Heights and quantitative arithmetic on stacky curves

Brett Nasserden and Stanley Yao Xiao

Abstract. In this paper, we investigate the theory of heights in a family of stacky curves following recent work of Ellenberg, Satriano, and Zureick-Brown. We first give an elementary construction of a height which is seen to be dual to theirs. We count rational points having bounded ESZ-B height on a particular stacky curve, answering a question of Ellenberg, Satriano, and Zureick-Brown. We also show that when the Euler characteristic of stacky curves is non-positive, the ESZ-B height coming from the anti-canonical divisor class fails to have the Northcott property. We prove that a stacky version of a conjecture of Vojta is equivalent to the *abc*-conjecture.

1 Introduction

Two of the outstanding conjectures in number theory are the so-called Batyrev–Manin conjecture [9] for the density of rational points on open subschemes of Fano varieties with respect to a Weil height, and Malle's conjecture [14] on the number of number fields having bounded discriminant, fixed degree, and fixed Galois group. Both conjectures assert, roughly, that the number of objects to be counted with an appropriate height at most X satisfy an asymptotic formula of the form

$$C \cdot X^{\alpha} (\log X)^{\beta},$$

where C, α, β are nonnegative numbers with $C, \alpha > 0$, and that C, α, β can be computed explicitly within their respective geometric and arithmetic frameworks.

In a recent article, J. Ellenberg, M. Satriano, and D. Zureick-Brown extend the theory of heights to algebraic stacks. They formulate bold conjectures that encompasses both the Batyrev–Manin and Malle conjectures as special cases [6, Main Conjecture]. While the Manin and Malle conjectures are well studied, comparatively little is known about the behavior of rational points on algebraic stacks.

In this article, we study heights on stacky curves, analogues of smooth projective algebraic curves defined over $\mathbb{Q}$. The height functions we develop on stacky curves are completely explicit and can be understood in an elementary manner. Further, we will see that natural questions involving stacky curves lead to an equivalent formulation of the *abc*-conjecture.

The Stacky Batyrev–Manin-Malle Conjecture is still open for smooth one-dimensional algebraic stacks. In this article, we focus on stacky curves whose coarse

moduli space is $\mathbb{P}^1$. A key example is $\mathbb{P}^1$ endowed with three stacky points of degree $\frac{1}{2}$ and the height given by the anti-canonical bundle. This algebraic stack may be thought of as the usual projective line, except the points $0, 1, \infty$ have been replaced with "stacky" points that have degree $\frac{1}{2}$ rather the 1. Alternatively, one may consider the M-curve as described by Darmon, here the points $0, 1, \infty$ have multiplicity 2. The key observation is that a point of multiplicity m has degree $\frac{1}{m}$.

In the case of $\mathbb{P}^1$ with three half points, rational points on this stack can be thought of as rational points on $\mathbb{P}^1$. The anti-canonical height function can now be explicitly written after normalization as

$$(1.1) \qquad H(a : b) = \operatorname{sqf}(|a|) \operatorname{sqf}(|b|) \operatorname{sqf}(|a + b|) \max\{|a|, |b|\}.$$

Put $N(T) = \{(a : b) \in \mathbb{P}^1(\mathbb{Q}) : H(a : b) \le T\}$. Then we have the following theorem.

Theorem 1.1 *There are positive numbers c_1, c_2, c_3 such that*

$$c_1 T^{\frac{1}{2}} \log(T)^3 < N(T) < c_2 T^{\frac{1}{2}} \log(T)^3$$

for all $T > c_3$.

The above estimate proves the stacky Batyrev–Manin conjecture for the anti-canonical height on $\mathbb{P}^1$ with three half points.

We note that P. Le Boudec has independently proved this statement in a private communication.

Let $\mathfrak{X}$ be a proper smooth stacky curve defined over $\mathbb{Q}$ that has coarse space $\mathbb{P}^1$. The rational points of this stack are the rational points of $\mathbb{P}^1$. As in the classical case, there is a canonical line bundle $K_{\mathfrak{X}}$ on $\mathfrak{X}$. The theory described in [6] provides a height function for every line bundle on $\mathfrak{X}$. Therefore, as in the case of algebraic curves, we may consider the anti-canonical height $H_{-K_{\mathfrak{X}}}$. This anti-canonical height is analogous to the anti-canonical height on a Fano variety in the usual Batyrev–Manin conjecture. A natural question is when does the anti-canonical height of $\mathfrak{X}$ satisfy the Northcott property. It turns out that the stacky Euler characteristic answers this question. The Euler characteristic of $\mathfrak{X}$ is defined as

$$(1.2) \qquad \chi(\mathfrak{X}) = \deg(-K_{\mathfrak{X}}).$$

If $\mathfrak{X}$ is $\mathbb{P}^1$ with stacky points $p_1, ..., p_n$ with $\deg p_i = \frac{1}{m_i}$, then one has the formula

$$(1.3) \qquad \chi(\mathfrak{X}) = 2 - \sum_{i=1}^{n} \left(1 - \frac{1}{m_i}\right).$$

The ESZ-B height associated with $-K_{\mathfrak{X}}$ is then given by the explicit formula

$$(1.4) \qquad H_{-K_{\mathfrak{X}}}(x, y) = \max\{|x|, |y|\}^{\chi(\mathfrak{X})} \prod_{i=1}^{r} \phi_{m_i}(\ell_i(x, y))^{\frac{1}{m_i}}.$$

The functions $\phi_{m_i}(\ell_i(x, y))$ are generalizations of the functions $\operatorname{sqf}(|x|)$, $\operatorname{sqf}(|y|)$, and $\operatorname{sqf}(|x + y|)$ that appear in equation (1.1) (see Section 3 for the precise definitions of the functions ℓ_i and ϕ_{m_i}). We obtain a similar explicit description for any line bundle on $\mathfrak{X}$ in terms of the functions ϕ_{m_i} and ℓ_i in Theorem 3.16. For the anti-canonical height, we obtain the following.

Theorem 1.2 *Let $\mathfrak{X}$ be a proper smooth stacky curve defined over $\mathbb{Q}$ that has coarse space $\mathbb{P}^1$ or is isomorphic to a smooth projective curve. Then the anti-canonical height $H_{-K_{\mathfrak{X}}}$ has the Northcott property if and only if $\chi(\mathfrak{X}) > 0$.*

The above result tells us that the ESZ-B heights when applied to the tangent bundle recovers the behavior of the Weil heights when applied to the tangent bundle, providing additional evidence that the ESZ-B theory of heights is the correct generalization of the classical theory. In particular, Theorem 1.2 demonstrates that the stacky Batyrev–Manin conjecture of [6] is a direct generalization of the classical Batyrev–Manin conjecture for Fano varieties.

Ellenberg, Satriano, and Zureick-Brown have proposed a generalized *Vojta's conjecture* applicable to the case of algebraic stacks [6, Conjecture 4.23]. In the case of stacky curves, the stacky Vojta conjecture is directly related to when does $H_{-K_{\mathfrak{X}}}$ to have the Northcott property.

In [6, Section 4.7], it is speculated that [6, Conjecture 4.23] for stacky curves should follow from some version of the *abc*-conjecture. In the case of algebraic curves, Vojta's conjecture is known to be equivalent to the *abc*-conjecture. We show that, much like the case of algebraic curves, the stacky analogue of Vojta's conjecture in the curve case is equivalent to the *abc*-conjecture. We formulate this as follows:

Theorem 1.3 *Let $\mathfrak{X}$ be a proper smooth stacky curve defined over $\mathbb{Q}$ that has coarse space $\mathbb{P}^1$ or is isomorphic to a smooth projective curve. Further, suppose that $\mathfrak{X}$ has negative Euler characteristic. Then the following statements are equivalent:*

(1) *The abc-conjecture holds; and*
(2) *For all $\mathfrak{X}$ satisfying the hypotheses of the theorem and for all $\delta > 0$, the function $\mathcal{H}_{-K_{\mathfrak{X}}} \cdot H^{\delta}$ has Northcott's property, where $H([x, y]) = \max\{|x|, |y|\}$ is the usual height function on $\mathbb{P}^1(\mathbb{Q})$.*

Theorem 1.3 shows that Conjecture 4.23 in [6] is equivalent to the *abc*-conjecture, answering a question of Ellenberg, Satriano, and Zureick-Brown. Their conjectures are motivated by the work of P. Vojita, see [21]

In [6], the authors wonder if the stacky Vojta conjecture might be more "in reach" for algebraic stacks obtained by rooting along a divisor D on a scheme X. The proof of Theorem 1.3 shows that if there is some $m \geq 4$ such that item (2) in Theorem 1.3 holds for $\mathfrak{X}_m = \mathfrak{X}(\mathbb{P}^1 : ((0, 1, \infty) : (m, m, m)))$, then a weak variant of the *abc*-conjecture can be derived. Specifically, there exists a positive number $c_m \geq 1$ such that for any co-prime $a, b, c \in \mathbb{Z}$ with $a + b = c$ and $\varepsilon > 0$ that

$$\max\{|a|, |b|, |c|\} \ll_{\varepsilon, m} \mathrm{rad}(abc)^{c_m + \varepsilon}.$$

In particular, any progress on the stacky Vojta conjecture for curves would lead to substantial progress on the *abc*-conjecture.

2 A further elaboration of our ideas

In this section, we motivate and describe our main results in more detail, as well as describe our grounds-up height construction.

2.1 An elementary height machine on stacky curves

We define our algebraic stacks in terms of a base variety along with some extra data which are enough to uniquely construct an algebraic stack. A stacky curve defined over a number field K is determined by the following data: A smooth variety X defined over K, and a finite number of *stacky* points $P_1, \ldots, P_r$ along with integer multiplicities $m_{P_i} = m_i > 1$ attached to each point P_i. We use the notation

$$\mathfrak{X} = (X : (P_1, m_1), \ldots, (P_r, m_r))$$

to denote the stacky curve with multiplicities $m_{P_i} = m_i$ at the points P_i. We will write $\mathfrak{X}(X : (\mathbf{a}, \mathbf{m}))$ as an abbreviation. We identify the rational points of the stack $\mathfrak{X}$ with those of the *coarse space*, which is just the variety X. That is, we require

$$(2.1) \qquad\qquad\qquad \mathfrak{X}(K) = X(K).$$

In [6], some care is taken to work with the locus of stacky points on an algebraic stack. In particular, one must contend with the accumulation of *infinitely new* stacky points. We ignore such difficulties since they are not important in our context.

To obtain an E-S-ZB height on a stacky curve $\mathfrak{X}(X : (\mathbf{a}, \mathbf{m}))$, we must choose a vector bundle on $\mathfrak{X}(X : (\mathbf{a}, \mathbf{m}))$. We will be primarily interested in the stacky curves $\mathfrak{X}(\mathbb{P}^1_{\mathbb{Q}} : (a_1, m_1), \ldots, (a_r, m_r))$ and the vector bundle being a line bundle. Unless otherwise mentioned, stacky curve $\mathfrak{X}$ will now be of this form. Associated with each stacky point, a_i is a line bundle $\mathcal{L}_{a_i}$ and it suffices to consider line bundles of the form

$$\mathcal{L} = L \otimes \prod_{i=1}^{s} \mathcal{L}_{a_i}^{\otimes c_i},$$

where L is a divisor on $\mathbb{P}^1$ and $0 \le c_i \le m_i - 1$. To associate a height on such a divisor, we associate a height to each $\mathcal{L}_{a_i}^{\otimes c_i}$ and extend linearly. Motivated by [6], we develop the following construction. For each stacky point $a_i = [\alpha_i : \beta_i]$ with α_i, β_i coprime integers, we associate to it the linear form $\ell_i(x, y) = \alpha_i y - \beta_i x$. For each m_i, define $\phi_{m_i}(n)$ is defined to be the smallest positive integer such that $n\phi_{m_i}(n)$ is a perfect m_i-th power. The height function associated with $\mathcal{L}_{a_i}^{\otimes c_i}$ is then

$$H_{\mathcal{L}_{a_i}^{\otimes c_i}}(x, y) = \phi_{m_i}\left(\ell_i(x, y)^{c_i}\right)^{\frac{1}{m_i}}.$$

The linear form ℓ_i takes into account the point a_i and ϕ_{m_i} accounts for the multiplicity of a_i, while the power c_i accounts for the multiple of $\mathcal{L}_{a_i}$. The introduction of the functions ϕ_{m_i} is due to [6] and working with these functions is a key feature of stacky curves with coarse space $\mathbb{P}^1$. We define a height function for any divisor $\mathcal{L} = L \otimes \prod_{i=1}^{s} \mathcal{L}_{a_i}^{\otimes c_i}$ on $\mathfrak{X}$ as

$$(2.2) \qquad H_{\mathcal{L}}(x, y) = \max\{|x|, |y|\}^{\deg \mathcal{L}} \cdot \prod_{i=1}^{r} \phi_{m_i}\left(\ell(x, y)^{c_i}\right)^{\frac{1}{m_i}}$$

whenever x, y are coprime integers. The Euler characteristic of the stacky curve is defined to be the degree of the anti-canonical divisor. If we wish to emphasize in our situation that $\chi(\mathfrak{X})$ only depends on the vector of multiplicities $\mathbf{m}$ and that the

anti-canonical height only depends on $(\mathbf{a}, \mathbf{m})$, we write

$$(2.3) \qquad 2 - \sum_{i=1}^{n} \left(1 - \frac{1}{m_i}\right) = \chi(\mathfrak{X}(\mathbb{P}^1_{\mathbb{Q}}, (a_1, m_1), \ldots, (a_r, m_r))) = \delta(\mathbf{m})$$

and

$$(2.4) \qquad \max\{|x|, |y|\}^{\delta(\mathbf{m})} \prod_{i=1}^{r} \phi_{m_i}(\ell_i(x, y))^{\frac{1}{m_i}} = H_{-K_{\mathfrak{X}}}([x : y]) = \mathcal{H}_{(\mathbf{a}, \mathbf{m})}(x, y).$$

2.2 Properties of the anti-canonical E-S-ZB height $H_{-K_{\mathfrak{X}}}$

The Northcott property of the naive height on $\mathbb{P}^1$ implies that the ESZ-B anti-canonical height $H_{-K_{\mathfrak{X}}}$ has the Northcott property whenever $\chi(\mathfrak{X}) = \delta(\mathbf{m}) > 0$. On the other hand, if $\chi(\mathfrak{X}) \leq 0$, then it is not at all obvious whether $H_{-K_{\mathfrak{X}}}$ should have the Northcott property. The following question is fundamental: Let $\mathcal{L}$ be a line bundle on $\mathfrak{X}(X : (\mathbf{a}, \mathbf{m}))$ and let $H_{\mathcal{L}}$ be the associated ESZ-B height. *When does $H_{\mathcal{L}}$ have the Northcott property?* We will tackle this question when $\mathcal{L} = -K_{\mathfrak{X}}$ and $X = \mathbb{P}^1$ leaving the general case for future study.

Theorem 2.1 *Let $\{a_1, \ldots, a_n\} \subset \mathbb{P}^1_{\mathbb{Q}}$ and $\mathbf{m} = (m_1, \ldots, m_n)$ be a vector of multiplicities. Then whenever*

$$\chi(\mathfrak{X}(\mathbb{P}^1 : (\mathbf{a}, \mathbf{m})) \leq 0,$$

the anti-canonical height $H_{-K_{\mathfrak{X}}}$ given by (2.2) *on the curve $\mathfrak{X}(\mathbb{P}^1 : (a_1, m_1), \ldots, (a_n, m_n))$ fails to have the Northcott property.*

If we assume that the ESZ-B theory should behave roughly like its classical counterpart, we can argue the converse: When $\chi(\mathfrak{X}) \leq 0$, one should have that $H_{-K_{\mathfrak{X}}}$ should fail to have the Northcott property. In particular, Theorem 2.1 shows that our arithmetic and geometric intuition prove to be correct when $\mathfrak{X}$ has coarse space $\mathbb{P}^1$. This answers a question posed by Ellenberg.

The proof of Theorem 2.1 uses the following theorem about elliptic curves.

Theorem 2.2 *Let $F \in \mathbb{Z}[x, y]$ be a non-singular binary quartic form. Then there exists square-free $d \in \mathbb{Z}$ such that the curve*

$$dz^2 = F(u, v)$$

has a rational point and such that its Jacobian has positive rank as an elliptic curve defined over $\mathbb{Q}$.

The proof of Theorem 2.2 is provided to us by Shnidman in [15], and we graciously acknowledge his assistance.

Combining these results gives the following uniform statement.

Corollary 2.3 *Let $\mathfrak{X}$ be a smooth proper stacky curve defined over $\mathbb{Q}$ such that $\mathfrak{X}$ has coarse space $\mathbb{P}^1_{\mathbb{Q}}$ or $\mathfrak{X}$ is a projective algebraic curve. Let $H_{\mathfrak{X}}$ be the height associated with the anti-canonical divisor $-K_{\mathfrak{X}}$. Then $\chi(\mathfrak{X}) > 0$ if and only if $H_{\mathfrak{X}}$ has the strong Northcott property.*

Proof   If $\mathfrak{X}$ is an algebraic stack and not an algebraic curve, then it is of the form $\mathfrak{X} = \mathfrak{X}(\mathbb{P}^1_{\mathbb{Q}}; (\mathbf{a}, \mathbf{m}))$ and (2.1) gives the desired result. On the other hand, if $\mathfrak{X}$ is a smooth projective and geometrically integral curve, then $\chi(\mathfrak{X}) \leq 0$ implies that $-H_C$ does not have the Northcott property.   ∎

In cases where we can prove that the Northcott property fails, we expect, according to [6], that there should be a stacky Vojta conjecture. In particular, we would like to know whether the anti-canonical height $\mathcal{H}_{(\mathbf{a}, \mathbf{m})}$ can be modified to recover the Northcott property. Difficulties arise because the ESZ-B height machine is not functorial in the usual sense; in the setting of algebraic varieties, one can work with a linear spaces of divisors and then apply the height machine which by functoriality will respect the linear structure. Such methods are not immediately available to us. Instead, we will apply the height machine, and then apply linear operations. We ask that for $\mathfrak{X} = \mathfrak{X}(\mathbb{P}^1 : (\mathbf{a}, \mathbf{m}))$ with $\chi(\mathfrak{X}) \leq 0$, what can be said about the quantity

$$\inf\{t \in \mathbb{R}_{\geq 0} : H^t_{\mathbb{P}^1} H_{-K_{\mathfrak{X}}} \text{ has the Northcott property}\}.$$

Clearly, if we change the exponent in the classical part of the height so that it is positive, then we will recover the Northcott property. In fact, we expect that something far less drastic suffices.

For a real number δ and the curve $\mathfrak{X}(\mathbb{P}^1 : (\mathbf{a}, \mathbf{m}))$, define the height

$$(2.5) \qquad \mathcal{H}^\delta_{(\mathbf{a}, \mathbf{m})}(x, y) = \prod_{i=1}^n \phi_{m_i}(\ell_i(x, y))^{1/m_i} \max\{|x|, |y|\}^\delta.$$

We then see that $\mathcal{H}_{(\mathbf{a}, \mathbf{m})} = \mathcal{H}^{\chi(\mathfrak{X})}_{(\mathbf{a}, \mathbf{m})}$. Next, put

$$(2.6) \qquad \gamma(\mathfrak{X}) = \inf\{\delta \in \mathbb{R} : \mathcal{H}^\delta_{(\mathbf{a}, \mathbf{m})} \text{ has the Northcott property on } \mathfrak{X}\}.$$

In fact, $\gamma(\mathfrak{X})$ depends only on $\mathbf{m}$, so we may also write it as $\gamma(\mathbf{m})$. We make the following conjecture.

Conjecture 2.4 (Northcott conjecture for stacky curves with coarse space $\mathbb{P}^1$)   *For all* $\mathfrak{X} = \mathfrak{X}(\mathbb{P}^1 : (\mathbf{a}, \mathbf{m}))$, *we have* $\gamma(\mathfrak{X}) = \min\{\chi(\mathfrak{X}), 0\}$.

Conjecture 2.4 is in fact a version of Vojta's conjecture for stacky curves, and agrees with a conjecture of Ellenberg, Satriano, and Zureick-Brown in [6]. Toward this conjecture, we have the following.

Theorem 2.5   *We have* $\gamma(\mathfrak{X}) = 0$ *if* $\chi(\mathfrak{X}) \geq 0$. *Moreover, the height* $\mathcal{H}^0_{(\mathbf{a}, \mathbf{m})}$ *has the Northcott property if and only if* $\chi(\mathfrak{X}) < 0$.

Combined with Theorem 2.1, the conjecture predicts that the set of $\delta \in \mathbb{R}$ such that $\mathcal{H}^\delta_{\mathbf{a}, \mathbf{m}}$ has the Northcott property is an interval of the form $(\chi(\mathfrak{X}), \infty)$ when $\chi(\mathfrak{X}) < 0$ and $(0, \infty)$ when $\chi(\mathfrak{X}) \geq 0$. Therefore, while Theorem 2.1 tells us we cannot count points with $\mathcal{H}_{\mathbf{a}, \mathbf{m}}$, Conjecture 2.4 predicts that we can count points using $\mathcal{H}^{\chi(\mathfrak{X})+\varepsilon}_{\mathbf{a}, \mathbf{m}}$ for any $\varepsilon > 0$.

Next, we prove that Conjecture 2.4 is a consequence of the *abc*-conjecture. However, it seems that we are very far from being able to prove such a result as strong as Conjecture 2.4 unconditionally.

Theorem 2.6 *Suppose that the abc-conjecture holds. Then, for any $\delta > \delta(\mathbf{m})$, the function $\mathcal{H}^{\delta}_{(\mathbf{a},\mathbf{m})}(x, y)$ on $\mathfrak{X}(\mathbb{P}^1 : (\mathbf{a}, \mathbf{m}))$ has Northcott's property.*

In fact, Conjecture 2.4 is *equivalent* to the *abc*-conjecture (see Theorem 1.3). The proof of the converse is quite different and so we give it in a separate subsection.

2.3 Quantitative arithmetic on stacky curves

In the positive Euler characteristic case, we consider a particular family of stacky curves, which includes an important example suggested by J. Ellenberg[1] and show that our theory of heights matches [6] in this instance. Finally, we verify a specific instance of the main conjecture in [6] given by Ellenberg, Satriano, and Zureick-Brown[2] using analytical methods. We remark that P. Le Boudec had obtained the same result as us in independent work (private communication).

We study the expression (2.3) a bit more carefully. It is easy to deduce that $\delta(\mathbf{m}) \geq 0$ if and only if $n \leq 4$, and $\delta(\mathbf{m}) > 0$ only if $n \leq 3$. We will not consider the case $n \leq 2$ in this paper.

