# Part III Essay:
# Computing Canonical Heights on Elliptic Curves

Filip Trenkić

2024

# Contents

# 1 Introduction

The theory of heights is an extremely important tool in the study of the arithmetic of elliptic curves. For $E/K$ an elliptic curve over a number field, the *canonical height* is a quadratic form $\hat{h} : E(K) \to \mathbb{R}$ with several desirable properties and far-reaching theoretical applications, playing a role in the celebrated Mordell-Weil Theorem, and constituting a term (the regulator) in the Birch-Swinnerton-Dyer conjecture, one of the most important open problems in mathematics.

The utility of heights extends to the computational theory of elliptic curves, where we aim to algorithmically compute the arithmetic invariants of a given elliptic curve $E$. This can be used to aid theory by providing evidence for conjectures (for example, searching for elliptic curves with large rank), with databases such as LMFDB [1] containing millions of elliptic curves over $\mathbb{Q}$ and small number fields.

Famously, there is no algorithm guaranteed to compute the rank of the Mordell-Weil group $E(K)$. However, practical methods to compute a Mordell-Weil basis exist, and depend heavily on height computations. For example, to check a set of points $P_1, \ldots, P_r \in E(\mathbb{Q})$ is linearly independent, we can check the height regulator matrix has nonzero determinant. The entries are defined by the height pairing

$$\langle P_i, P_j \rangle_{\hat{h}} = \hat{h}(P_i + P_j) - \hat{h}(P_i) - \hat{h}(P_j)$$

and this justifies the need for an efficient computation of $\hat{h}$.

Furthermore, to determine a Mordell-Weil basis from a set of generators for $E(K)/mE(K)$, one method involves to enumerating all points with height bounded by some constant (see [2]):

$$\{P \in E(\mathbb{Q}) \,:\, \hat{h}(P) \leq B\}.$$

Since the canonical height $\hat{h}$ is expensive to compute compared to the naive height $h$, we instead find an upper bound for their difference, say $h - \hat{h} \leq C$, and enumerate the set:

$$\{P \in E(\mathbb{Q}) \,:\, h(P) \leq B + C\}.$$

Because the height is a *logarithmic* measure of the arithmetic complexity of a point, the search space grows exponentially with $C$, so a good upper bound is essential for efficiency.

These examples illustrate the two key problems - we would like practical algorithms to compute the canonical height $\hat{h}(P)$ of a point $P \in E(K)$, and to compute good bounds for the height difference $h(P) - \hat{h}(P)$ over all points $P \in E(K)$. In both cases, the most fruitful approach has been to use the theory of *local heights* to decompose the problem into local terms, and then employ a variety of techniques to understand these individually. The purpose of the essay is to provide an expositional insight into this research, discussing the necessary theory, and explaining some of the algorithms used in practice.

First, in Chapter 2, we recall the theory of heights and establish our normalisations. In Chapter 3, we discuss the decomposition of $\hat{h}$ and $h - \hat{h}$ into local terms, studying the nonarchimedean and archimedean cases in detail. In Chapter 4 we will use this theory to explain and assess algorithms for height difference bounds; in Chapter 5 we will turn to algorithms for computing $\hat{h}$. The prequisite knowledge assumed by the essay should be familiar to a Masters-level student specialising in Number Theory, and we will provide comprehensive references throughout.

# 2 Global Heights

In this section, we briefly summarise the global theory of heights for elliptic curves defined over a number field $K$. In particular, this serves to establish the normalisations, which we shall use throughout the essay, for the absolute values on $K$, naive height $h$, and canonical height $\hat{h}$. For a more detailed treatment of this theory, see Silverman's book [3, Chapter VIII], or alternatively Lang's book [4, Chapter IV].

Throughout this essay, $E$ will denote an elliptic curve over a number field $K$, given by a Weierstrass equation:

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

with the 'usual quantities' $b_2$, $b_4$, $b_6$, $b_8$, $c_4$, $c_6$ and discriminant $\Delta$ defined as given in [3, Chapter III] without further comment.

## 2.1 Absolute Values

To make all our formulae explicit, we will fix a normalisation of absolute values as follows:

- We normalise the absolute values on $\mathbb{Q}$ such that
$$|x|_\infty = \max\{x, -x\}$$
$$\left|\frac{a}{b}p^n\right|_p = p^{-n}$$
where in the latter formula $a, b$ are integers coprime to $p$.
- For a number field $K/\mathbb{Q}$, we will write $M_K$ for the set of places of $K$. In particular, $M_K^0$ will denote the finite places, and $M_K^\infty$ the infinite places.
- For a place $v \in M_K$, we will write $|\cdot|_v$ to denote the corresponding absolute value, normalised so that its restriction is equal to an absolute value on $\mathbb{Q}$.
- The completion of $K$ at $v \in M_K$ will be denoted $K_v$. The *local degree at* $v$, denoted $n_v$, is defined by
$$n_v = [K_v : \mathbb{Q}_v]$$

With these definitions, the product formula then takes the form:

$$\prod_{v \in M_K} |x|_v^{n_v} = 1 \quad \forall x \in K^\times$$

## 2.2 The Naive Height

For an elliptic curve $E/\mathbb{Q}$, we recall the naive height $h : E(\mathbb{Q}) \to \mathbb{R}$ is defined by $h(P) = \log H(x(P))$, where $H$ is the height function on (rational) projective space,

$$H : \mathbb{P}^N(\mathbb{Q}) \to \mathbb{R}$$
$$H(x_0 : \ldots : x_N) = \max\{|x_0|, \ldots, |x_N|\}.$$

where we always choose $x_i \in \mathbb{Z}$ and $\gcd(x_0, \ldots, x_N) = 1$.

**Definition** (Height on Projective Space)**.** To extend the theory to general number fields $K/\mathbb{Q}$, we first extend the definition of $H$ as follows:

$$H : \mathbb{P}^N(K) \to \mathbb{R}$$
$$H(P) = \left(\prod_{v \in M_K} \max\{|x_0|_v, \ldots, |x_N|_v\}^{n_v}\right)^{1/[K:\mathbb{Q}]}$$

By the product forumla, this is independent of our choice of coordinates $x_i$ for $P$. In fact, we can also check that $H(P)$ is independent of the choice of field $K$ containing the coordinates of $P$, and thus gives a well-defined height function on all of $\mathbb{P}^N(\bar{\mathbb{Q}})$.

**Definition** (Naive Height)**.** For an elliptic curve $E/K$, the *naive height* of a point $P \in E(K)$ is defined by $h(P) = \log H(x(P))$. Explicitly, for $P \neq O_E$,

$$h(P) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M_K} n_v \log \max\{|x(P)|_v, 1\}$$

Just as with heights over $\mathbb{Q}$, we can prove the naive height exhibits several desirable properties which make it an appropriate theoretical tool for infinite descent.

**Proposition 2.1.** Let $E/K$ be an elliptic curve.

(i) If $\phi : E \to E'$ is an isogeny defined over $K$, then there exists a constant $c > 0$ such that $\forall P \in E(K)$,

$$|h(\phi(P)) - (\deg \phi)h(P)| \leq c$$

(ii) There exists a constant $c > 0$ such that $\forall P, Q \in E(K)$,

$$h(P+Q) + h(P-Q) \leq 2h(P) + 2h(Q) + c$$

(iii) For any $B > 0$, the set

$$\{P \in E(K) : h(P) \leq B\}$$

is finite.

**Remark.** The proofs of (i) and (ii) are similar to the proofs over the rationals. However, property (iii) is trivial over $\mathbb{Q}$, but much more delicate in the case of a general number field. The idea is to for each $x \in \bar{\mathbb{Q}}$ relate $H([x,1])$ to the height of the coefficients in the minimal polynomial for $x$, essentially reducing the case $K = \mathbb{Q}$. For details, consult [3, Theorem VIII.5.11].

In particular, these three properties are sufficient to deduce the Mordell-Weil Theorem from its weak form, as shown in [3, Theorem VIII.6.7] and [4, Theorem IV.2.1].

## 2.3 The Canonical Height

Although the naive height $h$ is sufficiently well-behaved to prove the Mordell-Weil theorem, it falls shy of being a quadratic form, and is dependent on the choice of Weierstrass equation for $E$. The *canonical height* (or *Néron-Tate height*) $\hat{h}$ has neither of these issues, and is a far more natural quantity with greater theoretical utility.

**Definition** (Canonical Height)**.** For an elliptic curve $E/K$, the *canonical height* of a point $P \in E(K)$ is defined as the limit of a Cauchy sequence:

$$\hat{h}(P) = \lim_{N \to \infty} \frac{1}{4^N} h(2^n P)$$

This 'averaging' construction means that $\hat{h}$ inherits the properties in Proposition 2.1 while ensuring the constants in (i),(ii) vanish in the limit; we state the corresponding properties for $\hat{h}$ below.

**Proposition 2.2.** Let $E/K$ be an elliptic curve.

(i) If $\phi : E \to E'$ is an isogeny defined over $K$, then $\forall P \in E(K)$,

$$\hat{h}(\phi(P)) = (\deg \phi)\hat{h}(P)$$

(ii) For all $P, Q \in E(K)$, $\hat{h}$ satisfies the parallelogram law,

$$\hat{h}(P+Q) + \hat{h}(P-Q) = 2\hat{h}(P) + 2\hat{h}(Q)$$

(iii) The difference $h - \hat{h}$ is bounded on $E(K)$. In particular, for any $B > 0$, the set

$$\{P \in E(K) : \hat{h}(P) \leq B\}$$

is finite.

It follows that $\hat{h}$ is a quadratic form, and independent of Weierstrass equation.

**Remark.** Some authors define $\hat{h}$ to be half of the value we define here. However, the normalisation we have chosen here is the natural choice for the Birch-Swinnerton-Dyer conjectures. For a comparison of various normalisations, see Appendix A.1.

# 3 Local Heights

For both computing the canonical height $\hat{h}$ and bounding the difference $h - \hat{h}$, our general strategy will be to decompose heights into sums of local terms. In this chapter, we will summarise the theory of local heights, and provide explicit formulae for both archimedean and non-archimedean places.

## 3.1 Decomposition of the Canonical Height

The naive height $h : E(K) \to \mathbb{R}$ is defined as a sum of local terms, and thus it is natural to ask whether there is a corresponding local decomposition of the canonical height $\hat{h}$.

In fact, in Néron's original construction of the canonical height, he first proves the existence of a *local height function* for each place, and proceeds to define the canonical height as a sum of local heights. Néron works with arbitrary abelian varieties; in the specific case of elliptic curves, the following formulation is due to Tate:

**Theorem 3.1.** (Néron, Tate) Let $(K_v, |\cdot|_v)$ be a complete valued field, and $E/K_v$ an elliptic curve. Fix a Weierstrass equation for $E$, with coefficients $a_1, \ldots, a_6$ and discriminant $\Delta$. Then:

(a) There exists a unique function, the *Néron local height*

$$\hat{\lambda}_v : E(K_v) \setminus \{O_E\} \to \mathbb{R}$$

satisfying the following three properties:

(i) $\hat{\lambda}_v$ is continuous on $E(K_v) \setminus \{O_E\}$, and bounded on the complement of any $v$-adic neighbourhood[a] of $O_E$.

(ii) The $v$-adic limit

$$\lim_{P \to O_E} \{\hat{\lambda}_v(P) - \log|x(P)|_v\}$$

exists.

(iii) For all $P = (x, y) \in E(K_v)$ with $[2]P \neq O_E$,

$$\hat{\lambda}_v(2P) = 4\hat{\lambda}_v(P) - 2\log|2y + a_1x + a_3|_v + \frac{1}{2}\log|\Delta|_v$$

(b) $\hat{\lambda}_v$ is independent of Weierstrass equation for $E$.

(c) For any finite extension $L/K$ with $|\cdot|_w$ extending $|\cdot|_v$, then the restriction of $\hat{\lambda}_w$ to $K$ is just $\hat{\lambda}_v$.

[a] The $v$-adic topology on $E$ is defined by the basis of open neighbourhoods

$$U((x_0, y_0), \varepsilon) = \{(x, y) \in E(K_v) \ ; \ |x - x_0|_v, |y - y_0|_v < \varepsilon\}$$
$$U(O_E, \varepsilon) = \{(x, y) \in E(K_v) \ ; \ |x|_v > \varepsilon^{-1}\} \cup \{O_E\}$$

*Proof.* See [5], Chapter VI, Theorem 1.1. Part (b) follows from the invariance of the quantity $(2y + a_1x + a_3)/\Delta^4$ arising in condition (iii), and uniqueness; part (c) also follows directly from uniqueness. □

**Remark.** There are several conventions for the normalisations of $\hat{\lambda}_v$ in the literature. For a comparison between these normalisations and how they relate to our chosen convention, see Appendix A.1.

If $K$ is a number field, then Theorem 3.1 tells us there is a local height function $\hat{\lambda}_v$ at each place. Although the $\hat{\lambda}_v$ are not quite quadratic, property (iii) gives a reasonable duplication formula. Writing $\ell_v = \log\max\{|x(P)|_v, 1\}$ for the local terms of the naive height, then property

(ii) says that, at least in the limit, $\hat{\lambda}_v$ behaves like $\ell_v$. In fact, as we will see in Lemma 3.3, more is true, and we have equality $\hat{\lambda}_v = \ell_v$ for all but finitely many places $v$.

