Height Difference Bounds For Elliptic Curves over Number Fields

J. E. Cremona

School of Mathematical Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, UK.

M. Prickett [1]

School of Mathematical Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, UK.

Samir Siksek [2]

Department of Mathematics and Statistics, College of Science, Sultan Qaboos University, P.O. Box 36, Al-Khod 123, Oman

Abstract

Let E be an elliptic curve over a number field K. Let h be the logarithmic (or Weil) height on E and $\hat{h}$ be the canonical height on E. Bounds for the difference $h - \hat{h}$ are of tremendous theoretical and practical importance. It is possible to decompose $h - \hat{h}$ as a weighted sum of continuous bounded functions $\Psi_v : E(K_v) \to \mathbb{R}$ over the set of places v of K. A standard method for bounding $h - \hat{h}$, (due to Lang, and previously employed by Silverman) is to bound each function Ψ_v and sum these local 'contributions'.

In this paper we give simple formulae for the extreme values of Ψ_v for non-archimedean v in terms of the Tamagawa index and Kodaira symbol of the curve at v.

For real archimedean v a method for sharply bounding Ψ_v was previously given by Siksek (1990). We complement this by giving two methods for sharply bounding Ψ_v for complex archimedean v.

Key words: Elliptic curves, heights, canonical height, height bounds
1991 MSC: Primary 11G50, 11G05, Secondary 11G07, 14G05

1 Introduction

Let K be a number field and let E be an elliptic curve defined over K. The canonical height $\hat{h}$ is a quadratic form on $E(K) \otimes \mathbb{R}$ whose difference from the logarithmic height h is bounded on $E(K)$. It is of tremendous importance both to the theoretical and to the explicit study of elliptic curves to have sharp bounds for the difference $h - \hat{h}$, particularly a small upper bound for this quantity. For example, explicit bounds on $h - \hat{h}$ are essential for the effective proof of the Mordell-Weil Theorem. Good bounds for this difference are an important part of algorithms for determining Mordell-Weil bases of elliptic curves, and for determining integral points on elliptic curves.

Let M_K be the set of places of K. It is possible to decompose the difference $h - \hat{h}$ as

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

where the Ψ_v are continuous bounded functions $\Psi_v : E(K_v) \to \mathbb{R}$. For all but a finite set of places v on K, the functions Ψ_v vanish identically. A reasonable approach to bounding $h - \hat{h}$, suggested by Lang in [7], is to bound each of the functions Ψ_v separately, and then to sum all of these local 'contributions' to obtain a bound for $h - \hat{h}$. This approach is adopted in Silverman's paper [13], where he derives concise bounds for $h - \hat{h}$ in terms of the coefficients and invariants of the curve. This Lang-Silverman approach is more conceptually attractive, and gives more precise bounds, than the earlier approach found in the papers (say) of Zimmer [17], and Demjanenko [4]; this earlier approach is based on estimating the difference $4h(P) - h(2P)$ and the use of Tate's 'telescoping' series.

Given the importance of the problem, it is highly desirable to take the Lang-Silverman approach to its logical extreme: rather than asking for bounds for each function Ψ_v, one should ask if the extrema of these real-valued functions can be determined. This idea first appears in a paper of Siksek [12], where he gives an algorithm - albeit a tedious one - for computing the suprema of the functions Ψ_v for non-archimedean v. For archimedean v, the functions Ψ_v are substantially more complicated and it seems hopeless to determine their extrema. It is however possible to write $\Psi_v = -\sum_{i=0}^{\infty} 4^{-i-1} \log \Phi_v(2^i P)$ for some (simpler, though still complicated) real-valued function $\Phi_v : E(K_v) \to \mathbb{R}$. Determining the extrema of Φ_v gives sharp bounds for Ψ_v. Siksek does this for

Email addresses: John.Cremona@nottingham.ac.uk (J. E. Cremona), martin.prickett@maths.nottingham.ac.uk (M. Prickett), siksek@squ.edu.om (Samir Siksek).

[1] Funded from 2000 to 2003 by a CASE studentship from the UK EPSRC

[2] Funded by a grant from Sultan Qaboos University (IG/SCI/DOMS/02/06).

real archimedean v, but gives a non-rigorous numerical method for estimating the extrema for complex archimedean v.

In this paper, using an exhaustive analysis of possible reduction types of elliptic curves, we give simple formulae for the extreme values of Ψ_v for non-archimedean v. These formulae depend only on the Kodaira symbol and Tamagawa index of the curve at v. We complement Siksek's determination of the extrema of Φ_v for real archimedean v by determining the extrema of this function for complex archimedean places v: the locations of the extrema are given as simultaneous zeros of some pairs of real bivariate polynomials. Thus the extrema can be determined by solving these pairs of polynomials using Groebner bases. We also give a second, very fast algorithm, which is numerical but completely rigorous, for computing the extrema of Φ_v to arbitrary desired accuracy (hence bounding Ψ_v for complex v).

The resulting bound for $h - \hat{h}$ we obtain in this paper has all the virtues of the bound in [12], but none of the vices. To summarize, whilst the final bound for $h - \hat{h}$ we give is numerically the same as that in [12], it has the following advantages

- better suited for theoretical investigations,
- (unlike [12]) entirely rigorous for number fields with complex embeddings,
- and almost trivial to implement/compute.

At the end of the paper we compute some examples and carry out a comparison between our bounds and those of Silverman.

The reader is warned at the outset that several normalizations of canonical and local heights appear in the literature; we say more on this in due course.

We are indebted to Professor Silverman for clarifying our ideas on local height normalizations and for useful comments on a previous version of this paper (including some of the history behind Proposition 5), to Professor Buchberger for useful discussions on numerical Groebner bases, and to Dr. Albaali for suggesting to us that the method of Lagrange multipliers might be useful in the proof of Lemma 12.

2 Statement of the Main Theorem

We fix once and for all the following notation.

K	a number field,
$\mathcal{O}_K$	the ring of integers of K,
M_K	the set of all places of K,
M_K^0	the set of non-archimedean places of K,
M_K^∞	the set of archimedean places of K,
v	a place of K,
K_v	the completion of K at v,
n_v	the local degree $[K_v : \mathbb{Q}_v]$.

We will use the notation v interchangeably for a place and for the associated normalized valuation. The following notation is relevant to places $v \in M_K^0$.

k_v	the residue field at v,
$\mathcal{O}_v$	ring of integers in K_v,
q_v	the cardinality of the residue field k_v.

Let E be an elliptic curve given by the Weierstrass equation

$$E : \qquad y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6, \tag{2}$$

where $a_1, \ldots, a_6 \in \mathcal{O}_K$. For each place v we denote by $E_0(K_v)$ the connected component of the identity in $E(K_v)$; for $v \in M_K^0$ this consists of the points of $E(K_v)$ with good reduction at v. The Tamagawa index at v, which is 1 for almost all v including those where E has good reduction, is the index $c_v = [E(K_v) : E_0(K_v)]$. The finite quotient $E(K_v)/E_0(K_v)$ is called the component group of E at v.

We define the usual associated constants (see [14, page 46]) as follows.

$$\begin{aligned}
b_2 &= a_1^2 + 4a_2,\\
b_4 &= 2a_4 + a_1a_3,\\
b_6 &= a_3^2 + 4a_6,\\
b_8 &= a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2,\\
\Delta &= -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6.
\end{aligned}$$

Let

$$\begin{aligned}
f(P) &= 4x(P)^3 + b_2x(P)^2 + 2b_4x(P) + b_6,\\
g(P) &= x(P)^4 - b_4x(P)^2 - 2b_6x(P) - b_8;
\end{aligned} \tag{3}$$

Table 1
Values of α_v

Kodaira type of $E_v^{\min}$ at v	Tamagawa index c_v	α_v
any	1	0
$\mathrm{I_m}$, m even	2 or m	$m/4$
$\mathrm{I_m}$, m odd	m	$(m^2-1)/4m$
III	2	1/2
IV	3	2/3
I_0^*	2 or 4	1
I_m^*	2	1
I_m^*	4	$(m+4)/4$
IV^*	3	4/3
III^*	2	3/2

so that $x(2P) = g(P)/f(P)$. Define the function $\Phi_v : E(K_v) \to \mathbb{R}$ by

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

It is straightforward to see that Φ_v is a continuous and hence bounded function on $E(K_v)$ (the boundedness follows immediately from the fact that $E(K_v)$ is compact with respect to the v-adic topology). Define

$$\epsilon_v^{-1} = \inf_{P \in E(K_v)} \Phi_v(P), \qquad \delta_v^{-1} = \sup_{P \in E(K_v)} \Phi_v(P), \tag{5}$$

where the exponents -1 have been chosen to simplify the formulae appearing later. In [12, Lemma 2.3] it shown that ϵ_v exists (i.e. the infimum appearing in its definition is non-zero) and satisfies $\epsilon_v \geq 1$.

