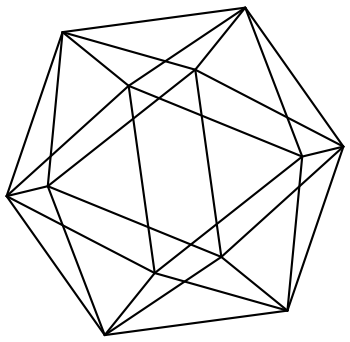
Max-Planck-Institut für Mathematik
Bonn

Bounds for the integral points on elliptic curves over function fields

by

Alisa Sedunova



Max-Planck-Institut für Mathematik
Preprint Series 2017 (69)

Bounds for the integral points on elliptic curves over function fields

Alisa Sedunova

Max-Planck-Institut für Mathematik
Vivatsgasse 7
53111 Bonn
Germany

MPIM 17-69

A. Sedunova. (2017) "Bounds for the integral points on elliptic curves over function fields,"
International Mathematics Research Notices, Vol. 2017, Article ID , 15 pages.
doi:10.1093/imrn/

Bounds for the integral points on elliptic curves over function fields

Alisa Sedunova[1]

[1]MPIM Bonn

Correspondence to be sent to: alisa.sedunova@phystech.edu

In this paper we give an upper bound for the number of integral points on an elliptic curve E over $\mathbb{F}_q[T]$ in terms of its conductor N and q. We proceed by applying the lower bounds for the canonical height that are analogous to those given by Silverman and extend the technique developed by Helfgott-Venkatesh to express the number of integral points on E in terms of its algebraic rank. We also use the sphere packing results to optimize the size of an implied constant. In the end we use partial Birch Swinnerton-Dyer conjecture that is known to be true over function fields to bound the algebraic rank by the analytic one and apply the explicit formula for the analytic rank of E.

1 Introduction

Let q be a prime power and $K = \mathbb{F}_q[T]$ be the field of polynomials in formal variable T with coefficients in a finite field $k = \mathbb{F}_q$ of order p. Our main goal here is to prove the following theorem.

Theorem 1.1. Let E be an elliptic curve over $\mathbb{F}_q[T]$ of a conductor N. Assume that the integral points on E are on minimal model. Then the number of integral points on E satisfies

$$\#E(\mathbb{F}_q[T]) \leq \exp\left(c\frac{\deg N_E}{\log\deg N_E}\right),$$

where c is an absolute constant and N_E is the degree of the conductor of E. □

Notice, that we work in the context where the analogue of Siegel's theorem is true (it is proven in [19]). In particular, if E is an ellipic curve over $\mathbb{F}_q[T]$ parametrized by $a, b \in \mathbb{F}_q$, then $E(\mathbb{F}_q(T)) = E(\mathbb{F}_q)$ and $\#E(\mathbb{F}_q[T]) \leq q + 1 + 2\sqrt{q}$. For a more general function field $\mathbb{F}_q(C)$ with ring of integers A we can have $E(A)$ infinite. Notice that if E is constant, i.e. defined over $\mathbb{F}_q$, then $E(\mathbb{F}_q(T)) = E(\mathbb{F}_q)$, therefore Siegel theorem holds in this case too. For the case of E being isotrivial (not defined over F_q and supersingular) Siegel theorem may be false.

The tools that allow us to proceed are that the necessary part of the famous Birch and Swinnerton-Dyer conjecture holds in the function field context, as well as the bounds for the analytic rank over a function field are known, thanks to the explicit formula given by Brumer in [3]. We also extend the technique of Helfgott (see [7]) to obtain an upper bound for the number of integral points on E in terms of its algebraic rank. However, this brings us to results that do depend on the curve. To get rid of this dependence we have to work with the estimation of the sort $\#E(\mathbb{F}_q[T]) \ll c^{\operatorname{rank} E+m}$ more carefully(here m stands for the number of multiplicative places). Namely, we extend the method developed by Helfgott-Venkatesh in [8] based on the ideas of Silverman [16]. We optimize the size of c by applying sphere packing results of Kabatiansky and Levenshtein [11].

