CANONICAL HEIGHTS ON GENUS TWO JACOBIANS

J. STEFFEN MÜLLER AND MICHAEL STOLL

Scale New Heights!

(Motto of International (now Jacobs) University Bremen,
the place where the first author started his PhD under the supervision of the second author)

ABSTRACT. Let K be a number field and let C/K be a curve of genus 2 with Jacobian variety J. In this paper, we study the canonical height $\hat{h}\colon J(K) \to \mathbb{R}$. More specifically, we consider the following two problems, which are important in applications:

(1) for a given $P \in J(K)$, compute $\hat{h}(P)$ efficiently;
(2) for a given bound $B > 0$, find all $P \in J(K)$ with $\hat{h}(P) \le B$.

We develop an algorithm running in polynomial time (and fast in practice) to deal with the first problem. Regarding the second problem, we show how one can tweak the naive height h that is usually used to obtain significantly improved bounds for the difference $h - \hat{h}$, which allows a much faster enumeration of the desired set of points.

Our approach is to use the standard decomposition of $h(P) - \hat{h}(P)$ as a sum of local 'height correction functions'. We study these functions carefully, which leads to efficient ways of computing them and to essentially optimal bounds. To get our polynomial-time algorithm, we have to avoid the factorization step needed to find the finite set of places where the correction might be nonzero. The main innovation at this point is to replace factorization into primes by factorization into coprimes.

Most of our results are valid for more general fields with a set of absolute values satisfying the product formula.

Date: March 31, 2016.

Contents

1. Introduction

Let K be a global field and let C/K be a curve of genus 2 with Jacobian variety J. There is a map $\kappa\colon J \to \mathbb{P}^3$ that corresponds to the class of twice the theta divisor on J; it identifies a point on J with its negative, and its image is the Kummer surface KS of J. Explicit versions of κ can be found in the book [CF96] by Cassels and Flynn for C given in the form $y^2 = f(x)$ and in the paper [Mül10] by the first author for general C (also in characteristic 2). Thus κ gives rise to a height function $h\colon J(K) \to \mathbb{R}$, which we call the *naive height* on J. It is defined by

$$h(P) = \sum_{v \in M_K} \log \max\{|\kappa_1(P)|_v, |\kappa_2(P)|_v, |\kappa_3(P)|_v, |\kappa_4(P)|_v\}\,,$$

where M_K is the set of places of K, $\kappa(P) = (\kappa_1(P) : \kappa_2(P) : \kappa_3(P) : \kappa_4(P))$, and $|\cdot|_v$ is the v-adic absolute value, normalized so that the product formula

$$\prod_{v \in M_K} |x|_v = 1 \qquad \text{for all } x \in K^\times$$

holds.

By general theory [HS00, Chapter B] the limit

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

exists; it is called the *canonical height* (or *Néron-Tate height*) of $P \in J(K)$. The difference $h - \hat{h}$ is bounded. The canonical height induces a positive definite quadratic form on $J(K)/J(K)_{\mathrm{tors}}$ (and on the $\mathbb{R}$-vector space $J(K) \otimes_{\mathbb{Z}} \mathbb{R}$).

In this paper, we tackle the following two problems:

Problem 1.1. *Find an efficient algorithm for the computation of $\hat{h}(P)$ for a given point $P \in J(K)$.*

Problem 1.2. *Find an efficient algorithm for the enumeration of all $P \in J(K)$ which satisfy $\hat{h}(P) \le B$, where B is a given real number.*

These problems are important because such algorithms are needed if we want to saturate a given finite-index subgroup of $J(K)$ (see the discussion at the end of Section 18). This, in turn, is required for the computation of generators of $J(K)$. Such generators are required, for instance, to carry out the method described in [BMS+08] for the computation of all integral points on a hyperelliptic curve over $\mathbb{Q}$. Furthermore, the regulator of $J(K)$ appearing in the conjecture of Birch and Swinnerton-Dyer is the Gram determinant of a set of generators of $J(K)/J(K)_{\mathrm{tors}}$ with respect to the canonical height. So Problem 1.1 and Problem 1.2 are also important in the context of gathering numerical evidence for this conjecture as in [FLS+01].

It is a classical fact, going back to work by Néron [Nér65], that $\hat{h}(P)$ and the difference $h(P) - \hat{h}(P)$ can be decomposed into a finite sum of local terms. In our situation, this can be done explicitly as follows. The duplication map $P \mapsto 2P$ on J induces a morphism $\delta\colon \mathrm{KS} \to \mathrm{KS}$, given by homogeneous polynomials $(\delta_1, \delta_2, \delta_3, \delta_4)$ of degree 4; explicit equations can again be found in [CF96] and [Mül10]. For a point $Q \in J(K_v)$, where K_v is the completion of K at a place $v \in M_K$, such that $\kappa(Q) = (x_1 : x_2 : x_3 : x_4) \in \mathrm{KS}(K_v)$, we set

$$\tilde{\varepsilon}_v(Q) = -\log \max\{|\delta_j(x_1, x_2, x_3, x_4)|_v : 1 \le j \le 4\} + 4 \log \max\{|x_j|_v : 1 \le j \le 4\}\,.$$

Note that this does not depend on the scaling of the coordinates. We can then write $\hat{h}(P)$ in the following form (compare Lemma 2.4):

$$\hat{h}(P) = h(P) - \sum_{v \in M_K} \sum_{n=0}^{\infty} 4^{-(n+1)} \tilde{\varepsilon}_v(2^n P)$$

We set, for $Q \in J(K_v)$ as above,

$$\tilde{\mu}_v(Q) = \sum_{n=0}^{\infty} 4^{-(n+1)} \tilde{\varepsilon}_v(2^n Q)\,, \tag{1.1}$$

and we deduce the decomposition

$$h(P) - \hat{h}(P) = \sum_{v \in M_K} \tilde{\mu}_v(P)\,, \tag{1.2}$$

which is valid for all points $P \in J(K)$. In addition, $\tilde{\varepsilon}_v = \tilde{\mu}_v = 0$ for all but finitely many v (the exceptions are among the places of bad reduction, the places where the given equation of C is not integral and the archimedean places). The maps $\tilde{\varepsilon}_v \colon J(K_v) \to \mathbb{R}$ are continuous maps (with respect to the v-adic topology) with compact domains, so they are bounded. Therefore $\tilde{\mu}_v$ is also bounded.

Let us first discuss Problem 1.1. Because of equation (1.2), it suffices to compute $h(P)$ (which is easy) and $\sum_{v \in M_K} \tilde{\mu}_v(P)$ in order to compute $\hat{h}(P)$ for a point $P \in J(K)$. Building on earlier work of Flynn and Smart [FS97], the second author introduced an algorithm for the computation of $\tilde{\mu}_v(P)$ in [Sto02]. One of the main problems with this approach is that we need integer factorization to compute the sum $\tilde{\mu}^{\mathrm{f}}(P) := \sum_v \tilde{\mu}_v(P)$, where v runs through the finite primes v such that $\tilde{\mu}_v(P) \neq 0$, because we need to find these primes, or at least a finite set of primes containing them.

We use an idea which was already exploited in [MS15] to find a polynomial-time algorithm for the computation of the canonical height of a point on an elliptic curves (in fact we first used this technique in genus 2 and only later realized that it also works, and is actually easier, for elliptic curves). When v is non-archimedean, then there is a constant $c_v > 0$ such that the function

$$\mu_v := \tilde{\mu}_v / c_v$$

maps $J(K_v)$ to $\mathbb{Q}$. In fact $\tilde{\mu}^{\mathrm{f}}(P)$ is a sum of rational multiples of logarithms of positive integers. As in [MS15], we find a bound on the denominator of μ_v that depends only on the valuation of the discriminant; this allows us to devise an algorithm that computes $\tilde{\mu}^{\mathrm{f}}(P)$ in quasi-linear time. We can compute $\tilde{\mu}_v(P)$ for archimedean v essentially from the definition of $\tilde{\mu}_v$. This leads to a factorization-free algorithm that computes $\hat{h}(P)$ in polynomial time. More precisely:

Theorem 1.3. *Let J be the Jacobian of a curve of genus 2 defined over $\mathbb{Q}$, and let $P \in J(\mathbb{Q})$. There is an algorithm that computes $\hat{h}(P)$ in time quasi-linear in the size of the coordinates of P and the coefficients of the given equation of C, and quasi-quadratic in the desired number of digits of precision.*

See Theorem 14.5 for a precise statement. We expect a similar result to be true for any number field K in place of $\mathbb{Q}$.

We now move on to Problem 1.2. If we have an upper bound β for $h - \hat{h}$, then the set of all points $P \in J(K)$ such that $h(P) \le B + \beta$ contains the set $\{P \in J(K) : \hat{h}(P) \le B\}$. Since the naive height h is a logarithmic height, β contributes exponentially to the size

of the box we need to search for the enumeration. Therefore it is crucial to keep β as small as possible.

We write $\tilde{\beta}_v = \max\{\tilde{\mu}_v(Q) : Q \in J(K_v)\}$, and we obtain the bound

$$h(P) - \hat{h}(P) \leq \sum_{v \in M_K} \tilde{\beta}_v$$

from (1.2). If we write $\tilde{\gamma}_v = \max\{\tilde{\varepsilon}_v(Q) : Q \in J(K_v)\}$, then clearly $\tilde{\gamma}_v/4 \leq \tilde{\beta}_v \leq \tilde{\gamma}_v/3$. In [Sto99], it is shown that for curves given in the form $y^2 = f(x)$, where f has v-adically integral coefficients, we have

$$\tilde{\gamma}_v \leq -\log |2^4 \operatorname{disc}(f)|_v = -\log |2^{-4}\Delta|_v \,,$$

with $\operatorname{disc}(f)$ denoting the discriminant of f considered as a polynomial of degree 6 and Δ denoting the discriminant of the given equation of C. When v is non-archimedean and the normalized additive valuation of Δ is 1, then we can take $\tilde{\gamma}_v = \tilde{\beta}_v = 0$ [Sto02].

The results of the present paper improve on this; they are based on a careful study of the functions $\tilde{\mu}_v$. It turns out that when v is non-archimedean, the set of points where μ_v (or equivalently, $\tilde{\mu}_v$) vanishes forms a group. Moreover, the function μ_v factors through the component group of the Néron model of J when the given model of C/K_v, which we assume to have v-integral coefficients in the following, has rational singularities; see Theorem 7.5. If the minimal regular model of C is semistable, then we can use results of Zhang and Heinz to give explicit formulas for μ_v in terms of the resistance function on the reduction graph of C (which is essentially the dual graph of the special fiber of the minimal regular model, suitably metrized). We use this to find simple explicit formulas for μ_v that apply in the most frequent cases of bad reduction, namely nodal or cuspidal reduction. These explicit formulas give us the optimal bounds for $\tilde{\mu}_v$ in these cases. By reducing to the semistable case and tracking how μ_v changes as we change the Weierstrass equation of C, we deduce the general upper bound

$$\tilde{\beta}_v \leq -\frac{1}{4}\log |\Delta|_v \tag{1.3}$$

for non-archimedean v; see Theorem 11.3.

When v is archimedean, we also get a new bound for $\tilde{\mu}_v$ by iterating the bound obtained by the second author in [Sto99], leading to vast improvements for $\tilde{\beta}_v$. Combining the archimedean and non-archimedean bounds, we find a nearly optimal bound β for $h - \hat{h}$.

To get even smaller search spaces for the enumeration, we make use of the observation that we can replace the naive height h by any function h' such that $|h' - h|$ is bounded. Using the results on nearly optimal bounds for μ_v and such a modified naive height h' (which is also better suited than h for the enumeration process itself) we get a much smaller bound on the difference $h' - \hat{h}$ than what was previously possible. This makes the enumeration feasible in many cases that were completely out of reach so far.

As an example, we compute explicit generators for the Mordell-Weil group of the Jacobian of the curve

$$\begin{aligned} C\colon y^2 = {} & 82342800x^6 - 470135160x^5 + 52485681x^4 \\ & + 2396040466x^3 + 567207969x^2 - 985905640x + 247747600 \end{aligned} \tag{1.4}$$

over $\mathbb{Q}$, conditional on the Generalized Riemann Hypothesis (which is needed to show that the rank is 22). See Proposition 19.1. This curve has at least 642 rational points, which is the current record for the largest number of known rational points on a curve of genus 2, see [Stob].

The paper is divided into four parts. In Part I, we first generalize the usual notion of the naive height on projective space and clarify the relation between these generalized naive heights and suitable canonical heights, all in Section 2. We then introduce local height correction functions ε and μ ($= \mu_v$ in the notation introduced above) on the Jacobian of a genus 2 curve over a non-archimedean local field in Section 3. This is followed in Section 4 by a study of certain canonical local heights constructed in terms of μ. We close Part I by introducing and investigating the notion of stably minimal Weierstrass models of curves of genus 2 in Section 5 and recalling some well-known results on Igusa invariants in Section 6.

Part II is in some sense the central part of the present paper. Here we study the local height correction function μ over a non-archimedean local field. Using Picard functors, we show in Section 7 that μ factors through the component group of the Néron model of the Jacobian when the given model of the curve has rational singularities. We then relate μ to the reduction graph of C in Section 8. Building on this, the following sections contain simple explicit formulas for μ when the reduction of the curve is nodal (Section 9), respectively cuspidal (Section 10). A simple argument then gives the improved general upper bound (1.3) for μ, see Section 11.

In Part III we describe our factorization-free algorithm for the computation of $\hat{h}(P)$ for $P \in J(K)$, where K is a global field. We start in Section 12 by showing how to compute $\mu_v(P)$ for non-archimedean v, using a bound on its denominator. The following section deals with archimedean places, before we finally combine these results in Section 14 into an algorithm for the computation of $\hat{h}(P)$ that runs in polynomial time; this proves Theorem 1.3. Some examples are discussed in Section 15.

In the final Part IV we turn to Problem 1.2. Section 16 contains two methods for bounding $\tilde{\mu}_v$ for archimedean v. In the following Section 17 we describe a modified naive height h' such that the bound on the difference $h' - \hat{h}$ becomes small. We use this, the results of Section 16, and our nearly optimal bounds for the non-archimedean height correction functions from Part II to give an efficient algorithm for the enumeration of the set of rational points with bounded canonical height in Section 18. In the final Section 19 we compute generators of the Mordell-Weil group of the record curve (1.4).

Acknowledgments. We would like to thank David Holmes for suggesting the proof of Lemma 7.1.

Part I: Generalities on Heights and Genus Two Jacobians

2. Generalized naive heights

Let K be a field with a set M_K of places v and associated absolute values $|\cdot|_v$ satisfying the product formula

$$\prod_{v\in M_K} |x|_v = 1 \qquad \text{for all } x \in K^\times \ .$$

We write K_v for the completion of K at v. For a tuple $x = (x_1, \ldots, x_m) \in K_v^m$ we set $\|x\|_v = \max\{|x_1|_v, \ldots, |x_m|_v\}$.

In the following we will introduce some flexibility into our notion of height on projective spaces. (This is similar to the framework of 'admissible families' in [Zar95].)

Definition 2.1.

(1) Let $v \in M_K$. A *local height function* on $\mathbb{P}^m$ at v is a map $h_v\colon K_v^{m+1} \setminus \{0\} \to \mathbb{R}$ such that
   (i) $h_v(\lambda x) = \log|\lambda|_v + h_v(x)$ for all $x \in K_v^{m+1} \setminus \{0\}$ and all $\lambda \in K_v^\times$, and
   (ii) $\big|h_v(x) - \log\|x\|_v\big|$ is bounded.

(2) A function $h\colon \mathbb{P}^m(K) \to \mathbb{R}$ is a *height* on $\mathbb{P}^m$ over K if there are local height functions h_v such that for all $x \in \mathbb{P}^m(K)$ we have

$$h\big((x_1 : x_2 : \ldots : x_{m+1})\big) = \sum_{v\in M_K} h_v(x_1, x_2, \ldots, x_{m+1})$$

and $h_v(x) = \log\|x\|_v$ for all but finitely many places v.

Note that property (i) of local height functions together with the product formula imply that h is invariant under scaling of the coordinates and hence is well-defined.

One example of such a height is the standard height h_{std}, which we obtain by setting $h_v(x) = \log\|x\|_v$ for all v. We then have the following simple fact.

Lemma 2.2. *Let h be any height on $\mathbb{P}^m$ over K and let h_{std} be the standard height. Then there is a constant $c = c(h)$ such that*

$$|h(P) - h_{\mathrm{std}}(P)| \le c \qquad \textit{for all } P \in \mathbb{P}^m(K).$$

Proof. This follows from property (ii) of local height functions. □

Example 2.3. Other examples of heights can be obtained in the following way. For each place v, fix a linear form $l_v(x_1, \ldots, x_{m+1}) = a_{v,1}x_1 + \ldots + a_{v,m+1}x_{m+1}$ with $a_{v,1}, \ldots, a_{v,m+1} \in K_v$ and $a_{v,m+1} \neq 0$, such that $l_v(x) = x_{m+1}$ for all but finitely many v. Then

$$h\big((x_1 : \ldots : x_m : x_{m+1})\big) = \sum_{v\in M_K} \log\max\{|x_1|_v, \ldots, |x_m|_v, |l_v(x_1, \ldots, x_{m+1})|_v\}$$

is a height on $\mathbb{P}^m$.

More generally, we could consider a family of automorphisms A_v of K_v^{m+1} with A_v equal to the identity for all but finitely many v, and take

$$h(x) = \sum_{v\in M_K} \log\max\|A_v(x)\|_v \, .$$

Now consider a variety $V \subset \mathbb{P}^m$ and an endomorphism $\varphi\colon V \to V$ of degree d (i.e., given by homogeneous polynomials of degree d). Then by general theory (see, e.g., [HS00, Thm. B.2.5]) $|h_{\mathrm{std}}(\varphi(P)) - dh_{\mathrm{std}}(P)|$ is bounded on $V(K)$. We write $\varphi^{\circ n}$ for the n-fold iteration of φ. Then the *canonical height*

$$\hat{h}(P) = \lim_{n\to\infty} d^{-n} h_{\mathrm{std}}(\varphi^{\circ n}(P))$$

exists (and satisfies $\hat{h}(\varphi(P)) = d\hat{h}(P)$) [HS00, Thm. B.4.1]. Let h be any height on $\mathbb{P}^m$. Since $|h - h_{\mathrm{std}}|$ is bounded, we can replace h_{std} by h in the definition of $\hat{h}$ without changing the result. We can then play the usual telescoping series trick in our more general setting.

Lemma 2.4. *Let $\varphi\big((x_1 : \ldots : x_{m+1})\big) = \big(\varphi_1(x) : \ldots : \varphi_{m+1}(x)\big)$ with homogeneous polynomials $\varphi_j \in K[x_1, \ldots, x_{m+1}]$ of degree d. We have*

$$\hat{h}(P) = h(P) - \sum_{v\in M_K} \tilde{\mu}_v(P)\,,$$

where

$$\tilde{\mu}_v(P) = \sum_{n=0}^{\infty} d^{-(n+1)} \tilde{\varepsilon}_v(\varphi^{\circ n}(P))$$

and, when $P = (x_1 : \ldots : x_{m+1})$ and $x = (x_1, \ldots, x_{m+1})$,

$$\tilde{\varepsilon}_v(P) = dh_v(x) - h_v\big(\varphi_1(x), \ldots, \varphi_{m+1}(x)\big)\,.$$

Proof. Note that $\tilde{\varepsilon}_v$ is well-defined: scaling x by λ adds $|\lambda|_v$ to $h_v(x)$ and $d|\lambda|_v$ to $h_v(\varphi_1(x), \ldots, \varphi_{m+1}(x))$. Let x be projective coordinates for P and write $x^{(n)}$ for the result of applying $(\varphi_1, \ldots, \varphi_{m+1})$ n times to $x = x^{(0)}$. Then

$$\begin{aligned}
\hat{h}(P) &= \lim_{n\to\infty} d^{-n} h(\varphi^{\circ n}(P)) \\
&= h(P) + \sum_{n=0}^{\infty} d^{-(n+1)} \big(h(\varphi^{\circ(n+1)}(P)) - dh(\varphi^{\circ n}(P))\big) \\
&= h(P) + \sum_{n=0}^{\infty} d^{-(n+1)} \sum_{v\in M_K} \big(h_v(x^{(n+1)}) - dh_v(x^{(n)})\big) \\
&= h(P) - \sum_{v\in M_K} \sum_{n=0}^{\infty} d^{-(n+1)} \tilde{\varepsilon}_v(\varphi^{\circ n}(P)) \\
&= h(P) - \sum_{v\in M_K} \tilde{\mu}_v(P)\,.
\end{aligned}$$

□

We call the functions $\tilde{\mu}_v\colon \mathbb{P}^m(K_v) \to \mathbb{R}$ *local height correction functions.*

Note that when K_v is a discretely valued field such that $|x|_v = \exp(-c_v v(x))$ for $x \in K^\times$ with a constant $c_v > 0$ (and where we abuse notation and write $v\colon K_v^\times \twoheadrightarrow \mathbb{Z}$ also for the normalized additive valuation associated to the place v) and $h = h_{\mathrm{std}}$, then we have

$$\tilde{\mu}_v(P) = c_v \mu_v(P) \qquad \text{and} \qquad \tilde{\varepsilon}_v(P) = c_v \varepsilon_v(P)\,,$$

where

$$\mu_v(P) = \sum_{n=0}^{\infty} d^{-(n+1)} \varepsilon_v(P)$$

and

$$\varepsilon_v(P) = \min\big\{v(\varphi_1(x)), \ldots, v(\varphi_{m+1}(x))\big\} - d\min\{v(x_1), \ldots, v(x_{m+1})\}\,,$$

if $x = (x_1, \dots, x_{m+1})$ are homogeneous coordinates for P. This is the situation that we will study in some detail in Part II of this paper, for the special case when $V \subset \mathbb{P}^3$ is the Kummer surface associated to a curve of genus 2 and its Jacobian J and φ is the duplication map (then $d = 4$).

To deal with Problem 1.1, we work with the standard height h_{std}. We use our detailed results on the local height correction functions to deduce a bound on the denominator of μ_v (its values are rational) in terms of the valuation of the discriminant of the curve. This is the key ingredient that leads to our new factorization-free and fast algorithm for computing $\hat{h}$, see Part III.

To deal with Problem 1.2, we use the flexibility in choosing the (naive) height h and modify the standard height in such a way that the sum $\sum_{v \in M_K} \sup \tilde{\mu}_v(J(K_v))$ that bounds the difference $h - \hat{h}$ is as small as we can make it. The local height functions we use are as in Example 2.3 above, with $l_v(x_1, x_2, x_3, x_4) = x_4/s_v$ for certain $s_v \in K_v^\times$ in most cases. This choice has the property that for any $P = (x_1 : x_2 : x_3 : x_4) \in \mathbb{P}^3(K)$ different from $(0 : 0 : 0 : 1)$ we have

$$0 \le h_{\mathrm{std}}\big((x_1 : x_2 : x_3)\big) \le h(P)\,.$$

This is relevant, since we can fairly easily enumerate all points P as above that are on the Kummer surface and satisfy $h_{\mathrm{std}}\big((x_1 : x_2 : x_3)\big) \le B$, see Part IV. Refinements of the standard height constructed using Arakelov theory were also used by Holmes [Hol14] to given an 'in principle' algorithm for the enumeration of points of bounded canonical heights on Jacobians of hyperelliptic curves over global fields.

3. Local height correction functions for genus 2 Jacobians

Until further notice, we let k be a non-archimedean local field with additive valuation v, normalized to be surjective onto $\mathbb{Z}$. Let $\mathcal{O}$ denote the valuation ring of k with residue class field $\mathfrak{k}$ and let π be a uniformizing element of $\mathcal{O}$. We consider a smooth projective curve C of genus 2 over k, given by a Weierstrass equation

$$Y^2 + H(X, Z)Y = F(X, Z) \tag{3.1}$$

in weighted projective space $\mathbb{P}_k(1, 3, 1)$, with weights 1, 3 and 1 assigned to the variables X, Y and Z, respectively. Here

$$F(X, Z) = f_0 Z^6 + f_1 X Z^5 + f_2 X^2 Z^4 + f_3 X^3 Z^3 + f_4 X^4 Z^2 + f_5 X^5 Z + f_6 X^6$$

and

$$H(X, Z) = h_0 Z^3 + h_1 X Z^2 + h_2 X^2 Z + h_3 X^3$$

are binary forms of degrees 6 and 3, respectively, such that the discriminant $\Delta(F, H)$ of the Weierstrass equation (3.1) is nonzero. In characteristic different from 2, this discriminant is defined as

$$\Delta(F, H) = 2^{-12} \operatorname{disc}(4F + H^2) \in \mathbb{Z}[h_0, \dots, h_3, f_0, \dots, f_6]\,,$$

and in general, we define it by the generic polynomial given by this formula. The curve defined by the equation is smooth if and only if $\Delta(F, H) \neq 0$.

For the remainder of this section we assume that $F, H \in \mathcal{O}[X, Z]$, so that equation (3.1) defines an *integral Weierstrass model* $\mathcal{C}$ of the curve in the terminology of Section 5 below. The discriminant of this model is then defined to be $\Delta(\mathcal{C}) := \Delta(F, H)$. We may assume that C is given by such an integral equation if k is the completion at a non-archimedean place of a number field K and C is obtained by base change from K, since we can choose a globally integral Weierstrass equation for the curve. But also in

general, we can always assume that C is given by an integral equation after applying a transformation defined over k, since we know from Corollary 4.6 in the next section how the local height correction function μ defined in Definition 3.1 below behaves under such transformations.

We now generalize the definition of ε given in [Sto02] to our more general setting ([Sto02] works with Weierstrass equations that have $H = 0$). As in the introduction, let J denote the Jacobian of C and let KS be its Kummer surface, constructed explicitly together with an explicit embedding into $\mathbb{P}^3$ in [CF96] in the case $H = 0$ and in [Mül10] in the general case. Also let $\kappa\colon J \to \mathbb{P}^3$ denote the composition of the quotient map from J to KS with this embedding; it maps the origin $O \in J(k)$ to the point $(0 : 0 : 0 : 1)$. A quadruple $x = (x_1, x_2, x_3, x_4) \in k^4$ is called a set of *Kummer coordinates* on KS if x is a set of projective coordinates for a point in $\mathrm{KS}(k)$; we denote the set of sets of Kummer coordinates on KS by $\mathrm{KS}_{\mathbb{A}}$ (this is the set of k-rational points on the pointed affine cone over KS). For $x \in \mathrm{KS}_{\mathbb{A}}$ we write $v(x) = \min\{v(x_1), \dots, v(x_4)\}$, and we say that x is *normalized* if $v(x) = 0$. If $P \in J(k)$, we say that $x \in \mathrm{KS}_{\mathbb{A}}$ is a set of *Kummer coordinates for* P if $\kappa(P) = (x_1 : x_2 : x_3 : x_4)$.

We let δ denote the duplication map on KS, which is given by homogeneous polynomials $\delta_1, \dots, \delta_4 \in \mathcal{O}[x_1, \dots, x_4]$ of degree 4 such that $\delta(0, 0, 0, 1) = (0, 0, 0, 1)$. We recall that there is a symmetric matrix $B = (B_{ij})_{1 \le i,j \le 4}$, where the $B_{ij} \in \mathcal{O}[x_1, \dots, x_4, y_1, \dots, y_4]$ are bi-homogeneous of degree 2 in $x_1, \dots, x_4$ and $y_1, \dots, y_4$ each and have the following properties, see [CF96, Chapter 3] and [Mül10].

(i) Let $x, y \in \mathrm{KS}_{\mathbb{A}}$ be Kummer coordinates for $P, Q \in J(k)$. Then there are Kummer coordinates $w, z \in \mathrm{KS}_{\mathbb{A}}$ for $P + Q$ and $P - Q$, respectively, such that
$$w * z := (w_i z_j + n_{ij} w_j z_i)_{1 \le i,j \le 4} = B(x, y)$$
and hence $v(w) + v(z) = v\big(B(x, y)\big)$; here $n_{ij} = 1$ if $i \ne j$ and $n_{ij} = 0$ if $i = j$.
(ii) If $x \in \mathrm{KS}_{\mathbb{A}}$, then $B(x, x) = \delta(x) * (0, 0, 0, 1)$.

We specialize the notions introduced in Section 2 to our situation: we consider the Kummer surface $\mathrm{KS} \subset \mathbb{P}^3$ with the duplication map δ of degree $d = 4$. We use the standard local height on $\mathbb{P}^3$.

Definition 3.1. Let $x \in \mathrm{KS}_{\mathbb{A}}$ be a set of Kummer coordinates on KS. Then we set
$$\varepsilon(x) = v(\delta(x)) - 4v(x) \in \mathbb{Z} \qquad \text{and} \qquad \mu(x) = \sum_{n=0}^{\infty} \frac{1}{4^{n+1}} \varepsilon(\delta^{\circ n}(x)) \,,$$
where $\delta^{\circ n}$ denotes the n-fold composition $\delta \circ \ldots \circ \delta$.

Because δ is given by homogeneous polynomials of degree 4, $\varepsilon(x)$ does not depend on the scaling of x, so it makes sense to define $\varepsilon(P) = \varepsilon(x)$ for points $P \in \mathrm{KS}(k)$, where $x \in \mathrm{KS}_{\mathbb{A}}$ is any set of Kummer coordinates for P, and to define $\varepsilon(P) = \varepsilon(\kappa(P))$ for points $P \in J(k)$. We likewise extend the definition of μ. Then we have
$$\mu(2P) - 4\mu(P) = -\varepsilon(P) \qquad \text{for all } P \in J(k).$$
Note that our assumption $F, H \in \mathcal{O}[X, Z]$ implies that $\varepsilon \ge 0$. If k is a local field (as we assume here), then $\mathrm{KS}(k)$ is compact in the v-adic topology, and ε is continuous, so ε is bounded.

Remark 3.2. More generally, if k is a field with a discrete valuation and not of characteristic 2, then the arguments in [Sto99] show that when $H = 0$, $\varepsilon \le v(2^4 \operatorname{disc}(F))$, so ε is bounded also for these more general fields.

If k is any field with a discrete valuation, then one can still conclude that ε is bounded, by making use of the fact that the duplication map is well-defined on KS, which implies that the ideal generated by the δ_j and the polynomial δ_0 defining KS contains a power of the irrelevant ideal. So for some $N > 0$, one can express every x_j^N as a linear combination of $\delta_0(x), \ldots, \delta_4(x)$ with coefficients that are homogeneous polynomials of degree $N-4$ with coefficients in k. The negative of the minimum of the valuations of these coefficients then gives a bound for ε.

Remark 3.3. If k is the completion of a global field at a place v, then for $\alpha \in k^\times$, $v(\alpha)/\log\|\alpha\|_v = -c_v$ is a negative constant. So for $P \in J(k)$ we have $\varepsilon(P) = c_v\tilde{\varepsilon}_v(P)$ and $\mu(P) = c_v\tilde{\mu}_v(P)$, where $\tilde{\varepsilon}_v$ and $\tilde{\mu}_v$ are as defined in the introduction.

We will also have occasion to use the following function. Let $x, y \in \mathrm{KS}_{\mathbb{A}}$ and define

$$\varepsilon(x,y) = v(B(x,y)) - 2v(x) - 2v(y). \tag{3.2}$$

In the same way as for $\varepsilon(x)$ above, we can extend this to points in $\mathrm{KS}(k)$ and $J(k)$.