If we assume $m_1 \leq m_2 \leq m_3$, then the only cases when we have positive Euler characteristic are when $m_1 = m_2 = 2$, $m_1 = 2, m_2 = m_3 = 3$ or $m_1 = 2, m_2 = 3, m_3 = 4$. In each of these three cases, the Northcott property for $\mathcal{H}_{(\mathbf{a},\mathbf{m})}$ holds trivially.

We now focus on the simplest cases, where $m_1 = m_2 = 2$ and $m_3 = m, m \geq 2$. Using that PGL_2 acts 3-transitively on $\mathbb{P}^1$, we reduce to the case $\{a_1, a_2, a_3\} = \{0, -1, \infty\}$. For $[x, y] \in \mathbb{P}^1$, we may then set

$$x = x_1 x_2^2, \; y = y_1 y_2^2$$

with x_1, y_1 square-free, and

$$x + y = z_1 z_2^2 \cdots z_{m-1}^{m-1} z_m^m,$$

with $z_1, \ldots, z_{m-1}$ square-free. In this notation, the E-S-ZB height is given by

$$(2.7) \qquad \mathcal{H}_{\mathbf{m}}(x, y) = |x_1|^{1/2} |y_1|^{1/2} |z_1^{m-1} \cdots z_{m-1}|^{1/m} \max\{|x_1 x_2^2|, |y_1 y_2^2|\}^{1/m}.$$

We normalize the height so that the exponent of the "classical part" is equal to one, to obtain the normalized height

$$(2.8) \qquad H_m(x, y) = |x_1|^{m/2} |y_1|^{m/2} |z_1^{m-1} \cdots z_{m-1}| \max\{|x_1 x_2^2|, |y_1 y_2^2|\}.$$

We now put

$$N_m(T) = \#\{(x_1, x_2), (y_1, y_2), (z_1, \ldots, z_m) : \gcd(x_1 y_1, x_2 y_2)$$

$$(2.9) \qquad\qquad = 1, x_1, y_1, z_1, \ldots, z_{m-1} \text{ square-free and pairwise co-prime,}$$

$$x_1 x_2^2 + y_1 y_2^2 = z_1 z_2^2 \cdots z_{m-1}^{m-1} z_m^m, |x_1|^{m/2} |y_1|^{m/2} |z_1^{m-1} \cdots z_{m-1}| \max\{|x_1 x_2^2|, |y_1 y_2^2|\} \leq T\}.$$

We prove the following theorem, which gives a crude upper bound for $N_m(T)$.

---

[1]"What's up in arithmetic statistics?" Number Theory Web Seminar, July 23, 2020.
[2]"What's up in arithmetic statistics?" Number Theory Web Seminar, July 23, 2020.

Theorem 2.7 *Let $\mathfrak{X} = \mathfrak{X}(\mathbb{P}^1 : (0, 2), (\infty, 2), (-1, m))$, and let H_m be the height function on $\mathfrak{X}$ defined by (2.8). Then, for any $\varepsilon > 0$, we have*

$$T^{1/m} \ll N_m(T) \ll_\varepsilon T^{2/m+\varepsilon}.$$

When $m = 2$, the upper bound of Theorem 2.7 is essentially the trivial bound, but it is nontrivial as soon as $m > 2$. In general, we expect the exponent in Theorem 2.7 to be equal to the lower bound. Indeed, this can be verified when $m = 2$. Even more, we can give an exact order of magnitude for $N_2(T)$

Theorem 2.8 *There exist positive numbers c_1, c_2, c_3 such that*

$$c_1 T^{1/2} (\log T)^3 < N_2(T) < c_2 T^{1/2} (\log T)^3$$

for all $T > c_3$.

In particular, we confirm the stacky Batyrev–Manin conjecture [6, Main Conjecture] for $\mathfrak{X}(\mathbb{P}^1_\mathbb{Q}, (a, 2), (b, 2), (c, 2))$. For this stacky curve, [6, Main Conjecture] predicts that $N_2(T) = O_\varepsilon\left(T^{1/2+\varepsilon}\right)$.[3] Our theorem gives an exact order of magnitude for $N_2(T)$. We remark, once again, that P. Le Boudec had obtained the same result. Further, our counting arguments are similar to those obtained by Le Boudec in [13] which studies the equation (7.4).

The other cases with positive Euler characteristic do not yield to the simple analytic counting arguments used to prove Theorem 2.8, though in principle counting rational points by height is a well-posed problem. We plan on returning to this issue in the future.

We illustrate how the stacky curve height machine (equation (2.2)) allows one to detect integral points on stacky curves. In this case, the standard height is given by $H_s(a, b) = \max\{|a|, |b|\}$ and the stacky height is given by (2.7). They are equal precisely when

$$|\operatorname{sqf}(a)\operatorname{sqf}(b)\operatorname{sqf}(a + b)| = 1,$$

or in the notation of (7.4), that $|x_1| = |x_2| = |x_3| = 1$. (7.4) then turns into

$$\pm y_1^2 \pm y_2^2 = \pm y_3^2,$$

and up to rearranging we are essentially counting points on the conic

$$(2.10) \qquad\qquad y_1^2 + y_2^2 = y_3^2.$$

Therefore, if we denote by $\mathcal{N}(T)$ the number of integral points (in the sense of Definition 3.19) on $\mathbb{P}^1_{2,2,2}$, then:

Corollary 2.9 *There exist positive numbers c_1, c_2, c_3 such that for all $T > c_3$ we have*

$$c_1 T^{1/2} < \mathcal{N}(T) < c_2 T^{1/2}.$$

The proof is elementary, since the curve can be explicitly parametrized by

$$y_1 = u^2 - v^2, y_2 = 2uv, y_3 = u^2 + v^2.$$

---

[3]"What's up in arithmetic statistics?" Number Theory Web Seminar, July 23, 2020.

The condition $\max\{|y_1|, |y_2|\} \leq T^{1/2}$ is subsumed by $u^2 + v^2 \leq 4T^{1/2}$ say, so the number of possible u, v's is $\asymp T^{1/2}$ as desired.

Theorem 2.8 and Corollary 2.9 imply that asymptotically 0% of the rational points on $\mathfrak{X}(\mathbb{P}^1 : (0,2), , (-1,2), (\infty,2))(\mathbb{Q})$ are integral, in the sense of Darmon (Definition 3.19).

To close off this subsection, we note that in [2], Bhargava and Poonen study situations where the rational and integral points of a stacky curve satisfy the Hasse Principle. Motivated by this work, we prove that the integral points on $\mathfrak{X}(\mathbb{P}^1 : (a_1, 2), \ldots, (a_n, 2))$ satisfy Hasse's principle.

Theorem 2.10 *Let*

$$\mathfrak{X} = \mathfrak{X}(\mathbb{P}^1_{\mathbb{Q}} : ([a_1 : -b_1], 2), ([a_2 : -b_2], 2), ([a_3, -b_3] : 2)).$$

Then $\mathfrak{X}$ has integral points if and only if the ternary quadratic form

$$Q_{\mathbf{a}}(u, v, w) = \det \begin{pmatrix} u^2 & v^2 & w^2 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}$$

defines a conic with a rational point.

Notation

We denote by $d_k(n)$ for the number of ways of writing n as a product of k (not necessarily distinct) positive integers, and write $d(n) = d_2(n)$ for the usual divisor function. We will also use the big-O notation as well as Landau's notation. In particular, we will denote in the subscripts any dependencies; if there are no subscripts, then the implied constants are absolute.

3 (Stacky) Heights on stacky curves

In this section, we give an alternative construction of the height functions constructed in [6] in a special case: We construct the ESZ-B heights associated with line bundles on stacky curves with coarse space $\mathbb{P}^1_{\mathbb{Q}}$. We use [20] as our main reference, though we made substantial use of [16].

Definition 3.1 (Definition 5.2.1 in [20]) A stacky curve $\mathfrak{X}$ over a field k of characteristic 0 is a smooth proper geometrically connected Deligne–Mumford stack over k of dimension 1 that contains an open dense subscheme.

A stacky curve can be thought of a smooth projective curve, along with a finite choice of points with integer multiplicities.

Theorem 3.2 (Classification of nice stacky curves: Lemma 5.3.10 in [20]) *Let $\mathfrak{X}$ be a stacky curve over k. Then the isomorphism class of $\mathfrak{X}$ is determined by the coarse moduli space X of $\mathfrak{X}$ and the orders of the stabilizer groups of points of $\mathfrak{X}$.*

Before continuing, let us fix some notation. We let $\mathfrak{X} = (X : (P_1, m_1), \ldots, (P_r, m_r))$ be the stacky curve with coarse space X and a μ_{m_i} stabilizer at P_i. In light of Theorem 3.2, this determines a unique stacky curve.

3.1 Construction of heights

We will give an alternative construction of heights on a stacky curve associated with line bundles. Our construction only depends on the coarse space, and the multiplicities of points. The ideas exposited in this section can be viewed as a "bottom up" construction, similar to the work of Geraschenko and Satriano in [10]. We then show that our height construction corresponds to the heights associated with line bundles in [6] when the coarse space is $\mathbb{P}^1_{\mathbb{Q}}$. As in the classical setting, we will work with height functions up to some bounded function. Line bundles on a stacky curve can be described as follows.

Lemma 3.3 ([7, Section 1.3]) *Let* $\mathfrak{X} = \mathfrak{X}(X : (P_1, \cdots, P_r), (m_1, \cdots, m_r))$. *Let* $\mathcal{O}_X(P)$ *be the line bundle associated with the divisor* P *on* X. *Then there are line bundles* $\mathcal{L}_{P_i}$ *on* $\mathfrak{X}$ *such that*

$$(3.1) \qquad \mathcal{L}_{P_i}^{\otimes m_i} \cong \pi_{\mathfrak{X}}^* \mathcal{O}_X(P),$$

where $\pi_{\mathfrak{X}} : \mathfrak{X} \to X$ *is the coarse space map. Moreover, we have that any line bundle* $\mathcal{L}$ *on* $\mathfrak{X}$ *can be uniquely written as*

$$(3.2) \qquad \mathcal{L} \cong \pi_{\mathfrak{X}}^* M \otimes \prod_{i=1}^r \mathcal{L}_{P_i}^{\otimes d_i},$$

where $0 \le d_i < m_i$ *and* M *is a line bundle on* X.

We will use the definition of the degree of a line bundle on a stacky curve.

Definition 3.4 Let $\mathfrak{X} = \mathfrak{X}(X : (P_1, \ldots, P_r), (m_1, \ldots, m_r)))$ and $\mathcal{L} = \pi_{\mathfrak{X}}^* D \otimes \prod_{i=1}^r \mathcal{L}_{P_i}^{\otimes d_i}$. Then we define

$$\deg_{\mathfrak{X}} \mathcal{L} = \deg_X D + \sum_{i=1}^r \frac{d_i}{m_i}.$$

In [6], the height is broken down into two parts: a so-called stable part and a local part. We now define the stable part in our setting.

Definition 3.5 Let $\mathfrak{X} = \mathfrak{X}(X : (P_1, \ldots, P_r), (m_1, \ldots, m_r)))$, and let $\mathcal{L}$ be a line bundle on $\mathfrak{X}$ with

$$\mathcal{L} \cong \pi_{\mathfrak{X}}^* M \otimes \prod_{i=1}^r \mathcal{L}_{P_i}^{\otimes d_i},$$

where $0 \le d_i < m_i$ and M is a line bundle on X with $\pi_{\mathfrak{X}}$ being the coarse space map. We define the *stable height* associated with $\mathcal{L}$ as

$$H_{\mathcal{L}}^{\mathrm{st}} = H_M \cdot \prod_{i=1}^r H_{P_i}^{\frac{d_i}{m_i}}.$$

Later in Proposition 4.3, we show that this definition matches the one given in [6]. The stable height should be thought of as the part of the height consisting of classical height functions.

We will use the notions introduced in Section 3 to define our heights. We further choose a finite set of primes S of $\mathcal{O}_K$ containing all the primes of bad reduction and all

infinite places of K. We further choose a smooth and proper model $\underline{X}$ of X over $\mathcal{O}_{K,S}$. Let P, Q be distinct points in $X(K)$ and place v a place in K with $v \notin S$. Take $\mathfrak{p}_v \subset \mathcal{O}_K$ to be the prime ideal associated with v.

Definition 3.6 (Darmon [4])   We define the *intersection multiplicity* of P and Q at v as follows:

$$(P \cdot Q)_v := \max\{m : \text{the images of } P, Q \text{ in } \underline{X}(\mathcal{O}_{K,S}/\mathfrak{p}_v^m) \text{ are equal.}\},$$

where the maximum over the empty set is defined to be 0.

We now package all the intersection multiplicities together while taking into account the arithmetic of the field extension $K \mid \mathbb{Q}$. We shall use the following notation for the remainder of this section.

Notation 3.7   Fix a stacky curve $\mathfrak{X} = (X : (P_1, m_1), \ldots, (P_r, m_r))$ defined over a number field K. Choose a finite set of primes S of $\mathcal{O}_K$ containing all the primes of bad reduction and all infinite places of K. We further choose a smooth and proper model $\underline{X}$ of X over $\mathcal{O}_{K,S}$. We define the following quantities.

(1) Given a prime $\mathfrak{p}_v \subseteq \mathcal{O}_K$, we let $\mathfrak{f}_v = [\mathcal{O}_K/\mathfrak{p}_v : \mathbb{Z}/(\mathfrak{p}_v \cap \mathbb{Z})]$.
(2) Fix $P \in X(K), t \in X(K)\backslash\{P\}$ and $v \notin S$, now put $(t \cdot P)_p = \sum_{v \notin S, v|p} \mathfrak{f}_v \cdot (t \cdot P)_v$.
(3) We set

$$\lambda_{S,\underline{X},v}(P, t) = \lambda_v(P, t) = \mathrm{N}(\mathfrak{p}_v)^{(t \cdot P)_v}$$

and

$$\lambda_{S,\underline{X}}(P, t) = \lambda(P, t) = \prod_{v \notin S} \lambda_v(P, t) = \prod_p p^{(t \cdot P)_p}.$$

The integer $\lambda(P, t)$ is an exponential version of the familiar looking intersection product

$$\sum_{v \notin S} (t \cdot P)_v \log(p) = \sum_p \left(\sum_{v \notin S, v|p} \mathfrak{f}_v \cdot (t \cdot P)_v \log(p) \right) = \sum_p (t \cdot P)_p \log(p).$$

We will also require the following basic functions.

Definition 3.8   For each integer $m \geq 1$, we let $[0], \ldots, [m-1]$ be a set of representatives of the equivalence classes of $\mathbb{Z}/m\mathbb{Z}$. Define

$$(3.3) \qquad\qquad N_{m,\mathrm{can}}([r]) = r$$

for $0 \leq r < m$ and

$$(3.4) \qquad\qquad N_{m,d}([r]) = N_{m,\mathrm{can}}([-rd])$$

for any $d \in \mathbb{Z}$. With this notation, $N_{m,-} = N_{m,1}$.

These functions are used to make the following definition of the height function associated with $\mathcal{L}_{P_i}^{d_i}$.

Definition 3.9   The stacky height function associated with $\mathcal{L}_{P_i}^{d_i}$ is a function

$$H_{\mathcal{L}_{P_i}^{d_i}} : X(K)\backslash P_i \to \mathbb{Z}_{\geq 0}$$

defined by

$$
(3.5) \qquad H_{\mathcal{L}_{P_i}^{d_i}}(t) = \left(\prod_p p^{N_{m_i,d_i}((t \cdot P_i)_p)} \right)^{\frac{1}{m_i}},
$$

where $N_{m_i,d_i} \colon \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ is the function defined by equation (3.4).

Putting this all together, we obtain the following.

Definition 3.10 (Definition of heights) Let $\mathcal{L}$ be the line bundle $\mathcal{L} \cong \pi_{\mathfrak{X}}^* M \otimes \prod_{i=1}^{r} \mathcal{L}_{P_i}^{\otimes d_i}$ on $\mathfrak{X}$. The stacky height associated with $\mathcal{L}$ is defined to be

$$
H_{\mathcal{L}}(t) = H_{\mathcal{L}}^{\mathrm{st}} \cdot \prod_{i=1}^{r} H_{\mathfrak{X}, \mathcal{L}_{P_i}^{\otimes d_i}}.
$$

Unwinding the definitions, we obtain

$$
(3.6) \qquad H_{\mathcal{L}}(t) = H_{\mathcal{L}}^{\mathrm{st}} \cdot \prod_{i=1}^{r} H_{\mathfrak{X}, \mathcal{L}_{P_i}^{\otimes d_i}}
$$

$$
(3.7) \qquad = H_M \cdot \prod_{i=1}^{r} \left(H_{P_i}^{d_i} \cdot \prod_p p^{N_{m_i,d_i}((t \cdot P_i)_p)} \right)^{\frac{1}{m_i}}
$$

$$
(3.8) \qquad = H_M \cdot \prod_{i=1}^{r} \left(H_{P_i}^{d_i} \cdot \prod_p p^{-d_i(t \cdot P_i)_p \mod m_i} \right)^{\frac{1}{m_i}}.
$$

This decomposition allows us to define what we call the classical and stacky part of a height function.

Definition 3.11 We call

$$
H_{\mathcal{L}}^{\mathrm{st}} = H_M \cdot \prod_{i=1}^{r} H_{P_i}^{\frac{d_i}{m_i}}
$$

the *classical part* of the height $H_{\mathcal{L}}$ and

$$
H_{\mathcal{L},\mathrm{stacky}} = \prod_{i=1}^{r} H_{\mathfrak{X}, \mathcal{L}_{P_i}^{\otimes d_i}}
$$

the *stacky* part of the height $H_{\mathcal{L}}$.

We will primarily work explicitly with stacky heights on $\mathbb{P}^1$, the formulas in that case are as follows.

Corollary 3.12 *Use the notation of Notation 3.7. The canonical height function may be computed as*

$$
H_{K_{\mathfrak{X}}} = H_{\mathbb{P}^1}^{\deg K_{\mathfrak{X}}} \prod_{i=1}^{r} \left(\prod_p p^{(t \cdot P_i)_p \mod m_i} \right)^{\frac{1}{m_i}}.
$$

The anti-canonical height function may be computed as

$$H_{-K_{\mathcal{X}}} = H_{\mathbb{P}^1}^{-\deg K_{\mathcal{X}}} \prod_{i=1}^{r} \left(\prod_p p^{-(t \cdot P_i)_p} \quad \mathrm{mod}\ m_i \right)^{\frac{1}{m_i}}.$$

Proof   This follows directly from the definition, Corollary 4.4, and the fact that $K_{\mathcal{X}}$ corresponds to the line bundle

$$\pi_{\mathcal{X}}^* \mathcal{O}_{\mathbb{P}^1}(K_{\mathbb{P}^1}) \otimes \prod_{i=1}^{r} \mathcal{L}_{P_i}^{\otimes m_i - 1}. \qquad \blacksquare$$

We now introduce two multiplicative functions ϕ_m and r_m that depend on an integer $m \geq 1$. The functions ϕ_m and r_m are *dual* to one another in a certain sense. This duality is key to understanding the nonlinear aspects of heights on stacky curves.

Let x be a positive integer with prime factorization $x = \prod_p p^{\mathrm{ord}_p(x)}$. We will work with the following functions.

(1) Using the division algorithm, we define integers $q_{p,m}(x), r_{p,m}(x)$ by the equation $\mathrm{ord}_p(x) = q_{p,m}(x)m + r_{p,m}(x)$ where $0 \leq r_{p,m}(x) < m$.
(2) $q_m(x) = \prod_p p^{q_{p,m}(x)}$ and $r_m(x) = \prod_p p^{r_{m,p}(x)}$.
(3) Set $\phi_m(x)$ to be the least positive integer such that $x\phi_m(x)$ is an mth power.
(4) We define the *m-radical* of x to be the product of all prime divisors of x whose order is not a factor of m. In other words,

$$\mathrm{rad}_m(x) = \prod_{p \text{ s.t. } \mathrm{ord}_p(x) \neq 0 \quad \mathrm{mod}\ m} p.$$

The r_m is related to $N_{m,\mathrm{can}}$ and ϕ_m is related to $N_{m,-}$ because of the following.

Proposition 3.13   *Let $x \in \mathbb{Z}_{\geq 0}$. Then we have*

$$r_m(x) = \prod_p p^{r_{p,m}(x)} = \prod_p p^{N_{m,\mathrm{can}}(\mathrm{ord}_p(x))},$$

$$\phi_m(x) = \prod_{p \text{ s.t. } \mathrm{ord}_p(x) \neq 0 \quad \mathrm{mod}\ m} p^{m - r_{p,m}(x)} = \prod_{p \mid \mathrm{rad}_m(x)} p^{m - r_{p,m}(x)} = \prod_p p^{N_{-,m}(\mathrm{ord}_p(x))}.$$

In particular, we have

$$H_{\mathcal{X}, \mathcal{L}_{P_i}^{-1}}(t) = r_{m_i}(\lambda(P_i, t))^{\frac{1}{m_i}},$$

$$H_{\mathcal{X}, P, \mathcal{L}_{P_i}^{\otimes d_i}}(t) = \phi_m(\lambda(P_i, t)^{d_i})^{\frac{1}{m_i}}.$$

From the above formulas, we obtain the following.

Proposition 3.14   *Fix and integer $m > 1$ and let $x \in \mathbb{Z}$. Then:*

(1) *Both r_m and ϕ_m are multiplicative functions.*
(2) *$\phi_m(x) r_m(x) = \mathrm{rad}_m(x)^m$.*
(3) *$\phi_m(x) = 1 \iff r_m(x) = 1$.*
(4) *If $m = 2$, then $r_m(x) = \phi_m(x)$.*

With these definitions in hand, we relate our height functions on the stacky curve $\mathfrak{X}$ to a function from the functions $\phi_{m_P}(\lambda(P,t))^{\frac{1}{m_P}}$, and the classical Weil heights on the coarse space X.

