywords: height, elliptic curve, quasiparallelogram law, canonical height, difference bounds.

Acknowledgements

I would like to express my deep gratitude to my advisor Professor Per Salberger for guiding me and sharing his insight and wisdom. I am also thankful to Professors John Cremona and Samir Siksek for providing useful remarks about their research.

Although mine is the only name written on the title page of this thesis, without my community I would be nobody and these pages blank. Therefore, I wish to thank my friends and family, especially Vincent Molin, for their support and encouragement.

Douglas Molin, Gothenburg, May 2021

Contents

1 Introduction **1**

2 Preliminaries **3**
2.1 Valuations on number fields 3
2.2 Elliptic curves 4

3 Heights **7**
3.1 Heights on projective space 7
3.2 Heights on elliptic curves 9

4 The Quasiparallelogram Law **11**
4.1 Modified height 11
4.2 Ordinary height 16
4.3 Naive height 20
4.4 Examples 22

5 The Canonical Height **25**
5.1 Construction of $\hat{h}$ 25
5.2 Difference bounds 28

Bibliography **31**

1

Introduction

Elliptic curves over number fields have played a central role in modern number theory and arithmetic geometry. They are curves defined by equations of the form

$$y^2 = x^3 + Ax + B, \tag{$\star$}$$

where A, B lie in some number field K. On such a curve, there is an addition law on the set of solutions $E(K)$ defined by a chord-and-tangent construction. The ultimate goal is to determine all points $(x, y) \in K^2$ satisfying the equation, and one hopes to use the group structure on $E(K)$ to get there. The celebrated Mordell-Weil theorem marks a crucial step in the process of accomplishing this.

Mordell-Weil Theorem. $E(K)$ is a finitely generated abelian group.

In essence, what the theorem says is that the data of some finite set of points determines $E(K)$. In other words, these generators of the *Mordell-Weil group* $E(K)$ form building blocks, and any point on the curve can be described by arithmetic formulas involving only them. The natural question to ask is how to determine these generators, and a complete answer is still to be found. Perhaps the most famous open problem in number theory aside from the Riemann Hypothesis is the conjecture due to Birch and Swinnerton-Dyer about generators of $E(K)$.

In order to derive the Mordell-Weil theorem from its weak form, André Weil conceived of the notion of *heights*, or *height functions*. These are real-valued functions defined on the curve that are thought of as measuring the arithmetic complexity of points. As an example, the numbers

$$\frac{1}{2} \text{ and } \frac{1000001}{2000000}$$

lie close on the number line, but the latter is certainly more complicated in terms of arithmetic because of its large numerator and denominator, so its height should be larger. This basic idea of comparing fractions leads to a naive definition of heights on varieties. When working with general number fields, ambiguities arise due to the lack of unique factorization, whence one is obliged to use the theory of valuations to generalize the naive definition.

The conception of height functions marks the beginning of the study of quantitative questions in diophantine geometry. In this thesis, two quantitative diophantine aspects of elliptic curves are investigated, and they both relate to the group law

mentioned above. The first of these is the *quasiparallelogram law* satisfied by height functions on elliptic curves. The name comes from an ancient theorem of geometry; the sum of the squares of the lengths of the four sides of a parallelogram equals the sum of the squares of the lengths of the diagonals. The quasiparallelogram is instead an inequality, and it reads as follows. Let E be the curve given in $(\star)$ above, and $h : E \to \mathbb{R}$ a height function. Then there is a constant $C_{A,B}$ such that for any $P, Q \in E(K)$

$$\big|h(P+Q) + h(P-Q) - 2h(P) - 2h(Q)\big| \leq C_{A,B}$$

Here $P \pm Q$ denotes the (chord-and-tangent) addition (subtraction) on E. Hence, the quasiparallelogram law rules that on an elliptic curve, the height function h satisfies the parallelogram law with respect to the group law, up to some bounded amount. There are at least two distinct interpretations. On the one hand, one may interpret the law to be essentially a statement about dynamics, encoding information about cancellation occurring in the formulas for addition. Another perspective, more abstract perhaps, is to regard the law as a statement about some form of arithmetic near-symmetry. This sounds vague at first, but becomes clear in light of the existence of the *canonical height*. This additional height function was discovered by John Tate and independently by André Néron, and it satisfies the parallelogram law exactly. Moreover, it differs from the classical height of Weil by a bounded amount. Consequently, we may perceive the quasiparallelogram law as an almost symmetric shadow cast by the perfectly regular canonical height.

The second aim of this thesis is to introduce the canonical height. Explicit bounds for the difference between the canonical height and the usual heights are of great interest as they are needed to compute generators of the Mordell-Weil group. The canonical height decomposes into a sum of so called local height functions, and one can find height difference bounds by working with this decomposition, which has been done by Silverman and others. The estimates in this thesis do not rely on the local decomposition; instead, we follow Zimmer who wrote extensively on the topic. Interestingly, the results of this thesis do not require any sophisticated techniques or results, yet are comparable to those found by Silverman.

The thesis is structured as follows. In chapter 2, we recall some basic concepts and theorems from the theories of number fields and elliptic curves. In the subsequent chapter, we construct and describe the naive height h on projective space and the ordinary height h_x on elliptic curves. Chapter 4 is devoted entirely to the proof of the quasiparallelogram laws of h and h_x. Completely explicit estimates are given for both and lastly, a number of examples are discussed. In the final chapter, we construct the canonical height $\hat{h}$ and derive bounds for the differences $h_x - \hat{h}, h - \frac{3}{2}\hat{h}$.

2

Preliminaries

In this chapter, we recall some basic facts and definitions from the theory of valuations on number fields and that of elliptic curves. Our exposition is brief, containing only the bare minimum of what we shall need. For a detailed account of the topics below, see [2] and [6].

2.1 Valuations on number fields

Throughout this thesis, K will be a number field of degree $n = [K : \mathbb{Q}]$ with fixed algebraic closure $\bar{K}$. We write M_K for the standard set of valuations on K, normalized so that every $v \in M_K$ restricts to some $-\log|\cdot|_p$ on $\mathbb{Q}$, where p is a prime or ∞. There are exactly n archimedean valuations, each arising from an embedding $\tau_v : K \hookrightarrow \mathbb{C}$ via the formula $v(x) = -\log|\tau_v(x)|$, where $|\cdot|$ denotes the usual absolute value on $\mathbb{C}$. We denote the set of archimedean valuations by M_K^∞. The set M_K^0 of nonarchimedean valuations is in bijection with the prime ideals of the ring of integers $\mathcal{O}_K$.

If L/K is a finite extension, we will write $w|v$ if w extends v. In other words, τ_w restricts to τ_v on K in the archimedean case, or the corresponding prime ideals satisfy $\mathfrak{p}_w\mathcal{O}_L \subseteq \mathfrak{p}_v\mathcal{O}_L$ in the nonarchimedean case. The local degree at a place v is denoted by $n_v = [K_v : \mathbb{Q}_v]$, and satisfies the following **extension formula**.

Proposition 2.1.1. *Let L/K be a finite extension, and let $v \in M_K$. Then*

$$\frac{1}{[L:K]} \sum_{\substack{w \in M_L \\ w|v}} n_w = n_v$$

Local degrees occur as multiplicities in the **sum formula** (or equivalently, the product formula).

Proposition 2.1.2. *Let $x \in K \setminus \{0\}$. Then*

$$\sum_{v \in M_K} -n_v v(x) = 0.$$

Equivalently,

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

Lastly, we recall the indispensable **triangle inequality**. In terms of valuations, it reads

$$v(x_1 + \cdots + x_k) \geq \min\{v(x_1), \ldots, v(x_k)\} + \epsilon_v \log k$$

where $\epsilon_v = 0$ if v is nonarchimedean, and -1 otherwise.

2.2 Elliptic curves

Throughout this thesis, we will work with elliptic curves defined by (short) **Weierstraß equations**

$$E : y^2 = x^3 + Ax + B,$$

with coefficients $A, B \in K$ satisfying $4A^3 + 27B^2 \neq 0$ to ensure smoothness. By setting $x = \frac{X}{Z}, y = \frac{Y}{Z}$ we pass to the projective closure of E, defined by the homogeneous Weierstraß equation

$$Y^2Z = X^3 + AXZ^2 + BZ^3.$$

Clearly, E has a single point at infinity $O := [0, 1, 0]$.

Proposition 2.2.1. *Short Weierstraß equations are unique up to isomorphisms of the form*

$$E : y^2 = x^3 + Ax + B \to E' : y^2 = x^3 + \rho^4 Ax + \rho^6 B$$
$$(x, y) \mapsto (\rho^2 x, \rho^3 y)$$

where $\rho \in K^*$.

The **discriminant** and j**-invariant** of E are defined as

$$\Delta = -16(4A^3 + 27B^2), \; j = -1728\frac{(4A)^3}{\Delta}.$$

The j-invariant is preserved under the isomorphisms described in Proposition 2.2.1, while the discriminant is multiplied by ρ^{12}.

Every elliptic curve carries the structure of an abelian group with O as its neutral element, and furthermore the addition law on E is defined in terms of regular functions on E with coefficients in K. In other words, the set of K-rational points $E(K)$ is closed under addition. In Chapter 4 we will analyze the group law and accordingly we will need a number of formulas and identities, which we record here.