Lemma 3.4. *Let $x, y, w, z \in \mathrm{KS}_{\mathbb{A}}$ be Kummer coordinates satisfying $w * z = B(x,y)$. Then we have*

$$\delta(w) * \delta(z) = B(\delta(x), \delta(y))\,.$$

Proof. The proof carries over verbatim from the proof of [Sto02, Lemma 3.2]. □

We deduce the following:

Lemma 3.5. *Let $x, y, w, z \in \mathrm{KS}_{\mathbb{A}}$ be Kummer coordinates satisfying $w * z = B(x,y)$. Then we have*

$$\varepsilon\big(\delta(x), \delta(y)\big) + 2\varepsilon(x) + 2\varepsilon(y) = \varepsilon(w) + \varepsilon(z) + 4\varepsilon(x,y)\,.$$

Proof. Using Lemma 3.4, relation (3.2), and property (i) above for $\delta(w)$, $\delta(z)$, $\delta(x)$ and $\delta(y)$, we obtain

$$v\big(\delta(w)\big) + v\big(\delta(z)\big) = v\big(B(\delta(x), \delta(y))\big) = \varepsilon\big(\delta(x), \delta(y)\big) + 2v\big(\delta(x)\big) + 2v\big(\delta(y)\big)\,.$$

Subtracting four times the corresponding relation for w, z, x and y, we get

$$\varepsilon(w) + \varepsilon(z) = \varepsilon\big(\delta(x), \delta(y)\big) - 4\varepsilon(x,y) + 2\varepsilon(x) + 2\varepsilon(y)\,,$$

which is the claim. □

We state a few general facts on the functions ε and μ.

Lemma 3.6. *For points $P, Q \in J(k)$, we have the relation*

$$\mu(P+Q) + \mu(P-Q) - 2\mu(P) - 2\mu(Q) = -\varepsilon(P,Q)\,.$$

Proof. Let x and y be Kummer coordinates for P and Q, respectively; then w and z as in Lemma 3.5 are Kummer coordinates for $P+Q$ and $P-Q$ (in some order). The claim

now follows from the formula in Lemma 3.5:

$$\begin{aligned}
&\mu(P+Q)+\mu(P-Q)-2\mu(P)-2\mu(Q)\\
&\qquad=\sum_{n=0}^{\infty} 4^{-n-1}\bigl(\varepsilon(2^nP+2^nQ)+\varepsilon(2^nP-2^nQ)-2\varepsilon(2^nP)-2\varepsilon(2^nQ)\bigr)\\
&\qquad=\sum_{n=0}^{\infty} 4^{-n-1}\bigl(\varepsilon(\delta^{\circ n}(w))+\varepsilon(\delta^{\circ n}(z))-2\varepsilon(\delta^{\circ n}(x))-2\varepsilon(\delta^{\circ n}(y))\bigr)\\
&\qquad=\sum_{n=0}^{\infty} 4^{-n-1}\bigl(\varepsilon(\delta^{\circ(n+1)}(x),\,\delta^{\circ(n+1)}(y))-4\varepsilon(\delta^{\circ n}(x),\,\delta^{\circ n}(y))\bigr)\\
&\qquad=-\varepsilon(x,y)=-\varepsilon(P,Q)\,.
\end{aligned}$$

□

Lemma 3.7. *If $P\in J(k)$ satisfies $\mu(P)=0$, then $\mu(P+Q)=\mu(Q)$ for all $Q\in J(k)$.*

Proof. We take $(P,Q)\leftarrow(Q+nP,P)$ in Lemma 3.6, where $n\in\mathbb{Z}$. Taking into account that $\mu(P)=0$ and writing a_n for $\mu(Q+nP)$, this gives

$$a_{n+1}-2a_n+a_{n-1}=-\varepsilon(P,Q+nP)\,.$$

As k is a non-archimedean local field, the multiples of P accumulate at the origin $O\in J(k)$. Recall that ε is locally constant. This implies that every value $\varepsilon(P,Q+nP)$ occurs for infinitely many $n\in\mathbb{Z}$, since $Q+(n+N)P$ will be close to $Q+nP$ for suitably chosen N. We have for any $m>0$

$$a_{m+1}-a_m-a_{-m}+a_{-m-1}=\sum_{n=-m}^{m}(a_{n+1}-2a_n+a_{n-1})=-\sum_{n=-m}^{m}\varepsilon(P,Q+nP)\,.$$

Since μ is bounded, the left hand side is bounded independently of m. We also know that $\varepsilon(P,Q+nP)\ge 0$. But if $\varepsilon(P,Q+nP)$ were nonzero for some n, then by the discussion above, the right hand side would be unbounded as $m\to\infty$. Therefore it follows that $\varepsilon(P,Q+nP)=0$ for all $n\in\mathbb{Z}$. This in turn implies $a_{n+1}-2a_n+a_{n-1}=0$ for all $n\in\mathbb{Z}$. The only bounded solutions of this recurrence are constant sequences. In particular, we have

$$\mu(P+Q)=a_1=a_0=\mu(Q)\,.$$

□

Proposition 3.8. *The subset $U=\{P\in J(k):\mu(P)=0\}$ is a subgroup of finite index in $J(k)$. The functions $P\mapsto\varepsilon(P)$ and $P\mapsto\mu(P)$ factor through the quotient $J(k)/U$.*

Proof. Lemma 3.7 shows that U is a subgroup. We have $\varepsilon(P)=0$ for $P\in J(k)$ sufficiently close to the origin. So taking a sufficiently small subgroup neighborhood U' of the origin in $J(k)$, we see that $\varepsilon(2^nP)=0$ for all $P\in U'$ and all $n\ge 0$. This implies that $\mu=0$ on U', so $U\supset U'$. Because k is a local field, U' and therefore also U have finite index in $J(k)$. By Lemma 3.7 again, μ factors through $J(k)/U$, and since $\varepsilon(P)=4\mu(P)-\mu(2P)$, the same is true for ε. □

We will now show that we actually have

$$U=\{P\in J(k):\varepsilon(P)=0\}$$

(the inclusion '$\subset$' is clear from the definition and Proposition 3.8.) This is equivalent to the implication $\varepsilon(x)=0\Longrightarrow\varepsilon(\delta(x))=0$ and generalizes [Sto02, Thm. 4.1]. For this we first provide a characteristic 2 analogue of [Sto02, Prop. 3.1(1)].

We temporarily let k denote an arbitrary field. Let $C_{F,H}$ be a (not necessarily smooth) curve in the weighted projective plane with respective weights 1, 3, 1 assigned to the variables X, Y, Z that is given by an equation

$$Y^2 + H(X,Z)Y = F(X,Z), \tag{3.3}$$

where $F, H \in k[X,Z]$ are binary forms of respective degrees 6 and 3. Let $\mathrm{KS}_{F,H}$ denote the subscheme of $\mathbb{P}^3$ given by the vanishing of the equation defining the Kummer surface of $C_{F,H}$ if $C_{F,H}$ is nonsingular. Then the construction of $\delta = (\delta_1, \delta_2, \delta_3, \delta_4)$ still makes sense in this context, but we may now have $\delta_i(x) = 0$ for all $1 \le i \le 4$ (which we abbreviate by $\delta(x) = 0$) for a set x of Kummer coordinates on $\mathrm{KS}_{F,H}$. We generalize Proposition 3.1 in [Sto02] (which assumes $H = 0$) to the case considered here.

Note that two equations (3.3) for $C_{F,H}$ are related by a transformation τ acting on an affine point (ξ, η) by

$$\tau(\xi,\eta) = \left(\frac{a\xi + b}{c\xi + d}, \frac{e\eta + U(\xi,1)}{(c\xi+d)^3} \right), \tag{3.4}$$

where $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(k)$, $e \in k^\times$ and $U \in k[X,Z]$ is homogeneous of degree 3. The transformation τ also acts on the forms F and H by

$$\begin{aligned} \tau^* F(X,Z) &= (ad-bc)^{-6} \left(e^2 F^A + (eH^A - U^A) U^A \right) \\ \tau^* H(X,Z) &= (ad-bc)^{-3} \left(eH^A - 2U^A \right), \end{aligned}$$

where we write

$$S^A = S(dX - bZ, -cX + aZ)$$

for a binary form $S \in k[X,Z]$.

Lemma 3.9. *Let $x \in \mathrm{KS}_{F,H}(k)$. If $\delta(\delta(x)) = 0$, then we already have $\delta(x) = 0$.*

Proof. If k has characteristic different from 2, we can apply a transformation so that the new Weierstrass equation will have $H = 0$; the statement is then [Sto02, Prop. 3.1(1)]. So from now on, k has characteristic 2. We may assume without loss of generality that k is algebraically closed. If the given curve is smooth, then the result is obvious, because the situation described in the statement can never occur. If it is not smooth, we can act on F and H using transformations of the form (3.4), so it is enough to consider only one representative of each orbit under such transformations. This is analogous to the strategy in the proof of [Sto02, Prop. 3.1]. We can, for example, pick the representatives listed in Table 1.

For these representatives, elementary methods as in the proof of [Sto02, Prop. 3.1] can be used to check that $\delta(x) = 0$ indeed follows from $\delta(\delta(x)) = 0$. □

We can use the above to analyze the group U.

Theorem 3.10. *Suppose that k is a non-archimedean local field and that J is the Jacobian of a smooth projective curve of genus 2, given by a Weierstrass equation (3.1) with integral coefficients. Then the set $\{P \in J(k) : \varepsilon(P) = 0\}$ equals the subgroup U in Proposition 3.8. In particular, U is a subgroup of finite index in $J(k)$ and ε and μ factor through the quotient $J(k)/U$. Moreover we have that $\varepsilon(-P) = \varepsilon(P)$ and U contains the kernel of reduction $J(k)^1$ with respect to the given model of J, i.e., the subgroup of points whose image in $\mathrm{KS}(\mathfrak{k})$ equals that of O.*

type	H	F	conditions
1	0	0	
2	Z^3	0	
3	Z^3	aXZ^5	$a \neq 0$
4	XZ^2	aXZ^5	$a \neq 0$
5	XZ^2	bX^3Z^3	$b \neq 0$
6	Z^3	$aXZ^5 + bX^3Z^3$	$ab \neq 0$
7	XZ^2	0	
8	$XZ(X+Z)$	0	
9	$XZ(X+Z)$	bX^3Z^3	$b(b+1) \neq 0$
10	$XZ(X+Z)$	$aXZ^5 + bX^3Z^3$	$a(a+b)(a+b+1) \neq 0$
11	XZ^2	$aXZ^5 + bX^3Z^3$	$ab \neq 0$
12	0	XZ^5	
13	0	X^3Z^3	

TABLE 1. Representatives in characteristic 2

Proof. The statement in Lemma 3.9 implies $\varepsilon(P) = 0 \Longrightarrow \varepsilon(2P) = 0$ for points $P \in J(k)$, since $\varepsilon(P) = 0$ is equivalent to $\delta(\tilde{x}) \neq 0$ if x are normalized Kummer coordinates for P, with reduction $\tilde{x}$. This shows that $\varepsilon(P) = 0$ implies $\mu(P) = 0$ (and conversely), so $\{P \in J(k) : \varepsilon(P) = 0\} = \{P \in J(k) : \mu(P) = 0\} = U$. The remaining statements now are immediate from Proposition 3.8, taking into account that for P in the kernel of reduction, we trivially have $\varepsilon(P) = 0$. □

An algorithm for the computation of $\mu(P)$ which is based on Theorem 3.10 (for $H = 0$) is given in [Sto02, §6]. Using the relation in Lemma 3.6, we obtain the following alternative procedure for computing $\mu(P)$.

1. Let x be normalized Kummer coordinates for P.
   Set $y_0 = (0, 0, 0, 1)$ and $y_1 = x$.
2. For $n = 1, 2, \ldots$, do the following.
   a. Using pseudo-addition (see [FS97, §4]), compute normalized Kummer coordinates y_{n+1} for nP from x, y_{n-1} and y_n; record $\varepsilon(P, nP)$, which is the shift in valuation occurring when normalizing y_{n+1}.
   b. If $\varepsilon(P, nP) = 0$, check whether $v(\delta(y_n)) = 0$ (by Theorem 3.10, this is equivalent to $nP \in U$). If yes, let $N = n$ and exit the loop.
3. Return
$$\mu(P) = \frac{1}{2N} \sum_{n=1}^{N-1} \varepsilon(P, nP)\,.$$

To see that this works, note that by Lemma 3.6 we have

$$\mu\big((n+1)P\big) - 2\mu(nP) + \mu\big((n-1)P\big) = 2\mu(P) - \varepsilon(P, nP)\,.$$

The sequence $\big(\mu(nP)\big)_{n\in\mathbb{Z}}$ is periodic with period N, where N is the smallest positive integer n such that $nP \in U$ (which exists according to Theorem 3.10). Taking the sum over one period gives

$$2N\mu(P) = \sum_{n=0}^{N-1} \varepsilon(P, nP) = \sum_{n=1}^{N-1} \varepsilon(P, nP)\,.$$

From the periodicity we can also deduce the possible denominators of $\mu(P)$. As ε has integral values, we see that $\mu(P) \in \frac{1}{2N}\mathbb{Z}$ if N is a period of $\big(\mu(nP)\big)_{n\in\mathbb{Z}}$. In fact, we can show a little bit more.

Corollary 3.11. *Let $P \in J(k)$ and $N = \min\{n \in \mathbb{Z}_{>0} : \mu(nP) = 0\}$. Then*

$$\mu(P) \in \frac{1}{N}\mathbb{Z} \quad \text{if } N \text{ is odd, and}$$
$$\mu(P) \in \frac{1}{2N}\mathbb{Z} \quad \text{if } N \text{ is even.}$$

Proof. The sequence $\big(\varepsilon(P, nP)\big)_{n\in\mathbb{Z}}$ has period N and is symmetric. So if N is odd, we actually have

$$\mu(P) = \frac{1}{2N}\sum_{n=1}^{N-1} \varepsilon(P, nP) = \frac{1}{N}\sum_{n=1}^{(N-1)/2} \varepsilon(P, nP) \in \frac{1}{N}\mathbb{Z}\,. \qquad \square$$

Analyzing the possible denominators of $\mu(P)$ will play a key role in Section 12, where we discuss another algorithm for the computation of $\mu(P)$.

4. Canonical local heights on Kummer coordinates

We now define a notion of canonical local height for Kummer coordinates. We keep the notation of the previous section.

Definition 4.1. Let $x \in \mathrm{KS}_{\mathbb{A}}$ be a set of Kummer coordinates on KS. The *canonical local height of x* is given by

$$\hat{\lambda}(x) = -v(x) - \mu(x)\,.$$

Remark 4.2. We can also define the canonical local height on an archimedean local field in an analogous way. Then, if K is a global field and x is a set of Kummer coordinates for a point $J(K)$, we have

$$\hat{h}(P) = \sum_{v\in M_K} \frac{1}{c_v}\hat{\lambda}_v(x)\,,$$

where c_v is the constant introduced in Remark 3.3 for a non-archimedean place v and $c_v = [K_v : \mathbb{R}]^{-1}$ if v is archimedean.

The canonical local height $\hat{\lambda}$ on Kummer coordinates has somewhat nicer properties than the canonical local height defined (for instance in [FS97] or, more generally, in [HS00, §B.9]) with respect to a divisor on J.

Proposition 4.3. *Let $x, y, z, w \in \mathrm{KS}_{\mathbb{A}}$. Then the following hold:*

(i) $\hat{\lambda}(\delta(x)) = 4\hat{\lambda}(x)$.
(ii) *If $w * z = B(x, y)$, then* $\hat{\lambda}(z) + \hat{\lambda}(w) = 2\hat{\lambda}(x) + 2\hat{\lambda}(y)$.
(iii) $\hat{\lambda}(x) = -\lim_{n\to\infty} 4^{-n} v\big(\delta^{\circ n}(x)\big)$.
(iv) *If k'/k is a finite extension of ramification index e and $\hat{\lambda}'$ is the canonical local height over k', then we have* $\hat{\lambda}'(x) = e \cdot \hat{\lambda}(x)$.

Proof.

(i) This follows easily from the two relations

$$v\big(\delta(x)\big) = 4v(x) + \varepsilon(x) \quad \text{and} \quad \mu\big(\delta(x)\big) = 4\mu(x) - \varepsilon(x)\,.$$

(ii) This is similar, using Lemma 3.6 and $\varepsilon(x, y) = v(w) + v(z) - 2v(x) - 2v(y)$.

(iii) This follows from (i) and the fact that $\mu(x)$ is a bounded function, implying
$$\hat\lambda(x) = 4^{-n}\hat\lambda\bigl(\delta^{\circ n}(x)\bigr) = -4^{-n}v\bigl(\delta^{\circ n}(x)\bigr) + O(4^{-n})\,.$$

(iv) This is obvious from the definition of $\hat\lambda$. □

The canonical local height on Kummer coordinates also behaves well under isogenies.

Proposition 4.4. *Let $\alpha\colon J \to J'$ be an isogeny of Jacobians of dimension 2 defined over k. Then α induces a map $\alpha\colon \mathrm{KS} \to \mathrm{KS}'$ between the corresponding Kummer surfaces; let d denote its degree. We also get a well-defined induced map $\alpha\colon \mathrm{KS}_{\mathbb{A}} \to \mathrm{KS}'_{\mathbb{A}}$ if we fix $a \in k^\times$ and require $\alpha(0,0,0,1) = (0,0,0,a)$. Then we have*
$$\hat\lambda\bigl(\alpha(x)\bigr) = d\hat\lambda(x) - v(a)$$
for all $x \in \mathrm{KS}_{\mathbb{A}}$.

Proof. All assertions except for the last one are obvious. By the definition of $\hat\lambda$, we can reduce to the case $a = 1$. Using part (iii) of Proposition 4.3 it is then enough to show that
$$v\bigl(\delta^{\circ n}(\alpha(x))\bigr) = dv(\delta^{\circ n}(x)) + O(1)\,.$$
However, we have $v(\alpha(x)) - dv(x) = O(1)$ by assumption, so it suffices to show that
$$v\bigl(\delta^{\circ n}(\alpha(x))\bigr) = v\bigl(\alpha(\delta^{\circ n}(x))\bigr)\,. \tag{4.1}$$
But since $\alpha\colon J \longrightarrow J'$ is an isogeny, $\delta^{\circ n}(\alpha(x))$ and $\alpha(\delta^{\circ n}(x))$ represent the same point on KS', hence they are projectively equal. Because they also have the same degree, the factor of proportionality is independent of x. It therefore suffices to check (4.1) for a single x; we take $x = (0,0,0,1) \in \mathrm{KS}_{\mathbb{A}}$. Because we have $\delta(x) = x$ and, by assumption, $\alpha(x) = x'$, where $x' = (0,0,0,1) \in \mathrm{KS}'_{\mathbb{A}}(k)$, we find
$$\delta^{\circ n}(\alpha(x)) = x' \quad\text{and}\quad \alpha(\delta^{\circ n}(x)) = x'\,,$$
thereby proving (4.1) and hence the proposition. □

Remark 4.5. Canonical local heights with similar functorial properties were constructed by Zarhin [Zar95] on total spaces of line bundles (without the zero section). See also [BG06] for an approach to canonical local heights using rigidified metrized line bundles.

The preceding proposition is particularly useful for analyzing the behavior of the canonical local height under a change of Weierstrass equation of the curve.

Recall that two Weierstrass equations for C are related by a transformation τ as in (3.4), specified by a triple (A, e, U), where $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(k)$, $e \in k^\times$ and
$$U = u_0Z^3 + u_1XZ^2 + u_2X^2Z + u_3X^3 \in k[X,Z]$$
is homogeneous of degree 3. Such a transformation induces a map on $\mathrm{KS}_{\mathbb{A}}$ as follows: Let $x = (x_1, x_2, x_3, x_4) \in \mathrm{KS}_{\mathbb{A}}$. Then $\tau(x)$ is given by the following quadruple:
$$\begin{aligned}(ad-bc)^{-1}\Bigl(&d^2x_1 + cdx_2 + c^2x_3,\\ &2bdx_1 + (ad+bc)x_2 + 2acx_3,\\ &b^2x_1 + abx_2 + a^2x_3,\\ &(ad-bc)^{-2}(e^2x_4 + l_1x_1 + l_2x_2 + l_3x_3)\Bigr)\,,\end{aligned}$$
where l_1, l_2, l_3 do not depend on x. More precisely, we can write
$$l_i = l_{i,1} + l_{i,2} + l_{i,3}\,,$$

where

$$l_{i,1} = \frac{e^2}{(ad-bc)^4} l'_{i,1} \quad \text{with } l'_{i,1} \in \mathbb{Z}[f_0, \dots, f_6, a, b, c, d],$$
$$l_{i,2} = \frac{e}{(ad-bc)^4} l'_{i,2} \quad \text{with } l'_{i,2} \in \mathbb{Z}[h_0, \dots, h_3, u_0, \dots, u_3, a, b, c, d],$$
$$l_{i,3} = \frac{1}{(ad-bc)^4} l'_{i,3} \quad \text{with } l'_{i,3} \in \mathbb{Z}[u_0, \dots, u_3, a, b, c, d]$$

for $i = 1, 2, 3$. All of the $l'_{i,j}$ are homogeneous of degree 8 in a, b, c, d and homogeneous in the other variables.

So we see that τ acts on k^4 as a linear map whose determinant has valuation

$$v(\tau) = 2v(e) - 3v(ad - bc)\,.$$

In this situation, Proposition 4.4 implies:

Corollary 4.6. *Let $\tau = ([a, b, c, d], e, U)$ be a transformation (3.4) between two Weierstrass equations $\mathcal{C}$ and $\mathcal{C}'$ of a smooth projective curve C/k of genus 2 and let* KS *be the model of the Kummer surface associated to $\mathcal{C}$. Then we have*

$$\hat{\lambda}(\tau(x)) = \hat{\lambda}(x) - v(\tau)$$

for all $x \in \mathrm{KS}_{\mathbb{A}}$. In particular,

$$\mu(x) = \mu(\tau(x)) + v(\tau(x)) - v(x) - v(\tau)\,.$$

This can be used to construct a canonical local height which does not depend on the choice of Weierstrass equation.

Definition 4.7. Let C/k be a smooth projective curve of genus 2 given by a Weierstrass equation (3.1) with discriminant Δ and let KS be the associated Kummer surface. We call the function

$$\tilde{\lambda}\colon \mathrm{KS}_{\mathbb{A}} \longrightarrow \mathbb{R}\,, \qquad x \longmapsto \hat{\lambda}(x) + \frac{1}{10} v(\Delta)$$

the *normalized canonical local height on* $\mathrm{KS}_{\mathbb{A}}$.

Corollary 4.8. *The normalized canonical local height is independent of the given Weierstrass equation of C.*

Proof. Let τ be a transformation (3.4) between two Weierstrass equations of C with respective discriminants Δ and Δ'. Then we have

$$v(\Delta') = v(\Delta) + 10v(\tau)\,, \tag{4.2}$$

see [Liu96, §2]. □

We will not need the normalized canonical local height in the remainder of this paper.

5. Stably minimal Weierstrass models

In this section, k continues to denote a non-archimedean local field with valuation ring $\mathcal{O}$ and residue field $\mathfrak{k}$. We build on results established by Liu [Liu96] in the more general context of hyperelliptic curves of arbitrary genus.

Recall that an equation of the form (3.1) defining a curve C over k of genus 2 is an *integral Weierstrass model* of C if the polynomials F and H have coefficients in $\mathcal{O}$.

(Note that this is slightly different from the notion of an 'integral equation' as defined in [Liu96, Définition 2], but the difference is irrelevant for our purposes, since any minimal Weierstrass model is actually given by an integral equation, see [Liu96, Remarque 4].) It is a *minimal Weierstrass model* of C if it is integral and the valuation of its discriminant is minimal among all integral Weierstrass models of C [Liu96, Définition 3]. We introduce the following variant of this notion.

Definition 5.1. An integral Weierstrass model of a smooth projective curve C over k of genus 2 is *stably minimal* if it is a minimal Weierstrass model for C over k' for every finite field extension k' of k.

Stably minimal Weierstrass models can be characterized in terms of the multiplicities of the points on the special fiber. The definition below is equivalent to [Liu96, Définition 9] when the curve is reduced, see [Liu96, Remarque 8].

Definition 5.2. Only for this definition let k be an arbitrary field, and let $C_{F,H}$ be a curve in $\mathbb{P}_k(1,3,1)$ given by an equation of the form (3.1) over k; we assume that $C_{F,H}$ is reduced. The *multiplicity* $m(P, C_{F,H})$ *of* $P \in C_{F,H}(\bar{k})$ is defined as follows:

- If P is a singular point of type A_n, then $m(P, C_{F,H}) = n + 1$.
- If P is fixed by the involution $\iota(X : Y : Z) = (X : -Y - H(X, Z) : Z)$ and is nonsingular, then $m(P, C_{F,H}) = 1$.
- Otherwise $m(P, C_{F,H}) = 0$.

An algorithm that computes the multiplicity was given by Liu [Liu96, §6.1]. Liu defines further multiplicities $\lambda_r(P)$ [Liu96, Définition 10] for points on the special fiber of an integral Weierstrass model (and $r \geq 1$) that allow to characterize when such a model is minimal. We note here that $\lambda_r(P)$ gives the value of $\lambda(P) = \lambda_1(P)$ after making a field extension of ramification index r. Also, Lemme 7(e) of [Liu96] states for r sufficiently large that $\lambda_r(P) = m(P)$ if the special fiber is reduced and implies that $\lambda_r(P) \geq r$ if the special fiber is non-reduced. In the reduced case, we also have $\lambda(P) \leq m(P)$.

Setting $\lambda = \lambda_1$, Corollaire 2 in [Liu96] states (for $g = 2$) that the model is minimal if and only if $\lambda(P) \leq 3$ and $\lambda'(P) \leq 4$ (and is the unique minimal Weierstrass model up to $\mathcal{O}$-isomorphism if and only if, in addition $\lambda'(P) \leq 3$) for all $\mathfrak{k}$-points P on the special fiber, where $\lambda'(P)$ is a number satisfying $\lambda'(P) \leq 2\lceil \lambda(P)/2 \rceil$, see [Liu96, Lemme 9(c)].

Lemma 5.3. *An integral Weierstrass model of a smooth projective curve C over k of genus 2 is stably minimal if and only if its special fiber is reduced and the multiplicity of every geometric point on the special fiber is at most* 3*.*

If the special fiber is reduced and all multiplicities are at most 2*, then the model is the unique minimal Weierstrass model of C over any finite extension k' of k, up to isomorphism over the valuation ring of k'.*

Proof. First note that the multiplicity of a point is a geometric property; it does not change when we replace k by a finite extension. If the special fiber of an integral Weierstrass model has the given properties, then it follows from Liu's results mentioned above that $\lambda(P) \leq m(P) \leq 3$ and therefore $\lambda'(P) \leq 4$ for all points P on the special fiber, even after replacing k by a finite extension. It follows that the model is stably minimal.

If $m(P) \leq 2$ for all P, then $\lambda(P) \leq 2$ and $\lambda'(P) \leq 2$, so by Liu's results, the model is the unique minimal Weierstrass model of C over k'.

Conversely, assume that the special fiber does not have the given properties. Then either the special fiber is non-reduced, or else there is a point P on the special fiber

of multiplicity $m(P) \geq 4$. If the special fiber is non-reduced, then after replacing k by a sufficiently ramified extension k', there is a point P on the special fiber such that $\lambda(P) > 3$ over k' (ramification index 4 is sufficient). If the special fiber is reduced and there is a (geometric) point P on the special fiber with $m(P) > 3$, then again after replacing k by a sufficiently large finite extension k' (such that P is defined over the residue field and the ramification index is at least $m(P)$), we have $\lambda(P) = m(P) > 3$ over k'. Liu's results then show that the model is not minimal over k'. □

Lemma 5.4. *If C is a smooth projective curve over k of genus 2, then there is a finite extension k' of k such that*

(i) *the minimal proper regular model of C over the valuation ring of k' has semistable reduction, and*
(ii) *each minimal Weierstrass model of C over k' is already stably minimal.*

Proof. That there is a finite extension with the first property is a special case of the semistable reduction theorem [DM69]. After a further unramified extension, we can assume that all components of the special fiber of the minimal proper regular model (which all have multiplicity 1) are defined over the residue field and that at least one component has a smooth point defined over the residue field. This implies by Hensel's Lemma that $C(k') \neq \emptyset$. It follows from [Liu96, Corollaire 5] that every minimal Weierstrass model of C over k' is dominated by the minimal proper regular model. Since the latter has reduced special fiber, the same is true for each minimal Weierstrass model.

Now assume that there exists a stably minimal Weierstrass model of C over k'. Then every minimal Weierstrass model of C over k' must already be stably minimal, since both models must have the same valuation of the discriminant, and the discriminant of the stably minimal model remains minimal over any finite field extension of k'. So it is enough to show that a stably minimal model exists.

We now consider the various possibilities for the special fiber of the minimal proper regular model. The possible configurations are shown in Figures 1, 2, 3 and 5 (on pages 30, 31, 32 and 37). If the reduction type is $[I_{m_1-m_2-m_3}]$ in the notation of [NU73], then the Weierstrass model whose special fiber contains the component(s) that are not (-2)-curves has the property that all points on the special fiber have multiplicity at most 2; this is then the unique minimal Weierstrass model, and it is stably minimal by Lemma 5.3. It remains to consider reduction type $[I_{m_1} - I_{m_2} - l]$. We see that the Weierstrass models that correspond to components in the chain linking the two polygons and also those coming from the component of one of the polygons that is connected to the chain satisfy the conditions of Lemma 5.3 and are thus stably minimal. On the other hand, Weierstrass models whose special fiber does not correspond to a component in the chain or to one of its neighbors have a point in the special fiber whose multiplicity is at least 4 and so cannot be stably minimal. □

6. Igusa invariants

In this section we describe how we can easily distinguish between different types of reduction using certain invariants of genus 2 curves introduced by Igusa in [Igu60]. The results of this section are essentially due to Liu [Liu93]; see also [Mes91].

Let k be an arbitrary field of characteristic not equal to 2 and consider the invariants $J_2, J_4, J_6, J_8, J_{10}$ defined in [Igu60], commonly called *Igusa invariants*. Then $J_{2i}(F)$ is an invariant of degree $2i$ of binary sextics, and if

$$F(X, Z) = f_0 Z^6 + f_1 X Z^5 + f_2 X^2 Z^4 + f_3 X^3 Z^3 + f_4 X^4 Z^2 + f_5 X^5 Z + f_6 X^6$$

is a binary sextic, then $J_{2i}(F) \in \mathbb{Z}[\frac{1}{2}, f_0, \dots, f_6]$. For example, $J_{10}(F) = 2^{-12}\operatorname{disc}(F)$. It is shown in [Igu60] that the invariants J_2, J_4, J_6, J_{10} generate the even degree part of the ring of invariants of binary sextics.

Now let F and H be the generic binary forms over $\mathbb{Z}$ of degrees 6 and 3, respectively, with coefficients $f_0, \dots, f_6$ and $h_0, \dots, h_3$ as before. It turns out that $J_{2i}\left(4F + H^2\right)$ is an element of $\mathbb{Z}[f_0, \dots, f_6, h_0, \dots, h_3]$.

Definition 6.1. Let k be an arbitrary field and let H, $F \in k[X, Z]$ be binary forms of respective degrees 3 and 6 over k. Let $C_{F,H}$ be the curve given by the equation $Y^2 + H(X, Z)Y = F(X, Z)$ in the weighted projective plane $\mathbb{P}_k(1, 3, 1)$. For $1 \le i \le 5$ we define the *Igusa invariant* $J_{2i}(C_{F,H})$ *of* $C_{F,H}$ as

$$J_{2i}(C_{F,H}) = J_{2i}\left(4F + H^2\right) .$$

Following Liu [Liu93], we also define two additional invariants, namely

$$I_4(C_{F,H}) = J_2(C_{F,H})^2 - 24J_4(C_{F,H})$$

and

$$\begin{aligned} I_{12}(C_{F,H}) = {} & -8J_4(C_{F,H})^3 + 9J_2(C_{F,H})J_4(C_{F,H})J_6(C_{F,H}) \\ & - 27J_6(C_{F,H})^2 - J_2(C_{F,H})^2 J_8(C_{F,H}) . \end{aligned}$$

The following is a consequence of [Liu93, Thm. 1].