Corollary 3.15 *Use the notation of Notation 3.7. Consider the line bundle*

$$\mathcal{L} \cong \pi_{\mathfrak{X}}^* M \otimes \prod_{i=1}^{r} \mathcal{L}_{P_i}^{\otimes d_i}$$

with $0 \le d_i \le m_i - 1$. *Then*

$$(3.9) \qquad H_{\mathcal{L}} = H_M \cdot \prod_{i=1}^{r} H_{P_i}^{\frac{d_i}{m_i}} \cdot \prod_{i=1}^{r} \phi_{m_i}(\lambda(P_i,t)^{d_i})^{\frac{1}{m_i}}.$$

In particular, when $X = \mathbb{P}^1$, we have

$$(3.10) \qquad H_{K_{\mathfrak{X}}}(t) = H_{\mathbb{P}^1}^{-\chi(\mathfrak{X})}(t) \cdot \prod_{i=1}^{r} r_{m_i}(\lambda(P_i,t))^{\frac{1}{m_i}},$$

$$(3.11) \qquad H_{-K_{\mathfrak{X}}}(t) = H_{\mathbb{P}^1}^{\chi(\mathfrak{X})}(t) \cdot \prod_{i=1}^{r} \phi_{m_i}(\lambda(P_i,t))^{\frac{1}{m_i}},$$

where $\chi(\mathfrak{X}) = -\deg K_{\mathfrak{X}}$ is the Euler characteristic of $\mathfrak{X}$.

One interesting feature of the heights given by (3.10) is that they differentiate between *rational* and *integral* points on stacky curves. The connection to [6] and our heights is the following, which is proved in 4.1.

Theorem 3.16 *Fix a stacky curve* $\mathfrak{X} = (\mathbb{P}_{\mathbb{Q}}^1 : (P_1,m_1), \ldots, (P_r,m_r))$. *Choose S to be the set of all finite primes of* $\mathbb{Z}$, *and let* $\mathbb{P}_{\mathbb{Z}}^1$ *be the canonical model of* $\mathbb{P}_{\mathbb{Q}}^1$ *over* $\mathbb{Z}$. *Let* $\mathcal{L}$ *be a line bundle on* $\mathfrak{X}$. *Let* $H_{\mathcal{L}}^{\text{ESZ-B}}$ *be the height constructed in [6] associated with* $\mathcal{L}$. *Then there is some constant $C > 0$ with*

$$(3.12) \qquad C^{-1} \cdot H_{\mathcal{L}}^{\text{ESZ-B}} \le H_{\mathcal{L}} \le C \cdot H_{\mathcal{L}}^{\text{ESZ-B}}.$$

That is to say, up to a constant, the stacky heights from Definition 3.10 agree with the ESZ-B heights in [6] when the coarse space is $\mathbb{P}_{\mathbb{Q}}^1$.

We now explain how the functions ϕ_m and r_m can be used to understand the difference between $H_{\mathcal{L}}$ and $H_{\mathcal{L}^{\otimes n}}$.

Proposition 3.17 *Let* $m \in \mathbb{Z}_{\ge 1}$ *and choose an integer* $d \ge 0$.

$$\phi_m(x^{-d \mod m}) = r_m(x^{d \mod m}).$$

Proof Since ϕ_m is multiplicative, it suffices to prove the statement for $x = p^a$ where p is some prime. Note that $\phi_m((p^a)^n = p^{m-na \mod m})$. Therefore, $\phi_m((p^a)^{-d \mod m}) = p^{m+da \mod m}$. On the other hand, $r_m((p^a)^d \mod m = p^{da \mod m} = p^{m+da \mod m}$ as needed.   ∎

The theory of heights is different from the classical theory of heights, as $H_{\mathcal{L}^{-1}} \ne \frac{1}{H_{\mathcal{L}}} + O(1)$ and $H_{\mathcal{L}^{\otimes n}} \ne (H_{\mathcal{L}})^n + O(1)$. The functions ϕ_m and r_m can be used to compute these quantities.

Theorem 3.18 (Duality theorem) *Let $\mathfrak{X} = \mathfrak{X}(X : (P_1, m_1), \ldots, (P_r, m_r))$ be a stacky curve, and let*

$$\mathcal{L} \cong \pi_{\mathfrak{X}}^* M \otimes \prod_{i=1}^r \mathcal{L}_{P_i}^{\otimes d_i}$$

with $0 \leq d_i \leq m_i - 1$. Fix an integer $n \neq 0$ and write $n_i = nd_i \mod m_i$.

(1) *Then we always have*

$$H_{\mathcal{L}^{\otimes n}} = (H_{\mathcal{L}}^{\mathrm{st}})^n \cdot \prod_{i=1}^r \phi_{m_i}(\lambda(P_i, t)^{n_i})^{\frac{1}{m_i}}$$

$$= (H_{\mathcal{L}}^{\mathrm{st}})^n \cdot \prod_{i=1}^r \phi_{m_i}(\lambda(P_i, t)^{nd_i \mod m_i})^{\frac{1}{m_i}}.$$

(2) *If $n > 0$, then*

$$H_{\mathcal{L}^{\otimes -n}} = (H_{\mathcal{L}}^{\mathrm{st}})^{-n} \cdot \prod_{i=1}^r r_{m_i}(\lambda(P_i, t)^{nd_i \mod m_i})^{\frac{1}{m_i}}.$$

In particular,

$$H_{\mathcal{L}^{-1}} = (H_{\mathcal{L}}^{\mathrm{st}})^{-1} \cdot \prod_{i=1}^r r_{m_i}(\lambda(P_i, t)^{d_i})^{\frac{1}{m_i}}.$$

Proof Write $nd_i = m_i q_i + r_i$ with $0 \leq r_i < m_i$. Note that $\mathcal{L}_{P_i}^{\otimes m_i} = \pi_{\mathfrak{X}}^* \mathcal{O}_X(P_i)$. So we have that

$$(3.13) \qquad \mathcal{L}^{\otimes n} = \pi_{\mathfrak{X}}^*(M^{\otimes n}) \cdot \prod_{i=1}^r \pi_{\mathfrak{X}}^*(\mathcal{O}_X(q_i P_i)) \cdot \prod_{i=1}^r \mathcal{L}_{P_i}^{\otimes r_i}.$$

Therefore, we have

$$H_{\mathcal{L}^{\otimes n}}^{\mathrm{st}} = H_m^n \cdot \prod_{i=1}^r H_{P_i}^{q_i} \cdot \prod H_{P_i}^{\frac{r_i}{m_i}} = H_m^n \cdot \prod_{i=1}^r H_{P_i}^{\frac{m_i q_i + r_i}{m_i}}$$

$$= H_m^n \cdot \prod_{i=1}^r H_{P_i}^{\frac{nd_i}{m_i}} = \left(H_M \cdot \prod_{i=1}^r H_{P_i}^{\frac{d_i}{m_i}} \right)^n = (H_{\mathcal{L}}^{\mathrm{st}})^n.$$

Now we have that

$$H_{\mathcal{L}^{\otimes n}} = H_{\mathcal{L}^{\otimes n}}^{\mathrm{st}} \prod_{i=1}^r \phi_{m_i}(\lambda(P_i, t)^{r_i})^{\frac{1}{m_i}}$$

$$= H_{\mathcal{L}^{\otimes n}}^{\mathrm{st}} \prod_{i=1}^r \phi_{m_i}(\lambda(P_i, t)^{nd_i \mod m_i})^{\frac{1}{m_i}}$$

by the definition of the r_i. Now let $n > 0$. By Proposition 3.17, we have $\phi_m(\lambda(P_i, t)^{-dn \mod m}) = r_m(\lambda(P_i, t)^{nd_i \mod m_i})$ giving the desired conclusion of (2). ∎

3.2   Integral points on stacky curves

Here, we show that the height functions defined in Definition 3.10) can be used to obtain information about integral points on $\mathfrak{X}$. Fix a stacky curve $\mathfrak{X} = \mathfrak{X}(X : (\mathbf{p}, \mathbf{m}))$.

Let $H_{\mathfrak{X}, P_i}$ be the height function associated with the line bundle $\mathcal{L}_{P_i}$ and put $\mathcal{D}_{\mathfrak{X}} = \prod_{i=1}^{r} \mathcal{L}_{P_i}$. Recall that we have $H_{\mathfrak{X}, P_i}(t) = \phi_{m_i}(\lambda(P_i, t))^{1/m_i}$. We will work with the height

$$H_{\mathcal{D}_{\mathfrak{X}}}(t) = \prod_{i=1}^{r} \phi_{m_i}(\lambda(P_i, t))^{1/m_i}.$$

This is the *stacky part* of the anti-canonical height. We will prove that the set of integral points is contained in the set of points where $H_{\mathcal{D}_{\mathfrak{X}}}(t) = 1$. When we take $K = \mathbb{Q}$, we see that this condition is sufficient. In other words, the S-integral points are those where the stacky part of the height is trivial. Following Darmon [4], we have the following notion of integral points on an stacky curve.

Definition 3.19 (Darmon)   Let $\mathfrak{X} = (X : (P_1, m_1), \ldots, (P_r, m_r))$ be a stacky curve over a number field K, S a finite set of places of K containing all primes of bad reduction. Let $\underline{X}$ be a smooth proper model for X over $\mathcal{O}_{K,S}$. The $(\underline{X}, S)$-integral points of $\mathfrak{X}$ (usually abbreviated to S-integral points of $\mathfrak{X}$) are the points $t \in X(K)$ such that

$$(3.14) \qquad\qquad\qquad (t \cdot P)_v \equiv 0 \mod m_P$$

for all $P \in X(K)$ and $v \notin S$.

We shall prove the following theorem.

Theorem 3.20   Let $\mathfrak{X} = (X : (P_1, m_1), \ldots, (P_r, m_r))$ be a stacky curve over K satisfying our assumptions and choose S and a model $\underline{X}$ as we have specified. Then we have the following conclusions.

(1)
$$\mathfrak{X}(\mathcal{O}_{K,S,\underline{X}}) \subseteq \bigcap_{i=1}^{r} \mathfrak{X}(P_i; K),$$

where $\mathfrak{X}(P_i; K) = \{t \in X(K) : H_{\mathfrak{X}, P_i}(t) = 1\}$.

(2) If $K = \mathbb{Q}$, then

$$\mathfrak{X}(\mathcal{O}_{K,S,\underline{X}}) = \bigcap_{i=1}^{r} \mathfrak{X}(P_i; K).$$

In particular, the set of S-integral points of $\mathfrak{X}$ is precisely the set of points where $H_{\mathfrak{X}, P_i}(t) = 1$ for all $i = 1, \ldots, r$.

Fix a prime $v \notin S$ and write $(t \cdot P)_v = m_P^{e_{v,P}(t)} \cdot q_{v,P}(t)$ where $e_{v,P}(t) \geq 0$ and $q_{v,P}(t) \geq 0$ is not divisible by m_P. In other words, $q_{v,P}(t)$ is the m_P-free part of $(t \cdot P)_v$. Set $\mathrm{N}(\mathfrak{p}_v) = p_v^{f(v)}$. Then

$$(3.15) \qquad\qquad \lambda_v(P, t) = p_v^{f(v)(t \cdot P)_v} = p_v^{m_P^{e_{v,P}(t)} \cdot q_{v,P}(t) \cdot f(v)}$$

and

$$(3.16) \qquad \lambda(P_i, t) = \prod_{v \notin S} p_v^{m_{P_i}^{e_{v,P_i}(t)} \cdot q_{v,P_i}(t) \cdot f(v)}.$$

Using the functions $\lambda(P, t)$, we can find subsets of the rational points that contain all integral points.

Proposition 3.21 *Suppose that $m_P > 1$. Define $\mathfrak{X}(P; K) = \{t \in X(K) : H_{\mathfrak{X},P}(t) = 1\}$. Then*

$$\mathfrak{X}(\mathcal{O}_{K,S,\underline{X}}) \subseteq \mathfrak{X}(P; K).$$

Proof Note that $H_{\mathfrak{X},P}(t) = 1$ is equivalent to $\phi_{m_P}(\lambda(P, t)) = 1$, which is then equivalent to $r_{m_P}(\lambda(P, t)) = 1$. Suppose now that t is an S-integral point. Then $(t \cdot P)_v \equiv 0$ mod $m_P \Rightarrow e_{v,P}(t) > 0$ for all $v \notin S$. Thus,

$$\lambda(P, t) = \prod_{v \notin S} p_v^{m_{P_i}^{e_{v,P_i}(t)} \cdot q_{v,P_i}(t) \cdot f(v)} = \left(\prod_{v \notin S} p_v^{m_{P_i}^{e_{v,P_i}(t)-1} \cdot q_{v,P_i}(t) \cdot f(v)} \right)^{m_P},$$

whence $H_{\mathfrak{X},P}(t) = 1$ as $\lambda(P, t)$ is an m_P-power. ∎

We see that each point P with multiplicity $m_P > 1$ imposes a height dropping condition on the set of integral points. Thus, to study integral points, it suffices to study

$$\mathfrak{X}(\mathcal{O}_{K,S,\underline{X}}) \subseteq \bigcap_{m_P > 1} \mathfrak{X}(P; K).$$

We obtain the following, that the stacky part of the anti-canonical height cuts out the integral points.

Corollary 3.22 *Let $\mathfrak{X} = \mathfrak{X}(X : (\mathbf{p}, \mathbf{m}))$ be a stacky curve and $\mathcal{D}_{\mathfrak{X}} = \prod_{i=1}^{r} \mathcal{L}_{P_i}$. Then*

$$\mathfrak{X}(\mathcal{O}_{K,S}) \subseteq \{t \in X(K) : H_{\mathcal{D}_{\mathfrak{X}}}(t) = 1\}.$$

In other words, the integral points are precisely those points where the stacky part of the anti-canonical height vanishes.

Example 3.23 Let $X = C$ be an elliptic curve. Then $H_{\mathcal{D}_{\mathfrak{X}}} = H_{-K_{\mathfrak{X}}}$. In other words, the integral points of the stacky elliptic curve are precisely the points where the anti-canonical height vanishes. Since $\phi_m(x) = 1 \iff r_m(x) = 1$, we have that the S-integral points of a stacky elliptic curve are also precisely the points where the stacky canonical height vanishes. If $\mathfrak{X}$ is a scheme and X is an elliptic curve, then this is certainly true, as the canonical height is trivial and every rational point is integral, and vice versa.

Proof of Theorem 3.20 We have already shown part (1) of Theorem 3.20 in (3.21). We turn to part (2) and assume that $K = \mathbb{Q}$. We know that $\bigcap_{m_P > 1} \mathfrak{X}(P; \mathbb{Q}) \subseteq \mathfrak{X}(\mathcal{O}_{\mathbb{Q},S,\underline{X}})$ by (3.21). We now show the reverse inclusion. Let $t \in X(\mathbb{Q})$ with $H_{\mathfrak{X},P}(t) = 1$ for all P with $m_P > 1$. Since $K = \mathbb{Q}$, we have that $N(\mathfrak{p}_v) = p_v$ and $f(v) = 1$ for all finite places v. Fix P with $m_P > 1$. Toward a contradiction suppose that $(t \cdot P)_{v_0} \not\equiv 0 \mod m_P$ for some $v_0 \notin S$. Then $e_{v_0,P}(t) = 0$. Notice that $H_{\mathfrak{X},P}(t) = 1$

means that $\lambda(P, t)$ is an m_P-power. Since if $v \neq v'$ we have that $p_v \neq p_{v'}$, we have by unique factorization of integers that

$$\lambda(P, t) = \prod_{v \notin S} p_v^{m_P^{e_{v,P}(t)} \cdot q_{v,P}(t)} = \left(\prod_{v \notin S} p_v^{z_v(t)}\right)^{m_P}$$

for some integers $z_v(t)$. In particular, for v_0, we have

$$p_{v_0}^{m_P^{e_{v_0,P}(t)} \cdot q_{v_0,P}(t)} = p_{v_0}^{q_{v_0,P}(t)} = p_{v_0}^{z_{v_0}(t)m_P}.$$

Thus, $z_{v_0}(t)m_P = q_{v_0,P}(t)$, which contradicts $q_{v_0,P}(t)$ being indivisible by m_P. Thus, for all $m_P > 1$ and $v \notin S$, we have $(t \cdot P)_v \equiv 0 \mod m_P$ and t is an S-integral point of $\mathfrak{X}$ by definition. ∎

4 Stacky curves with coarse space $\mathbb{P}^1$

We focus on the situation when the base curve is $\mathbb{P}^1$. Let $X = \mathbb{P}^1_{\mathbb{Q}}$ and $S = \{v_\infty\}$ and take $\mathcal{L}$ to be $\mathcal{O}_{\mathbb{P}^1}(1)$, so the ample height is the usual one. We consider the stacky curve

$$\mathfrak{X} = (\mathbb{P}^1_{\mathbb{Q}} : (P_1, m_1), ..., (P_r, m_r)).$$

In this situation, the $\lambda(P, t)$ can be easily computed.

Proposition 4.1 *Let $t = [x : y]$ and suppose that $P_i = (a_i : b_i)$ where a_i, b_i are coprime integers. Then $\lambda(P_i, t) = |a_i y - b_i x|$.*

Proof We have that $(t \cdot P_i)_p = \max_n\{[x : y] \equiv [a_i : b_i] \mod p^n\}$. Note that this means there is some $\lambda \neq 0 \mod p^n$ with $(x, y) = \lambda(a_i, b_i) \mod p^n$. Since a_i, b_i have been taken coprime, we may assume that p does not divide a_i or b_i. Suppose that $p \nmid a_i$ (the other case is similar). Then $\lambda = \frac{x}{a_i} \mod p^n$ and therefore $b_i x - a_i y = 0 \mod p^n$. Thus, $(t \cdot P_i)_p = \mathrm{ord}_p(b_i x - a_i y)$. Then we have that

$$\lambda(P_i, t) = \prod_p p^{\mathrm{ord}_p(b_i x - a_i y)} = |a_i y - b_i x|,$$

as needed. ∎

Definition 4.2 (Euler characteristic of stacky curves [4]) Let $\mathfrak{X} = (X : (P_1, m_1), \ldots, (P_r, m_r))$ be a stacky curve. The *Euler characteristic* of $\mathfrak{X}$ is defined by the formula

$$\chi(\mathfrak{X}) = 2 - 2g(X) - \sum_{i=1}^r \left(1 - \frac{1}{m_i}\right) = 2 - 2g(X) - r + \sum_{i=1}^r \frac{1}{m_i},$$

where $g(X)$ is the genus of the curve X. We define the *genus* of a stacky curve by the formula $\chi(\mathfrak{X}) = 2 - 2g(\mathfrak{X})$.

We now begin assembling the necessary ingredients to compare our heights constructed in Section 3 to those in [6]. We first work with $H_{\mathcal{L}}^{\mathrm{st}}$.

Proposition 4.3 Let $\mathfrak{X} = \mathfrak{X}(X : (P_1, \ldots, P_r), (m_1, \ldots, m_r))$, and let $\mathcal{L}$ be a line bundle on $\mathfrak{X}$ with

$$\mathcal{L} \cong \pi_\mathfrak{X}^* M \otimes \prod_{i=1}^r \mathcal{L}_{P_i}^{\otimes d_i},$$

where $0 \le d_i < m_i$ and M is a line bundle on X with $\pi_\mathfrak{X}$ being the coarse space map. Let $H_\mathcal{L}^{\mathrm{st},\mathrm{ESZ\text{-}B}}$ be the stable height as constructed in [6]. Then $H_\mathcal{L}^{\mathrm{st}} = H_\mathcal{L}^{\mathrm{st},\mathrm{ESZ\text{-}B}}$.

Proof In [6], a general definition of the stable height is given. Let $m = \prod_{i=1}^r m_i$. By the properties of the stable height described in [6], we have that

$$H_\mathcal{L}^{\mathrm{st},\mathrm{ESZ\text{-}B}} = \left(H_{\mathcal{L}^{\otimes m}}^{\mathrm{st},\mathrm{ESZ\text{-}B}}\right)^{\frac{1}{m}}.$$

On the other hand,

$$\mathcal{L}^{\otimes m} = \pi_\mathfrak{X}^* M^{\otimes m} \otimes \prod \pi_\mathfrak{X}^* \mathcal{O}_X \left(\frac{d_i m}{m_i} P_i\right).$$

It is a fact that if L is a line bundle on X, then $H_{\pi_\mathfrak{X}^* L}^{\mathrm{st},\mathrm{ESZ\text{-}B}} = H_L \circ \pi_\mathfrak{X}$. Therefore, we have

$$H_{\mathcal{L}^{\otimes m}}^{\mathrm{st},\mathrm{ESZ\text{-}B}} = H_{\mathcal{L}^{\otimes m}} = H_M^m \cdot \prod_{i=1}^r H_{P_i}^{\frac{d_i m}{m_i}}.$$

Taking mth roots gives the desired inequality. $\blacksquare$

Corollary 4.4 Let $\mathfrak{X} = \mathfrak{X}(\mathbb{P}^1 : (P_1, \ldots, P_r), (m_1, \ldots, m_r)))$, and let $\mathcal{L}$ be a line bundle on $\mathfrak{X}$. Then

(4.1) $$H_\mathcal{L}^{\mathrm{st}} = H_{\mathbb{P}^1}^{\deg_\mathfrak{X} \mathcal{L}}.$$

Proof Let $\mathcal{L}$ be a line bundle on $\mathfrak{X}$. Then we may write

$$\mathcal{L} \cong \pi_\mathfrak{X}^* \mathcal{O}_{\mathbb{P}^1}(d) \otimes \prod_{i=1}^r \mathcal{L}_{P_i}^{\otimes d_i}.$$

On $\mathbb{P}^1$, we have that $\mathcal{O}_{\mathbb{P}^1}(P_i) \cong \mathcal{O}_{\mathbb{P}^1}(1)$. So, by definition, the stable height is

$$H_{\mathbb{P}^1}^d \cdot \prod H_{\mathbb{P}^1}^{\frac{d_i}{m_i}} = H_{\mathbb{P}^1}^{d + \sum_{i=1}^r \frac{d_i}{m_i}} = H_{\mathbb{P}^1}^{\deg_\mathfrak{X} \mathcal{L}},$$

as needed. $\blacksquare$

We can now precisely define the heights on a stacky curve with coarse space $\mathbb{P}^1_\mathbb{Q}$.

Definition 4.5 Let $\mathfrak{X} = (\mathbb{P}^1_\mathbb{Q}, (P_1, m_1), \ldots, (P_r, m_r))$ be a stacky curve. Set $P_i = [a_i : b_i]$ with a_i, b_i coprime integers and $\ell_i(t) = ax - by$ when $t = [x : y]$ for x, y coprime integers. Let $\mathcal{L} = \pi_\mathfrak{X}^* M \otimes \prod_{i=1}^r \mathcal{L}_{P_i}^{\otimes d_i}$ with $0 \le d_i < m_i$. Then define

$$H_\mathcal{L} = \max\{|x|, |y|\}^{\deg_\mathfrak{X} \mathcal{L}} \cdot \prod_{i=1}^r \phi_{m_i}(\ell_i(x, y)^{d_i})^{\frac{1}{m_i}}.$$

In particular, we have that

$$H_{-K_\mathfrak{X}}(t) = \max\{|x|, |y|\}^{\chi(\mathfrak{X})} \cdot \prod_{i=1}^r \phi_{m_i}(\ell_i(x, y))^{\frac{1}{m_i}}$$

and

$$H_{K_{\mathfrak{X}}}(t) = \max\{|x|, |y|\}^{-\chi(\mathfrak{X})} \cdot \prod_{i=1}^{r} r_{m_i}(\ell_i(x,y))^{\frac{1}{m_i}}.$$

When $P_1 = [1,0]$, $P_2 = [1,-1]$, $P_3 = [1:0]$, and $m_1 = m_2 = m_3 = 2$, we obtain the square root of the height function Question (2.7). We now have enough to prove our main comparison theorem.