The previously known bounds of such a type (see Theorem 1 of [15]) give us $\#E(\mathbb{Z} \cap I^2) \ll |I|^{\frac{1}{3}+\varepsilon}$, where we are restricted to counting integral points lying in a small box of size $|I|$, where I is an 'interval' of polynomials defined in [15]. This result is analogous to Bombieri-Pila theorem [2], that gives the upper bound $\ll N^{\frac{1}{d}+\varepsilon}$, where d is the degree of a curve and is equal to 3 in the case of elliptic curves, however the method of getting it is different and mainly based on the ideas of Helfgott-Venkatesh [8] and the interpolation part used by Heath-Brown [6]. Here we take the approach proposed by Helfgott in [8] and further developed by Helfgott-Venkatesh in [8], but it turns out that this way of doing things is closely related to the one used in [2].

The paper is organized as follows. In Section 2 we review some basic definitions, that are going to be used throughout the paper as well as some important facts (see (5) and (6), also (7)) that are crucial in our proof.

Received 1 Month 20XX; Revised 11 Month 20XX; Accepted 21 Month 20XX

Then we prove several standard results regarding canonical height on an elliptic curve E. Based on this we show how to get a cheap, but useless bound for the number of points in $E(\mathbb{F}_q[T])$ of a bounded height. We introduce local heights $\lambda_v(\cdot)$ to get rid of this problem and prove lower bounds for $\lambda_v(\cdot)$ under some 'good' slicing, that will bring us to another bound for the canonical height, namely Lemma 3.5, that is proved in the spirit of [8, Proposition 3.4]. We also need a lower bound for the canonical height on E due to Silverman, see [16].

Further, in Section 3 we prove the bound for the number of S-integral points on E in terms of algebraic rank of E using Lemmas from previous sections together with sphere packing results by Kabatiansky and Levenstein [11]. Finally, in Section 5 we prove the main result by taking an advantage of working in function fields, where Birch and Swinnerton-Dyer conjecture partly holds (see (5)) and apply the explicit formula for an analytic rank, given in the expression (6) by Brumer.

2 Auxiliary results

We briefly review some tools that we use during the proof. For more detailed survey see the work of Ulmer [18]. Let $k = \mathbb{F}_q$ be the finite field of cardinality q, with its characteristics $\mathrm{char}(k) = p$. We write K for the function field of a smooth, projective absolutely irreducible curve $\mathcal{C}$ over k. In what follows we consider $\mathcal{C} = \mathbb{P}^1$, thus $K = \mathbb{F}_q[T]$ is the field of polynomials in a formal variable T with coefficients lying in k. For $X \in K$ we denote by $|X|$ its norm: $|X| = q^{\deg X}$. We recall that an elliptic curve over K is a smooth, projective, absolutely irreducible algebraic curve of genus 1 over K with a K-rational point $\mathcal{O}$ that plays the role of identity element in the group $E(K)$ of K-rational points lying on E (Mordell-Weil group of E). Lang and Néron generalized the result of Mordell-Weil and proved that for a funstion field K $E(K)$ is a finitely generated abelian group. As a consequence of this result the torsion group $E(K)_{\mathrm{tors}}$ (i.e. the group of K-points on E of finite order) is finite and isomorphic to a group of the form

$$\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z},$$

where m divides n and p does not divide m. Define an algebraic rank(E) of an elliptic curve E/K as the number of independent points of infinite order in $E(K)$, so to say the number of copies of $\mathbb{Z}$ in $E(K)$.

An equivalent definition of an elliptic curve E/K can be given due to the Riemann-Roch theorem: an elliptic curve E/K can always be described as a projective plane curve of degree 3 with a (homogeneous) Weierstrass equation

$$y^2z + a_1xyz + a_3yz^2 = x^3 + a_2x^2z + a_4xz^2 + a_6z^3, \tag{1}$$

where all a_i belong to K. As usually, the origin is the point at infinity, namely $\mathcal{O} = [0 : 1 : 0]$. The condition of smoothness of E is equivalent to the fact that its discriminant Δ is not zero. The equation above can be also given in an affine form by the change of variables $(x, y) \to (x/z, y/z)$

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6. \tag{2}$$

Let v be an equivalence class of valuations of K. Recall that a valuation on a field K is a generalization of the p-adic norm. Concretely, it is a function $|\cdot|_v$ from a field K to the real numbers $\mathbb{R}$ such that the following properties hold for all $x, y \in K$:

- $|x|_v \geq 0$, $|x| = 0$ if and only if $x = 0$;
- $|xy|_v = |x|_v \cdot |y|_v$;
- $|x|_v \leq 1$ implies $|1 + x|_v \leq C$ for some constant $C \geq 1$ independent of x.