In particular, any given point $P \in E(K)$ only has finitely many nonzero local heights $\hat{\lambda}_v(P) \neq 0$. With an appropriate weighted sum of these nonzero terms, we obtain the canonical height:

**Theorem 3.2.** Let $K$ be a number field, and $E/K$ an elliptic curve. For each $v \in M_K$ let $\hat{\lambda}_v : E(K_v) \setminus \{O_E\} \to \mathbb{R}$ be the local Néron height function at $v$. Then,

$$\hat{h}(P) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M_K} n_v \hat{\lambda}_v(P)$$

*Proof Sketch.* (See [5], Chapter VI, Theorem 2.1 for details.) Define $L(P)$ to be the right-hand-side, and note by the discussion above there are only finitely many nonzero terms. Using properties (i) and (ii) from Theorem 3.1, each difference $\hat{\lambda}_v - \ell_v$ is bounded. Summing over the nonzero terms, $L(P) - h(P)$ is bounded, thus $L(P) - \hat{h}(P)$ is bounded. Using property (iii) for $\lambda_v$, one also verifies that $L(2P) = 4L(P)$. Finally, defining $F = L - \hat{h}$, we deduce that $F$ is bounded and quadratic, hence identically zero. □

Although the Néron local height $\hat{\lambda}_v$ is normalised as above for theoretical purposes, we will define a slight modification of the local height which is preferred in the literature for the purposes of explicit computation, simplifying many formulae.

**Definition** (Modified Local Height)**.** For each place $v \in M_K$, define the *modified local height* at $v$ by:

$$\lambda_v = \hat{\lambda}_v + \frac{1}{6} \log|\Delta|_v$$

The duplication formula then becomes slightly simpler:

$$\lambda_v(2P) = 4\lambda_v(P) - 2\log|2y + a_1 x + a_3|_v.$$

Summing over $v \in M_K$, we see (by taking logs in the product formula) that the contribution of $\frac{1}{6}\log|\Delta|_v$ cancels, and thus we still have a decomposition:

$$\hat{h}(P) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M_K} n_v \lambda_v(P).$$

Unlike $\hat{\lambda}_v$, note $\lambda_v$ is no longer invariant under change of Weierstrass equation $E \to E'$. But, if we denote the corresponding modified local heights by $\lambda_v$, $\lambda'_v$ and the discriminants by $\Delta$, $\Delta'$, then

$$\lambda_v = \lambda'_v + \frac{1}{6}\log|\Delta/\Delta'|_v.$$

## 3.2 Local Error Functions

The decomposition of $\hat{h}$ into local heights provides a natural decomposition of $h - \hat{h}$,

$$h - \hat{h} = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M_K} n_v \Psi_v$$

where $\Psi_v$ is the *local error function* at $v$,

$$\Psi_v : E(K_v) \to \mathbb{R}$$
$$\Psi_v(P) = \begin{cases} 0 & P = O_E \\ \log \max\{|x(P)|_v, 1\} - \lambda_v & P \neq O_E \end{cases}$$

which, by the properties of $\hat{\lambda}_v$ given in Theorem 3.1, is continuous and bounded on $E(K_v)$. Further, by Lemma 3.3, $\Psi_v$ is identically zero for all but finitely many places. Our approach for bounding the height difference $h - \hat{h}$ will thus be to bound $\Psi_v$ at each place, and sum these local bounds.

## 3.3 Nonarchimedean Local Heights

The behaviour of local heights $\hat{\lambda}_v$ at nonarchimedean places $v \in M_K^0$ is well-understood and can be described with a set of explicit formulae depending on the reduction type of $E/K_v$.

We fix the following notation:

| | |
|---|---|
| $K_v$ | the completion of number field $K$ at a nonarchimedean place $v$, |
| $n_v$ | the local degree, |
| $v$ | the normalised valuation, |
| $\lvert\cdot\rvert_v$ | the absolute value extending one on $\mathbb{Q}$, |
| $\pi$ | a uniformiser, |
| $\mathcal{O}_{K_v}$ | the valuation ring, |
| $k_v$ | the residue field. |

For an elliptic curve $E/K_v$ which is minimal at $v$, we denote its reduction mod $\pi$ by $\tilde{E}/k_v$. We write $E_0(K_v)$ to mean the subgroup of points in $E(K_v)$ reducing to a nonsingular point on $\tilde{E}$.

We call the finite quotient $E(K_v)/E_0(K_v)$ the *component group*. Its order is the *Tamagawa number*, for which we write $c_v$.

### 3.3.1 Reduction Types & Tate's Algorithm

The *reduction type* of $E/K_v$ is a classification of the *special fiber* of a *minimal proper regular model*. These geometric notions are beyond the scope of this essay, but are covered in great detail in Silverman's *Advanced Topics* book ([5], Chapters III, IV, V). We will use Kodaira's notation to denote the reduction types: Type $\mathrm{I}_0$ for good reduction, Type $\mathrm{I}_n$ for multiplicative reduction, and Types III, IV, $\mathrm{I}_0^*$, $\mathrm{I}_n^*$, $\mathrm{IV}^*$, $\mathrm{III}^*$ and $\mathrm{II}^*$ for cases of additive reduction.

Given an integral Weierstrass equation for $E$, Tate's algorithm [6] gives a method to compute the reduction type (amongst other quantities), by checking a series of arithmetic conditions on the Weierstrass coefficients. Furthermore, since $K_v$ is the completion of a number field and thus complete, Tate's algorithm will simultaneously compute the Tamagawa number $c_v$. Thus, by tracing the steps of Tate's algorithm, we essentially have an arithmetic classification of the reduction types and Tamagawa number, which we will use extensively throughout this essay.

Silverman gives a complete description of the algorithm in [5, IV.9.4], which we adapt into a flowchart in Appendix A.2. Following Silverman, we label Steps 1-11 of the algorithm, also indicated in the flowchart. Finally, we note that Step 11 is only reached if the initial Weierstrass was not minimal (and thus, if we assume the input to be minimal, we must terminate in Steps 1-10).

### 3.3.2 Silverman's Explicit Formulae

Assume $E/K$ is given by a Weierstrass equation which is minimal at $v$. In this section, we provide explicit formulae for the local height at a nonarchimedean place $v \in M_0^K$, splitting into cases by reduction type. First, for points $P \in E_0(K_v)$ the local height is easy to understand (regardless of the reduction type or even minimality):

**Lemma 3.3.** Let $E/K_v$ be an elliptic curve with $v$-integral coefficients and discriminant $\Delta$. Then, for each $P \in E_0(K_v)$ the Néron local height is given by the formula:

$$\hat{\lambda}_v(P) = \log \max\{|x(P)|_v, 1\} - \frac{1}{6}\log|\Delta|_v$$

*Proof.* See [5, Theorem VI.4.1]. □

Thus, it remains to consider cases where $P \notin E_0(K)$ (thus $c_v > 1$).

If $E/K_v$ has *split* multiplicative reduction, then by Tate's $p$-adic uniformisation theorem [5, Theorem V.5.3], $E$ is isomorphic over $K_v$ to a Tate curve. That is to say, there exists $q \in K_v^\times$ such that $|q|_v < 1$ and we have a parameterisation

$$\phi : K_v^\times / q^{\mathbb{Z}} \xrightarrow{\sim} E.$$

Moreover, the parameterisation identifies the subgroups $\mathcal{O}_{K_v}^\times \cong E_0(K_v)$ and hence induces an isomorphism:

$$E(K_v)/E_0(K_v) \xrightarrow{\sim} K^\times/q^{\mathbb{Z}}\mathcal{O}_{K_v}^\times$$

Then, the normalised valuation $v : K_v^\times \to \mathbb{Z}$ incudes a further isomorpism

$$\kappa : E(K_v)/E_0(K_v) \xrightarrow{\sim} \mathbb{Z}/N\mathbb{Z}$$

where $N = v(q) = v(\Delta(E))$, and thus $N = c_v$ is the Tamagawa number. In particular, the component group is cyclic and we have fixed an isomorphism with $\mathbb{Z}/N\mathbb{Z}$. We say that $P$ *lies in the $n$-th component* if $\kappa(P) = n$. With this in mind, we can state the result:

**Lemma 3.4.** For a Tate curve as above and $P \notin E_0(K_v)$, suppose $P$ lies in the $n$-th component. Then, the *modified* local height is given by the explicit formula:

$$\lambda_v(P) = \frac{n(N-n)}{N^2}\log|\Delta|_v = -\frac{n(N-n)}{N}\frac{\log \#k_v}{n_v}$$

*Proof.* This is a restatement of [5, Proposition VI.4.2], or [4, Theorem III.5.1]. Note that in our normalisation, the absolute value and discrete valuation are related by

$$v(\,\cdot\,) = -\frac{n_v}{\log \#k_v}\log|\cdot|_v$$

and that $N = v(q) = v(\Delta)$ and $n = v(\phi^{-1}(P))$. □

**Remark.** If this formula is to be computationally useful, we should justify that the component $n = \kappa(P)$ of $P$ can be easily computed. In [7], Cremona provides a detailed discussion on explicity computing the isomorphism $\kappa$. However, for our purposes we need only know $\kappa$ 'up to sign', since the formula is invariant under the automorphism $n \mapsto -n$ of $\mathbb{Z}/N\mathbb{Z}$. To this end, we may use a simple formula given by Silverman in [8, Lemma 5.1],

$$n = \pm\min\left\{v(2y + a_1x + a_3),\ \frac{1}{2}v(\Delta)\right\}.$$

Now suppose $P \notin E_0(K)$ and $E$ has *non-split* multiplicative reduction. Since $c_v > 1$, we are left with the case $c_v = 2$ and reduction type $\mathrm{I}_N^*$, $N$ even. Since $\lambda_v$ is invariant under field extension, we may consider the unramified quadratic extension $L_w/K_v$ in which $E$ has *split* multiplicative reduction. Then we simply use the formula in Lemma 3.4, noting $P$ lies in the component of order 2 and thus $n = N/2$. (Note also $(\log \#k_v)/n_v$ is invariant under *unramified* field extension.)

It remains to consider cases of additive reduction with $c_v > 1$, where the reduction type is thus one of III, IV, $\mathrm{I}_m^*$, IV$^*$, or III$^*$. We state the formulae in Theorem 3.5, including the results above for completeness.

**Theorem 3.5** (Silverman's Explicit Formulae)**.** Suppose $E/K_v$ is given by a *minimal* Weierstrass equation at $v \in M_0^K$, with coefficients $a_1, \ldots, a_6$.

Suppose that $P = (x, y) \in E(K_v) \setminus \{O_E\}$. The *modified* local height $\lambda_v$ satisfies:

(i) The value of $\lambda_v(P)$ depends on the image of $P$ in $E(K_v)/E_0(K_v)$.

(ii) If $P \in E_0(K_v)$ then

$$\lambda_v(P) = \log\max\{|x(P)|_v, 1\}$$

(iii) If $E$ has Kodaira type $I_N$ and $P \in E(K_v) \setminus E_0(K_v)$ lies in the $n$-th component,

$$\lambda_v(P) = -\frac{n(N-n)}{N}\frac{\log \#k_v}{n_v}$$

(iv) If $E$ has Kodaira type IV or IV$^*$ and $P \notin E_0(K_v)$,

$$\lambda_v(P) = \frac{2}{3}\log|2y + a_1x + a_3|_v$$

(v) If $E$ has Kodaira type III, $\mathrm{I}_m^*$, or III$^*$ and $P \notin E_0(K_v)$,

$$\lambda_v(P) = \frac{1}{4}\log|3x^4 + b_2x^3 + 3b_4x^2 + 3b_6x + b_8|_v$$

*Proof.* Parts (ii), (iii) are Lemma 3.3, Lemma 3.4. Parts (iv), (v) are given by Silverman in [8]. See also [9, Proposition 5] for further references and a summary of the history of this result. □

### 3.3.3 Exhaustive Analysis

Based on Silverman's explicit formulae, we can go further and give an exhaustive list of values taken by the local error function $\Psi_v$ at any nonarchimedan place where $E$ is minimal. This will be an essential tool for sharply bounding the nonarchimedean contribution to the height difference $h - \hat{h}$, and also for computing the nonarchimedean contribution to $\hat{h}$ in the Muller-Stoll algorithm.

**Theorem 3.6.** Suppose $E/K_v$ is given by a *minimal* Weierstrass equation at $v \in M_0^K$, with coefficients $a_1, \ldots, a_6$. Then,

(i) For each $P \in E(K)$,

$$\Psi_v(P) = 0 \iff P \text{ has nonsingular reduction at } v$$

(ii) The list of nonzero values taken by $\Psi_v(P)$, as $P$ ranges over $E(K)$, depends only on the reduction type and Tamagawa number, and is given exactly by Table 1 in terms of

$$\mu_v(P) = \Psi_v(P) \cdot \frac{n_v}{\log \# k_v}$$

which is always rational.

**Table 1:** Nonzero values of $\mu_v(P)$, for $E$ minimal at $v$ with prescribed reduction type and Tamagawa number.

| Reduction Type | Tamagawa Number $c_v$ | Nonzero values of $\mu_v(P)$ |
|---|---|---|
| Any | 1 | None |
| $\mathrm{I}_m$ | $m$ | $n(m-n)/n$, $0 < n < m$ |
| $\mathrm{I}_m$ | 2 | $m/4$ |
| III | 2 | $1/2$ |
| IV | 3 | $2/3$ |
| $\mathrm{I}_0^*$ | 2, 4 | 1 |
| $\mathrm{I}_m^*$ | 2 | 1 |
| $\mathrm{I}_m^*$ | 4 | 1, $(m+4)/4$ |
| IV$^*$ | 3 | $4/3$ |
| III$^*$ | 2 | $3/2$ |

*Proof.* (cf. [9, Proposition 6]) We exhaust over the cases by following Tate's algorithm, so we accumulate arithmetic conditions on the quantities $a_i$, $b_i$, etc. and use these to evaluate Silverman's explicit formulae from Theorem 3.5. Here, we will give a flavour of the proof by exhibiting a few cases.