4.1 Proof of Theorem 3.16

We follow the argument given in [6, p. 45]. Write $\mathcal{L} \cong \pi_{\mathfrak{X}}^* M \otimes \prod_{i=1}^{r} \mathcal{L}_{P_i}^{\otimes d_i}$. From [6, Section 2.3], we have a decomposition

$$(4.2) \qquad \frac{H_{\mathcal{L}}^{\text{ESZ-B}}}{H_{\mathcal{L}}^{\text{st,ESZ-B}}} = \prod_p \delta_{\mathcal{L},p},$$

where the $\delta_{\mathcal{L},p}$'s are the local discrepancies associated with $H_{\mathcal{L}}^{\text{ESZ-B}}$. On the other hand, we have that

$$(4.3) \qquad \frac{H_{\mathcal{L}}(P)}{H_{\mathcal{L}}^{\text{st}}(P)} = \prod_{i=1}^{r} \phi_{m_i}(\ell_i(P)^{d_i})^{\frac{1}{m_i}}.$$
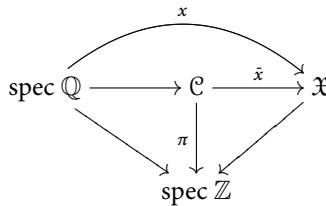
By Proposition 4.3, we have that

$$(4.4) \qquad H_{\mathcal{L}}^{\text{st,ESZ-B}} = H_{\mathcal{L}}^{\text{st}}$$

up to some positive constant. Therefore, it suffices to show that

$$\prod_p \delta_{\mathcal{L},p}(P) = \prod_{i=1}^{r} \phi_{m_i}(\ell_i(P)^{d_i})^{\frac{1}{m_i}}.$$

Let $x \colon \text{spec } \mathbb{Q} \to \mathfrak{X}$ be a rational point whose image is not any of the stacky points P_i. Then, in [6], there is a one-dimensional stack $\mathcal{C}$ called the *tuning stack* and a diagram:

Moreover, the local discrepancies can be computed at a prime p by

$$p^{-\deg \pi_* \tilde{x}^* \mathcal{L}^{-1} - (-\deg \tilde{x}^* \mathcal{L}^{-1})}.$$

These degrees can be computed locally on $\mathfrak{X}$ in terms of the stacky points P_i. In other words,

$$\deg \bar{x}^* \mathcal{L}^{-1} = \sum_{i=1}^{r} \deg_{P_i} \bar{x}^* \mathcal{L}^{-1},$$

$$\deg \pi_* \bar{x}^* \mathcal{L}^{-1} = \sum_{i=1}^{r} \deg_{P_i} \pi_* \bar{x}^* \mathcal{L}^{-1}.$$

The local degree of $\mathcal{L}$ at P_i at $\mathbb{Q}/\mathbb{Z}$ is $\frac{d_i}{m_i}$. Following [6], we have the local degrees

$$\deg_{P_i} \bar{x}^* \mathcal{L}^{-1} = -\frac{d_i \operatorname{ord}_p(\ell_i(x))}{m_i},$$

$$\deg_{P_i} \pi_* \bar{x}^* \mathcal{L}^{-1} = \left\lfloor -\frac{d_i \operatorname{ord}_p(\ell_i(x))}{m_i} \right\rfloor.$$

We obtain that the contribution at P_i to the local discrepancy at p can be written as

$$p^{\left\lceil \frac{d_i \ell_i(x)}{m_i} \right\rceil - \frac{d_i \ell_i(x)}{m_i}}.$$

Now write $d_i \ell_i(x) = q_i m_i + r_i$ with $0 \le r_i \le m_i - 1$. First, suppose that $r_i = 0$. Then

$$\left\lceil \frac{d_i \ell_i(x)}{m_i} \right\rceil - \frac{d_i \ell_i(x)}{m_i} = \left\lceil \frac{q_i m_i}{m_i} \right\rceil - \frac{q_i m_i}{m_i} = q_i - q_i = 0.$$

Now suppose that $r_i \ne 0$. Then we have that

$$\left\lceil \frac{d_i \ell_i(x)}{m_i} \right\rceil - \frac{d_i \ell_i(x)}{m_i} = \left\lceil \frac{q_i m_i + r_i}{m_i} \right\rceil - \frac{q_i m_i + r_i}{m_i}$$

$$= q_i + \left\lceil \frac{r_i}{m_i} \right\rceil - \frac{r_i}{m_i} - q_i$$

$$= 1 - \frac{r_i}{m_i} = \frac{m_i - r_i}{m_i}.$$

In other words, we have shown that

$$\left\lceil \frac{d_i \ell_i(x)}{m_i} \right\rceil - \frac{d_i \ell_i(x)}{m_i} = \begin{cases} 0, & d_i \ell_i(x) = 0 \mod m_i, \\ -d_i \ell_i(x) \mod m_i, & d_i \ell_i(x) \ne 0 \mod m_i, \end{cases}$$

$$= \frac{-d_i \ell_i(x) \mod m_i}{m_i} = \frac{N_{m_i, d_i}(\operatorname{ord}_p(\ell_i(x)))}{m_i}.$$

Thus, the local discrepancies are given by

$$\prod_p \delta_{\mathcal{L}, p}(x) = \prod_{i=1}^{r} \cdot \left(\prod_p p^{\frac{N_{m_i, d_i}(\operatorname{ord}_p(\ell_i(x)))}{m_i}} \right)$$

$$= \prod_{i=1}^{r} \phi_{m_i} (\ell_i(x)^{d_i})^{\frac{1}{m_i}}.$$

Combining equations 4.2–4.4 gives the desired conclusion.

Question 4.6 Consider a stacky curve $\mathfrak{X} = (X : (P_1, m_1), \ldots, (P_r, m_r))$. A line bundle M on X with a chosen height function H_M and integers $0 \le d_i < m_i$. Then does

$$H_{\mathcal{L}}(t) = H_M(t) \cdot \prod_{i=1}^{r} H_{P_i}^{\frac{d_i}{m_i}}(t) \cdot \prod_{i=1}^{r} \phi_{m_i}(\lambda(P_i, t)^{d_i})^{\frac{1}{m_i}}$$

agree with the height constructed in [6] up to a bounded constant when $X \ne \mathbb{P}^1_{\mathbb{Q}}$. In other words, does our stacky height machine recover the heights in [6] for all stacky curves over all number fields? If not, can one define *different* size functions $\tilde{N}_{m_i, d_i} \colon \mathbb{Z}_{\ge 0} \to \mathbb{Z}_{\ge 0}$ so that the ESZ-B height associated with $\mathcal{L}$ is of the form

$$H_{\mathcal{L}}(t) = \left(H_M(t) \cdot \prod_{i=1}^{r} H_{P_i}^{\frac{d_i}{m_i}}(t) \right) \cdot \left(\prod_{i=1}^{r} \left(\prod_{p} p^{\tilde{N}_{m_i, d_i}(t \cdot P_i)_p} \right)^{\frac{1}{m_i}} \right).$$

This result will follow provided one can show that the local degree of $\bar{x}^* \mathcal{L}$ with respect to P_i over a prime p is

$$\frac{d_i \operatorname{ord}_p(\lambda(P_i, t))}{m_i} = \frac{d_i(t \cdot P_i)_p}{m_i}.$$

In this case, the argument given in Theorem 3.16 would give the desired result. Further, one might ask if these methods could be extended to compute the height functions of line bundles certain higher dimensional analogues of stacky curves.

4.2 Morphisms of stacky curves

We will require some results on morphisms between stacky curves.

Definition 4.7 (Darmon [4]) Let $\mathfrak{X}_1 = (X_1, (P, m_P)), \mathfrak{X}_2 = (X_2, (Q, m_Q))$ be M curves defined over a number field K. A morphism of stacky curves over K is a morphism of algebraic curves $\pi \colon X_1 \to X_2$ defined over K such that for all $P \in X_1(K)$ with $\pi(P) = Q$, we have that

$$m_Q \mid e_\pi(P) m_P,$$

where $e_\pi(P)$ is the ramification index of π at P. We also define $e_{\underline{\pi}}(P) = \frac{e_\pi(P) m_P}{m_Q}$ the ramification index of $\underline{\pi}$ at P.

Now let $\mathfrak{X} = (X : \mathbb{Q}; (P_1, m_1), \ldots, (P_r, m_r))$ be an M curve. For any $s < r$, choose positive divisors d_i of m_i for $i = 1, \ldots, s$. Then there is a multiplicity lowering morphism

$$\underline{\pi}(d_1, \ldots, d_s) \colon \mathfrak{X}(X : (P_1, m_1), \ldots, (P_r, m_r)) \to \mathfrak{X}(X : (P_1, d_1), \ldots, (P_s, d_s))$$

defined by the identity morphism on X. The usefulness of this notion can be seen by the following.

Proposition 4.8 (Darmon [4]) *Let $\underline{\pi} \colon \mathfrak{X}_1 \to \mathfrak{X}_2$ be a morphism of stacky curves defined over K. Then*

$$\underline{\pi}(\mathfrak{X}_1(\mathcal{O}_{K,S})) \subseteq \mathfrak{X}_2(\mathcal{O}_{K,S}).$$

In other words, a morphism of stacky curves preserves the notion of S-integral points.

Lemma 4.9 *Let K be a field and*

$$U = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Define a bilinear form $L(\mathbf{v}, \mathbf{x}) = \mathbf{v}^T U^T \mathbf{x}$. Let T be a non-singular matrix with entries in K. Then

$$L(T\mathbf{v}, T\mathbf{x}) = (\det T)L(\mathbf{v}, \mathbf{x}).$$

Proof We have that

$$L(T\mathbf{v}, T\mathbf{x}) = (T\mathbf{v})^T U^T T\mathbf{x} = \mathbf{v}T^T U^T T\mathbf{x} = \mathbf{v}(UT)^T T\mathbf{x}.$$

Direct computation shows that $(UT)^T T = (\det T)U^T$. We then have

$$L(T\mathbf{v}, T\mathbf{x}) = \mathbf{v}(UT)^T T\mathbf{x} = \mathbf{v}^T(\det T)U^T \mathbf{x} = (\det T)L(\mathbf{v}, \mathbf{x}). \qquad \blacksquare$$

Note that if $P = [a : b]$ and $t = [x, y]$ where a, b and x, y are coprime integers, then $\lambda(P, t) = ay - bx$ by (4.1). In other words, $\lambda(P, t) = L(P, t)$ as defined in (4.9).

Lemma 4.10 *Let $P = [a : b], t = [x, y] \in \mathbb{P}^1_{\mathbb{Q}}$ with a, b and x, y coprime integers. Fix an integer $m > 1$. Let $\alpha \colon \mathbb{P}^1_{\mathbb{Q}} \to \mathbb{P}^1_{\mathbb{Q}}$ be an automorphism. Let $\det \alpha$ be the smallest possible nonnegative determinant of an integral representation of α and similarly for $\det \alpha^{-1}$. Then we have*

(4.5)
$$(\mathrm{rad}_m(\det \alpha^{-1}))^{-(m-1)} \phi_m(\lambda(P, t)) \le \phi_m(\lambda(\alpha(P), \alpha(t))) \le \mathrm{rad}_m(\det \alpha)^{m-1} \phi_m(\lambda(P, t)).$$

Proof Let L be as (4.1). Then we have that

(4.6) $$\lambda(P, t) = L\left(\frac{\alpha(P)}{d_1}, \frac{\alpha(t)}{d_2}\right) = \frac{(\det \alpha)}{d_1 d_2} L(P, t) = \frac{(\det \alpha)}{d_1 d_2} \lambda(P, t)$$

for some integers d_1, d_2 which account for common factors of $\alpha(P)$ and $\alpha(t)$. Let $n_1 = q_1 m + r_1, n_2 = q_2 m + r_2$ be integers with $0 \le r_i < m$.

(4.7)

$$\phi_m(p^{n_1 + n_2}) = \begin{cases} p^{m - r_1 - r_2} = \dfrac{\phi_m(p^{n_1})\phi_m(p^{n_2})}{p^m}, & \text{if } r_1 + r_2 < m, \\ p^{m - (r_1 + r_2 - m)} = p^{2m - r_1 - r_2} = \phi_m(p^{n_1})\phi_m(p^{n_2}), & \text{if } r_1 + r_2 \ge m. \end{cases}$$

Since ϕ_m is multiplicative, we have that $\phi_m(zw) \le \phi_m(z)\phi_m(w) \le \mathrm{rad}_m(z)^{m-1}\phi_m(w)$. Therefore, using (4.6),

$$\phi_m(\lambda(\alpha(P), \alpha(t))) \le \mathrm{rad}_m(\det \alpha)^{m-1}\phi_m(\lambda(P, t)).$$

Applying the same reasoning using α^{-1}, we have that

$$\phi_m(\lambda(P, t)) \le \mathrm{rad}_m(\det \alpha^{-1})^{m-1}\lambda(\phi(P), \phi(t)).$$

Therefore, we have

(4.8)
$$\left(\mathrm{rad}_m(\det \alpha^{-1})\right)^{-(m-1)} \phi_m(\lambda(P,t)) \le \phi_m(\lambda(\alpha(P),\alpha(t))) \le \mathrm{rad}_m(\det \alpha)^{m-1} \phi_m(\lambda(P,t)),$$

as required. ∎

Corollary 4.11   *Let $\mathfrak{X} = \mathfrak{X}(\mathbb{P}^1; (P_1,m_1),\dots,(P_r,m_r))$. Let $\alpha\colon \mathbb{P}^1_{\mathbb{Q}} \to \mathbb{P}^1_{\mathbb{Q}}$ be an automorphism of $\mathbb{P}^1_{\mathbb{Q}}$. Let $Q_i = \alpha(P_i)$ and $\mathfrak{Y} = \mathfrak{Y}(\mathbb{P}^1; (Q_1,m_1),\dots,(Q_r,m_r))$. Then α induces an isomorphism $\underline{\alpha}\colon \mathfrak{X} \to \mathfrak{Y}$. Let $\det \alpha$ and $\det \alpha^{-1}$ be as in (4.10). Let $D(\alpha,\mathfrak{X}) = \prod_{i=1}^r \mathrm{rad}_{m_i}(\det \alpha)^{1-1/m_i}$ and similarly define $D(\alpha^{-1},\mathfrak{X})$. Suppose that C_ϕ is a constant (such a constant always exists) such that*

(4.9)
$$C_\alpha^{-1} H_{\mathbb{P}^1}(t)^{\chi(\mathfrak{X})} \le H_{\mathbb{P}^1}(\phi(t))^{\chi(\mathfrak{X})} \le C_\alpha H_{\mathbb{P}^1}(t)^{\chi(\mathfrak{X})}.$$

Then

(4.10)
$$D(\alpha^{-1},\mathfrak{X})^{-1} C_\alpha^{-1} H_{-K_{\mathfrak{X}}}(t) \le H_{\mathbb{T}_{\mathfrak{Y}}}(\phi(t)) \le D(\alpha,\mathfrak{X}) C_\alpha H_{-K_{\mathfrak{X}}}(t).$$

Proof   By assumption, for each $i = 1,\dots,r$, we have from (4.10) and our assumption that

$$
\begin{aligned}
H_{\mathcal{T}_{\mathfrak{Y}}}(\phi(t)) &= H_{\mathbb{P}^1}(\phi(t))^{\chi(\mathfrak{Y})} \prod_{i=1}^r \phi_{m_i}(\lambda(Q_i,\phi(t)))^{\frac{1}{m_i}} \\
&\le C_\alpha H_{\mathbb{P}^1}(t)^{\chi(\mathfrak{X})} \prod_{i=1}^r \mathrm{rad}_{m_i}(\det(\alpha))^{1-1/m_i} \phi_{m_i}(\lambda(P_i,t))^{\frac{1}{m_i}} \\
&= C_\alpha D(\alpha,\mathfrak{X}) H_{-K_{\mathfrak{X}}}(t).
\end{aligned}
$$

Similarly, we have

$$
\begin{aligned}
H_{\mathcal{T}_{\mathfrak{Y}}}(\phi(t)) &= H_{\mathbb{P}^1}(\phi(t))^{\chi(\mathfrak{Y})} \prod_{i=1}^r \phi_{m_i}(\lambda(Q_i,\phi(t)))^{\frac{1}{m_i}} \\
&\ge C_\alpha^{-1} H_{\mathbb{P}^1}(t)^{\chi(\mathfrak{X})} \prod_{i=1}^r \mathrm{rad}_{m_i}(\det(\alpha^{-1}))^{-(1-1/m_i)} \phi_{m_i}(\lambda(P_i,t))^{\frac{1}{m_i}} \\
&= C_\alpha^{-1} D(\alpha^{-1},\mathfrak{X}) H_{-K_{\mathfrak{X}}}(t). \qquad \blacksquare
\end{aligned}
$$

We can get slightly worse, but more understandable bounds as follows. We always have $\mathrm{rad}_m(x) \le \mathrm{rad}(x)$. Note that we have that $\sum_{i=1}^r (1-1/m_i) = 2 - \chi(\mathfrak{X}) = 2g(\mathfrak{X})$. Thus, we in fact have

(4.11)
$$C_\alpha^{-1} \mathrm{rad}(\det \alpha^{-1})^{-2g(\mathfrak{X})} H_{-K_{\mathfrak{X}}}(t) \le H_{\mathcal{T}_{\mathfrak{Y}}}(\phi(t)) \le C_\alpha \mathrm{rad}(\det \alpha)^{2g(\mathfrak{X})} H_{-K_{\mathfrak{X}}}(t).$$

Of particular note is that we see that when studying the Northcott property, we may change the height by an automorphism. Thus, the Northcott property is stable under isomorphism as expected.

4.3 Northcott property of the canonical height on stacky curves

We now investigate the properties of the canonical height on stacky curves, given by

$$(4.12) \qquad H_{(\mathbf{a},\mathbf{m})}(x,y) = \prod_{i=1}^{n} r_{m_i}(\ell_i(x,y))^{\frac{1}{m_i}} \cdot \max\{|x|,|y|\}^{-\delta(\mathbf{m})}.$$

When $\delta(\mathbf{m}) = 0$, we see that the canonical height exhibits a clear duality with the anti-canonical height, and so the same argument shows that Northcott's property will fail. When $\delta(\mathbf{m}) < 0$, we then see at once that $H_{(\mathbf{a},\mathbf{m})}$ will have Northcott's property as a consequence that the Weil height having Northcott's property. It remains to consider Northcott's property when $\delta(\mathbf{m}) > 0$. In this case, we have $\mathbf{m} = (2, m_2, m_3)$ with

$$\frac{1}{m_2} + \frac{1}{m_3} > \frac{1}{2}.$$

It suffices to show that for any such pair (m_2, m_3), there exist integers a, b, c such that the curve

$$ax^2 + by^{m_2} + cz^{m_3} = 0, \quad \frac{1}{2} + \frac{1}{m_2} + \frac{1}{m_3} > 1$$

has infinitely many primitive integral solutions. This is the content of Beukers' paper [1], and we are done.

5 On the Northcott property of canonical and anti-canonical heights on stacky curves

In this section, we prove Theorem 1.2, starting with Theorem 2.1. We start with a reduction procedure of a curve $\mathfrak{X}(\mathbb{P}^1 : (\mathbf{a}, \mathbf{m}))$ which we describe colloquially. By convention, we shall take our weight vectors $\mathbf{m}$ to have the property that

$$m_1 \le m_2 \le \cdots \le m_n.$$

Definition 5.1 Consider a stacky curve $\mathfrak{X}(X : (\mathbf{a}, \mathbf{m}))$ with $\mathbf{a} = (P_1, \dots, P_r)$. Let $\mathbf{i} = i_1, \dots, i_k$ be a sub-sequence of $1, 2, \dots, r$. Then there is a morphism $\pi_{\mathbf{i}} \colon \mathfrak{X}(X : (\mathbf{a}, \mathbf{m})) \to \mathfrak{X}(X : (\mathbf{a}', \mathbf{m}'))$ where $\mathbf{a}' = (P_{i_1}, \dots, P_{i_k})$ and $\mathbf{m}' = (m_{i_1}, \dots, m_{i_k})$. The map is defined by taking the identity morphism on the coarse space X. We call such a morphism a *totally ramified canonical covering*.

The above construction defines a morphism by the definition of a morphism of M-curves. It is totally ramified in the sense that if i is some index that does not appear in $\mathbf{i}$, then $\pi_{\mathbf{i}}$ has ramification index m_i at P_i. We use the term canonical this type of construction can be used for any stacky curve. In particular, by taking $\mathbf{i}$ to be the empty set, we obtain the coarse space morphism. We will show that if Theorem 2.1 holds for a totally ramified canonical covering of the shape $\mathfrak{X}(\mathbb{P}^1 : (\mathbf{a}', \mathbf{m}'))$ where $\mathbf{a}', \mathbf{m}'$ is obtained from $\mathbf{a}, \mathbf{m}$, respectively, by removing a subset of indices, then it also holds for $\mathfrak{X}(\mathbb{P}^1 : (\mathbf{a}, \mathbf{m}))$ (see Theorem 5.2 below).