Notice that if a valuation $|\cdot|_v$ satisfies the last condition above with $C = 2$, then it satisfies the triangle inequality

$$|x + y|_v \leq |x|_v + |y|_v$$

for all $x, y \in K$ and such a valuation is called archimedean. If the condition is satisfied with $C = 1$, then $|\cdot|_v$ satisfies the stronger ultrametric inequality:

$$|x + y|_v \leq \max(|x|_v, |y|_v)$$

for all $x, y \in K$ and we call this valuation non-archimedean. Here we work only with non-archimedean valuations.

For every v denote by $\mathcal{O}_{(v)}$ the ring of rational functions on $\mathcal{C}$ regular at v. In our case ($\mathcal{C} = \mathbb{P}^1$) the finite places correspond to monic irreducible polynomials $f \in K = \mathbb{F}_q[T]$. If such a place v corresponds to f, then

$$\mathcal{O}_{(v)} = \{g/h. \text{ s.t. } g, h \in K, \deg(g) < \deg(h)\}.$$

Assume that the degree of $v=\infty$ is 1. Write $\mathcal{M}_v \subset \mathcal{O}_{(v)}$ for the maximal ideal (its elements are the functions vanishing at v) and $\kappa_v = \mathcal{O}_{(v)}/\mathcal{M}_v$ for the residue field at v. Set $\deg(v) = [\kappa_v : k]$, $q_v = q^{\deg(v)}$ for the norm of v. Choose a minimal integral model for E in the form (2). Let $\bar{a}_i \in \kappa_v$ be the reductions of the coefficients at v and define the reduced curve E_v by

$$E_v:\ y^2 + \bar{a}_1 xy + \bar{a}_3 y = x^3 + \bar{a}_2 x^2 + \bar{a}_4 x + \bar{a}_6 \tag{3}$$

over the residue field κ_v. We say that E_v has

- a good reduction at v if E_v defines an elliptic curve over κ_v ($v \nmid \Delta$),
- a multiplicative (nodal) reduction at v if E_v has a node at v. If the tangent lines at the node are rational over the residue field κ_v, then we call this type of reduction split multiplicative. Otherwise non-split multiplicative.
- an additive (cuspidal) reduction at v if E_v has a cusp at v.

Notice that terms multiplicative and additive are used here to emphasize that the non-singular part of the reduced curve defined by $E_v^* = E_v/\{\text{singular point}\}$ is isomorphic to $\mathbb{G}_m$ (or $\mathbb{G}_m[\cdot]$ for the non-split case) and $\mathbb{G}_a$ respectively (here $\mathbb{G}_m$ stands for the multiplicative group, $\mathbb{G}_m[\cdot]$ for the twisted multiplicative group and $\mathbb{G}_a$ for the additive group). Elliptic curves, $\mathbb{G}_a$, $\mathbb{G}_m$ and $\mathbb{G}_m[\cdot]$ over K are the only irreducible algebraic curves over K having group structures given by regular maps.

The reduced curve E_v may be singular, but yet the set of nonsingular points of $\tilde{E}_v(K_v)$ forms a group. Moreover $E(K)$ admits the following filtration of abelian groups

$$E_1(K) \subset E_0(K) \subset E(K),$$

where $E_0(K) = \{P \in E(K) : P_v \in \tilde{E}_v(K_v)\}$ and $E_1(K) = \{P \in E(K) : P_v = O_v\}$ with P_v taken to be the image of $P \in E(K)$ under the reduction map $E(K) \to \tilde{E}_v(K_v)$.