If $c_v = 1$ then $E(K) = E_0(K)$ and thus by Lemma 3.3, $\Psi_v(P) = 0$ identically. Thus, we may henceforth assume $c_v > 1$ and only consider points with singular reduction, $P \notin E_0(K_v)$. Then, we have $|x(P)|_v \leq 1$ and hence $\Psi_v(P) = -\lambda_v(P)$.

- **Type $\mathrm{I}_m$.** By Lemma 3.4, if $P$ lies in the $n$-th component,

$$\Psi_v(P) = \frac{n(m-n)}{m}\frac{\log \# k_v}{n_v}$$

In the case of split multiplicative reduction, there is a point in each component $0 < n < m$. In the case of non-split multiplicative reduction, then since we assumed $c_v > 1$, we know $m$ is even and $c_v = 2$. Each point of good reduction lies in the trivial component; each point of bad reduction lies in component $n = m/2$. The values of $\mu_v(P)$ follow.

At this point, following Tate's algorithm, we may assume under a translation $P = (0, 0)$ which is valid within this proof since $\lambda_v$ is invariant under translation. Since $c_v > 1$ and we have eliminated case $\mathrm{I}_m$, we reach Step 4, and thus have the assumptions:

$$a_6 = 0, \quad \pi \mid a_3, a_4, b_2, \quad \pi^2 \mid b_6, b_8$$

- **Type III, $c_v = 2$.** In this case we gain the further condition $\pi^3 \nmid b_8$, so $v(b_8) = 2$ exactly. Since $x(P) = 0$ the explicit formula simplifies to:

$$\lambda_v(P) = \frac{1}{4}\log|b_8|_v = -\frac{1}{2}\frac{\log \#k_v}{n_v}$$

- **Type IV, $c_v = 3$.** In this case, $\pi^3 \nmid b_6 = a_3^2 + 4a_6 = a_3^2$. Thus $\pi^2 \nmid a_3$ and so $v(a_3) = 1$. Again the explicit formula simplifies to:

$$\lambda_v(P) = \frac{2}{3}\log|a_3|_v = -\frac{2}{3}\frac{\log \#k_v}{n_v}$$

- **Types $\mathrm{IV}^*$, $\mathrm{I}_0^*$, $\mathrm{III}^*$** all follow in a similar fashion, using the cumulative assumptions to determine the valuation of $b_8$ or $a_3$ as appropriate.

- **Type $\mathrm{I}_m^*$, $m \geq 2$ even, $c_v = 2$.** Following Tate's algorithm, we terminated in iteration $i = (m-2)/2$ of the Step 7 subroutine, and so we have translated $E$ such that

$$\pi \mid a_1, a_2, \quad \pi^{\frac{m+4}{2}} \mid a_3, a_4, \quad \pi^{m+3} \mid a_6$$

and the polynomial

$$\mathcal{P}(T) = T^3 + a_{2,1}T^2 + a_{4,2}T + a_{6,3} \equiv T^3 + a_{2,1}T^2 \pmod{\pi}$$

has a double and single root mod $\pi$. This forces $v(a_2) = 2$ exactly, and $-a_{2,1} \not\equiv 0 \pmod{\pi}$ is the simple root.

Since $c_v = 2$ it suffices to find just one point with singular reduction and evaluate $\lambda_v(P)$. To do this, we use Hensel's lemma to lift $-a_{2,1}$ to a root $\alpha \in \mathcal{O}_K^*$ of $\mathcal{P}(T)$. Rewriting the Weierstrass equation as $y^2 = \pi^3\mathcal{P}(\pi^{-1}T)$, it follows that $P = (\pi\alpha, 0)$ lies on $E(K_v)$ and reduces to the singular point $\tilde{P} = (0, 0)$. Finally, we evaluate:

$$3x^4 + b_2x^3 + 3b_4x^2 + 3b_6x + b_8 \equiv (-\pi\alpha)^4 \pmod{\pi}$$

$$\implies \lambda_v(P) = \frac{\log \#k_v}{n_v}$$

The remaining cases for $\mathrm{I}_m^*$ are checked in [9]. □

## 3.4 Archimedean Local Heights

In this section, we describe how the uniformisation of elliptic curves over $\mathbb{C}$ can be used to provide explicit formulae for the local height and local error, in terms of elliptic functions.

**Uniformisation of the Curve.** Starting with an arbitrary Weierstrass equation, we are free to make the usual translation of $y$ (not affecting the naive height $h$) for an equation of the form (sometimes known as a $b$-model):

$$E : y^2 = 4x^3 + b_2x^2 + 2b_4x + b_6$$

Considering the substitution $x = x' - \frac{b_2}{12}$, we have

$$y^2 = 4(x')^3 - g_2(x') - g_4 \quad \text{where} \quad g_2 = \frac{1}{12}c_4 \quad g_4 = \frac{c_6}{216}$$

so by the Uniformisation Theorem [3, VI.6] there exists a lattice $\Lambda$ such that $(x', y)$ is parameterised by $z \mapsto (\wp(z), \wp(z'))$, where the Weierstrass $\wp$-function is given by:

$$\wp(z;\Lambda) = \frac{1}{z^2} + \sum_{w \in \Lambda \setminus \{0\}} \left( \frac{1}{(z-w)^2} - \frac{1}{w^2} \right)$$

So, $E$ is parameterised by:

$$\begin{aligned} \phi : \mathbb{C}/\Lambda &\to E(\mathbb{C}) \\ z &\mapsto \left( \wp(z) - \frac{b_2}{12},\ \wp'(z) \right) \end{aligned}$$

For computational purposes, we will need to explicitly obtain a basis $\{\omega_1, \omega_2\}$ for $\Lambda$ from $g_2$ and $g_4$, and we will need to be able to evaluate the isomorphism in both directions (its inverse is called the *elliptic logarithm*). Good algorithms for both of these problems are known, described in Cohen's book [10, Algorithms 7.4.7-8], and included in computer algebra packages such as `PARI/GP`.

**The Local Height.** Once we have uniformised the curve, we have an explicit formula for the Néron local height:

$$\hat{\lambda}_v(z) = \Re(z\eta(z)) - 2\log|\sigma(z)| - \frac{1}{6}\log|\Delta|$$

where the Weierstrass $\sigma$-function is given by:

$$\sigma(z;\Lambda) = \prod_{w \in \Lambda \setminus \{0\}} \left(1 - \frac{z}{w}\right) e^{\left(\frac{z}{w} + \frac{1}{2}\frac{z}{w}^2\right)}$$

and $\eta$ is the *quasi-period homomorphism* of the lattice; see [5, Theorem IV.3.2] for details.

**The Local Error.** In what follows, it will be convenient to use the *Néron local height* $\hat{\lambda}_v$ instead of the modified local height $\lambda_v$. We thus adjust our normalisation of $\Psi_v$ slightly, defining

$$\hat{\Psi}_v(P) = \log\max\{|x(P)|_v, 1\} - \hat{\lambda}_v = \Psi_v(P) + \frac{1}{6}\log|\Delta|_v$$

This new error function may thus be viewed as an function on $\mathbb{C}/\Lambda$ via

$$\hat{\Psi}_v(z) = \log\max\left\{ \left| \wp(z) - \frac{b_2}{12} \right|, 1 \right\} - \hat{\lambda}_v(z)$$

Following Bruin [11], let $t_1, t_2 \in \mathbb{C}/\Lambda$ be the roots of $\wp(z) - \frac{b_2}{12}$, which correspond to the inverse points on $E$ with $x(P) = 0$, and so $t_1 + t_2 = 0$. Then the two functions:

$$\begin{aligned} &\log\left| \wp(z) - \frac{b_2}{12} \right| \\ &\hat{\lambda}(z) - \frac{1}{2}\hat{\lambda}(z - t_1) - \frac{1}{2}\hat{\lambda}(z - t_2) \end{aligned}$$

have the same image under the Laplacian operator $\partial\bar{\partial}$, namely $2\pi(\delta_{t_1} + \delta_{t_2} - 2\delta_0)$, where $\delta_{z_0}$ denotes the Dirac delta function at $z = z_0$. Thus, they differ by a constant $I$, which allows us to write

$$\hat{\Psi}(z)_v = \begin{cases} -\frac{1}{2}\hat{\lambda}_v(z - t_1) - \frac{1}{2}\hat{\lambda}_v(z - t_2) + I & \left|\wp(z) - \frac{b_2}{12}\right| \geq 1 \\ -\hat{\lambda}_v & \left|\wp(z) - \frac{b_2}{12}\right| \leq 1 \end{cases} \qquad (\star)$$

This gives a practical formula for evaluating $\hat{\Psi}_v(z)$ at specific $z$, assuming we have a good method to evaluate the local height $\hat{\lambda}_v$ (see Chapter 5).

We note that the constant $I$ can be determined by ensuring the two cases are equal for some $z$ lying on the boundary $\left|\wp(z) - \frac{b_2}{12}\right| = 1$.

# 4 Computing Height Difference Bounds

Armed with the theory of local heights, we are now ready to address the first main problem of the essay, and describe how to algorithmically bound the height difference. The strategy is to use the decomposition of the height difference into local error functions:

$$h - \hat{h} = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M_K} n_v \Psi_v$$

and bound each function $\Psi_v$ individually. For each nonarchimedean place $v \in M_K^0$, we can give exact, sharp bounds using the exhaustive analysis of values taken by $\Psi_v$. The archimedean contribution is more mysterious, but we will give two different numerical methods which are used to obtain bounds. Finally, we will justify this approach by proving that sharp *local* bounds lead to tight *global* bounds for the height difference on the set of algebraic points $E(\bar{\mathbb{Q}})$.

## 4.1 Nonarchimedean Contribution

We begin with an easy corollary of the exhaustive analysis of $\Psi_v$ given in Theorem 3.6. Note that, in the hypothesis for Theorem 3.6, we assumed $E$ is minimal at $v$. However, unless $E$ is a globally minimal Weierstrass equation (the existence of which is not guaranteed over an arbitrary number field $K$, see [3, VIII.8]), we have some more work to do. Thankfully, we can easily drop the minimality assumption in exchange for an extra term in the upper bound.

**Corollary 4.1.** For every nonarchimedean place $v \in M_K^0$ and every $P \in E(K_v)$,

$$0 \leq \Psi_v(P) \leq \frac{\log \# k_v}{n_v} \left( \alpha_v + \frac{1}{6} v(\Delta / \Delta_v^{\min}) \right)$$

where $\alpha_v$ depends on the reduction type and Tamagawa number of a minimal model $E_v^{\min}$ for $E$ at $v$, given in Table 2, and $\Delta_{\min}^v$ is the discriminant of this model. Furthermore, each bound is attained by some point in $E(K_v)$.

**Table 2:** Definition of constant $\alpha_v$ appearing in Theorem 4.1.

| Reduction Type | Tamagawa Number $c_v$ | $\alpha_v$ |
|---|---|---|
| Any | 1 | 0 |
| $\mathrm{I}_m$, $m$ even | $2, m$ | $m/4$ |
| $\mathrm{I}_m$, $m$ odd | $m$ | $(m^2-1)/4m$ |
| III | 2 | 1/2 |
| IV | 3 | 2/3 |
| $\mathrm{I}_0^*$ | $2, 4$ | 1 |
| $\mathrm{I}_m^*$ | 2 | 1 |
| $\mathrm{I}_m^*$ | 4 | $(m+4)/4$ |
| $\mathrm{IV}^*$ | 3 | 4/3 |
| $\mathrm{III}^*$ | 2 | 3/2 |

*Proof.* In the case where $E$ is minimal, this follows immediately from Theorem 3.6. It remains to understand how $\Psi_v$ changes when $E$ is not minimal. In this case, let $E'$ be a minimal model, and denote all quantities relating to $E'$ decorated with a prime. By standard theory $E$, $E'$ are related by a substitution

$$x = u^2 x' + r \ , \quad y = u^3 y' + s x' + t$$

where $u, r, s, t \in \mathcal{O}_{K_v}$ and $v(\Delta_v^{\min}) = v(u^{-12}\Delta)$. Using the definition of $\Psi_v$, and the transformation formula for the modified local height $\lambda_v$, we obtain:

$$\Psi_v(P) = \Psi_v'(P) + \log \left( \frac{\max\{1, |u^2 x'(P) + r|_v\}}{\max\{1, |x'(P)|_v\}} \right) - \frac{1}{6} \log |\Delta / \Delta_v^{\min}|_v$$

Investigating the central term, its maximum is 0 (whenever $|x'(P)|_v \leq 1$) and minimum is $\log|u^2|_v$ (for $|x'(P)|_v$ sufficiently large). The bounds for $\Psi_v$ follow. The lower bound is attained by some point with non-singular reduction and $|x'(P)|_v$ large, and the upper bound is attained by some point with singular reduction. □

The algorithm is now simple to describe:

**Method 4.2.**

1. Factorise the discriminant ideal $(\Delta) = \mathfrak{p}_1, \ldots, \mathfrak{p}_r$ into primes of $\mathcal{O}_K$.
2. For each prime $v = \mathfrak{p}_i$,
   (i) Use Tate's algorithm on $E/K_v$ to compute a minimal model, its discriminant $\Delta_v^{\min}$, its reduction type, and its Tamagawa number.
   (ii) Look up the corresponding value of $\alpha_v$ in Table 2.
3. The nonarchimedean contribution to the height difference bound is then:

$$0 \leq \sum_{v \in M_K^0} n_v \Psi_v \leq \sum_{v \in M_K^0} \left( \alpha_v + \frac{1}{6} v(\Delta/\Delta_v^{\min}) \right) \log \# k_v$$

From a complexity standpoint, any algorithm which begins with factorisation of a potentially large number is not particularly efficient. However, in practice, this is feasible for reasonably sized elliptic curves, and is not the limiting factor in computing height difference bounds; computing a good bound for the archimedean contribution, as we will see, is much more time-consuming.