For each valuation $v \in M_K^0$ let $E_v^{\min}$ be a minimal model for E over K_v, and let $\Delta_v^{\min}$ be the discriminant of $E_v^{\min}$. Thus we can take $E_v^{\min} = E$ and $\Delta_v^{\min} = \Delta$ for almost all $v \in M_K^0$, and they are always equal if the model E is globally minimal. For $v \in M_K^0$ we define the constants α_v according to the Kodaira type of $E_v^{\min}$ and the Tamagawa index c_v as in Table 1. We can now state our main theorem:

Theorem 1 *For all $P \in E(K)$,*

$$\frac{1}{3[K:\mathbb{Q}]}\sum_{v\in M_K^\infty} n_v \log\delta_v \le h(P) - \hat{h}(P) \le \frac{1}{3[K:\mathbb{Q}]}\sum_{v\in M_K^\infty} n_v \log\epsilon_v$$
$$+\frac{1}{[K:\mathbb{Q}]}\sum_{v\in M_K^0}\left(\alpha_v + \frac{1}{6}\operatorname{ord}_v(\Delta/\Delta_v^{\min})\right)\log(q_v).$$

The next theorem is a by-product of the proof of Theorem 1. In essence it says that the bounds are sharper if we restrict ourselves to points that have everywhere good reduction. Although this result is less general than Theorem 1, we suspect it may be useful for some applications.

Theorem 2 *Suppose $P \in E(K)$. If $P \in E_0(K_v)$ for all non-archimedean valuations v then*

$$\frac{1}{3[K:\mathbb{Q}]}\sum_{v\in M_K^\infty} n_v \log\delta_v \le h(P) - \hat{h}(P) \le \frac{1}{3[K:\mathbb{Q}]}\sum_{v\in M_K^\infty} n_v \log\epsilon_v$$

3 Definitions of Heights and Local Heights

In this section we review the definitions of logarithmic and canonical heights, as well as their decompositions into local components. If $P \in E(K)$ then the naive heights of P and $2P$ are respectively given by

$$H_K(P) = \prod_{v\in M_K}\max\{1, |x(P)|_v\}^{n_v}, \quad H_K(2P) = \prod_{v\in M_K}\max\{|f(P)|_v, |g(P)|_v\}^{n_v},$$

where f, g are the polynomials defined in (3).

The logarithmic height is given by

$$h(P) = \frac{1}{[K:\mathbb{Q}]}\log H_K(P) = \frac{1}{[K:\mathbb{Q}]}\sum_{v\in M_K} n_v \log\max\{1, |x(P)|_v\}. \qquad (6)$$

It is then easy to see that for $P \in E(K)$ we have

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

where Φ_v is defined in (4). Using the usual 'telescoping' series we see that

$$\begin{aligned}\hat{h}(P) =& \lim_{i\to\infty}\frac{1}{4^i}h(2^iP))\\ =& h(P) + (\frac{1}{4}h(2P) - h(P)) + (\frac{1}{4^2}h(2^2P) - \frac{1}{4}h(2P)) + \ldots\\ =& \frac{1}{[K:\mathbb{Q}]}\sum_{v\in M_k} n_v\left(\log\max\{1, |x(P)|_v\} + \sum_{i=0}^{\infty}\frac{1}{4^{i+1}}\log\Phi_v(2^iP)\right).\end{aligned}$$

We define the local height $\lambda_v : E(K_v)\backslash\{O\} \to \mathbb{R}$ by

$$\lambda_v(P) = \log\max\{1, |x(P)|_v\} + \sum_{i=0}^{\infty}\frac{1}{4^{i+1}}\log\Phi_v(2^iP). \tag{7}$$

The canonical and local heights are then related by the formula,

$$\hat{h}(P) = \frac{1}{[K:\mathbb{Q}]}\sum_{v\in M_K} n_v\lambda_v(P). \tag{8}$$

Let $\Psi_v : E(K_v) \to \mathbb{R}$ be given by

$$\Psi_v(P) = \begin{cases} 0 & \text{if} \quad P = O,\\ \log\max\{1, |x(P)|_v\} - \lambda_v(P) & \text{otherwise.}\end{cases} \tag{9}$$

Combining (6) and (8) we deduce the validity of the decomposition of the height difference $h - \hat{h}$ in (1).

Although the following proposition is not used later on, it is helpful to bear in mind and does motivate our approach to bounding $h - \hat{h}$.

Proposition 3 *Suppose $v \in M_K$ (archimedean or non-archimedean). Ψ_v is continuous bounded function on $E(K_v)$. Moreover, if $v \in M_K^0$ and E has good reduction at v then Ψ_v vanishes identically on $E(K_v)$.*

Proof: By the explicit formula for the local height (7) the function Ψ_v can be rewritten as

$$\Psi_v(P) = -\sum_{i=0}^{\infty}\frac{1}{4^{i+1}}\log\Phi_v(2^iP). \tag{10}$$

We recall that $\log\Phi_v$ is continuous and bounded on $E(K_v)$. For this see the proof of [15, Lemma VI.1.2]. The continuity of Ψ_v follows by the Weierstrass M-test, and its boundedness from the compactness of $E(K_v)$.

Suppose that $v \in M_K^0$. If P has good reduction then its local height is simply given by $\lambda_v(P) = \log\max\{1, |x(P)|_v\}$ (see for example [15, Theorem VI.4.1]). The proposition follows. □

4 An Important Warning about Canonical and Local Heights

It is crucial when comparing formulae for canonical heights and local heights to bear in mind the different normalizations of these quantities appearing in the literature.

We have chosen our normalization of the canonical height to agree with that taken in Cremona's book [2]. This has the merit of being the most natural normalization for the conjectures of Birch and Swinnerton-Dyer (see [2, page 72]). Our canonical height is double that found in the papers and books of Silverman that we cite ([15], [16], [13]).

The situation with local heights is more complicated. Again our local heights agree with those in Cremona's book. Moreover, if we denote by $\lambda_v^{\mathrm{SilP}}$ the local height in Silverman's paper [16], and by $\lambda_v^{\mathrm{SilB}}$ the local height in Silverman's book [15, Chapter IV], then these are related to our normalization by

$$\lambda_v = 2\lambda_v^{\mathrm{SilP}} = 2\lambda_v^{\mathrm{SilB}} + \frac{1}{6}\log|\Delta_v|. \tag{11}$$

The normalizations λ_v and $\lambda_v^{\mathrm{SilP}}$ are easier to use for explicit purposes. The normalizations $\lambda_v^{\mathrm{SilB}}$ (and $2\lambda_v^{\mathrm{SilB}}$) are better suited for theoretical purposes. In particular $\lambda_v^{\mathrm{SilB}}$ has the advantage of being independent of the choice of Weierstrass model for the curve (see [15, Theorem VI.1.1]). We would like to record here the corresponding fact for our normalization, which can be readily deduced from the relationship expressed in (11).

Lemma 4 *Suppose that E and E' are different models for the same elliptic curve, and let λ_v and λ_v' be the corresponding local heights for valuation v. Then*

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

where Δ and Δ' are respectively the discriminants of the models E and E'.

5 The local height λ_v for non-archimedean valuations

Throughout this section $v \in M_K^0$ is a non-archimedean valuation. Our aim in this section is to compile an exhaustive table of the values of local heights at the points of bad reduction. This is the most difficult step in the proof of Theorem 1. To do this we will need the following proposition, due to Silverman, which gives explicit formulae for non-archimedean local heights.

Proposition 5 (Silverman's Explicit Formulae for Local Heights) *Let $v \in M_K^0$, and suppose that E is minimal at v. Suppose that $P \in E(K_v) \backslash \{O\}$.*

(a) The value of $\lambda_v(P)$ depends only on the image of P in $E(K_v)/E_0(K_v)$.
(b) If $P \in E_0(K_v)$ then

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

(c) Suppose E has Kodaira type $\mathrm{I_m}$ at v. If $P \in E(K_v) \backslash E_0(K_v)$ lies on the i-th component of $E(K_v)/E_0(K_v)$ then

$$\lambda_v(P) = -\frac{i(m-i)}{m} \frac{\log(q_v)}{n_v}.$$

(d) If E has Kodaira type IV or IV^* then

$$\lambda_v(P) = -\frac{2}{3} \operatorname{ord}_v(2y(P) + a_1 x(P) + a_3) \frac{\log(q_v)}{n_v}$$

for all $P \notin E_0(K_v)$.
(e) If E has Kodaira type III, III^*, I_0^*, or I_m^* then

$$\lambda_v(P) = -\frac{1}{4} \operatorname{ord}_v(3x(P)^4 + b_2 x(P)^3 + 3b_4 x(P)^2 + 3b_6 x(P) + b_8) \frac{\log(q_v)}{n_v}$$

for all $P \notin E_0(K_v)$.