A model for E given by E_v with its coefficients $\bar{a}_i \in \mathcal{O}_{(v)}$ is called integral at v. The minimal integral model at v is the model E_v with the valuation of the discriminant Δ of E being minimal. The local exponent n_v of the conductor at v is given by

$$n_v = \begin{cases} 0, & \text{if } E \text{ has good reduction at } v, \\ 1, & \text{if } E \text{ has multiplicative reduction at } v, \\ 2+\delta_v, & \text{if } E \text{ has additive reduction at } v, \end{cases}$$

where δ_v is the wild ramification

$$\delta_v = \begin{cases} 0, & \text{if } p > 3, \\ \geqslant 0, & \text{if } p = 2, 3. \end{cases}$$

Thus n_v has the information about the ramification in the field extensions generated by the points of finite order in the group law of the elliptic curve E. The conductor of E/K is given by a product of prime ideals and associated exponents n_v. The (global) conductor of E is a divisor $N = \sum_v n_v[v]$. The degree of the conductor is $\deg N = \sum_v n_v \deg v$. N is an effective divisor on $\mathbb{P}^1$ which is divisible only by the places v of bad reduction of E. The L-function of E is defined be the Euler product

$$L(E,s) = \prod_{v \nmid N}^{\text{good}} \left(1 - \frac{a_v}{q_v^s} + \frac{q_v}{{q_v}^{2s}}\right)^{-1} \times \prod_{v|N}^{\text{mult}} \left(1 - \frac{1}{q_v^s}\right)^{-1} \tag{4}$$

where "good" stands for "E has a good reduction at v", "mult" – for the case of either split multiplicative or non split multiplicative reduction at v and, finally, a_v is an integer defined as

$$a_v = \begin{cases} q_v + 1 - \#E_v(k_v), & \text{if } E \text{ has good reduction at } v, \\ \pm 1, & \text{if } E \text{ has multiplicative reduction at } v, \\ 0, & \text{if } E \text{ has additive reduction at } v. \end{cases}$$

($a_v = 1$ for the split multiplicative reduction and to -1 for the non split multiplicative reduction). Due to the Hasse bound on a_v the first product of (4) converges absolutely for $\operatorname{Re} s > 3/2$ and admits a meromorphic continuation on $\mathbb{C}$. As usually we define an analytic rank of E/K as the order of vanishing of its L-function at $s=1$

$$\operatorname{rank}_{an}(E) = \operatorname{ord}_{s=1} L(E,s).$$

We recall that an elliptic curve E/K is called constant if it can be defined by a Weierstrass equation (2) with coefficients belong to k. It is called non-constant if it is not constant. Also E/K is called isotrivial if it becomes constant over some finite extension of K, otherwise – non-isotrivial.

Remark 1. In the non-constant case of E Theorem 9.3 of [18] gives us an upper bound of a type $\operatorname{rank}_{an} E \leqslant N$. □

The famous conjecture of Birch and Swinnerton-Dyer connects the analytic behaviour of L-functions of elliptic curves with the group of K-rational points on E/K, in particular (among some other relations) it predicts that

$$\operatorname{rank}_{an}(E) \stackrel{?}{=} \operatorname{rank}(E).$$

While the original conjecture remains unsolved, much more is known in this context for the case of function fields.

Theorem (Tate [17], Milne [12])**.** Let E be an elliptic curve over a function field K. Then

$$\operatorname{rank} E \leq \operatorname{rank}_{an} E. \tag{5}$$

□

The usual technique for obtaining upper bounds of an analytic rank is using so-called explicit formula. We refer here to the result given by [3].