## 4.2 Archimedean Contribution

Since the local height at $v \in M_K^\infty$ is given by a transcendental function, sharp bounds for $\Psi_v$ are difficult to obtain explicitly as in the nonarchimedean case. One approach is to further decompose $\Psi_v = -\sum_{n=1}^{\infty} 4^{-n-1} \log \Phi_v(2^n P)$ and appeal to numerical methods to bound the function $\Phi_v$. Alternatively, one can obtain directly obtain an approximation for the extrema of $\Psi_v$, to arbitrary precision, via a numerical algorithm of Bruin. In this section, we summarise these two methods, and provide a comparison of their advantages and disadvantages in practical use.

### 4.2.1 Cremona-Pickett-Siksek (CPS) Method

To describe this method, we begin by decomposing the local error $\Psi_v$.

*Derivation.* First, we define the polynomials:

$$f(x) = 4x^3 + b_2 x^2 + 2b_4 x + b_6$$
$$g(x) = x^4 - b_4 x^2 - 2b_6 x - b_8$$

such that for $P \in E(K_v)$, the duplication formula reads:

$$x(2P) = \frac{g(x(P))}{f(x(P))}$$

(Note that we will slightly abuse notation and write $f, g$ as functions of $x$ or $P$ interchangeably.) Now, we see that:

$$h(2P) - 4h(P) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M_K} n_v \Phi_v(P)$$

$$\Phi_v(P) = \begin{cases} 1 & \text{if } P = O_E \\ \dfrac{\max\{|f(x(P))|_v, |g(x(P))|_v\}}{\max\{1, |x(P)|_v\}^4} & \text{if } P \neq O_E \end{cases}$$

And note each $\Phi_v$ is continuous and bounded on on $E(K_v)$. Now we apply Tate's telescoping trick: replace $P$ by $2^n P$, divide by $4^n$, and take the limit $n \to \infty$ to get:

$$\hat{h}(P) - h(P) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M_K} n_v \sum_{i=0}^{\infty} \frac{1}{4^{i+1}} \log \Phi_v(2^i P)$$

$$\implies \Psi_v(P) = -\sum_{i=0}^{\infty} \frac{1}{4^{i+1}} \log \Phi_v(2^i P) \qquad \square$$

**Real Places.** Now, we explain how to bound $\Phi_v$ in the real case $K_v = \mathbb{R}$. We begin by splitting into the two regions $|x(P)| \leq 1$ and $|x(P)| \geq 1$ on $E(\mathbb{R})$.

Note that $x$ is an $x$-coordinate on $E(\mathbb{R})$ iff $f(x) \geq 0$. Thus, in the first region,

$$d = \sup_{|x(P)| \leq 1} \Phi_v(P) = \sup_{x \in D} \max\{|f(x)|, |g(x)|\}$$

where

$$D = \{x \in [-1, 1] \,:\, f(x) \geq 0\}$$

is a finite set of closed intervals. So we have reduced the problem to finding the maximum of two polynomials on a closed interval, which is easily done by considering the endpoints, the extrema of $f, g$, and the points where they are equal.

In the latter region, we use the parameter $t = 1/x$. Then,

$$d' = \sup_{|x(P)| \geq 1} \Phi_v(P) = \sup_{x \in D'} \max \left\{ \left| t^4 f\left(\frac{1}{t}\right) \right|, \left| t^4 g\left(\frac{1}{t}\right) \right| \right\}$$

where

$$D' = \{t \in [-1, 1] : f(1/t) \geq 0\}$$

is a finite set of closed intervals, and $t^4 f(1/t)$, $t^4 g(1/t)$ are polynomials, thus the bound as easily computed as in the first region.

To finish, we simply take $\max(d, d')$ as our upper bound for $\Phi_v$. The lower bound is computed in almost exactly the same way. For further details, see [9, §7].

**Complex Places.** The CPS method extends to complex places, beginning with the exact same idea, but in the complex case we see that we should take the regions $D = D' = \{z \in \mathbb{C} \,:\, |z| \leq 1\}$ to be the unit disk. It remains to describe a method to find the extrema of the maximum of two polynomials $F, G \in K[X]$ on the unit disk, where $K \subseteq \mathbb{C}$ is a number field.

For a more theoretical approach, Cremona et al. describe a 'Grobner method' [9, §8]. This involves using Lagrange multipliers and the maximum modulus principle to determine a set of simultaneous polynomial equations for the suprema; these can then be solved using Grobner bases. However, the authors note that in practice, a simple numerical approach of repeated quadrisection on the unit circle [9, §9] yields a faster algorithm.

### 4.2.2 Bruin's Method

For a complex place $K_v = \mathbb{C}$ (which we are free to assume upon extending $K$), Bruin's paper [11] avoids the decomposition of $\Psi_v$, and instead provides an algorithm to directly compute the extrema of $\hat{\Psi}_v$ to arbitrary precision. This is again a numerical algorithm, using repeated bisection, but this time on a fundamental domain $\mathbb{C}/\Lambda$ of the period lattice of the elliptic curve $E$.

**Bounding the Derivative.** Recall that we previously described (Section 3.4) a method to evaluate $\hat{\Psi}_v = \Psi_v + \frac{1}{6} \log|\Delta|_v$ at some given $z$. To ensure the bisection algorithm finds the extrema to within a desired precision, we need to bound the derivative of $\hat{\Psi}_v$.

Since the local height $\hat{\lambda}_v : \mathbb{C}/\Lambda \setminus \{0\} \to \mathbb{R}$ is smooth and takes real values, we can take $Z : \mathbb{C}/\Lambda \setminus \{0\} \to \mathbb{C}$ to be the unique smooth function such that:

$$d\hat{\lambda}_v(z) = -\left(Z(z)dz + \overline{Z(z)}d\bar{z}\right)$$

In fact,

$$Z(z) = \zeta_\Lambda(z) - Cz - D\bar{z}$$

where $C, D$ are known invariants of the lattice (see [11]). So, differentiating the explicit formula for $\hat{\Psi}_v$,

$$d\hat{\Psi}(z) = W(z)dz + \overline{W(z)}d\bar{z} \quad \text{for } \left|\wp(z) - \frac{b_2}{12}\right| \neq 1$$

$$W(z) = \begin{cases} \frac{1}{2}Z(z-t_1) + \frac{1}{2}Z(z-t_2) & \left|\wp(z) - \frac{b_2}{12}\right| > 1 \\ Z(z) & \left|\wp(z) - \frac{b_2}{12}\right| < 1 \end{cases}$$

and furthermore, using the expression above for $Z(z)$ and that $\zeta'_\Lambda(z) = -\wp(z)$, we check that $\partial\bar{\partial}W = 0$ in each region, i.e. $W$ is *harmonic* and thus $|W|$ attains its maximum on the boundary $\left|\wp(z) - \frac{b_2}{12}\right|$.

We have reduced the problem to bounding our expressions in $Z$ on the boundary. We know:

$$dZ(z) = -(\wp(z) + C)dz - Dd\bar{z}$$

and we can use this to find expressions for $dZ(z-t_1)$ and $dZ(z-t_2)$. Then, summing these and using the group law for $\wp$, a somewhat tedious calculation yields

$$d\left[\frac{1}{2}Z(z-t_1) + \frac{1}{2}Z(z-t_2)\right] = \left(-C - \frac{b_2}{12} - \frac{1}{2}\frac{b_4}{\wp(z) - \frac{b_2}{12}} - \frac{1}{2}\frac{b_6}{\left(\wp(z) - \frac{b_2}{12}\right)^2}\right)dz - Dd\bar{z}$$

We now bound $Z$ by considering two points $p, q$ on the boundary. We choose a path $\gamma$ from $p$ to $q$, whose image $\gamma'$ under the function $\wp(z) - \frac{b_2}{12}$, is identified with an arc on the unit circle. Then, by a basic integral estimate,

$$\begin{aligned} Z(q) - Z(p) &= \int_\gamma dZ \\ &= -\int_\gamma \left[(\wp(z) + C)\,dz + Dd\bar{z}\right] \\ &\leq \left(\underbrace{\left|\wp(z) - \frac{b_2}{12}\right|}_{=1} + \left|C + \frac{b_2}{12}\right| + |D|\right) \text{length}(\gamma) \end{aligned}$$

By the definition of the period lattice (see for example [3, Proposition 5.2]), $dz$ pulls back to the invariant differential on $E$, so we get the estimate

$$\text{length}(\gamma) = \int_\gamma |dz| = \int_{\gamma'} \frac{|dx|}{|2y + a_1 + a_3|} \leq \int_{\theta=0}^{2\pi} \frac{d\theta}{|4e^{i\theta} + b_2e^{2i\theta} + 2b_4e^{i\theta} + b_6|^{1/2}}$$

So overall we have an estimate $|Z(q)| \leq |Z(p)| + M_1 J$. Here $M_1$ depends only on $C, D, b_2$, and $J$ is the length estimate depending only on $b_2$, $b_4$, $b_6$.

In a similar fashion, we bound

$$\left|\frac{1}{2}Z(q-t_1) + \frac{1}{2}Z(q-t_2)\right| \leq \left|\frac{1}{2}Z(q-t_1) + \frac{1}{2}Z(q-t_2)\right| + M_2 J$$

where now $M_2 = M_1 - 1 + \frac{|b_4|}{2} + \frac{|b_6|}{2}$ depends only on $C$, $D$ and the $b_i$. Combining these, we finally obtain our bound for $W$:

$$|W(z)| \leq \max\left\{|Z(p)| + M_1 J\ , \quad \left|\frac{1}{2}Z(p-t_1) + \frac{1}{2}Z(p-t_2)\right| + M_2 J\right\} \qquad (\star\star)$$

**The Recursive Bisection.** The algorithm for finding the extrema of $\hat{\Psi}_v$ is now a simple matter of repeatedly bisecting the fundamental domain (a parallelogram) for $\Lambda$, evaluating $\hat{\Psi}_v$ at the centre of each smaller parallelogram and recursively bisecting these, until eventually the bound on $|W|$ guarantees we are within a certain accuracy. We give a detailed description to compute the maximum below; computing the minimum is analogous.

**Method 4.3** (Bruin's Algorithm)**.** We notate a parallelogram in the complex plane as follows:

$$R(z_0, z_1, z_2) = \{z_0 + sz_1 + tz_2 \ : \ s, t \in [-1/2, 1/2]\} \subseteq \mathbb{C}$$

and call $z_0$ the 'centre'.

To compute the maximum of $\hat{\Psi}_v$ on $\mathbb{C}/\Lambda$ to within a desired precision $\varepsilon > 0$,

1. Begin with a fundamental domain $R(z_0, z_1, z_2)$ for $\mathbb{C}/\Lambda$, where $z_1$, $z_2$ are a basis for the lattice $\Lambda$. Explicitly evaluate $\hat{\Psi}(z_0)$ using $(\star)$, and set $\mu \leftarrow \hat{\Psi}(z_0)$, a variable which will record the largest value of $\hat{\Psi}$ encountered so far.

2. Use $(\star\star)$ to evaluate an upper bound $M$ for $|W|$ (using $p = z_0$). Then, for $z \in R(z_0, z_1, z_2)$, using the expression for $d\hat{\Psi}$ in terms of $W$, a basic integral estimate gives

$$\begin{aligned} |\hat{\Psi}(z) - \hat{\Psi}(z_0)| &\leq 2M|z - z_0| \\ &\leq M \cdot \max\{|z_1 - z_2|, |z_1 + z_2|\} \end{aligned}$$

   which we can compute easily.

3. Suppose that we have the inequality:

$$\hat{\Psi}(z_0) + M \cdot \max\{|z_1 - z_2|, |z_1 + z_2|\} \leq \mu + \varepsilon$$

   Then, on the parallelogram $R(z_0, z_1, z_2)$, we know $\hat{\Psi}_v$ never exceeds the largest value $\mu$ we already found (within precision $\varepsilon$), and there is no need to bisect the parallelogram further.

4. Otherwise, bisect $R(z_0, z_1, z_2)$ along a line parallel to its shorter edges. Say the resulting parallelograms have centres $z_0'$, $z_0''$, then set $\mu \leftarrow \max\{\mu, \hat{\Psi}(z_0'), \hat{\Psi}(z_0'')\}$. Now, repeat steps (2)-(5) recursively on each of the new parallelograms.

When the algorithm terminates (the bisection must halt for some sufficiently small parallelograms), the final value of $\mu$ is guaranteed to be within $\varepsilon$ of the true upper bound.

Note that the algorithm can be made slightly more efficient; we bounded $W$ using two cases, but if we know the rectangle $R(z_0, z_1, z_2)$ lies entirely on one side of the boundary $|\wp(z) - \frac{b_2}{12}|$, we know exactly which case to use.