Assume $P \neq \pm Q$. Then

$$x_{P\pm Q} = \frac{x_P^2 x_Q + x_P x_Q^2 + A(x_P + x_Q) \mp 2y_P y_Q + 2B}{(x_P - x_Q)^2} \tag{2.1}$$

$$x_{P+Q}x_{P-Q} = \frac{(x_P x_Q - A)^2 - 4B(x_P + x_Q)}{(x_P - x_Q)^2} \tag{2.2}$$

$$x_{P+Q} - x_{P-Q} = \frac{4y_P y_Q}{(x_P - x_Q)^2}. \tag{2.3}$$

If $2P := P + P \neq O$, we have

$$x_{2P} = \frac{x_P^4 - 2Ax_P^2 - 8Bx_P + A^2}{4y_P^2}. \tag{2.4}$$

Furthermore

$$P = (x, y) \iff -P = (x, -y), \tag{2.5}$$

and in particular,

$$2P = O \iff y = 0. \tag{2.6}$$

Remark 2.2.2. It should be noted that practically everything we will do can be done over general global fields (at least of characteristic not 2 or 3). In other words, K could be replaced by the function field of a curve over a finite field k. In this case, an elliptic curve over K is a certain kind of surface over k. This allows for a more geometric approach to some of the topics we shall be discussing in the sequel. For this reason, together with the fact that the cases $\operatorname{char} K = 2, 3$ require special treatment and finally, since the primary case of interest is $K = \mathbb{Q}$ anyway, we restrict our attention to number fields.

Remark 2.2.3. Even though any elliptic curve over a number field is isomorphic to a curve with short Weierstraß equation, there are important parts of the general theory that necessitate the consideration of long Weierstraß equations

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

We discuss long Weierstraß equations briefly in Remark 5.2.8.

3

Heights

In this chapter, we introduce the classical height functions. These real-valued functions are thought of as measuring the arithmetic complexity or size of a point in some space defined by polynomial equations. Such a function yields a language for quantitative aspects of diophantine equations, in particular asymptotics. There is no all-encompassing formal definition of the term *height*. Instead, the term is used for a number of different functions that in some way fit the description given above. In this sense, the concept of height is defined in terms of what it does, not what it is.

3.1 Heights on projective space

Let $P \in \mathbb{P}^n(\mathbb{Q})$. Then we can find homogeneous coordinates $[x_0, \dots, x_n] = P$ such that $x_j \in \mathbb{Z}$ and $\gcd(x_0, \dots, x_n) = 1$. These coordinates are unique up to a sign, and a natural measure of the size of P is

$$H_{\mathbb{Q}}(P) = \max\{|x_0|, \dots, |x_n|\}.$$

Now, we run into trouble when we want to extend this naive definition to number fields with class number different from 1, meaning their rings of integers are not unique factorization domains. Indeed, in this case there is no non-arbitrary choice of homogeneous coordinates as above. To accommodate this ambiguity, we need to take into account all embeddings into $\mathbb{C}$, or rather all valuations on K. Additionally, it will be convenient in the sequel to have a height that behaves additively instead of multiplicatively.

Definition 3.1.1. Let $P = [x_0, \dots, x_n] \in \mathbb{P}^n(K)$. The (logarithmic) height of P relative to K is

$$h_K(P) = \sum_{v \in M_K} -n_v \min\{v(x_0), \dots, v(x_n)\}.$$

To see that this is a reasonable definition, a few things have to be verified.

Proposition 3.1.2.

(a) *h_K is a well-defined function $\mathbb{P}^n(K) \to \mathbb{R}_{\geq 0}$.*
(b) *Let L/K be a finite extension. Then $h_L = [L : K]h_K$.*
(c) *In the case $K = \mathbb{Q}$, the definition of $h_{\mathbb{Q}}$ agrees with $\log H_{\mathbb{Q}}$ where $H_{\mathbb{Q}}$ is defined as above.*
(d) *For any $c \in \mathbb{R}$, the set $\{P \in \mathbb{P}^n(K) \mid h_K(P) \leq C\}$ is finite.*

Proof. Let $P = [x_0, \dots, x_n] \in P^n(K)$.
(a) Any other choice of homogeneous coordinates is of the form $[\lambda x_0, \dots, \lambda x_n]$ where $\lambda \neq 0$. By the sum formula,

$$\begin{aligned}\sum_{v \in M_K} -n_v \min\{v(\lambda x_0), \dots, v(\lambda x_n)\} &= \sum_{v \in M_K} -n_v \big(v(\lambda) - \min\{v(x_0), \dots, v(x_n)\}\big) \\ &= \sum_{v \in M_K} -n_v \min\{v(x_0), \dots, v(x_n)\}.\end{aligned}$$

To prove nonnegativitity, note that since at least one of $x_0, \dots, x_n$ is non-zero, by scaling we may assume one of them to be 1. It follows that $\min\{v(x_0), \dots, v(x_n)\} \leq 0$ for every $v \in M_K$, which proves the claim.
(b) The claim follows from the extension formula 2.1.1. Indeed,

$$\begin{aligned}h_L(P) &= \sum_{w \in M_L} -n_w \min\{w(x_0), \dots, w(x_n)\} \\ &= \sum_{v \in M_K} \sum_{\substack{w \in M_L \\ w|v}} -n_w \min\{v(x_0), \dots, v(x_n)\} \\ &= \sum_{v \in M_K} -[L:K] n_v \{v(x_0), \dots, v(x_n)\} = [L:K] h_K(P).\end{aligned}$$

(c) If $P \in \mathbb{P}^n(\mathbb{Q})$ we can assume $x_0, \dots, x_n \in \mathbb{Z}$ and furthermore $\gcd(x_0, \dots, x_n) = 1$. Hence for every discrete v, we have $v(x_j) = 0$ for at least one j. Thus, only the archimedean valuation v contributes a non-zero amount in the sum, whereby

$$\begin{aligned}h_{\mathbb{Q}}(P) &= -\min\{-\log|x_0|, \dots, -\log|x_n|\} \\ &= \max\{\log|x_0|, \dots, \log|x_n|\} = \log H_{\mathbb{Q}}(P).\end{aligned}$$

(d) See [6, Theorem VIII.5.11]. □

Often, it is convenient to have a height function that is independent of the underlying field. In light of Proposition 3.1.2(c), the following definition makes sense.

Definition 3.1.3. Let $P \in \mathbb{P}^n(\bar{K})$. The (absolute logarithmic) height of P is

$$h(P) = \frac{1}{[L:\mathbb{Q}]} h_L(P),$$

where L is any field such that $P \in \mathbb{P}^n(L)$.

When $\alpha \in K$, we write $h(\alpha) = h([\alpha, 1])$.

Theorem 3.1.4. [6, Theorem VIII.5.9] *Let*

$$f = \sum_{k=0}^{d} a_k T^k = a_d \prod_{\ell=1}^{d} (T - \alpha_\ell) \in \bar{K}[T]$$

be a polynomial of degree d. Then

$$\sum_{j=1}^{d} h(\alpha_j) - d \log 2 \leq h([a_0, \dots, a_d]) \leq \sum_{j=1}^{d} h(\alpha_j) + (d-1) \log 2.$$

To conclude this section, we relate heights to maps. Recall that a morphism $g : \mathbb{P}^m \to \mathbb{P}^n$ is defined by

$$g\big([x_0, \dots, x_m]\big) = \big[g_0([x_0, \dots, x_m]), \dots, g_n([x_0, \dots, x_m])\big]$$

where $g_0, \dots, g_n \in \bar{K}[X_0, \dots, X_m]$ are homogeneous polynomials of the same degree d, with no common zero. We say that g has degree d.

Theorem 3.1.5. [6, Theorem VIII.5.6] *Let $g : \mathbb{P}^m \to \mathbb{P}^n$ be a morphism of degree d. Then for all $P \in \mathbb{P}^m(\bar{K})$,*

$$h(g(P)) = dh(P) + O_g(1).$$

Remark 3.1.6. Bounding $h(g(P))$ from above is straightforward since if the coordinates of P are small, $g_j(P)$ must be too and this is so regardless of whether g is a morphism or not. On the other hand, the lower bound is more subtle; of course, the value of a polynomial at a point P can be small even if the coordinates are large. Here the assumption that g is a morphism comes in, allowing us to use Hilbert's Nullstellensatz to express the coordinates x_j in terms of the g_j.

3.2 Heights on elliptic curves

The homogeneous Weierstraß equation defines an embedding $E \hookrightarrow \mathbb{P}^2$ and by simply restricting the $\mathbb{P}^2$-height h we obtain the **naive height** on E, also denoted by h.

The projection map

$$\begin{aligned} E(\bar{K}) \setminus \{O\} &\to \bar{K} \\ [x, y, 1] &\mapsto x \end{aligned}$$

is regular, and induces a morphism $x : E \to \mathbb{P}^1$ by setting $x(O) = [0, 1]$. We define the **ordinary height** of a point $P \in E(\bar{K})$ as

$$h_x(P) = h(x(P))$$

where h denotes the height on $\mathbb{P}^1$. In particular, when $K = \mathbb{Q}$ and $P \in E(\mathbb{Q}) \setminus \{O\}$ we have $x(P) = [\frac{r}{s}, 1] = [r, s]$ for some coprime $r, s \in \mathbb{Z}$, so $h_x(P) = h(\frac{r}{s}) = \log \max\{|r|, |s|\}$.

It follows from (2.5) that h and h_x are even functions, i.e. $h(-P) = h(P)$ and $h_x(-P) = h_x(P)$.