Theorem (Brumer [3])**.** Let E be an elliptic curve over $\mathbb{F}_q[T]$. Then its analytic rank is bounded by

$$\operatorname{rank}_{an} E \leq \frac{(b_E - 4)\log q}{2\log b_E} + O\left(\frac{n_E \log^2 q}{\sqrt{q}\log^2 b_E}\right), \tag{6}$$

where b_E is the degree of L-function as a polynomial in q^{-s}. □

For the case of $\mathbb{F}_q[T]$ we have

$$b_E = n_E - 4,$$

where $n_E = \deg N$ and N is the conductor of an elliptic curve E/K. We note that if E has a additive reductions and m multiplicative reductions, then

$$n_E \leq 2a + m.$$

This result is interesting if and only if n_E is rather big, since the trivial bound for the rank is $n_E + 4g_X - 4$. We thus have

$$\operatorname{rank}_{an} E \leq \frac{(\deg N - 8)\log q}{2\log\deg N} + O\left(\frac{\deg N \log^2 q}{\sqrt{q}\log^2 \deg N}\right). \tag{7}$$

The easy bound is

$$\operatorname{rank} E \leq \operatorname{rank}_{an} E \leq b_E = n_E - 4.$$

If E is constant, then $\operatorname{rank} E = 0$.

3 Heights and their properties

Here we investigate some properties of height function on an elliptic curve E over a field $K = \mathbb{F}_q[T]$. The crucial fact here is that $|\hat{h} - \frac{1}{2}h_x|$ and $|\hat{h}^E - \frac{1}{3}h_y|$ are bounded on the set of all points of E. This allows us to give a lower bound for $\hat{h}^E(P)$ as well as to estimate the number of points with $\hat{h}^E < c_2$ under condition that E does not have any non torsion points P with $\hat{h}^E(P) > c_1$. However, this path leads us to a problem that the bound would depend on the curve. To avoid this difficulty we will use local heights as in [4] and establish the bound $\lambda_v(P-Q) \geq \min(\lambda_v(P), \lambda_v(Q))$ that fails only in the case of bad reduction with which we will deal separately. We subdivide $E(K_v)$ into small enough number of slices, so that $\lambda_v(P-Q) \geq \min(\lambda_v(P), \lambda_v(Q))$ still holds true on these slices with P, Q belong to the same slice (for more details see Lemma 3.3 and Lemma 3.4). Using that we prove that integral points we wish to count are far apart from each other in the Mordell-Weil lattice. Recall that any elliptic curve over K can be written in the following form

$$E:\ y^2 = f(x), \tag{8}$$

where $f(x) \in K$ is a cubic polynomial defined by Weierstrass equation. We say that $d \in K$ is square free if it has no factor of the form g^2 with $g \in K$ and $\deg g \geq 1$. For any $d \in K$ square free define a quadratic twist of E as

$$E_d:\ dy^2 = f(x). \tag{9}$$

Note that we restrict to the case of square free d, since if d has a squared factor, then by a change of variables in (9) one can find a curve E_d^* isomorphic to E_d. We write $\hat{h}^E$ for the canonical height on an elliptic curve E, and h_x, h_y for the height on E with respect to x and y:

$$\hat{h}^E((x,y)) = \lim_{n\to\infty} \frac{1}{n^2} h_x([n](x,y)), \tag{10}$$

where we use the notation $[n]P = \underbrace{P + \ldots + P}_{n \text{ times}}$ and

$$h_x((x,y)) = \begin{cases} 0, & \text{if } P = \mathcal{O}, \\ \log_q H(x), & \text{otherwise,} \end{cases}$$
$$h_y((x,y)) = \begin{cases} 0, & \text{if } P = \mathcal{O}, \\ \log_q H(y), & \text{otherwise.} \end{cases}$$

For any $x \in K$ define its norm by $|x| = q^{\deg x}$. We notice that $\hat{h}^E$ is defined on all points of $E(\bar{K})$ and $\hat{h}$ is a positive definite quadratic form on $E(\bar{K})$ as well as on $E(K)$(in the sense that it maps non-torsion elements to positive numbers).

For $x = x_0/x_1$ with $x_0, x_1 \in K$ not having as polynomials any common factor other than a constant polynomial in K (we encrypt this fact by $(x_0, x_1)_K = 1$), one can write $H(x) = \max(|x_0|, |x_1|)$. Let L be any algebraic field extension of $\mathbb{F}_q[T]$. Define $H(y)$ by

$$H(y) = (H_L(y))^{[L:K]^{-1}},\ H_L(y) = \prod_w \max(|y|_w^{n_w}, 1),$$

where $y \in L$, the product is taken over all places w of L, n_w stands for the degree of quotient field $L_w/K_w[T]$. For example, if $y = \frac{y_0}{y_1}$ with $y_0, y_1 \in K$, then $y \in \mathbb{F}_q(T)$ and for $L = \mathbb{F}_q(T)$ $H(y) = H_L(y) = \max(|y_0|, |y_1|)$. We list some important properties of the canonical height in the following lemma.