Theorem 5.2 *Let $\mathfrak{X}(\mathbb{P}^1 : (a_1, m_1), \dots, (a_n, m_n))$ be a stacky curve. If the Northcott property fails for the height (2.2) for some totally ramified canonical cover of $\mathfrak{X}$, then it will also fail for $\mathfrak{X}$.*

Proof Let $\mathfrak{X}$ be given as in the statement of Theorem 5.2. We may assume, after reindexing the points if necessary, that the Northcott property for the ESZ-B height fails for the totally ramified canonical cover given by

$$\mathfrak{X}^{(1)} = \mathfrak{X}(\mathbb{P}^1 : (a_1, m_1), \ldots, (a_k, n_k))$$

for some $k \leq n$. This implies that, for some positive number C_k depending at most on $a_1, \ldots, a_k$ and $m_1, \ldots, m_k$, there are infinitely many integers x, y such that

$$H_{(\mathbf{a}^{(k)}, \mathbf{m}^{(k)})} = \prod_{i=1}^{k} \phi_{m_i}(\ell_i(x, y))^{1/m_i} \max\{|x|, |y|\}^{\delta(\mathbf{m}^{(k)})} < C_k.$$

Next, note that the quotient

$$\mathcal{Q} = \frac{H_{(\mathbf{a}, \mathbf{m})}(x, y)}{H_{(\mathbf{a}^{(k)}, \mathbf{m}^{(k)})}(x, y)} = \prod_{i=k+1}^{n} \phi_{m_i}(\ell_i(x, y))^{1/m_i} \cdot \max\{|x|, |y|\}^{\sum_{i=k+1}^{n}(-1+1/m_i)}.$$

Observe that $\phi_m(s) \leq |s|^{m-1}$ for any integer s, with equality if and only if s is square-free. It follows that

$$\mathcal{Q} \leq \prod_{i=k+1}^{n} |\ell_i(x, y)|^{1-1/m_i} \cdot \max\{|x|, |y|\}^{\sum_{i=k+1}^{n}(-1+1/m_i)},$$

and from here we immediately see from the triangle inequality that

$$\mathcal{Q} \ll_{\mathbf{a}} 1.$$

Thus, by replacing C_k with a larger positive number if necessary, we see that the Northcott property also fails for $H_{(\mathbf{a}, \mathbf{m})}$ on $\mathfrak{X}$. ∎

Now, given Theorem 5.2, it remains to consider certain *minimal* choices of $\mathbf{m}$. We say that $\delta(\mathbf{m})$ is *minimally nonnegative* if there is no subsequence $\mathbf{m}'$ of $\mathbf{m}$ such that $\delta(\mathbf{m}') \leq 0$. We have the following lemma characterizing the minimally nonnegative tuples.

Lemma 5.3 *Suppose* $\mathbf{m} = (m_1, \ldots, m_n)$ *with* $2 \leq m_1 \leq \cdots \leq m_n$ *is minimally nonnegative. Then* $n \leq 4$.

Proof Suppose $n \geq 5$. If there exist $m_i, m_j, m_k \geq 3$, then the sub-sequence (m_i, m_j, m_k) satisfies $\delta((m_i, m_j, m_k)) \leq 0$, so $\mathbf{m}$ is not minimally nonnegative. If $m_3 \geq 3$, then such a choice of i, j, k exists, since $n \geq 5$. Therefore, we may assume that $m_1 = m_2 = m_3 = 2$. But then, $(2, 2, 2, m_4)$ satisfies $\delta((2, 2, 2, m_4)) \leq 0$, so $\mathbf{m}$ is not minimally nonnegative. ∎

It remains to deal with minimally nonnegative tuples with $n = 3, 4$. Before we proceed, we will require Theorem 2.2, which we prove now.

Proof of Theorem 2.2 As we remarked earlier, the proof given here is provided to us by Shnidman in [15].

For given non-singular binary quartic form $F \in \mathbb{Z}[x, y]$ given by

$$F(x, y) = a_4 x^4 + a_3 x^3 y + a_2 x^2 y^2 + a_1 x y^3 + a_0 y^4,$$

we write C_F for the curve defined by

$$C_F : z^2 = F(u, v).$$

The Jacobian of the genus one curve $C_{a,b}$ is the elliptic curve $E_{a,b}$ given by

$$E_F : y^2 = x^3 - \frac{I(F)}{3}x - \frac{J(F)}{27},$$

where I, J are the basic invariants given by

$$I(F) = 12a_4a_0 - 3a_3a_1 + a_2^2, J(F) = 72a_4a_2a_0 + 9a_3a_2a_1 - 27a_0a_3^2 - 27a_4a_1^2 - 2a_2^3.$$

By 2-descent, we see that C_F corresponds to a class c in $H^1(\mathbb{Q}, E_{a,b}[2])$. Note that for any integer d, the group $H^1(\mathbb{Q}, E_F^{(d)}[2])$ is canonically isomorphic to $H^1(\mathbb{Q}, E_F^{(d)}[2])$ such that c is the class of $C_F^{(d)}$ in $H^1(\mathbb{Q}, E_F^{(d)}[2])$.

We now consider two cases. First suppose that c does not come from 2-torsion. In this case, it is immediate that $E_F^{(d_0)}$ has positive rank, where

$$d_0 = F(u_0, v_0), u_0, v_0 \in \mathbb{Z}.$$

This is because in this case $C_F^{(d_0)}(\mathbb{Q}) \neq \varnothing$.

If c comes from 2-torsion, then we note that $C_F^{(d)}(\mathbb{Q})$ is non-empty for all $d \in \mathbb{Z}$. That is, for all $d \in \mathbb{Z}$, we have $C_F^{(d)}$ is isomorphic to $E_F^{(d)}$ over $\mathbb{Q}$. We can then choose a class c' in $H^1(\mathbb{Q}, E_F[2])$, represented by a different binary quartic form G, and choose d such that the twist of the genus one curve $\mathcal{C} : z^2 = G(u, v)$ given by $\mathcal{C}^{(d)}$ has a rational point. This implies that $E_F^{(d)}$ has positive rank. Then, with this choice of d, we find that $C_F^{(d)}(\mathbb{Q}) \neq \varnothing$ and $E_F^{(d)}$ has positive rank, which completes the proof. ∎

With Theorem 2.2, we proceed to handle minimally nonnegative tuples, starting with the case $n = 4$.

5.1 Minimally nonnegative tuples with $n = 4$

We begin with the case $\mathbf{m} = (2, 2, 2, 2)$, and we will need Theorem 2.2. By 3-transitivity of the action of PGL_2 on $\mathbb{P}^1$ and Lemma 4.10, we may assume that three of the points are $0, 1, \infty$, corresponding to the linear forms $x, y, x + y$ in the variables x, y. We then write $\ell(x, y) = ax + by$ for the linear form representing the fourth half-point.

We prove the following as a warm-up.

Lemma 5.4　*There exist integers a, b such that the stack $\mathfrak{X}(\mathbb{P}^1_\mathbb{Q} : (0, 2), (1, 2), (\infty, 2), (a/b, 2))$ has infinitely many rational points of E-S-ZB height equal to one.*

Proof　In this case, the height is given by

$$H(x, y) = \mathrm{sqf}(x)\,\mathrm{sqf}(y)\,\mathrm{sqf}(x + y)\,\mathrm{sqf}(ax + by),$$

so this is equal to one if and only if each of $x, y, x + y, ax + by$ is a square. To wit, we set

$$x = x_1^2, y = x_2^2, x + y = x_3^2.$$

This induces the equation

$$x_1^2 + x_2^2 = x_3^2,$$

which is solvable and whose (primitive) integral solutions are parametrized by

$$x_1 = 2uv, x_2 = u^2 - v^2, x_3 = u^2 + v^2.$$

Inserting this into $ax + by$ gives

$$a(2uv)^2 + b(u^2 - v^2)^2 = F_{a,b}(u, v).$$

We then fix $u = u_0, v = v_0$ so that $2u_0 v_0$ and $u_0^2 - v_0^2$ are co-prime, then solve the linear diophantine equation

$$a(2u_0 v_0)^2 + b(u_0^2 - v_0^2) = 1.$$

Given a solution (a, b) to this diophantine equation, one obtains a genus curve defined by

$$w^2 = F_{a,b}(u, v)$$

which is isomorphic to an elliptic curve, since it has a rational point given by $w_0^2 = F_{a,b}(u_0, v_0)$. In particular, it must be isomorphic to its Jacobian. A simple calculation shows that the Jacobian of this genus one curve is given by the equation

$$(5.1) \qquad E_{a,b} : y^2 = x^3 - \frac{16(a^2 - ab + b^2)}{3} - \frac{64(a + b)(2a - b)(a - 2b)}{27}$$
$$= \left(x - \frac{4b - 8a}{3} \right) \left(x - \frac{4a - 8b}{3} \right) \left(x - \frac{4b + 4a}{3} \right),$$

so it suffices to find a, b such that $E_{a,b}$ has positive rank. We find that setting $u_0 = 1, v_0 = 5$ and $a = 17, b = -118$ that the curve $E_{a,b}$ has positive rank, and therefore $w^2 = F_{a,b}(u, v)$ will have infinitely many integral solutions (u, v, w). This gives infinitely many pairs u, v such that $F_{17,-118}(u, v)$ is a square. This implies our result, since

$$\mathrm{sqf}(x)\,\mathrm{sqf}(y)\,\mathrm{sqf}(x + y)\,\mathrm{sqf}(17x - 118y)$$
$$= \mathrm{sqf}((2uv)^2)\,\mathrm{sqf}((u^2 - v^2)^2)\,\mathrm{sqf}((u^2 + v^2)^2)\,\mathrm{sqf}(F_{17,-118}(u, v)) = 1. \qquad \blacksquare$$

The general case will follow by applying the same ideas in tandem with Theorem 2.2. Indeed, Theorem 2.2 gives that for any $a, b \in \mathbb{Z}$ such that $F_{a,b}(u, v) = a(u^2 - v^2)^2 + 4bu^2v^2$ is non-singular that there exists $d \in \mathbb{Z}$ such that $C_F(\mathbb{Q}) \neq \varnothing$ and E_F has positive rank. Fixing such a d, we see that there are infinitely many co-prime integers u, v, z such that

$$dz^2 = F(u, v).$$

Recall that in this setup we have

$$x = (u^2 - v^2)^2, y = 4u^2v^2, x + y = (u^2 + v^2)^2,$$

whence

$$H(x, y) = 1 \cdot 1 \cdot 1 \cdot \operatorname{sqf}(F(u, v)) \le |d|.$$

This concludes the proof for the $\mathbf{m} = (2, 2, 2, 2)$ case.

Note that if $\mathbf{m} = (m_1, m_2, m_3, m_4)$ is minimally nonnegative, then $m_1 = m_2 = 2$, since $\delta((3, 3, 3)) = 0$. Thus, we may write

$$x = x_1 x_2^2, \, y = y_1 y_2^2$$

with x_1, y_1 square-free. We then write

$$\ell_3(x, y) = x + y = z_1 z_2^2 \cdots z_{m_3}^{m_3}, \, \ell_4(x, y) = w_1 w_2^2 \cdots w_{m_4}^{m_4}.$$

Again, we have z_i, w_j are square-free for $1 \le i \le m_3 - 1$ and $1 \le j \le m_4 - 1$.

We now specialize to the points where $z_i = w_j = 1$ except for $i = 2$ and $j = 1, 2$, as well as $x_1 = y_1 = 1$. Applying Theorem 2.2 and using the same argument as in the $(2, 2, 2, 2)$ case, we see that there is a choice of $w_1 = d$ such that there are infinitely many choices of x_2, y_2, w_2, z_2 satisfying

$$x = x_2^2, \, y = y_2^2, \, x + y = z_2^2, \, \ell_4(x, y) = dw_2^2.$$

The height of such a point is given by

$$\phi_2(x_2^2)\phi_2(y_2^2)\phi_{m_3}(z_2^2)\phi_{m_4}(dw_2^2) \max\{x_2^2, y_2^2\}^{1/m_3 + 1/m_4 - 1}$$

$$\ll |z_2|^{1 - 2/m_3} |w_2|^{1 - 2/m_4} \max\{|z_2|, |w_2|\}^{(2/m_3 - 1) + (2/m_4 - 1)} \ll 1.$$

It follows that there are infinitely many points of bounded height, and so Northcott's property fails.

5.2 Minimally nonnegative tuples with $n = 3$

To complete the proof of Theorem 2.1, it remains to handle the cases when $n = 3$ and $\chi(\mathfrak{X}) \le 0$. We shall assume that $m_1 \le m_2 \le m_3$. We then note that $\delta(\mathbf{m}) \le 0$ if and only if one of the following conditions is satisfied:

(1) $m_1 \ge 3$.
(2) $m_1 = 2, m_2 = 3, m_3 \ge 6$.
(3) $m_1 = 2, m_2 \ge 4$.

We deal with the first case. We then write

$$(5.2) \qquad x = x_{1,1} x_{1,2}^2 \cdots x_{1,m_1}^{m_1}, \, y = x_{2,1} x_{2,2}^2 \cdots x_{2,m_2}^{m_2},$$

$$x + y = x_{3,1} x_{3,2}^2 \cdots x_{3,m_3}^{m_3}.$$

Now set

$$x_{i,j} = 1 \text{ for } (i, j) \notin \{(1, 3), (2, 3), (3, 3), (3, 1)\}$$

and

$$x_{1,3} = z_1, x_{2,3} = z_2, x_{3,3} = z_3, x_{3,1} = d.$$

Then the value of the height $H_{(\mathbf{a},\mathbf{m})}(x,y)$ in this case is given by

$$H(z_1^3, z_2^3) = \phi_{m_1}(z_1^3)^{1/m_1}\phi_{m_2}(z_2^3)^{1/m_2}\phi_{m_3}(dz_3^3)^{1/m_3}\max\{|z_1|^3, |z_2|^3\}^{1/m_1 + 1/m_2 + 1/m_3 - 1}$$

$$\leq d^{m_3-1}|z_1|^{\frac{m_1-3}{m_1}}|z_2|^{\frac{m_2-3}{m_2}}|z_3|^{\frac{m_3-3}{m_3}}\max\{|z_1|, |z_2|\}^{\left(\frac{3}{m_1}-1\right)+\left(\frac{3}{m_2}-1\right)+\left(\frac{3}{m_3}-1\right)}.$$

Observe that

$$|z_i|^{\frac{m_i-3}{3}}\max\{|z_1|, |z_2|\}^{-1+\frac{3}{m_i}} \ll 1$$

for $i = 1, 2, 3$.

It remains to choose d so that the plane cubic curve

$$z_1^3 + z_2^3 = dz_3^3$$

has infinitely many rational points. This is an easy consequence of the seminal work of Stewart and Top [18, Theorem 7], which in turn depends on the important work of Stewart in [17]. In particular, they showed that the number of cube-free integers d with $|d| \leq X$ such that the equation

$$x^3 + y^3 = d$$

defines an elliptic curve with rank at least 2 is asymptotically greater than $X^{1/3}$. We, of course, do not need such a strong statement; indeed, we only need one such d. This completes the proof for the case $m_1 \geq 3$.

We proceed to handle the case $m_1 = 2, m_2 \geq 4$. Using the same notation as in (5.2), we then set

$$x_{i,j} = 1 \text{ for all } (i,j) \notin \{(1,1), (1,2), (2,4), (3,4)\},$$

and set

$$x_{1,2} = z_1, x_{2,4} = z_2, x_{3,4} = z_3, x_{1,1} = d.$$

This gives a curve

$$dz_1^2 = z_3^4 - z_2^4.$$

We need to choose square-free d so that this curve has infinitely many integral solutions, and such a d exists by Theorem 2.2. The height $H_{(\mathbf{a},B\mathbf{m})}(x,y)$ is given by

$$H_{(\mathbf{a},\mathbf{m})}(x,y) = \phi_2(dz_1^2)^{1/2}\phi_{m_2}(z_2^3)^{1/m_i}\phi_{m_3}(dz_3^3)^{1/m_3}\max\{d|z_1|^2, |z_2|^4\}^{1/m_2 + 1/m_3 - 1/2}$$

$$\leq d|z_2|^{1-4/m_2}|z_3|^{1-4/m_3}\max\{d|z_1|^2, |z_2|^4\}^{1/m_2 + 1/m_3 - 1/2}.$$

Note that

$$|z_3|^4 \asymp \max\{dz_1^2, z_2^4\},$$

so we obtain the upper bound

$$\left(\frac{|z_2|}{\max\{|z_2|, |z_3|\}}\right)^{1-\frac{4}{m_2}} \cdot \left(\frac{|z_3|}{\max\{|z_2|, |z_3|\}}\right)^{1-\frac{4}{m_3}} \ll 1.$$

It follows that there are infinitely many integers x, y such that $H_{(\mathbf{a},\mathbf{m})}(x, y)$ remains bounded.

Finally, we resolve the case $m_1 = 2, m_2 = 3, m_3 \geq 6$. In this case, we use the fact that there exist integers a, b, c with a square-free, b cube-free, and c sixth power-free such that the equation

$$(5.3) \qquad ax^2 + by^3 + cz^6 = 0$$

has infinitely many primitive solutions (see [5, Section 6.3]). Thus, by fixing such a triple (a, b, c) and setting

$$x_{1,1} = a, x_{2,1}x_{2,2}^2 = b, x_{3,1} \cdots x_{3,5} = c,$$

$$x_{1,2} = u_1, x_{2,3} = u_2, x_{3,6} = u_3$$

and

$$x_{5,j} = 1 \text{ for all } j \geq 7,$$

we specialized a point on $\mathfrak{X}(\mathbb{P}^1 : (0, 2), (\infty, 3), (-1, 6))$ to the curve given by (5.3). The height of such a point is then bounded in terms of a, b, c only, and is thus absolutely bounded. This shows that the Northcott property fails in this case as well.

This concludes the proof of Theorem 2.1.

5.3 Proof of Theorems 2.5

We proceed to prove Theorem 2.5. The claim when $\delta(\mathbf{m}) = 0$ is covered in Theorem 2.1, so we will not discuss it again. When $\delta(\mathbf{m}) > 0$, we note that $n = 3$, and that in each such case, there exist integers $a_{\mathbf{m}}, b_{\mathbf{m}}, c_{\mathbf{m}}$ such that the equation

$$a_{\mathbf{m}}x^{m_1} + b_{\mathbf{m}}y^{m_2} + c_{\mathbf{m}}z^{m_3} = 0$$

has infinitely many primitive integral solutions (see, for example, [1]). This shows that the Northcott property fails for H^0.

We may now work with the case when $\delta(\mathbf{m}) < 0$. We see that the height H^0 is bounded below by

$$(5.4) \qquad \phi_{m_1}(x)^{1/m_1}\phi_{m_2}(y)^{1/m_2}\phi_{m_3}(x + y)^{1/m_3},$$

so it suffices to show that this quantity necessarily goes to infinity. We then use the notation from (5.2), to obtain the equation

$$(5.5) \qquad x_{1,m_1}^{m_1}\prod_{j=1}^{m_1-1} x_{1,j}^j + x_{2,m_2}^{m_2}\prod_{j=1}^{m_2-1} x_{2,j}^j = x_{3,m_3}^{m_3}\prod_{j=1}^{m_3-1} x_{3,j}^j.$$

By convention, we have that $x_{i,j}$ is square-free for $1 \leq i \leq 3$ and $1 \leq j \leq m_i - 1$. Thus, (5.4) is equal to

$$\prod_{j=1}^{m_1-1} x_{1,j}^{\frac{m_1-j}{m_1}}\prod_{j=1}^{m_2-1} x_{2,j}^{\frac{m_2-j}{m_2}}\prod_{j=1}^{m_3-1} x_{3,j}^{\frac{m_3-j}{m_3}}.$$

Viewing these products as coefficients in (5.5), we see that if H^0 is to be bounded, these coefficients must be bounded. Therefore, it suffices to check that for a fixed triple of integers a, b, c, the equation

$$ax^{m_1} + by^{m_2} + cz^{m_3} = 0$$

has finitely many primitive integer solutions when $1/m_1 + 1/m_2 + 1/m_3 < 1$. But this is exactly the content of Darmon and Granville's paper [5], so we are done.

6 Northcott property of perturbed anti-canonical heights and the abc-conjecture

In this section, we prove Theorem 1.3, starting with Theorem 2.6. We consider the property of recovering Northcott's property on a modified ESZ-B anti-canonical height on the stacky curve

$$\mathfrak{X} = \mathfrak{X}(\mathbb{P}^1_{\mathbb{Q}} : (\mathbf{a}, \mathbf{m})).$$

Here, the modified height takes the shape

$$H^\delta_{(\mathbf{a},\mathbf{m})}(x, y) = \prod_{i=1}^n \phi_{m_i}(\ell_i(x, y))^{1/m_i} \max\{|x|, |y|\}^\delta.$$

Since $\chi(\mathfrak{X}) = \delta(\mathbf{m}) = 2 - \sum_{i=1}^n (1 - \frac{1}{m_i}) \le 0$, our goal is to show that

$$\gamma(\mathfrak{X}) = \inf\{\delta \in \mathbb{R} : H^\delta_{(\mathbf{a},\mathbf{m})} \text{ has the Northcott property}\} = \chi(\mathfrak{X})$$

assuming the abc-conjecture. Recall that we have shown that $H^{\chi(\mathfrak{X})}_{(\mathbf{a},\mathbf{m})}$ does not have the Northcott property unconditionally. Thus, we must show that $H^{\chi(\mathfrak{X})+\kappa}_{(\mathbf{a},\mathbf{m})}$ has the Northcott property for all $\kappa > 0$. First, assume that $\chi(\mathfrak{X}) = 0$. The Northcott property for the standard height implies that $H^\delta_{(\mathbf{a},\mathbf{m})}(x, y)$ has the Northcott property whenever $\delta > 0$. So $\inf\{\delta \in \mathbb{R} : H^\delta_{(\mathbf{a},\mathbf{m})} \text{ has the Northcott property}\} = 0 = \chi(\mathfrak{X})$ as needed. Now suppose that $\chi(\mathfrak{X}) < 0$. Assume, without loss of generality, that $m_1 \le m_2 \le \cdots \le m_n$. We then write

$$\ell_i(x, y) = z_{i,1} z_{i,2}^2 \cdots z_{i,m_i-1}^{m_i-1} z_{i,m_i}^{m_i}.$$

We have

$$\phi_{m_i}(\ell_i(x, y))^{\frac{1}{m_i}} = \prod_{j=1}^{m_i-1} z_{i,j}^{\frac{m_i-j}{m_i}}.$$

It follows that

$$H^{\chi(\mathfrak{X})+\delta}_{(\mathbf{a},\mathbf{m})} = \max\{|x|, |y|\}^{\chi(\mathfrak{X})+\delta} \prod_{i=1}^n \prod_{j=1}^{m_i-1} z_{i,j}^{\frac{m_i-j}{m_i}}.$$

Suppose that the following inequality holds for any $\varepsilon > 0$:

$$(6.1) \qquad \prod_{i=1}^{n} \prod_{j=1}^{m_i-1} z_{i,j}^{\frac{m_i-j}{m_i}} \gg_{\varepsilon} \max\{|x|, |y|\}^{-\chi(\mathfrak{X})-\varepsilon}.$$

Then, by multiplying both sides of the equation by $\max\{|x|, |y|\}^{\chi(\mathfrak{X})+\kappa}$, we obtain

$$(6.2) \qquad \max\{|x|, |y|\}^{\chi(\mathfrak{X})+\kappa} \prod_{i=1}^{n} \prod_{j=1}^{m_i-1} z_{i,j}^{\frac{m_i-j}{m_i}} \gg_{\varepsilon} \max\{|x|, |y|\}^{\kappa-\varepsilon}.$$

Taking $\varepsilon = \frac{\kappa}{2}$, we have that

$$(6.3) \qquad H_{(\mathbf{a},\mathbf{m})}^{\chi(\mathfrak{X})+\kappa}(x, y) \gg_{\varepsilon} \max\{|x|, |y|\}^{\frac{\kappa}{2}}.$$

Thus, $H_{(\mathbf{a},\mathbf{m})}^{\chi(\mathfrak{X})+\kappa}(x, y)$ must have the Northcott property as it cannot remain bounded by the usual Northcott property for $\mathbb{P}^1$. Therefore, we are done if we can confirm inequality (6.1). To do so, we require the following proposition, due to Granville [11].