### 4.2.3 Examples & Comparison

We provide numerical examples and discuss the difference in tightness of bounds, and runtime, for the two algorithms. For simplicity we will work with elliptic curves over $\mathbb{Q}$, so we use the Cremona-Pickett-Siksek method for real places we have re-implemented in Sage. We perform Bruin's method on $E(\mathbb{C})$ (which of course gives a bound over $E(\mathbb{R})$), using the `PARI/GP` script attached to his paper [11].

**Example 1.** Consider the rank 2 elliptic curve (taken from LMFDB [1]) with Cremona label `389a1`,

$$E/\mathbb{Q} : y^2 + y = x^3 + x^2 - 2x$$

Both programs terminated instantaneously. The CPS method (shifted to give a bound for $\hat{\Psi}$ instead of $\Psi$) gives:

$$0.226 \leq \hat{\Psi} \leq 0.994$$

whereas Bruin's method (with $\varepsilon = 10^{-3}$) gives:

$$0.284 \leq \hat{\Psi} \leq 0.994$$

**Example 2.** Consider the rank 3 elliptic curve with Cremona label `228920b1`,

$$E/\mathbb{Q} : y^2 = x^3 + x^2 - 14385x + 219283$$

The CPS method gives:

$$-0.924 \leq \hat{\Psi} \leq 5.460$$

whereas Bruin's method, taking 5 minutes[1] with $\varepsilon = 10^{-2}$, produced:

$$0.63 \leq \hat{\Psi} \leq 5.46$$

**Example 3.** Consider the rank 2 elliptic curve with Cremona label `338062e1`,

$$E/\mathbb{Q} : y^2 + xy = x^3 - 22332x + 1282576$$

The CPS method gives:

$$-2.478 \leq \Psi \leq 6.374$$

whereas Bruin's method, taking a grand total of 114 minutes with $\varepsilon = 10^{-2}$, produced:

$$-0.58 \leq \Psi \leq 5.83$$

**Conclusion.** Although Bruin's method is finding the extrema on $E(\mathbb{C})$ instead of $E(\mathbb{R})$, we have largely found that the bounds are superior to those given by the CPS method: the error caused by decomposing $\Psi$ is too significant. However, we note that Bruin's method is far slower than the CPS method and thus may be less useful in practice.

## 4.3 Optimality over $\bar{\mathbb{Q}}$-points

To conclude this section, we give a theoretical justification for the approach we have taken of bounding $h - \hat{h}$ by bounding each local term. In particular, we show that attaining sharp bounds for each local term actually gives tight *global bounds* over the $\bar{\mathbb{Q}}$-points $E(\bar{\mathbb{Q}})$ (although not necessarily on the $K$-points).

**Theorem 4.4.** Let $a_1, \ldots, a_6$ be algebraic integers defining a Weierstrass equation for an elliptic curve $E/\bar{\mathbb{Q}}$, discriminant. Let $K$ be any number field containing the $a_i$. Then,

$$[K:\mathbb{Q}] \inf_{P \in E(\bar{\mathbb{Q}})} \Big( h(P) - \hat{h}(P) \Big) = \sum_{v \in M_K^\infty} n_v \inf_{P \in E(\bar{K}_v)} \hat{\Psi}_v - \frac{1}{6} \log(N_{K/\mathbb{Q}} \Delta)$$

$$[K:\mathbb{Q}] \sup_{P \in E(\bar{\mathbb{Q}})} \Big( h(P) - \hat{h}(P) \Big) = \sum_{v \in M_K^\infty} n_v \sup_{P \in E(\bar{K}_v)} \hat{\Psi}_v + \frac{[K:\mathbb{Q}]}{12} \log \Delta^{\text{stable}}$$

where $\hat{\Psi}_v$ is the local error function (using $\hat{\lambda}$), and $\Delta^{\text{stable}}$ is the *stable discriminant*, which is given in terms of the *minimal discriminant* $\mathcal{D}_{E/K}$ (see [3, VIII §8]) by

$$\Delta^{\text{stable}} = \big( N_{K/\mathbb{Q}} \mathcal{D}_{E/K} \big)^{1/[K:\mathbb{Q}]}$$

[1] All runtimes recorded on Dell XPS 15 9560, Intel Core i7

*Proof.* (cf. [11, Theorem 2.1]) First, it is an exercise to show $\Delta^{\text{stable}}$ is invariant under field extension $K$. Further, $\hat{\Psi}_v = \log \max\{|x(P)|_v, 1\} - \hat{\lambda}_v$ is also invariant under field extension. Considering how the set of places $M_K$ and the indices $n_v$ and $[K:\mathbb{Q}]$ change under field extension, to prove the equation we are in fact free to take an arbitrary field extension of $K$.

Replacing $K$ by an appropriate extension, we may assume:

(i) $E/K$ has multiplicative reduction at all primes of bad reduction;

(ii) $K$ is totally complex;

(iii) All local degrees $n_v$ are even.

(it is possible to obtain such an extension by the *semi-stable reduction theorem* [3, Proposition VII.5.4] and general ramification theory for number fields). Write the nonarchimedean contribution as:

$$\sum_{v \in M_K^0} n_v \hat{\Psi}_v = \sum_{v \in M_K^0} n_v \left( \Psi_v + \frac{1}{6} \log |\Delta|_v \right)$$

Now note that at each $v \in M_0^K$ we have Kodaira type $I_m$ with $m = n_v = v(\Delta^{\min})$. Applying Corollary 4.1 with all $\alpha_v = n_v/4$, and after some fiddling with the definitions, we obtain the stated bounds:

$$-\frac{1}{6} \log(N_{K/\mathbb{Q}} \Delta) \leq \sum_{v \in M_K^0} n_v \hat{\Psi}_v \leq \frac{[K:\mathbb{Q}]}{12} \log \Delta^{\text{stable}}$$

To prove the theorem, it remains to show the global bounds we have attained are indeed the supremum and the infimum, which we do by the Weak Approximation Theorem.

Suppose $v$ is an nonarchimedean place. If there is good reduction at $v$, then $\hat{\Psi}_v$ is constant and so its upper bound is attained everywhere. Otherwise, if there is bad reduction, its upper bound is attained by at least one point (see Corollary 4.1). Furthermore, we know from Theorem 3.5 that the value of $\hat{\Psi}(P)$ will depend only on $|x(P)|_v$ and the image of $P$ in $E(K_v)/E_0(K_v)$. Recall also that $E_0(K)$ is $v$-adically open. Thus, there exists an open set $U_v \subseteq \mathbb{P}^1(K_v)$ such that the maximum is attained for all points with $x$-coordinate lying in $U_v$.

Suppose instead $v$ is a nonarchimedean place. Then $\hat{\Psi}_v$ is continuous and so there exists an open set $U_v \subseteq \mathbb{P}^1(K_v)$ on which $\hat{\Psi}_v$ is always within $\varepsilon$ of the supremum.

Now, using the approximation theorem for the finitely many $v$ which are archimedean or have bad reduction, we obtain $x \in \mathbb{P}^1(K)$ such that $x \in U_v \subseteq E(K_v)$ at all $v$. Replacing $K$ by a quadratic extension if necessary, there exists $y$ such that $P = (x, y) \in E(K)$. By construction, this point has $h(P) - \hat{h}(P)$ within $\varepsilon$ of the stated bound, and taking $\varepsilon \to 0$ proves it is indeed the supremum.

The proof for the infimum is analogous. □

# 5 Computing Canonical Heights

In this chapter, we reach the second main goal of the essay: to describe an efficient algorithm to compute the canonical height $\hat{h}(P)$ of a point $P \in E(K)$. The limit definition of the canonical height

$$\hat{h}(P) = \lim_{N\to\infty} \frac{1}{4^N} h(2^N P)$$

has poor convergence properties and requires computing large multiplies of $P$, so does not give a practical algorithm we can use for computation. Instead, once again, we tackle the archimedean and nonarchimedean contributions separately. The archimedean contribution can be calculated with a variety of efficient algorithms; here we exposit Tate's series, and Silverman's refinement thereof. The nonarchimedean contribution, however, poses an issue. The simple approach is to identify the primes of bad reduction by factoring $\Delta$, and then apply Silverman's explicit formulae (Theorem 3.5) at each. This has been used in practice (see [8]), but factorisation of $\Delta$ becomes extremely slow for curves with large coefficients. Instead, we will present a factorisation-free algorithm of Müller and Stoll, which generally gives a far superior runtime.

## 5.1 Archimedean Contribution

### 5.1.1 Tate's Series

Originally described in a letter from Tate to Serre [12], Tate's series gives an elegant algorithm to compute the modified local height function $\lambda_v = \hat{\lambda}_v + \frac{1}{6}\log|\Delta|_v$ at a place $v \in M_K$, which is valid provided there are no points on $E(K_v)$ with $x$-coordinate equal to 0.

For our computational purposes, we are only interested in the archimedean cases.[2] If $v$ is a real place, then we can always ensure the condition holds by first making an appropriate shift $x' = x + r$. However, if $v$ is a complex place then $K_v = \mathbb{C}$ is algebraically closed, and there will always exist points with $x(P) = 0$. Thus, Tate's series works for real places (and in particular, suffices for computations over $\mathbb{Q}$), but we will need more work later to extend this to complex places with Silverman's series.

**Derivation.** Let $E/K$ be an elliptic curve, and fix a Weierstrass equation with coefficients $a_1, \ldots, a_6 \in K$. Recall the duplication formula:

$$x(2P) = \frac{x^4 - b_4x^2 - 2b_6x - b_8}{4x^3 + b_2x^2 + 2b_4x + b_6} = \frac{g(P)}{f(P)}$$

In Theorem 3.1, Property (a)(ii) of $\lambda_v$ tells us that in the limit $P \to O_E$ we expect it to behave like $\log|x|_v$, so it is natural to study the function:

$$\rho(P) = \begin{cases} 4\left(\lambda_v(P) - \log|x(P)|_v\right) & P \neq O_E \\ 0 & P = O_E \end{cases}$$

where the factor of 4 is admittedly added in hindsight to clean things up later. We note $\rho$ is *not* just a scaling of the local error function $\Psi$, as no maximum is taken between $|x(P)|_v$ and 1.

Now using the identity $f(P) = (2y + a_1x + a_3)^2$ and the duplication formulae for $\lambda_v$ and $x$, we can derive a duplication formula for $\rho$:

$$\rho(2P) = 4\rho(P) - 4\log\left|\frac{g(P)}{x(P)^4}\right|_v$$

Then, by repeatedly substituting $\rho$ into itself $N$ times, and recalling the assumption that $x \neq 0$ on $E(K_v)$, we obtain the summation:

$$\rho(P) = \sum_{n=0}^{N} 4^{-n} \log\left|\frac{g(2^nP)}{x(2^nP)^4}\right|_v + 4^{-N}\rho(2^NP)$$

[2] However, for theoretical purposes, note that Tate's series can be independently proven to converge and satisfy the characteristic properties of $\lambda_v$ (see [12]) thus providing an existence proof for Theorem 3.1 in all cases except $K_v = \mathbb{C}$.

The plan now is to take $N \to \infty$ and hope the series converges to $\rho(P)$, which we will justify in Theorem 5.1. However, for this series to be of any computational use, we must understand how one can efficiently compute its coefficients. In particular, we want to compute the sequence:

$$z(2^n P) = \frac{g(2^n P)}{x(2^n P)^4}$$

The trick here is to use $t = 1/x$ as a parameter, which is bounded on $E(K_v)$ by the assumption that $x(P) \neq 0$ for all $P \in E(K_v)$. Thus, we may compute the sequence $t(2^n P)$ without being concerned that the terms will grow 'too large'. Defining $z = t^4 g(1/t)$ as above, and $w = t^4 f(1/t)$, the duplication formula for $x(2P)$ yields

$$t(2P) = \frac{w(P)}{z(P)}$$

which can then be applied repeatedly to efficiently compute the sequence $t(2^n P)$, and thus obtain $z(2^n P)$.

**Theorem 5.1** (Tate's Series)**.** Let $E/K$ be an elliptic curve, $v \in M_K$, and suppose that $x(P) \neq 0$ for all $P \in E(K_v)$. Then, with $z$ defined as above, $\lambda_v$ is given by the series:

$$\lambda_v(P) = \log|x(P)|_v + \frac{1}{4}\sum_{n=0}^{\infty} 4^{-n} \log|z(2^n P)|_v$$

Furthermore, the error if one restricts the sum to $n \leq N$ is $\mathcal{O}(4^{-N})$ as $N \to \infty$.

*Proof.* This follows from the derivation above, provided we can justify the sequences $\log|z(2^n P)|_v$ and $\rho(2^n P)$ are bounded. Since $\log|z(P)|_v = \rho(P) - \frac{1}{4}\rho(2P)$, it will suffice to show the latter is bounded. Indeed, since $\lim_{P \to O_E} \rho(P) = 0$ exists, $\rho$ is bounded in a neighbourhood of $O_E$. Then, since $\lambda_v$ and $\log|x|_v$ are bounded away from $O_E$, we conclude $\rho$ is bounded everywhere. □

In particular, the condition $x(P) \neq 0$ implies that $\exists\, \varepsilon > 0$ such that $|x(2^n P)|_v > \varepsilon$ for all $n$, and the big-$\mathcal{O}$ constant grows with $\log 1/\varepsilon$ as $\varepsilon \to 0$. Thus, Tate's series converges poorly when the $x$-coordinates $x(2^n P)$ are 'small'. It is this observation, along with the need to compute $\lambda_v$ when $K_v = \mathbb{C}$, that motivates Silverman's refinement of Tate's series.