Lemma 3.1. Let $f(x) \in K = \mathbb{F}_q[T]$ be a monic polynomial of non-zero discriminant in (8). Let also d be a square-free polynomial $d \in K$ and $P = (x, y)$ be a K-point on the quadratic twist E_d of E. Let $P' = (x, d^{1/2}y)$ be a point on $E_1 = E$ associated to P. Then

1. $\hat{h}^{E_d}(P) = \hat{h}^E(P')$, where the canonical heights are defined on E_d and E, respectively and, of course, $\deg f = 3$.
2. The height h_y ($y \neq 0$) is bounded on E, namely $h_y(P') \geq \frac{3}{8} \deg d$.
3. If $\deg f = 3$, then $\hat{h}^{E_d}(P) \geq \frac{1}{8} \deg d + c_f$, where c_f is a constant depending only on f.

□

Proof. **1.** We do not put any change in the x-coordinate, so clearly $h_x(P) = h_x(P')$. For the sake of simplicity we consider the case of $\operatorname{char} k \neq 2, 3$. The proof goes analogously in the characteristics 2 and 3. Under this assumption we can write an equation of E in so-called short Weierstrass form (see, for example, Theorem 2.1 in [13])

$$E : y^2 = x^3 + ax + b,\ a, b \in K. \tag{11}$$

Then the duplication law on E is given by

$$[2]P = P + P = \left(\frac{(3x^2 + a)^2 - 8xy^2}{4y^2}, \frac{F_{a,b}(x)}{(2y)^3} \right), \tag{12}$$

where $F_{a,b}(x) = x^6 + 5ax^4 + 20bx^3 - 5a^2x^2 - 4abx - a^3 - 8b^2$. The short Weierstrass equation for the twisted curve E_d is given by the change of variables $(x, y) \to (dx, d^2y)$

$$E_d : y^2 = x^3 + ad^2x + bd^3.$$

Write $X(P)$ and $Y(P)$ for the coordinate functions of P. Then

$$X([2]P') = \frac{(3x^2 + a)^2 - 8dxy^2}{4dy^2} \text{ and } X([2]P) = \frac{(3x^2 + a)^2 - 8dxy^2}{4y^2}.$$

Thus $X([2]P') = X((P+P)')$. Further,

$$Y([2]P) = \frac{F_{a,b}(x)}{(2y)^3} \text{ and } Y([2]P') = \frac{F_{a,b}(x)}{d^{\frac{3}{2}}(2y)^3},$$

which shows that $Y([2]P') = Y((P+P)')$. We conclude that $(P+P)' = P' + P'$. Notice that here the addition is made on E_d on the left hand side and on E on the right hand side. Iterating this and using (10) we get

$$\hat{h}(P) = \frac{1}{2}\lim_{n\to\infty}\frac{h_x([2^n]P)}{2^{2n}} = \frac{1}{2}\lim_{n\to\infty}\frac{h_x([2^n]P')}{2^{2n}} = \hat{h}(P').$$

2. Write $y = \frac{y_0}{y_1}$ for $y_0, y_1 \in K$, such that they do not have any common factor $g \in K$ of a positive degree. For $a, b \in K$ we denote by $\langle a, b\rangle = \langle a, b\rangle_K$ the biggest common factor (in the sense that there is no other polynomial $g \in K$ of a bigger degree, such that g is a factor of both a and b) of polynomials a, b. We have $\langle y_0, y_1\rangle_K = 1$ and we call such polynomials coprime. If g is a monic irreducible polynomial, such that g is a factor of $\langle d, y_1^2\rangle$, then g^2 can not be a factor of $\langle d, y_1^2\rangle$ (by the fact that d is a square free polynomial), but it is a factor of y_1. Hence, if g is not a factor of $\langle d, y_1^2\rangle$, then write