Proposition 6.1 (Granville) *Suppose that the abc-conjecture holds. Then, for any binary form F with nonzero discriminant and $\varepsilon > 0$, we have*

$$\operatorname{rad}(F(m, n)) = \prod_{p | F(m,n)} p \gg_{F,\varepsilon} \max\{|m|, |n|\}^{\deg F - 2 - \varepsilon}.$$

In other words, if the *abc*-conjecture holds, then the radical of $F(m, n)$ will be quite large compared to the variables m, n (provided that the degree is at least 3).

We will apply Proposition 6.1 to reduce the proof of Theorem 2.6 to a *linear programming problem*.

6.1 A linear program bound

Observe that for each $1 \leq i \leq n$,

$$\prod_{j=1}^{m_i} z_{i,j} \geq \operatorname{rad}\left(\prod_{j=1}^{m_i} z_{i,j}\right).$$

Applying Proposition 6.1 to the binary form

$$Q_{\mathbf{a}}(x, y) = \prod_{i=1}^{n} \ell_i(x, y)$$

in conjunction with the above observation, we obtain

$$(6.4) \qquad \prod_{i=1}^{n} \prod_{j=1}^{m_i} z_{i,j} \geq \operatorname{rad}\left(\prod_{i=1}^{n} \ell_i(x, y)\right) \gg_{\varepsilon} \max\{|x|, |y|\}^{n-2-\varepsilon}.$$

Similarly, for each i, we have the bound

$$(6.5) \qquad |\ell_i(x, y)| \ll \max\{|x|, |y|\}.$$

Taking logarithms and writing $y_{i,j} = \log|z_{i,j}|$, we then have an optimization problem:

$$(6.6) \qquad \min \sum_{i=1}^{n} \frac{1}{m_i} \sum_{j=1}^{m_i-1} (m_i - j) y_{i,j}$$

subject to

$$(6.7) \qquad \sum_{i=1}^{n} \sum_{j=1}^{m_i} y_{i,j} \ge (n-2-\varepsilon)\log B$$

and

$$(6.8) \qquad \sum_{j=1}^{m_i} j y_{i,j} \ll \log B,$$

where $B = \max\{|x|, |y|\}$. Further, we have $y_{i,j} \ge 0$ for all i, j.

We emphasize that, at this point, integrality no longer plays a role, and neither does the syzygies relating the $z_{i,j}$'s. Indeed, we only need to solve the above linear program allowing arbitrary real inputs.

Now put

$$c_{ij} = \frac{m_i - j}{m_i}$$

for $1 \le i \le n$ and $1 \le j \le m_i - 1$. Write $c_{i,m_i} = 0$ and let $\mathbf{c} = (c_{i,j})$ to be the column vector with

$$\mathbf{c}^T = [c_{1,1}, c_{1,2}, \ldots, c_{1,m_1-1}, 0, c_{2,1}, c_{2,2}, \ldots, c_{n,1}, \ldots c_{n,m_n-1}, 0].$$

We have that $\mathbf{c} \in \mathbb{R}^N$ where $N = \sum_{i=1}^{n} nm_i$.

Let A be the matrix with rows representing the constraints,

$$(6.9) \qquad C_0 : \sum_{i=1}^{n} \sum_{j=1}^{m_i} y_{i,j} \ge (n-2-\varepsilon)\log B$$

$$(6.10) \qquad C_i : -\sum_{j=1}^{m_i} j y_{i,j} \gg -\log B.$$

If we have taken $\mathbf{e}_{i,j}$ to be a basis of $\mathbb{R}^N$, then we have that the rows of A are given by

$$(6.11) \qquad \mathbf{a}_0 = \sum_{i=1}^{n} \sum_{j=1}^{m_i} \mathbf{e}_{i,j},$$

$$(6.12) \qquad \mathbf{a}_i = \sum_{j=1}^{m_i} -j\mathbf{e}_{i,j}.$$

Finally, let $\mathbf{b}$ be the column vector with $n+1$ entries representing the constraints given by (6.7) and (6.8). In other words, we have

$$(6.13) \qquad b_0 = (n-2-\varepsilon)\log B,$$

$$(6.14) \qquad b_i = -\log B.$$

Our linear programming problem is then the following: Let $\mathbf{y} = (y_{ij})$ ordered as above.

(6.15) Minimize: $\mathbf{c}^T \mathbf{y}$,

subject to: $A\mathbf{y} \geq \mathbf{b}$ and $\mathbf{y} \geq 0$.

The dual linear program is

(6.16) Maximize: $\mathbf{b}^T \mathbf{x}$,

subject to: $A^T \mathbf{x} \leq \mathbf{c}$ and $\mathbf{x} \geq 0$,

where $\mathbf{x} = [x_0, x_1, ..., x_n]$. We call a vector $\mathbf{x}$ dual feasible if $A^T \mathbf{x} \leq \mathbf{c}$ and vector $\mathbf{y}$ primal feasible if $A\mathbf{y} \geq \mathbf{b}$. We have the following well-known *weak duality statement*.

Lemma 6.2 (Weak duality)   *Let A be an $m \times n$ matrix with real entries and $\mathbf{c}$ an $n \times 1$ real vector and $\mathbf{b}$ an $m \times 1$ real vector. Consider the primal linear program*

Minimize: $\mathbf{c}^T$,

subject to: $A\mathbf{y} \geq \mathbf{b}$ and $\mathbf{y} \geq 0$

and the dual linear program

Maximize: $\mathbf{b}^T \mathbf{x}$,

subject to: $A^T \mathbf{x} \leq \mathbf{c}$ and $\mathbf{x} \geq 0$.

Let $\mathbf{y}$ be any primal feasible vector and $\mathbf{x}$ a dual feasible vector. Then

$$\mathbf{c}^T \mathbf{y} \geq \mathbf{b}^T \mathbf{x}.$$

Proof   Let $A = (a_{i,j})$. Because $\mathbf{y}$ is primal feasible, we have $A\mathbf{y} \geq \mathbf{b}$. Therefore, for all $1 \leq i \leq m$, we have

$$\sum_{j=1}^{n} a_{i,j} y_j \geq b_i.$$

Multiplying by x_i and summing over all i, we have

(6.17) $$\sum_{i=1}^{m} \sum_{j=1}^{n} a_{i,j} y_j x_i \geq \sum_{i=1}^{n} b_i x_i = \mathbf{b}^T \mathbf{x}.$$

On the other hand, because $\mathbf{x}$ is dual feasible, we have that $A^T \mathbf{x} \leq \mathbf{c}$, so for each $1 \leq j \leq n$, we have

$$\sum_{i=1}^{m} a_{j,i} x_i \leq c_j.$$

Multiplying by y_j and summing over all j gives

(6.18) $$\sum_{j=1}^{n} \sum_{i=1}^{m} a_{i,j} x_i y_j \leq \sum_{j=1}^{n} y_j c_j = \mathbf{c}^T \mathbf{y}.$$

Combining inequality (6.17) and inequality (6.18) gives

$$\mathbf{c}^T\mathbf{y} \geq \sum_{j=1}^{n}\sum_{i=1}^{m} a_{i,j}x_i y_j \geq \mathbf{b}^T\mathbf{x}.$$ ∎

Returning to our problem, the weak duality theorem tells us that it suffices to find a dual feasible solution $\mathbf{x} = [x_0, \ldots x_n]$ such that $\mathbf{b}^T\mathbf{x} \geq -\chi(\mathfrak{X}) + \varepsilon$. In other words, we seek $\mathbf{x} = [x_0, \ldots, x_n]$ with

$$\mathbf{b}^T\mathbf{x} = \log B\left((n-2-\varepsilon)x_1 - \sum_{j=1}^{n} x_i\right) \geq -\chi(\mathfrak{X}) + \varepsilon,$$

$$A^T\mathbf{x} \leq \mathbf{c},$$

$$\mathbf{x} \geq 0.$$

Take $\mathbf{x} = \left[1, \frac{1}{m_1}, \frac{1}{m_2}, \ldots, \frac{1}{m_n}\right]$. We first show that $\mathbf{x}$ is dual feasible. In this case, A is an $(n+1) \times \sum_{i=1}^{n} m_i$ matrix. So a row of A^T is indexed by a pair (i, j) with $1 \leq i \leq n$ and $1 \leq j \leq m_i$. We have that the (i, j) entry of $A^T\mathbf{x} = \mathbf{x}^T A$ can be computed as

$$x_0 - jx_i.$$

Therefore, to show that $\mathbf{x}$ is dual feasible for an arbitrary $\mathbb{X}$, we need that

$$x_0 - jx_i \leq c_{i,j} = \frac{m_i - j}{m_i} = 1 - \frac{j}{m_i}.$$

In our case, the (i, j) entry of $A^T\mathbf{x}$ is given by

$$1 - \frac{j}{m_i} = c_{i,j},$$

so $\mathbf{x}$ is dual feasible. We then compute

$$\mathbf{b}^T\mathbf{x} = \log B(n-2-\varepsilon)x_0 - \sum_{i=1}^{n} x_i \log B$$

$$= \log B\left(n-2-\varepsilon - \sum_{i=1}^{n} \frac{1}{m_i}\right)$$

$$= \log B\left(-(2 - \sum_{i=1}^{n}(1 - \frac{1}{m_i})) - \varepsilon)\right)$$

$$= \log B(-\chi(\mathfrak{X}) - \varepsilon).$$

Therefore, $\mathbf{x} = \left(1, \frac{1}{m_1}, \ldots, \frac{1}{m_n}\right)$ is a dual feasible solution and

$$\mathbf{b}^T\mathbf{x} = \log B(-\chi(\mathfrak{X}) - \varepsilon).$$

By the weak duality theorem, we have that

$$\sum_{i=1}^{n} \frac{1}{m_i} \sum_{j=1}^{m_i-1} (m_i - j)y_{i,j} \geq \log B(-\chi(\mathfrak{X}) - \varepsilon).$$

Exponentiating gives

$$\prod_{i=1}^{n} \prod_{j=1}^{m_i-1} z_{i,j}^{\frac{m_i-j}{m_i}} \geq B^{(-\chi(\mathfrak{X})-\varepsilon)}.$$

As $B = \max\{|x|, |y|\}$, we have verified inequality (6.1) and consequently we have that conditional on the *abc*-conjecture that

$$\gamma(\mathfrak{X}) = \chi(\mathfrak{X})$$

when $\chi(\mathfrak{X}) \leq 0$.

6.2 Proof of Theorem 1.3

One direction of the theorem is provided by Theorem 2.6, which we proved in the previous subsection. It suffices to prove the converse.

Actually, for the converse, we only need the assertion that for any $\kappa > 0$ and $m \geq 4$, the function $H_{-K_{\mathfrak{X}_m}}(\mathbf{x}) \cdot H(\mathbf{x})^\kappa$ has Northcott's property, where $\mathfrak{X}_m = \mathfrak{X}(\mathbb{P}^1 : ((0, 1, \infty), (m, m, m)))$. To see this, let us fix $\varepsilon > 0$. Choose $0 < \kappa < \varepsilon/3$ and choose $m \in \mathbb{N}$ sufficiently large so that

$$\frac{m-3}{m-1} + \frac{\kappa m}{2(m-1)} > \frac{1}{1+\varepsilon}.$$

By hypothesis, we have

(6.19)
$$H_{-K_{\mathfrak{X}_m}}(\mathbf{x}) = \phi_m(x)^{1/m} \phi_m(y)^{1/m} \phi_m(x+y)^{1/m} \max\{|x|, |y|\}^{\frac{3}{m}-1} \gg_\kappa \max\{|x|, |y|\}^{1-\frac{3}{m}-\frac{\kappa}{2}}.$$

Trivially, we see that

$$\phi_m(u) \leq \mathrm{rad}(u)^{m-1}$$

for all $u \in \mathbb{Z}$. Hence, (6.19) implies

(6.20) $$\mathrm{rad}(x)^{\frac{m-1}{m}} \mathrm{rad}(y)^{\frac{m-1}{m}} \mathrm{rad}(x+y)^{\frac{m-1}{m}} \gg_\kappa \max\{|x|, |y|\}^{1-\frac{3}{m}+\frac{\kappa}{2}}.$$

Since $x, y, x+y$ are pairwise co-prime, we have $\mathrm{rad}(x)\,\mathrm{rad}(y)\,\mathrm{rad}(x+y) = \mathrm{rad}(xy(x+y))$; hence,

$$\mathrm{rad}(xy(x+y))^{\frac{m-1}{m}} \gg_\kappa \max\{|x|, |y|\}^{1-\frac{3}{m}+\frac{\kappa}{2}}.$$

Raising both sides to the $m/(m-1)$ power, we have

$$\mathrm{rad}(xy(x+y)) \gg_\kappa \max\{|x|, |y|\}^{\frac{m-3}{m-1}+\frac{\kappa m}{2(m-1)}} \gg_\kappa \max\{|x|, |y|\}^{\frac{1}{1+\varepsilon}},$$

by our hypotheses on m, κ. It follows that

$$\mathrm{rad}(xy(x+y))^{1+\varepsilon} \gg_\varepsilon \max\{|x|, |y|\},$$

which is plainly equivalent to the *abc*-conjecture, provided we adjust the implied constant.

7 Quantitative arithmetic of stacky curves

7.1 Crude bound for $N_{\mathbf{m}}(T)$, with $\mathbf{m} = (2, 2, m)$

Here, we deal with the case $\mathbf{m} = (2, 2, m)$. The Euler characteristic is equal to

$$\delta(\mathbf{m}) = 2 - \frac{1}{2} - \frac{1}{2} - 1 + \frac{1}{m} = \frac{1}{m}.$$

The height $H(x, y)$ is given by

$$H(x, y) = |x_1|^{1/2}|y_1|^{1/2}|z_1^{m-1}\cdots z_{m-1}|^{1/m} \max\{|x_1 x_2^2|, |y_1 y_2^2|\}^{1/m},$$

where $x = x_1 x_2^2$, $y = y_1 y_2^2$ and

$$(7.1) \qquad x_1 x_2^2 + y_1 y_2^2 = z_1 z_2^2 \cdots z_{m-1}^{m-1} z_m^m,$$

with $x_1, y_1, z_1, \ldots, z_{m-1}$ square-free. We normalize the height by raising it to the mth power, obtaining the bound

$$(7.2) \qquad |x_1 y_1|^{m/2}|z_1^{m-1}\cdots z_{m-1}| \max\{|x_1 x_2^2|, |y_1 y_2^2|\} \leq T.$$

From here, we see that

$$|z_1 z_2^2 \cdots z_m^m| \ll \max\{|x_1 x_2^2|, |y_1 y_2^2|\} \ll \frac{T}{|x_1 y_1|^{m/2}|z_1^{m-1}\cdots z_{m-1}|},$$

whence we conclude that

$$|z_m| \ll \frac{T^{1/m}}{|x_1 y_1|^{1/2}|z_1 \cdots z_{m-1}|}.$$

This bound and $|z_m| \geq 1$ implies that

$$|x_1 y_1|^{1/2}|z_1 \cdots z_{m-1}| \ll T^{1/m}.$$

From here, we obtain a crude upper bound for $N_{\mathbf{m}}(T)$, which proves Theorem 2.7. Indeed, having chosen $x_1, y_1, z_1, \ldots, z_{m-1}$, there are then $O(T^{1/m}/(|x_1 y_1|^{1/2} |z_1 \cdots z_{m-1}|))$ possibilities for z_m. Having chosen z_m as well, there are then $O_\varepsilon(T^\varepsilon)$ possibilities for x_2, y_2, since x_2, y_2 are polynomially bounded, so they will be determined by the norm-equation (7.1) up to a log factor. Thus, there are

$$(7.3) \qquad \sum_{|x_1 y_1|^{1/2}|z_1 \cdots z_{m-1}| \leq T^{1/m}} O_\varepsilon\left(\frac{T^{1/m+\varepsilon}}{|x_1 y_1|^{1/2}|z_1 \cdots z_{m-1}|}\right)$$

possible solutions to (7.1) satisfying the height bound (7.2). We evaluate this as

$$\sum_{|x_1 y_1|^{1/2}|z_1 \cdots z_{m-1}| \le T^{1/m}} O_\varepsilon \left(\frac{T^{1/m+\varepsilon}}{|x_1 y_1|^{1/2}|z_1 \cdots z_{m-1}|} \right)$$

$$= \sum_{|x_1 y_1| \le T^{2/m}} \frac{1}{|x_1 y_1|^{1/2}} \sum_{|z_1 \cdots z_{m-1}| \le T^{1/m}/|x_1 y_1|^{1/2}} O_\varepsilon \left(\frac{T^{1/m+\varepsilon}}{|z_1 \cdots z_{m-1}|} \right)$$

$$\ll_\varepsilon \sum_{|x_1 y_1| \le T^{2/m}} \frac{T^{1/m+\varepsilon}}{|x_1 y_1|^{1/2}} \ll_\varepsilon T^{2/m+\varepsilon}.$$

To give a lower bound, we choose square-free integers a, b, c so that the curve

$$ax^2 + by^2 = cz^m$$

has a primitive integral solution. Such a triple is guaranteed to exist (see [1]). Then we can parametrize (some) of the solutions by a triple of integral binary forms (F, G, h) where $\deg F = \deg G = m$ and $\deg h = 2$. By

$$x = F(u, v), y = G(u, v), z = h(u, v).$$

The height is

$$|a|^{m/2}|b|^{m/2}|c|^{m-1} \max\{|ax^2|, |by^2|\},$$

so if we treat a, b, c as constants, then

$$\max\{|x|, |y|\} \ll_{a,b,c} T^{1/2}.$$

Therefore, we are looking for solutions to the Thue inequality

$$\max\{|F(u, v)|, |G(u, v)|\} \ll_{a,b,c} T^{1/2}.$$

If we restrict u, v so that

$$\max\{|u|, |v|\} \ll_{a,b,c} T^{1/(2m)},$$

then we see that the above height bound is satisfied. Thus, $N_m(T) \gg T^{1/m}$.

7.2 Proof of Theorem 2.8

In this section, we prove Theorem 2.8. To do so, we will show that $N_2(T) = O\left(T^{1/2}(\log T)^3\right)$ and give a separate argument to show that $N_2(T) \gg T^{1/2}(\log T)^3$. The incompatibility of these two arguments represents the main obstacle as to why an asymptotic formula for $N_2(T)$ remains elusive.

We count rational points of bounded height on the curve $\mathfrak{X}(\mathbb{P}^1_{\mathbb{Q}}; (0, 2), (-1, 2), (\infty, 2))$ with the height on $\mathbb{P}^1$ given by (2.7). On writing

$$a = x_1 y_1^2, b = x_2 y_2^2, x_1, x_2 \text{ square-free}$$

(note that this differs from the notation used elsewhere in the paper), we then have

$$H(a, b) = |x_1 x_2| \operatorname{sqf}(x_1 y_1^2 + x_2 y_2^2) \max\{|x_1 y_1^2|, |x_2 y_2^2|\},$$

and the max on the right-hand side is dependent only on the relative size of $|a|, |b|$. If we write

$$(7.4) \qquad x_1 y_1^2 + x_2 y_2^2 = x_3 y_3^2,$$

then we further obtain the expression

$$H(a, b) = \max\{|x_2 x_3 (x_1 y_1)^2|, |x_1 x_3 (x_2 y_2)^2|\}.$$

We may assume without loss of generality that $|x_1 y_1^2| \geq |x_2 y_2|^2$ and $x_1 > 0$, so that

$$H(a, b) = |x_2 x_3 (x_1 y_1)^2|.$$

We consider the problem of counting integral points on the variety defined by (7.4), subject to the constraint

$$(7.5) \qquad 0 < |x_2 x_3 (x_1 y_1)^2| \leq T, |x_1 y_1^2| \geq |x_2 y_2^2|.$$

To obtain the upper bound, we must dissect (7.5) into suitable ranges. When $|x_1 x_2 x_3| \leq T^{1/2}$, we fix x_1, x_2, x_3 and treat (7.4) as a diagonal ternary quadratic form, say $Q_{\mathbf{x}}$. It is then the case that

$$(7.6) \qquad |y_i| \leq \frac{T}{|x_1 x_2 x_3| \cdot |x_i|}$$

for $i = 1, 2, 3$, and by Corollary 2 of [3], we then have the estimate

$$O\left(d(x_1 x_2 x_3)\left(\frac{T^{1/2}}{|x_1 x_2 x_3|} + O(1)\right)\right)$$

for the number of $\mathbf{y} \in \mathbb{Z}_{\neq 0}^3$ satisfying (7.5) and (7.4) provided that the quadratic form $Q_{\mathbf{x}}$ has a rational zero. Otherwise, it is clear that there will be no contribution. Thus, we must estimate

$$\sum_{\substack{1 \leq |x_1 x_2 x_3| \leq T^{1/2} \\ Q_{\mathbf{x}} \text{ has a rational zero}}} d(x_1 x_2 x_3).$$

This is similar to the work of Guo in [12], except he counted with respect to the height $\|\mathbf{x}\|_\infty$. Nevertheless, the techniques are similar, and again this may be of independent interest.

Next, we must deal with the case when $|x_1 x_2 x_3| \geq T^{1/2}$. For this, it suffices to observe from (7.6) that $|x_1 x_2 x_3| \geq T^{1/2}$ implies

$$|y_1 y_2 y_3| \leq \frac{T^{3/2}}{(x_1 x_2 x_3)^2} \leq T^{1/2}.$$

We then treat (7.4) as a linear form $L_{\mathbf{y}}$ in $\mathbf{x}$. We use this to show that the contribution for each $\mathbf{y}$ is $O\left(T^{1/2}|y_1 y_2 y_3|^{-1} + 1\right)$, which gives an acceptable contribution upon summing over $\mathbf{y}$.

For the lower bound, we first restrict $y_1, y_2, y_3 \in \mathbb{Z}_{\neq 0}$ satisfying

$$|y_1 y_2 y_3| \leq T^\delta$$

for some explicit $\delta > 0$ to be specified later. We note that to obtain the correct order of magnitude, it is permissible to choose any $\delta > 0$.