Proposition 6.2. *Let k be a field and let $C_{F,H}/k$ be the curve given by the equation*

$$Y^2 + H(X, Z)Y = F(X, Z)$$

in $\mathbb{P}_k(1, 3, 1)$, where H, $F \in k[X, Z]$ are binary forms of degree 3 and 6, respectively. For $1 \le i \le 5$ and $j \in \{4, 12\}$ we set $J_{2i} = J_{2i}(C_{F,H})$ and $I_j = I_j(C_{F,H})$.

(i) *$C_{F,H}$ is smooth $\iff$ $J_{10} \ne 0$.*
(ii) *$C_{F,H}$ has a unique node and no point of higher multiplicity $\iff$ $J_{10} = 0$ and $I_{12} \ne 0$.*
(iii) *$C_{F,H}$ has exactly two nodes $\iff$ $J_{10} = I_{12} = 0$, $I_4 \ne 0$, and $J_4 \ne 0$ or $J_6 \ne 0$.*
(iv) *$C_{F,H}$ has three nodes $\iff$ $J_{10} = I_{12} = J_4 = J_6 = 0$ and $I_4 \ne 0$.*
(v) *$C_{F,H}$ has a cusp $\iff$ $J_{10} = I_{12} = I_4 = 0$ and $J_{2i} \ne 0$ for some $i \le 4$.*
(vi) *$C_{F,H}$ is non-reduced or has a point of multiplicity at least 4 $\iff$ $J_{2i} = 0$ for all i.*

When C is a curve of genus 2 over a non-archimedean local field, then Igusa invariants can also be used to obtain information on the reduction type of C, see [Liu93, Thm. 1, Prop. 2].

Proposition 6.3. *Let k be a non-archimedean local field with normalized additive valuation $v\colon k^\times \twoheadrightarrow \mathbb{Z}$ and valuation ring $\mathcal{O}$, and let C/k be a smooth projective genus 2 curve, given by a minimal Weierstrass model with reduced special fiber. Suppose that the minimal regular model $\mathcal{C}^{\min}$ of C over $\operatorname{Spec}\mathcal{O}$ is semistable and has reduction type $\mathcal{K}$ in the notation of [NU73]. We set $J_{2i} = J_{2i}(C)$ for $i \in \{1, \dots, 5\}$ and $I_4 = I_4(C)$, $I_{12} = I_{12}(C)$.*

(i) *If $\mathcal{K} = [I_{m-0-0}]$, where $m > 0$, then $m = v(J_{10})$.*
(ii) *If $\mathcal{K} = [I_{m_1-m_2-0}]$, where $0 < m_1 \le m_2$, then*

$$m_1 = \min\left\{v(I_{12}), \tfrac{1}{2}v(J_{10})\right\} \quad \text{and} \quad m_2 = v(J_{10}) - m_1 .$$

(iii) *If* $\mathcal{K} = [I_{m_1-m_2-m_3}]$, *where* $0 < m_1 \le m_2 \le m_3$, *then*

$$\begin{aligned} m_1 &= \min\left\{v(J_4), \tfrac{1}{3}v(J_{10}), \tfrac{1}{2}v(I_{12})\right\}, \\ m_2 &= \min\left\{v(I_{12}) - m_1, \tfrac{1}{2}(v(J_{10}) - m_1)\right\} \quad \textit{and} \\ m_3 &= v(J_{10}) - m_1 - m_2\,. \end{aligned}$$

(iv) *If* $\mathcal{K} = [I_0 - I_0 - l]$, *then* $l = \frac{1}{12}v(J_{10})$.

(v) *If* $\mathcal{K} = [I_{m_1} - I_0 - l]$, *where* $m_1 > 0$, *then*

$$l = \tfrac{1}{12}v(I_{12}) \quad \textit{and} \quad m_1 = v(J_{10}) - v(I_{12})\,.$$

(vi) *If* $\mathcal{K} = [I_{m_1} - I_{m_2} - l]$, *where* $m_2 \ge m_1 > 0$ *and* $l > 0$, *then*

$$\begin{aligned} l &= \tfrac{1}{4}v(I_4)\,, \\ m_1 &= \min\left\{v(I_{12}) - 3v(I_4)\,, \tfrac{1}{2}(v(J_{10}) - 3v(I_4))\right\} \quad \textit{and} \\ m_2 &= v(J_{10}) - 3v(I_4) - m_1\,. \end{aligned}$$

Part II: Study of Local Height Correction Functions

In Part II of the paper, k will always denote a non-archimedean local field with residue field $\mathfrak{k}$, valuation ring $\mathcal{O}$ and normalized additive valuation $v\colon k^\times \twoheadrightarrow \mathbb{Z}$. We let C be a curve of genus 2 over k, given by an integral Weierstrass model $\mathcal{C}$, which we consider as a subscheme of the weighted projective plane $\mathbb{P}_S(1,3,1)$, where $S = \operatorname{Spec}(\mathcal{O})$. In the following five sections we find explicit formulas and bounds for the local height correction function μ for the most frequent cases of bad reduction and use these to deduce a general bound on μ. We denote the minimal regular model of C over S by $\mathcal{C}^{\min}$. Let J be the Jacobian of C; we denote its Néron model over S by $\mathcal{J}$. We write $\mathcal{C}_v$, $\mathcal{C}_v^{\min}$ and $\mathcal{J}_v$ for the respective special fibers of $\mathcal{C}$, $\mathcal{C}^{\min}$ and $\mathcal{J}$.

7. The 'kernel' of μ

By Theorem 3.10, the set

$$U = \{P \in J(k) : \varepsilon(P) = 0\}$$

is a group and the local height correction function μ factors through the quotient $J(k)/U$. In this section we relate U to the Néron model of J when $\mathcal{C}$ has rational singularities. See [Art86] for a brief account of the theory of rational singularities on arithmetic surfaces.

For the remainder of this section we assume that $\mathcal{C}/S$ is normal and reduced. We let $\mathcal{J}^0$ denote the identity component (the scheme with generic fiber $\mathcal{J}_k \cong J$ and special fiber the connected component of the identity $\mathcal{J}_v^0$ of the special fiber $\mathcal{J}_v$) of the Néron model $\mathcal{J}$. If $\mathcal{C}' \to \mathcal{C}$ is a desingularization of $\mathcal{C}$, then the identity components $\operatorname{Pic}^0_{\mathcal{C}'/S}$ and $\operatorname{Pic}^0_{\mathcal{C}/S}$ of the respective relative Picard functors of $\mathcal{C}'$ and $\mathcal{C}$ can both be represented by separated schemes, see [BLR90, Thm. 9.7.1]. There are canonical S-group scheme morphisms

$$\operatorname{Pic}^0_{\mathcal{C}/S} \longrightarrow \operatorname{Pic}^0_{\mathcal{C}'/S} \xrightarrow{\sim} \mathcal{J}^0\,; \tag{7.1}$$

the latter map is an isomorphism by [BLR90, Thm. 9.4.2]. Let $\alpha\colon \operatorname{Pic}^0_{\mathcal{C}/S} \to \mathcal{J}^0$ denote the composition of the morphisms from (7.1); note that α does not depend on the choice of the desingularization $\mathcal{C}'$.

Let $\mathcal{C}_{\mathrm{sm}}$ be the smooth locus of $\mathcal{C}$. Essentially following [BLR90, §9.3], we define W as the subscheme of the smooth S-scheme $\mathcal{C}_{\mathrm{sm}}^{(2)}$ containing the points $w = \{P_1, P_2\} \in W$ such that the hyperelliptic involution ι maps the component containing P_1 to the component containing P_2 and such that $H^1(\mathcal{C}_w, \mathcal{O}_{\mathcal{C}_w}(D_w)) = 0$, where D is the universal Cartier divisor $D \subset \mathcal{C} \times_S \mathcal{C}_{\mathrm{sm}}^{(2)}$ induced by the canonical map $\mathcal{C}_{\mathrm{sm}}^{(2)} \to \operatorname{Div}^2_{\mathcal{C}/S}$. The latter condition implies that if $w = \{P_1, P_2\}$ lies on the generic fiber of W, then the divisor $(P_1) + (P_2)$ is non-special. Then W has the following properties:

(i) W is an open subscheme of $\mathcal{C}_{\mathrm{sm}}^{(2)}$.
(ii) There is a strict S-birational group law on W, induced by the group law on $\operatorname{Pic}_{\mathcal{C}/S}$.
(iii) $\operatorname{Pic}^0_{\mathcal{C}/S}$ is the S-group scheme associated with this strict S-birational group law.

For (ii) and (iii) see the discussion preceding [BLR90, Thm. 9.3.7].

Let $\operatorname{Pic}^{(2)}_{\mathcal{C}/S}$ be the open subfunctor of $\operatorname{Pic}_{\mathcal{C}/S}$ whose elements have total degree 2. Let $\rho\colon W \to \operatorname{Pic}^{(2)}_{\mathcal{C}/S}$ be the canonical map induced by D; by [BLR90, Lemma 9.3.5] it is an open immersion. We can find a section $x_0 \in \mathbb{P}^1_S(S)$ such that its pullback D_0 under the

covering map $\mathcal{C} \to \mathbb{P}^1_S$ is horizontal and does not intersect the singular locus of $\mathcal{C}$. We denote by c_0 the class of D_0 in $\mathrm{Pic}^{(2)}_{\mathcal{C}/S}$. Let $w = \{P_1, P_2\} \in W$; using the condition on the action of ι on the components P_1 and P_2 lie on, we find that

$$\rho_0(w) := \rho(w) - c_0 \in \mathrm{Pic}^0_{\mathcal{C}/S} \, .$$

In fact ρ_0 defines an open immersion $\rho_0 \colon W \to \mathrm{Pic}^0_{\mathcal{C}/S}$, see [BLR90, Lemma 9.3.6].

Lemma 7.1. *If $P \in J(k)$ is in the image of α and the residue characteristic of k is not 2, then $\varepsilon(P) = \mu(P) = 0$.*

Proof. We may assume that $\mathcal{C} : Y^2 = F(X, Z)$. Let J_F denote the model of J in $\mathbb{P}^{15}$ constructed in [CF96, Chapter 2] and let $\mathcal{J}_F/S$ denote the model it defines over S. Furthermore, we denote by $\mathcal{J}^0_F$ the scheme whose generic fiber is $\mathcal{J}_F$ and whose special fiber $\mathcal{J}^0_{F,v}$ is the connected component of the identity of the smooth locus of the special fiber $\mathcal{J}_{F,v}$. We have a morphism $\psi \colon \mathcal{C}^{(2)}_{\mathrm{sm}} \to \mathcal{J}^0_F$, defined using the expressions for the coordinates on J_F in [CF96, Chapter 2], see the proof of [BS10, Lemma 5.7]. We also denote the restriction of this morphism to W by ψ.

By the Néron mapping property, there is a morphism $\varphi \colon \mathcal{J}^0_F \to \mathcal{J}$. In general, its image can be a proper subset of $\mathcal{J}^0$. Nevertheless, the following diagram of S-scheme morphisms is commutative by [Liu02, Prop. 3.3.11], since W is reduced, $\mathcal{J}^0$ is separated and the diagram is commutative when restricted to generic fibers:

$$\begin{array}{ccc} W & \xrightarrow{\psi} & \mathcal{J}^0_F \\ {\scriptstyle \rho_0}\downarrow & & \downarrow{\scriptstyle \varphi} \\ \mathrm{Pic}^0_{\mathcal{C}/S} & \xrightarrow{\alpha} & \mathcal{J}^0 \end{array} \tag{7.2}$$

We claim that $\mathrm{im}(\alpha) \subset \mathrm{im}(\varphi)$. To prove this, we will show that φ is surjective onto the image of α as a morphism of sheaves on the big étale site over S. Let T be an S-scheme and let $x \in \mathrm{Pic}^0_{\mathcal{C}/S}(T)$. By property (ii) and (iii) of W, there is an étale cover T'/T and $w_1, \ldots, w_n \in W(T')$ such that

$$x = \rho_0(w_1) + \ldots + \rho_0(w_n) \, ,$$

where the sum is taken with respect to the group law on $\mathrm{Pic}^0_{\mathcal{C}/S}$. In fact we can take $n = 2$; this follows from [BLR90, Lemma 5.1.4] and the discussion following [BLR90, Lemma 5.2.4]. Setting

$$y = \psi(w_1) + \psi(w_2) \, ,$$

where now the sum is taken with respect to the group law on $\mathcal{J}^0_F$, we immediately deduce $\varphi(y) = \alpha(x)$ from the commutativity of the diagram (7.2).

Let $P \in J(k)$. It follows from [BS10, Prop. 5.10] that $\varepsilon(P) = 0$ if and only if P maps to $\mathcal{J}^0_{F,v}(\mathfrak{k})$. Since $\varepsilon(P) = 0$ implies $\mu(P) = 0$, the result follows. □

Now suppose that the residue characteristic is 2. In this situation no explicit analogue of the group scheme $\mathcal{J}_F$ has been constructed yet. Instead, we will show directly that there is an open subscheme W' of W such that a point $P \in J(k)$ in the image of $\alpha \circ \rho_0|_{W'} \colon W' \to \mathcal{J}^0$ must satisfy $\mu(P) = 0$. This will suffice for our intended application.

In analogy with [BS10, Definition 5.1], we define a subscheme $\tilde{D}_{\tilde{F},\tilde{H}}$ of $\mathbb{A}^3_{\mathfrak{k}} \times \mathbb{A}^4_{\mathfrak{k}} \times \mathbb{A}^5_{\mathfrak{k}}$ consisting of all triples

$$(A, B, C) = \big((a_0, a_1, a_2),\ (b_0, b_1, b_2, b_3),\ (c_0, c_1, c_2, c_3, c_4)\big) \in \mathbb{A}^3_{\mathfrak{k}} \times \mathbb{A}^4_{\mathfrak{k}} \times \mathbb{A}^5_{\mathfrak{k}}$$

such that

$$AC = F - B^2 - BH\,,$$

where

$$\begin{aligned} A &= a_0Z^2 + a_1XZ + a_2X^2, \\ B &= b_0Z^3 + b_1XZ^2 + b_2X^2Z + b_3X^3, \\ C &= c_0Z^4 + c_1XZ^2 + c_2X^2Z^2 + c_3X^3Z + c_4X^4. \end{aligned}$$

Moreover, we denote by $D_{\tilde{F},\tilde{H}}$ the subscheme of $\mathbb{P}^2_{\mathfrak{k}} \times \mathbb{A}^4_{\mathfrak{k}}$ defined as $(\pi_2 \times \mathrm{id})\big(\mathrm{pr}_{12}(\tilde{D}_{\tilde{F},\tilde{H}})\big)$, where pr_{12} is the projection onto the first two factors and π_2 is the canonical map $\mathbb{A}^3_{\mathfrak{k}} \setminus \{(0,0,0)\} \to \mathbb{P}^2_{\mathfrak{k}}$.

Note that if the curve $C_{\tilde{F},\tilde{H}}$ defined by $Y^2 + \tilde{H}(X,Z)Y = \tilde{F}(X,Z)$ in $\mathbb{P}_{\mathfrak{k}}(1,3,1)$ is nonsingular, then $D_{\tilde{F},\tilde{H}}$ is in bijective correspondence with the possible Mumford representations of effective divisors of degree 2 on $C_{\tilde{F},\tilde{H}}$.

Let $D'_{\tilde{F},\tilde{H}}$ be the subset of all $(A,B) \in D_{\tilde{F},\tilde{H}}$ such that A does not vanish at the image in $\mathbb{P}^1$ of a singular point of $\mathcal{C}$. We define a map $\zeta\colon D'_{\tilde{F},\tilde{H}} \to \mathcal{C}^{(2)}$ such that if $\zeta((A,B)) = \{\tilde{P}_1, \tilde{P}_2\}$, then there are representatives (X_i, Y_i, Z_i) of $\tilde{P}_i$ $(i = 1, 2)$ satisfying

(i) $A(X,Z) = (Z_1X - X_1Z)(Z_2X - X_2Z)$;
(ii) $Y_i = B(X_i, Z_i)$ for $i = 1, 2$.

Consider the subscheme $\tilde{W}$ of $D_{\tilde{F},\tilde{H}}$ given by those points (A,B) which satisfy one of the following conditions:

(1) $(A,B) \in D'_{\tilde{F},\tilde{H}}$, and if $\zeta((A,B)) = \{\tilde{P}_1, \tilde{P}_2\}$, then ι maps the component containing $\tilde{P}_1$ to the component containing $\tilde{P}_2$.
(2) $A = c\tilde{L}(X,Z)^2$, where $\tilde{L}$ vanishes at the projection of a singular point of $\mathcal{C}$ to $\mathbb{P}^1$ and $\tilde{L}^3$ does not divide $\tilde{F} - B^2 - B\tilde{H}$.

If $C_{\tilde{F},\tilde{H}}$ is nonsingular, and $(A,B) \in D_{\tilde{F},\tilde{H}}$, then we can compose the natural surjection $D_{\tilde{F},\tilde{H}} \to \mathrm{Jac}(C_{\tilde{F},\tilde{H}}) \setminus \{O\}$ with the quotient map $\mathrm{Jac}(C_{\tilde{F},\tilde{H}}) \to \mathrm{KS}_{\tilde{F},\tilde{H}}$. In the general case one can also define a surjection $\omega\colon D_{\tilde{F},\tilde{H}} \to \mathrm{KS}_{\tilde{F},\tilde{H}} \setminus \{(0:0:0:1)\}$, where the image of a pair $(A,B) \in D_{\tilde{F},\tilde{H}}$ under ω is of the form $(a_0 : -a_1 : a_2 : x_4)$. We only give the form of x_4 in two special cases: If $b_3 = b_2 = 0$, then we have

$$x_4 = b_1^2 + b_1S_1(a_0,a_1,a_2) + S_0(a_0,a_1,a_2). \tag{7.3}$$

and if $b_0 = b_1 = 0$, then

$$x_4 = b_2^2 + b_2T_1(a_0,a_1,a_2) + T_0(a_0,a_1,a_2). \tag{7.4}$$

Here S_0, S_1, T_0, T_1 are homogeneous forms in a_0, a_1, a_2; T_0 and S_0 have coefficients in $\mathbb{F}_2[f_0, \ldots, f_6]$, whereas S_1 and T_1 have coefficients in $\mathbb{F}_2[h_0, \ldots, h_3]$.

Lemma 7.2. *If $x \in \omega(\tilde{W}) \subset \mathrm{KS}_{\tilde{F},\tilde{H}}$, then we have $\delta(x) \neq 0$.*

Proof. Changing the given model if necessary, we can assume that $\tilde{H}$ and $\tilde{F}$ are as in the list of representatives 1–13 in Table 1. Table 2 contains conditions on x which are equivalent to the vanishing of $\delta(x)$ for each representative and additional conditions which a point $x = (x_1 : x_2 : x_3 : x_4) \in \mathbb{P}^3$ satisfying $\delta(x) = 0$ must satisfy in order to lie on $\mathrm{KS}_{\tilde{F},\tilde{H}}$. Finally, we have listed the multiplicities that $C_{\tilde{F},\tilde{H}}$ has at the points

type	condition	additional	multiplicities
1	$x_4 = 0$		
2	$x_4 = 0$		(6)
3	$x_4 = 0$	$x_1 = 0$	(5)
4	$x_4 = 0$	$x_1 = 0$	(4)
5	$x_1x_3 = x_4 = 0$		(3, 2)
6	$x_1 = x_4 = 0$		(3)
7	$x_4 = 0$		(4, 2)
8	$x_4 = 0$		(2, 2, 2)
9	$x_1x_3 = x_4 = 0$		(2, 2)
10	$x_1 = x_4 = 0$		(2)
11	$x_1 = x_4 = 0$		(3)
12	$x_4 = 0$	$x_1 = 0$	(5)
13	$x_4 = 0$	$x_1x_3 = 0$	(3, 3)

TABLE 2. Conditions for the vanishing of $\delta(x)$

$(X : Z) = (1 : 0)$, $(X : Z) = (0 : 1)$ and $(X : Z) = (1 : 1)$, respectively, in case the multiplicities there are greater than 1.

Let $(A, B) \in \tilde{W}$ be such that $\omega((A, B)) = x$. If $A(X, Z)$ does not vanish at the image in $\mathbb{P}^1$ of a singular point, then in particular $x_1 \neq 0$ and, if $(0, 0)$ is a singular point, also $x_3 \neq 0$. Using Table 2, this already implies that $\delta(x) \neq 0$ whenever $\mathcal{C}$ is irreducible. In the reducible cases 2, 7 and 8, $\mathcal{C}$ has two irreducible components and one checks easily that x_4 does not vanish if $\zeta((A, B)) = \{\tilde{P}_1, \tilde{P}_2\}$ such that ι maps the component containing $\tilde{P}_1$ to the component containing $\tilde{P}_2$. Hence $\delta(x) \neq 0$ by Table 2.

We now assume that $A(X, Z)$ vanishes at the image of a singular point. It suffices to consider the cases $\tilde{L} = Z$ and $\tilde{L} = X$. We will show that $x_4 \neq 0$, which is enough to prove $\delta(x) \neq 0$ using Table 2.

If $\tilde{L} = Z$, then we may assume $B = b_2X^2Z + b_3X^3$ and we have $x = (0 : 0 : 1 : x_4)$. Because A divides $\tilde{F} + B^2 + B\tilde{H}$, we get $b_3 = 0$. A simple computation using (7.4) shows that in all cases 1–13 we have $x_4 = b_2(b_2 + h_2)$, but $(A, B) \in \tilde{W}$ implies $b_2(b_2 + h_2) \neq 0$.

If $\tilde{L} = X$, then we may assume $B = b_1XZ^2 + b_0Z^3$ and find $x = (1 : 0 : 0 : x_4)$. Using an argument analogous to the one employed for $\tilde{L} = Z$ we get $b_0 = 0$ and $x_4 = b_1(b_1 + h_1) \neq 0$ using (7.3). □

Define the subscheme W' of W as the intersection $W \cap \zeta(\tilde{W}) \subset \mathcal{C}^{(2)}$.

Lemma 7.3. *If $P \in J(k)$ is in the image of α and the residue characteristic of k is 2, then $\varepsilon(P) = \mu(P) = 0$.*

Proof. First suppose that $P \in J(k)$ is a point such that $(\tilde{A}, \tilde{B}) \in \tilde{W}$, where (A, B) is a Mumford representative of (the effective degree 2 divisor corresponding to) P. If $x \in \mathrm{KS}_{\mathbb{A}}$ is a set of normalized Kummer coordinates for P, then the reduction $\tilde{x} \in \mathrm{KS}_{\tilde{F},\tilde{H}}$ of x satisfies $\tilde{x} \in \omega(\tilde{W})$. Lemma 7.2 implies that $\delta(\tilde{x}) \neq 0$ and hence $\varepsilon(x) = 0$.

To finish the proof, we argue in a similar way as in the proof of Lemma 7.1. Note that $\tilde{W}$ is defined in such a way that we have $W' = W$ if $\mathcal{C}$ is irreducible. If $\mathcal{C}$ is reducible, then a point $\{P_1, P_2\}$ in the special fiber of W lies in the special fiber of W' if and only if ι maps the component containing P_1 to the component containing P_2. Hence W' is an open subscheme of W (recall that $\mathcal{C}$ is assumed to be reduced).

This implies that for an S-scheme T and $x \in \mathrm{Pic}^0_{\mathcal{C}/S}(T)$ we can find an étale cover T'/T and finitely many $w_1, \dots, w_n \in W'(T')$ such that $x = \rho_0(w_1) + \dots + \rho_0(w_n)$. We already showed that $\mu(\alpha(\rho_0(w_i))) = 0$ for all $i = 1, \dots, n$; by Lemma 3.7, this suffices to prove the lemma. □

Taking Lemmas 7.1 and 7.3 together, we have now proved the following.

Proposition 7.4. *Let $\alpha\colon \mathrm{Pic}^0_{\mathcal{C}/S} \to \mathcal{J}^0$ be the canonical homomorphism. If $P \in J(k)$ is in the image of α, then $\varepsilon(P) = \mu(P) = 0$.*

Let $J_0(k)$ denote the subgroup of $J(k)$ consisting of points whose image on the special fiber of $\mathcal{J}$ is in $\mathcal{J}^0(\mathfrak{k})$. Then the group $\Phi(\mathfrak{k})$ of $\mathfrak{k}$-rational points in the component group Φ of J satisfies

$$\Phi(\mathfrak{k}) \cong J(k)/J_0(k)\,.$$

We can now give a criterion for when ε and μ factor through $\Phi(\mathfrak{k})$.

Theorem 7.5. *Let C be a smooth projective curve of genus 2 defined over a non-archimedean local field k, given by an integral Weierstrass model $\mathcal{C}$ with rational singularities. Then ε and μ factor through $\Phi(\mathfrak{k})$.*

Proof. First note that if $\mathcal{C}$ has rational singularities, then $\mathcal{C}$ is normal and reduced. Moreover, according to [BLR90, Thm. 9.7.1(b)], the homomorphism α is an isomorphism if and only if $\mathcal{C}$ has rational singularities. This implies that the image of α, restricted to the generic fiber, is $J_0(k)$. By Proposition 7.4, we have $\varepsilon(P) = \mu(P) = 0$ for P in the image of α. Theorem 3.10 implies that μ and ε factor through $\Phi(\mathfrak{k})$. □

Remark 7.6. A non-minimal Weierstrass model cannot have rational singularities. However, there are minimal (even stably minimal) Weierstrass models of curves of genus 2 that have non-rational singularities. See Example 10.4 for a stably minimal Weierstrass model having $\mu(P) \neq 0$ for some points $P \in J_0(k)$.

This behavior cannot occur for elliptic curves; here μ always factors through $\Phi(\mathfrak{k})$, provided the given Weierstrass model is minimal, see [Sil88]. This is crucial for the usual algorithms to compute canonical heights on elliptic curves. Note that a Weierstrass model of an elliptic curve is minimal if and only if it has rational singularities by [Con05, Corollary 8.4].

8. Néron functions and reduction graphs

Our next goal is to derive a formula for $\mu(P)$ in the case when the minimal proper regular model of C is semistable and μ factors through $\Phi(\mathfrak{k})$. To this end, we need the notion of Néron functions. The following result is due to Néron; see [Lan83, §11.1].

Proposition 8.1. *Let A be an abelian variety defined over a local field k. Then we can associate to any divisor $D \in \mathrm{Div}(A)$ a function $\lambda_D\colon A \setminus \mathrm{supp}(D) \to \mathbb{R}$ such that the following conditions are satisfied, where we write $\lambda \equiv \lambda'$ mod const. to indicate that the functions λ and λ' differ by a constant.*

(1) *If $D, E \in \mathrm{Div}(A)$, then $\lambda_{D+E} \equiv \lambda_D + \lambda_E$ mod const.*
(2) *If $D = \mathrm{div}(f) \in \mathrm{Div}(A)$ is principal, then $\lambda_D \equiv v \circ f$ mod const.*
(3) *If $D \in \mathrm{Div}(A)$ and $T_P\colon A \to A$ is the translation map by a point $P \in A(k)$, then we have $\lambda_{T_P^* D} \equiv \lambda_D \circ T_P$ mod const.*

Also, λ_D is uniquely determined up to adding a constant.

We call a function λ_D as in Proposition 8.1 a *Néron function associated with* D.

We can use local heights on Kummer coordinates to construct Néron functions on the Jacobian J of our genus 2 curve C. If $P_0 \in C(\bar{k})$, then we have an embedding $C \to J$, defined over $\bar{k}$, that maps a point $P \in C$ to the divisor class $[(P) - (P_0)] \in \mathrm{Pic}^0_C = J$. Its image is the theta divisor Θ_{P_0}. We set $\Theta^{\pm}_{P_0} = \Theta_{P_0} + \Theta_{\iota(P_0)}$; then $\Theta^{\pm}_{P_0}$ is symmetric and in the linear equivalence class of 2Θ (where Θ is a theta divisor coming from taking a Weierstrass point as base-point). For the following, fix a point $\infty \in C(\bar{k})$ at infinity. For $i \in \{1, \dots, 4\}$, we set

$$D_i = \Theta^{\pm}_{\infty} + \mathrm{div}\left(\frac{\kappa_i}{\kappa_1}\right)$$

and we define a function $\hat{\lambda}_i \colon J(k) \setminus \mathrm{supp}(D_i) \to \mathbb{R}$ by

$$\hat{\lambda}_i(P) = \hat{\lambda}\left(\frac{\kappa(P)}{\kappa_i(P)}\right).$$

Lemma 8.2. *Let $\infty \in C(\bar{k})$ be a point at infinity as above and let $i \in \{1, \dots, 4\}$. Then D_i is defined over k and the function $\hat{\lambda}_i$ is a Néron function associated with D_i.*

Proof. If $\infty \notin C(k)$, then we have $\infty \in C(k')$ for some quadratic extension k' of k and the nontrivial element of the Galois group $\mathrm{Gal}(k'/k)$ maps ∞ to $\iota(\infty)$, proving the first assertion. For a proof of the second assertion, see [Uch11, Thm. 5.3]. □

Definition 8.3. Assume that C has semistable reduction over k. The *reduction graph* $R(C)$ of C is a graph with vertex set the set of irreducible components of the special fiber of $\mathcal{C}^{\min}$; two vertices Γ_1 and Γ_2 are connected by n edges, where n is the number of intersection points of Γ_1 and Γ_2 if $\Gamma_1 \neq \Gamma_2$, and n is the number of nodes of Γ_1 if $\Gamma_1 = \Gamma_2$.

We consider $R(C)$ as a metric graph by giving each edge length 1. For two vertices Γ_1 and Γ_2, we define $r(\Gamma_1, \Gamma_2)$ as the resistance between the vertices, when $R(C)$ is considered as an electric network with unit resistance along every edge.

Remark 8.4. We can compute $r(\Gamma_1, \Gamma_2)$ as follows. Order the vertices of $R(C)$ in some way and let M be the intersection matrix with respect to this ordering. Since all components of the special fiber have multiplicity one, the kernel of M is spanned by the 'all-ones' vector and the image of M consists of the vectors whose entries sum to zero. Let v be the vector with entries zero except that the entry corresponding to Γ_1 is 1 and the entry corresponding to Γ_2 is -1. Then there is a vector g with rational entries such that $Mg = v$, and

$$r(\Gamma_1, \Gamma_2) = -g \cdot v$$

is, up to sign, the standard inner product of the two vectors. (Note that g is not unique, but adding a vector in the kernel of M to it will not change the result.) See for instance [Cin11, Lemma 6.1].

Note that the linear map given by M on the space of functions on the vertices can be interpreted as the discrete Laplace operator on the graph $R(C)$. It is then easy to see that g, viewed as a function on the vertices, is piecewise linear along sequences of edges not containing Γ_1, Γ_2 or a vertex of degree at least 3. This makes it quite easy to find g and to compute $r(\Gamma_1, \Gamma_2)$.

The reduction graph is unchanged when we replace k by an unramified extension. If we base-change to a ramified extension k' of k with ramification index e, then the new reduction graph is obtained by subdividing the edges of $R(C)$ into e new edges. We can

give these new edges length $1/e$; then the underlying metric space remains the same. In particular, $r(\Gamma_1,\Gamma_2)$ does not depend on k'. This allows us to replace k by a finite extension if necessary. The scaling of the length corresponds to extending the valuation $v\colon k^\times \twoheadrightarrow \mathbb{Z}$ to $\bar{k}^\times \to \mathbb{Q}$ instead of considering the normalized valuation on k'. All notions defined in terms of the valuation (for example, intersection numbers) are then scaled accordingly.