Proof: See Silverman's paper [16, pages 351–354]. Parts (a) and (b) are implicit in Néron's original paper [10]; part (c) may have originally appeared in a letter from Tate to Serre, while parts (d) and (e) are essentially in Silverman's thesis, subsequently published in [16]. □

Next we use the above explicit formulae to calculate an exhaustive table of values of local heights at points of bad reduction.

Proposition 6 *Suppose that E is minimal at $v \in M_K^0$, and that the Tamagawa index $c_v > 1$. The values of $\lambda_v(P)$ as P ranges over $E(K_v) \backslash E_0(K_v)$ (that is, the v-adic points of bad reduction) are given by Table 2.*

The remainder of this section is devoted to the proof of this proposition. In the next section we use it to deduce Theorem 1.

We need separate proofs for different Kodaira types and Tamagawa indices. In the course of the proof we will need to make unimodular changes of the Weierstrass model for E: these are standard changes of variable of the form $x = x' + r$ and $y = y' + sx' + t$ where r, s, t are v-adic integers. Note that

Table 2
Values of λ_v

Kodaira type of $E_v^{\min}$ at v	Tamagawa index c_v	$-\left(\frac{n_v}{\log q_v}\right)\lambda_v$
$\mathrm{I_m}$, m even	m	$i(m-i)/m \quad i = 1, \ldots m-1$
$\mathrm{I_m}$, m even	2	$m/4$
$\mathrm{I_m}$, m odd	m	$i(m-i)/m \quad i = 1, \ldots m-1$
III	2	1/2
IV	3	2/3
I_0^*	2 or 4	1
$\mathrm{I_m^*}$	2	1
$\mathrm{I_m^*}$	4	1, $\quad (m+4)/4$
IV*	3	4/3
III*	2	3/2

such changes do not affect the minimality of E nor its discriminant, and we deduce from Lemma 4 that the values of λ_v are also unchanged by these model changes. In other words, for the purpose of proving the proposition, such changes are harmless.

Let $\pi = \pi_v$ be a uniformiser for v; write ord for ord_v. Recall our assumption that the Tamagawa index $c_v > 1$.

Proof for Kodaira type $\mathrm{I_m}$, m odd and $c_v = m$: If the Kodaira type is $\mathrm{I_m}$ with m odd then there are in general two possibilities for the Tamagawa index c_v: either $c_v = 1$ or $c_v = m$. Since we have excluded the former possibility, we may assume the latter holds. Then, from Proposition 5, we know that the possible values for λ_v at the points of bad reduction are

$$-\frac{i(m-i)}{m}\frac{\log(q_v)}{n_v}$$

with $i = 1, 2, \ldots, m-1$ in agreement with the table. □

Proof for Kodaira type $\mathrm{I_m}$, m even, $c_v = 2$ or m: If the Kodaira type is $\mathrm{I_m}$ with m even then there are again two possibilities for the Tamagawa index c_v; either $c_v = 2$ or $c_v = m$. For $c_v = m$ the proof is exactly as above.

Suppose now that $c_v = 2$. Then the points with bad reduction have order 2 in $E(K_v)/E_0(K_v)$, and so lie in the $m/2$-th component. If P is a point of bad reduction then $\lambda_v(P) = -(m/4)(\log q_v/n_v)$ as required.

Note that here we benefit from knowing that $P \in E(K_v)$ and not just $E(\overline{K_v})$, which is cyclic of order m.

□

We will henceforth suppose that the reduction is additive and that $c_v > 1$. We note from the table in [15, page 365] that the assumption $c_v > 1$ excludes the possibility of Kodaira types II and II*.

Suppose that $P \in E(K_v)$ is a point of bad reduction. By making an appropriate translation we can suppose that $P = (0, 0)$, and so $a_6 = 0$. We will follow the steps of Tate's algorithm as in [15, pages 366–369]. Since we have advanced beyond Kodaira types I_0, I_m, II, we see that

$$a_6 = 0, \qquad \pi \mid b_2, \qquad \pi^2 \mid b_6, \qquad \pi^2 \mid b_8.$$

As $x(P) = y(P) = 0$, Proposition 5 gives

$$\lambda_v(P) = -\frac{2}{3}\operatorname{ord}(a_3)\frac{\log(q_v)}{n_v}$$

if the Kodaira type is IV or IV*, and

$$\lambda_v(P) = -\frac{1}{4}\operatorname{ord}(b_8)\frac{\log(q_v)}{n_v}$$

if the Kodaira type is III, III*, I_0^* or I_m^*. For most of the remaining cases we content ourselves with evaluating $\operatorname{ord}(a_3)$ and $\operatorname{ord}(b_8)$, whichever is relevant, and leave the rest to the reader. We note that the assumption that $c_v > 1$ forces $c_v = 2,\ 3,\ 3,\ 2$ for Kodaira types III, IV, IV*, III* respectively. For Kodaira types I_0^* and I_m^*, if $c_v > 1$ then $c_v = 2$ or 4. We state this to demonstrate that we have covered all the possibilities in Table 2.

Proof of the Lemma for Kodaira type III **and** $c_v = 2$**:** From Tate's algorithm we know that $\pi^3 \nmid b_8$. Hence $\operatorname{ord}(b_8) = 2$. □

We resume following the steps of Tate's algorithm. Thus suppose

$$a_6 = 0, \qquad \pi^2 \mid b_6, \qquad \pi^3 \mid b_8.$$

Proof for Kodaira type IV **and** $c_v = 3$**:** From Tate's algorithm we know that $\pi^3 \nmid b_6$. But $b_6 = a_3^2 + 4a_6 = a_3^2$. Thus $\operatorname{ord}_v(a_3) = 1$. □

Suppose now that $\pi^3 \mid b_6$. Tate's algorithm now (Step 6) requires us to make a certain change of variables, and since we have $a_6 = 0$, it is sufficient to make a translation of the form $y' = y + \alpha x$ (where $\alpha \in \mathcal{O}_v$ is the double root of $Y^2 + a_1 Y - a_2 \equiv 0 \pmod{\pi}$). This translation does not move the point $P = (0,0)$, and we get

$$\pi \mid a_1, \quad \pi \mid a_2, \quad \pi^2 \mid a_3, \quad \pi^2 \mid a_4, \quad a_6 = 0.$$

Next we must consider the factorization of the polynomial

$$P(T) = T^3 + a_{2,1}T^2 + a_{4,2}T,$$

modulo π, where by definition $a_{i,r} := \pi^{-r} a_i$.

Proof for Kodaira type I_0^* with $c_v = 2$ or 4: From Tate's algorithm we know that the Kodaira type is I_0^* if and only if $P(T)$ has distinct roots modulo π. In particular $\pi \nmid a_{4,2}$, or equivalently $\mathrm{ord}(a_4) = 2$. Since $a_6 = 0$, we see that $b_8 = -a_1 a_3 a_4 + a_2 a_3^2 - a_4^2$. It follows from the assumptions so far that $\mathrm{ord}(b_8) = 4$. □

Now if $P(T)$ has a double root and a simple root modulo π then the Kodaira type is I_m^*, and if it has a triple root modulo π then the Kodaira type is IV^* or III^*. The case I_m^* is complicated and we will leave it to the end.

Thus suppose that $P(T)$ has a triple root modulo π. This together with previous assumptions implies that

$$\pi \mid a_1, \quad \pi^2 \mid a_2, \quad \pi^2 \mid a_3, \quad \pi^3 \mid a_4, \quad a_6 = 0.$$

Proof for Kodaira type IV^* and $c_v = 3$: From Tate's algorithm we know that this type occurs when $Y^2 + a_{3,2}Y$ has distinct roots modulo π. Thus $\mathrm{ord}(a_3) = 2$. □

Now we suppose that $Y^2 + a_{3,2}Y$ has a double root modulo π; so $\pi^3 \mid a_3$.