$$\langle d, y_1^2\rangle = \frac{dy_1^2}{\{d, y_1^2\}},$$

where $\{d, y_1^2\}$ is a minimal polynomial that has both d and y_1^2 as factors. Then using the fact that y_0 and y_1 are taken to be coprime we conclude that g has a power -1 as a factor of $dy^2 = dy_0y_1^{-2} = d^2y_0^2\langle d, y_1^2\rangle^{-1}\{d, y_1^2\}^{-1}$. Recall that P lies on our curve E, so $dy^2 = f(x)$ and if g has a non-negative degree as a factor of x, then it also has a non-negative degree as a factor of dy^2. But if g has a negative degree as a factor of x, then its degree in dy^2 drops to ≤ -3 leaving us with a contradiction. Therefore we conclude that $|y_1| \geq \langle d, y_1^2\rangle^2$. Since $y \in K$ we can write by the definition of $H(y)$ and considering the Euclidean norm

$$\begin{aligned} H(y) &= \max\left(|y_0||d^{-1}\langle d, y_1^2\rangle|^{-\frac{1}{2}}, |y_1||\langle d, y_1^2\rangle|^{-\frac{1}{2}}\right) \\ &\geq \max\left(|y_0||d^{-1}\langle d, y_1^2\rangle|^{-\frac{1}{2}}, |\langle d, y_1^2\rangle|^{\frac{3}{2}}\right) \geq |d|^{\frac{3}{8}}, \end{aligned}$$

where we used the fact that max gets its minimal value when $|\langle d, y_1^2\rangle| = |d|^{\frac{1}{4}}$. Finally, $h_y(P) = \log H(P) \geq \frac{3}{8}\log q^{\deg d} = \frac{3}{8}\deg d$.

3. It is a simple consequence of 1 and 2. If P' is a point on $E = E_1$, then, by 1 $\hat{h}^{E_d}(P) = \hat{h}^E(P')$. The difference $|\hat{h}^E - h_x^E|$ is bounded on E, thus by application of second part of 3.1 the result follows. ∎

Corollary 3.2. Let E be an elliptic curve over $K = \mathbb{F}_q[T]$. If there are no non-torsion points $P \in E(K)$ of a canonical height $\hat{h}(P) > c_1$, then there are at most

$$O\left(\left(1 + 2\sqrt{\frac{c_2}{c_1}}\right)^{\operatorname{rank} E}\right)$$

points in $E(K)$ of a canonical height $< c_2$. □

Proof. Let's take our canonical height to the square of the Euclidean norm. There is one to one correspondence $f : K^{\operatorname{rank} E} \to K^{\operatorname{rank} E}$ such that $\hat{h}^E(\mathbb{P}) = |f(\mathbb{P})|^2$ for all vectors $\mathbb{P} \in K^{\operatorname{rank} E}$ of the length $\operatorname{rank} E$ with coordinates in K. Since $\hat{h}^E(P) > c_1$ for all non-zero $P \in K$, then we are equipped by $f(K^{\operatorname{rank} E})$ with a lattice L, such that for every element $l \in L$ different from 0 we have $|l| \geq c_1^{\frac{1}{2}}$. For every point $l \in L$ draw a sphere Sp_l centred at l of the radius $\frac{1}{2}c_1^{\frac{1}{2}}$, so that they do not overlap. Each of the spheres Sp_l is contained in the bigger one Sp with the radius $c_2^{\frac{1}{2}} + \frac{1}{2}c_1^{\frac{1}{2}}$ centred at the origin. By bounding the total volume of all spheres by $\operatorname{vol}(Sp) \leq (c_2^{\frac{1}{2}} + \frac{1}{2}c_1^{\frac{1}{2}})^{\operatorname{rank} E}$ we end the proof. ∎

The implied constants c_1, c_2 do not have any dependency on the twist, but depend on the curve. This would bring us to a problem once we want to bound the canonical height in terms of naive height (namely, we want something of the sort $h(P) \leq c_3$, where $h(P)$ is the naive height and c_3 is an absolute constant), because then the constant inside big O will change to $(1 + 2\sqrt{c_3/c_1})^{\operatorname{rank} E}$, where c_1 depends only on the curve, whilst c_3 depends on both the curve and c_2 (say, $c_2 = c_3 + O_E(1)$). To avoid this difficulty we have to exclude the hidden dependency by the method proposed in [8].