Proposition 8.5. *We assume that $\mathcal{C}^{\min}$ is semistable. Let $P = [(P_1) - (P_2)] \in J(k)$, with $P_1, P_2 \in C(k)$ mapping to components Γ_1 and Γ_2, respectively, of the special fiber of $\mathcal{C}^{\min}$. We make the following further assumptions.*

(i) *If $Q_1, Q_2 \in C(k)$ map to Γ_1 and Γ_2, respectively, then $\mu(P) = \mu([(Q_1) - (Q_2)])$.*
(ii) *There is a constant $\mu_1 \in \mathbb{Q}$ such that $\mu([(Q_1) - (Q_1')]) = \mu_1$ for all $Q_1, Q_1' \in C(k)$ mapping to Γ_1 such that the images of Q_1 and Q_1' on the special fiber of $\mathcal{C}^{\min}$ are distinct.*
(iii) *There is a constant $\mu_2 \in \mathbb{Q}$ such that $\mu([(Q_2) - (Q_2')]) = \mu_2$ for all $Q_2, Q_2' \in C(k)$ mapping to Γ_2 such that the images of Q_2 and Q_2' on the special fiber of $\mathcal{C}^{\min}$ are distinct.*

Then we have

$$\mu(P) = r(\Gamma_1,\Gamma_2) + \frac{\mu_1+\mu_2}{2}\,.$$

Proof. By the discussion preceding the statement of the theorem, we can assume that k is sufficiently large for $C(k)$ to contain all points we might be interested in.

Let $P_0 \in C(k)$. The embedding with respect to P_0 is obtained from the 'difference map' $\psi\colon C \times C \to J$ that sends a pair of points (P_1, P_2) to $[(P_1) - (P_2)]$ by specializing the second argument to P_0. One easily checks that

$$\psi^*\Theta_{P_0} = \Delta_C + (\{\iota(P_0)\} \times C) + (C \times \{P_0\})\,,$$

where Δ_C denotes the diagonal and ι is the hyperelliptic involution on C. We then have

$$\psi^*\Theta^{\pm}_{P_0} = 2\Delta_C + \mathrm{pr}_1^*\, D_0 + \mathrm{pr}_2^*\, D_0\,,$$

where $D_0 = (P_0) + (\iota(P_0))$. By the results in [Hei04] this implies that, taking λ_0 to be a Néron function associated to $\Theta^{\pm}_{P_0}$,

$$\lambda_0\big([(P_1) - (P_2)]\big) = 2\langle P_1, P_2\rangle + \langle P_1 + P_2, P_0 + \iota(P_0)\rangle + c$$

for all points $P_1, P_2 \in C(k^{\mathrm{nr}})$ with $P_1 \neq P_2$ and $\{P_1, P_2\} \cap \{P_0, \iota(P_0)\} = \emptyset$, where $\langle\cdot,\cdot\rangle$ is the pairing in [Hei04, Thm. 4.4] and $c \in \mathbb{R}$ is a constant.

If $\mathcal{C}^{\min}$ has semistable reduction, then, by [Hei04, Remark 4.6], the pairing $\langle\cdot,\cdot\rangle$ coincides with Zhang's admissible pairing $(\cdot,\cdot)_a$ defined in [Zha93] in terms of harmonic analysis on the reduction graph $R(C)$. In these terms, we have for $Q, Q' \in C(k^{\mathrm{nr}})$:

$$\langle Q, Q'\rangle = (Q, Q')_a = i(\overline{Q}, \overline{Q}') + g_\nu(\Gamma, \Gamma')\,,$$

where $i(\overline{Q}, \overline{Q}')$ is the intersection multiplicity of the sections $\overline{Q}, \overline{Q}' \in \mathcal{C}^{\min}(\mathcal{O}^{\mathrm{nr}})$ induced by Q and Q', respectively, and $g_\nu(\Gamma,\Gamma')$ is the Green's function associated to a certain measure ν on $R(C)$, with Γ and Γ' being the respective components of the special fiber of $\mathcal{C}^{\min}$ that Q and Q' reduce to. See [Zha93, §4]. We extend g_ν to a bilinear map on the free abelian group generated by the vertices of $R(C)$.

Lemma 8.2 gives, for $P_0 = \infty$ and $P = [(P_1)-(P_2)]$ with normalized Kummer coordinates $x(P) = (x_1(P), \dots, x_4(P))$,

$$\begin{aligned}\mu(P) &= v(x_1(P)) - \hat{\lambda}_1(P) \\ &= v(x_1(P)) - 2i(\overline{Q}_1, \overline{Q}_2) - i(\overline{Q}_1 + \overline{Q}_2, \overline{P}_0 + \overline{\iota(P_0)}) \\ &\qquad - 2g_\nu(\Gamma_1, \Gamma_2) - g_\nu(\Gamma_1 + \Gamma_2, \Gamma_0 + \Gamma_0') - c\,,\end{aligned}$$

where Γ_1 and Γ_2 are the respective components that P_1 and P_2 reduce to, and Γ_0 and Γ_0' are the respective components that P_0 and $\iota(P_0)$ reduce to. We assume for a moment that the images of P_1 and P_2 on the special fiber of the original model $\mathcal{C}$ are distinct from the images of the points at infinity. By assumption (i), $\mu(P)$ is unchanged when we replace the points P_1 and P_2 by other points still mapping to Γ_1 and Γ_2, respectively. We can therefore assume that the images of P_1 and P_2 on the special fiber of $\mathcal{C}^{\min}$ are distinct from each other and also from the images of P_0 and $\iota(P_0)$. This implies that $v(x_1(P)) = 0$. Then the intersection numbers in the formula above are zero. We can choose further points Q_1 and Q_2 that also reduce to Γ_1 and Γ_2 with reductions on the special fiber of $\mathcal{C}$ distinct from those of P_0 and $\iota(P_0)$ and such that P_1, P_2, Q_1 and Q_2 all reduce to distinct points on the special fiber of $\mathcal{C}^{\min}$. Using assumptions (ii) and (iii), we obtain the following relations.

$$\begin{aligned}-\tfrac{1}{2}\mu_1 = -\tfrac{1}{2}\mu\big([(P_1) - (Q_1)]\big) &= g_\nu(\Gamma_1, \Gamma_1) + g_\nu(\Gamma_1, \Gamma_0 + \Gamma_0') + \tfrac{1}{2}c \\ \mu(P) = \mu\big([(P_1) - (P_2)]\big) &= -2g_\nu(\Gamma_1, \Gamma_2) - g_\nu(\Gamma_1 + \Gamma_2, \Gamma_0 + \Gamma_0') - c \\ -\tfrac{1}{2}\mu_2 = -\tfrac{1}{2}\mu\big([(P_2) - (Q_2)]\big) &= g_\nu(\Gamma_2, \Gamma_2) + g_\nu(\Gamma_2, \Gamma_0 + \Gamma_0') + \tfrac{1}{2}c\end{aligned}$$

Adding them together gives

$$\mu(P) - \tfrac{1}{2}(\mu_1 + \mu_2) = g_\nu(\Gamma_1 - \Gamma_2, \Gamma_1 - \Gamma_2) = r(\Gamma_1, \Gamma_2)\,,$$

as desired. See [Zha93, §3] for the last equality.

If our assumption that the images of P_1 and P_2 on the special fiber of the original model $\mathcal{C}$ are distinct from the images of the points at infinity is not satisfied, then we choose another point P_0 for which the assumption is satisfied. We can then perform a change of coordinates τ over $\mathcal{O}$ that moves P_0 to infinity and apply the result above. By Corollary 4.6 (note that $v(\tau) = 0$ in this case) and the fact that $v(\tau(x)) = v(x)$, $\mu(P)$ is unchanged by τ. □

Remark 8.6. We see from the proof that for two points Q, Q' both having image on a component Γ, but with distinct reductions that are also distinct from those of P_0 and $\iota(P_0)$, we always have

$$\mu([(Q) - (Q')]) = -2g_\nu(\Gamma, \Gamma) - 2g_\nu(\Gamma, \Gamma_0 + \Gamma_0') - c\,.$$

So the assumption that this value does not depend on the choice of Q and Q' is not really necessary.

Theorem 8.7. *Let C be a smooth projective curve of genus 2 defined over a non-archimedean local field k, given by an integral Weierstrass model. Let J be the Jacobian of C and $\mathcal{J}$ its Néron model over $S = \operatorname{Spec} \mathcal{O}$. Assume that the minimal proper regular model $\mathcal{C}^{\min}$ of C over S is semistable and that μ factors through the component group $\Phi(\mathfrak{k})$ of $\mathcal{J}$. Let $P \in J(k)$ be such that its image in $\Phi(\mathfrak{k})$ is $[\Gamma_1 - \Gamma_2]$, where Γ_1 and Γ_2 are components of the special fiber of $\mathcal{C}^{\min}$. Then we have*

$$\mu(P) = r(\Gamma_1, \Gamma_2)\,.$$

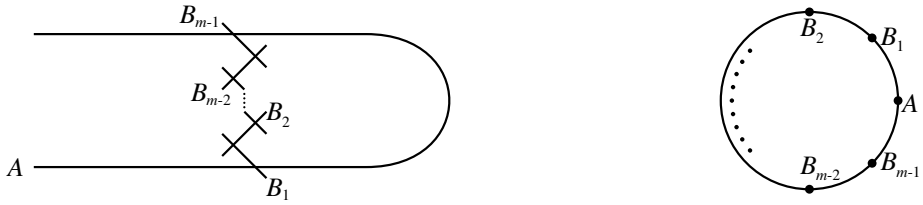

FIGURE 1. The special fiber of reduction type $[I_{m-0-0}]$ and its reduction graph

Proof. Since μ factors through $\Phi(\mathfrak{k})$, it follows that $\mu([(P_1)-(P_2)])$ vanishes when P_1 and P_2 map to the same component on the special fiber of $\mathcal{C}^{\min}$ and in general depends only on the components P_1 and P_2 map to. This shows that assumptions (i) to (iii) in Proposition 8.5 are satisfied with $\mu_1 = \mu_2 = 0$. The claim follows. □

9. Formulas and bounds for $\mu(P)$ in the nodal reduction case

In this section and the next, we will deduce explicit formulas for $\mu(P)$ when we have a stably minimal Weierstrass model $\mathcal{C}$.

In this section we shall use Theorem 8.7 and Remark 8.4 to find explicit formulas for $\mu(P)$ whenever C/k has nodal reduction, i.e., the special fiber $\mathcal{C}_v$ of $\mathcal{C}$ is reduced and all multiplicities are at most 2. In this case $\mathcal{C}$ is semistable and therefore it has rational singularities. Let $\Delta = \Delta(\mathcal{C})$ denote the discriminant of $\mathcal{C}$; we assume that there is at least one node, so that $v(\Delta) > 0$.

Since there are at most three nodes in the special fiber of $\mathcal{C}$, we have to consider three different cases. We assume that $x = (x_1, x_2, x_3, x_4)$ are normalized Kummer coordinates for a point $P \in J(k)$.

First suppose that there is a unique node in the special fiber of $\mathcal{C}$ and set $m = v(\Delta)$. In the notation of Namikawa and Ueno [NU73] this is reduction type $[I_{m-0-0}]$. If $m = 1$, then $\mathcal{C}$ is regular over S. In general, there is a unique component, which we denote by A, of genus 1 in the special fiber of $\mathcal{C}^{\min}$. As in the case of multiplicative reduction of elliptic curves (see for example [Sil94]), the singular point on the special fiber is replaced by a string of $m-1$ components of $\mathcal{C}^{\min}$, all of genus 0 and multiplicity 1. We choose one of the two components intersecting A and call it B_1 and number the other components $B_2, \dots, B_{m-1}$ consecutively as in Figure 1.

Using [BLR90, Thm. 9.6.1], it is easy to see that the geometric component group $\Phi(\bar{\mathfrak{k}})$ of the Néron model is generated by $[B_1 - A]$ and is isomorphic to $\mathbb{Z}/m\mathbb{Z}$. We have $[B_j - A] = j \cdot [B_1 - A]$ in $\Phi(\bar{\mathfrak{k}})$.

We set $B_0 := B_m := A$. Then we have the following result.

Proposition 9.1. *Suppose that there is a unique node in the special fiber of $\mathcal{C}$; let m and the notation for the components of the special fiber of $\mathcal{C}^{\min}$ be as above. If $P \in J(k)$ maps to $[B_i - A]$ in the component group, then we have*

$$\mu(P) = \frac{i(m-i)}{m}\,.$$

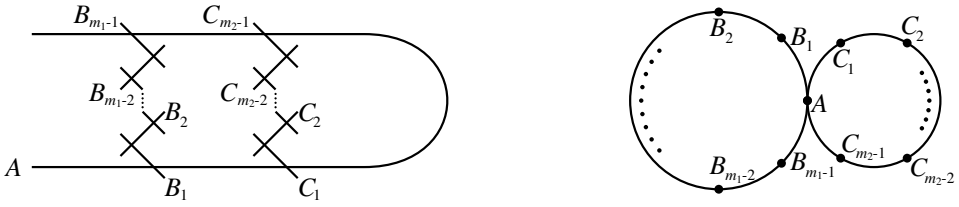

FIGURE 2. The special fiber of reduction type $[I_{m_1-m_2-0}]$ and its reduction graph

Proof. Since the given model is semistable, we can use Theorem 8.7 and Remark 8.4. One choice of g as in Remark 8.4 is given by

$$g(B_j) = \begin{cases} -\dfrac{j(m-i)}{m} & \text{if } 0 \le j \le i, \\ -\dfrac{i(m-j)}{m} & \text{if } i \le j \le m. \end{cases}$$

Then

$$\mu(P) = r(B_i, A) = -\big(g(B_i) - g(A)\big) = \frac{i(m-i)}{m},$$

as claimed. □

Remark 9.2. Proposition 9.1 resembles the formula for the canonical local height on an elliptic curve with split multiplicative reduction given, for instance, in [Sil88].

Now suppose that there are precisely two nodes in the special fiber of $\mathcal{C}$. The reduction type is $[I_{m_1-m_2-0}]$ in the notation of [NU73], where $m_1, m_2 \ge 1$ and $m_1 + m_2 = v(\Delta)$. The special fiber of $\mathcal{C}^{\min}$ is obtained by blowing up the two singular points of the special fiber of $\mathcal{C}$ repeatedly and replacing them with a chain of $m_1 - 1$ and $m_2 - 1$ curves of genus 0, respectively. We call these components $B_1, \dots, B_{m_1-1}, C_1, \dots, C_{m_2-1}$, numbered as in Figure 2, where A contains all images of points reducing to a nonsingular point and we pick components B_1 and C_1 intersecting A as in the case of a unique node. The component group $\Phi(\bar{\mathfrak{k}})$ is isomorphic to $\mathbb{Z}/m_1\mathbb{Z} \times \mathbb{Z}/m_2\mathbb{Z}$ and is generated by $[B_1 - A]$ and $[C_1 - A]$; this follows again using [BLR90, Thm. 9.6.1]. If we have $m_1 = 1$ or $m_2 = 1$, then the corresponding singular point on the special fiber of $\mathcal{C}$ is regular and is therefore not blown up.

We set $B_0 := B_{m_1} := C_0 := C_{m_2} := A$. Then every element of the component group has a representative of the form $[B_i - C_j]$ with $0 \le i \le m_1$ and $0 \le j \le m_2$. The following result expresses $\mu(P)$ in terms of this representative.

Proposition 9.3. *Suppose that there are exactly two nodes in the special fiber of $\mathcal{C}$; let m_1 and m_2 and the notation for the components of the special fiber of $\mathcal{C}^{\min}$ be as above. If $P \in J(k)$ maps to $[B_i - C_j]$ in the component group, then we have*

$$\mu(P) = \frac{i(m_1 - i)}{m_1} + \frac{j(m_2 - j)}{m_2}.$$

Proof. This is an easy computation along the same lines as in the proof of Proposition 9.1. □

The final case that we have to consider is the case of three nodes in the special fiber of $\mathcal{C}$, which then has two components. We call these components A and E. The special

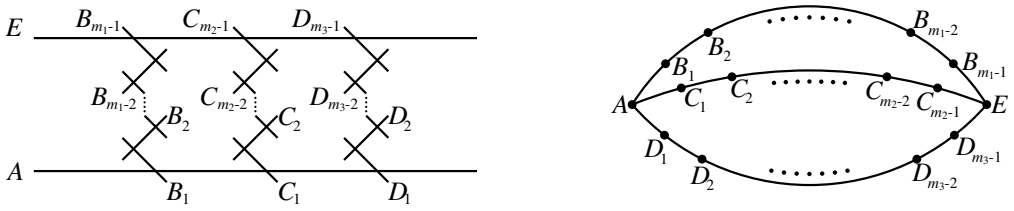

FIGURE 3. The special fiber of reduction type $[I_{m_1-m_2-m_3}]$ and its reduction graph

fiber of the minimal proper regular model is obtained using a sequence of blow-ups of the singular points; they are replaced by a chain of m_i-1 curves of genus 0 and multiplicity 1, respectively, where $v(\Delta) = m_1 + m_2 + m_3$. Hence the special fiber of $\mathcal{C}^{\min}$ contains the two components A and E, connected by three chains of curves of genus 0 that we call $B_1, \dots, B_{m_1-1}$, $C_1, \dots, C_{m_2-1}$ and $D_1, \dots, D_{m_3-1}$, respectively, where B_1, C_1 and D_1 intersect A, as shown in Figure 3. The reduction type is $[I_{m_1-m_2-m_3}]$.

By [BLR90, Prop. 9.6.10], the group $\Phi(\bar{\mathfrak{k}})$ is isomorphic to $\mathbb{Z}/d\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$, where

$$d = \gcd(m_1, m_2, m_3) \quad \text{and} \quad n = \frac{m_1m_2 + m_1m_3 + m_2m_3}{d}\,.$$

We set $B_0 := C_0 := D_0 := A$ and $B_{m_1} := C_{m_2} := D_{m_3} := E$. Then it is not hard to see that each element of $\Phi(\bar{\mathfrak{k}})$ can be written in one of the forms

$$[B_i - C_j], \qquad [C_j - D_l] \qquad \text{or} \qquad [D_l - B_i]$$

with $0 \le i \le m_1$, $0 \le j \le m_2$, $0 \le l \le m_3$. The following result allows us to express $\mu(P)$ for any $P \in J(k)$ in terms of the component P maps to.

Proposition 9.4. *Suppose that there are three nodes in the special fiber of $\mathcal{C}$; let m_1, m_2, m_3 and the notation for the components of the special fiber of $\mathcal{C}^{\min}$ be as above. If P maps to $[B_i - C_j]$ in the component group for some $0 \le i \le m_1$ and $0 \le j \le m_2$, then we have*

$$\mu(P) = \frac{m_2 i(m_1 - i) + m_3(i + j)(m_1 - i + m_2 - j) + m_1 j(m_2 - j)}{m_1m_2 + m_1m_3 + m_2m_3}\,.$$

The formulas for $[C_j - D_l]$ and $[D_l - B_i]$ are analogous.

Proof. The proof is analogous to those of Propositions 9.1 and 9.3. To find g, use that it is piecewise linear on the segments $AB_1 \dots B_i$, $B_i \dots B_{m_1-1}E$, $AC_1 \dots C_j$, $C_j \dots C_{m_2-1}E$ and $AD_1 \dots D_{m_3-1}E$ and the relations at the vertices A, E, B_i and C_j. □

Remark 9.5. Using the relation $\varepsilon(P) = 4\mu(P) - \mu(2P)$, one can show by a somewhat tedious computation involving a number of different cases that if the image of P in $\Phi(\mathfrak{k})$ is $[\Gamma_1 - \Gamma_2]$, where Γ_1 and Γ_2 are components of the special fiber of $\mathcal{C}^{\min}$, then $\varepsilon(P)$ is the 'distance' between Γ_1 and Γ_2 in the reduction graph, where the 'length' of the path between B_i and B_j (say, analogously for C_i, C_j and D_i, D_j) is $\min\{2|i - j|, m_1\}$ and otherwise, 'lengths' are additive. In particular, if $\Phi(\mathfrak{k}) = \Phi(\bar{\mathfrak{k}})$, then

$$\gamma = \max\{\varepsilon(P) : P \in J(k)\} = \max\{m_i + m_j - \delta_{ij} : 1 \le i < j \le 3\}\,,$$

where $\delta_{ij} = 0$ if both m_i and m_j are even, and $\delta_{ij} = 1$ otherwise.

Remark 9.6. In order to use the results of this section to actually compute $\mu(P)$ for a given point $P \in J(k)$, we need to be able to find the component of $\mathcal{J}_v$ that P reduces to. One approach is to find P_1 and $P_2 \in C$ such that $P = [(P_1) - (P_2)]$ and find the reductions of P_1 and P_2 to $\mathcal{C}_v^{\min}$. Another approach is to use a transformation (possibly defined over an unramified extension of k) to move the singular points to $\infty, (0,0)$ and $(1,0)$, respectively. Then we can (possibly after applying another transformation) read off the component that P maps to directly from the Kummer coordinates of P.

The discussion of this section shows that we get the following results on the local height constant $\beta = \max\{\mu(P) : P \in J(k)\}$. Recall that $\gamma = \max\{\varepsilon(P) : P \in J(k)\}$ and that $\gamma/4 \le \beta \le \gamma/3$. We will see that in many cases the lower bound is attained.

Let P be a node on $\mathcal{C}_v$; it is defined over a finite extension of $\mathfrak{k}$. We say that the node P is *split* if the two tangent directions of the branches at P are defined over every extension that P is defined over, otherwise P is *non-split*. We say that P is *even* if its contribution m_i to the valuation of the discriminant is even, and *odd* otherwise.

Corollary 9.7. *Suppose that C/k is a smooth projective curve of genus 2 given by an integral Weierstrass model $\mathcal{C}$ such that there is a unique node in the special fiber of $\mathcal{C}$ and let $m = v(\Delta)$. Then we have*

$$\beta = \frac{1}{2m}\left\lfloor \frac{m^2}{2} \right\rfloor \le \frac{v(\Delta)}{4}.$$

if the node is split or even, and $\beta = 0$ otherwise.

Proof. This follows from Proposition 9.1, taking into account that if m is odd and the node is non-split, then the group $\Phi(\mathfrak{k})$ is trivial. □

Remark 9.8. Using the relation $\varepsilon(P) = 4\mu(P) - \mu(2P)$, one can check that

$$\varepsilon(P) = 2\min\{i, m-i\} \qquad \text{if } P \text{ maps to } [B_i - A] \text{ in } \Phi(\mathfrak{k}).$$

If m is even (and $\beta > 0$), then $\beta = m/4 = \gamma/4$. If m is odd, then $\beta = (m - 1/m)/4$ and $\gamma = m - 1$, so $\beta/\gamma = (1 + 1/m)/4$ approaches $1/4$ as $m \to \infty$, but for $m = 3$ (the worst case), we have $\beta = \gamma/3$.

Corollary 9.9. *Suppose that C/k is a smooth projective curve of genus 2 given by an integral Weierstrass model $\mathcal{C}$ such that there are exactly two nodes in the special fiber of $\mathcal{C}$. Let $v(\Delta) = m_1 + m_2$ as above. Then we have*

$$\beta = \frac{1}{2m_1}\left\lfloor \frac{m_1^2}{2} \right\rfloor + \frac{1}{2m_2}\left\lfloor \frac{m_2^2}{2} \right\rfloor \le \frac{v(\Delta)}{4}$$

if each of the nodes is split or even,

$$\beta = \frac{1}{2m_i}\left\lfloor \frac{m_i^2}{2} \right\rfloor$$

if the node corresponding to m_i is split or even and the other node is non-split and odd, and $\beta = 0$ if both nodes are non-split and odd.

Proof. This follows from Proposition 9.3, taking into account the action of Frobenius on $\Phi(\bar{\mathfrak{k}})$. □

If we have three nodes, then it helps to take the field of definition of the nodes into account.

Corollary 9.10. *Suppose that C/k is a smooth projective curve of genus 2 given by an integral Weierstrass model $\mathcal{C}$ such that there are three nodes in the special fiber of $\mathcal{C}$. We say that $\mathcal{C}$ is split if the two components A and E of the special fiber of $\mathcal{C}^{\min}$ are defined over $\mathfrak{k}$, otherwise $\mathcal{C}$ is non-split. Let $v(\Delta) = m_1 + m_2 + m_3$ as above and set $M = m_1m_2 + m_1m_3 + m_2m_3$.*

(a) *If all nodes are $\mathfrak{k}$-rational, $\mathcal{C}$ is split, and we have $m_1 \geq m_3$ and $m_2 \geq m_3$, then*

$$\beta = \frac{1}{2M}\left(m_2\left\lfloor\frac{m_1^2}{2}\right\rfloor + m_3\left\lfloor\frac{(m_1+m_2)^2}{2}\right\rfloor + m_1\left\lfloor\frac{m_2^2}{2}\right\rfloor\right) \leq \frac{m_1+m_2}{4} < \frac{v(\Delta)}{4}\,.$$

(b) *If all nodes are $\mathfrak{k}$-rational, but $\mathcal{C}$ is non-split, then*

$$\beta = \max\{0\} \cup \left\{\frac{m_i+m_j}{4} : 1 \leq i < j \leq 3,\ m_i \text{ and } m_j \text{ even}\right\}.$$

(c) *If two of the nodes lie in a quadratic extension of $\mathfrak{k}$ and are conjugate over $\mathfrak{k}$ and one is $\mathfrak{k}$-rational, then*

$$\beta = \begin{cases} \dfrac{m_1}{M}\max\left\{\left\lfloor\dfrac{m_1^2}{2}\right\rfloor + m_1m_3, \left\lfloor\dfrac{m_3^2}{2}\right\rfloor + m_1\left\lfloor\dfrac{m_3}{2}\right\rfloor\right\}, & \text{if } \mathcal{C} \text{ is split,} \\ \dfrac{m_1}{2}, & \text{if } \mathcal{C} \text{ is non-split and } m_1 \text{ is even,} \\ 0, & \text{otherwise.} \end{cases}$$

where m_3 corresponds to the rational node (and $m_1 = m_2$).

(d) *If all nodes are defined over a cubic extension of $\mathfrak{k}$ and are conjugate over $\mathfrak{k}$, then $m_1 = m_2 = m_3 = v(\Delta)/3$ and*

$$\beta = \begin{cases} \dfrac{v(\Delta)}{9}, & \text{if } \mathcal{C} \text{ is split,} \\ 0, & \text{otherwise.} \end{cases}$$

Proof. The proof of (a) follows easily from Proposition 9.4.

For the other cases, note that in the non-split case, some power of Frobenius acts as negation on the component group $\Phi(\bar{\mathfrak{k}})$, so the only elements of $\Phi(\mathfrak{k})$ are elements of order 2 in $\Phi(\bar{\mathfrak{k}})$, which correspond to $[B_{m_1/2} - C_{m_2/2}]$ if m_1 and m_2 are even (where μ takes the value $(m_1+m_2)/4$), and similarly with the obvious cyclic permutations.

In the situation of (c), we must have $m_1 = m_2$. If $P = [(P_1) - (P_2)] \in J(k)$ and $P_1 \in C(\bar{k})$ maps to one of the conjugate nodes, then P_2 must map to the other, so all $P \in J(k)$ must map to a component of the form $[B_i - C_j]$ or $[D_i - D_j]$. Now the result in the split case follows from a case distinction depending on whether $m_1 \leq m_3$ or not. In the non-split case, the only element of order 2 that is defined over $\mathfrak{k}$ is $[B_{m_1/2} - C_{m_1/2}]$ if it exists.

In the situation of (d), the group $\Phi(\mathfrak{k})$ is of order 3 (generated by $[E - A]$) in the split case and trivial in the non-split case. □

Extending the valuation $v\colon k^\times \twoheadrightarrow \mathbb{Z}$ to $\bar{v}\colon \bar{k}^\times \to \mathbb{Q}$, we get extensions of ε and μ to $J(\bar{k})$. Denote $\max\{\mu(P) : P \in J(\bar{k})\}$ by $\bar{\beta}$ and $\max\{\varepsilon(P) : P \in J(\bar{k})\}$ by $\bar{\gamma}$. Then by the discussion at the beginning of Section 8 and the results above, we find that

$$\bar{\beta} = \frac{\bar{\gamma}}{4} = \frac{v(\Delta)}{4}\,,$$

when there are one or two nodes, and

$$\frac{v(\Delta)}{6} \leq \bar{\beta} = \frac{\bar{\gamma}}{4} = \frac{v(\Delta) - \min\{m_1,m_2,m_3\}}{4} < \frac{v(\Delta)}{4}\,,$$

when there are three nodes.

10. Formulas and bounds for $\mu(P)$ in the cuspidal reduction case

In this section we consider the case of a stably minimal Weierstrass model $\mathcal{C}$ such that there are (one or two) points of multiplicity 3 on the special fiber. In the notation of Namikawa and Ueno [NU73], the reduction type is of the form $[\mathcal{K}_1 - \mathcal{K}_2 - l]$, where $l \geq 0$ and $\mathcal{K}_j$ is an elliptic Kodaira type for $j \in \{1, 2\}$. By [Liu94], we have

$$\Phi(\bar{\mathfrak{k}}) \cong \Phi_1(\bar{\mathfrak{k}}) \times \Phi_2(\bar{\mathfrak{k}}),$$

where Φ_j is the component group of an elliptic curve with Kodaira type $\mathcal{K}_j$. As in the previous section, we write $\Delta = \Delta(\mathcal{C})$ for the discriminant of the model $\mathcal{C}$.

Lemma 10.1. *Suppose that the residue characteristic of k is not 2. Let C be given by a stably minimal Weierstrass model with reduction type $[\mathcal{K}_1 - \mathcal{K}_2 - l]$. Then after at most a quadratic unramified extension of k there is a stably minimal Weierstrass model*

$$\mathcal{C}\colon Y^2 = F(X, Z) = f_6X^6 + f_5X^5Z + f_4X^4Z^2 + X^3Z^3 + f_2X^2Z^4 + f_1XZ^5 + f_0Z^6$$

of C, isomorphic to the given model of C, such that the elliptic curve E_1 with Weierstrass model

$$\mathcal{E}_1\colon Y^2Z = X^3 + f_2X^2Z + f_1XZ^2 + f_0Z^3$$

has Kodaira type $\mathcal{K}_1$ and the elliptic curve E_2 with Weierstrass model

$$\mathcal{E}_2\colon Y^2Z = X^3 + f_4X^2Z + f_5XZ^2 + f_6Z^3$$

has Kodaira type $\mathcal{K}_2$.

Proof. After possibly making a quadratic unramified extension and applying a transformation, we can suppose that one cusp on the special fiber is $P = (0, 0)$ and that the other singularity (if it exists) is the unique point at infinity on the special fiber.