Proof for Kodaira type III^* and $c_v = 2$: From Tate's algorithm, this case is equivalent to $\mathrm{ord}(a_4) = 3$, which implies that $\mathrm{ord}(b_8) = 6$. □

We are now left with proving the proposition for Kodaira type I_m^*.

Proof for Kodaira type I_m^*, m even, and $c_v = 2$ or 4: We have already dealt with I_0^*, so we may suppose that $m \geq 2$. By step 7 of Tate's algorithm in [15, pages 367–368] and its proof in [15, pages 373-374] we can make a translation such that

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

(It is no longer convenient to maintain the assumption $a_6 = 0$). Note that $-a_{2,1}$ is a simple root, modulo π, of the polynomial $X^3 + a_{2,1}X^2 + a_{4,2}X + a_{6,3}$. Hence, by Hensel's Lemma, this polynomial has a (unit) root $\alpha \in \mathcal{O}_v^*$ such that $\alpha \equiv -a_{2,1} \pmod{\pi}$. Then $P_1 = (\pi\alpha, 0) \in E(K_v)$ has bad reduction.

We want to evaluate $\lambda_v(P_1)$ using the formula in Proposition 5. Now $x(P_1) = \pi\alpha \equiv -a_2 \pmod{\pi^2}$. We also have $\pi^3 \mid b_4$, $\pi^5 \mid b_6$ and $\pi^6 \mid b_8$. Hence it follows for $x = x(P_1)$ that

$$\begin{aligned} 3x^4 + b_2x^3 + 3b_4x^2 + 3b_6x + b_8 &\equiv (\pi\alpha)^3(3\pi\alpha + a_1^2 + 4a_2) + 3b_4(\pi\alpha)^2 \\ &\equiv -(\pi\alpha)^4 \pmod{\pi^5}, \end{aligned}$$

which has valuation 4. Thus $\lambda_v(P_1) = -\log(q_v)/n_v$, using Proposition 5.

Part (a) of Proposition 5 tells us that λ_v factors through $E(K_v)/E_0(K_v)$. If $c_v = 2$, then the only value λ_v takes at points of bad reduction is $-\log(q_v)/n_v$. This proves the proposition when $c_v = 2$ (and Kodaira type is I_m^* with m even).

Suppose that $c_v = 4$. There are now three nontrivial cosets of $E_0(K_v)$ in $E(K_v)$ to be considered, including the point $P_1 = (\pi\alpha, 0)$ defined above. We make a further translation that simplifies our model for E, taking a point representing either of the two extra cosets to $(0, 0)$. The special fiber of the Néron model in our case is defined by

$$a_{2,1}x_\mu^2 + a_{4,\mu+1}x_\mu + a_{6,2\mu+1} \equiv 0 \pmod{\pi}$$

where $m = 2\mu - 2$ and $x_\mu = x\pi^{-\mu}$ (see [15, page 374]). The fact that $c_v = 4$ is equivalent to saying that the polynomial on the left has distinct roots modulo π, so these lift to roots in $\mathcal{O}_v$. The corresponding lifted values of x are the x-coordinates of the points representing the other two cosets in the component group.

Hence, making a suitable translation of x modulo π^μ, we may move either of these two roots to 0, so that $P_2 = (0, 0)$ is a new point of bad reduction. Note that this extra translation has not changed P_1 essentially, since $x(P_1)$ was only determined modulo π^2. After the extra translation, we have

$$\pi \mid a_1, \quad \pi \mid\mid a_2, \quad \pi^{\frac{m+4}{2}} \mid a_3, \quad \pi^{\frac{m+4}{2}} \mid\mid a_4, \quad a_6 = 0.$$

It is straightforward to verify that $\mathrm{ord}(b_8) = m+4$, so the corresponding value in our table is $(m+4)/4$.

Since P_1, P_2 are points of bad reduction that have differing local heights they must belong to different non-trivial cosets of $E(K_v)/E_0(K_v)$. Hence we know all of the values of λ_v at the points of bad reduction. □

Proof for Kodaira type I_m^*, m odd, and $c_v = 2$ or 4: This case is similar to the previous case. We can make a translation such that

$$\pi \mid a_1, \quad \pi \mid\mid a_2, \quad \pi^{\frac{m+3}{2}} \mid a_3, \quad \pi^{\frac{m+5}{2}} \mid a_4, \quad \pi^{m+3} \mid a_6.$$

Again we have a point $P_1 \in E(K_v)$ (of order 2) such that $P_1 \equiv (-a_2, 0) \pmod{\pi^2}$ and again $\lambda_v(P_1) = -\log(q_v)/n_v$. So if $c_v = 2$ then we are finished.

Suppose that $c_v = 4$. Now for m odd Tate's algorithm tells us that the component group is cyclic (of order 4). Let $P_2 \in E(K_v)$ represent the element of exact order 4 in $E(K_v)/E_0(K_v)$. Thankfully, Silverman has determined ([16, page 353]) that $\lambda_v(P_2) = -(m+4)\log q_v/4n_v$. The three non-trivial elements of the component group are represented by P_1 and $\pm P_2$. Since $\lambda_v(-P_2) = \lambda_v(P_2)$, we see that the nonzero values for λ_v are precisely

$$-\frac{\log(q_v)}{n_v}, \quad -\frac{(m+4)}{4}\frac{\log(q_v)}{n_v}.$$

This completes the proof of the proposition in this case. □

6 Proof of Theorem 1

In this section we deduce Theorem 1 from two propositions.

Proof of Theorem 1: Recall the decomposition of the height difference $h - \hat{h}$ in (1) in terms of the functions Ψ_v. Theorem 1 follows immediately from Propositions 7, 8 below. □

Proposition 7 *Let $v \in M_K$ (archimedean or non-archimedean), and define ϵ_v and δ_v by (5). Then for all $P \in E(K_v)$ we have*

$$\frac{\log(\delta_v)}{3} \le \Psi_v(P) \le \frac{\log(\epsilon_v)}{3}.$$

Proof: The proposition follows immediately from the expansion of Ψ_v in terms of Φ_v given in (10) and the definitions of ϵ_v and δ_v in (5). □

Proposition 8 *Suppose $v \in M_K^0$. Then*

$$\inf_{P \in E(K_v)} \Psi_v(P) = 0,$$

and

$$\begin{aligned}\sup_{P \in E(K_v)} \Psi_v(P) &= \left(\alpha_v + \frac{1}{6}\operatorname{ord}_v(\Delta/\Delta_v^{\min})\right)\frac{\log(q_v)}{n_v}\\ &= \alpha_v \frac{\log(q_v)}{n_v} - \frac{1}{6}\log\left|\Delta/\Delta^{\min}\right|_v,\end{aligned}$$

where α_v is given by Table 1, and $\Delta^{\min}$ is the discriminant of a minimal model for E at v.

Proof: First we make two claims.

Claim I.

$$\Psi_v(P) = \begin{cases} 0 & \text{if} \quad P \in E_0(K_v) \\ -\lambda_v(P) & \text{otherwise.}\end{cases} \tag{12}$$

To see this, recall that

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

for all $P \in E_0(K_v)\setminus\{O\}$, by Proposition 5(a). It is then immediate from (9) that $\Psi_v(P) = 0$ for all $P \in E_0(K_v)$. Suppose that $P \notin E_0(K_v)$. Then $|x(P)|_v \leq 1$ (since all the points with $|x(P)|_v > 1$ have good reduction). From the definition of Ψ_v in (9) we deduce that $\Psi_v(P) = -\lambda_v(P)$. This proves our claim.

Claim II. The proposition is true under the assumption that E is minimal at v. In fact a little more is true: if E is minimal at non-archimedean place v, then

$$\inf_{P \in E(K_v)} \Psi_v(P) = 0,$$

with the infimum attained at all points $P \in E_0(K_v)$. Moreover,

$$\sup_{P \in E(K_v)} \Psi_v(P) = \alpha_v \frac{\log(q_v)}{n_v},$$

where α_v is given by Table 1, with the supremum attained at some point $P \in E(K_v)$ of bad reduction.

To see this, recall that Proposition 6 gives us all the values of $\lambda_v(P)$ for points of bad reduction. From this and (12), we are able to write down for each Tamagawa index and reduction type a complete list of values of $\Psi_v(P)$ for $P \in E(K_v)$, and so verify the claimed infima and suprema. We leave the details to the reader. This proves our second claim.