### 5.1.2 Silverman's 'Switching' Series

In this section, we derive Silverman's refinement of Tate's series, to compute the modified local height $\lambda_v = \hat{\lambda}_v + \frac{1}{6}\log|\Delta|_v$. Continuing with the notation of the previous section, the problems with Tate's series arise when the $x$-coordinates $x(2^n P)$ are 'small', and so Silverman's idea is to avoid this by dynamically switching between two Weierstrass equations for $E$.

Silverman's method begins by computing the coefficients of Tate's series as usual, until we hit a point where $x(2^n P)$ is 'small', say $t(2^n P) > 2$. Then, we make the substitution $x' = x + 1$, and switch to computing the Tate series coefficients on the new curve $E'$. If at any point $t'(2^n P) > 2$, then we switch back to the original curve $E$.

Although $\lambda_v$ is invariant under shifts of $x$, *a priori* there is no reason why intertwining two different series should help us to compute $\lambda_v$. However, by appropriately adjusting the $n$-th term each time we switch, we ensure resulting 'switching series' converges to $\lambda_v$.

**Derivation.** Let $E/K$ be the given elliptic curve, with Weierstrass coefficients $a_1, \ldots, a_6$. As in the derivation of Tate's series, we have:

$$\begin{aligned} t &= \frac{1}{x} \\ z &= t^4 g\left(\frac{1}{t}\right) = 1 - b_4 t^2 - 2b_6 t^3 - b_8 t^4 \\ w &= t^4 f\left(\frac{1}{t}\right) = 4t + b_2 t^2 + 2b_4 t^3 + b_6 t^4 \end{aligned}$$

Now, write $E'$ for the curve obtained after the substitution $x' = x + 1$. The new parameter $t' = 1/x'$ is related the original parameter $t$ by

$$t' = \frac{t}{1+t} \quad , \quad t = \frac{t'}{1-t'}$$

and the new quantities $b'_2, b'_4, b'_6, b'_8$ of $E'$ are easily computed from well-known explicit formulae in the $b_i$ (see [3], Chapter III, Table 1.2), thus we also obtain formulae for $w'$ and $z'$ in terms of $t'$.

To obtain Tate's series we derived a duplication formula for $\rho = 4(\hat{\lambda}_v - \log|x|_v)$, then repeatedly substituted this into itself, and took the limit. For Silverman's series, when we 'switch' from $E$ to $E'$, we will use a 'mixed' duplication formulae, giving $\rho$ in terms of $\rho' = 4(\lambda_v - \log|x'|_v)$. To derive this, first note $\lambda_v$ is invariant under translations, so

$$\frac{1}{4}(\rho - \rho') = \log\Big|\frac{x}{x'}\Big|_v \implies \rho = \rho' + 4\log|1+t|_v$$

Thus, our original duplication formula for $\rho$ becomes

$$\begin{aligned}\rho(P) &= \log|z(P)|_v + \frac{1}{4}\rho(2P) \\ &= \log|z(P)|_v + \frac{1}{4}\left(\rho'(2P) + 4\log|1+t(2P)|_v\right) \\ &= \log|z(P) + w(P)| + \frac{1}{4}\rho'(2P)\end{aligned}$$

Similarly, we obtain a mixed formula for switching from $E'$ to $E$,

$$\rho'(P) = \log|z'(P) - w'(P)|_v + \frac{1}{4}\rho(2P)$$

Now, Silverman's series is simple to describe. Assuming initially $x(P) > 1/2$, we write $\lambda_v = \frac{1}{2}\log|x(P)|_v + \frac{1}{8}\rho(P)$. We repeatedly expand $\rho(P)$ using the normal duplication formula, until we see that $t(2^{n+1}P) > 2$. Now, we use the mixed duplication formula to write $\rho(2^nP)$ in terms of $\rho'(2^{n+1}P)$. We then continue expanding $\rho'$ with its normal duplication formula, until such a time that we need to switch back to $\rho$. After $N$ steps, we have derived a series:

$$\lambda_v(P) = \log|x(P)|_v + \frac{1}{4}\sum_{n=0}^{N-1} 4^{-n}c_n + 4^{-N}d_N$$

where the $c_n$ are easily computed quantites in terms of $z, z', w, w'$, and $d_N$ is either $\rho(2^NP)$ or $\rho'(2^NP)$. This method is summarised in Theorem 5.2, where we also justify convergence.

**Theorem 5.2** (Silverman's Switching Series)**.** Let $E/K$ be an elliptic curve, $v \in M_K$, and $P \in E(K_v)$. Let $t, t', w, w', z, z'$ all be defined as in the derivation above.

Define a sequence of boolean values $\beta_{-1}, \beta_0$ as follows:

$$\beta_{-1} = \begin{cases} 1 & \text{if } |t(P)|_v \le 2 \\ 0 & \text{if } |t(P)|_v > 2 \end{cases}$$

$$\beta_n = \begin{cases} 1 & \text{if } \beta_{n-1} = 1 \text{ and } |t(2^{n+1}P)|_v \le 2 \\ 0 & \text{if } \beta_{n-1} = 1 \text{ and } |t'(2^{n+1}P)|_v > 2 \\ 0 & \text{if } \beta_{n-1} = 0 \text{ and } |t'(2^{n+1}P)|_v \le 2 \\ 1 & \text{if } \beta_{n-1} = 0 \text{ and } |t'(2^{n+1}P)|_v > 2 \end{cases}$$

(Then $\beta_n$ tells us whether $c_n$ should be computed by substituting for $\rho$, or for $\rho'$.) Define the

coefficients $c_{-1}, c_0, \ldots$ as follows:

$$c_{-1} = \begin{cases} -\log|t(P)|_v & \text{if } \beta_{-1} = 1 \\ -\log|t'(P)|_v & \text{if } \beta_{-1} = 0 \end{cases}$$

$$c_n = \begin{cases} \log|z(2^n P)|_v & \text{if } \beta_{n-1} = 1, \beta_n = 1 \\ \log|z(2^n P) + w(2^n P)|_v & \text{if } \beta_{n-1} = 1, \beta_n = 0 \\ \log|z'(2^n P)|_v & \text{if } \beta_{n-1} = 0, \beta_n = 0 \\ \log|z'(2^n P) - w'(2^n P)|_v & \text{if } \beta_{n-1} = 0, \beta_n = 1 \end{cases}$$

(Where cases 1 and 3 come from the usual duplication formulae, and cases 2 and 4 from the mixed duplication formulae.)

Then, the local height $\lambda_v(P)$ is given by the series:

$$\lambda_v(P) = c_{-1} + \frac{1}{4} \sum_{n=0}^{\infty} 4^{-n} c_n$$

and furthermore, the error in taking only $N$ terms is $\mathcal{O}(4^{-N})$ as $N \to \infty$.

*Proof Sketch.* (see [8], Theorem 2.2 for details) Following the derivation previously, the error after taking only $N$ terms is:

$$\begin{cases} 4^{-N} \rho(2^N P) & \text{if } \beta_N = 1 \\ 4^{-N} \rho'(2^N P) & \text{if } \beta_N = 0 \end{cases}$$

Thus, it suffices to argue that:

$$\rho \text{ is bounded on } \{2^N P \ : \ \beta_N = 1\}$$
$$\rho' \text{ is bounded on } \{2^N P \ : \ \beta_N = 0\}$$

(The $c_n$ will then also be bounded, since they come from formulae in $\rho$ and $\rho'$ evaluated at such points.) Now, simply note that:

$$\beta_N = 1 \implies |x(2^N P)|_v > \frac{1}{2}$$
$$\beta_N = 0 \implies |x(2^N P)|_v > \frac{1}{2}$$

so the result follows by showing $\rho$, and $\rho'$ are bounded on sets with no small $x$-coordinates, as we did in the proof of Tate's series, Theorem 5.1. □

To use the series in practice, we would like an explicit estimate for the $\mathcal{O}(4^{-N})$ term, which is given by Silverman in [8, §4].

The merit of the Silverman series for $\lambda_v$ comes from its simplicity, good convergence, and ease of implementation. However, one can also use the explicit formulae in terms of elliptic functions; see [10, Algorithm 7.5.7].

## 5.2 Nonarchimedean Contribution

### 5.2.1 Local Errors without Minimality

To obtain a factorisation-free algorithm for computing the nonarchimedean heights, following Muller and Stoll [13], we will now devise an algorithm for computing local errors *without* requiring a minimal model for $E$ and determining its reduction type; instead we use the fact $\mu(P) = \Psi_v \cdot (n_v)/(\log \# k_v)$ is always rational and bound its denominator. Then, if we know it lies in a sufficiently small interval, we can deduce its value.

Recall the decomposition of the local error,

$$\Psi_v(P) = -\sum_{n=0}^{\infty} \frac{1}{4^{n+1}} \log \Phi_v(2^n P)$$

To clean up normalisations, we introduce the new functions $E(K) \to \mathbb{R}$:

$$\mu(P) = \Psi_v(P) \cdot \frac{n_v}{\log \# K_v}$$

$$\varepsilon(P) = -\log \Phi_v(P) \cdot \frac{n_v}{\log \# k_v} = \min\{v(f(P)), v(g(P))\} - 4\min\{v(x(P)), 0\}$$

and note the corresponding decomposition:

$$\mu(P) = \sum_{n=0}^{\infty} \frac{1}{4^{n+1}} \varepsilon(2^n P)$$

The functions $\varepsilon$ and $\mu$ will be extremely important for the remainder of the essay, so let us acquaint ourselves with their properties.

**Proposition 5.3.** For all $P \in E(K)$,

(i) $\mu(P) \in \mathbb{Q}$ with denominator at most $v(\Delta)$,

(ii) $0 \leq \mu(P) \leq \frac{1}{4} v(\Delta)$,

(iii) $0 \leq \varepsilon(P) \leq v(\Delta)$,

*Proof.*

(i) From the proof of Corollary 4.1, we see that if $E$ is not minimal at $v$, this changes $\mu$ by adding some integer. Thus, we may simply prove the case where $E$ is minimal. Then by Theorem 3.6, $\mu(P)$ is zero when $P$ has non-singular reduction, and some positive rational otherwise. For the denominator condition, we follow Tate's algorithm to obtain a lower bound for $v(\Delta)$ at each reduction type where it is possible that $c_v > 1$. The result then follows from direct comparison with the possible values of $\mu$, as shown in Table 3.

(ii) Note we can no longer assume $E$ is minimal at $v$, but from Corollary 4.1 we have the bounds:

$$0 \leq \mu(P) \leq \alpha_v + \frac{1}{6}\left(v(\Delta) - v(\Delta_v^{\min})\right)$$

and it again remains to check the result case-by-case. For example, in the case $\mathrm{I}_m$,

$$\alpha_v \leq \frac{m}{4}, \quad v(\Delta_v^{\min}) = m \leq v(\Delta)$$
$$\implies \mu(P) \leq \frac{m}{12} + \frac{1}{6} v(\Delta) \leq \frac{1}{4} v(\Delta)$$

(iii) The fact $\varepsilon$ is non-negative follows from the explicit formula:

$$\varepsilon(P) = \min\{v(g(P)), v(f(P))\} - 4\min\{v(x(P)), 0\}$$

and considering the cases $v(x(P)) < 0$ and $v(x(P)) \geq 0$ separately. The upper bound follows from part (ii) and the relation:

$$\varepsilon(P) = 4\mu(P) = \mu(2P)$$

□

**Table 3:** Values of $v(\Delta)$ and $\mu(P)$, for a minimal Weierstrass equation $E$ with prescribed reduction type.

| Reduction Type | $v(\Delta)$ | Possible nonzero values for $\mu(P)$ |
|---|---|---|
| $\mathrm{I}_m$ | $= m$ | $n(m-n)/n$, $0 < n < m$ |
| III | $\geq 3$ | $1/2$ |
| IV | $\geq 4$ | $2/3$ |
| $\mathrm{I}_m^*$ | $\geq m+6$ | $1$, $(m+4)/4$ |
| $\mathrm{IV}^*$ | $\geq 8$ | $4/3$ |
| $\mathrm{III}^*$ | $\geq 9$ | $3/2$ |

Furthermore, we show that, even though we have dropped the assumption $E$ is minimal, we still have a zero contribution for points with nonsingular reduction:

**Proposition 5.4.** The following are equivalent:

(i) $P$ has nonsingular reduction (here, we mean the *naive* reduction of the possibly non-minimal, but integral, Weierstrass equation)

(ii) $\mu(P) = 0$

(iii) $\varepsilon(P) = 0$

*Proof.* (i) $\iff$ (iii): For illustration we assume $\operatorname{char} k_v \neq 2$. Let $P = (x, y) \in E(K_v)$. In the case $v(x(P)) < 0$ (and thus nonsingular reduction) we see $\varepsilon(P) = 0$, so it remains to prove the result for $v(x(P)) \geq 0$.

Indeed $P = (x, y) \in E(K_v)$ has singular reduction if and only if the polynomials $f(x)$ and $h(x) = 6x^2 + b_2x + b_4$ both vanish mod $\pi$. Using the identity:

$$(h(x))^2 = 4g(x) + (8x + b_2)f(x)$$

we see the vanishing of $f$ and $h$ mod $\pi$ is equivalent to the vanishing of $f$ and $g$ mod $\pi$, which is equivalent to $\varepsilon(P) > 0$.

(ii) $\implies$ (iii): Follows from the decomposition $\mu(P) = \sum_{n=0}^{\infty} 4^{-n-1}\varepsilon(2^nP)$ and the fact that $\varepsilon$ is nonnegative.