Because the residue characteristic is not 2, we may assume that $\mathcal{C}$ has $H = 0$ and $f_3 = 1$ (the latter can be achieved by scaling the variables in a suitable way; note that f_3 must be a unit). By Hensel's Lemma there is a factorization $F = F_1F_2$, where F_1 is a cubic form reducing to X^3. Similarly, F_2 reduces to Z^3 if there is a cusp at infinity and to $Z^2(aX + Z)$ with $a \neq 0$ if there is a node at infinity. Consider the elliptic curves given by the Weierstrass models

$$\mathcal{D}_1\colon Y^2Z = F_1(X, Z)$$

and

$$\mathcal{D}_2\colon Y^2Z = F_2(Z, X).$$

Then we claim that $\mathcal{D}_1$ has Kodaira type $\mathcal{K}_1$. To see this, note that in order to desingularize $\mathcal{C}$ (respectively, $\mathcal{D}_1$) in P (respectively, in the singular point $Q = (0, 0)$ on the special fiber of $\mathcal{D}_1$) we first blow up $\mathcal{C}$ (respectively, $\mathcal{D}_1$) in P (respectively, Q). Under this blow-up map, the strict transform of P (respectively, Q) is contained in the chart $\mathcal{C}^1$ (respectively, $\mathcal{D}_1^1$) obtained by dividing the x- and y-coordinates by π. But in fact $\mathcal{C}^1$ and $\mathcal{D}_1^1$ have identical special fiber and this continues to hold after further blow-ups (if any are necessary), which suffices to prove that $\mathcal{D}_1$ has Kodaira type $\mathcal{K}_1$. An analogous argument shows that $\mathcal{D}_2$ has Kodaira type $\mathcal{K}_2$.

To complete the proof of the lemma, we therefore only need to make sure that if F factors as $F = F_1F_2$, then $\mathcal{E}_i$ has the same reduction type as $\mathcal{D}_i$ for $i = 1, 2$. This is certainly satisfied if the coefficients of $\mathcal{E}_i$ and $\mathcal{D}_i$ agree modulo π^{N_i+1}, where N_i is the number of blow-ups needed to construct the minimal desingularization of $\mathcal{D}_i$. Writing out the coefficients of F in terms of the coefficients of F_1 and F_2, we see that this can

be achieved by acting on a stably minimal Weierstrass model (with $H = 0$) for C via a suitable element of $\mathrm{GL}_2(\mathcal{O})$. □

Remark 10.2. If the residue characteristic is 2, then it is not hard to see that one can also construct a stably minimal Weierstrass model $\mathcal{C}$ and elliptic Weierstrass models $\mathcal{E}_1$ and $\mathcal{E}_2$ as in the lemma in a similar way. The construction is more cumbersome, since we cannot assume $H = 0$.

In view of Theorem 7.5 we want a condition for $\mathcal{C}$ to have rational singularities.

Lemma 10.3. *The model $\mathcal{C}$ has rational singularities if and only if $l = 0$.*

Proof. We may assume that $\mathcal{C}$ is as in Lemma 10.1 or Remark 10.2. By [Art66, Prop. 1], the assertion that P is a rational singularity depends only on the strict transform of its preimage under an arbitrary desingularization. So let $\xi\colon \mathcal{C}' \to \mathcal{C}$ be the minimal desingularization of P and let $\xi_1\colon \mathcal{E}_1' \to \mathcal{E}_1$ be the minimal desingularization of the singular point $Q = (0,0)$ on the Weierstrass model $\mathcal{E}_1$ (cf. Lemma 10.1 and Remark 10.2). By the proof of Lemma 10.1, the strict transforms of P and Q under ξ and ξ_1, respectively, are isomorphic to each other. In particular, P is a rational singularity if and only if Q is a rational singularity. If ∞ is a cusp as well (if it is a node, then it is a rational singularity anyway), then we get the analogous statement for $\mathcal{E}_2$ by symmetry.

However, by [Con05, Corollary 8.4] a Weierstrass model of an elliptic curve has rational singularities if and only if it is minimal. But it is easy to see that $\mathcal{E}_1$ and $\mathcal{E}_2$ are simultaneously minimal if and only if $l = 0$. □

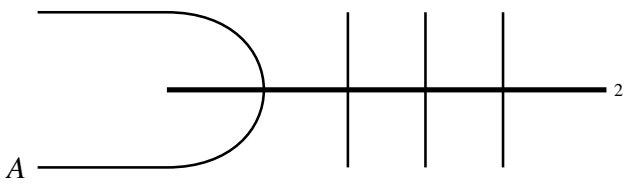

FIGURE 4. The special fiber of reduction type $[I_0 - I_0^* - 0]$

Suppose $l = 0$. If both $\mathcal{K}_1$ and $\mathcal{K}_2$ are good or multiplicative, then we are in the situation $[I_{m_1-m_2-0}]$ for some $m_1, m_2 \geq 0$, which we have discussed in the previous section. So we may assume that at least one of the $\mathcal{K}_j$ is additive, say $\mathcal{K}_2$. Then the special fiber of the minimal proper regular model $\mathcal{C}^{\min}$ of C looks like the special fiber of $\mathcal{E}_2$, but with one of the rational curves replaced by (see [NU73])

- a curve A of genus 1, if $\mathcal{K}_1 = I_0$, see Figure 4 for the case $\mathcal{K}_2 = I_0^*$;
- one of the rational components of $\mathcal{K}_1$, otherwise.

Suppose now that $l > 0$. Then the special fiber of $\mathcal{C}^{\min}$ consists of the Kodaira types $\mathcal{K}_1$ and $\mathcal{K}_2$, connected by a chain of $l - 1$ rational curves, see for example Figure 5. According to Lemma 10.3, not all singularities of the given stably minimal Weierstrass model $\mathcal{C}$ are rational. The following example shows that in this situation $\varepsilon(P) \neq 0$ and hence $\mu(P) \neq 0$ can indeed occur for $P \in J_0(k)$.

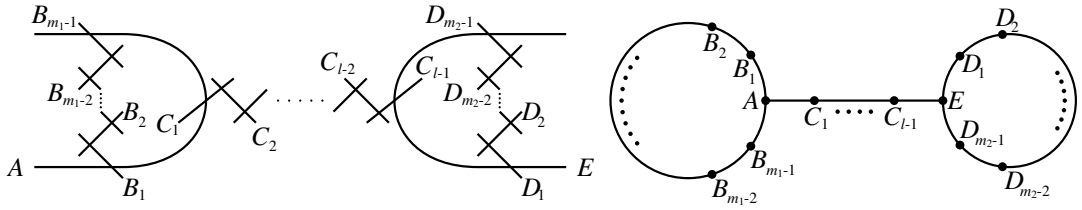

FIGURE 5. The special fiber of reduction type $[I_{m_1} - I_{m_2} - l]$ and its reduction graph

Example 10.4. Let p be an odd prime and let $C/\mathbb{Q}_p$ be given by

$$Y^2 = Z(X^2 + Z^2)(X^3 + p^5 X Z^2 + p^8 Z^3)\,.$$

Let $P_1 = (0, p^4) \in C(\mathbb{Q}_p)$ and $P_2 = \iota(P_1)$. We have reduction type $[I_0 - III - 1]$ and hence $\#\Phi(\bar{\mathfrak{k}}) = 2$. It turns out that both P_1 and P_2 map to the same component and so we have $P = [(P_1) - (P_2)] \in J_0(k)$. The image of P on the Kummer surface is of the form $(x_1 : 0 : 0 : x_4)$, where $v(x_4) - v(x_1) = 2$. We get $\varepsilon(P) = \varepsilon(2P) = 6$ and $\mu(P) = \mu(2P) = 2$.

The case of semistable reduction, corresponding to reduction type $[I_{m_1} - I_{m_2} - l]$, see Figure 5, deserves special attention. Here $l \geq 1$, by the discussion above. Note that $m_1 = 0$ (or $m_2 = 0$) is possible; in that case A (or E) is a curve of genus 1 and there are no components B_i (or D_i). If $m_1 = 1$ (or $m_2 = 1$), then A (or E) is a nodal curve (and again there are no B_i or D_i). After perhaps an unramified quadratic extension, we can assume that all components in the 'chain' that connects the two polygons in the special fiber of $\mathcal{C}^{\min}$ are defined over $\mathfrak{k}$. There are then $l+1$ different (meaning pairwise non-isomorphic over $\mathcal{O}$) minimal Weierstrass models of the curve, compare the proof of Lemma 5.4. Explicitly, these models can be taken to have the form

$$\begin{aligned} \mathcal{C}_j \colon Y^2 + (h_0\pi^{3j}Z^3 + h_1\pi^j Z^2 X + h_2\pi^{l-j} Z X^2 + h_3\pi^{3(l-j)}X^3)Y \\ = f_0\pi^{6j}Z^6 + f_1\pi^{4j}XZ^5 + f_2\pi^{2j}X^2Z^4 + X^3Z^3 \\ + f_4\pi^{2(l-j)}X^4Z^2 + f_5\pi^{4(l-j)}X^5Z + f_6\pi^{6(l-j)}X^6 \end{aligned} \tag{10.1}$$

for $j = 0, 1, \ldots, l$, where

$$\begin{aligned} y^2 + h_1xy + h_0y &= x^3 + f_2x^2 + f_1x + f_0 \qquad \text{and} \\ y^2 + h_2xy + h_3y &= x^3 + f_4x^2 + f_5x + f_6 \end{aligned}$$

are minimal Weierstrass equations of elliptic curves of reduction types I_{m_1} and I_{m_2}, respectively. Such a model corresponds to the vertex C_j of the reduction graph (where we set $C_0 = A$ and $C_l = E$); the corresponding component of the special fiber of $\mathcal{C}^{\min}$ is the one that is visible in the special fiber of $\mathcal{C}_j$. The valuation of the discriminant of $\mathcal{C}_j$ is $m_1 + m_2 + 12l$ and does not depend on j.

Let $P_1, P_2 \in C(k)$ reduce to components Γ_1 and Γ_2 of the special fiber of $\mathcal{C}^{\min}$, respectively. Consider the model $\mathcal{C}_j$ of C. If there is a simple path from Γ_1 to Γ_2 in the reduction graph that passes through C_j, then we say that $\mathcal{C}_j$ *lies between* P_1 and P_2. We denote the μ-function computed with respect to $\mathcal{C}_j$ by μ_j.

Proposition 10.5. *Let $P_1, P_2 \in C(k)$ be points reducing to components Γ_1 and Γ_2 of the special fiber of $\mathcal{C}^{\min}$ and let $j \in \{0, 1, \ldots, l\}$. Define $j_{\min}$ and $j_{\max}$ to be the*

smallest, respectively largest, $j' \in \{0,1,\dots,l\}$ such that $\mathcal{C}_{j'}$ lies between P_1 and P_2. Let $P = [(P_1) - (P_2)] \in J(k)$. Then

$$r(\Gamma_1,\Gamma_2) + j_{\max} - j_{\min} \le \mu_j(P) \le r(\Gamma_1,\Gamma_2) + |j - j_{\max}| + |j - j_{\min}|\,.$$

If $\mathcal{C}_j$ lies between P_1 and P_2, then the inequalities are equalities.

Proof. First note that the last statement follows from the first, since $j_{\min} \le j \le j_{\max}$ implies $j_{\max} - j_{\min} = |j - j_{\max}| + |j - j_{\min}|$.

Let $B_0 = B_{m_1} = A$ and $D_0 = D_{m_2} = E$. We prove a number of lemmas.

Lemma 10.6. *If $j = j_{\max} = j_{\min} \in \{0,l\}$, then $\mu_j(P) = r(\Gamma_1,\Gamma_2)$.*

Proof. We assume that $j = j_{\max} = j_{\min} = l$; the other case is analogous. Then Γ_1 and Γ_2 are both of the form D_i, and we consider the model $\mathcal{C}_l$. We first claim that $\mu(P) = 0$ if $\Gamma_1 = \Gamma_2$, but the images of P_1 and P_2 on Γ_1 are distinct. This is clear if $\Gamma_1 = D_0 = E$, since in this case P is in the image of α, compare Lemmas 7.1 and 7.3. Otherwise, we note that the multiplicities on the special fiber of $\mathcal{C}_l$ are 1, 2 and 3. Transforming the equation over $\mathcal{O}$ if necessary, we can assume that its reduction is case 7 in Table 1 of [Sto02] or (if the residue characteristic is 2) case 5 in Table 2 here.

Recall that $\Gamma_1 = \Gamma_2 = D_i$, where we can assume $0 < i \le m_2/2$. Applying a transformation, we may assume that the points $P_1 = (\xi_1 : \eta_1 : 1)$ and $P_2 = (\xi_2 : \eta_2 : 1)$ both reduce to $(0:0:1)$ modulo π and that $m_2 = \min\{v(f_0), 2v(f_1)\}$. First suppose that $i < m_2/2$. We then have $v(\xi_1) = v(\xi_2) = v(\xi_1 - \xi_2) = i$. Normalizing the Kummer coordinates x of P so that $x_1 = 1$, we can check that $v(x_2)$ and $v(x_3)$ are positive, but that $v(x_4) = 0$. This follows because $\Gamma_1 = D_i = \Gamma_2$ implies that $v(f_2\xi_1\xi_2 + 2\eta_1\eta_2) = 2i$ if $\mathrm{char}(\mathfrak{k}) \ne 2$ and $H = 0$ and that $v(\xi_1\eta_2 + \xi_2\eta_1) = 2i$ if $\mathrm{char}(\mathfrak{k}) = 2$. By a similar argument, the reduction of the image of P on the Kummer surface has non-vanishing last coordinate if m_2 is even and $i = m_2/2$. According to the tables, this implies that $\varepsilon(P) = 0$ and therefore also $\mu(P) = 0$.

Now consider the case that Γ_1 and Γ_2 do not necessarily coincide. The considerations above imply that the assumptions of Proposition 8.5 are satisfied with $\mu_1 = \mu_2 = 0$ (where we use Lemma 3.7 for the first assumption); the proposition then establishes the claim. □

Lemma 10.7. *Assume that $\Gamma_1 = \Gamma_2 = C_j$ with $0 < j < l$. Then $\mu_j(P) = 0$.*

Proof. In this case, P is in the image of α, so the claim follows by Proposition 7.4. □

Note that Lemmas 10.6 and 10.7 establish the claim of Proposition 10.5 in all cases such that $j = j_{\min} = j_{\max}$.

Lemma 10.8. *Assume that both $\mathcal{C}_j$ and $\mathcal{C}_{j+1}$ lie between P_1 and P_2, where $0 \le j < l$. Then $\mu_j(P) = \mu_{j+1}(P)$.*

Proof. Let $\tau\colon (\xi : \eta : \zeta) \mapsto (\pi\xi : \eta : \pi^{-1}\zeta)$; then τ gives an isomorphism from the generic fiber of $\mathcal{C}_j$ to that of $\mathcal{C}_{j+1}$. The induced map on Kummer coordinates is

$$(x_1,x_2,x_3,x_4) \longmapsto (\pi^{-2}x_1, x_2, \pi^2 x_3, x_4)\,;$$

we have $v(\tau) = 0$. Since both $\mathcal{C}_j$ and $\mathcal{C}_{j+1}$ lie between P_1 and P_2, assuming that Γ_1 is to the left and Γ_2 to the right of C_j and C_{j+1}, we must have that the x-coordinate of P_1 on $\mathcal{C}_j$ does not reduce to infinity, whereas that of P_2 does. For normalized Kummer coordinates $x = (x_1,x_2,x_3,x_4)$ of P on the Kummer surface associated to $\mathcal{C}_j$, this implies $v(x_2) = 0$ (the point is not in the kernel of reduction, so $v(x_4) \ge \min\{v(x_1), v(x_2), v(x_3)\}$) and

$v(x_1) > 0$. Comparing valuations in the equation of $\mathcal{C}_j$, we see that $P_2 = (1 : \eta : \zeta)$ must have $v(\zeta) \geq 2$, which implies $v(x_1) \geq 2$. It follows that $v(\tau(x)) = 0 = v(x)$. By Corollary 4.6 we also have $\hat{\lambda}(\tau(x)) = \hat{\lambda}(x)$ (recall that $v(\tau) = 0$). Since

$$-v(x) - \mu_j(P) = \hat{\lambda}(x) = \hat{\lambda}(\tau(x)) = -v(\tau(x)) - \mu_{j+1}(P)\,,$$

the claim follows. □

Lemma 10.9. *If $\mathcal{C}_j$ lies between P_1 and P_2, then $\mu_j(P)$ depends only on Γ_1 and Γ_2.*

Proof. Let $P_1', P_2' \in C(k)$ be points also mapping to Γ_1 and Γ_2, respectively. We assume without loss of generality that Γ_1 is to the left of Γ_2. By Lemmas 10.6 or 10.7, we have that $\mu_{j_{\min}}([(P_1) - (P_1')]) = 0$ and $\mu_{j_{\max}}([(P_2) - (P_2')]) = 0$. Using Lemmas 10.8 and 3.7, we obtain

$$\begin{aligned}\mu_j\big([(P_1') - (P_2')]\big) &= \mu_{j_{\min}}\big([(P_1') - (P_2')]\big) = \mu_{j_{\min}}\big([(P_1) - (P_2')]\big) \\ &= \mu_{j_{\max}}\big([(P_1) - (P_2')]\big) = \mu_{j_{\max}}\big([(P_1) - (P_2)]\big) = \mu_j(P)\,.\end{aligned}$$ □

Lemma 10.10. *Let $P_1', P_2' \in C(k)$ be points mapping to distinct points on the same component of the special fiber of $\mathcal{C}^{\min}$ and let $P' = [(P_1') - (P_2')] \in J(k)$. Let j_0 be the unique index such that $\mathcal{C}_{j_0}$ lies between P_1' and P_2'. Then $\mu_j(P') = 2|j - j_0|$.*

Proof. By Lemmas 10.6 and 10.7, we have $\mu_{j_0}(P') = 0$. Since the images of P_1' and P_2' on the special fiber of $\mathcal{C}^{\min}$ are distinct, P' is not in the kernel of reduction with respect to $\mathcal{C}_{j_0}$. If

$$x^{(j_0)} = (x_1^{(j_0)}, x_2^{(j_0)}, x_3^{(j_0)}, x_4^{(j_0)})$$

are normalized Kummer coordinates for P' on the Kummer surface associated to $\mathcal{C}_{j_0}$, we therefore have

$$0 = v(x^{(j_0)}) = \min\{v(x_1^{(j_0)}), v(x_2^{(j_0)}), v(x_3^{(j_0)})\}\,.$$

Applying a suitable power of τ (see the proof of Lemma 10.8), we find that

$$x^{(j)} = (\pi^{2(j_0-j)}x_1^{(j_0)}, x_2^{(j_0)}, \pi^{2(j-j_0)}x_3^{(j_0)}, x_4^{(j_0)})$$

are (not necessarily normalized) Kummer coordinates for P' on the Kummer surface associated to $\mathcal{C}_j$. For definiteness, assume that $j > j_0$, the case $j = j_0$ being clear. Similarly to the proof of Lemma 10.8, we find that $0 = v(x^{(j_0)}) = v(x_1^{(j_0)})$, which implies that $v(x^{(j)}) = -2(j - j_0)$. In the same way as in the proof of Lemma 10.8, we deduce $\mu_j(P') = 2(j - j_0) = 2|j - j_0|$. □

To continue the proof of the proposition, we now first consider the case that $\mathcal{C}_j$ lies between P_1 and P_2. In this case, Lemmas 10.9 and 10.10 show that the assumptions in Proposition 8.5 hold with $\mu_1 = 2|j - j_{\min}|$ and $\mu_2 = 2|j - j_{\max}|$ or conversely. So the statement follows from Proposition 8.5 and $|j - j_{\max}| + |j - j_{\min}| = j_{\max} - j_{\min}$.

Now assume that $\mathcal{C}_j$ does not lie between P_1 and P_2. We assume for definiteness that $j > j_{\max}$. For normalized Kummer coordinates $x^{(j_{\max})}$ for $P = [(P_1) - (P_2)]$ on the Kummer surface associated to $\mathcal{C}_{j_{\max}}$, we have $v(x_2^{(j_{\max})}) \leq \min\{v(x_1^{(j_{\max})}), v(x_3^{(j_{\max})})\}$, compare the proof of Lemma 10.8 above. Then $x^{(j)} = \tau^{j-j_{\max}}(x^{(j_{\max})})$ are Kummer coordinates for $[(P_1) - (P_2)]$ on the Kummer surface associated to $\mathcal{C}_j$, and we have

$$v(x^{(j_{\max})}) - 2(j - j_{\max}) \leq v(x^{(j)}) \leq v(x^{(j_{\max})})\,.$$

It follows that

$$\begin{aligned}\mu_j(P) - \mu_{j_{\max}}(P) &\\ &= \bigl(-\hat\lambda(x^{(j)}) - v(x^{(j)})\bigr) - \bigl(-\hat\lambda(x^{(j_{\max})}) - v(x^{(j_{\max})})\bigr)\\ &= v(x^{(j_{\max})}) - v(x^{(j)}) \in \{0, 1, \dots, 2(j - j_{\max})\}.\end{aligned}$$

As $\mu_{j_{\max}}(P) = r(\Gamma_1, \Gamma_2) + j_{\max} - j_{\min}$ by the case already discussed, the result follows. □

Corollary 10.11. *Let $\mathcal{C}$ be a stably minimal Weierstrass model of C with discriminant Δ; assume that C has reduction type $[I_{m_1} - I_{m_2} - l]$ with $l > 0$. As usual, let*

$$\beta(\mathcal{C}) = \max\{\mu(P) : P \in J(k)\} \qquad \text{and} \qquad \bar\beta(\mathcal{C}) = \max\{\mu(P) : P \in J(\bar k)\}\,,$$

where μ is computed with respect to $\mathcal{C}$. Then we have

$$\beta(\mathcal{C}) \le \bar\beta(\mathcal{C}) = \frac{m_1 + m_2}{4} + 2l < \frac{v(\Delta)}{4} \qquad \text{and} \qquad \bar\beta \ge \frac{v(\Delta)}{6}\,.$$

Proof. The assumption on the reduction type implies that the model is equivalent to one of the form (10.1). Proposition 10.5 then gives upper bounds for $\mu([(P_1) - (P_2)])$, with $P_1, P_2 \in C(\bar k)$, depending on the images Γ_1 and Γ_2 of P_1 and P_2 in the reduction graph. The maximizing case occurs for $\Gamma_1 = B_{m_1/2}$ and $\Gamma_2 = D_{m_2/2}$, giving

$$\mu([(P_1) - (P_2)]) = r(B_{m_1/2}, D_{m_2/2}) + l = \tfrac14 m_1 + l + \tfrac14 m_2 + l\,.$$

For the remaining inequalities, recall that $v(\Delta) = m_1 + m_2 + 12l$ and that $l > 0$. □

We state a technical lemma, which will be needed for the proof of Theorem 10.13 below.

Lemma 10.12. *Suppose that the residue characteristic of k is not 2. Consider a degenerate Weierstrass equation of the form*

$$\mathcal{C}\colon Y^2 = f_0 Z^6 + f_1 X Z^5 + f_2 X^2 Z^4 + X^3 Z^3$$

and let

$$\mathcal{E}\colon y^2 = f_0 + f_1 x + f_2 x^2 + x^3$$

be an elliptic Weierstrass equation. If $Q_1 = (x_1, y_1)$ and $Q_2 = (x_2, y_2)$ are points in $\mathcal{E}(k)$, then $P_1 = (x_1 : y_1 : 1)$ and $P_2 = (x_2 : y_2 : 1)$ are points in $\mathcal{C}(k)$, and if $x_1, x_2 \in \mathcal{O}$, then $\mu_{\mathcal{C}}([(P_1) - (P_2)]) \le \mu_{\mathcal{E}}(Q_1 - Q_2)$.

Here $\mu_{\mathcal{E}}$ is the height correction function for the elliptic curve $\mathcal{E}$ and $\mu_{\mathcal{C}}$ denotes the height correction function defined in the same way as μ in the smooth case in terms of the equation $\mathcal{C}$.

Proof. Let $\underline{\delta}_{\mathcal{C}} = (\delta_{\mathcal{C},1}, \delta_{\mathcal{C},2}, \delta_{\mathcal{C},3}, \delta_{\mathcal{C},4})$ be the duplication polynomials on the Kummer surface associated to $\mathcal{C}$, and let $\underline{\delta}_{\mathcal{E}} = (\delta_{\mathcal{E},1}, \delta_{\mathcal{E}_2})$ be the duplication polynomials for the numerator and denominator of the x-coordinate associated to $\mathcal{E}$. Then a generic computation shows that, if $(\xi_1 : \xi_2 : \xi_3 : \xi_4)$ is the image of $[(P_1) - (P_2)]$ on the Kummer surface, we have $(\xi_4 : \xi_1) = x(Q_1 - Q_2)$. In addition, we find that (as polynomials in the ξ_j) $\delta_{\mathcal{C},1}(\xi_1, \xi_2, \xi_3, \xi_4) = \delta_{\mathcal{E},2}(\xi_4, \xi_1)$ and $\delta_{\mathcal{C},4}(\xi_1, \xi_2, \xi_3, \xi_4) = \delta_{\mathcal{E},1}(\xi_4, \xi_1)$.

That $P_1, P_2 \in \mathcal{C}(k)$ is obvious from the equations. For the last statement, we observe that $\min\{v(\xi_1), v(\xi_2), v(\xi_3), v(\xi_4)\} = \min\{v(\xi_1), v(\xi_4)\}$ (this is where we use that x_1 and x_2 are integral), which implies

$$\begin{aligned}\mu_{\mathcal{C}}([(P_1) - (P_2)]) &= \lim_{n\to\infty} 4^{-n} v\bigl(\underline{\delta}_{\mathcal{C}}^{\circ n}(\underline{\xi})\bigr) - v(\underline{\xi})\\ &\le \lim_{n\to\infty} 4^{-n} v\bigl(\underline{\delta}_{\mathcal{E}}^{\circ n}(\xi_4, \xi_1)\bigr) - \min\{v(\xi_1), v(\xi_4)\}\\ &= \mu_{\mathcal{E}}(Q_1 - Q_2)\,.\end{aligned}$$

□

The following consequence is useful for practical purposes. For simplicity, we state it for the case of residue characteristic $\neq 2$, but we expect that the statement remains true for residue characteristic 2.

Theorem 10.13. *Suppose that the residue characteristic of k is not 2. Let $\mathcal{C}$ be a stably minimal Weierstrass model of C such that C has reduction type $[\mathcal{K}_1 - \mathcal{K}_2 - l]$. Then*

$$\beta(\mathcal{C}) \leq \beta(\mathcal{K}_1) + \beta(\mathcal{K}_2) + 2l\,,$$

where $\beta(\mathcal{K})$ denotes the maximum of μ for an elliptic curve of reduction type $\mathcal{K}$ (taking the action of Frobenius into account), see Table 1 in [CPS06].

Proof. We may assume that the point(s) of multiplicity 3 on the special fiber are defined over $\mathfrak{k}$, at the cost of an at most quadratic unramified extension of k. Then we can move these points to have x-coordinates 0 and ∞, respectively, and so we can assume that our model $\mathcal{C}$ is as in Lemma 10.1. Let $P \in J(k)$; we write $P = [(P_1) - (P_2)]$ with points $P_1, P_2 \in C(k')$ for a finite extension k' of k such that the reduction of C over k' is semistable. We can find $\mathcal{C}_0$, $\mathcal{C} = \mathcal{C}_j$ and $\mathcal{C}_l$ as vertices in the reduction graph of the minimal proper regular model of C over k'. Then the part of the graph to the left of $\mathcal{C}_0$ corresponds to the reduction graph of $\mathcal{E}_1$ over k', in the sense that we consider a semistable model that dominates $\mathcal{E}_1$ (and is minimal with that property); the graph then is either a line segment (potentially good reduction) or a line segment joined to a circle (potentially multiplicative reduction), with $\mathcal{E}_1$ corresponding to the end of the line segment joined to the remaining graph of $\mathcal{C}$. Similarly, the part of the graph to the right of $\mathcal{C}_l$ corresponds to the reduction graph of $\mathcal{E}_2$ over k'.

Now assume that both P_1 and P_2 map (strictly) to the left of $\mathcal{C}_0$ in the reduction graph. This means that the x-coordinates of the points have positive valuation. We can then find points P_1' and P_2' in $\mathcal{E}_1(k')$ with the same x-coordinates as P_1 and P_2 and nearby y-coordinates. Then $P_1' - P_2'$ is in $\mathcal{E}_1(k)$ and P_1' and P_2' have the same images as P_1 and P_2 in the reduction graph. By our previous results for the semistable case, the value of (or at least the upper bound given in Proposition 10.5 for) $\mu_0(P)$ depends only on the part of the graph to the left of $\mathcal{C}_0$. We can therefore let l tend to infinity; then Lemma 10.12 and the discussion preceding Lemma 10.3 show that $\mu_0(P)$ is bounded by the value of $\mu_{\mathcal{E}_1}$ on the difference $P_1' - P_2'$. By the arguments in the proof of Proposition 10.5, we have that

$$\mu_{\mathcal{C}}(P) = \mu_j(P) \leq \mu_0(P) + 2j \leq \beta(\mathcal{K}_1) + 2l\,.$$

The case that P_1 and P_2 both map to the right of $\mathcal{C}_l$ is similar.

If (say) P_1 maps to the left of $\mathcal{C}_0$ and P_2 maps to the right of $\mathcal{C}_0$, but not to the right of $\mathcal{C}_l$, then by the formula of Proposition 10.5, we can bound $\mu_{\mathcal{C}}(P)$ by $\mu_1 + 2l$, where μ_1 comes from the part of the graph between P_1 and $\mathcal{C}_0$. By an argument similar to the one used in the previous paragraph, μ_1 can be bounded by $\mu_{\mathcal{E}_1}(P_1')$, where P_1' is the point on $\mathcal{E}_1$ corresponding to P_1 and we take the second point to be on the component visible in $\mathcal{C}_0$. If P_2 maps to the right of $\mathcal{C}_l$, then we similarly obtain a bound of the form $\mu_1 + \mu_2 + 2l \leq \beta(\mathcal{K}_1) + \beta(\mathcal{K}_2) + 2l$. The remaining cases are similar or follow directly from Proposition 10.5. □

The example in Section 19 demonstrates the effect of the improved bounds on β as given in the preceding section. For other examples the bounds established in this section will be similarly useful.

11. General upper and lower bounds for $\bar\beta$

In this section we derive an upper bound for the geometric height constant $\bar\beta(\mathcal{C})$ in the general case by reducing to the semistable situation. We also give a lower bound of the same order of magnitude. We note the following consequence of the results obtained so far, see the discussion at the end of Section 9 and Corollary 10.11.

Corollary 11.1. *Assume that $\mathcal{C}$ is a stably minimal Weierstrass model of C over k and that the minimal proper regular model $\mathcal{C}^{\min}$ of C over k has semistable reduction. Denoting the discriminant of $\mathcal{C}$ by Δ and writing $\bar\beta(\mathcal{C}) = \max\{\mu_{\mathcal{C}}(P) : P \in J(\bar k)\}$, where $\mu_{\mathcal{C}}$ denotes μ with respect to the model $\mathcal{C}$ and J is the Jacobian of C, we have*

$$\frac{v(\Delta)}{6} \le \bar\beta(\mathcal{C}) \le \frac{v(\Delta)}{4}\,.$$

When $\mathcal{C}^{\min}$ does not have semistable reduction, the idea is to pass to a suitable field extension k'/k and apply Corollary 11.1 over k'. In order to compare the corresponding geometric height constants $\bar\beta$, we need to analyze how μ changes under minimization. We first prove the following key lemma:

Lemma 11.2. *There exists a transformation $\tau\colon \mathcal{C} \to \mathcal{C}'$, defined over k, such that $\mathcal{C}'$ is a minimal Weierstrass model and*

$$v(\tau(x)) + v(\tau) \le v(x) \text{ for all } x \in \mathrm{KS}_{\mathbb{A}}\,.$$

Proof. If $\mathcal{C}$ is already minimal, then there is nothing to prove. Otherwise, [Liu96, Remarque 11] implies that we can compute a minimal Weierstrass model by going through the following steps for finitely many points P on the special fiber of $\mathcal{C}$.