Since we have already covered the minimal case, we may suppose that the model E is non-minimal. Thus there is a change of variable

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

with $u,\ r,\ s,\ t \in \mathcal{O}_v$ and $\mathrm{ord}_v(u) \geq 1$, such that the resulting model $E^{\min}$ is minimal. Denote by Ψ_v' the function corresponding to Ψ_v on $E^{\min}$, and by $\Delta^{\min}$ the discriminant of $E^{\min}$, so that $\Delta = u^{12}\Delta^{\min}$. Then, from Lemma 4 and the definition of Ψ_v in (9),

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

It is helpful to take a closer look at the middle term of the right-hand side of the above equation. It is a straightforward exercise to show that

$$\inf_{P \in E(K_v)} \log\left(\frac{\max\{1, |u^2 x'(P) + r|_v\}}{\max\{1, |x'(P)|_v\}}\right) = \log|u^2|_v,$$

with the infimum attained for large $x'(P)$, and also

$$\sup_{P \in E(K_v)} \log\left(\frac{\max\{1, |u^2 x'(P) + r|_v\}}{\max\{1, |x'(P)|_v\}}\right) = 0,$$

with the supremum attained whenever $|x'(P)|_v \leq 1$.

Now the infimum of Ψ_v', which is 0 by the case already proved, is attained for points with large $|x'(P)|_v$ (since these have good reduction); and the supremum of Ψ_v', which is $\alpha_v \log(q_v)/n_v$, is attained at some point P with $|x'(P)|_v \leq 1$. We deduce that

$$\begin{aligned} \inf_{P \in E(K_v)} \Psi_v(P) &= \log|u^2|_v - \frac{1}{6}\log\left|\Delta/\Delta^{\min}\right|_v \\ &= -\frac{1}{6}\log\left|(u^{-12}\Delta/\Delta^{\min})\right|_v = 0, \end{aligned}$$

and

$$\sup_{P \in E(K_v)} \Psi_v(P) = \alpha_v \frac{\log(q_v)}{n_v} - \frac{1}{6}\log\left|\Delta/\Delta^{\min}\right|_v$$

as required. □

7 The Real Contributions

To be able to compute the bounds in our Theorem 1 we need a method for determining δ_v and ϵ_v for archimedean places v. In this section we give such a method for real places v. Thus suppose that v is a real place; in other words, there is an embedding $\sigma : K \hookrightarrow \mathbb{R}$ such that $|a|_v = |\sigma(a)|$ for all $a \in K$. To ease the notation, we will henceforth identify K with its image in $\mathbb{R}$ under σ, and thus view elements of K as real numbers.

Write

$$\begin{aligned} f(x) &= 4x^3 + b_2x^2 + 2b_4x + b_6, \\ g(x) &= x^4 - b_4x^2 - 2b_6x - b_8. \end{aligned}$$

and let

$$F(x) = x^4 f(1/x), \qquad G(x) = x^4 g(1/x).$$

Define

$$\begin{aligned} D &= \{x \in [-1,1] : f(x) \geq 0\}, \\ D' &= \{x \in [-1,1] : F(x) \geq 0\}. \end{aligned}$$

The following lemma is elementary.

Lemma 9 *Define constants e, e' by*

$$\begin{aligned} e &= \inf_{x\in D} \max\{|f(x)|, |g(x)|\}, \\ e' &= \inf_{x\in D'} \max\{|F(x)|, |G(x)|\}, \end{aligned}$$

and constants d, d' by

$$\begin{aligned} d &= \sup_{x\in D} \max\{|f(x)|, |g(x)|\}, \\ d' &= \sup_{x\in D'} \max\{|F(x)|, |G(x)|\}. \end{aligned}$$

Then $\epsilon_v = \min(e, e')^{-1}$ and $\delta_v = \max(d, d')^{-1}$.

Proof: The lemma follows from the definitions of ϵ_v and δ_v made in (5) and the fact that $(x, y) \in E(\mathbb{R})$ if and only if $f(x) = (2y + a_1x + a_3)^2$. □

It is clear that D, D' are finite unions of closed intervals. Moreover the problem of determining δ_v and ϵ_v has been reduced to the problem of determining d, d', e, e'. This is straightforward by the following lemma.

Lemma 10 *If P, Q are continuous real functions and $I \subset \mathbb{R}$ is a closed interval, then the extrema of the continuous function* $\max\{|P(X)|, |Q(X)|\}$ *over the interval I are attained at one of the following points:*

(i) an end point of I;
(ii) one of the roots of $P+Q$, $P-Q$ in the interval I;
(iii) a turning point of one of the functions P, Q.

Proof: We simply note that at any point in I not listed in (i) or (ii), the function $\max\{|P(X)|, |Q(X)|\}$ is equal to one of $\pm P$, $\pm Q$ and its supremum or infimum must be a local supremum or infimum of P, or Q. □

It is easy to turn this lemma into an algorithm. To compute e, for example, let S be the set of zeros of f, g, f', g', and $f \pm g$, together with $\{\pm 1\}$. Then e is the minimum of $\max\{|f(x)|, |g(x)|\}$ over those $x \in S$ for which $|x| \leq 1$ and $f(x) \geq 0$.

8 The Complex Contributions I. Groebner Approach

In this and the next section we consider the determination of δ_v and ϵ_v for complex archimedean places v. As in the previous section, we regard all elements of K as lying in $\mathbb{C}$ via a suitable embedding.

Let f, g, F, G be as in the previous section, and $D = \{z \in \mathbb{C} : |z| \leq 1\}$ be the unit disc. If P, Q are polynomials with complex coefficients, we define

$$\alpha(P,Q) = \inf_{z\in D} \max\{|P(z)|, |Q(z)|\},$$
$$\beta(P,Q) = \sup_{z\in D} \max\{|P(z)|, |Q(z)|\}.$$

We note in passing that since D is compact, the infimum and supremum exist and are attained at some points of D.

The following lemma is elementary.

Lemma 11 *With notation as above*

$$\epsilon_v = \min(\alpha(f,g), \alpha(F,G))^{-1}, \qquad \delta_v = \max(\beta(f,g), \beta(F,G))^{-1}.$$

In this section we give a method for computing $\alpha(P,Q)$, $\beta(P,Q)$ for polynomials P, Q with complex coefficients that do not vanish simultaneously (it is noted that this condition is satisfied by both of our pairs f, g and F, G). The method is based on real multivariate calculus and Groebner bases. In the next section we give an alternative method for computing the complex contributions using a fairly simple numerical method based on repeated quadrisection of the unit disc D (with fast convergence).

So let P, $Q \in \mathbb{C}[z]$ be polynomials with complex coefficients that do not vanish simultaneously. Write $z = x + iy$ and $P = P_1 + iP_2$, $Q = Q_1 + iQ_2$ where P_j, Q_j are real polynomials in x, y.

Lemma 12 *The supremum of the function* $\max\{|P(z)|, |Q(z)|\}$ *on the region* D *is attained at a point* $z_0 = x_0 + iy_0$ *that satisfies one of the following pairs of simultaneous equations:*

$$\text{(i)} \quad y\frac{\partial(P_1^2+P_2^2)}{\partial x} - x\frac{\partial(P_1^2+P_2^2)}{\partial y} = 0, \qquad x^2+y^2=1,$$
$$\text{(ii)} \quad y\frac{\partial(Q_1^2+Q_2^2)}{\partial x} - x\frac{\partial(Q_1^2+Q_2^2)}{\partial y} = 0, \qquad x^2+y^2=1.$$

The infimum of the function $\max\{|P(z)|, |Q(z)|\}$ *on* D *is attained at a point* $z_0 = x_0 + iy_0$ *such that one of the following holds:*

(a) the point z_0 *satisfies one of the above pairs of simultaneous equations (i) or (ii),*

(b) the point z_0 *satisfies the simultaneous equations*

$$P_1^2+P_2^2 = Q_1^2+Q_2^2, \qquad x^2+y^2=1,$$

(c) the point z_0 *belongs to the interior* $x^2+y^2<1$ *and satisfies these two simultaneous equations:*

$$P_1^2+P_2^2 = Q_1^2+Q_2^2, \qquad \frac{\partial(P_1^2+P_2^2)}{\partial x}\frac{\partial(Q_1^2+Q_2^2)}{\partial y} - \frac{\partial(P_1^2+P_2^2)}{\partial y}\frac{\partial(Q_1^2+Q_2^2)}{\partial x} = 0.$$

Proof: Note that

$$\sup\max\{|P(z)|, |Q(z)|\} = \max\{\sup|P(z)|, \sup|Q(z)|\},$$

and, by the maximum modulus theorem [1, page 134], the suprema of $|P(z)|$ and $|Q(z)|$ are attained at the boundary $x^2+y^2=1$. Suppose that the supremum is attained at some point $z_0 = x_0 + iy_0$. Then z_0 must be a local supremum for either $|P(z)|$ or $|Q(z)|$ restricted to the unit circle. Let us suppose that it is a local supremum for $|P(z)|$ restricted to the unit circle. Then (x_0, y_0) represents a local maximum for the function $P_1^2+P_2^2$ on the (analytic) curve $x^2 + y^2 = 1$. By the method of Lagrange multipliers, the two vectors $\nabla(P_1^2 + P_2^2)$, $\nabla(x^2 + y^2 - 1)$ must be linearly dependent when evaluated at (x_0, y_0). It is then easy to verify that (x_0, y_0) satisfies

$$y\frac{\partial(P_1^2 + P_2^2)}{\partial x} - x\frac{\partial(P_1^2 + P_2^2)}{\partial y} = 0.$$

Thus (x_0, y_0) satisfies (i). Similarly, if z_0 is a local supremum for $|Q(z)|$ restricted to the unit circle then (x_0, y_0) satisfies (ii). This proves the first part of the Lemma.