The duplication formula (2.4) and Theorem 3.1.5 together imply that $h_x(2P) - 4h_x(P)$ is bounded from above. In fact, this is a consequence of a much more general fact called the **quasiparallelogram law**, the proof of which will occupy us for the entire next chapter. In noneffective form, it reads as follows (for the effective versions, see Theorems 4.2.4 and 4.3.3).

Theorem 3.2.1. *For every $P, Q \in E(\bar{K})$,*

$$h_x(P+Q) + h_x(P-Q) - 2h_x(P) - 2h_x(Q) = O_E(1),$$
$$h(P+Q) + h(P-Q) - 2h(P) - 2h(Q) = O_E(1).$$

The quasiparallelogram law states that for a given curve E, the left-hand expressions are uniformly bounded from above and below. Setting $P = Q$ we see that $h_x(2P) \approx 4h_x(P)$. Note that the duplication map does not extend to a morphism $\mathbb{P}^2 \to \mathbb{P}^1$, but the conclusion of Theorem 3.1.5 still holds.

4

The Quasiparallelogram Law

In this chapter, we prove the quasiparallelogram law for both h and h_x. Our main tool will be the function d introduced by Zimmer (with a different normalization). We first show that this *modified height* satisfies the quasiparallelogram law and having done so, we estimate the differences $h_x - d$ and $h - \frac{3}{2}d$ to obtain the quasiparallelogram laws. All bounds are completely explicit, and when practical we keep close track of the constants appearing due to our repeated use of the triangle inequality. At the end, a number of examples are investigated to give an idea of how sharp the bounds are.

When f is a real-valued function on E, we will use the shorthand notation

$$f(P,Q) = f(P+Q) + f(P-Q) - 2f(P) - 2f(Q).$$

All theorems from now on will be stated for arbitrary $P \in E(\bar{K})$. In order to not have to preface every argument with choosing a large enough field L/K so that $P \in E(L)$, we adopt the following convention: *all theorems are proved for points in* $E(K)$. Extending the proofs is just a matter of passing to an extension.

4.1 Modified height

In this section, we study the modified height function d. Here and in following sections, all estimates involving the modified height are slight modifications of those found in [9]. Before introducing d, we derive some inequalities from the Weierstraß equation which will be used tacitly in the sequel.

Most of our bounds will be expressed in terms of the following quantities associated to (the Weierstraß equation of) E.

$$\begin{aligned} \mu &= \frac{1}{6}h\big([A^3, B^2]\big) &&= \frac{1}{n}\sum_{v\in M_K} -n_v\mu_v, \\ \nu &= \frac{1}{6}h\big([A^3, B^2, 1]\big) &&= \frac{1}{n}\sum_{v\in M_K} -n_v\nu_v. \end{aligned}$$

In other words, $\mu_v = \min\{\frac{1}{2}v(A), \frac{1}{3}v(B)\}$ and $\nu_v = \min\{0, \frac{1}{2}v(A), \frac{1}{3}v(B)\} = \min\{0, \mu_v\}$. Thus, $\nu_v \le \mu_v$, $\nu_v \le 0$ and $\nu \ge \mu \ge 0$. Moreover, $2\mu_v \le v(A)$ and $3\mu_v \le v(B)$.

Remark 4.1.1. The quantity 6ν is known in the literature as the naive height of the curve E. Since a change of equation transforms A, B into $A' = \rho^4 A, B' = \rho^6 B$, we have $\mu' = \mu$, just as $j' = j$. In other words, μ is independent of the choice of Weierstraß equation. The same does not hold for ν.

Applying the triangle inequality directly to the Weierstraß equation gives

$$\begin{aligned} v(y) &= \frac{1}{2}v(x^3 + Ax + B) \\ &\geq \frac{3}{2}\min\{v(x), \frac{1}{3}\big(v(x) + v(A)\big), \frac{1}{3}v(B)\} + \epsilon_v \log 3 \\ &\geq \frac{3}{2}\min\{v(x), \nu_v\} + \epsilon_v \frac{1}{2}\log 3. \end{aligned} \tag{4.1}$$

Rewriting the equation as $x^3 = y^2 - Ax - B$ we obtain in a similar way

$$\frac{3}{2}v(x) \geq \min\{v(y), \frac{1}{2}v(x) + \nu_v\} + \epsilon_v \frac{1}{2}\log 3. \tag{4.2}$$

Equivalently,

$$\frac{3}{2}v(x) \geq v(y) + \epsilon_v \frac{1}{2}\log 3 \qquad \text{if } v(y) < \nu_v + \frac{1}{2}v(x), \tag{4.3}$$

$$\frac{3}{2}v(x) \geq \nu_v + \frac{1}{2}v(x) + \epsilon_v \frac{1}{2}\log 3 \text{ if } v(y) \geq \nu_v + \frac{1}{2}v(x). \tag{4.4}$$

Definition 4.1.2. Let $P \in E(\bar{K})$. The modified height of P is defined as

$$d(P) = \frac{1}{6}h\big([x^6, A^3, B^2]\big) = \frac{1}{n}\sum_{v \in M_K} -n_v \min\{v(x), \mu_v\}$$

if $P = [x, y, 1]$, and $d(O) = \mu$.

Lemma 4.1.3. *Let $P, Q \in E(\bar{K}) \setminus \{O\}$ and assume $P \neq \pm Q$. Then*

$$\begin{aligned} -6\mu - \log 367 - \frac{1}{2}\log 24 \leq d(P, Q) \leq \log 24, \\ -6\mu - \log 367 \leq d(2P) - 4d(P) \leq \log 12. \end{aligned}$$

Proof. The first step is to bound $d(2P) - 4d(P) =: \frac{1}{n}\sum n_v \delta_v$ from above, and we begin with the case $2P = O$. By the definition of d, we are interested in bounding the terms

$$\delta_v = -\mu_v + 4\min\{v(x_P), \mu_v\}.$$

Since $\delta_v \leq 3\mu_v$, we have

$$d(2P) - 4d(P) \leq -3\mu \leq 0,$$

which proves our upper bound. To prove the lower bound, first note that $y_P = 0$ must hold since P is of order 2. Using (4.4) and (4.3), we see that $v(x) \geq \mu_v + \epsilon_v \frac{1}{2}\log 3$. Summing over v we get $d(2P) - 4d(P) \geq -\mu - \frac{1}{2}\log 3$.

Assuming now that $2P \neq O$, we instead have

$$\delta_v = -\min\{v(x_{2P}), \mu_v\} + 4\min\{v(x_P), \mu_v\}.$$

Case 1: $v(x_{2P}) \geq \mu_v$. Here,

$$\delta_v = -\mu_v + 4\min\{v(x_P), \mu_v\} \leq 3\mu_v,$$

so $d(2P) - 4d(P) \leq -3\mu \leq 0$.
Case 2: $v(x_{2P}) < \mu_v \wedge v(x_P) \geq \mu_v$. In this case, $\delta_v = -v(x_{2P}) + 4\mu_v$. Applying v to the duplication formula (2.4) and using our assumptions, we get

$$\begin{aligned} v(x_{2P}) &\geq \min\{4v(x_P), 2v(x_P) + v(A), v(x_P) + v(B), 2v(A)\} \\ &\quad + \epsilon_v \log 12 - 2v(2y_P). \\ &\geq 4\mu_v + \epsilon_v \log 12 - 2v(2) - 2v(2y_P). \end{aligned}$$

It follows that

$$\delta_v \leq \epsilon_v \log 12 - 2v(2) - 2v(y_P),$$

and applying the sum formula we arrive at

$$d(2P) - 4d(P) \leq \log 12.$$

Case 3: $v(x_{2P}) < \mu_v \wedge v(x_P) < \mu_v$. In this case, $\delta_v = -v(x_{2P}) + 4v(x_P)$, and similarly as in Case 2 we get that

$$v(x_{2P}) \geq 4v(x_P) + \epsilon_v \log 12 - 2v(2) - 2v(y_P),$$

whence $d(2P) - 4d(P) \leq \log 12$.

The second step is to bound $d(2P) - 4d(P)$ from below. First, let

$$\phi_2(x_P) = x_P^4 - 2Ax_P^2 - 8Bx_P + A^2.$$

Then we may rewrite (2.4) as $\phi_2(x_P) = 4x_{2P}y_P^2$, so

$$v(\phi_2(x_P)) = v(x_{2P}) + 2v(y_P) + 2v(2). \tag{4.5}$$

We will need two relations satisfied by the discriminant. For any $P = (x, y) \neq O$,

$$\begin{aligned} -\frac{1}{16}\Delta x^7 &= \left((4A^3 + 27B^2)x^3 - A^2Bx^2 + (3A^4 + 22AB^2)x + (3A^3B + 24B^3)\right)\phi_2(x) \\ &\quad + \left(A^2Bx^3 + (5A^4 + 32AB^2)x^2 + (26A^3B + 192B^3)x - (3A^5 + 24A^2B^2)\right)y^2 \\ -\frac{1}{16}\Delta &= (3x^2 + 4A)\phi_2(x) - (3x^3 - 5Ax - 27B)y_P^2 \end{aligned}$$