Having fixed $\mathbf{y} = (y_1, y_2, y_3)$, we consider the simultaneous conditions (7.4) and (7.5). This gives rise to a binary form inequality of the shape

$$(7.7) \qquad |x_1^2 x_2 (y_1^2 x_1 + y_2^2 x_2)| \leq T y_3^2 y_1^{-2}.$$

Because $|y_1 y_2 y_3|$ is small, we can count the number of solutions $\mathbf{x}$ to this inequality with reasonable precision. However, even with $|y_1 y_2 y_3|$ counting the number of solutions $\mathbf{x}$ with enough uniformity appears to still be a challenging task, because the binary form in (7.7) is singular. This difficulty is exacerbated by the fact that we will need to apply a square-free sieve eventually to produce triples $\mathbf{x}$ with each coordinate square-free.

To get around this issue, we simply count solutions to (7.7) with x_1, x_2 satisfying the inequalities

$$|x_i y_i^2| \leq c_i T^{1/4} |y_1 y_2 y_3|^{1/2}, i = 1, 2$$

for some positive numbers c_1, c_2. This has the effect that the long cusps inherent in (7.7) are removed, and reduce the problem to a more straightforward geometry of numbers question.

7.2.1 Upper bounds

To obtain upper bounds, it is crucial to view (7.4) as a plane in x_1, x_2, x_3 when $|y_1 y_2 y_3| \leq T^{1/2}$ and viewing (7.4) as a conic in y_1, y_2, y_3 when $|x_1 x_2 x_3| \leq T^{1/2}$. We call the former the *linear case* and the latter the *quadratic case*. We proceed to deal with the linear case below.

We shall first suppose that $|y_1 y_2 y_3| \leq T^{1/2}$ is fixed, and count the triples (x_1, x_2, x_3) and (y_1, y_2, y_3) for which (7.4) holds.

The key is the following lemma on counting points in sublattices of $\mathbb{Z}^2$.

Lemma 7.1 *Let $\Lambda \subset \mathbb{Z}^2$ be a lattice. Then, for all positive real numbers R_1, R_2, the number of primitive integral points $\mathbf{x} \in \Lambda$ satisfying $|x_i| \leq R_i, i = 1, 2$ is at most $O\left(R_1 R_2 / \det(\Lambda) + 1\right)$.*

Proof If the rectangle $[-R_1, R_2] \times [-R_2, R_2]$ contains at least two primitive vectors in Λ, say $\mathbf{x}_1, \mathbf{x}_2$, then since this rectangle is convex it contains the parallelogram with end points $\pm\mathbf{x}_1, \pm\mathbf{x}_2$. The area of this parallelogram is at least as large as $\det\Lambda$, since the lattice spanned by $\mathbf{x}_1, \mathbf{x}_2$ is a sublattice of Λ. It thus follows that

$$R_1 R_2 \gg \det\Lambda.$$

Otherwise, the rectangle $[-R_1, R_1] \times [-R_2, R_2]$ contains at most one primitive vector in Λ. This completes the proof. ∎

The strength of this lemma is that it gives a strong upper bound even in lopsided boxes.

Given (7.4), it follows that there is at least one $i \in \{2, 3\}$ such that

$$|x_i y_i^2|/2 \leq x_1 y_1^2 \leq 2|x_i y_i^2|,$$

whence

$$\frac{x_1 y_1^2}{2 y_i^{-2}} \le |x_i| \le \frac{2 x_1 y_1^2}{y_i^2}.$$

Without loss of generality, we assume that this holds for $i = 2$. Suppose that $M_1 \le x_1 < 2M_1$. By (7.5), we have

$$|x_3| \le \frac{T}{|x_2 x_1^2 y_1^2|},$$

whence

$$|x_3| \le T \cdot \frac{2 y_2^2}{(x_1 y_1^2)(x_1^2 y_1^2)}$$
$$\le \frac{2 y_2^2 T}{M_1^3 y_1^4}.$$

Applying Lemma 7.1 to the lattice defined by the congruence $y_1^2 x_1 - y_3^2 x_3 \equiv 0 (\mathrm{mod}\ y_2^2)$ which has determinant equal to y_2^2, there are

$$O\left(M_1 \cdot \frac{T y_2^2}{M_1^3 y_1^4} \cdot \frac{1}{y_2^2} + 1\right) = O\left(\frac{T}{M_1^2 y_1^4} + 1\right)$$

possibilities for x_1, x_3, which then determines $x_2 = (y_1^2 x_1 - y_3^2 x_3)/y_2^2$. Similarly, applying Lemma 7.1 to the lattice defined by $y_1^2 x_1 + y_2^2 x_2 \equiv 0 (\mathrm{mod}\ y_3^2)$, with determinant equal to y_3^2, gives the estimate

$$O\left(M_1 \cdot \frac{y_1^2 M_1}{y_2^2} \frac{1}{y_3^2} + 1\right) = O\left(\frac{M_1^2 y_1^2}{y_2^2 y_3^2} + 1\right)$$

for the number of x_1, x_2, which then also determine x_3. The two bounds coincide when

$$M_1 = \frac{T^{1/4} |y_2 y_3|^{1/2}}{|y_1|^{3/2}},$$

and we get the bound

$$O\left(\frac{T^{1/2} |y_2 y_3| y_1^2}{y_2^2 y_3^2 |y_1|^3} + 1\right) = O\left(\frac{T^{1/2}}{|y_1 y_2 y_3|} + 1\right)$$

for the number of x_1, x_2, x_3 given y_1, y_2, y_3. Thus, we obtain an acceptable estimate whenever $|y_1 y_2 y_3| \ll T^{1/2}$, since

$$\sum_{1 \le |y_1 y_2 y_3| \le T^{1/2}} \frac{T^{1/2}}{|y_1 y_2 y_3|} + 1 \ll T^{1/2} \sum_{n \le T^{1/2}} \frac{d_3(n)}{n} + \sum_{n \le T^{1/2}} d_3(n)$$

It is well known that

$$\sum_{n \le Z} d_3(n) = Z (\log Z)^2 + O(Z \log Z).$$

By partial summation, we have

$$\sum_{n \leq Z} \frac{d_3(n)}{n} = Z^{-1} \sum_{n \leq Z} d_3(n) + \int_1^Z \left(\sum_{n \leq t} d_3(n) \right) \frac{dt}{t^2}$$
$$\ll (\log Z)^2 + \int_1^Z \frac{(\log t)^2 dt}{t}$$
$$\ll (\log Z)^3.$$

It follows that

$$T^{1/2} \sum_{n \leq T^{1/2}} \frac{d_3(n)}{n} + \sum_{n \leq T^{1/2}} d_3(n) \ll T^{1/2} (\log T)^3.$$

It remains to deal with the case when $|y_1 y_2 y_3| \gg T^{1/2}$, where we instead fiber over $\mathbf{x}$ and consider zeros of the corresponding diagonal quadratic forms $Q_{\mathbf{x}}$. Since

$$|x_i y_i^2| \ll x_1 y_1^2$$

for $i = 1, 2$ by assumption, it follows that

$$|x_1 x_2 x_3 y_1^2 y_2^2 y_3^2| \leq x_1^3 y_1^6;$$

hence,

$$|y_1^2 y_2^2 y_3^2| \ll \frac{x_1^3 y_1^6}{x_1 |x_2 x_3|}.$$

If $|x_1 x_2 x_3| \gg T^{1/2}$, then

$$x_1^3 y_1^6 \gg T^{3/2} \Leftrightarrow x_1 y_1^2 \gg T^{1/2}.$$

This implies that

$$|x_1 x_2 x_3| \cdot x_1 y_1^2 \gg T,$$

which violates (7.5) if the implied constants are sufficiently large. It thus follows that we must have $|x_1 x_2 x_3| \ll T^{1/2}$ in this case.

We now fix x_1, x_2, x_3 and consider (7.4) as a ternary quadratic form in y_1, y_2, y_3. We shall require the following version of Corollary 2 in [3], which is an analogue of Lemma 7.1.

Lemma 7.2 *Let x_1, x_2, x_3 be pairwise co-prime square-free integers. Let R_1, R_2, R_3 be positive real numbers. Then the number of primitive solutions y_1, y_2, y_3 to the equation*

$$x_1 y_1^2 + x_2 y_2^2 = x_3 y_3^2$$

with $|y_i| \leq R_i$ is bounded by

$$O\left(d(x_1 x_2 x_3) \left(\left(\frac{R_1 R_2 R_3}{|x_1 x_2 x_3|} \right)^{1/3} + 1 \right) \right).$$

Since $|x_i y_i^2| \ll x_1 y_1^2$ for $i = 1, 2$, it follows that

$$|x_1 x_2 x_3 (x_i y_i^2)| \ll |x_1 x_2 x_3 (x_1 y_1^2)| \leq T$$

for $i = 1, 2$, whence

$$|(x_1 y_1)^2 x_2 x_3|, |(x_2 y_2)^2 x_1 x_3|, |(x_3 y_3)^2 x_1 x_2| \ll T.$$

This implies that

$$(y_1 y_2 y_3)^2 (x_1 x_2 x_3)^4 \ll T^3;$$

hence

$$|y_1 y_2 y_3| \ll \frac{T^{3/2}}{(x_1 x_2 x_3)^2}.$$

Lemma 7.2 then implies that for fixed x_1, x_2, x_3, the number of primitive $\mathbf{y} = (y_1, y_2, y_3)$ satisfying (7.4) is

$$O\left(d(x_1 x_2 x_3) \left(\frac{T^{1/2}}{|x_1 x_2 x_3|} + 1 \right) \right).$$

We now sum over primitive $\mathbf{x} \in \mathbb{Z}^3$ satisfying $|x_1 x_2 x_3| \ll T^{1/2}$, with the property that the quadratic form $Q_\mathbf{x}$ given by (7.4) has a rational zero. By the Hasse–Minkowski theorem, this is tantamount to the form $Q_\mathbf{x}(\mathbf{y}) = x_1 y_1^2 + x_2 y_2^2 - x_3 y_3^2$ being everywhere locally soluble. The estimation of this is interesting on its own right and will be handled in a separate subsection.

7.2.2 Counting soluble ternary quadratic forms

In this section, we consider the set

$$\mathcal{S} = \{(x_1, x_2, x_3) \in \mathbb{Z}^3 : x_1, x_2, x_3 > 0, \gcd(x_1, x_2) = \gcd(x_1, x_3) = \gcd(x_2, x_3) = 1,$$
$$x_i \text{ square-free for } i = 1, 2, 3, x_1 y_1^2 + x_2 y_2^2 - x_3 y_3^2 \text{ is everywhere locally soluble}\}.$$

By a well-known theorem of Legendre (see [12]), the indicator function for $\mathcal{S}$ is given by

(7.8)
$$f_\mathcal{S}(x_1, x_2, x_3) = \left(2^{-\omega(x_1)} \sum_{a_1 | x_1} \left(\frac{x_2 x_3}{a_1} \right) \right) \left(2^{-\omega(x_2)} \sum_{a_2 | x_2} \left(\frac{x_1 x_3}{a_2} \right) \right) \left(2^{-\omega(x_3)} \sum_{a_3 | x_3} \left(\frac{-x_1 x_2}{a_3} \right) \right).$$

We will now combine the ideas given in [12] and those in [8].

Put

$$\mathcal{S}(X) = \sum_{1 \le x_1 x_2 x_3 \le X} \sum_{(x_1, x_2, x_3) \in \mathcal{S}} \frac{d(x_1 x_2 x_3)}{x_1 x_2 x_3}$$

$$= \sum_{1 \le |x_1 x_2 x_3| \le X} \frac{d(x_1 x_2 x_3)}{x_1 x_2 x_3} f_\mathcal{S}(x_1, x_2, x_3).$$

Since x_1, x_2, x_3 are pairwise coprime and square-free, it follows that

$$d(x_1 x_2 x_3) = 2^{\omega(x_1 x_2 x_3)} = 2^{\omega(x_1)} \cdot 2^{\omega(x_2)} \cdot 2^{\omega(x_3)},$$

where $\omega(n)$ is the number of distinct prime factors of n. It follows that

$$(7.9) \quad \mathcal{S}(X) = \sum_{1 \le x_1 x_2 x_3 \le X} \frac{2^{\omega(x_1 x_2 x_3)}}{x_1 x_2 x_3} f_{\mathcal{S}}(x_1, x_2, x_3)$$

$$= \sum_{1 \le x_1 x_2 x_3 \le X} \frac{1}{x_1 x_2 x_3} \left(1 + \left(\frac{x_2 x_3}{x_1} \right) \left(\frac{x_1 x_3}{x_2} \right) \left(\frac{-x_1 x_2}{x_3} \right) + \sum_g g(x_1, x_2, x_3) \right),$$

where g expresses a product of Jacobi symbols. The sum

$$(7.10) \quad \mathcal{S}_1(X) = \sum_{1 \le |x_1 x_2 x_3| \le X} \frac{1}{x_1 x_2 x_3} \left(1 + \left(\frac{x_2 x_3}{x_1} \right) \left(\frac{x_1 x_3}{x_2} \right) \left(\frac{-x_1 x_2}{x_3} \right) \right)$$

is expected to contribute the main term, while the sum

$$(7.11) \quad \mathcal{S}_2(x) = \sum_{1 \le x_1 x_2 x_3 \le X} \frac{1}{x_1 x_2 x_3} \sum_g g(x_1, x_2, x_3)$$

is expected to be negligible, due to the cancellation of characters.

By partial summation, we obtain

$$(7.12) \quad \mathcal{S}_i(X) = \frac{1}{X} \Sigma_i(X) + \int_1^X \Sigma_i(t) \frac{t}{t^2},$$

where

$$\Sigma_1(X) = \sum_{\substack{1 \le |x_1 x_2 x_3| \le X \\ x_1 x_2 x_3 \text{ square-free} \\ Q_{(x_1, x_2, x_3)} \text{ is soluble}}} \left(1 + \left(\frac{x_2 x_3}{x_1} \right) \left(\frac{x_1 x_3}{x_2} \right) \left(\frac{x_1 x_2}{x_3} \right) \right)$$

and

$$\Sigma_2(X) = \sum_{1 \le x_1 x_2 x_3 \le X} \sum_g g(x_1, x_2, x_3).$$

Our situation differs from that of Guo in [12] since we are counting over triples with $|x_1 x_2 x_3| \le X$ rather than $\max\{|x_1|, |x_2|, |x_3|\} \le X$, which introduces some difficulties. However, this is exactly analogous to the situation encountered by Fouvry and Kluners in [8].

Our key proposition will be the following.

Proposition 7.3 *We have the asymptotic upper bound*

$$\mathcal{S}(X) = O\left((\log X)^3 \right).$$

In fact, we can refine Proposition 7.3 to give an asymptotic formula, but this is unnecessary for our purposes.

We proceed to prove Proposition 7.3 in the remainder of the section. We begin by showing that triples (x_1, x_2, x_3) with $\mu^2(x_1 x_2 x_3) = 1$ and $\omega(x_1 x_2 x_3)$ large contribute negligibly. To wit, put

$$\mathcal{S}_2^{(r)}(X) = \sum_{\substack{1 \le x_1 x_2 x_3 \le X \\ \omega(x_1 x_2 x_3) = r}} \frac{1}{x_1 x_2 x_3} \sum_g g(x_1, x_2, x_3).$$

By the triangle inequality, it is clear that

$$\left| \mathcal{S}_2^{(r)}(X) \right| \ll \sum_{\substack{n \leq X \\ \mu^2(n)=1, \omega(n)=r}} \frac{d_3(n)}{n}.$$

By partial summation, we have

$$\sum_{\substack{n \leq X \\ \mu^2(n)=1, \omega(n)=r}} \frac{d_3(n)}{n} = X^{-1} \sum_{\substack{n \leq X \\ \mu^2(n)=1, \omega(n)=r}} d_3(n) + \int_1^X \left(\sum_{\substack{n \leq t \\ \mu^2(n)=1, \omega(n)=r}} d_3(n) \right) \frac{dt}{t^2}.$$

To estimate the latter sum, we will need the following result, which is Lemma 11 in [8].

Lemma 7.4 *There exists an absolute constant $B_0 \geq 1$ such that for every $r \geq 0$, we have*

$$\left| \{ n \leq X : \omega(n) = r, \mu^2(n) = 1 \} \right| \leq B_0 \cdot \frac{X}{\log X} \cdot \frac{(\log\log X + B_0)^r}{r!}.$$

Applying the lemma, we have for $\Omega = 30(\log\log X + B_0)$

$$\sum_{\substack{n \leq X \\ \mu^2(n)=1, \omega(n) \geq \Omega}} d_3(n) \ll \frac{X}{\log X} \sum_{r \geq \Omega} 3^r \cdot \frac{(\log\log X + B_0)^r}{r!}$$

$$\ll \frac{X}{\log X} \sum_{r \geq \Omega} \left(\frac{3e(\log\log X + B_0)}{r} \right)^r$$

$$\ll \frac{X}{\log X} \sum_{r \geq \Omega} \left(\frac{3e}{10} \right)^r,$$

the final sum a convergent geometric series. Hence,

$$\sum_{r \geq \Omega} \left(\frac{3e}{10} \right)^r \ll \left(\frac{3e}{10} \right)^\Omega \ll \frac{1}{\log X}.$$

We thus conclude that

$$(7.13) \qquad \sum_{r \geq \Omega} \left| \mathcal{S}_2^{(r)}(X) \right| \ll 1 + (\log X)^{-2} + \int_1^X \frac{dt}{t(\log t)^2}$$

$$= O(1)$$

and is thus negligible.

Note that $x_1, x_2, -x_3$ cannot all be the same sign; otherwise, (7.4) will only have a trivial real solution. Hence, the signs of (x_1, x_2, x_3) must be $(+, +, +)$, or $(+, -, +)$, since we assumed $x_1 > 0$ and $x_1 y_1^2 \geq |x_2 y_2^2|$. By rearranging, we must thus assume $x_1, x_2, x_3 > 0$.

We then expand (7.11) by writing $x_i = x_{i1} x_{i2}$ for $i = 1, 2, 3$, and

$$\sum_{\substack{1 \leq x_1 x_2 x_3 \leq X \\ \mu^2(x_1 x_2 x_3)=1}} \sum_g g(x_1, x_2, x_3)$$

$$= \sum_{\substack{(x_{11}x_{12})(x_{21}x_{22})(x_{31}x_{32}) \leq X \\ 1 < x_{i1} < x_i \text{ for } 1 \leq i \leq 3}} \left(\frac{x_{21}x_{22}x_{31}x_{32}}{x_{11}} \right) \left(\frac{x_{11}x_{12}x_{31}x_{32}}{x_{21}} \right) \left(\frac{x_{11}x_{12}x_{21}x_{22}}{x_{31}} \right).$$

We now follow the strategy outlined in [8] and break up the set

$$\{(x_{11}, x_{12}, x_{21}, x_{22}, x_{31}, x_{32}) \in \mathbb{N}^6 : x_{11}x_{12}x_{21}x_{22}x_{31}x_{32} \le X\}$$

by restricting the x_{ij}'s to intervals of the form

$$[A_{ij}, \Delta A_{ij}),$$

where

$$\Delta = 1 + (\log X)^{-3}.$$

For a given $\mathbf{A} = (A_{11}, A_{12}, A_{21}, A_{22}, A_{31}, A_{32})$, put

$$S_2(X; \mathbf{A}) = \sum_{\substack{x_{ij} \in [A_{ij}, \Delta A_{ij}) \\ \mu^2(x_{11}x_{12}x_{21}x_{22}x_{31}x_{32})=1 \\ \prod_{i,j} x_{ij} \le X}} \left(\frac{x_{21}x_{22}x_{31}x_{32}}{x_{11}}\right)\left(\frac{x_{11}x_{12}x_{31}x_{32}}{x_{21}}\right)\left(\frac{x_{11}x_{12}x_{21}x_{22}}{x_{31}}\right).$$

We then have the following lemma.

Lemma 7.5 *We have the bound*

$$\sum_{\prod A_{ij} \ge \Delta^{-6}X} |S_2(X; \mathbf{A})| = O\left(X(\log X)^{-1}\right).$$

Proof We have

$$\sum_{\prod A_{ij} \ge \Delta^{-6}X} |S_2(X; \mathbf{A})| \le \sum_{\substack{\Delta^{-6}X \le n \le X \\ \mu^2(n)=1}} d_3(n)$$

$$\ll \sum_{\Delta^{-6}X \le n \le X} 3^{\omega(n)}$$

$$\ll (1 - \Delta^{-6})X(\log X)^2.$$

By Taylor's theorem, we have

$$\Delta^{-6} = \left(1 + (\log X)^{-3}\right)^{-6} = 1 - 6(\log X)^{-3} + O\left((\log X)^{-6}\right).$$

The proof then follows. ∎

To proceed, we shall require the following well-known lemma regarding character sums.

Lemma 7.6 (Double Oscillation Lemma) *Let $\{\alpha_n\}, \{\beta_m\}$ be two sequences of complex numbers with each term having absolute value bounded by 1. Let M, N be positive real numbers. Then we have*

$$\sum_{m \le M} \sum_{n \le N} \alpha_m \beta_n \mu^2(2m)\mu^2(2n) \left(\frac{m}{n}\right)$$

$$\ll \min\left\{\left(M^{-1/2} + (N/M)^{-1/2}\right), \left(N^{-1/2} + (M/N)^{-1/2}\right)\right\}$$

and for every $\varepsilon > 0$,

$$\sum_{m \le M} \sum_{n \le N} \alpha_m \beta_n \mu^2(2m)\mu^2(2n) \ll_\varepsilon MN \left(M^{-1/2} + N^{-1/2}\right)(MN)^\varepsilon.$$

We will also need the following variant of the Siegel–Walfisz theorem.

Lemma 7.7 *Let χ_q be a primitive character modulo $q \ge 2$. Then, for every $A > 1$, we have*

$$\sum_{Y \le p \le X} \chi_q(p) = O_A\left(\sqrt{q} \cdot X(\log X)^{-A}\right)$$

uniformly for $X \ge Y \ge 2$.