For the second part, again suppose that the infimum is attained at some point $z_0 = x_0 + iy_0$. Suppose first that z_0 belongs to the boundary. If $|P(z_0)| = |Q(z_0)|$ then (b) is satisfied. Suppose that $|P(z_0)| > |Q(z_0)|$. Then in some small neighbourhood of z_0, we see that $\max\{|P(z)|, |Q(z)|\} = |P(z)|$ and so z_0 must be a local infimum of the function $|P(z)|$ restricted to the unit circle. By a trivial modification of the above argument we show that $z_0 = x_0 + iy_0$ satisfies (i). Similarly if $|P(z_0)| < |Q(z_0)|$ then (ii) is satisfied. Thus either (a) or (b) is satisfied if z_0 belongs to the boundary.

We are now left to consider the case where z_0 belongs to the interior of D. We want to first show that $|P(z_0)| = |Q(z_0)|$. Suppose not; without loss of generality we may suppose that $|P(z_0)| > |Q(z_0)|$. Then for some small disc around z_0 and contained in D we have $|P(z)| > |Q(z)|$ implying $\max\{|P(z)|, |Q(z)|\} = |P(z)|$. Thus the holomorphic function P attains a non-zero infimum that is in the interior of this small disc. Applying the maximum modulus theorem to P^{-1} immediately gives a contradiction.

Hence $|P(z_0)| = |Q(z_0)|$. Consider the (analytic) curve in $\mathbb{R}^2$ defined by the equation $P_1^2 + P_2^2 = Q_1^2 + Q_2^2$. Then we are saying that (x_0, y_0) is on this curve and moreover is a point where the infimum of the function

$$\max\left\{P_1^2 + P_2^2, Q_1^2 + Q_2^2\right\} = P_1^2 + P_2^2$$

is attained. The proof can now be completed using Lagrange multipliers as before. □

Thus from the lemma, to compute $\alpha(P, Q)$, $\beta(P, Q)$, we need to solve a few pairs of polynomial equations in two variables. In theory these can be solved using elimination theory. There are two alternatives here:

(1) The first is to recall that our coefficients are contained in a number field and do the elimination using Groebner bases algorithms over this field, to obtain the points in some extension field, and then specialize using the complex embeddings. We have found this (exact arithmetic) approach

extremely slow in practice. Note that while $P, Q \in K[z]$, their real and imaginary parts P_j, Q_j are in general defined over a larger field $K(i)$.

(2) The second approach is to use numerical (that is floating-point) Groebner basis packages available in some computer algebra systems. To the best of our knowledge, the theory behind these floating-point packages is not documented and they may not be entirely rigorous. Thus there is perhaps a risk of missing a solution.

Fortunately, there is now a rigorous numerical Groebner basis algorithm due to Aleksey Kondratyev [6] which could be used for our purpose.

9 The Complex Contributions II. The Repeated Quadrisection Approach

Recall that our objective is to compute $\log(\epsilon_v)$ and $\log(\delta_v)$ to a certain desired accuracy. Suppose $\mu > 0$ is given. We give an algorithm to compute, for any pair of polynomials P, Q with complex coefficients that do not vanish simultaneously, constants $\alpha^*(P,Q)$, $\beta^*(P,Q)$ such that

$$\alpha^*(P,Q)e^{-\mu} \leq \alpha(P,Q) \leq \alpha^*(P,Q). \tag{13}$$

and

$$\beta^*(P,Q) \leq \beta(P,Q) \leq \beta^*(P,Q)e^{\mu}. \tag{14}$$

Thus if f, g, F, G are as before then

$$-\log\min(\alpha^*(f,g),\alpha^*(F,G)) \leq \log\epsilon_v \leq -\log\min(\alpha^*(f,g),\alpha^*(F,G)) + \mu,$$

and

$$-\log\max(\beta^*(f,g),\beta^*(F,G)) - \mu \leq \log\delta_v \leq -\log\max(\beta^*(f,g),\beta^*(F,G));$$

meaning that we can compute the contribution at complex places to arbitrary accuracy μ.

Now fix complex polynomials P, Q that do not vanish simultaneously. To ease notation, let

$$h(z) = \max\{|P(z)|, |Q(z)|\}.$$

Thus

$$\alpha(P,Q) = \inf_{z\in D} h(z), \qquad \beta(P,Q) = \sup_{z\in D} h(z). \tag{15}$$

Given $\eta > 0$ we define

$$\mathcal{E}(z,\eta) = \max\left\{\sum_{n=1}^{d_1}\frac{\eta^n}{n!}|P^{(n)}(z)|, \sum_{n=1}^{d_2}\frac{\eta^n}{n!}|Q^{(n)}(z)|\right\},$$

where d_1, d_2 are the degrees of P, Q respectively.

We naturally identify $\mathbb{R}^2$ and $\mathbb{C}$.

Lemma 13 *Let S be the square $S = [a, a+r] \times [b, b+r]$. Then*

$$h(u) - \mathcal{E}(u, \eta) \leq h(z) \leq h(u) + \mathcal{E}(u, \eta),$$

for all $z \in S$, where either

- *u is the centre of S and $\eta = r/\sqrt{2}$, or*
- *u is a corner of S and $\eta = r\sqrt{2}$.*

Proof: This follows from Taylor's Theorem applied to the polynomials P, Q. $\square$

Now we give a method of computing $\alpha^*(P, Q)$, $\beta^*(P, Q)$. Let H be the set

$$H = \left\{ h\left(\frac{m + ni}{10}\right) : m, n \in \mathbb{Z},\ m^2 + n^2 \leq 100 \right\}.$$

We start with $S = [-1, 1] \times [-1, 1]$ and the initial values

$$\alpha^* = \min H, \qquad \beta^* = \max H.$$

This gives (fairly crude) upper and lower bounds for $\alpha(P, Q)$ and $\beta(P, Q)$; we repeatedly refine these until we obtain values $\alpha^*(P, Q)$ and $\beta^*(P, Q)$ satisfying (13) and (14) respectively. To do this for α^* we use the following recursive procedure, starting with S and α^* as above. When the procedure returns (possibly after many recursive function calls), we will have a value of α^* that we can take as $\alpha^*(P, Q)$.

RefineAlphaBound(P, Q, μ, S, α^*)

```
INPUT:     $P$, $Q \in \mathbb{C}[z]$, $\mu > 0$, square $S = [a, a+r] \times [b, b+r] \subset \mathbb{C}$, $\alpha^*$
OUTPUT:    $\alpha^*$ (possibly modified)
BEGIN
IF $S \cap D = \emptyset$ THEN RETURN($\alpha^*$); ENDIF;
IF $a + r/2 + (b + r/2)i \in D$ THEN $u = a + r/2 + (b + r/2)i$ AND $\eta = r/\sqrt{2}$;
ELSE $u$ is any corner of $S$ in $D$ AND $\eta = r\sqrt{2}$; ENDIF.
IF $h(u) - \mathcal{E}(u, \eta) > \alpha^* e^{-\mu}$ THEN RETURN($\alpha^*$); ENDIF;
LET $\alpha^* = \min(\alpha^*, h(u))$;
LET $S_1 = [a, a + r/2] \times [b, b + r/2]$, $S_2 = [a, a + r/2] \times [b + r/2, b + r]$,
$S_3 = [a + r/2, a + r] \times [b, b + r/2]$, $S_4 = [a + r/2, a + r] \times [b + r/2, b + r]$;
```

```
 7.   LET $\alpha^* = \mathrm{RefineAlphaBound}(P, Q, \mu, S_1, \alpha^*)$;
 8.   LET $\alpha^* = \mathrm{RefineAlphaBound}(P, Q, \mu, S_2, \alpha^*)$;
 9.   LET $\alpha^* = \mathrm{RefineAlphaBound}(P, Q, \mu, S_3, \alpha^*)$;
10.   LET $\alpha^* = \mathrm{RefineAlphaBound}(P, Q, \mu, S_4, \alpha^*)$;
11.   RETURN($\alpha^*$);
12.   END
```

The procedure for calculating $\beta^*(P, Q)$ is slightly different since by Lemma 12 the supremum of $h(z)$ is attained on the boundary of D. Thus let ∂D be the boundary of D (that is the circle $x^2 + y^2 = 1$). To obtain $\beta^*(P, Q)$ apply the following procedure to the value of β^* above with $S = [-1, 1] \times [-1, 1]$ again.