Write $\Delta' = -\frac{1}{16}\Delta$ and $(x, y) = (x_P, y_P)$. Assume for the moment that $v(x) < \mu_v$. The first identity above is a sum of 367 terms, so by the triangle inequality

$$v(\Delta') + 7v(x) \geq \min\{v(\phi_2(x)) + M_1, 2v(y) + M_2\} + \epsilon_v \log 367,$$

$$\begin{aligned}
v(\phi_2(x)) + M_1 &= \\
&v(\phi_2(x)) + \min\{3v(x) + 3v(A), 3v(x) + 2v(B), 2v(x) + 2v(A) + v(B), \\
&\qquad v(x) + 4v(A), v(x) + v(A) + 2v(B), 3v(A) + v(B), 3v(B)\} \\
&\geq v(\phi_2(x)) + \min\{3v(x) + 6\mu_v, 2v(x) + 7\mu_v, v(x) + 8\mu_v, 9\mu_v\} \\
&= v(\phi_2(x)) + 3v(x) + 6\mu_v,
\end{aligned}$$

$$\begin{aligned}
2v(y) + M_2 &= \\
&2v(y) + \min\{3v(x) + 2v(A) + v(B), 2v(x) + 4v(A), 2v(x) + v(A) + 2v(B), \\
&\qquad v(x) + 3v(A) + v(B), v(x) + 3v(B), 5v(A), 2v(A) + 2v(B)\} \\
&\geq 2v(y) + \min\{3v(x) + 7\mu_v, 2v(x) + 8\mu_v, v(x) + 9\mu_v, 10\mu_v\} \\
&= 2v(y) + 3v(x) + 7\mu_v.
\end{aligned}$$

Applying (4.5) and rewriting, we arrive at

$$v(\Delta') + 7v(x) \geq 6\mu_v + 3v(x) + \min\{\mu_v, v(x_{2P})\} + 2v(y) + \epsilon_v \log 367.$$

To deal with the case $v(x) \geq \mu_v$, we use the second of the discriminant identities above in a similar way. We have

$$\begin{aligned}
v(\Delta') &\geq \min\{v(\phi_2(x)) + M_1', 2v(y) + M_2'\} + \epsilon_v \log 42 \\
v(\phi_2(x)) + M_1' &= v(\phi_2(x)) + \min\{2v(x), v(A)\} \\
&\geq v(\phi_2(x)) + 2\mu_v. \\
2v(y) + M_2' &= 2v(y) + \min\{3v(x), v(x) + v(A), v(B)\} \\
&\geq 2v(y) + 3\mu_v \\
\implies v(\Delta') &\geq \min\{v(\phi_2(x)) + 2\mu_v, 2v(y) + 3\mu_v\} + \epsilon_v \log 42.
\end{aligned}$$

Applying (4.5) again, we obtain

$$v(\Delta') \geq 2v(y) + 2\mu_v + \min\{\mu_v, v(x_{2P})\} + \epsilon_v \log 42.$$

It now follows in any case, i.e. without assumption on $v(x)$, that

$$v(\Delta') - 6\mu_v + 4\min\{v(x), \mu_v\} - 2v(y) - \epsilon_v \log 367 \geq \min\{v(x_{2P}), \mu_v\}.$$

Applying this to $\delta_v = -\min\{v(x_{2P}), \mu_v\} + 4\min\{v(x), \mu_v\}$ and using the the sum formula proves the lower bound. Indeed, it suffices to note that

$$\begin{aligned}
\delta_v &\geq -\big((v(\Delta') - 6\mu_v) + 4\min\{v(x), \mu_v\} - 2v(y) - \epsilon_v \log 367\big) + 4\min\{v(x_P), \mu_v\} \\
&= 6\mu_v - v(\Delta') + 2v(y) + \epsilon_v \log 367.
\end{aligned}$$

Turning our attention to $d(P, Q)$ with $P \neq \pm Q$ instead, we wish to estimate

$$\begin{aligned}
\delta_v = &-\min\{v(x_{P+Q}), \mu_v\} - \min\{v(x_{P-Q}), \mu_v\} \\
&+ 2\min\{v(x_P), \mu_v\} + 2\min\{v(x_Q), \mu_v\}.
\end{aligned}$$

Case 1: $v(x_{P+Q}) < \mu_v \wedge v(x_{P-Q}) < \mu_v$. To deal with this case, we apply v to both sides of (2.2) and use the triangle inequality to get

$$\begin{aligned}
&v(x_{P+Q}) + v(x_{P-Q}) \geq \\
&\min\{2v(x_P x_Q - A), v(x_P + x_Q) + v(B)\} + \epsilon_v \log 5 - 2v(x_P - x_Q).
\end{aligned}$$

We have

$$\begin{aligned} 2v(x_P x_Q - A) &\geq 2\min\{v(x_P) + v(x_Q), v(A)\} + \epsilon_v \log 4 \\ &\geq 2\big(\min\{v(x_P), \mu_v\} + \min\{v(x_Q), \mu_v\}\big) + \epsilon_v \log 4, \\ v(x_P + x_Q) + v(B) &\geq 3\mu_v + \min\{v(x_P), v(x_Q)\} + \epsilon_v \log 2 \\ &\geq 2\big(\min\{v(x_P), \mu_v\} + \min\{v(x_Q), \mu_v\}\big) + \epsilon_v \log 2. \end{aligned}$$

It follows that

$$\delta_v \leq 2v(x_P - x_Q) - \epsilon_v \log 20.$$

Case 2: $v(x_{P+Q}) < \mu_v \wedge v(x_{P-Q}) \geq \mu_v$. To handle this case, we apply the triangle inequality to (2.1) to see that

$$\begin{aligned} v(x_{P+Q}) + \mu_v \geq \min\{&v(x_P^2 x_Q), v(x_P x_Q^2), v(A(x_P + x_Q)), v(y_P y_Q), v(B)\} \\ &+ \epsilon_v \log 8 - 2v(x_P - x_Q) + \mu_v. \end{aligned}$$

We estimate each term inside the $\min\{\dots\}$ together with μ_v:

$$\begin{aligned} 2v(x_P) + v(x_Q) + \mu_v &\geq 2\big(\min\{v(x_P), \mu_v\} + \min\{v(x_Q), \mu_v\}\big) \\ v(x_P) + v(x_Q) + \mu_v &\geq 2\big(\min\{v(x_P), \mu_v\} + \min\{v(x_Q), \mu_v\}\big) \\ v(A) + v(x_P + x_Q) + \mu_v &\geq 3\mu_v + \min\{v(x_P), v(x_Q)\} + \epsilon_v \log 2 \\ &\geq 2\big(\min\{v(x_P), \mu_v\} + \min\{v(x_Q), \mu_v\}\big) + \epsilon_v \log 2 \\ v(y_P) + v(y_Q) + \mu_v &\geq \frac{3}{2}\big(\min\{v(x_P), \mu_v\} + \min\{v(x_Q), \mu_v\}\big) + \epsilon_v \log 3 + \mu_v \\ &\geq 2\big(\min\{v(x_P), \mu_v\} + \min\{v(x_Q), \mu_v\}\big) + \epsilon_v \log 3 \\ v(B) + \mu_v &\geq 4\mu_v \end{aligned}$$

From the above, it follows that

$$\delta_v \leq 2v(x_P - x_Q) - \epsilon_v \log 24.$$

Case 3: $v(x_{P+Q}) \geq \mu_v \wedge v(x_{P-Q}) < \mu_v$. By symmetry, this follows from Case 2.

Case 4: $v(x_{P+Q}) \geq \mu_v \wedge v(x_{P-Q}) \geq \mu_v$. It follows from our assumption that

$$v(x_P - x_Q) \geq \min\{v(x_P), \mu_v\} + \min\{v(x_Q), \mu_v\} - \mu_v + \epsilon_v \log 2$$

and hence multiplying by 2 and rearranging we get

$$\delta_v \leq 2v(x_P - x_Q) - \epsilon_v \log 4.$$

Taking the worst bound (i.e. Case 3) and summing we get, by the sum formula,

$$d(P, Q) \leq \frac{1}{n}\left(2\sum_{v \in M_K} n_v v(x_P - x_Q) - \sum_{v \in M_K} n_v \epsilon_v \log 24\right) = \log 24.$$

It remains to prove the second lower bound, i.e. when $P \neq \pm Q$. A clever trick due to Zagier allows us to bypass further calculations. Combining the bounds proved thus far, we see that

$$\begin{aligned} 2d(P+Q) + 2d(P-Q) &\geq d(2P) + d(2Q) - \log 24 \\ &\geq 4d(P) + 4d(Q) - 12\mu - 2\log 367 - \log 24. \end{aligned}$$

Rearranging and dividing by 2 completes the proof. □

4.2 Ordinary height

We now turn our attention to the ordinary height h_x defined in section 3.2. The proof of the upper bound is an effective version of that of [6, Theorem VII.6.2]. It is possible to make that proof of the lower bound effective too, using a Nullstellensatz argument. However, the obtained lower bound is very poor. Instead, we derive a vastly superior lower bound by means of a comparison between h_x and the modified height d.