(iii) $\implies$ (ii): If $\varepsilon(P) = 0$, then $P$ has nonsingular reduction. Since $E_0(K)$ is a group (even under naive reduction), then each multiple $2^nP$ has nonsingular reduction, so all $\varepsilon(2^nP) = 0$, and hence $\mu(P) = 0$. □

With these properties in hand, we are ready to describe the algorithm for computing $\mu(P)$.

**Method 5.5.** To compute $\mu(P)$,

1. If $v(\Delta) \leq 1$, then $E$ is minimal at $v$ with Tamagawa number $c_v = 1$, and so $\mu(P) = 0$.
2. Set $B = v(\Delta)$. By Proposition 5.3, this is an upper bound both for $\varepsilon(P)$ and the denominator of $\mu(P)$.
3. Else, set $m = \lfloor \frac{\log B^3/3}{\log 4} \rfloor$ and compute the sum:

$$\mu_0 = \sum_{n=0}^{m} \frac{1}{4^{n+1}}\varepsilon(2^nP)$$

Then, by the decomposition of $\mu$,

$$\mu_0 \leq \mu(P) \leq \mu_0 + \sum_{n>m} \frac{1}{4^{n+1}}B = \mu_0 + \frac{B}{3 \cdot 4^{m+1}} \leq \mu_0 + \frac{1}{B^2}$$

Thus, $\mu(P)$ is the unique fraction with denominator $\leq B$ in the interval

$$[\mu_0\,,\ \mu_0 + 1/B^2]$$

**Efficient Calculation of $\mu_0$.** To implement this method in practice, we describe an efficient subroutine to compute the terms $\varepsilon(2^n P)$ in the sum for $\mu_0$.

(i) Choose *primitive* projective coordinates $(x_1 : x_2)$ for $x(P) \in \mathbb{P}^1(K_v)$. That is to say, scale the coordinates such that $\min\{v(x_1), v(x_2)\} = 0$. Ensure that we have a precision of at least $m(B+1)+1$ $v$-adic digits.

(ii) Let $f^*$ and $g^*$ be the degree 4 homogenisations of the duplication polynomials $f, g$. Compute the values of $f^*(x_1, x_2)$ and $g^*(x_1, x_2)$ to $B(m+1)+1$ $v$-adic digits.

(iii) Since we used primitive coordinates, then $\varepsilon(P)$ is easily computed as:

$$\varepsilon(P) = \min\{v(g^*(x_1, x_2)), v(f^*(x_1, x_2))\}$$

(iv) Furthermore, by the duplication formula, we now have a pair of primitive coordinates for $x(2P)$,

$$x(2P) = (\pi^{-\varepsilon(P)} g^*(x_1, x_2) : \pi^{-\varepsilon(P)} f^*(x_1, x_2))$$

with at least $(m+1)B + 1 - \varepsilon(P) \leq mB + 1$ digits of $v$-adic precision.

(v) Iterate the algorithm, to compute $\varepsilon(P)$, $\varepsilon(2P)$, $\ldots$, $\varepsilon(2^m P)$. At each stage, our primitive coordinates for $x(2^i P)$ will have at most $(m+1-i)B + 1 \geq 1$ digits of $v$-adic precision, so the values of $\varepsilon(2^n P)$ will all be correct.

### 5.2.2 Müller-Stoll Algorithm

For simplicity we now suppose $K = \mathbb{Q}$. Fix some $P \in E(\mathbb{Q})$, for which we want to compute the canonical height. Since the naive height $h(P)$ is easy to compute, we will take advantage of the decomposition:

$$\hat{h}(P) = h(P) - \Psi_\infty(P) - \sum_p \mu_p(P) \log p$$

Using any previously described method to compute the archimedean local height $\lambda_\infty(P)$, we then easily compute the local error $\Psi_\infty(P) = \log \max\{|x(P)|_v, 1\}$. It remains to describe an efficient method to compute the sum $\sum_p \mu_p(P) \log p$.

Although Method 5.5 gives an efficient algorithm for calculating the individual terms $\mu_p(P)$, it requires the knowledge of $v_p(\Delta)$ at each $p$, and thus we would have still have to factorise $\Delta$. However, the idea employed by Müller and Stoll [13] is to see whether we can somehow *simultaneously* run Method 5.5 on multiple primes at once and circumvent the factorisation.

To avoid prime factorisation, we will use the weaker notion of *factorisation into coprimes.*

**Definition** (Coprime Base)**.** For a list of positive integers $a_1, \ldots, a_s$, a *coprime base* is a list of positive integers $(q_1, \ldots, q_r)$ which are *pairwise coprime*, and such that each $a_j$ is some product of their powers:

$$a_j = \prod_{i=1}^{s} {q_i}^{e_{i,j}}$$

Consider for each point $2^n P$ a pair of integral primitive coordinates:

$$x(2^n P) = \left(x_1^{(n)} : x_2^{(n)}\right) \in \mathbb{P}^1(\mathbb{Q})\,, \quad \gcd(x_1^{(n)}, x_2^{(n)}) = 1$$

and define the quantity:

$$g_n = \gcd\left(g^*(x_1^{(n)}, x_2^{(n)})\,,\ f^*(x_1^{(n)}, x_2^{(n)})\right)$$

Then, recalling the remark after Method 5.5, we have $v_p(g_n) = \varepsilon(2^n P)$ and so:

$$\begin{aligned}\sum_p \mu_p(P)\log p &= \sum_p \sum_{n=0}^{\infty} \frac{1}{4^{n+1}} \varepsilon_p(2^n P)\log p \\ &= \sum_{n=0}^{\infty} \frac{1}{4^{n+1}} \sum_p \min\left\{ v_p\left(g^*(x_1^{(n)}, x_2^{(n)}\right)\ ,\ v_p\left(f^*(x_1^{(n)}, x_2^{(n)})\right)\right\}\log p \\ &= \sum_{n=0}^{\infty} \frac{1}{4^{n+1}} \log g_n\end{aligned}$$

and furthermore, since $\varepsilon_p(2^n P) \leq v_p(\Delta)$ for all $p$ (Proposition 5.3), we have that all $g_n \mid \Delta$.

The core idea is as follows: first, we compute terms of the series up to some $m$ sufficiently large. This truncated series can be written as some rational linear combination of logarithms of primes. Now, analogously to Method 5.5, $\mu_p(P)$ can be identified as the unique rational number of bounded denominator which is sufficiently close to the coefficient of $\log p$. Note the coefficient of $\log p$ comes from the terms where $p \mid g_n$. Thus, the value of $\mu_p(P)$ is the same for any two primes $p, q$ which always appear together in the factorisation of the $g_n$, for $n \leq m$. So, taking a *coprime base* $(q_1, \ldots, q_r)$ for the integers $(g_0, \ldots, g_m)$, then for each $i$ we actually simultaneously compute the value of $\mu_p(P)$ at all primes $p \mid q_i$.

We now give a descripton of the algorithm, prove its validity, and assess its runtime.

**Method 5.6** (Müller-Stoll Algorithm)**.** To compute the total nonarchimedean local error at $P \in E(\mathbb{Q})$,

1. Let $(x_1 : x_2)$ be primitive coordinates for $x(P)$.
2. Compute $(x_1^{(1)}, x_2^{(1)}) = (g^*(x_1, x_2), f^*(x_1, x_2))$ and set $g_0 = \gcd(x_1^{(1)}, x_2^{(1)})$.
3. Set $D = \max\{\gcd(\Delta, g_0^n) : n \in \mathbb{N}\} = \gcd(\Delta, g_0^{\infty})$, the largest divisor of $\Delta$ built from primes dividing $g_0$.
4. Set $B = \lfloor \log D / \log 2 \rfloor$. If $B \leq 1$, then return 0. Otherwise, set:
$$m = \left\lfloor \frac{\log(B^5/3)}{\log 4} \right\rfloor$$
5. For each $0 \leq n \leq m$,
   (i) Compute $g_n = \gcd\left(g^*(x_1^{(n)}, x_2^{(n)})\,,\ f^*(x_1^{(n)}, x_2^{(n)})\right)$ modulo $D^{m+1} g_0$
   (ii) Set $(x_1^{(n+1)}, x_2^{(n+1)}) = \left(\dfrac{g^*(x_1, x_2)}{g_n} : \dfrac{f^*(x_1, x_2)}{g_n}\right)$
6. Compute a coprime base $(q_1, \ldots, q_r)$ for the integers $(g_0, \ldots, g_m)$ such that
$$g_n = \prod_{i=1}^{r} q_i^{e_{i,n}}$$
7. For each $1 \leq i \leq r$,
   (i) Compute $a = \sum_{n=0}^{m} 4^{-n-1} e_{i,n}$
   (ii) Let $\mu_i$ be the unique fraction with denominator at most $B^2$ in the interval
$$[a, a + 1/B^4]$$

8. The total nonarchimedean local error, is:

$$\sum_{i=1}^{r} \mu_i \log q_i$$

This is an exact expression, given as a formal sum of logarithms.

There are several small modifications which can be made to improve runtime, see [13], but the most important detail is to only compute the coordinates $x_1^{(n)}$, $x_2^{(n)}$ in Step 5 modulo $D^{m+1}g_0$ (to see this has sufficient $p$-adic accuracy for each $p \mid g_0$, recall the discussion after Method 5.5). This avoids working with impractically large $x$-coordinates, one of the main reasons the limit definition of $\hat{h}$ was not viable for computation in the first place.

**Validity.** First, we eliminate the trivial case, when $B \leq 1$, so $D \in \{1, 2, 3\}$.

- If $D = 1$, recall $g_0 \mid \Delta$ and so this forces $g_0 = 1$. Then for each prime $\varepsilon_p(P) = 0$, so $\varepsilon_p(2^i P) = 0$ for all $i$, and thus $\mu_p(P) = 0$. Hence the output of zero is correct.
- If $D = 2$, then $g_0$ is some power of 2 and $v_2(\Delta) = 1$, but now our upper bound on $\varepsilon$ (Proposition 5.3) gives $\varepsilon_2(P) \leq 1$, so $g_0 = 1$ and we have a contradiction.
- If $D = 3$, we reach a contradiction analogously.

From now on, suppose $B > 1$, and let $p$ be a prime. If $p \nmid g_0$, then $\varepsilon_p(P) = 0$, and thus $\mu_p(P) = 0$. On the other hand, we claim that for each prime $p$ dividing $g_0$,

$$\mu_i = \frac{\mu_p(P)}{v_p(q_i)}$$

where $i = i(p)$ is the unique index such that $p \mid q_i$. If the claim is true, then:

$$\sum_p \mu_p(P) \log p = \sum_{i=1}^{r} \sum_{p \mid q_i} \mu_i \cdot v_p(q_i) \cdot \log p = \sum_{i=1}^{r} \mu_i \log q_i$$

and thus the algorithm is valid. Now, proving the claim is analogous to justifying Method 5.5. We have that $B = \lfloor \log D / \log 2 \rfloor \geq v_p(\Delta)$ is an upper bound for both $\varepsilon_p(P)$ and the denominator of $\mu_p(P)$. Also, we know $v_p(q_i) \leq \lfloor \log q_i / \log 2 \rfloor \leq \lfloor \log D / \log 2 \rfloor = B$. Thus, the denominator of $\mu_p(P)/v_p(q_i)$ is at most $B^2$. Finally, we show it lies in the right interval, and hence must be equal to $\mu_i$. Recalling that $\varepsilon(2^n P) = v_p(g_n)$,

$$\frac{\mu_p(P)}{v_p(q_i)} = \sum_{n=0}^{\infty} \frac{1}{4^{n+1}} \frac{\varepsilon(2^n P)}{v_p(q_i)} = \sum_{n=0}^{\infty} \frac{1}{4^{n+1}} e_{i,n}$$

$$\implies a \leq \frac{\mu_p(P)}{v_p(q_i)} \leq a + \sum_{n=m+1}^{\infty} \frac{1}{4^{n+1}} \varepsilon(2^n P) \leq a + \frac{B}{3 \cdot 4^{m+1}} \leq a + \frac{1}{B^4}$$

**Runtime.** A detailed complexity analysis is given in the original paper [13], which we discuss here. We focus on Step 5 and Step 6, since the remainder of the steps are negligible in comparison. Write $\mathbf{M}(d)$ for the time taken to perform elementary operations on $d$-bit integers; using fast multiplication algorithms, $\mathbf{M}(d)$ is *quasi-linear*, i.e. has complexity

$$\mathcal{O}(d(\log d)^k), \quad k \in \mathbb{N}$$

and furthermore, the GCD of two $d$-bit integers can be computed in time

$$\mathcal{O}(\mathbf{M}(d) \log d).$$

Thus Step 5, which $m$ times computes the gcd of $(\log_2 D^{m+1} g_0)$-digit integers, can be done in quasi-linear time:

$$\mathcal{O}\left(\mathbf{M}((\log \Delta)(\log\log \Delta)) \cdot (\log\log \Delta)^2\right)$$

where we used that $m = \mathcal{O}(\log\log\Delta)$ and $D = \mathcal{O}(\Delta)$.

In the original paper, the authors argue that Step 6 also takes only quasi-linear time, but this argument hinges on a preprint of Bernstein [14] claiming an algorithm to find a coprime base in effectively linear time

$$b(\log b)^{4+o(1)}$$

where $b$ is the number of bits in the input. In practice, a simpler algorithm of Buchmann and Lenstra [15, Proposition 6.5] is used.

Nevertheless, the algorithm is extremely efficient; impressively capable of computing heights on curves with extremely large coefficients almost instantaneously, where other algorithms fail to terminate in a practical timeframe [13, §7].