RefineBetaBound(P, Q, μ, S, β^*)

```
      INPUT:      $P$, $Q \in \mathbb{C}[z]$, $\mu > 0$, square $S = [a, a+r] \times [b, b+r] \subset \mathbb{C}$, $\beta^*$
      OUTPUT:     $\beta^*$ (possibly modified)
 1.   BEGIN
 2.   IF $S \cap \partial D = \emptyset$ THEN RETURN($\beta^*$); ENDIF;
 3.   IF $a + r/2 + (b + r/2)i \in D$ THEN $u = a + r/2 + (b + r/2)i$ AND $\eta = r/\sqrt{2}$;
      ELSE $u$ is any corner of $S$ in $D$ AND $\eta = r\sqrt{2}$; ENDIF.
 4.   IF $h(u) - \mathcal{E}(u, \eta) < \beta^* e^{\mu}$ THEN RETURN($\beta^*$); ENDIF;
 5.   LET $\beta^* = \max(\beta^*, h(u))$;
 6.   LET $S_1 = [a, a + r/2] \times [b, b + r/2]$, $S_2 = [a, a + r/2] \times [b + r/2, b + r]$,
      $S_3 = [a + r/2, a + r] \times [b, b + r/2]$, $S_4 = [a + r/2, a + r] \times [b + r/2, b + r]$;
 7.   LET $\beta^* = \mathrm{RefineBetaBound}(P, Q, \mu, S_1, \beta^*)$;
 8.   LET $\beta^* = \mathrm{RefineBetaBound}(P, Q, \mu, S_2, \beta^*)$;
 9.   LET $\beta^* = \mathrm{RefineBetaBound}(P, Q, \mu, S_3, \beta^*)$;
10.   LET $\beta^* = \mathrm{RefineBetaBound}(P, Q, \mu, S_4, \beta^*)$;
11.   RETURN($\beta^*$);
12.   END
```

Remark. For Step 2 in the procedure `RefineAlphaBound` we need a method of deciding if $S \cap D = \emptyset$. Of course our initial square S is $[-1, 1] \times [-1, 1]$ for which we know that the intersection is not empty. All other squares S will be contained in one of the four quadrants, and then all we need to check is whether the corner closest to the origin is in D.

For Step 2 in the procedure `RefineBetaBound`, if S is a square contained in the four quadrants then $S \cap \partial D = \emptyset$ if and only if all its corners are outside D, or all its corners are strictly inside D.

The reader might be surprised to find that we are using a two-dimensional method for estimating the infimum β, when we have shown that the infimum is attained on the boundary. It is true that the boundary $x^2 + y^2 = 1$ can be parametrized by trigonometric (or algebraic) functions of one parameter, and so it should be possible to find the infimum using repeated bisection of an interval rather than repeated quadrisection of the unit disc. But it is then much harder to obtain an error term $\mathcal{E}$ that is simultaneously rigorous and small. The reader will note that our error term $\mathcal{E}(z, \eta)$ (which follows from Taylor's Theorem) is the maximum of finite sums; this is because all sufficiently high derivatives of the polynomials P, Q vanish. The same would not be true for trigonometric or rational functions.

Proposition 14 *The above algorithms terminate giving values $\alpha^*(P, Q)$ $\beta^*(P, Q)$ that satisfy the inequalities (13) and (14) respectively.*

Proof: We prove the proposition for $\alpha^*(P, Q)$; the proof for $\beta^*(P, Q)$ is similar. Recall that

$$\alpha(P, Q) = \inf_{z \in D} h(z).$$

We note that the initial value of α^* is obtained by taking the minimum of values of h at some points in D, and throughout the algorithm α^* is only changed in Step 5, where we replace α^* with $\min(\alpha^*, h(u))$ with u being some point in D. Clearly the resulting value of $\alpha^*(P, Q)$ satisfies

$$\alpha^*(P, Q) \geq \inf_{z \in D} h(z) = \alpha(P, Q).$$

Moreover, when we leave any square S without quadrisecting it, either it is outside D completely (and there is nothing to prove), or

$$h(u) - \mathcal{E}(u, \eta) > \alpha^* e^{-\mu}.$$

But for all $z \in S \cap D$ we know that $h(z) \geq h(u) - \mathcal{E}(u, \eta)$. Moreover, the final value $\alpha^*(P, Q)$ satisfies $\alpha^*(P, Q) \leq \alpha^*$. Thus

$$h(z) > \alpha^*(P, Q) e^{-\mu}$$

for all z belonging to $S \cap D$. If the algorithm terminates then we will have covered D with squares S so that this inequality is satisfied for all points on the overlaps $S \cap D$; thus this inequality is satisfied for all $z \in D$.

All that remains to prove now is that the algorithm terminates. Suppose otherwise; then there is a convergent sequence $\{u_n\}_{n=1}^{\infty} \subset D$, and a sequence of real numbers $\{\eta_n\}$ converging to zero, such that

$$h(u_n) - \mathcal{E}(u_n, \eta_n) \leq \alpha_n^* e^{-\mu},$$

where

$$\alpha_n^* = \min(h(u_i) \ : \ i = 1, \ldots, n).$$

From the formula for $\mathcal{E}$ we see that $\lim \mathcal{E}(u_n, \eta_n) = 0$. Thus

$$\lim \alpha_n^* \leq \lim h(u_n) \leq \lim \alpha_n^* e^{-\mu}.$$

Since $\mu > 0$ we deduce that $\lim \alpha_n^* = \lim h(u_n) = 0$. This is impossible since then P, Q will have a common zero. □

10 Silverman's Archimedean Contributions

In [13, page 737] Silverman gives an estimate for the archimedean contributions that is proved using the complex parameterization of the curve. We give this here as it is occasionally better than our own archimedean estimates, and is indeed simpler to calculate for complex places.

Theorem 15 (Silverman) *Let $E/\mathbb{C}$ be an elliptic curve given by a Weierstrass equation (2), and let $\lambda : E(\mathbb{C}) \backslash \{O\} \to \mathbb{R}$ be the complex local height function. Then for all $P \in E(\mathbb{C})$,*

$$\begin{aligned} & -\frac{1}{6}\log^+|\Delta| - \frac{1}{6}\log^+|j| - \log^+|\frac{b_2}{12}| - \log 2^* - 2.14 \\ & \leq \log\max\{1, x(P)\} - \lambda(P) \\ & \leq \frac{1}{6}\log^+|\Delta^{-1}| + \frac{1}{4}\log^+|j| + \log^+|\frac{b_2}{12}| + \log 2^* + 1.946, \end{aligned}$$

where $\log^+ x = \log\max\{1, x\}$ *and*

$$2^* = \begin{cases} 2 \text{ if } b_2 \neq 0, \\ 1 \text{ if } b_2 = 0. \end{cases}$$

11 Examples and Numerical Comparisons

In this section we give some examples based on our implementations of the height bound formula in Theorem 1. In the case $K = \mathbb{Q}$ we have implementations in `PARI/GP` (see [11]), MAGMA (see [8]) and `C++`, the latter being part of the first-named author's package of elliptic curve programs including the 2-descent program `mwrank`. For general number fields, the algorithms were implemented in MAGMA by the second-named author, and by the third-named author in `Mathematica` (see [9]), using the floating-point Groebner basis method which is implemented in that package for computation of the complex contributions.