Lemma 4.2.1. *Let $g : \mathbb{P}^2 \to \mathbb{P}^2$ be the map (in fact, morphism) defined by*

$$[x_0, x_1, x_2] \mapsto [x_1^2 - 4x_0x_2,\ 2x_1(Ax_0 + x_2),\ (Ax_0 - x_2)^2 - 4Bx_0x_1].$$

Then for every $P \in \mathbb{P}^2(\bar{K})$,

$$h(g(P)) \leq 2h(P) + h\big([1, A^2, B]\big) + 3\log 2.$$

Proof. Let $P = [x_0, x_1, x_2] \in \mathbb{P}^2(K)$, and write

$$v(P) = \min\{v(x_0), v(x_1), v(x_2)\},$$

so that $h(P) = \frac{1}{n}\sum -n_v v(P)$. With this notation, the sum we wish to bound from above is

$$\frac{1}{n} \sum_{v \in M_K} -n_v v(g(P)),$$

and we do so by applying the triangle inequality termwise and to each component of g.

$$\begin{aligned} v(x_1^2 - 4x_0x_2) &\geq 2v(P) + \epsilon_v \log 5 \\ v(2x_1(Ax_0 + x_2)) &\geq 2v(P) + \min\{0, v(A)\} + \epsilon_v \log 4 \\ v((Ax_0 - x_2)^2 - 4Bx_0x_1) &\geq 2v(P) + \min\{0, v(A), 2v(A), v(B)\} + \epsilon_v \log 8 \\ &= 2v(P) + \min\{0, 2v(A), v(B)\} + \epsilon_v \log 8 \end{aligned}$$

Combining the three, multiplying by $-\frac{1}{n}n_v$ and summing over v, we get

$$h(g(P)) \leq 2h(P) + h\big([1, A^2, B]\big) + 3\log 2$$

as claimed. □

Proposition 4.2.2. *Let $P, Q \in E(\bar{K})$. Then*

$$h_x(P,Q) \leq h\big([1, A^2, B]\big) + 7\log 2.$$

Proof. Since $h_x(O) = 0$ and $h_x(-P) = h_x(P)$, we have $h_x(P,Q) = 0$ whenever $P = O$ or $Q = O$. Moreover, if $2P = O$ then $h_x(P,P) \leq 0$. Assume now that $P, Q \neq O$. With some slight rearrangement, (2.2) and (2.3) become

$$x_{P+Q}x_{P-Q} = \frac{(x_P x_Q - A)^2 - 4B(x_P + x_Q)}{(x_P + x_Q)^2 - 4x_P x_Q},$$
$$x_{P+Q} + x_{P-Q} = \frac{2(x_P + x_Q)(A + x_P x_Q) + 4B}{(x_P + x_Q)^2 - 4x_P x_Q}.$$

Defining g as in Lemma 4.2.1, we obtain a commutative diagram

$$\begin{array}{ccc} E \times E & \xrightarrow{G} & E \times E \\ {\scriptstyle\sigma}\downarrow & & \downarrow{\scriptstyle\sigma} \\ \mathbb{P}^2 & \xrightarrow{g} & \mathbb{P}^2 \end{array}$$

where $G : (P,Q) \mapsto (P+Q, P-Q)$ and σ is the composition of the maps

$$\begin{aligned} E \times E &\to \mathbb{P}^1 \times \mathbb{P}^1 \\ (P,Q) &\mapsto (x(P), x(Q)) \end{aligned}$$
$$\begin{aligned} \mathbb{P}^1 \times \mathbb{P}^1 &\to \mathbb{P}^2 \\ \big([s,t],[u,v]\big) &\mapsto [tv, sv+tu, su] \end{aligned}$$

Lemma 4.2.1 and the commutativity of the diagram imply that

$$h\big(\sigma(P+Q, P-Q)\big) \leq 2h\big(\sigma(P,Q)\big) + h\big([1, A^2, B]\big) + 3\log 2. \tag{4.6}$$

To complete the proof, we investigate $h\big(\sigma(R,S)\big)$. If R or S is O, then clearly

$$h\big(\sigma(R,S)\big) = h_x(R) + h_x(S).$$

If not, writing

$$x(R) = [\alpha_1, 1],\ x(S) = [\alpha_2, 1]$$

we get

$$\begin{aligned} h\big(\sigma(R,S)\big) &= h\big([1, \alpha_1+\alpha_2, \alpha_1\alpha_2]\big), \\ h_x(R) + h_x(S) &= h(\alpha_1) + h(\alpha_2). \end{aligned}$$

Thus, applying Theorem 2 to the polynomial $(T+\alpha_1)(T+\alpha_2)$ gives

$$h(\alpha_1) + h(\alpha_2) - 2\log 2 \leq h\big([1, \alpha_1+\alpha_2, \alpha_1\alpha_2]\big) \leq h(\alpha_1) + h(\alpha_2) + \log 2. \tag{4.7}$$

Applying (4.7) to each side of (4.6) we get

$$h_x(P+Q) + h_x(P-Q) \leq 2h_x(P) + 2h_x(Q) + h\big([1, A^2, B]\big) + 7\log 2$$

as needed. □

To simplify the statement of the next lemma, we define the *narrow valuations* and the *denominator valuations* of the pair A, B over the field K as

$$N_K^\infty = \{v \in M_K^\infty \mid \mu_v \leq 0\},$$
$$D_K^0 = \{v \in M_K^0 \mid \mu_v \leq 0\}.$$

Note that $\mu_v \geq 0$ precisely when v is non-archimedean and *both* A and B are v-integral or v is archimedean and *both* A and B are mapped to the unit disk under the embedding $\tau_v : K \hookrightarrow \mathbb{C}$ associated to v. To these sets, we associate the quantities

$$\mu_N = \frac{1}{n} \sum_{v \in N_K^\infty} -n_v \mu_v, \quad \mu_D = \frac{1}{n} \sum_{v \in D_K^0} -n_v \mu_v.$$

When $K = \mathbb{Q}$ and $A, B \in \mathbb{Z}$, we have $N_\mathbb{Q}^\infty = \emptyset$ as long as $A, B \neq \pm 1$ and $\mu_v \leq 0$ for every $v \in M_\mathbb{Q}^0$. Therefore, in this case $\mu_N = 0$ and $\mu_D = \nu$. Moreover, these quantities are invariant under extensions of K, whereby it makes sense to speak of μ_D, μ_N without reference to K.

Lemma 4.2.3. *Let $P \in E(\bar{K})$. Then*

$$-\nu \leq h_x(P) - d(P) \leq -\mu + \mu_D + \mu_N$$

Proof. When $P = O$, we have $h_x(P) - d(P) = -\mu$ so the statement holds. Now assume $P = [x, y, 1] \in E(K) \setminus \{O\}$. We have

$$\begin{aligned} h_x(P) - d(P) &= \frac{1}{n} \sum_{v \in M_K} n_v \big(- \min\{v(x), 0\} + \min\{v(x), \mu_v\} \big) \\ &= \frac{1}{n} \sum_{v \in M_K} n_v \delta_v. \end{aligned}$$

To estimate δ_v, we make a division into three cases.
Case 1: $v(x) \geq 0$. In this case, $\delta_v = \min\{v(x), \mu_v\}$, whence

$$\nu_v \leq \delta_v \leq \mu_v.$$

Case 2: $v(x) < 0 \wedge v(x) \geq \mu_v$. Here, $\delta_v = -v(x) + \mu_v$, so

$$\nu_v \leq \mu_v \leq \delta_v \leq 0.$$

Case 3: $v(x) < 0 \wedge v(x) < \mu_v$. Here, $\delta_v = 0$.

In total, we have the following global estimate:

$$-\nu \leq \frac{1}{n} \sum n_v \delta_v \leq \frac{1}{n} \sum_{v \in M_K} -n_v \min\{0, -\mu_v\},$$

and the righthand side is

$$\begin{aligned} \frac{1}{n} \sum_{v \in M_K} -n_v \min\{0, -\mu_v\} &= -\mu - \sum_{v \in D_K^0} n_v \mu_v - \sum_{v \in N_K^\infty} n_v \mu_v. \\ &= -\mu + \mu_D + \mu_N \end{aligned}$$

as claimed. □

Theorem 4.2.4. *Let $P, Q \in E(\bar{K})$. Then*

$$-2\nu - 2\mu - 4\mu_D - 4\mu_N - \log 367 - \frac{1}{2}\log 24 \leq h_x(P,Q) \leq h\big([1, A^2, B]\big) + 7\log 2$$

Proof. When P or Q is O, the statement is trivially true, so assume $P, Q \neq O$. The upper bound is Proposition 4.2.2. The lower bound follows from Lemmas 4.2.3 and 4.1.3. Indeed, if $P \neq \pm Q$ one has

$$\begin{aligned} h_x(P,Q) &= h_x(P,Q) - d(P,Q) + d(P,Q) \\ &= h_x(P+Q) - d(P+Q) + h_x(P-Q) - d(P-Q) \\ &\quad - 2\big(h_x(P) - d(P)\big) - 2\big(h_x(Q) - d(Q)\big) + d(P,Q) \end{aligned}$$

and estimating each term, we obtain the theorem. For the case $P = \pm Q$, we do the same thing replacing $d(P,Q)$ with $d(2P) - 4d(P)$. □

The case $K = \mathbb{Q}$ and A, B integers deserves special mention.

Corollary 4.2.5. *Let*

$$E : y^2 = x^3 + Ax + B$$

be an elliptic curve defined over $\mathbb{Z}$. Then

$$-6\nu - 2\mu - 7.495 \leq h_x(P,Q) \leq \log\max\{|A|^2, |B|\} + 4.852.$$

Remark 4.2.6. When $P = Q$, one gets the slightly stronger lower bound

$$-5\nu - 2\mu - 5.906 \leq h_x(P,P).$$

In this case, one can also in a straightforward way improve a lot on the constant in the upper bound by working directly with the duplication formula (2.4).