(a) Move P to $(0,0)$.
(b) Scale x by $1/\pi$.
(c) Replace $\mathcal{C}$ by the normalization of the resulting model.

As transformations of the form (a) do not change $v(x)$ and have determinant of valuation 0, it suffices to prove

$$v(\tau(x)) + v(\tau) \le v(x) \text{ for all } x \in \mathrm{KS}_{\mathbb{A}}$$

for a transformation $\tau = \sigma \circ \rho$, where ρ is as in (b) and σ is as in (c). Note that such a transformation decreases the valuation of the discriminant, cf. [Liu96, Lemme 9] and [Liu96, Corollaire 2]. By the discussion following Proposition 4.4, the transformation ρ maps $x \in \mathrm{KS}_{\mathbb{A}}$ to $(\pi x_1, x_2, \pi^{-1}x_3, \pi^3 x_4)$.

Suppose $v(2) = 0$ and, without loss of generality, $H = 0$. According to [Liu96, Remarque 2], the normalization can be computed using the transformation σ mapping an affine point (ξ,η) to $\sigma(\xi,\eta) = (\xi, \eta\pi^{-s})$ for some nonnegative integer s. As $v(\tau) = 3 - 2s$, we must have $s \ge 2$, since otherwise τ would increase the valuation of the discriminant. Because $\tau(x) = (\pi x_1, x_2, \pi^{-1}x_3, \pi^{3-2s}x_4)$ for $x \in \mathrm{KS}_{\mathbb{A}}$, we find that $v(\tau(x)) \le v(x) + 1$, implying

$$v(\tau(x)) + v(\tau) - v(x) \le -2s + 4 \le 0\,.$$

The case $v(2) > 0$ is slightly more complicated. Here one computes the normalization by repeatedly applying transformations

$$(\xi,\eta) \mapsto \left(\xi, \frac{\eta + R(\xi,1)}{\pi}\right), \tag{11.1}$$

where $R \in \mathcal{O}[X, Z]$ is a certain cubic form, until the minimum of the valuations of the coefficients of $F + RH - R^2$ is equal to 1. See [Liu96, Remarque 2]. Such a transformation maps Kummer coordinates $x = (x_1, x_2, x_3, x_4)$ to

$$\big(x_1, x_2, x_3, \pi^{-2}x_4 + l_1x_1 + l_2x_2 + l_3x_3\big)$$

and the expressions for the l_i given in Section 4 show that $v(l_i) \geq -2$ for all i. As the determinant of a transformation (11.1) has valuation -2, we need to apply at least two such transformations, because otherwise the valuation of the discriminant would increase. In other words, $\sigma = \sigma_s \circ \cdots \circ \sigma_1$ where $s \geq 2$ and every σ_i is of the form (11.1).

By the properties of the transformations (11.1), it suffices to show the desired inequality for the case $s = 2$, since further applications of transformations σ_i will only make the left hand side of the desired inequality smaller and will not change the right hand side. So suppose that $\sigma = \sigma_2 \circ \sigma_1$; then $\tau = \sigma \circ \rho$ maps $x \in \mathrm{KS}_{\mathbb{A}}$ to

$$\tau(x) = \big(\pi x_1,\, x_2,\, \pi^{-1}x_3,\, \pi^{-1}x_4 + \pi l_1x_1 + \pi l_2x_2 + \pi l_3x_3 + \pi l_1'x_1 + l_2'x_2 + \pi^{-1}l_3'x_3\big)\ ,$$

where the l_i arise from σ_1 and the l_i' arise from σ_2. As $v(\tau) = -1$, it clearly suffices to prove that

$$v(\tau(x)) \leq v(x) + 1. \tag{11.2}$$

But if (11.2) is false, then $v(x) = v(x_4) < \min\{v(x_1), v(x_2) + 1, v(x_3) + 2\}$. In this situation it follows from the lower bounds $v(l_i) \geq -2$ and $v(l_i') \geq -2$ that we get

$$v\big(\pi l_1x_1 + \pi l_2x_2 + \pi l_3x_3 + \pi l_1'x_1 + l_2'x_2 + \pi^{-1}l_3'x_3\big) > v(x_4) - 1\,.$$

This implies (11.2) and therefore finishes the proof of the lemma. □

Theorem 11.3. *Let C be a smooth projective curve of genus 2 defined over a non-archimedean local field k, given by an integral Weierstrass model $\mathcal{C}$. Then we have*

$$\bar{\beta}(\mathcal{C}) \leq \frac{v(\Delta(\mathcal{C}))}{4}\,.$$

Proof. By Lemma 5.4 there is a finite extension k'/k such that the minimal proper regular model of C over k' is semistable and such that all minimal Weierstrass models of C over k' are stably minimal. By Corollary 11.1, the claim therefore holds for any minimal Weierstrass model of C over k'.

It follows from Lemma 11.2 that there is a transformation $\tau\colon \mathcal{C} \to \mathcal{C}'$ defined over k' such that $\mathcal{C}'$ is a minimal (and hence stably minimal) Weierstrass model over k' and such that

$$v(\tau(x)) + v(\tau) \leq v(x) \tag{11.3}$$

for all $x \in \mathrm{KS}_{\mathbb{A}}$.

Then by the above we have

$$\mu(\tau(x)) \leq \frac{v(\Delta(\mathcal{C}'))}{4}\,.$$

Now using Corollary 4.6 and the relation (4.2), we find

$$\begin{aligned}
\mu(x) &= \mu(\tau(x)) - v(x) + v(\tau(x)) - v(\tau) \\
&\leq \frac{v(\Delta(\mathcal{C}'))}{4} - v(x) + v(\tau(x)) - v(\tau) \\
&= \frac{v(\Delta(\mathcal{C}))}{4} - v(x) + v(\tau(x)) + \frac{3}{2}v(\tau) \\
&\leq \frac{v(\Delta(\mathcal{C}))}{4}\,,
\end{aligned}$$

where we have used (11.3) and $v(\tau) \le 0$. □

Remark 11.4. When the residue characteristic is not 2, then we can easily show that $\bar\beta(\mathcal{C})$ is indeed always comparable to $v(\Delta(\mathcal{C}))$. We can assume that $H = 0$ and write $F = cF_0$ with F_0 primitive. We consider the points of order 2 on J. Such a point P is given by a factorization $F_0 = G_1G_2$ with G_1 and G_2 primitive of degrees 2 and 4, respectively. An explicit computation shows that

$$\varepsilon(P) = 4v(c) + 2v(R(P))\,,$$

where $R(P)$ denotes the resultant of G_1 and G_2, and we have $4\mu(P) = \varepsilon(P)$. Since $v(\Delta(\mathcal{C})) = v(\operatorname{disc}(F)) = 10v(c) + v(\operatorname{disc}(F_0))$ and $4v(\operatorname{disc}(F_0))$ is the sum of the valuations of the 15 resultants $R(P)$, we find that

$$\begin{aligned}\bar\beta(\mathcal{C}) &\ge \frac{1}{4}\max_{O\ne P\in J[2]}\bigl(4v(c) + 2v(R(P))\bigr) \ge v(c) + \frac{1}{30}\sum_{O\ne P\in J[2]} v(R(P))\\ &= v(c) + \frac{2}{15}v(\operatorname{disc}(F_0)) \ge \frac{1}{10}v(\Delta(\mathcal{C}))\,.\end{aligned}$$

A similar statement should be true when the residue characteristic is 2.

Recall that we denote $\max\{\varepsilon(P) : P \in J(\bar k)\}$ by $\bar\gamma(\mathcal{C})$.

Corollary 11.5. *Let C be a smooth projective curve of genus 2 defined over a non-archimedean local field k, given by an integral Weierstrass model $\mathcal{C}$. Then we have*

$$\bar\gamma(\mathcal{C}) \le v(\Delta(\mathcal{C}))\,.$$

If $H = 0$ and $\operatorname{char}(k) \ne 2$, then this can be improved to

$$\bar\gamma(\mathcal{C}) \le v(2^{-4}\Delta(\mathcal{C}))\,.$$

Proof. The first inequality follows from 11.3 and from $\varepsilon(P) = 4\mu(P) - \mu(2P)$. The second inequality is Theorem 6.1 of [Sto99]. □

Question 11.6. *If $\mathcal{C}$ is a minimal Weierstrass model, does $\bar\beta(\mathcal{C})$ only depend on the special fiber of $\mathcal{C}^{\min}$?*

Note that the corresponding statement holds for elliptic curves [CPS06]. In our situation, however, there may be several non-isomorphic minimal Weierstrass models, which complicates the picture.

Part III: Efficient Computation of Canonical Heights

In this part we show how to compute the canonical height $\hat{h}(P)$ efficiently for a point P over a number field, global function field or more general field with a system of absolute values as in Section 2. We first explain how to compute the local height correction functions. We use $\mathsf{M}(d)$ to denote the time needed to multiply two d-bit integers.

12. Computing μ at non-archimedean places

In this section, k is a non-archimedean local field again, with valuation ring $\mathcal{O}$, uniformizer π, normalized valuation v and residue class field $\mathfrak{k}$. Let $\mathcal{C}$ be an integral Weierstrass model for a genus 2 curve C over k. We make no assumptions on the reduction type of C. We already discussed a method for the computation of $\mu(P)$ for a given point $P \in J(k)$ in Section 3. In this section, we provide an alternative fast algorithm and show that its running time is $\ll (\log v(\Delta))\,\mathsf{M}\big((\log v(\Delta))v(\Delta)(\log \#\mathfrak{k})\big)$, where $\Delta = \Delta(\mathcal{C})$.

Lemma 12.1. *Assume that M is a positive integer such that $M\mu(P) \in \mathbb{Z}$. Further assume that $\max\{\varepsilon(P) : P \in J(k)\} \le B$. Then*

$$\mu(P) = \frac{1}{M}\Big\lceil M \sum_{n=0}^{\lfloor \log(BM/3)/\log(4)\rfloor} 4^{-n-1}\varepsilon(2^n P)\Big\rceil .$$

Proof. This follows from $M\mu(P) \in \mathbb{Z}$ and from

$$0 \le M \sum_{n \ge m} 4^{-n-1}\varepsilon(2^n P) \le \frac{BM}{3\cdot 4^m} .$$ □

If we know that the reduction is nodal, then we get an upper bound B for $\varepsilon(P)$ and all possible denominators of $\mu(P)$ from the results of Section 9. More generally, if we know the smallest positive period N of the sequence $(\mu(nP))_n$, then we can take $M = N$ (respectively, $M = 2N$) if N is odd (respectively, even) by Corollary 3.11. Also note that we can always take $B = v(\Delta)$ (or even $B = v(2^{-4}\Delta)$ if $\mathrm{char}(k) \ne 2$ and the equation of the curve has $H = 0$), see Corollary 11.5.

If we only know an upper bound for the denominator of $\mu(P)$, then the following alternative approach can be used. This is analogous to [MS15, Lemma 4.2].

Lemma 12.2. *Assume that $M \ge 2$ is an integer such that $M'\mu(P) \in \mathbb{Z}$ for some $0 < M' \le M$. Assume in addition that $\max\{\varepsilon(P) : P \in J(k)\} \le B$, and set*

$$m = \Big\lfloor \frac{\log(BM^2/3)}{\log 4} \Big\rfloor .$$

Then $\mu(P)$ is the unique fraction with denominator $\le M$ in the interval $[\mu_0, \mu_0 + 1/M^2]$, where

$$\mu_0 = \sum_{n=0}^{m} 4^{-n-1}\varepsilon(2^n P) .$$

Proof. Note that

$$\mu_0 \le \mu(P) \le \mu_0 + \sum_{n>m} 4^{-n-1}B < \mu_0 + 1/M^2 .$$

But since $M \ge 2$, the interval $[\mu_0, \mu_0 + 1/M^2]$ contains at most one fraction with denominator bounded by M; by assumption, $\mu(P)$ is such a fraction. □

In order to apply Lemma 12.2, we now find a general upper bound M on the possible denominators of μ. Let $\mathcal{J}$ denote the Néron model of J over $S = \operatorname{Spec}(\mathcal{O})$ and write Φ for the component group of $\mathcal{J}$.

Proposition 12.3. *Let N denote the exponent of $\Phi(\bar{\mathfrak{k}})$ and let $P \in J(k)$. Then we have*

$$\mu(P) \in \frac{1}{2N}\mathbb{Z}\,.$$

If N is odd or if C has a k^{nr}-rational Weierstrass point, then we have

$$\mu(P) \in \frac{1}{N}\mathbb{Z}\,.$$

Proof. Let $i \in \{1,\dots,4\}$ be such that $\kappa_i(P) \neq 0$. Recall from Lemma 8.2 that the function $\hat{\lambda}_i = \hat{\lambda} \circ \frac{\kappa}{\kappa_i}$ is a Néron function with respect to the divisor D_i. As $P \notin \operatorname{supp} D_i$, we find

$$\mu(P) \equiv \hat{\lambda}(x) \equiv \hat{\lambda}_i(P) \pmod{\mathbb{Z}}$$

for any set of Kummer coordinates x for P. It follows from the results of [Nér65] and [Lan83, §11.5] that

$$\hat{\lambda}_i(P) \equiv j(D_i, (P) - (O)) \pmod{\mathbb{Z}}\,,$$

where $j(\ ,\ )$ denotes Néron's bilinear j-pairing, defined in [Nér65, §III.3].

By [Nér65, Prop. III.2], the values of the j-pairing lie in $\frac{1}{2N'}\mathbb{Z}$, where $N' = \#\Phi(\bar{\mathfrak{k}})$ It is easy to see we can replace N' by the exponent N in the proof of [Nér65, Prop. III.2], so the first statement of the proposition follows.

For the second statement, note that the j-pairing takes values in $\frac{1}{N}\mathbb{Z}$ if N is odd, again by [Nér65, Prop. III.2] and its proof. If C has a k^{nr}-rational Weierstrass point P_0, then the divisor D_i is linearly equivalent over k^{nr} to $2\Theta_{P_0}$, where Θ_{P_0} is the theta divisor with respect to P_0. The Néron model does not change under unramified extensions, and $\mu(P) \bmod \mathbb{Z}$ does not depend on the Weierstrass model of C by Corollary 4.6. Hence we can assume that $i = 1$ and $D_1 = 2\Theta_{P_0}$, so the linearity of the j-pairing in the first variable proves the claim. □

Remark 12.4. In the notation of Namikawa-Ueno [NU73], the only reduction types for which Proposition 12.3 does not show that $\mu(P) \in 1/N\mathbb{Z}$ (where N is the exponent of $\Phi(\bar{\mathfrak{k}})$), are $[2III - l]$ and $[2III^* - l]$ for $l \geq 0$; $[2I_n^* - l]$ for $n, l \geq 0$; and $[2I_n - l]$ for $n > 0$ even and $l \geq 0$. We have not found an example where $\mu(P) \notin 1/N\mathbb{Z}$.

We can compute the group $\Phi(\bar{\mathfrak{k}})$ in practice using [BLR90, §9.6]. For this we need to know the intersection matrix of the special fiber of a regular model of C over S. This is implemented in Magma, but can be rather slow. If the residue characteristic is not 2, then we can apply Liu's algorithm [Liu94] to compute the reduction type and read off $\Phi(\bar{\mathfrak{k}})$.

In general, an upper bound for the exponent of $\Phi(\bar{\mathfrak{k}})$ suffices to apply Lemma 12.2. We give a bound which only depends on the valuation of the discriminant $\Delta = \Delta(\mathcal{C})$.

Lemma 12.5. *The exponent of $\Phi(\bar{\mathfrak{k}})$ is bounded from above by*

$$M := \max\left\{2, \left\lfloor \frac{v(\Delta)^2}{3} \right\rfloor\right\}\,.$$

Moreover, the denominator of $\mu(P)$ is bounded from above by M for all $P \in J(k)$.

Proof. This follows from a case-by-case analysis, using the list of groups $\Phi(\bar{\mathfrak{k}})$ from [Liu94, §8] for all reduction types in [NU73], and Proposition 12.3. □

Remark 12.6. By going through all reduction types, it is possible to obtain better upper bounds for the denominator M' of $\mu(P)$ from the Igusa invariants discussed in Section 6. First note that if the special fiber of $\mathcal{C}$ is non-reduced, then we have

(i) $M' \leq 4$ if $v(\Delta) \leq 12$;
(ii) $M' \leq \max\{12, v(\Delta) - 15\}$ otherwise.

Suppose that $\mathcal{C}$ is reduced; then, by Proposition 6.2, we can use the Igusa invariants of the special fiber to distinguish between the multiplicities of its singularities.

(i) If all points on the special fiber of $\mathcal{C}$ have multiplicity at most 2, then we can bound M' using Proposition 6.3 (i–iii) and Propositions 9.1, 9.3, 9.4.
(ii) If there is a point of multiplicity 3 on the special fiber, then we have
- $M' \leq \min\{6, v(\Delta) + 1\}$ if $v(\Delta) \leq 10$;
- $M' \leq 12$, if $v(\Delta) \leq 20$;
- $M' \leq \left\lfloor \frac{(v(\Delta)-12)^2}{4} \right\rfloor$ otherwise.

(iii) If there is a point of multiplicity ≥ 4 on the special fiber, then we have
- $M' \leq 3v(\Delta) - 10$ if $v(\Delta) \leq 10$;
- $M' \leq 4v(\Delta) - 20$ if $v(\Delta) > 10$ and the model is minimal;
- $M' \leq \left\lfloor \frac{(v(\Delta)-10)^2}{3} \right\rfloor$ if the model is not minimal.

The results of this section lead to an efficient algorithm for the computation of $\mu(P)$, which is analogous to Algorithm 4.4 of [MS15]. We assume that the coefficients of F and H and the coordinates of P are given to sufficient v-adic precision (in practice, they will be given exactly as elements of a number field or function field).

1. If $\mathrm{char}(k) \neq 2$ and $H = 0$, set $B := v(2^{-4}\Delta)$. Otherwise, set $B := v(\Delta)$.
2. Set $M := \max\left\{2, \lfloor v(\Delta)^2/3 \rfloor\right\}$.
3. Set $m := \lfloor \log(BM^2/3)/\log(4) \rfloor$.
4. Set $\mu_0 := 0$. Let x be normalized Kummer coordinates for P with $(m+1)B+1$ v-adic digits of precision.
5. For $n := 0$ to m do:
   a. Compute $x' := \delta(x)$ (to $(m+1)B+1$ v-adic digits of precision).
   b. If $v(x') = 0$, then return μ_0.
   c. Set $\mu_0 := \mu_0 + 4^{-n-1}v(x')$.
   d. Set $x := \pi^{-v(x')}x'$
6. Return the unique fraction with denominator at most M in the interval between μ_0 and $\mu_0 + 1/M^2$.

The fraction in the final step can be computed easily, for instance using continued fractions.

For the complexity analysis in the following proposition, we assume that elements of $\mathcal{O}$ are represented as truncated power series in π, whose coefficients are taken from a complete set of representatives for the residue classes. Operations on these coefficients can be performed in time $\ll \mathsf{M}(\log \#\mathfrak{k})$.

Proposition 12.7. *The algorithm above computes $\mu(P)$. Its running time is*

$$\ll (\log v(\Delta))\, \mathsf{M}\big((\log v(\Delta))v(\Delta)(\log \#\mathfrak{k})\big)$$

as $v(\Delta) \to \infty$, with an absolute implied constant.

Proof. The following proof is analogous to the proof of [MS15, Prop. 4.5]. Corollary 11.5 shows that B is a suitable upper bound for ε and Lemma 12.5 shows that M is an upper bound for the denominator of μ. Because $M \geq 2$, the loop in step 5 computes the sum in Lemma 12.2. Note that when $v(x') = 0$ in step 5b, then $\mu(P) = \mu_0$ by Theorem 3.10. At each duplication step, the precision loss is $\varepsilon(2^n P) \leq B$, so that with our choice of starting precision, after the $m+1$ steps in the loop the resulting x still has at least one digit of precision. This proves the correctness of the algorithm.

Clearly the running time of the algorithm is dominated by the running time of the loop in step 5. Step 5a consists of a fixed number of additions and multiplications of elements of $\mathcal{O}$ which are given to a precision of $(m+1)B+1$ digits. Because steps 5b–5d take negligible time compared to step 5a, each pass through the loop takes

$$\ll \mathsf{M}\big(((m+1)B+1)(\log \#\mathfrak{k})\big)$$

operations, leading to a total running time that is

$$\begin{aligned} &\ll (m+1)\,\mathsf{M}\big(((m+1)B+1)(\log \#\mathfrak{k})\big) \\ &\ll m\,\mathsf{M}(mB(\log \#\mathfrak{k})) \\ &\ll (\log v(\Delta))\,\mathsf{M}\big((\log v(\Delta))v(\Delta)(\log \#\mathfrak{k})\big) \end{aligned}$$

as $v(\Delta) \to \infty$. Here we use that $B \ll v(\Delta)$ and $M \ll v(\Delta)^2$, so that $m \ll \log v(\Delta)$. □

Remark 12.8. In step 2, we can use Remark 12.6 to compute a sharper upper bound for the denominator of μ. See also the discussion following Remark 12.4. Of course, if we want to find $\mu(P)$ for several points P, the quantities M, B and m only have to be computed once.

Remark 12.9. We can compute $\mu(P)$ using the algorithm above in more general situations. Suppose that k is any discretely valued field with valuation ring $\mathcal{O}$ and uniformizer π. In that case, the sequence $(\mu(nP))_n$ might not have a finite period, so the method for the computation of $\mu(P)$ discussed in Section 3 might not be applicable. However, Lemma 12.1, Lemma 12.2, Proposition 12.3 and Lemma 12.5 remain valid. If $\mathrm{char}(k) \neq 2$ and if $H = 0$, then we have the upper bound $\varepsilon(P) \leq v(2^{-4}\Delta)$ (cf. Remark 3.2), so the algorithm above can be used and Proposition 12.7 remains valid as well, in the sense that the computation can be done using $\ll \log v(\Delta)$ operations with elements of $\mathcal{O}/\pi^n\mathcal{O}$, where $n \ll v(\Delta)\log v(\Delta)$. In the remaining cases, we can compute an upper bound B on ε as in Remark 3.2, and we can apply the algorithm with this choice of B.

13. Computing μ at archimedean places

In this section, k is an archimedean local field, so $k = \mathbb{R}$ or $k = \mathbb{C}$. We assume that the curve C is given by a Weierstrass equation $\mathcal{C}$ with $H = 0$. In the following, $\log_+ x = \max\{0, \log x\}$.

Let $x \in k^4$ be a set of Kummer coordinates. Recall that

$$\tilde{\varepsilon}(x) = -[k:\mathbb{R}]\,(\log \|\delta(x)\|_\infty - 4\log \|x\|_\infty)$$

and

$$\tilde{\mu}(x) = \sum_{n=0}^{\infty} 4^{-n-1}\tilde{\varepsilon}(\delta^{\circ n}(x))\,.$$

We easily obtain a lower bound for $\tilde{\varepsilon}$ using the standard estimate for $\|\delta(x)\|_\infty$. Since the coefficients of the duplication polynomials δ_j are universal polynomials of degree at most 4 in the coefficients of F, this gives

$$-\tilde{\varepsilon} \ll 1 + \log_+ \|F\|_\infty \,,$$

where $\|F\|_\infty$ is the maximum norm of the coefficient vector of F. We recall that the method described in Section 7 of [Sto99], leading to equation (7.1) there, provides an upper bound $\tilde{\gamma}$ for $\tilde{\varepsilon}$ that can be explicitly computed for any given Weierstrass equation $\mathcal{C}$ of the curve (provided $H = 0$). It is given by

$$\begin{aligned}\tilde{\gamma} &= \log \max_i \left(\sum_{\{S,S'\}} |a_{i,\{S,S'\}}| \sqrt{\sum_{j=1}^{4} |b_{\{S,S'\},j}|} \right)^2 \\ &\le \log 400 + 2 \log \max_{i,\{S,S'\}} |a_{i,\{S,S'\}}| + \log \max_{\{S,S'\},j} |b_{\{S,S'\},j}|\end{aligned}$$

with certain numbers $a_{i,\{S,S'\}}$, $b_{\{S,S'\},j}$, where $i, j \in \{1,2,3,4\}$ and $\{S, S'\}$ runs through the ten partitions of the set of roots of F into two sets of three. Using the formulas in [Sto99, §10] and Mignotte's bound (see for example [vzGG99, Cor. 6.33]), we see that

$$\log \max_{\{S,S'\},j} |b_{\{S,S'\},j}| \ll 1 + \log_+ \|F\|_\infty$$

and

$$\log \max_{i,\{S,S'\}} |a_{i,\{S,S'\}}| \ll 1 + \log_+ \|F\|_\infty + \log_+ \max_{\{S,S'\}} |R(S,S')|^{-1} \,,$$

where $R(S, S')$ is the resultant of the two factors G, G' of F corresponding to the partition of the roots. Using Mignotte's bound again, we find that

$$|R(S,S')|^{-1} = \frac{\sqrt{|\operatorname{disc} G|\,|\operatorname{disc} G'|}}{\sqrt{|\operatorname{disc} F|}} \ll \|F\|_\infty^2 |\Delta(\mathcal{C})|^{-1/2} \,,$$

leading finally to the estimate

$$|\tilde{\varepsilon}| \ll 1 + \log_+ \|F\|_\infty + \log_+ |\Delta(\mathcal{C})|^{-1} =: s(F) \,.$$

If $|\tilde{\varepsilon}(x)| \le \tilde{\eta}$ for all $x \in \mathrm{KS}_{\mathbb{A}}$, then we have

$$\left| \sum_{n \ge N} 4^{-n-1} \tilde{\varepsilon}(\delta^{\circ n}(x)) \right| \le \frac{\tilde{\eta}}{3} 4^{-N} \,,$$

so we need to sum the first

$$N = \left\lceil \frac{d}{2} + \frac{\log(\tilde{\eta}/3)}{\log 4} \right\rceil \ll d + \log s(F)$$

terms to obtain an accuracy of 2^{-d}. Comparing the largest term in any of the δ_j and the lower bound on $\|\delta(x)\|_\infty$, we obtain a bound $\tilde{\theta}$ on the loss of relative precision (in terms of bits) in the computation of $\delta(x)$; we have $\tilde{\theta} \ll s(F)$. To achieve the desired precision at the end, we therefore need to compute with an initial precision of

$$d + 1 + N\tilde{\theta} \ll (d + \log s(F)) s(F)$$

bits. The time needed for each duplication is then

$$\ll \mathsf{M}\big((d + \log s(F)) s(F)\big) \,.$$

A logarithm can be computed to d bits of precision in time $\ll (\log d)\, \mathsf{M}(d)$ by one of several quadratically converging algorithms, see for example [BB98, Chapter 7], so we obtain the following result.

Proposition 13.1. *Given Kummer coordinates x of a point P in $J(k)$ (or $\mathrm{KS}(k)$) to sufficient precision, we can compute $\tilde{\mu}(P)$ to an accuracy of d bits in time*

$$\ll \big(d + \log s(F)\big)(\log d)\,\mathsf{M}\big((d + \log s(F))s(F)\big)\,,$$

where

$$s(F) = 1 + \log_+ \|F\|_\infty + \log_+ |\Delta(\mathcal{C})|^{-1}\,.$$

In the applications k will be the completion of a number field at a real or complex place. If the number field is $\mathbb{Q}$ and the given equation $\mathcal{C}$ of C is integral, then $|\Delta(\mathcal{C})| \geq 1$ and we have $s(F) = 1 + \log \|F\|_\infty = 1 + h(F)$, where $h(F)$ denotes the (logarithmic) height of the coefficient vector of F as a point in affine space. In general, we have the estimate (denoting the value of $s(F)$ for a place v by $s_v(F)$)

$$\begin{aligned}\sum_{v|\infty} s_v(F) &\leq [K:\mathbb{Q}] + \sum_{v|\infty} \log_+ \|F\|_v + \sum_{v|\infty} \log_+ |\Delta(\mathcal{C})|_v^{-1} \\ &\leq [K:\mathbb{Q}] + h(F) + h(\Delta(\mathcal{C})) \ll h(F)\end{aligned}$$

for $h(F)$ large. This implies that we can compute the infinite part of the height correction function in time

$$\ll \big(d + \log h(F)\big)(\log d)\,\mathsf{M}\big((d + \log h(F))h(F)\big)\,,$$

which is polynomial in d and $h(F)$.

14. Computing the canonical height of rational points

The first algorithm for computing the canonical height on a genus 2 Jacobian over $\mathbb{Q}$ was introduced by Flynn and Smart [FS97]. It does not require any integer factorization, but can be impractical even for simple examples, see the discussion in [Sto02, §1]. A more practical algorithm was introduced by the second author in [Sto02]; here the local height correction functions are computed separately, so some integer factorization is required. Uchida [Uch11] later introduced a similar algorithm. De Jong and the first author [dJM14] used division polynomials for a different approach.

Building on the Arakelov-theoretic Hodge index theorem for arithmetic surfaces due to Faltings and Hriljac, Holmes [Hol12] and the first author [Mül14] independently developed algorithms for the computation of canonical heights of points on Jacobians of hyperelliptic curves of arbitrary genus over global fields. While these algorithms can be used to compute canonical heights for genus as large as 10 (see [Mül14, Example 6.2]), they are much slower than the algorithm from [Sto02] when the genus is 2.

In this section we now combine the results of Sections 12 and 13 into an efficient algorithm for computing the canonical height of a point on the Jacobian of a curve of genus 2 over a global field K.

When K is a function field, then there are no archimedean places and factorization is reasonably cheap. So in this case, the best approach seems to be to first find the places v of K such that $\mu_v(P)$ is possibly non-zero (this includes the places at which the given equation of the curve is non-integral) and then compute the corrections $\mu_v(P)$ for each place separately as in the algorithm of Proposition 12.7, if necessary changing first to an integral model and correcting for the transformation afterwards. In fact this approach can be used whenever K is a field with a set of absolute values that satisfy the product formula, because the algorithm before Proposition 12.7 is applicable over any discretely valued field, see Remark 12.9. This includes function fields such as $\mathbb{Q}(t)$ and $\mathbb{C}(t)$.

If K is a number field, then we compute the contribution from the archimedean places as described in Section 13. The finite part of our algorithm is analogous to our quasi-linear algorithm for the computation of the finite part of the canonical height of a point on an elliptic curve in [MS15]; see Proposition 14.3 below. For simplicity, we take K to be $\mathbb{Q}$ in the following. We write ε_p and μ_p for the local height correction functions over $\mathbb{Q}_p$ as given by Definition 3.1 and $\tilde{\mu}_\infty$ for the local height correction function over $\mathbb{R}$ as defined in equation (1.1).

We assume that our curve is given by a model $\mathcal{C}\colon Y^2 = F(X,Z)$ with $F \in \mathbb{Z}[X,Z]$, and we set $\Delta = \Delta(\mathcal{C})$. Our goal is to devise an algorithm for the computation of $\hat{h}(P)$ that runs in time polynomial in $\log \|F\|_\infty$, $h(P)$ and the required precision d (measured in bits after the binary dot). We note that $h(P)$ can be computed in time

$$\ll \log(h(P)+d)\,\mathsf{M}(h(P)+d)\,,$$

since it is just a logarithm. By Proposition 13.1, the height correction function $\tilde{\mu}_\infty(P)$ can be computed in polynomial time. So we only have to find an efficient algorithm for the computation of the 'finite part' $\tilde{\mu}^{\mathrm{f}}(P) := \sum_p \mu_p(P)\log p$ of the height correction.