Example 1. Consider the curve

$$E: \qquad y^2 = x^3 + (1+5i)x + (3+i)$$

over the field $K = \mathbb{Q}(i)$. Using our `Mathematica` program we get the bound

$$-1.37727 \le \Psi_\infty \le 0.114857$$

for the complex contribution. In comparison, Silverman's theorem gives

$$-4.89012 \le \Psi_\infty \le 3.96119.$$

We now complete the computation of the bound for $h - \hat{h}$ using our Theorem 1. The discriminant of the curve is

$$1280 + 4448i = -i(1+i)^{10}(40+139i)$$

where the last factor is prime. Since the discriminant is not divisible by any 12–th powers we see that the curve is globally minimal. The Tamagawa indices at the two bad primes are both 1. From Theorem 1 it follows that

$$-1.37727 \le h(P) - \hat{h}(P) \le 0.114857$$

for all $P \in E(K)$. Silverman's bounds for the same curve are

$$-4.89012 \le h(P) - \hat{h}(P) \le 5.75838.$$

We also programmed our repeated quadrisection method for computing the complex contributions in `PARI/GP`. Taking $\mu = 0.01$ we find that

$$0.34456246612 \ldots\ldots \le \log \epsilon_v \le 0.34556246612 \ldots$$

Thus we know $\log(\epsilon_v)$ to three decimal places (the computation took 1.03 seconds). Thus the complex contribution

$$\frac{n_v}{3[K:\mathbb{Q}]} \log \epsilon_v \le 0.1151.$$

Example 2. This example comes from the paper [5] of Halberstadt and Kraus. Let $K = \mathbb{Q}(\theta)$ be the degree 5 number field generated by a root θ of the polynomial $x^5 + 5x^3 + 5x - 1$. Let E be the curve defined over K with equation

$$y^2 = x^3 + (-30\theta^3 - 100\theta + 30)x^2 + (500\theta^4 - 600\theta^3 + 500\theta^2 - 1700\theta + 300)x + (4000\theta^4 - 28000\theta^3 - 57000\theta + 11000).$$

K has one real embedding and two (pairs of) complex embeddings. The associated values of ϵ_v are (approximately) 2.21, 25.11 and 20.52 respectively, giving a total contribution of 0.8856 for the archimedean contribution to the upper bound. Similarly the values of δ_v are 0.01808, 0.0000000358 and 0.000000167, giving a lower bound of -4.634.

There are three primes of bad reduction:

- one of norm 2, Kodaira type II, discriminant valuation 16 for the given model and 4 for the minimal model; this contributes $\frac{2}{5}\log(2)$ to the upper bound;
- one of norm 16, Kodaira type II, discriminant valuation 16 for the given model and 4 for the minimal model; this contributes $\frac{8}{5}\log(2)$ to the upper bound;
- one of norm 5, Kodaira type III*, discriminant valuation 45 for the given model and 9 for the minimal model; this contributes $\frac{3}{2}\log(5)$ to the upper bound;

giving a total contribution from the non-archimedean valuations of 3.800.

Putting these together we find that the height bounds for this curve are

$$-4.634 \leq h(P) - \hat{h}(P) \leq 4.686.$$

This computation took about 6 minutes, most of the time being spent computing the complex contributions to the bounds.

For the same curve, the Silverman bounds are

$$-11.01053 \leq h(P) - \hat{h}(P) \leq 11.42791.$$

Example 3.

As a comprehensive test of our bounds, we considered the 33355 isomorphism classes of elliptic curves defined over $\mathbb{Q}$ with conductors N in the range $20000 \leq N \leq 25000$ (see [3]). For each curve we computed the height bounds given by our Theorem 1, and the height bounds given by Silverman's paper [13]. Our average lower and upper bounds for the difference $h - \hat{h}$ are -3.483 and 5.218, whereas the respective averages for Silverman bounds are -9.011 and 11.251. Indeed we found that both our upper and lower bounds are better than Silverman's bounds for all curves in the given range except the following two:

- For the curve 20449G3 our bounds are $-12.594 \leq h - \hat{h} \leq 17.251$ whereas Silverman's bounds are $-14.214 \leq h - \hat{h} \leq 17.205$.
- For the curve 23622G1 our bounds are $-20.056 \leq h - \hat{h} \leq 23.525$ whereas Silverman's bounds are $-19.811 \leq h - \hat{h} \leq 28.082$.

Since our non-archimedean contributions are best possible (in the sense explained in the introduction), it is sensible to compare our archimedean contributions with those given by Silverman's Theorem 15. For the curves in the above range our average archimedean contributions to the lower and upper bounds are respectively -3.483 and 1.029, whereas the corresponding average contributions to Silverman's bounds are -9.011 and 5.048. Our archimedean contributions are better than Silverman's for all except 28 curves in the above range.

Example 4.

It is reasonable to ask how our bounds compare with actual values of the difference $h-\hat{h}$ on rational points. We recall that we have decomposed the difference $h - \hat{h}$ as a weighted sum of continuous bounded functions Ψ_v, with v running over the set of places of M_K. For non-archimedean v the value of Ψ_v depends only on the image of the point in the component group $E(K_v)/E_0(K_v)$. Thus one does not expect our bounds to be sharp unless the map

$$E(K) \to \prod_{v \in M_K^0} E(K_v)/E_0(K_v)$$

is surjective. One the other hand, if the map is surjective, then the total non-archimedean contribution is attained at some rational point, and the only discrepancy that can arise comes from the archimedean contributions which in general are rather small.

Consider for example the elliptic curve $E/\mathbb{Q}$ given by

$$E \; : \quad y^2 = x^3 - 459x^2 - 3478x + 169057.$$

Here all the component groups are trivial, and so we are trivially in the case where the above map is surjective. Our bounds give

$$-6.5319247238\ldots \leq h(P) - \hat{h}(P) \leq 0.4620981203\ldots$$

The elliptic curve E has rank 4; its Mordell-Weil group has basis

$$P_1 = [16, -1], \quad P_2 = [-4, -419], \quad P_3 = [-22, -113], \quad P_4 = [566, -5699].$$

We computed that values of $h(P)-\hat{h}(P)$ on points $P=\sum m_iP_i$ with $|m_i|\le 3$; the maximum and minimum values for $h(P)-\hat{h}(P)$ for these points are as follows:

$$P = 2P_1, \qquad h(P)-\hat{h}(P) = 0.4620980788\ldots,$$
$$P = P_1 - 3P_2 + P_3 + 3P_4, \qquad h(P)-\hat{h}(P) = -4.9001533427\ldots.$$

We leave it to the reader to draw his own conclusions.

References

[1] L. V. Ahlfors, *Complex analysis*, third edition, McGraw–Hill, 1979.

[2] J. E. Cremona, *Algorithms for modular elliptic curves*, second edition, Cambridge University Press, 1996.

[3] J. E. Cremona, Elliptic Curve Data, `http://www.maths.nott.ac.uk/personal/jec/ftp/data/INDEX.html`.

[4] V. A. Demjanenko, *An estimate for the remainder term in Tate's formula*, Mat. Zametki **3** (1968), 271–278. (Russian)

[5] E. Halberstadt and A. Kraus, *Une conjecture de V.A. Lebesgue*, J. London Math. Soc. (2) **69** (2004), no. 2, 291–302.

[6] A. Kondratyev, *Numerical computation of Groebner bases*, Ph.D. thesis, University of Linz, Austria, 2003.

[7] S. Lang, *Conjectured Diophantine estimates on elliptic curves*, Progr. Math. **35** (1983), 155–171.

[8] MAGMA is described in W.Bosma, J.Cannon and C. Playoust: *The Magma algebra system I: The user language*, J. Symbolic Comput. **24**, 235–265 (1997). (Also see the Magma home page at `http://www.maths.usyd.edu.au:8000/u/magma/`.)

[9] Mathematica, Version 5.0, Wolfram Research, Inc., Champaign, IL (2003).

[10] A. Néron, *Quasi-fonctions et hauteurs sur les variétés abéliennes*, Ann. of Math. (2), **82** (1965), 249–331.

[11] PARI/GP, version 2.2.8, Bordeaux, 2004, `http://pari.math.u-bordeaux.fr/`.

[12] S. Siksek, *Infinite descent on elliptic curves*, Rocky Mountain Journal of Mathematics **25**, number 4 (Fall 1990), 1501–1538.

[13] J. H. Silverman, *The difference between the Weil height and the canonical height on elliptic curves*, Math. Comp. **55** (1990), 723-743.

[14] J. H. Silverman, *The arithmetic of elliptic curves*, GTM 106, Springer-Verlag, 1986.

[15] J. H. Silverman, *Advanced topics in the arithmetic of elliptic curves*, GTM 151, Springer-Verlag, 1994.

[16] J. H. Silverman, *Computing heights on elliptic curves*, Math. Comp. **51** (1988), 339-358.

[17] H. G. Zimmer, *On the difference between the Weil height and the Néron–Tate height*, Math. Z. **174** (1976), 35–51.