4.3 Naive height

In order to simplify calculations, we follow Zimmer [9] in defining two auxilliary functions h_ν and d_ν. Set $h_\nu(O) = \frac{3}{2}d_\nu(O) = \frac{3}{2}\nu$, and for $P \neq O$ let

$$\begin{aligned} h_\nu(P) &= \frac{1}{4}h\big([x^4, y^4, A^3, B^2, 1]\big) \\ &= \frac{1}{n}\sum_{v\in M_K} -n_v \min\{v(x), v(y), \frac{3}{2}\nu_v\}, \\ d_\nu(P) &= \frac{1}{6}h\big([x^6, A^3, B^2, 1]\big) \\ &= \frac{1}{n}\sum_{v\in M_K} -n_v \min\{v(x), \nu_v\}. \end{aligned}$$

Lemma 4.3.1. *Let $P \in E(\bar{K})$. Then*

$$\begin{aligned} -\frac{3}{2}\nu \leq h(P) - h_\nu(P) \leq 0 \\ 0 \leq d_\nu(P) - d(P) \leq \frac{3}{2}(\nu - \mu). \end{aligned}$$

Proof. The statements follow directly from the definitions and trivial termwise estimations. □

Lemma 4.3.2. *Let $P \in E(\bar{K})$. Then*

$$-\frac{3}{4}\log 3 \leq h_\nu(P) - \frac{3}{2}d_\nu(P) \leq \frac{1}{2}\log 3$$

Proof. We want to estimate

$$\begin{aligned} h_\nu(P) - \frac{3}{2}d_\nu(P) &= \frac{1}{n}\sum_{v\in M_K} n_v\left(\frac{3}{2}\min\{\nu_v, v(x)\} - \min\{v(x), v(y), \frac{3}{2}\nu_v\}\right) \\ &= \frac{1}{n}\sum_{v\in M_K} n_v\delta_v. \end{aligned}$$

It suffices to show that $\epsilon_v\frac{3}{4}\log 3 \leq \delta_v \leq -\epsilon_v\frac{1}{2}\log 3$, and we make a division into two cases.
Case 1: $v(x) \geq \nu_v$. Since $\nu_v \leq 0$, we have $v(x) \geq \frac{3}{2}\nu_v$ and

$$\delta_v = \frac{3}{2}\nu_v - \min\{v(y), \frac{3}{2}\nu_v\}.$$

Applying (4.1), we obtain

$$0 \leq \delta_v \leq -\epsilon_v\frac{1}{2}\log 3.$$

Case 2: $v(x) < \nu_v$. In this case,

$$\delta_v = \frac{3}{2}v(x) - \min\{v(x), v(y), \frac{3}{2}\nu_v\}.$$

Now, if $\min\{v(x), v(y), \frac{3}{2}\nu_v\} = \frac{3}{2}\nu_v$, it must also hold that $v(y) \geq \frac{3}{2}\nu_v > \nu_v + \frac{1}{2}v(x)$ by our assumption on $v(x)$. Hence, (4.4) implies

$$\epsilon_v \frac{3}{4}\log 3 \leq \delta_v < 0.$$

Secondly, if instead $\min\{v(x), v(y), \frac{3}{2}\nu_v\} = v(x)$ we have $\delta_v = \frac{1}{2}v(x)$, and (4.3) and (4.4) simplify to

$$\begin{aligned}
&\frac{3}{2}v(x) \geq v(x) + \epsilon_v \frac{1}{2}\log 3 \\
&\frac{3}{2}v(x) \geq \nu_v + \frac{1}{2}v(x) + \epsilon_v \frac{1}{2}\log 3 \geq \frac{7}{6}v(x) + \epsilon_v \frac{1}{2}\log 3
\end{aligned}$$

respectively. In any case we have $\frac{1}{2}v(x) \geq \epsilon_v \frac{3}{4}\log 3$, and we obtain

$$\epsilon_v \frac{3}{4}\log 3 \leq \delta_v \leq \frac{1}{2}\nu_v \leq 0.$$

Thirdly, if $\min\{v(x), v(y), \frac{3}{2}\nu_v\} = v(y)$ then $\delta_v = -v(y) + \frac{3}{2}v(x) \leq -\epsilon_v \frac{1}{2}\log 3$ by (4.1), and for the lower bound we note that (4.3) and (4.4) imply

$$\begin{aligned}
&-v(y) + \frac{3}{2}v(x) \geq \epsilon_v \frac{1}{2}\log 3, \\
&-v(y) + \frac{3}{2}v(x) \geq -v(y) + \nu_v + \frac{1}{2}v(x) + \epsilon_v \frac{1}{2}\log 3 \geq \epsilon_v \frac{3}{4}\log 3,
\end{aligned}$$

In both cases, we observe that $\delta_v \geq \epsilon_v \frac{3}{4}\log 3$ and thereby the proof is complete. □

Theorem 4.3.3. *Let $P, Q \in E(\bar{K})$. Then*

$$-9\nu - 3\mu - 15.087 \leq h(P, Q) \leq 9\nu - 3\mu + 9.162.$$

Note that in terms of ν, μ, the lower bound is exactly $\frac{3}{2}$ times that of the ordinary height, but the upper bounds differ.

Proof. When P or Q is O, the theorem is trivially true, so assume $P, Q \neq O$. Adding up the three inequalities of Lemmas 4.3.1 and 4.3.2 we obtain

$$-\frac{3}{2}\nu - \frac{3}{4}\log 3 \leq h(P) - \frac{3}{2}d(P) \leq \frac{3}{2}(\nu - \mu) + \frac{1}{2}\log 3. \tag{4.8}$$

Just as in the proof of 4.2.4, when $P \neq \pm Q$ we add and subtract $\frac{3}{2}d(P, Q)$:

$$h(P, Q) = h(P, Q) - \frac{3}{2}d(P, Q) + \frac{3}{2}d(P, Q).$$

Applying (4.8) and Lemma 4.1.3 proves the theorem. The case $P = \pm Q$ is proved the same way. □

4.4 Examples

In this section, we work out a few examples in order to gauge the sharpness of the bounds proved in this chapter. The general strategy we employ is constructing one-parameter families $\{E_t \mid t \in \mathbb{N}\}$ with easy-to-find points $P_t \in E_t$, and then letting t tend to infinity in the expressions $h_x(P_t, P_t)$ and $h(P_t, P_t)$. The upper bounds are not too difficult to test, while finding pairs (P, Q) such that $h(P, Q)$ and $h_x(P, Q)$ are small is more challenging. This echos the points made in Remark 3.1.6, that bounding polynomials from below is subtle while upper bounds are straightforward.

Example 4.4.1. Let $t \in \mathbb{N}$ and define

$$E_t : y^2 = x^3 + (t^2 - 1)x.$$

E_t contains the point $P_t = (1, t)$, which satisfies $h_x(P_t) = 0$. Doubling the point gives

$$x(2P_t) = \frac{t^4 - 4t^2 + 4}{4t^2}$$

and hence $h_x(P_t, P_t) = h_x(2P_k) = \log(t^4 - 4t^2 + 4)$ when t is large enough and odd. Theorem 4.2.4 states that

$$h_x(P_t, P_t) \leq 2\log(t^2 - 1) + 7\log 2, \tag{4.9}$$

and by computing

$$\lim_{\substack{t \to \infty \\ t \text{ odd}}} \frac{2\log(t^2 - 1) + 7\log 2}{\log(t^4 - 4t^2 + 4)} = 1,$$

we see that the dependence on A in (4.9) is optimal.

Example 4.4.2. Let $t \in \mathbb{N}$ and define

$$E_t : y^2 - t^2 = x^3 - 1,$$

we have $P_t = (1, t) \in E_t$. In this case, for t large enough

$$x(2P_t) = \frac{-4t^2 + 1}{2t^2} \implies h_x(2P_t) = \log(4t^2 - 1).$$

Since $A = 0$, the upper bound in Theorem 4.2.4 is

$$h_x(P_t, P_t) \leq \log(t^2 - 1) + 7\log 2 \tag{4.10}$$

Comparing our two inequalities, we see that as $t \to \infty$,

$$\frac{\log(t^2 - 1) + 7\log 2}{2\log(4t^2 - 1)} \to 1.$$

We conclude that (4.10) has optimal dependence on B.

Example 4.4.3. Let $t \in \mathbb{N}$ and define

$$E_t : y^2 = x^3 - tx + 1.$$

Then $P_t = (0, 1) \in E_t$ satisfies $h(P_t) = 0$, and

$$2P_t = \left(\frac{t^2}{4}, \frac{t^3 - 8}{64}\right) \implies h(P_t) = \log(t^3 - 8)$$

for t large enough and odd. For this curve $\nu = \mu = \frac{1}{2}\log t$, so the upper bound in Theorem 4.3.3 is

$$h(P_t, P_t) \leq 3\log t + 9.162. \tag{4.11}$$

Finally, we have

$$\lim_{\substack{t\to\infty \\ t \text{ odd}}} \frac{3\log t + 9.162}{\log(t^3 - 8)} = 1,$$

which proves that the dependence on A in (4.11) is optimal.

Example 4.4.4. Let $t \in \mathbb{N}$ and define

$$E_t : y^2 = x^3 - t^3$$

The point $P_t = (t, 0) \in E(\mathbb{Q})$ satisfies $2P_t = O$ and $h_x(P_t) = \log t$. Hence,

$$h_x(P_t, P_t) = -4\log t, \tag{4.12}$$

The lower bound from Theorem 4.2.4 is

$$-6\log t - 7.495 \leq h_x(P_t, P_t).$$

Letting $t \to \infty$, we see that at worst, the dependence on B in this case is at worst $\frac{3}{2}$ times optimal.