Fix $P \in J(\mathbb{Q})$. We call a set x of Kummer coordinates for P *primitive* if $x \in \mathbb{Z}^4$ and $\gcd(x) = 1$. We set $g_n = \gcd(\delta(x^{(n)}))$, where $x^{(n)}$ is a primitive set of Kummer coordinates for $2^n P$. Then

$$\tilde{\mu}^{\mathrm{f}}(P) = \sum_{n=0}^{\infty} 4^{-n-1}\log g_n\,.$$

We also know by [Sto99] that g_n divides $D = |\Delta|/2^4 = 2^4|\operatorname{disc}(F)|$, which implies that $\log g_n \le \log D$ for all n. To achieve a precision of 2^{-d}, it is therefore enough to take the sum up to

$$n = m := \left\lfloor \frac{d}{2} + \log\frac{\log D}{3} \right\rfloor \ll d + \log\log D \ll d + \log\log\|F\|_\infty\,.$$

Since at each duplication step, we have to divide by g_n to obtain primitive coordinates again, it suffices to do the computation modulo D^{m+2}. This leads to the following algorithm.

1. Let $D = |\Delta|/16$ and set $m := \lfloor d/2 + \log\log D - \log 3\rfloor$.
2. Let x be primitive Kummer coordinates for P.
3. Set $\mu := 0$.
4. For $n := 0$ to m do:
   a. Compute $x' := \delta(x) \bmod D^{m+2}$.
   b. Set $g_n := \gcd(D, \gcd(x'))$ and $x := x'/g_n$.
   c. Set $\mu := \mu + 4^{-n-1}\log g_n$ (to d bits of precision).
5. Return $\tilde{\mu}^{\mathrm{f}}(P) \approx \mu$.

Proposition 14.1. *This algorithm computes $\tilde{\mu}^{\mathrm{f}}(P)$ to d bits of precision in time*

$$\ll (d+\log\log D)\,\mathsf{M}\big((d+\log\log D)\log D\big) + h(P)\,.$$

Proof. The discussion preceding the algorithm shows that it is correct. The duplication in step 4a and the gcd in step 4b can be computed in time

$$\ll \mathsf{M}((m+2)\log D) \ll \mathsf{M}((d+\log\log D)\log D)\,;$$

the division is even faster, since g_n is small. The computation of the logarithm takes time $\ll \log(d+\log D)\,\mathsf{M}(d+\log D)$; this is dominated by the time for duplication. This gives a time complexity of

$$\ll (d+\log\log D)\,\mathsf{M}\big((d+\log\log D)\log D\big) + h(P)\,,$$

where the last term comes from processing the input x. □

Note that $\log D \ll \log \|F\|_\infty$, so this bound is similar to (and even better by a factor of $\log d$ than) the complexity for computing $\tilde{\mu}_\infty(P)$.

Remark 14.2. We note that an alternative way to proceed is to compute $x' = \delta^{\circ(m+1)}(x)$ mod D^{m+2} (without dividing out gcd's in between) and then use $\mu = 4^{-m-1} \log \gcd(x')$. The advantage of the algorithm above is that we can actually work mod D^{m+2-n}, which makes the computation more efficient. The advantage of the alternative is that it can also be used when working over a number field with non-trivial class group (replacing $\log \gcd(x')$ by the logarithm of the ideal norm of the ideal generated by x'). The resulting complexity is similar, with the implied constant depending on the base field.

We now show that we can in fact do quite a bit better than this, by using the strategy already employed in [MS15]. Note that $\tilde{\mu}^{\mathrm{f}}(P)$ is a rational linear combination of logarithms of positive integers. We can compute such a representation exactly and efficiently by the following algorithm. We again assume that x is a set of primitive Kummer coordinates for P.

1. Set $x' := \delta(x)$, $g_0 := \gcd(x')$ and $x := x'/g_0$.
2. Set $D := \gcd(2^4 \operatorname{disc}(F), g_0^\infty)$ and $B := \lfloor \log D / \log 2 \rfloor$.
3. If $B \le 1$, return 0.
   Otherwise, set $M := \max\{2, \lfloor (B+4)^2/3 \rfloor\}$ and $m := \lfloor \log(B^3 M^2/3)/\log 4 \rfloor$.
4. For $n := 1$ to m do:
   a. Compute $x' := \delta(x)$ mod $D^{m+1} g_0$.
   b. Set $g_n := \gcd(D, \gcd(x'))$ and $x := x'/g_n$.
5. Using the algorithm in [DJB04] (or in [DJB05]), compute a sequence $(q_1, \dots, q_r)$ of pairwise coprime positive integers such that each g_n (for $n = 0, \dots, m$) is a product of powers of the q_i: $g_n = \prod_{i=1}^r q_i^{e_{i,n}}$.
6. For $i := 1$ to r do:
   a. Compute $a := \sum_{n=0}^m 4^{-n-1} e_{i,n}$.
   b. Let μ_i be the simplest fraction between a and $a + 1/(B^2 M^2)$.
7. Return $\sum_{i=1}^r \mu_i \log q_i$ (a formal linear combination of logarithms).

Proposition 14.3. *The preceding algorithm computes $\tilde{\mu}^{\mathrm{f}}(P)$ in time*

$$\ll (\log\log D)\, \mathsf{M}((\log\log D)(\log D)) + \mathsf{M}(h(P))\,.$$

Note that $D \le |\Delta|/16$ and $\log D \ll \log \|F\|_\infty$.

Proof. If $B \le 1$ in step 3, then we either have $g_0 = 1$ and $\tilde{\mu}^{\mathrm{f}}(P) = 0$, or we have $D \in \{2, 3\}$. In the latter case, g_0 is a power of $p = 2$ or 3 and $v_p(\Delta) = 1$, which would imply that $\varepsilon_p(P) = 0$ by [Sto02, Prop. 5.2], so $g_0 = 1$, and we get a contradiction.

If a prime p does not divide g_0, then $\varepsilon_p(P) = 0$, implying $\mu_p(P) = 0$. Suppose now that p divides g_0; then we have $v_p(D) \le B$ and $v_p(\Delta) \le B + 4$, so B, M and m are suitable values for Lemma 12.2. We have $v_p(g_n) = \varepsilon_p(2^n P)$ for all $n \le m$, because $p^{(m+1)v_p(D)+1} \mid D^{m+1} g_0$ (compare the proof of Proposition 12.7). All the g_n are power products of the q_i, so there will be exactly one $i = i(p) \in \{1, \dots, r\}$ such that $p \mid q_{i(p)}$. Setting $b_p = v_p(q_{i(p)})$ and $a = \sum_{n=0}^m 4^{-n-1} e_{i(p),n}$, we have

$$\sum_{n=0}^m 4^{-n-1} \varepsilon_p(2^n P) = \sum_{n=0}^m 4^{-n-1} v_p(g_n) = b_p a\,,$$

implying

$$\mu_p(P) = \sum_{n=0}^{\infty} 4^{-n-1}\varepsilon_p(2^n P) = b_p a + \sum_{n=m+1}^{\infty} 4^{-n-1}\varepsilon_p(2^n P)\,.$$

Here the last sum is in $[0, 1/(B^2M^2)]$ by the definition of m (compare the proof of Lemma 12.2). Therefore

$$a \le \mu_p(P)/b_p \le a + 1/(b_p B^2 M^2) \le a + 1/(B^2M^2)\,.$$

Since the denominator of $\mu_p(P)$ is at most M and since we have $b_p \le v_p(D) \le B$, the denominator of $\mu_p(P)/b_p$ is at most BM. Hence $\mu_p(P)/b_p$ is the unique fraction in $[a, a + 1/(B^2M^2)]$ with denominator bounded by BM, so $\mu_p(P)/b_p = \mu_{i(p)}$ by Step 6b. Now

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

so the algorithm is correct.

The complexity analysis is as in the proof of [MS15, Prop. 6.2]. Namely, the computations in step 1 can be done in time $\ll \mathsf{M}(h(P))$. The computations in steps 2 and 3 take negligible time. Each pass through the loop in step 4 takes time $\ll \mathsf{M}\big((m+2)\log D\big)$, so the total time for step 4 is $\ll m\,\mathsf{M}(m(\log D)) \ll (\log\log D)\,\mathsf{M}((\log\log D)(\log D))$, because $m \ll \log\log D$. The coprime factorization algorithm in [DJB04] (or in [DJB05]) computes suitable q_i for a pair (a, b) of positive integers in time $\ll (\log ab)(\log\log ab)^2$. We iterate this algorithm, applying it first to g_0 and g_1, then to each of the resulting q_i and g_2, and so on. There are always $\ll \log D$ terms in the sequence of q_i's and we have $g_n \le D$ for all n. Hence step 5 takes time $\ll \log D(\log\log D)^3$. Because this is dominated by the time for the loop and because the remaining steps take negligible time, the result follows. □

Note that the complexity of the algorithm above is quasi-linear in $\log D$ and $h(P)$. In practice, the efficiency of this approach can be improved somewhat:

- We can split off the contributions of all sufficiently small primes p by choosing a suitable bound T and trial factoring Δ up to T; the corresponding μ_p can then be computed using the algorithm of Proposition 12.7; see also Remark 12.8. In step 3, we can then set $B := \lfloor \log D'/\log T \rfloor$, where D' is the unfactored part of D, and replace $B+4$ by B in the definition of M. If the coefficients of F are sufficiently large, then this trial division can become quite expensive (even for small values of T). So when $h(F)$ is large, it is usually preferable to avoid trial division altogether.
- We can update the q_i after each pass through the loop in step 4 using the new g_n; we can also do the computation in step 4a modulo suitable powers of the q_i instead of modulo $D^{m+1}g_0$. Moreover, it is possible to use separate values of B, M and m for each q_i; these will usually be smaller than the one computed in step 2 and 3. In this way, we can integrate steps 4, 5 and 6 into one loop.

Remark 14.4. Over a more general number field K in place of $\mathbb{Q}$ the algorithm as stated does not quite work, since we cannot always divide out greatest common divisors. In this case we first compute $x^{(1)} = \delta(x)$ and the ideal g_0 generated by D and the entries of $x^{(1)}$. Then we compute $x^{(2)} = \delta(x^{(1)}), \ldots, x^{(m+1)} = \delta(x^{(m)})$ modulo the ideal $D^{m+1}g_0$. Let G_j be the ideal generated by the entries of $x^{(j)}$ and D^{m+1} and set

$$g_1 = g_0^{-4}G_2, \quad g_2 = G_2^{-4}G_3, \quad g_3 = G_3^{-4}G_4, \quad \ldots, \quad g_m = G_m^{-4}G_{m+1}\,.$$

The coprime factorization algorithms in [DJB04] and [DJB05] also work for ideals. In the final result, $\log q_i$ has to be replaced by $\log N(q_i)$, where $N(q_i)$ is the norm of the

ideal q_i. This should result in a complexity similar to that over $\mathbb{Q}$ (with the implied constant depending on K), or at least one that is dominated by the complexity of computing the naive height and the contributions from the archimedean places. Unfortunately, no complexity analysis for standard operations with ideals in number fields seems to be available in the literature; this prevents us from making a precise statement. Alternatively, we can take the approach described in Remark 14.2.

Combining this with the results for archimedean places, we obtain an efficient algorithm for computing the canonical height $\hat{h}(P)$ of a point $P \in J(\mathbb{Q})$. As mentioned above, we expect a similar result to hold for any number field K in place of $\mathbb{Q}$, with the implied constant depending on K.

Theorem 14.5. *Let C be given by the model $Y^2 = F(X,Z)$ with $F \in \mathbb{Z}[X,Z]$ and let $P \in J(\mathbb{Q})$ be given by primitive Kummer coordinates x (i.e., the coordinates are coprime integers). We can compute $\hat{h}(P)$ to d bits of precision in time*

$$\ll \log(d+h(P))\,\mathsf{M}(d+h(P)) + (d+\log\log\|F\|_\infty)(\log d)\,\mathsf{M}\big((d+\log\log\|F\|_\infty)\log\|F\|_\infty\big)\,.$$

Proof. The first term comes from computing $h(P)$. The second term is the complexity bound for $\tilde{\mu}_\infty(P)$ from Proposition 13.1. The complexity of computing $\tilde{\mu}^{\mathrm{f}}(P)$ using the algorithm of Proposition 14.3 is dominated by their sum, since we have $D \le |\Delta|/16$ and $\log D \ll \log\|F\|_\infty$. □

Note that the complexity is quasi-linear in $\log\|F\|_\infty$ and in $h(P)$, and quasi-quadratic in d. The latter is caused by the (only) linear convergence of the computation of $\tilde{\mu}_\infty(P)$. For elliptic curves one can use a quadratically convergent algorithm due to Bost and Mestre [BM93], see also [MS15]; such an algorithm in the genus 2 case would lead to a complexity that is quasi-linear in d as well.

In Section 15 below we illustrate the efficiency of our algorithm by applying it to a family of curves and points with the property that the number g_0 above is large, so that the previously known algorithms have problems factoring it.

15. Examples

We have implemented our algorithm using the computer algebra system Magma [BCP97]. For the factorization into coprimes we have implemented a simple quadratic algorithm due to Buchmann and Lenstra [BL94, Prop. 6.5] instead of the quasi-linear, but more complicated, algorithms of [DJB04] or [DJB05]

Since the estimates for the required precision in the computation of the archimedean contribution as given in Section 13 are too wasteful in practice, we instead compute this contribution repeatedly using a geometrically increasing sequence of digits of precision until the results agree up to the desired number of bits.

We now compare our implementation with Magma's built-in `CanonicalHeight` (version 2.21-2), which is based on [FS97] and the second author's paper [Sto02], for a family of genus 2 curves. In `CanonicalHeight`, the duplication on the Kummer surface is done using arithmetic over $\mathbb{Q}$, making the implementation slow when points with large coordinates show up during the computation. No factorization of the discriminant is required. However, to find a set of primes such that $\mu_p(P) \neq 0$ for every prime p not in the set, `CanonicalHeight` factors the integer $\gcd(\delta(x))$, where x are primitive Kummer coordinates for P.

Example 15.1. For an integer $a \neq 0$, consider the curve C_a of genus 2 defined by the integral Weierstrass model

$$y^2 = x^5 + a^2x + a^2\,.$$

Let J_a denote the Jacobian of C_a. Then the point $P = [((0,a)) - (\infty)] \in J_a(\mathbb{Q})$ is non-torsion. A set of primitive Kummer coordinates is given by $x = (0,1,0,0)$ and we have $\delta(x) = (4a^2, 0, 0, a^4)$. Hence `CanonicalHeight` needs to factor a^2.

Consider $a =$ 58076586049885709421603671222868245057879201906396781960722099044468153398453014079361023706360328 2, with partial factorization $2 \cdot 7 \cdot 643 \cdot 804743 \cdot a'$, where a' has 89 decimal digits, and its smallest prime factor has 34 decimal digits. Our implementation computes $\hat{h}(P)$ in 0.51 seconds, whereas Magma's `CanonicalHeight` needs about 15 minutes.

Next, we look at $a =$ 20040377295605948895028978950785361771970176052862676844566933718565237900274022252385435405754315284683055562000693599990660880918217466228207807628635725503145772718577795819689 20 . This factors as $a = 2^3 \cdot 5 \cdot 17 \cdot a'$, where a' has 178 decimal digits and no prime divisor with less than 50 decimal digits. Here, our implementation took 1.04 seconds to compute $\hat{h}(P)$, whereas `Magma` did not terminate in 8 weeks.

For $a = p \cdot q$, where p (respectively, q) is the smallest prime larger than 10^{200} (respectively, 10^{250}), the canonical height of P was computed in 5.87 seconds using our implementation.

For the computations in these examples, we used a single core Xeon CPU E7-8837 having 2.67GHz. All heights were computed to 30 decimal digits of precision.

We conclude this part with an example over the rational function field $\mathbb{Q}(t)$.

Example 15.2. Consider the curve $C/\mathbb{Q}(t)$ given by the equation

$$\begin{aligned} y^2 = {} & x^6 - 2t(t+1)x^5 + (t+1)(t^3 - 5t^2 + 4t - 2)x^4 + 2t(t+1)^2(3t^2+1)x^3 \\ & - (t+1)(3t^4 - 2t^2 + 4t - 1)x^2 - 4t^2(t+1)^3(t^2 + 2t - 1)x + 4t^4(t+1)^4\,. \end{aligned}$$

It has the points

$$P_1 = (1:1:0), \quad P_2 = \big(0, -2t^2(t+1)^2\big), \quad P_3 = \big(t+1, 2t(t-1)(t+1)\big)$$

(and also points with x-coordinate $t(t+1)$ and a Weierstrass point $(-t-1, 0)$). Let $Q = [(P_1) - 2(P_2) + (P_3)] \in J(\mathbb{Q}(t))$. Its image on the Kummer surface has coordinates

$$(1 : -t+1 : -2t^2(t+1) : 0)\,.$$

Applying the duplication polynomials and looking at the gcd of the result, we see that we have to compute the height correction functions at the places given by $t = 0$, $t = 1$ and $t = -1$. We also have to consider the place at infinity, since our model of C is not integral there. We use the algorithms of Section 12. Consider the place $t = 0$. From the valuations of the Igusa invariants (see Section 6) we can deduce that the reduction type is $[I_{7-3-2}]$, which gives us $M = 41$ for the exponent of the component group and a bound $B = 10$ for ε. We follow Lemma 12.1 and compute

$$\mu_0(Q) = \frac{1}{41}\Big\lceil 41 \sum_{n=0}^{3} 4^{-n-1}\varepsilon_0(2^nQ)\Big\rceil = \frac{1}{41}\Big\lceil 41\Big(\frac{8}{4} + \frac{4}{4^2} + \frac{7}{4^3} + \frac{6}{4^4}\Big)\Big\rceil = \frac{98}{41}\,.$$

At $t = 1$, the model is not stably minimal. We can deduce from the Igusa invariants that there is a stably minimal model over an extension of ramification index 4, which has reduction type $[I_{12-2-2}]$. This shows that the denominator of μ_1 is divisible by $4 \cdot 26 = 104$. With $M = 104$ and $B = 9$ we get $m = 4$ in Lemma 12.1; we obtain

$$\mu_1(Q) = \frac{1}{104}\Big\lceil 104 \sum_{n=0}^{4} 4^{-n-1}\varepsilon_1(2^nQ)\Big\rceil = \frac{1}{104}\Big\lceil 104\Big(\frac{4}{4} + \frac{4}{4^2} + \frac{3}{4^3} + \frac{2}{4^4} + \frac{2}{4^5}\Big)\Big\rceil = \frac{17}{13}\,.$$

At $t=-1$, the situation is similar. There is a stably minimal model over an extension with ramification index 4 again, which has reduction type $[I_{20-0-0}]$. This leads to $M = 4 \cdot 20 = 80$ and $B = 20$, so $m = 4$, and

$$\mu_{-1}(Q) = \frac{1}{80}\Big\lceil 80 \sum_{n=0}^{4} 4^{-n-1} \varepsilon_{-1}(2^n Q)\Big\rceil = \frac{1}{80}\Big\lceil 80\Big(\frac{7}{4} + \frac{10}{4^2} + \frac{8}{4^3} + \frac{10}{4^4} + \frac{8}{4^5}\Big)\Big\rceil = \frac{51}{20}\,.$$

Finally, at the infinite place, there is a stably minimal integral model over an extension with ramification degree 2, which has reduction type $[I_{8-0-0}]$. In a similar way as for $t=-1$ and taking into account a shift of -8 coming from making the model integral, we obtain $\mu_\infty(Q) = 19/4 - 8 = -13/4$. This results in

$$\hat{h}(Q) = h(Q) - \mu_0(Q) - \mu_1(Q) - \mu_{-1}(Q) - \mu_\infty(Q) = 3 - \frac{98}{41} - \frac{17}{13} - \frac{51}{20} + \frac{13}{4} = \frac{11}{5330}\,.$$

To our best knowledge, the point Q is the point of smallest known nonzero canonical height on the Jacobian of a curve of genus 2 over $\mathbb{Q}(t)$. The curve was found by Andreas Kühn (a student of the second author) in the course of a systematic search for curves with many points mapping into a subgroup of rank 1 in the Jacobian.

Part IV: Efficient Search for Points With Bounded Canonical Height

16. Bounding the height difference at archimedean places

We now describe two approaches for getting a better upper bound $\tilde{\beta}$ on $\tilde{\mu}$ than the one coming from the bound on $\tilde{\varepsilon}$ given in [Sto99, Equation (7.1)], when k is an archimedean local field and C/k is a smooth projective curve of genus 2, given by a Weierstrass equation $Y^2 = F(X, Z)$ in $\mathbb{P}_K(1, 3, 1)$.

We write $\|x\|_\infty = \max\{|x_1|, |x_2|, |x_3|, |x_4|\}$ for the maximum norm.

16.1. **Bounding $\tilde{\varepsilon}$ closely.**

For the first approach we assume that $k = \mathbb{R}$. We describe how to approximate $\max\{\tilde{\varepsilon}(P) : P \in J(\mathbb{R})\}$ to any desired accuracy, which gives us an essentially optimal bound $\tilde{\gamma}$. Recall that

$$\tilde{\varepsilon}(P) = -\log \frac{\max\{|\delta_1(x_1, x_2, x_3, x_4)|, \ldots, |\delta_4(x_1, x_2, x_3, x_4)|\}}{\max\{|x_1|, |x_2|, |x_3|, |x_4|\}^4},$$

where $(x_1 : x_2 : x_3 : x_4)$ is the image of $P \in J(\mathbb{R})$ on the Kummer surface. We can normalize the Kummer coordinates in such a way that $\|x\|_\infty = 1$ and one of the coordinates is 1. We then have to minimize $\max\{|\delta_1|, \ldots, |\delta_4|\}$ over four three-dimensional unit cubes, restricted to the points on the Kummer surface that are in the image of $J(\mathbb{R})$. This means that the relevant points satisfy the equation defining the Kummer surface and in addition the value of (at least) one of four further auxiliary polynomials is positive. (In general, the values of these polynomials are squares if the point comes from the Jacobian, and the converse holds for any one of the polynomials when its value is non-zero. One can choose four such polynomials in such a way that they do not vanish simultaneously on the Kummer surface.)

The idea is now to successively subdivide the given cubes. For each small cube, we check if it may contain points in the image of $J(\mathbb{R})$, by evaluating the various polynomials at the center of the cube and bounding the gradient on the cube. If it can be shown that the defining equation cannot vanish on the cube or that one of the auxiliary polynomials takes only negative values on the cube, then the cube can be discarded. Otherwise, we find upper and lower estimates for $\max\{|\delta_1|, \ldots, |\delta_4|\}$ in a similar way. If the lower bound is larger than our current best upper bound for the minimum, the cube can also be discarded. (At the beginning, we have a trivial upper bound of 1 for the minimum, coming from the origin.) Otherwise, we keep it and subdivide it further. We continue until the difference of the upper and lower bounds for $\tilde{\varepsilon}$ on the cube with the smallest lower bound for $\max\{|\delta_1|, \ldots, |\delta_4|\}$ becomes smaller than a specified tolerance. The upper bound for $\tilde{\varepsilon}$ on that cube is then our bound $\tilde{\gamma}$, and we take (as before) $\tilde{\beta} = \tilde{\gamma}/3$.

We have implemented this approach in Magma [BCP97]. After a considerable amount of fine-tuning, our implementation usually takes a few seconds to produce the required bound. In many cases the new bound, which is essentially optimal as a bound on $\tilde{\varepsilon}$, is considerably better than the bound of [Sto99, (7.1)], but there are also cases for which it turns out that the old bound is actually pretty good.

We used the following tricks to get the implementation reasonably fast.

- We keep the polynomials shifted and rescaled so that the cube under consideration is $[-1, 1]^3$.

- The shifting and scaling is done using linear algebra (working with vectors of coefficients and matrices) and not using polynomial arithmetic.
- The coordinates of the centers and vertices of all cubes are dyadic fractions. We scale everything (by $2^4 = 16$ at each subdivision step — note that the polynomials involved are of degree 4) so that we can compute with integers instead.

16.2. **Iterating Stoll's bound.**

We now describe a different approach that also works for complex places. Instead of trying to get an optimal bound on $\tilde{\varepsilon}$, we aim at a bound on $\tilde{\mu}$ by iterating the bound obtained from equation (7.1) in [Sto99]. We recall how this bound was obtained. There is an elementary abelian group scheme G of order 32 that maps onto $J[2]$ and acts on the space of quadratic forms in the coordinates of the $\mathbb{P}^3$ containing the Kummer surface. This representation splits into a direct sum of ten one-dimensional representations that correspond to the ten partitions $\{S, S'\}$ of the set of ramification points of the double cover $C \to \mathbb{P}^1$ into two sets of three. We write $y_{\{S,S'\}}$ for suitably normalized generators of these eigenspaces ([Sto99] gives explicit formulas in the case $H = 0$). We can then express the squares x_i^2 as linear combinations of these quadratic forms:

$$x_i^2 = \sum_{\{S,S'\}} a_{i,\{S,S'\}} y_{\{S,S'\}}(x)$$

for certain complex numbers $a_{i,\{S,S'\}}$ that can be explicitly determined. On the other hand, $y_{\{S,S'\}}^2$ is a quartic form invariant under the action of $J[2]$ (the representation of G on quartic forms descends to a representation of $J[2]$) and is therefore a linear combination of the duplication polynomials δ_j and the quartic defining the Kummer surface. So there are complex numbers $b_{\{S,S'\},j}$ that can also be explicitly determined such that

$$y_{\{S,S'\}}(x)^2 = \sum_{j=1}^{4} b_{\{S,S'\},j} \delta_j(x)$$

if x is a set of Kummer coordinates. Taking absolute values and using the triangle inequality, we obtain

$$|x_i|^4 \le \left(\sum_{\{S,S'\}} |a_{i,\{S,S'\}}||y_{\{S,S'\}}(x)| \right)^2 \le \left(\sum_{\{S,S'\}} |a_{i,\{S,S'\}}| \sqrt{\sum_{j=1}^{4} |b_{\{S,S'\},j}||\delta_j(x)|} \right)^2$$

for all $(x_1 : x_2 : x_3 : x_4) \in \mathrm{KS}(\mathbb{C})$. This gives a bound for $\tilde{\varepsilon}$ in terms of the $a_{i,\{S,S'\}}$ and $b_{\{S,S'\},j}$ as in equation (7.1) of [Sto99].

We refine this as follows. Define a function

$$\varphi \colon \mathbb{R}^4_{\ge 0} \longrightarrow \mathbb{R}^4_{\ge 0}, \quad (d_1, d_2, d_3, d_4) \longmapsto \left(\sqrt{\sum_{\{S,S'\}} |a_{i,\{S,S'\}}| \sqrt{\sum_{j=1}^{4} |b_{\{S,S'\},j}| d_j}} \right)_{1 \le i \le 4}.$$

Lemma 16.1. *Define a sequence $(b_n)_n$ in $\mathbb{R}^4_{\ge 0}$ by*

$$b_0 = (1, 1, 1, 1) \qquad \text{and} \qquad b_{n+1} = \varphi(b_n)\,.$$

Then (b_n) converges to a limit b and we have

$$\tilde{\mu}(P) \le \frac{4^N}{4^N - 1} \log \|b_N\|_\infty$$

for all $N \ge 1$ and all $P \in J(\mathbb{C})$. In particular, $\sup \tilde{\mu}(J(\mathbb{C})) \le \log \|b\|_\infty$.

Proof. By our previous considerations, it is clear that $|\delta_j(x)| \le d_j$ for all j implies $|x_i| \le \varphi_i(d_1, d_2, d_3, d_4)$ for all i. We deduce by induction on N that

$$\log \|x\|_\infty \le \log \|b_N\|_\infty + 4^{-N} \log \|\delta^{\circ N}(x)\|_\infty$$

for all $N \ge 1$. Writing

$$\tilde{\mu}(P) = -\sum_{m=0}^{\infty} 4^{-mN} \left(\log \|\kappa(2^{mN} P)\|_\infty - 4^{-N} \log \|\delta^{\circ N}(\kappa(2^{mN} P))\|_\infty \right),$$

we obtain an upper bound of $\log \|b_N\|_\infty$ for each of the terms in parentheses, which gives the desired bound.

To see that (b_n) converges, we consider $\Phi(x) = \big(\log \varphi_i(\exp(x_1), \dots, \exp(x_4))\big)_{1 \le i \le 4}$. It is easy to see that the partial derivatives $\frac{\partial \Phi_i}{\partial x_j}$ are positive and that for each i, summing them over j gives $\frac{1}{4}$. (This comes from the fact that φ_i is homogeneous of degree $\frac{1}{4}$.) This implies that $\|\Phi(x') - \Phi(x)\|_\infty \le \frac{1}{4} \|x' - x\|_\infty$, so that Φ is contracting with contraction factor $\le \frac{1}{4}$. The Banach Fixed Point Theorem then guarantees the existence of a unique fixed point of Φ, which every iteration sequence converges to. This implies the corresponding statement for φ. □

If we are dealing with a real place, then we may gain a little bit more by making use of the fact that the $\delta_j(x)$ are real, while some of the coefficients $b_{\{S,'S\},j}$ may be genuinely complex. This can lead to a better bound on $|y_{\{S,S'\}}|$.

For example, considering the curve with the record number of known rational points, we get an improvement from 7.726 to 0.973 for the upper bound on $-\tilde{\mu}$ using Lemma 16.1. See Section 19 for more details. In practice it appears that this second approach is at the same time more efficient and leads to better bounds than the approach described in Section 16.1 above.

The approach described here can also be applied in the context of heights on genus 3 hyperelliptic Jacobians, see [Sto14].

17. Optimizing the naive height

We now consider an arbitrary local field k, with absolute value $|\cdot|$. Let C be given by an equation

$$Y^2 = F(X, Z),$$

and let W be the canonical class on C. The first three coordinates of the image of a point $P = [(X_1 : Y_1 : Z_1) + (X_2 : Y_2 : Z_2)] - W \in J$ on the Kummer surface are given by $Z_1 Z_2$, $X_1 Z_2 + Z_1 X_2$, $X_1 X_2$, whereas the fourth coordinate is homogeneous of degree 1 in the coefficients f_j of F (if we consider Y_1 and Y_2 to be of degree $1/2$). This has the effect that the fourth coordinate usually differs by a factor of about $\|F\| := \max\{|f_0|, |f_1|, \dots, |f_6|\}$ from the other three, which gives this last coordinate a much larger (when $\|F\|$ is large; this is usually the case when k is archimedean) or smaller (this may occur when k is non-archimedean) influence on the local contribution to the naive height when $k = K_v$ and K is a global field. This imbalance tends to increase the difference $h_{\mathrm{std}} - \hat{h}$ between naive and canonical height. This observation suggests to modify the naive height in the following way, so as to give all coordinates roughly the same weight. Compare Section 2 for the general set-up. Let x be a set of Kummer coordinates over a global field K and set

$$h'(x) := \sum_{v \in M_K} \log \max\big\{ |x_1|_v, |x_2|_v, |x_3|_v, |x_4|_v / \|F\|_v \big\}.$$

This is a height as in Example 2.3.

If v is a non-archimedean place, then this has the effect of using the Kummer surface associated to a quadratic twist C_0 of C (which has a primitive polynomial on the right hand side of its equation), since we essentially divide the coefficients by their gcd. The two Kummer surfaces are isomorphic, and so we can use the bound $\bar{\beta}(C_0)$ for the height correction function associated to the quadratic twist as a bound for the height correction function μ_v relative to the modified naive height. (We cannot necessarily use $\beta(C_0)$, since the isomorphism between the Jacobians of C and of C_0 is in general only defined over a quadratic extension.) Note that $v(\Delta(C)) = v(\Delta(C_0)) + 10v(F)$, which leads to bounds for β on C that are quite a bit larger than what we can get for C_0. Note also that this allows us to deal with non-integral equations; in this case, we again implicitly scale to make the polynomial on the right integral and primitive.