# 6 Conclusion

For our two goals, bounding the height difference $h - \hat{h}$ and computing the canonical height $\hat{h}$, we have demonstrated algorithms which are currently viable for practical purposes. But, as the computational study of elliptic curves progresses to curves with larger and larger coefficients, we will need new ideas to improve efficiency, or new algorithms altogether. For the archimedean height difference bounds, Müller and Stumpe provide a new method in [16], and suggest that it should be combined with the methods described in this essay; for the archimedean local height, an unpublished manuscript of Bost & Mestre gives another efficient method with quadratic convergence. Additionally, extensions of the algorithms here to Jacobians of genus-two curves are discussed in [17].

# A Appendix

## A.1 Comparison of Normalisations

Here we compare some of the normalisations of heights, across the main sources of this essay. We write $\hat{h}$, $\hat{\lambda}_v$, $\lambda_v$ for our values of the canonical height, Néron local height, and modified local height. We will write, say, $\hat{h}^{(\text{Paper})}$ to mean the value of $\hat{h}$ as defined in the given paper.

- **Silverman's Paper** [8] uses a canonical height which is half the value of ours, and by the 'local height' refers to half of our modified local height:

$$\hat{h} = 2\hat{h}^{(\text{SilP})}\ , \quad \lambda_v = 2\lambda_v^{(\text{SilP})}$$

- **Silverman's Books** [3] [5] also halves the canonical height, but instead uses a local height which is half our Néron local height,

$$\hat{h} = 2\hat{h}^{(\text{SilB})}\ , \quad \hat{\lambda}_v = 2\lambda_v^{(\text{SilB})}$$

- **Cremona, Pickett & Siksek** [9] use our canonical height and modified local height:

$$\hat{h} = \hat{h}^{(\text{CPS})}\ , \quad \lambda_v = \lambda_v^{(\text{CPS})}$$

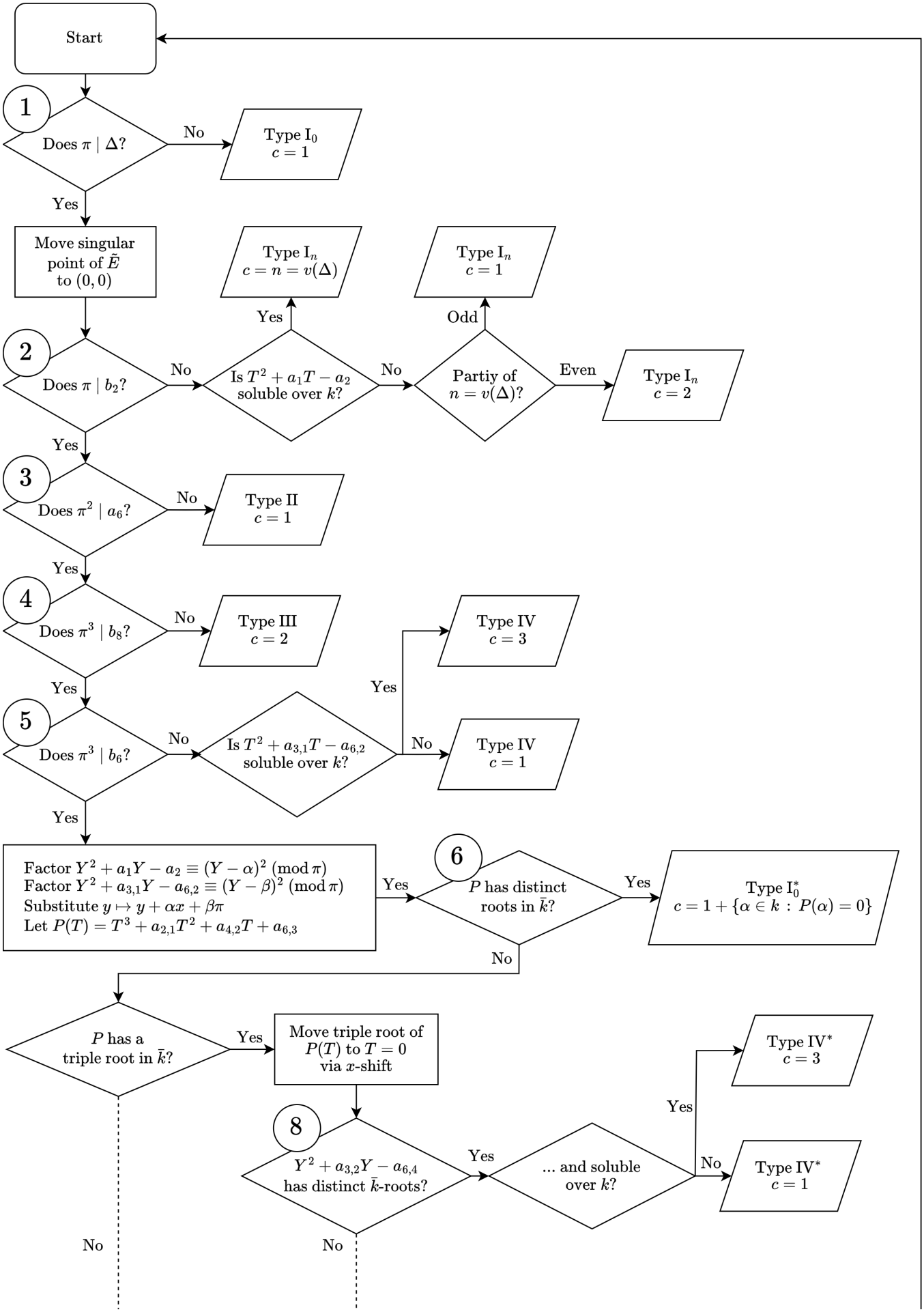
- **Bruin** [11] uses the same canonical height, but a local height which is half the Néron local height,

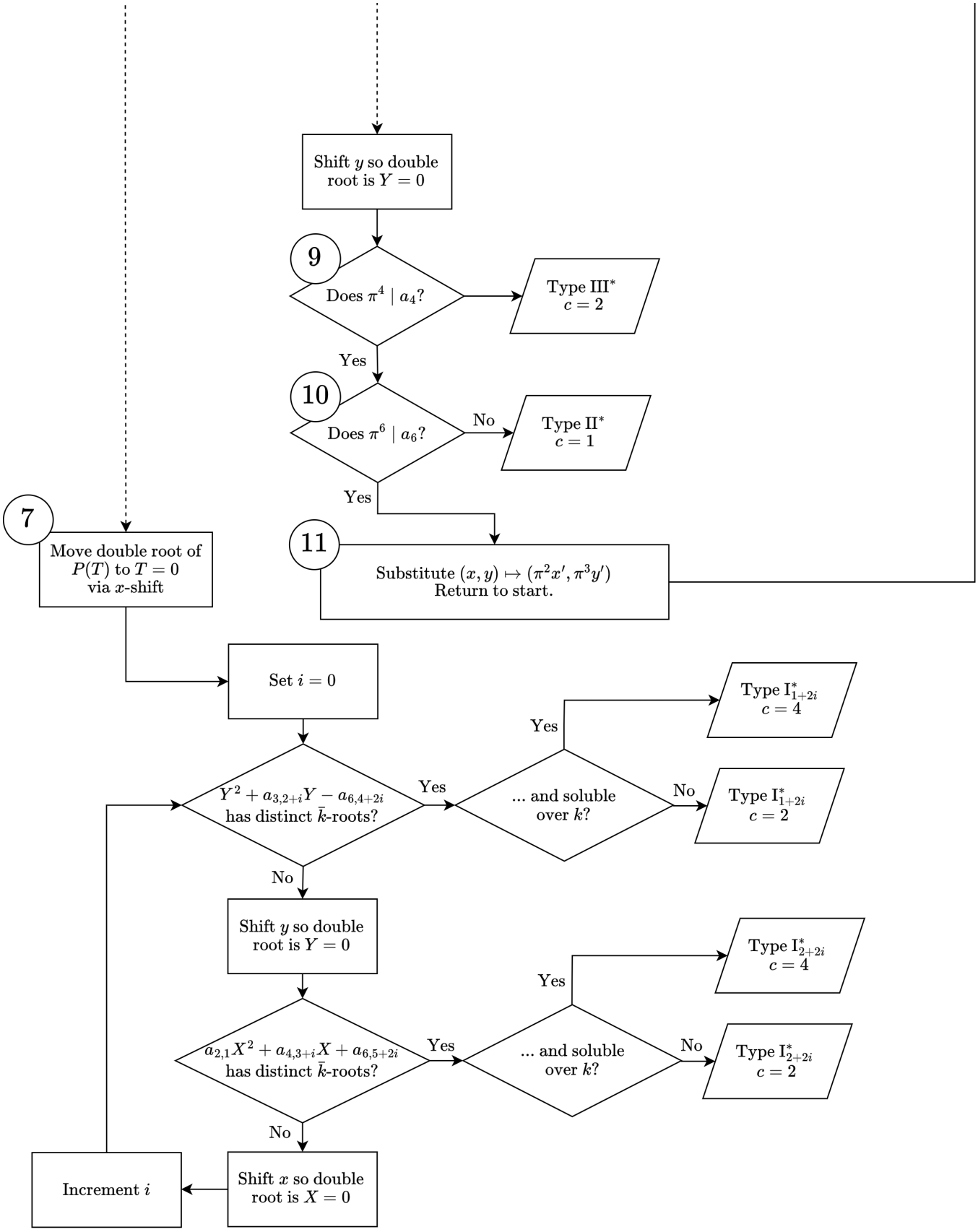
$$\hat{h} = \hat{h}^{(\text{Bruin})}\ , \quad \hat{\lambda}_v = 2\lambda_v^{(\text{Bruin})}$$

## A.2 Tate's Algorithm

We include in Figure 1 a flowchart of Tate's algorithm, with steps numbered as in Silverman's book [5, IV.9.4].

**Figure 1:** Flowchart summarising Tate's Algorithm.




Shift $y$ so double root is $Y = 0$
9
Does $\pi^4 \mid a_4$?
Type III$^*$ $c = 2$
Yes
10
Does $\pi^6 \mid a_6$?
No
Type II$^*$ $c = 1$
Yes
7
Move double root of $P(T)$ to $T = 0$ via $x$-shift
11
Substitute $(x, y) \mapsto (\pi^2 x', \pi^3 y')$ Return to start.
Set $i = 0$
$Y^2 + a_{3,2+i}Y - a_{6,4+2i}$ has distinct $\bar{k}$-roots?
Yes
... and soluble over $k$?
Yes
Type I$^*_{1+2i}$ $c = 4$
No
Type I$^*_{1+2i}$ $c = 2$
No
Shift $y$ so double root is $Y = 0$
$a_{2,1}X^2 + a_{4,3+i}X + a_{6,5+2i}$ has distinct $\bar{k}$-roots?
Yes
... and soluble over $k$?
Yes
Type I$^*_{2+2i}$ $c = 4$
No
Type I$^*_{2+2i}$ $c = 2$
No
Shift $x$ so double root is $X = 0$
Increment $i$

# References


[1] The LMFDB Collaboration. The L-functions and modular forms database. `https://www.lmfdb.org`, 2024. [Online; accessed April 2024].

[2] S. Siksek. Infinite descent on elliptic curves. *The Rocky Mountain Journal of Mathematics*, 25(4):1501–1538, 1995.

[3] J.H. Silverman. *The Arithmetic of Elliptic Curves*. Graduate Texts in Mathematics. Springer New York, 2009.

[4] S. Lang. *Elliptic Curves: Diophantine Analysis*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2013.

[5] J.H. Silverman. *Advanced Topics in the Arithmetic of Elliptic Curves*. Graduate texts in mathematics. Springer-Verlag, 1994.

[6] J. Tate. Algorithm for determining the type of a singular fiber in an elliptic pencil. In *Modular Functions of One Variable IV*, pages 33–52, Berlin, Heidelberg, 1975. Springer Berlin Heidelberg.

[7] J. Cremona. Computing in component groups of elliptic curves. In *Algorithmic Number Theory*, pages 118–124, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.

[8] J.H. Silverman. Computing heights on elliptic curves. *Mathematics of Computation*, 51(183):339–358, 1988.

[9] J. Cremona, M. Prickett, and S. Siksek. Height difference bounds for elliptic curves over number fields. *Journal of Number Theory*, 116:42–68, 01 2006.

[10] H. Cohen. *A Course in Computational Algebraic Number Theory*. Graduate Texts in Mathematics. Springer Berlin Heidelberg, 2013.

[11] P. Bruin. Bornes optimales pour la différence entre la hauteur de weil et la hauteur de Néron–Tate sur les courbes elliptiques sur q. *Acta Arithmetica*, 160(4):385–397, 2013.

[12] J. Tate. An oft cited letter from Tate to Serre on computing local heights on elliptic curves, 2012.

[13] J.S. Müller and M. Stoll. Computing canonical heights on elliptic curves in quasi-linear time. *LMS Journal of Computation and Mathematics*, 19(A):391–405, 2016.

[14] D.J. Bernstein. Research announcement: Faster factorisation into coprimes. *Preprint*, 2004.

[15] J.A. Buchmann and H.W. Lenstra. Approximatting rings of integers in number fields. *Journal de théorie des nombres de Bordeaux*, 6(2):221–260, 1994.

[16] J.S. Müller and C. Stumpe. Archimedean local height differences on elliptic curves. *Acta Arithmetica*, 190(3):293–303, July 2019.

[17] J.S. M̈ller and M. Stoll. Canonical heights on genus-2 jacobians. *Algebra and Number Theory*, 10:2153–2234, 12 2016.