Example 4.4.5. Let $t \in \mathbb{N}$ and define

$$E_t : y^2 = x^3 + 4t^4 x.$$

The point $(2t^2, 4t^3) \in E_t(\mathbb{Q})$ satisfies $h(P_t) = 3\log t + 2\log 2$ and $2P_t = (0, 0)$. Hence, $h(P_t, P_t) = -12\log t - 8\log 2$. Discarding the constant, the lower bound from Theorem 4.3.3 is $-9\nu = -18\log t$. We conclude that the lower bound is in this case at worst $\frac{3}{2}$ times optimal. Similarly, the lower bound for the ordinary height is at worst $\frac{3}{2}$ times optimal in this family.

5

The Canonical Height

In this chapter, we construct the canonical (Néron-Tate) height $\hat{h}$ on E using Tate's averaging procedure. After some fundamental properties of $\hat{h}$ are proved, we apply the results of Chapter 4 to bound the differences $h_x - \hat{h}$ and $h - \frac{3}{2}\hat{h}$. The obtained estimates are then compared to well-known bounds from the literature.

There is an extensive amount of literature on the topics of this chapter, and it is difficult to say anything new. Nevertheless, the results in section 5.2 are easy consequences of the quasiparallelogram law, and a thesis on heights without a treatment of the canonical height would surely be lacking.

5.1 Construction of $\hat{h}$

A real-valued function f defined on an abelian group G is called a **quadratic form** if it is even: $f(-g) = f(g)$, and the map

$$\begin{aligned} \langle \cdot, \cdot \rangle_f : G \times G &\to \mathbb{R} \\ (g, h) &\mapsto f(g+h) - f(g) - f(h) \end{aligned}$$

is bilinear. A quadratic form is said to be **positive semidefinite** if $f(g) \geq 0$ for all g, and **positive definite** if equality holds only when $g = 0$. It is a standard fact that $\langle \cdot, \cdot \rangle_f$ is bilinear if and only if f satisfies the parallelogram law. Indeed, with notation as before, $f(g, h) = 0$ for all $g, h \in G$ implies

$$f(a+c, b) + f(a+b, c) - f(a, b-c) - f(b, c) = 0.$$

Expanding and rearranging yields

$$\langle a+b, c \rangle_f = \langle a, c \rangle_f + \langle b, c \rangle_f.$$

This proves one direction, and the other is immediate from the definitions.

The quasiparallelogram laws tell us that the naive and ordinary heights are quadratic forms on $E(\bar{K})$ 'up to $O(1)$'. It was discovered by Tate that by means of an averaging procedure, it is possible to define a new height function which is a quadratic form. To define the canonical height, we use Tate's averaging procedure on Zimmer's modified height. We will see in Theorem 5.1.3(b) that all definitions of $\hat{h}$ are equivalent up to a normalizing factor.

Definition-Theorem 5.1.1. *Let $P \in E(\bar{K})$. The limit*

$$\hat{h}(P) := \lim_{k\to\infty} \frac{d(2^k P)}{4^k}$$

exists, and $\hat{h}$ is called the canonical height on E.

Proof. The convergence follows from the quasiparallelogram law d. Indeed,

$$\begin{aligned} & \frac{d(2^k P)}{4^k} = \frac{d(2^{k-1}P)}{4^{k-1}} + \frac{d(2^k P) - 4d(2^{k-1}P)}{4^k} \\ \implies & \frac{d(2^k P)}{4^k} = d(P) + \sum_{\ell=1}^{k} \frac{d(2^\ell P) - 4d(2^{\ell-1}P)}{4^\ell} \\ \implies & \hat{h}(P) = d(P) + \sum_{\ell=1}^{\infty} \frac{d(2^\ell P) - 4d(2^{\ell-1}P)}{4^\ell}. \end{aligned} \tag{5.1}$$

By Theorem 4.1.3, the numerators of the summands can be uniformly bounded (see Theorem 5.2.1 below for details), whereby the series converges. □

Remark 5.1.2. Since the difference $h_x - d$ is bounded, we could just as well have defined $\hat{h}$ in terms of h_x instead. Working with d is easier, and will lead to a stronger bound on $\hat{h} - h_x$ in the next section.

The canonical height possesses a number of interesting properties, and the following theorem should serve to motivate the name.

Theorem 5.1.3. *Let E/K be an elliptic curve with canonical height $\hat{h}$.*

(a) *The canonical height $\hat{h}$ is a positive semidefinite quadratic form. Equivalently, $\hat{h}$ is even, non-negative and satisfies the parallelogram law:*

$$\hat{h}(P, Q) = 0 \text{ for all } P, Q \in E(\bar{K}).$$

(b) *For any $m \in \mathbb{Z}$ and $P \in E(\bar{K})$,*

$$\hat{h}(mP) = m^2 \hat{h}(P).$$

(c) *For all $P \in E(\bar{K})$,*

$$\hat{h}(P) - h_x(P) = O_E(1).$$

Moreover, this together with the parallelogram law determines $\hat{h}$ completely.

(d) *$\hat{h}(P) = 0$ if and only if P is torsion.*

(e) *The canonical height is independent of choice of short Weierstrass equation. That is, let $E \to E'$ be a birational change of coordinates, and let $\hat{h}, \hat{h}'$ denote the respective canonical heights, defined in terms of d, d'. Assume $P \mapsto P'$. Then*

$$\hat{h}'(P') = \hat{h}(P).$$

Proof. (a) Since d is even and non-negative, $\hat{h}$ is even and positive semidefinite. Moreover, Theorem 4.1.3 states that

$$d(2^k P, 2^k Q) = O_E(1).$$

Dividing both sides by 4^k and letting $k \to \infty$, we see that $\hat{h}(P, Q) = 0$. In other words, $\hat{h}$ is a quadratic form.
(b) Let $m \in \mathbb{Z}$. Then

$$\hat{h}(mP) = \lim_{k\to\infty} \frac{d(2^k mP)}{4^k} = \lim_{k\to\infty} m^2 \frac{d(2^{\log_2 m+k} P)}{4^{\log_2 m+k}} = m^2 \hat{h}(P).$$

(c) The identity (5.1) together with the quasiparallelogram law for d (Theorem 4.1.3) allows us to conclude that $\hat{h}(P) - d(P) = O_E(1)$. Lemma 4.2.3 tells us that $h_x(P) - d(P) = O_E(1)$, and the first claim follows. To prove the second part, let f be a function as in (a) such that the difference $f - h_x$ is bounded. Then

$$f(2^k P) = 4^k f(P) \text{ for } k \geq 1.$$

Since the difference $\hat{h} - h_x$ is bounded, so is $f - \hat{h}$. If $P \in E(\bar{K})$, we see that

$$\begin{aligned} f(P) &= 4^{-k} f(2^k P) \\ &= 4^{-k}\big(\hat{h}(2^k P) + O_E(1)\big) \\ &= \hat{h}(P) + O_E(4^{-k}). \end{aligned}$$

Since this holds for any $k \geq 1$, we have $f(P) = \hat{h}(P)$.
(d) Let P satisfy $mP = O$ for some $m \neq 0$. It follows that $d(2^k P)$ takes only a finite number of values, and hence the limit $\hat{h}(P)$ must be 0. Conversely, let $Q \in E(L)$ satisfy $\hat{h}(Q) = 0$, where L/K is an extension. It follows from (b) that for any $m \in \mathbb{Z}$,

$$\hat{h}(mQ) = m^2 \hat{h}(Q) = 0.$$

It then follows from (c) that $h_x(mQ) \leq C$ for some C. Moreover, $mQ \in E(L)$ so we have an inclusion

$$\{mQ \in E(L) \mid m \in \mathbb{Z}\} \subseteq \{P \in \mathbb{P}^2(L) \mid h_x(P) \leq C\}.$$

By Theorem 3.1.2(e), the latter set is finite. Hence, Q is torsion.
(e) By Proposition 2.2.1, we know that $A' = \rho^4 A$, $B' = \rho^6 B$ and $x_{P'} = \rho^2 x_P$ for some $\rho \neq 0$. Hence,

$$\begin{aligned} d'(P') &= \frac{1}{6} h\big([\rho^{12} x_P^6, \rho^{12} A^3, \rho^{12} B^2]\big) \\ &= d(P), \end{aligned}$$

which implies $\hat{h}'(P') = \hat{h}(P)$. □

5.2 Difference bounds

In this section we derive estimates of the differences $h_x - \hat{h}$ and $h - \frac{3}{2}\hat{h}$. Such estimates are necessary in order to calculate the generators of the Mordell-Weil group $E(K)$. Theoretically, one could use the naive height h instead, but since it is generally about $\frac{3}{2}h_x$, doing so requires a larger search range. Such computational considerations are the main reason why h_x is the most commonly used height.

It is crucial to note that the problem of determining difference bounds for a single given curve has been completely solved by work of Cremona, Prickett, Siksek [1] and Uchida [8]. For families of curves, there is Silverman's bound [7]. Both are implemented in Magma under the names `CPSHeightBounds` and `SilvermanHeightBounds`, and they apply to curves with long Weierstraß equations with K-integral coefficients. In addition, the former requires the additional assumption that the equation is *minimal*, which is a kind of reducedness property (see [6] for details). Our results are applicable only to curves with short Weierstrass equations, but without integrality or minimality assumptions.