We now consider, as in [8], the quantities

(7.14) $$X^\dagger = (\log X)^9, \quad X^\ddagger = \exp\left((\log X)^{1/8}\right).$$

We now consider those $\mathbf{A}$ with the property that at most two entries larger than $X^\ddagger$. We dissect the sum according to the number $r \le 2$ of terms A_{ij} greater than $X^\ddagger$. Let n be the product of those x_{ij} which are larger than $X^\ddagger$, and m the product of the remaining ones. We sum over $\mathbf{A}$ with such properties to obtain

$$\sum_{\mathbf{A}}^{(2)} |\mathcal{S}_2(X; \mathbf{A})| \le \sum_{r \le 2} \sum_{m \le (X^\ddagger)^{6-r}} \mu^2(m) d_{6-r}(m) \sum_{n \le X/m} \mu^2(n) d_r(n)$$

$$\ll \sum_{r \le 2} \sum_{m \le (X^\ddagger)^{6-r}} \mu^2(m) d_{6-r}(m) \left(\frac{X}{m}\right)(\log X)^{r-1}$$

$$\ll X \left(\sum_{r \le 2}(\log X)^{r-1}\right)\left(\sum_{m \le (X^\ddagger)^6} \frac{d_6(m)}{m}\right)$$

$$\ll X(\log X)\left(\log \exp\left((\log X)^{1/8}\right)\right)^7$$

$$\ll X(\log X)^{15/8}.$$

This is sufficiently small for our purposes.

We may now assume that $A_{ij} \ge X^\ddagger$ for at least three pairs i, j with $1 \le i \le 3, 1 \le j \le 2$. We now suppose that there exist $a \ne b$ such that

$$A_{a,2}, A_{b,1} \ge X^\dagger.$$

The sum over $\mathbf{A}$ satisfying these properties can be bounded by

$$\sum_{\mathbf{A}} |\mathcal{S}_2(X; \mathbf{A})| \le \sum_{x_{ij}, (i,j) \ne (a,2),(b,1)} \prod_{(i,j) \ne (a,2),(b,1)} \left| \sum_{x_{a,2}} \sum_{x_{b,1}} \alpha_{(a,2)} \beta_{(b,1)} \left(\frac{x_{a,2}}{x_{b,1}}\right)\right|,$$

where α, β have modulus at most one. Lemma 7.6 then applies, and since our variables $x_{a,2}, x_{b,1}$ range over intervals exceeding $X^\dagger$ in length, it follows that

$$|\mathcal{S}_2(X; \mathbf{A})| \ll \left(\prod_{(i,j) \ne (a,2),(b,1)} A_{ij}\left(A_{a,2}A_{b,1}\left(A_{a,2}^{-1/3} + A_{b,1}^{-1/3}\right)\right)\right) \ll X(X^\dagger)^{-1/3} = O\left(X(\log X)^{-3}\right),$$

which is again enough.

Next, consider the family where the two previous conditions do not hold, and in addition there exist $a \neq b$ such that $2 \leq A_{b,1} \leq X^\dagger$ and $A_{a,2} > X^\ddagger$. Under these conditions, we see that

$$|\mathcal{S}_2(X; \mathbf{A})| \ll \sum_{x_{ij}, (i,j) \neq (a,2), (b,1)} \sum_{x_{a,2}} \left| \sum_{x_{b,1}} \mu^2 \left(\prod_{(i,j) \neq (a,2), (b,1)} x_{ij} \right) \left(\frac{x_{a,2}}{x_{b,1}} \right) \right|,$$

where $A_{ij} \leq x_{ij} \leq \Delta A_{ij}$ and $\omega(x_{ij}) \leq \Omega$ for $1 \leq i \leq 3, 1 \leq j \leq 2$. Now put $\ell = \omega(x_{a,2})$, writing

$$x_{a,2} = p_1 \cdots p_\ell$$

with $p_1 < p_2 < \cdots < p_\ell$, we obtain

$$|\mathcal{S}_2(X; \mathbf{A})| \ll \sum_{\substack{x_{ij} \\ (i,j) \neq (a,2), (b,1)}} \sum_{x_{b,1}} \sum_{0 \leq \ell \leq \Omega} \left| \sum_{\omega(x_{a,2}) = \ell} \mu^2 \left(\prod_{i,j} x_{ij} \right) \left(\frac{x_{a,2}}{x_{b,1}} \right) \right|,$$

the inner sum being bounded by

$$\sum_{p_1 \cdots p_{\ell-1}} \left| \sum_{p_\ell} \left(\frac{p_\ell}{x_{b,1}} \right) \right|,$$

and $p_1, \ldots, p_\ell$ satisfy $A_{a,2} \leq p_1 \cdots p_\ell \leq \Delta A_{a,2}$. Note that

$$p_\ell \geq A_{a,2}^{1/\ell} \geq \exp\left((\log X)^{1/9}\right).$$

We may now apply Lemma 7.7 to obtain the bound

$$\left| \sum_{p_\ell} \left(\frac{p_\ell}{x_{b,1}} \right) \right| \ll_A A_{b,1}^{1/2} \frac{A_{a,2}}{p_1 \cdots p_{\ell-1}} (\log X)^{-A/9} + \Omega,$$

with A arbitrarily large. Note that $p_1 \cdots p_{\ell-1} \leq X$, and hence

$$\sum_{p_1 \cdots p_{\ell-1} \leq X} (p_1 \cdots p_{\ell-1})^{-1} \ll \sum_{n \leq X} \frac{1}{n} \ll \log X.$$

Hence,

$$\sum_{\mathbf{A}}^{(3)} |\mathcal{S}_2(X; \mathbf{A})| \ll A_{b,1}^{1/2} \prod_{i,j} A_{ij} (\log X)^{-A/9+1} \ll X (\log X)^{-A/9+11/2}.$$

Choosing A large shows that this contribution is negligible.

The remaining case can be summarized by the following properties:

(1) $\prod_{i,j} A_{ij} \leq \Delta^{-6} X$.
(2) $A_{ij} \geq X^\ddagger$ for at least three pairs of indices (i,j).
(3) If $A_{ij}, A_{k\ell} \geq X^\dagger$, then $j = \ell$.
(4) If $A_{ij} \leq A_{k\ell}$ with $j \neq \ell$, then either $A_{ij} = 1$ or $2 \leq A_{ij} \leq X^\dagger$ and $A_{k\ell} < X^\ddagger$.

We now show that the second option in (4) cannot happen. This will imply that we have accounted for all possibilities for (7.11), and hence reduced our problem to estimating $\mathcal{S}_1(X)$.

Suppose, without loss of generality, that $2 \leq A_{11} \leq X^\dagger$ and $A_{22} < X^\ddagger$. Since $A_{ij} \geq X^\ddagger$ for at least three pairs of indices (i, j), one of A_{12} or A_{32} must exceed $X^\ddagger$. We then have $A_{11} \leq X^\dagger$ and A_{32}, say, exceeds $X^\ddagger$, which means that our earlier estimation covers this case.

The upshot now is that

(7.15) $$\Sigma_2(X) \ll_A X(\log X)^{15/8}$$

for some $\kappa(A) > 0$. It follows from (7.12) that

$$\begin{aligned}
S_2(X) &= X^{-1}\Sigma_2(X) + \int_1^X \Sigma_2(t)\frac{dt}{t^2} \\
&\ll (\log X)^{15/8} + \int_1^X \frac{(\log t)^{15/8}\,dt}{t} \\
&= (\log X)^{23/8},
\end{aligned}$$

which is sufficiently small for our purposes.

Finally, we may evaluate the main term, which is given by (7.10). By the triangle inequality, we have

$$S_1(X) \ll \sum_{x_1 x_2 x_3 \leq X} \frac{1}{x_1 x_2 x_3} = \sum_{n \leq X} \frac{d_3(n)}{n},$$

which is $O((\log X)^3)$. This completes the proof of the proposition.

7.2.3 Lower bounds

For the lower bound, it suffices to give an accurate count for some subset of the points enumerated by the quantity $N_2(T)$. The arguments used here are inspired by the work of the second author and C.L. Stewart in [19], though the situation here is slightly simpler. To wit, we shall consider the subset of points $(\mathbf{x}, \mathbf{y})$ satisfying the condition

(7.16) $$1 \leq |y_1 y_2 y_3| \leq T^\delta,$$

where δ is some explicit positive number which we shall specify later. Next, we suppose that x_1, x_2 satisfy

(7.17) $$|x_i y_i^2| \leq \frac{T^{1/4}}{2}|y_1 y_2 y_3|^{1/2}, \quad i = 1, 2$$

Note that

$$|x_3 y_3^2| = |x_1 y_1^2 + x_2 y_2^2| \leq |x_1 y_1^2| + |x_2 y_2^2| \leq \left(\frac{1}{2} + \frac{1}{2}\right) T^{1/4}|y_1 y_2 y_3|^{1/2},$$

whence

$$|x_1 x_2 x_3|(y_1 y_2 y_3)^2 = |x_1 y_1^2||x_2 y_2^2||x_3 y_3^2| \leq \frac{T^{3/4}}{4}|y_1 y_2 y_3|^{3/2}.$$

Thus,

$$|(x_1 y_1^2)x_1 x_2 x_3| \leq \left(\frac{T^{1/4}}{2}|y_1 y_2 y_3|^{1/2}\right)\left(\frac{T^{3/4}}{4}|y_1 y_2 y_3|^{-1/2}\right) < T.$$

Therefore, every pair (x_1, x_2) satisfying (7.17) with x_1, x_2 both square-free and $x_3 = (y_1^2 x_1 + y_2^2 x_2) y_3^{-2} \in \mathbb{Z}$ square-free will contribute to $N_2(T)$.

We now count pairs (x_1, x_2) such that:

(1) (x_1, x_2) satisfies (7.17);
(2) $\gcd(x_1, x_2) = 1$;
(3) x_1, x_2 are square-free; and
(4) $y_1^2 x_1 + y_2^2 x_2 \equiv 0 \pmod{y_3^2}$, $(y_1^2 x_1 + y_2^2 x_2) y_3^{-2}$ is square-free.

For each prime p, we interpret conditions (2)–(4) modulo p^2. Condition (2) is the assertion that $p | x_1 \Rightarrow p \nmid x_2$, Condition (3) is the assertion that for all primes p we have $p^2 \nmid x_1, x_2$, and condition (4) is stating $y_3^2 | y_1^2 x_1 + y_2^2 x_2$, and if $p^s \| y_3$, then $p^{2s+2} \nmid y_1^2 x_1 + y_2^2 x_2$. Let

$$\rho_{\mathbf{y}}(m) = \#\{(x_1, x_2) \pmod{m} : (2) \text{ to } (4) \text{ holds for all } p | m\}.$$

It is apparent that $\rho_{\mathbf{y}}(\cdot)$ is multiplicative. Put

$$N^*(\mathbf{y}; T) = \#\{(x_1, x_2) \in \mathbb{Z}^2 : (1) \text{ to } (4) \text{ hold}\}$$

and

$$N_b^*(\mathbf{y}; T) = \#\{(x_1, x_2) \in \mathbb{Z}^2 : (7.17) \text{ holds}, (2) \text{ to } (4) \text{ holds mod } b\}.$$

By standard arguments using the inclusion–exclusion sieve, we have

$$N^*(\mathbf{y}; T) = \prod_{p \leq Y} \left(1 - \frac{\rho_{\mathbf{y}}(p^{2k})}{p^{2k}}\right) \frac{T^{1/2}}{|y_1 y_2 y_3|} + O\left(\sum_{Y < p < T^{1/8} |y_1 y_2 y_3|^{1/4} \max\{|y_1|^{-1}, |y_2|^{-1}\}} \left(\frac{T^{1/2}}{p^2 |y_1 y_2 y_3|} + 1\right)\right),$$

the error term being bounded by

$$O\left(\frac{T^{1/2}}{Y |y_1 y_2 y_3|} + \frac{T^{1/8} |y_1 y_2 y_3|^{1/2}}{\min\{|y_1|, |y_2|\}}\right).$$

Since $|y_1 y_2 y_3| \leq T^\delta$, we obtain an acceptable error term provided that $\delta < 1/4$. This shows that

$$N(T) \gg \sum_{1 \leq |y_1 y_2 y_3| \leq T^\delta} N^*(\mathbf{y}; T) \gg \sum_{1 \leq |y_1 y_2 y_3| \leq T^\delta} \frac{T^{1/2}}{|y_1 y_2 y_3|}.$$

Since

$$\sum_{1 \leq |y_1 y_2 y_3| \leq Z} |y_1 y_2 y_3|^{-1} \gg \sum_{n \leq Z} d_3(n) n^{-1} \gg (\log Z)^3,$$

this confirms the lower bound.

7.3 Counting points with respect to the canonical height when $\chi(\mathfrak{X}) < 0$

In this section, we first prove that the number of *quadratic points* on a hyperelliptic curve given by the model

$$C_F : z^2 = F(x, y)$$

where F is an integral, non-singular binary form having degree $2g + 2$ with $g \geq 2$, is dominated by the "obvious" points given by triples $(x, y, \sqrt{F(x,y)})$. To show that the *proper* quadratic points are negligible, we note that when $g = 2$ the proper quadratic points, which come in conjugate pairs, are in bijection with the rational points of the Jacobian $\mathrm{Jac}(C_F)$ via the correspondence $[P] \mapsto [P_1 + P_2] - K_{C_F}$, where K_{C_F} is the canonical divisor. Thus, in this case, the proper quadratic points of bounded height are given by the rational points of bounded height in $\mathrm{Jac}(C_F)(\mathbb{Q})$, for which there are $O_F((\log T)^{r_F})$ many, where r_F is the Mordell–Weil rank of $\mathrm{Jac}(C_F)$. For $g \geq 3$, the proper quadratic points are finite by Faltings' theorem. Thus, the number of quadratic points on C_F is asymptotically equal to the number of rational points in $\mathbb{P}^1_{\mathbb{Q}}$ of bounded height.

To the contrary, for $\mathfrak{X} = \mathfrak{X}(\mathbb{P}^1 : (\mathbf{a}, \mathbf{m}))$ with $\chi(\mathfrak{X}) < 0$, we get a much less reasonable result. This is because we have little control over the set of integers x, y such that $\ell_i(x, y)$ is divisible by a large square for $i = 1, \ldots, n$. Even with the *abc*-conjecture, there is only so much that can be shown. In the case when $\mathbf{m} = (2, \ldots, 2)$, we have the following.

Theorem 7.8   *Let $\mathfrak{X} = \mathfrak{X}(\mathbb{P}^1 : (\mathbf{a}, \mathbf{m}))$ be a stacky curve with $\mathbf{m} = \underbrace{(2, \ldots, 2)}_{n}$ with $n \geq 5$. Let $N_{\mathbf{a},n}(T)$ be the number of rational points on $\mathfrak{X}$ satisfying $H_{(\mathbf{a},\mathbf{m})}(x, y) \leq T$. Assume that the abc-conjecture holds. Then, for any $\varepsilon > 0$, we have*

$$N_{(\mathbf{a},\mathbf{m})}(T) \ll_\varepsilon T^{\frac{1}{n-3}+\varepsilon}.$$

Proof   This is similar to the proof of Theorem 2.6. We conclude from that proof that

$$\prod_{i=1}^{n} |x_i y_i| \geq \mathrm{rad}\left(\prod_{i=1}^{n} \ell_i(x, y)\right) \gg_\varepsilon \max\{|x|, |y|\}^{n-2-\varepsilon},$$

and

$$\prod_{i=1}^{n} |x_i y_i^2| \ll \max\{|x|, |y|\}^n$$

by the triangle inequality. Comparing, we conclude that

$$\prod_{i=1}^{n} |y_i| \ll_\varepsilon \max\{|x|, |y|\}^{2+\varepsilon}$$

and in turn

$$\prod_{i=1}^{n} |x_i| \gg_\varepsilon \max\{|x|, |y|\}^{n-4-\varepsilon}.$$

It follows that

$$H_{(\mathbf{a},\mathbf{m})}(x, y) = \max\{|x|, |y|\}^{n-4} \prod_{i=1}^{n} |x_i| \gg_\varepsilon \max\{|x|, |y|\}^{2n-8-\varepsilon}.$$

Hence, $N_{(\mathbf{a},\mathbf{m})}(T)$ is bounded by the number of rational points in $\mathbb{P}^1_{\mathbb{Q}}$ having height at most $O_\varepsilon\left(T^{\frac{1}{2n-8-\varepsilon}}\right)$, which is $O_\varepsilon\left(T^{\frac{1}{n-4-\varepsilon}}\right)$. By adjusting ε, we see that

$$N_{(\mathbf{a},\mathbf{m})}(T) \ll_\varepsilon T^{\frac{1}{n-4}+\varepsilon}. \qquad \blacksquare$$

Remark 7.9 We do not expect the upper bound given in Theorem 7.8 to be sharp. Indeed, the bound we obtain essentially comes from the scenario that for almost all integers $m \ll_\varepsilon T^{2+\varepsilon}$ that there exist x, y with $\max\{|x|, |y|\} \ll_\varepsilon T^{2+\varepsilon}$ with $Q(x, y)$ divisible by m^2. We expect that this should not be the case.

7.4 Hasse principle for integral points when $\mathbf{m} = (2, 2, 2)$

We now consider the question of whether the Hasse principle holds for integral points on stacky curves of the shape $\mathfrak{X} = \mathfrak{X}(\mathbb{P}^1_{\mathbb{Q}} : (a_1, a_2, a_3), (2, 2, 2))$. By Theorem 3.20 it suffices to consider when it is possible for the stacky part of the height to be equal to one. This is tantamount to requiring the existence of co-prime integers x, y and integers y_1, y_2, y_3 for which

$$|\ell_i(x, y)| = y_i^2 \text{ for } i = 1, 2, 3.$$

Here, as we recall, $\ell_i(x, y) = \alpha_i x - \beta_i y$, with $a_i = [\alpha_i : \beta_i]$. For $i = 1, 2$, we obtain a system of linear equations

$$\begin{bmatrix} \alpha_1 & -\beta_1 \\ \alpha_2 & -\beta_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y_1^2 \\ y_2^2 \end{bmatrix}.$$

Inverting, we find that

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{\alpha_1\beta_2 - \alpha_2\beta_1} \begin{bmatrix} \beta_2 & -\beta_1 \\ \alpha_2 & -\alpha_1 \end{bmatrix} \begin{bmatrix} y_1^2 \\ y_2^2 \end{bmatrix}.$$

It follows that

$$(\alpha_1\beta_2 - \alpha_2\beta_1)y_3^2 = \alpha_3\left(\beta_2 y_1^2 - \beta_1 y_2^2\right) - \beta_3\left(\alpha_2 y_1^2 - \alpha_1 y_2^2\right),$$

which we can write as

$$(\alpha_2\beta_3 - \alpha_3\beta_2)y_1^2 - (\alpha_1\beta_3 - \alpha_3\beta_1)y_2^2 + (\alpha_1\beta_2 - \alpha_2\beta_1)y_3^2 = \begin{vmatrix} y_1^2 & y_2^2 & y_3^2 \\ \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \end{vmatrix} = 0.$$

Therefore, the existence of the integers y_1, y_2, y_3, and hence x, y, depends on whether this conic has a rational point.

Competing interests   As far as we are aware, there are no competing interests involved with this manuscript.

Data availability Statement   Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

References

[1] F. Beukers, *The Diophantine equation $Ax^p + By^q = Cz^r$*. Duke. Math. J. **91**(1998), no. 1, 61–88.

[2] M. Bhargava and B. Poonen, *The local–global principle for integral points on stacky curves*. Preprint, 2020,   arXiv:2006.00167 [math.NT].

[3] T. D. Browning and D. R. Heath-Brown, *Counting rational points on hypersurfaces*. J. Reine Angew. Math. **584**(2005), 83–115.

[4] H. Darmon, *Faltings plus epsilon, wiles plus epsilon, and the generalized Fermat equation*. C. R. Math. Rep. Acad. Sci. Canada **97**(1997), 3–14.

[5] H. Darmon and A. Granville, *On the equations $z^m = F(x, y)$ and $Ax^p + By^q = Cz^r$*. Bull. Lond. Math. Soc. **27**(1995), no. 6, 513–543.

[6] J. Ellenberg, M. Satriano, and D. Zureick-Brown, *Heights on stacks and a generalized Batyrev–Manin–Malle conjecture*. Forum Math. Sigma **11**(2023), e14.

[7] B. Fantechi, E. Mann, and F. Nironi, *Smooth toric Deligne–Mumford stacks*. J. Reine Angew. Math. **648**(2010), 201–244.

[8] E. Fouvry and J. Kluners, *On the 4-rank of class groups of quadratic number fields*. Invent. Math. **167**(2007), 455–513.

[9] J. Franke, Y. I. Manin, and Y. Tschinkel, *Rational points of bounded height on Fano varieties*. Invent. Math. **95**(1989), 421–435.

[10] A. Geraschenko and M. Satriano, *A "bottom up" characterization of smooth Deligne–Mumford stacks*. Int. Math. Res. Notices. IMRN **21**(2017), 6469–6483.

[11] A. Granville, *ABC allows us to count square-frees*, Int. Math. Res. Not. IMRN **19**(1998), 991–1009.

[12] C. R. Guo, *On solvability of ternary quadratic forms*. Proc. Lond. Math. Soc. (2) **70**(1995), 241–263.

[13] P. le Boudec, *Density of rational points on a certain smooth bihomogeneous threefold*. Int. Math. Res. Not. **2015**(2015), no. 21, 10703–10715.

[14] G. Malle, *On the distribution of Galois groups, II*. Exp. Math. **13**(2004), 129–136.

[15] A. Shnidman, *Rank of Jacobians of twists of hyperelliptic curves of genus one*. MathOverflow. https://mathoverflow.net/questions/380099/rank-of-jacobians-of-twists-of-hyperelliptic-curves-of-genus-one.

[16] Stacks Project Authors, *The stacks project*. https://stacks.math.columbia.edu/tag/04V3.

[17] C. L. Stewart, *On the number of solutions of polynomial congruences and Thue equations*. J. Amer. Math. Soc. (4) **4**(1991), 793–835.

[18] C. L. Stewart and J. Top, *On ranks of twists of elliptic curves and power-free values of binary forms*. J. Amer. Math. Soc. (4) **8**(1995), 943–973.

[19] C. L. Stewart and S. Y. Xiao, *On the representation of k-free integers by binary forms*. Rev. Mat. Iberoam. **37**(2021), no. 2, 723–748.

[20] J. Voight and D. Zureick-Brown, *The canonical ring of a stacky curve*. Memoirs of the American Mathematical Society. **277**(2022), no. 1362 DOI: https://doi.org/10.1090/memo/1362

[21] P. Vojta, *A more general ABC conjecture*. Int. Math. Res. Not. IMRN **1998**(1998), 1103–1116.

Department of Pure Mathematics, University of Waterloo, Waterloo, ON, Canada
e-mail: bnasserd@uwaterloo.ca

Department of Mathematics and Statistics, University of Northern British Columbia, 3333 University Way, Prince George, BC V2N 4Z9, Canada
e-mail: sxiao@unbc.ca