When $k = K_v \cong \mathbb{Q}_2$ (say) and we can write $F = 4F_1 + H^2$ with binary forms F_1 and H with integral coefficients, then C is isomorphic to the curve C' given by the Weierstrass equation

$$Y^2 + H(X,Z)Y = F_1(X,Z)\,,$$

and we can use the Kummer surface of the latter to define the local contribution to the naive height. The isomorphism between the Kummer surfaces is given by (see [Mül10, p. 53]; note that this is the inverse of the map given there)

$$(x_1 : x_2 : x_3 : x_4) \longmapsto \big(x_1 : x_2 : x_3 : \tfrac{1}{4}x_4 + \tfrac{1}{2}(h_0h_2x_1 + h_0h_3x_2 + h_1h_2x_3)\big)\,,$$

so defining the local component at v of $h'(x)$ to be

$$\log\max\big\{|x_1|_v, |x_2|_v, |x_3|_v, \big|\tfrac{1}{4}x_4 + \tfrac{1}{2}(h_0h_2x_1 + h_0h_3x_2 + h_1h_2x_3)\big|_v\big\}\,,$$

we can replace the bound for μ_v by the bound we get on C'. If we use this at the places above 2 where it applies (instead of, or combined with, the scaling described above), we still obtain a height as in Example 2.3.

If v is an archimedean place, then the approach described in Section 16.2 above can easily be adapted to the modified naive height. We just have to replace $b_{\{S,S'\},4} = 1$ by $\|F\|_v$ and $a_{4,\{S,S'\}}$ by $a_{4,\{S,S'\}}/\|F\|_v^2$. This will usually lead to a *negative* upper bound for $\tilde{\mu}_v$, which is fairly close to $-\log\|F\|_v$, at least when F is reduced in the sense of [SC03] and its roots are not too close together. This is because the scaled $a_{i,\{S,S'\}}$ are now all of size $\approx \|F\|_\infty^{-2}$ and the scaled $b_{\{S,S'\},j}$ are all of size $\approx \|F\|_\infty$, so Φ as in the proof of Lemma 16.1 roughly satisfies $\|\Phi(x)\|_\infty \approx -\frac{3}{4}\log\|F\|_\infty + \frac{1}{4}\|x\|_\infty$, which has $-\log\|F\|_\infty$ as its fixed point.

Note that for a point $(0:0:0:1) \neq P = (x_1 : x_2 : x_3 : x_4) \in \mathrm{KS}(K)$ we have, for all versions h' of the modified height,

$$h_{\mathrm{std}}\big((x_1 : x_2 : x_3)\big) \le h'(P)\,.$$

We will therefore find all points P with $h'(P) \le B$, if we can enumerate all P with $h_{\mathrm{std}}((x_1 : x_2 : x_3)) \le B$. This can be done (over $\mathbb{Q}$) by using the `-a` option of the second author's program `j-points`, which is available at [Stoa]. In this way, enumerating all points as above with B up to roughly $\log 50\,000$ is feasible. See the discussion in Section 18 below.

Note that it is quite possible that we end up with a bound

$$h_{\mathrm{std}}\big((x_1 : x_2 : x_3)\big) \le h'(P) \le \hat{h}(P) + \tilde{\beta} \qquad \text{for all } P \in J(\mathbb{Q}) \setminus \{O\}$$

with $\tilde{\beta} < 0$. In this case $-\tilde{\beta}$ is a lower bound on the canonical height of any nontrivial point in $J(\mathbb{Q})$; in particular, the torsion subgroup of $J(\mathbb{Q})$ must be trivial. To give an

indication of when we can expect $\tilde{\beta}$ to be close to zero or negative, write $|\operatorname{disc}(F)| = DD'$ with D and D' coprime and D' squarefree. Then the contribution of the finite places to $\tilde{\beta}$ can be bounded by $\frac{1}{4}\log D$, and we get $\tilde{\beta} \approx -\log\|F\|_\infty + \frac{1}{4}\log D$. So if $D \ll \|F\|_\infty^4$, we are in good shape. Note that $|\operatorname{disc}(F)| \ll \|F\|_\infty^{10}$, so this means that 60% or more of $\log|\operatorname{disc}(F)|$ comes from primes p dividing the discriminant exactly once. For curves that are not very special this is very likely to be the case.

In Section 19 we show how this approach can be used to get a very small bound for the height difference even for a curve with ten-digit coefficients.

18. Efficient enumeration of points of bounded canonical height

Let $C\colon y^2 = f(x)$ be a curve of genus 2 over $\mathbb{Q}$ with Jacobian J. In this section we describe the algorithm for enumerating all points $P \in J(\mathbb{Q})$ with $\hat{h}(P) \le B$ that follows from the considerations above. We assume that $f \in \mathbb{Z}[x]$ and proceed as follows.

1. Compute the complex roots of f numerically.
2. Compute the coefficients $a_{i,\{S,S'\}}$ and $b_{\{S,S'\},j}$ from the roots and the leading coefficient of f according to the formulas given in [Sto99, Section 10].
3. Multiply all $a_{4,\{S,S'\}}$ by $\|f\|_\infty^{-2}$ and multiply all $b_{\{S,S'\},4}$ by $\|f\|_\infty$.
4. Iterate the function φ from Section 17 (but using the modified coefficients) a number of times, starting at $(1,1,1,1)$, until there is little change; let $\tilde{\beta}_\infty$ be the upper bound for $\tilde{\mu}_\infty$ as in Lemma 16.1.
5. Factor the discriminant of f.
   Let g be the gcd of the coefficients of f and set $f_1(x) = f(x)/g$.
6. For each prime divisor p of $2\operatorname{disc}(f)$, do the following.
   a. Let e_p be the p-adic valuation of g.
   b. If $p = 2$ and $f_1 = h^2 + 4f_2$ for polynomials $f_2, h \in \mathbb{Z}[x]$, set $C_1\colon y^2 + h(x)y = f_2(x)$; otherwise set $C_1\colon y^2 = f_1(x)$. Let J_1 be the Jacobian of C_1.
   c. If e_p is even, let β_p be the bound for μ_p on $J_1(\mathbb{Q}_p)$ as obtained in Part II. Otherwise, let β_p be the bound for μ_p on $J_1(\bar{\mathbb{Q}}_p)$.
7. Set $\tilde{\beta} = \tilde{\beta}_\infty + \sum_p \beta_p \log p$.
8. Use `j-points` with the `-a` option to enumerate all points $O \ne P \in J(\mathbb{Q})$ such that $h_{\mathrm{std}}\big((\kappa_1(P) : \kappa_2(P) : \kappa_3(P))\big) \le B + \tilde{\beta}$.
9. Add O to this set and return it.

It follows from the discussion in the previous sections that the set returned by this algorithm contains all points with canonical height at most B. If necessary, one can compute the actual canonical heights using the algorithm from Part III and discard the points whose height is too large.

The actual enumeration is done by running through all points $(x_1 : x_2 : x_3) \in \mathbb{P}^2$ of (standard) height at most $B + \tilde{\beta}$ and checking whether there are rational numbers x_4 such that $(x_1 : x_2 : x_3 : x_4)$ is on the Kummer surface. For each of these points on the Kummer surface, we then check if it lifts to the Jacobian. Both these conditions are equivalent to some expression in the coordinates (and the coefficients of f) being a square. `j-points` tries to do this efficiently by using information modulo a number of primes to filter out triples that do not lift to rational points on J. Let $N = \lfloor \exp(B + \tilde{\beta}) \rfloor$. Then `j-points` usually takes a couple of seconds when $N = 1000$, a few minutes when $N = 5\,000$ and a few days when $N = 50\,000$. The running time scales with N^3, but

the scaling factor depends on how effective the sieving mod p is. For Jacobians of high rank, the program tends to take longer than for 'random' Jacobians.

Since the running time depends exponentially on $B+\tilde{\beta}$, it is very important to obtain a small bound $\tilde{\beta}$ for the difference between naive and canonical height. The improvement at the infinite place that we can achieve by considering a modified naive height is crucial for making the enumeration feasible also in cases when the defining polynomial has large coefficients. This is demonstrated by the example in Section 19 below.

If the discriminant of f is too large to be factored, then one can use

$$\tilde{\beta} = \tilde{\beta}_\infty + \frac{1}{4}\log|\operatorname{disc}(f_1)|$$

(or use information from small prime divisors as in the algorithm above and $\frac{1}{4}\log D$ for the remaining primes, where D is the unfactored part of the discriminant). But note that it is usually a great advantage to know the bad primes, since we can take $\beta_p = 0$ for primes p such that $v_p(\operatorname{disc}(f)) = 1$. In most cases, this leads to a much smaller bound $\tilde{\beta}$.

One of the most important applications of this enumeration algorithm is its use in saturating a given finite-index subgroup of $J(\mathbb{Q})$, which gives (generators of) the full group $J(\mathbb{Q})$. This is a necessary ingredient for the method for obtaining all integral points on C developed in [BMS+08], for example, and for computing the regulator of $J(\mathbb{Q})$.

There are essentially two ways of performing the saturation. Let $G \subset J(\mathbb{Q})$ denote the known subgroup.

(i) Let ρ be (an upper bound for) the covering radius of the lattice $\Lambda = (G/G_{\mathrm{tors}}, \hat{h})$. Then $J(\mathbb{Q})$ is generated by G together with all points $P \in J(\mathbb{Q})$ that satisfy $\hat{h}(P) \le \rho^2$, see [Sto02, Prop. 7.1]. This approach is feasible when $\tilde{\beta}+\rho^2$ is sufficiently small.

(ii) Let $I = (J(\mathbb{Q}) : G)$ denote the index; we assume that $J(\mathbb{Q})_{\mathrm{tors}} \subset G$. If $m_1, \dots, m_r$ are the successive minima of Λ and there are no points $P \in J(\mathbb{Q})\backslash G$ with $\hat{h}(P) < B$, then

$$I \le \sqrt{\frac{R \cdot \gamma_r^r}{\prod_{j=1}^r \min\{m_j, B\}}};$$

see [FS97, Section 7]. Here γ_r is (an upper bound for) the Hermite constant for lattices of rank r and R is the regulator of G (i.e., the determinant of the Gram matrix of any basis of Λ). This can be used to get a bound on I whenever B is strictly positive, so for the enumeration we only need $\tilde{\beta}$ to be sufficiently small. (If $\tilde{\beta} < 0$, then we can do entirely without enumeration to get an index bound.) In a second step, one then has to check that G is p-saturated in $J(\mathbb{Q})$ (or find the largest group $G \subset G' \subset J(\mathbb{Q})$ with $(G' : G)$ a power of p) for all primes p up to the index bound. This can be done by considering the intersection of the kernels of the maps $J(\mathbb{Q})/pJ(\mathbb{Q}) \to J(\mathbb{F}_q)/pJ(\mathbb{F}_q)$ for a set of good primes q (such that the group on the right is nontrivial). If this intersection is trivial, then G is p-saturated; otherwise it tells us where to look for points that are potentially divisible by p. Since the index bound gets smaller with increasing B (as long as $B < m_r$), it makes sense to pick B in such a way as to balance the time spent in the two steps of this approach.

19. Example

As an example that demonstrates the use of our nearly optimal upper bound for the difference $h - \hat{h}$ between naive and canonical height (which is based on the optimal

bounds for the μ_p obtained in Sections 9, 10 and 11 and the variation of the naive height discussed in Section 17), we consider the curve

$$C\colon y^2 = 82342800x^6 - 470135160x^5 + 52485681x^4 + 2396040466x^3 + 567207969x^2 - 985905640x + 247747600\,.$$

This curve is of interest, since it holds the current record for the largest number of known rational points (which is 642 for this curve), see [Stob]. A 2-descent on its Jacobian J (assuming GRH) as described in [Sto01] and implemented in Magma gives an upper bound of 22 for the rank of $J(\mathbb{Q})$, and the differences of the known rational points generate a group of rank 22. The latter statement can be checked by computing the determinant R of the height pairing matrix of the following 22 points in $J(\mathbb{Q})$, which is fairly fast using the algorithm for computing canonical heights described in Section 14. The points are given in Mumford representation $(a(x), b(x))$, which stands for $[(\theta_1, b(\theta_1)) + (\theta_2, b(\theta_2))] - W$, where θ_1, θ_2 are the two roots of $a(x)$ and W is the canonical class. Not all of these points are differences of rational points, but they are linear combinations of such differences.

$$\begin{array}{rr}
(x^2 + x, 18868x + 15740), & (x^2 - \tfrac{1}{3}x, \tfrac{216800}{3}x - 15740),\\
(x^2 + \tfrac{2}{3}x - \tfrac{1}{3}, \tfrac{11747}{3}x + \tfrac{21131}{3}), & (x^2 + 5x + 4, 276256x + 273128),\\
(x^2 + \tfrac{4}{3}x - \tfrac{5}{9}, 16315x + \tfrac{26195}{9}), & (x^2 + \tfrac{53}{12}x + \tfrac{5}{3}, \tfrac{1433669}{6}x + \tfrac{371650}{3}),\\
(x^2 - 3x - 4, 34104x + 30976), & (x^2 - 4x - 5, 65987x + 69115),\\
(x^2 + \tfrac{8}{5}x + \tfrac{3}{5}, 67671x + 64543), & (x^2 - 5x - 6, \tfrac{883626}{7}x + \tfrac{905522}{7}),\\
(x^2 - \tfrac{3}{4}x - \tfrac{7}{4}, 31875x + 35003), & (x^2 + \tfrac{5}{7}x - \tfrac{2}{7}, \tfrac{432898}{49}x + \tfrac{279626}{49}),\\
(x^2 + \tfrac{29}{6}x - \tfrac{178}{9}, \tfrac{3014179}{6}x - \tfrac{10824742}{9}), & (x^2 + \tfrac{19}{84}x - \tfrac{65}{84}, \tfrac{4287373}{294}x + \tfrac{5207005}{294}),\\
(x^2 + \tfrac{97}{42}x - \tfrac{37}{42}, \tfrac{23742013}{294}x - \tfrac{5459431}{294}), & (x^2 - \tfrac{5}{11}x, \tfrac{1089388}{121}x - 15740),\\
(x^2 + \tfrac{325}{84}x - \tfrac{11}{21}, \tfrac{30014567}{147}x - \tfrac{2230444}{147}), & (x^2 - \tfrac{683}{140}x - \tfrac{279}{140}, \tfrac{45519013}{490}x + \tfrac{5478709}{490}),\\
(x^2 - \tfrac{91}{769}x - \tfrac{584}{769}, \tfrac{6911886712}{591361}x + \tfrac{16665656516}{591361}), & (x^2 - \tfrac{259}{96}x + \tfrac{163}{72}, \tfrac{52305719}{768}x - \tfrac{13101271}{576}),\\
(x^2 - \tfrac{3073}{2307}x - \tfrac{1252}{769}, \tfrac{54505985456}{1774083}x + \tfrac{25990632928}{591361}), & (x^2 - \tfrac{137}{51}x + \tfrac{40}{51}, \tfrac{47131040}{867}x - \tfrac{8471860}{867})
\end{array}$$

TABLE 3. Generators of the known part of $J(\mathbb{Q})$.

The discriminant of C factors as

$$\Delta = 2^{47}\cdot 3^5\cdot 5^9\cdot 11^2\cdot 13^2\cdot 17^6\cdot 19^4\cdot 23^2\cdot 41^4\cdot 73^3 \cdot 2707\cdot 43579\cdot 108217976921\cdot 87232835173157510 77\,.$$

The results of [Sto99, Sto02] lead to a bound of

$$\frac{1}{3}\big(43\log 2 + 3\log 3 + 9\log 5 + 2\log 11 + 2\log 13 + 6\log 17 + 4\log 19 + 2\log 23 + 4\log 41 + 3\log 73\big) \approx 40.1$$

for the contribution of the finite places to the height difference bound. When trying to get a better bound (for γ_p) by essentially doing an exhaustive search over the p-adic points of the Kummer surface, Magma gets stuck at $p = 2$ for a long while, but eventually finishes with a contribution of 26.434 from the finite places and a total bound of 34.163. This contribution turns out to be $(\gamma_p/3)\log p$ in all cases except for $p = 73$,

where it is $\frac{2}{3}\log 73$ instead of $\frac{1}{3}\log 73$. Our new results from this paper give the following bounds on the local contributions. Φ_p is the component group (ε and μ factor through it in all cases) and 'gain' gives the gain in the bound on the height difference obtained by using the optimal bound on μ versus the bound $\gamma/3$, where γ is the maximum of the values of ε.

p	reduction type	Φ_p	β_p	$\gamma_p/3$	gain
2	$[I_{10-9-8}]$	$\mathbb{Z}/242\mathbb{Z}$	$2+1145/242$	$26/3$	1.341
3	$[I_0-IV-0]$	$\mathbb{Z}/3\mathbb{Z}$	$2/3$	$2/3$	0.000
5	$[I_{4-3-2}]$	$\mathbb{Z}/26\mathbb{Z}$	$22/13$	2	0.495
11	$[I_{2-0-0}]$	$\mathbb{Z}/2\mathbb{Z}$	$1/2$	$2/3$	0.400
13	$[I_{2-0-0}]$	$\mathbb{Z}/2\mathbb{Z}$	$1/2$	$2/3$	0.427
17	$[I_{2-2-2}]$	$\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/6\mathbb{Z}$	1	$4/3$	0.944
19	$[I_{2-1-1}]$	$\mathbb{Z}/5\mathbb{Z}$	$3/5$	$2/3$	0.196
23	$[I_{2-0-0}]$	$\mathbb{Z}/2\mathbb{Z}$	$1/2$	$2/3$	0.523
41	$[I_{2-1-1}]$	$\mathbb{Z}/5\mathbb{Z}$	$3/5$	$2/3$	0.248
73	$[I_{1-1-1}]$	$\mathbb{Z}/3\mathbb{Z}$	$1/3$	$1/3$	0.000

This now gives a bound of ≈ 20.429 for the contribution of the finite places. The optimization of the naive height does not give any improvement at the odd finite places, since the polynomial f defining the curve is primitive. On the other hand, we note that f is congruent to a square mod 4, so we can use the Kummer surface of the curve $y^2+(x^2+x)y=f_1(x)$ (where $f(x)=4f_1(x)+(x^2+x)^2$) for the local height at 2; this reduces the local height difference bound at 2 by $2\log 2$, so the contribution of the finite places to the height difference $h-\hat{h}$ can be bounded by ≈ 19.043.

Now we consider the contribution of the infinite place. The bound obtained from [Sto99, (7.1)] is 7.726. Using Lemma 16.1 with $N=10$ improves this to 0.973; increasing N further gives no significant improvement. However, modifying the local height at the infinite place by scaling the contribution of the fourth coordinate by $\|f\|_\infty^{-1}$ reduces this bound considerably to $\tilde{\mu}_\infty \le -19.25654$ (compare this to $-\log\|f\|_\infty \approx -21.59708$). This finally gives

$$h'(P) \le \hat{h}(P) - 0.21357$$

for our modified (at the places 2 and ∞) naive height h'. Note that this is an instance where we immediately get a strictly positive lower bound on $\hat{h}(P)$ for $P \ne O$ (since then $h'(P)\ge 0$); this shows that $J(\mathbb{Q})$ has trivial torsion subgroup (which is also easy to see by computing the order of $J(\mathbb{F}_p)$ for a few good primes p).

So if we enumerate all points $P\in J(\mathbb{Q})$ with $h'(P)\le \log N$ and do not find points that are not in the known subgroup G, then we obtain a bound for the index $I=(J(\mathbb{Q}):G)$ as follows (see the discussion at the end of Section 18).

$$I \le \sqrt{\frac{R\cdot\gamma_{22}^{22}}{\prod_{j=1}^{22}\min\{m_j, \log N + 0.21357\}}},$$

where R is the regulator of G and $m_1, m_2, \dots, m_{22}$ are the successive minima of the lattice $(G,\hat{h})$, which are

$$8.5276,\ 8.5668,\ 8.5956,\ 8.8594,\ 9.0256,\ 9.0776,\ 9.1426,\ 9.1753,$$
$$9.4456,\ 9.7428,\ 9.7747,\ 9.9047,\ 9.9465,\ 9.9611,\ 9.9704,\ 10.1408,$$
$$10.3472,\ 10.3784,\ 10.5284,\ 10.5356,\ 10.6318,\ 10.9287\,.$$

With $N = 10\,000$ we obtain $I \leq 1516$, with $N = 20\,000$ we get $I \leq 1024$ and with $N \geq 45\,033$ we obtain the best possible bound $I \leq 900$. We checked that there are no unknown points P with $\kappa(P) = (x_1 : x_2 : x_3 : x_4)$ such that $h_{\mathrm{std}}((x_1 : x_2 : x_3)) \leq \log 20\,000$ and verified that the index is not divisible by any prime $p \leq 1024$. Both computations took about two days each on a single core. This implies the following.

Proposition 19.1. *Assume the Generalized Riemann Hypothesis. Let*

$$C\colon y^2 = 82342800x^6 - 470135160x^5 + 52485681x^4 + 2396040466x^3 + 567207969x^2 - 985905640x + 247747600\,.$$

and denote by J the Jacobian of C. Then $J(\mathbb{Q})$ is a free abelian group of rank 22, freely generated by the points listed in Table 3. In particular, $J(\mathbb{Q})$ is generated by the differences of rational points on C.

References

[Art66] Michael Artin, *On isolated rational singularities of surfaces*, Amer. J. Math. **88** (1966), 129–136. ↑10

[Art86] M. Artin, *Lipman's proof of resolution of singularities for surfaces*, Arithmetic geometry (Storrs, Conn., 1984), Springer, New York, 1986, pp. 267–287. ↑7

[DJB04] Daniel J. Bernstein, *Research announcement: Faster factorization into coprimes*, 2004. Preprint. ↑5, 14, 14.4, 15

[DJB05] ———, *Factoring into coprimes in essentially linear time*, Journal of Algorithms **54** (2005), 1–30. ↑5, 14, 14.4, 15

[BG06] Enrico Bombieri and Walter Gubler, *Heights in Diophantine geometry*, New Mathematical Monographs, vol. 4, Cambridge University Press, Cambridge, 2006. ↑4.5

[BB98] Jonathan M. Borwein and Peter B. Borwein, *Pi and the AGM*, Canadian Mathematical Society Series of Monographs and Advanced Texts, 4, John Wiley & Sons, Inc., New York, 1998. A study in analytic number theory and computational complexity; Reprint of the 1987 original; A Wiley-Interscience Publication. ↑13

[BLR90] Siegfried Bosch, Werner Lütkebohmert, and Michel Raynaud, *Néron models*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 21, Springer-Verlag, Berlin, 1990. ↑7, 7, 7, 7, 7, 9, 9, 9, 12

[BCP97] Wieb Bosma, John Cannon, and Catherine Playoust, *The Magma algebra system. I. The user language*, J. Symbolic Comput. **24** (1997), no. 3-4, 235–265. Computational algebra and number theory (London, 1993). ↑15, 16.1

[BM93] Jean-Benoît Bost and Jean-François Mestre, *Calcul de la hauteur archimédienne des points d'une courbe elliptique par un algorithme quadratiquement convergent et application au calcul de la capacité de l'union de deux intervalles*, 1993. Unpublished Manuscript. ↑14

[BS10] Nils Bruin and Michael Stoll, *The Mordell-Weil sieve: proving non-existence of rational points on curves*, LMS J. Comput. Math. **13** (2010), 272–306. ↑7, 7

[BL94] J. A. Buchmann and H. W. Lenstra Jr., *Approximating rings of integers in number fields*, J. Théor. Nombres Bordeaux **6** (1994), no. 2, 221–260 (English, with English and French summaries). ↑15

[BMS+08] Yann Bugeaud, Maurice Mignotte, Samir Siksek, Michael Stoll, and Szabolcs Tengely, *Integral points on hyperelliptic curves*, Algebra Number Theory **2** (2008), no. 8, 859–885. ↑1, 18

[CF96] J. W. S. Cassels and E. V. Flynn, *Prolegomena to a middlebrow arithmetic of curves of genus 2*, London Mathematical Society Lecture Note Series, vol. 230, Cambridge University Press, Cambridge, 1996. ↑1, 1, 3, 7

[Cin11] Zubeyir Cinkir, *Zhang's conjecture and the effective Bogomolov conjecture over function fields*, Invent. Math. **183** (2011), no. 3, 517–562. ↑8.4

[Con05] B. Conrad, *Minimal models for elliptic curves*, 2005, unpublished manuscript. Available at http://math.stanford.edu/∼conrad/papers/minimalmodel.pdf. ↑7.6, 10

[CPS06] J. E. Cremona, M. Prickett, and Samir Siksek, *Height difference bounds for elliptic curves over number fields*, J. Number Theory **116** (2006), no. 1, 42–68. ↑10.13, 11

[DM69] P. Deligne and D. Mumford, *The irreducibility of the space of curves of given genus*, Inst. Hautes Études Sci. Publ. Math. **36** (1969), 75–109. ↑5

[FLS+01] E. Victor Flynn, Franck Leprévost, Edward F. Schaefer, William A. Stein, Michael Stoll, and Joseph L. Wetherell, *Empirical evidence for the Birch and Swinnerton-Dyer conjectures for modular Jacobians of genus 2 curves*, Math. Comp. **70** (2001), no. 236, 1675–1697 (electronic). ↑1

[FS97] E. V. Flynn and N. P. Smart, *Canonical heights on the Jacobians of curves of genus* 2 *and the infinite descent*, Acta Arith. **79** (1997), no. 4, 333–352. ↑1, 2a, 4, 14, 15, ii

[vzGG99] Joachim von zur Gathen and Jürgen Gerhard, *Modern computer algebra*, Cambridge University Press, New York, 1999. ↑13

[Hei04] Niels Heinz, *Admissible metrics for line bundles on curves and abelian varieties over non-Archimedean local fields*, Arch. Math. (Basel) **82** (2004), no. 2, 128–139. ↑8

[HS00] Marc Hindry and Joseph H. Silverman, *Diophantine geometry*, Graduate Texts in Mathematics, vol. 201, Springer-Verlag, New York, 2000. An introduction. ↑1, 2, 4

[Hol12] David Holmes, *Computing Néron-Tate heights of points on hyperelliptic Jacobians*, J. Number Theory **132** (2012), no. 6, 1295–1305. ↑14

[Hol14] ———, *An Arakelov-theoretic approach to naïve heights on hyperelliptic Jacobians*, New York J. Math. **20** (2014), 927–957. MR3272917 ↑2

[Igu60] Jun-ichi Igusa, *Arithmetic variety of moduli for genus two*, Ann. of Math. (2) **72** (1960), 612–649. ↑6

[dJM14] Robin de Jong and J. Steffen Müller, *Canonical heights and division polynomials*, Math. Proc. Cambridge Philos. Soc. **157** (2014), no. 2, 357–373. ↑14

[Lan83] Serge Lang, *Fundamentals of Diophantine geometry*, Springer-Verlag, New York, 1983. ↑8, 12

[Liu93] Qing Liu, *Courbes stables de genre* 2 *et leur schéma de modules*, Math. Ann. **295** (1993), no. 2, 201–222 (French). ↑6, 6.1, 6, 6

[Liu94] ———, *Modèles minimaux des courbes de genre deux*, J. Reine Angew. Math. **453** (1994), 137–164 (French). ↑10, 12, 12

[Liu96] ———, *Modèles entiers des courbes hyperelliptiques sur un corps de valuation discrète*, Trans. Amer. Math. Soc. **348** (1996), no. 11, 4577–4610 (French, with English summary). ↑4, 5, 5, 5, 5, 11, 11, 11

[Liu02] ———, *Algebraic geometry and arithmetic curves*, Oxford Graduate Texts in Mathematics, vol. 6, Oxford University Press, Oxford, 2002. Translated from the French by Reinie Erné; Oxford Science Publications. ↑7

[Mes91] Jean-François Mestre, *Construction de courbes de genre* 2 *à partir de leurs modules*, Effective methods in algebraic geometry (Castiglioncello, 1990), Progr. Math., vol. 94, Birkhäuser Boston, Boston, MA, 1991, pp. 313–334 (French). ↑6

[Mül10] J. Steffen Müller, *Explicit Kummer surface formulas for arbitrary characteristic*, LMS J. Comput. Math. **13** (2010), 47–64. ↑1, 1, 3, 17

[Mül14] ———, *Computing canonical heights using arithmetic intersection theory*, Math. Comp. **83** (2014), no. 285, 311–336. ↑14

[MS15] J. Steffen Müller and Michael Stoll, *Computing canonical heights on elliptic curves in quasi-linear time*, December 22, 2015. Preprint, arXiv:1509.08748v2 [math.NT]. ↑1, 12, 12, 12, 14, 14, 14, 14

[NU73] Yukihiko Namikawa and Kenji Ueno, *The complete classification of fibres in pencils of curves of genus two*, Manuscripta Math. **9** (1973), 143–186. ↑5, 6.3, 9, 9, 10, 10, 12.4, 12

[Nér65] A. Néron, *Quasi-fonctions et hauteurs sur les variétés abéliennes*, Ann. of Math. (2) **82** (1965), 249–331 (French). ↑1, 12

[Sil88] Joseph H. Silverman, *Computing heights on elliptic curves*, Math. Comp. **51** (1988), no. 183, 339–358. ↑7.6, 9.2

[Sil94] ———, *Advanced topics in the arithmetic of elliptic curves*, Graduate Texts in Mathematics, vol. 151, Springer-Verlag, New York, 1994. ↑9

[Sto99] Michael Stoll, *On the height constant for curves of genus two*, Acta Arith. **90** (1999), no. 2, 183–201. ↑1, 1, 3.2, 11, 13, 14, 16, 16.1, 16.2, 2, 19

[Sto01] ———, *Implementing 2-descent for Jacobians of hyperelliptic curves*, Acta Arith. **98** (2001), no. 3, 245–277. ↑19

[Sto02] ———, *On the height constant for curves of genus two. II*, Acta Arith. **104** (2002), no. 2, 165–182. ↑1, 1, 3, 3, 3, 3, 3, 3, 3, 10, 14, 14, 15, i, 19

[Sto14] ———, *An Explicit Theory of Heights for Hyperelliptic Jacobians of Genus Three*, May 15, 2014. Preprint, http://www.mathe2.uni-bayreuth.de/stoll/schrift.html#AG51. ↑16.2

[Stoa] ———, *`j-points`, a program for searching rational points on genus 2 Jacobians*. http://www.mathe2.uni-bayreuth.de/stoll/programs/index.html. ↑17

[Stob] ———, *A genus 2 curve with at least 642 rational points.* http://www.mathe2.uni-bayreuth.de/stoll/recordcurve.html. ↑1, 19

[SC03] Michael Stoll and John E. Cremona, *On the reduction theory of binary forms*, J. Reine Angew. Math. **565** (2003), 79–99. ↑17

[Uch11] Yukihiro Uchida, *Canonical local heights and multiplication formulas for the Jacobians of curves of genus 2*, Acta Arith. **149** (2011), no. 2, 111–130. ↑8, 14

[Zar95] Yu. G. Zarkhin, *Local heights and Néron pairings*, Trudy Mat. Inst. Steklov. **208** (1995), no. Teor. Chisel, Algebra i Algebr. Geom., 111–127 (Russian); English transl., Proc. Steklov Inst. Math. **208** (1995), 100-114. ↑2, 4.5

[Zha93] Shouwu Zhang, *Admissible pairing on a curve*, Invent. Math. **112** (1993), no. 1, 171–193. ↑8

Institut für Mathematik, Carl von Ossietzky Universität Oldenburg, 26111 Oldenburg, Germany

E-mail address: jan.steffen.mueller@uni-oldenburg.de

Mathematisches Institut, Universität Bayreuth, 95440 Bayreuth, Germany.

E-mail address: Michael.Stoll@uni-bayreuth.de