Theorem 5.2.1. *Let $P \in E(\bar{K})$. Then*

$$-\nu - \frac{1}{3}\log 12 \leq h_x(P) - \hat{h}(P) \leq \mu + \mu_D + \mu_W + \frac{1}{3}\log 367$$

Proof. This is just a matter of making explicit what we have already stated above. Combining the identity (5.1), the bounds of Theorem 4.1.3, Lemma 4.2.3 and the fact that $\sum_{\ell=1}^{\infty} 4^{-\ell} = \frac{1}{3}$ proves the theorem. □

Theorem 5.2.2. *Let $P \in E(\bar{K})$. Then*

$$-\frac{3}{2}\nu - \frac{1}{2}\big(\frac{3}{2}\log 2 + \log 12\big) \leq h(P) - \frac{3}{2}\hat{h}(P) \leq \frac{3}{2}\nu + \frac{3}{2}\mu + \frac{1}{2}\big(\log 3 + \log 367\big)$$

Proof. Same as 5.2.1. □

We record the special cases $K = \mathbb{Q}, A, B \in \mathbb{Z}$.

Corollary 5.2.3. *Let E be an elliptic curve defined over $\mathbb{Z}$ and $P \in E(\bar{K})$. Then*

$$\begin{aligned} -\nu - 0.829 \leq h_x(P) - \hat{h}(P) \ &\leq \nu + \mu + 1.969 \\ -\frac{3}{2}\nu - 1.763 \leq h(P) - \frac{3}{2}\hat{h}(P) &\leq \frac{3}{2}\nu + \frac{3}{2}\mu + 3.502. \end{aligned}$$

Note that the dependencies on ν, μ are the same up to the normalizing constant $\frac{3}{2}$.

To conclude the chapter, we make some comparisons between the available difference bounds. For a curve with short Weierstrass equation with integral coefficients, Silverman [7], obtained the following bound.

$$-\frac{1}{6}h(\Delta) - \frac{1}{12}h_\infty(j) - 2.14 \leq h_x(P) - \hat{h}(P) \leq \frac{1}{6}h(\Delta) + \frac{1}{12}h(j) + \frac{1}{6}h_\infty(j) + 1.946.$$

Here, h_∞ denotes the archimedean contribution to the height. Furthermore, there is the Cremona-Prickett-Siksek-Uchida bound mentioned above. For a given curve with minimal long Weierstrass equation, it gives the best possible bound.

Example 5.2.4. Let $t \in \mathbb{N}$ and define

$$E_t : y^2 = x^3 - t^2 x.$$

For this curve, $j = -1728$, $\Delta = 64t^6$, $\mu = 0$ and $\nu = \frac{1}{6}\log t$. Hence, the Silverman bound and Corollary 5.2.3 are

$$\begin{aligned} -\log t - 3.455 \leq h_x(P) - \hat{h}(P) \leq \log t + 4.503, \\ -\log t - 0.829 \leq h_x(P) - \hat{h}(P) \leq \log t + 1.969. \end{aligned}$$

The point $P_t = (t, 0) \in E_t(\mathbb{Q})$ is of order 2 and has ordinary height $h_x(P_t) = \log t$, whereby we see that the dependence on t in the upper bounds are optimal (since $\hat{h}(P_t) = 0$). Restricting t to primes ensures that E_t is minimal, allowing us to compare with `CPSHeightBounds`. For example, setting $t = 2, 101, 1009$ yields the following upper bounds.

t	2	101	1009
CPSU	1.034	4.962	7.264
5.2.3	2.663	6.585	8.886
Silverman	5.197	9.119	11.420

Example 5.2.5. Let $t \in \mathbb{N}$ and define

$$E_t : y^2 = x^3 + (t^2 - 1)x.$$

For this curve, $\mu = 0$ and $\nu = \frac{1}{2}\log(t^2 - 1)$, so Corollary 5.2.3 for the ordinary height reads

$$-\frac{1}{2}\log(t^2 - 1) - 0.829 \leq h_x(P) - \hat{h}(P) \leq \frac{1}{2}\log(t^2 - 1) + 1.969.$$

We claim that for this family, the dependency on t in the lower bound is optimal. Let $P_t = (1, t) \in E_t$. One can use the techniques in [5] and [4] to show that

$$\lim_{t\to\infty} \frac{\hat{h}(P_t)}{\log t} = 1,$$

from which it follows that

$$\lim_{t\to\infty} \frac{h_x(P_t) - \hat{h}(P_t)}{-\frac{1}{2}\log(t^2 - 1) - 0.829} = 1,$$

which proves the claim.

Example 5.2.6. Let $t \in \mathbb{N}$ and define

$$E_t : y^2 = x^3 + t^2.$$

For this curve, $\mu = 0$ and $\nu = \frac{2}{3}\log t$, so for the naive height Corollary 5.2.3 reads

$$-\log t - 1.763 \leq h(P) - \frac{3}{2}\hat{h}(P) \leq \log t + 3.502.$$

The point $P_t = (0, t) \in E_t(\mathbb{Q})$ is of order 3 and has naive height $h(P_t) = \log t$. Since $\hat{h}(P_t) = 0$, we see that the dependence on t in the upper bound is optimal.

Remark 5.2.7. Since $\hat{h}$ satisfies the parallelogram law, Silverman's bound implies the quasiparallelogram law for h_x. We have $h_x(P, Q) = h_x(P, Q) - \hat{h}(P, Q)$ and therefore

$$-h(\Delta) - \frac{1}{3}h(j) - \frac{5}{6}h_\infty(j) - 12.064 \leq h_x(P, Q) \leq h(\Delta) + \frac{1}{6}h(j) + \frac{2}{3}h_\infty(j) + 12.452.$$

As Silverman's result relies on a different construction of $\hat{h}$ than ours, it is interesting to compare this with Theorem 4.2.4. The constant terms differ a great deal of course, but in terms of dependencies on A, B, the lower bounds are quite similar in shape, while our upper bound is superior.

Remark 5.2.8. Any elliptic curve over a number field is isomorphic to a curve defined by a short Weierstrass equation. Indeed, given a curve in long Weierstrass form

$$E_{\text{long}} : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6,$$

the 'shortening map' f defined by

$$f\big([X, Y, Z]\big) = [36X - 18b_2Z,\ a_1X + Y - a_3Z,\ 216Z]$$

is an isomorphism between E_{long} and the curve

$$E_{\text{short}} : y^2 = x^3 - 27c_4x - 54c_6$$

where c_4, c_6 are certain polynomials in the coefficients $a_1, \ldots, a_6$. One can define a canonical height satisfying Theorem 5.1.3 on curves in long Weierstrass form. Letting $\hat{h}^l, h^l_x, \hat{h}^s$ and h^s_x denote the various heights, we have

$$\hat{h}^l \approx h^l_x \approx h^s_x \circ f \approx \hat{h}^s \circ f,$$

where $\approx$ means equality up to $O(1)$. Consequently, $\hat{h}^l = \hat{h}^s \circ f$ is an equality, and for any $P \in E_{\text{long}}(\bar{K})$,

$$\hat{h}^l(P) - h^l_x(P) \leq \hat{h}^s(P) - h^s_x(P) + h^s_x(f(P)) - h^l_x(P).$$

It is straightforward to find an upper bound for the difference $h^s_x(f(P)) - h^l_x(P)$, and this together with Theorem 5.2.1 yields an upper bound for the difference $\hat{h}^l - h^l_x$. It is poor in general since it is a combination of two estimates, and a better bound is obtained either by the method of Silverman or carrying out analogous arguments to those of Chapter 3 with the addition formulas and modified height for long equations (see [3]). The latter method appears to lead to a weaker result for curves in short form, which serves to motivate the restriction to short equations in this thesis.

Bibliography

[1] J.E. Cremona, M. Prickett, and S. Siksek. "Height difference bounds for elliptic curves over number fields". In: *Journal of Number Theory* 116.1 (2006), pp. 42–68. DOI: 10.1016/j.jnt.2005.03.001.

[2] S. Lang. *Algebraic Number Theory*. Springer, 2014.

[3] S. Schmitt and H.G. Zimmer. *Elliptic Curves: A Computational Approach*. Vol. 31. De Gruyter studies in mathematics. De Gruyter.

[4] J.H. Silverman. "Computing heights on elliptic curves". In: *Mathematics of Computation* 51.183 (1988), pp. 339–339. DOI: 10.1090/s0025-5718-1988-0942161-4.

[5] J.H. Silverman. "Heights and the specialization map for families of abelian varieties." In: *Journal für die reine und angewandte Mathematik* 342 (1983), pp. 197–211. DOI: 10.1515/crll.1983.342.197.

[6] J.H. Silverman. *The Arithmetic of Elliptic Curves*. Springer, 2009.

[7] J.H. Silverman. "The difference between the Weil height and the canonical height on elliptic curves". In: *Mathematics of Computation* 55.192 (1990), pp. 723–723. DOI: 10.1090/s0025-5718-1990-1035944-5.

[8] Y. Uchida. "The difference between the ordinary height and the canonical height on elliptic curves". In: *Journal of Number Theory* 128.2 (2008), pp. 263–279. DOI: 10.1016/j.jnt.2007.10.002.

[9] H.G. Zimmer. "On the difference of the Weil height and the Néron-Tate height". In: *Mathematische Zeitschrift* 147.1 (1976), pp. 35–51. DOI: 10.1007/bf01214273.