Recall that κ_v is the residue field at v and $d_v = \deg(v) = [\kappa_v : k]$. Let M_k be the set of places v on K. For each place $v \in K$, there exists a natural local height function λ_v such that the canonical height on E can be given in terms of λ_v

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

We say that an elliptic curve E over a non-archimedean local field K has potentially good reduction if it has a model with good reduction in some extension of K. Similarly, E has potentially multiplicative reduction if it does not have potentially good reduction.

Lemma 3.3. Let E be an elliptic curve over a non-archimedean local field K_v with potentially good reduction. Let $P, Q \in E(K_v)$ be two distinct points. Then

$$\lambda_v(P - Q) \geq \min(\lambda_v(P), \lambda_v(Q)).$$

□

Proof. Consider an extension L_w of K_v on which E has good reduction. Choose a Weierstrass equation for E over L_w such that $v(\Delta) = 0$. Then by [4, Proposition 2] we find that

$$\lambda_v(P) = \lambda_w(P) = \frac{1}{2} \max(\log |x(P)|_w, 0).$$

Since v is non-archimedean, then $|x + y|_v \leq \max(|x|_v, |y|_v)$ and the claim follows. ■

The following lemma is [8, Lemma 3.2] and the proof is completely analogous.

Lemma 3.4. Let E be an elliptic curve over a non-archimedean local field K_v with potentially multiplicative reduction. Then for any $\varepsilon > 0$ small enough, there is a subdivision

$$E(K_v) = W_{v,0} \cup W_{v,1} \cup \ldots \cup W_{v,d_v} \ll |\log \varepsilon|,$$

such that for any two distinct points $P, Q \in W_{v,0}$ we have

$$\lambda_v(P - Q) \geq \min(\lambda_v(P), \lambda_v(Q)), \quad \lambda_v(P_1), \lambda_v(P_2) \geq 0,$$

and for any two distinct points $P, Q \in W_{v,j}$, where $1 \leq j \leq d_v$ we have

$$\lambda_v(P - Q) \geq (1 - \varepsilon) \max(\lambda_v(P), \lambda_v(Q)),$$
$$\lambda_v(P - Q) \geq (1 - 2\varepsilon) \max(\lambda_v(P), \lambda_v(Q)),$$

where the implied constant is absolute. □

Now we have to adapt [8, Proposition 3.4], that will serve us for as a bound for the canonical height that does not depend on the curve any longer. Here we assume that our two points are of the same reduction as well as that they fall into the same W-class, so we can apply Lemma 3.4.

Since we are working in $K = \mathbb{F}_q[T]$, we don't have any archimedean valuations and thus, the proof can be significantly simplified.

Lemma 3.5. Let E be an elliptic curve over K. Let S be a finite set of places of $K = \mathbb{F}_q[T]$, that includes all irreducible divisors of the discriminant Δ of E. Let P_1, P_2 be two distinct integral points on E that belong to the same set $W_{v,i}$ for any place v among the ones with potentially multiplicative reduction. Suppose that

$$\sum_{v \in T} d_v |\lambda_v(P_1) - \lambda_v(P_2)| \leq \varepsilon \max_{j=1,2} \sum_{v \in T} d_v \lambda_v(P_j),$$

where $\varepsilon > 0$ sufficiently small and

$$T = \{v \in S : \lambda_v(P_1), \lambda_v(P_2) \geq 0\}.$$

Assume that P_1 and P_2 have the same reduction modulo I, where I is any ideal not divisible by irreducible elements of S. Then

$$\hat{h}(P_1 - P_2) \geq (1 - 2\varepsilon) \max(\hat{h}(P_1), \hat{h}(P_2)) + \frac{\log NI}{[K:L]}.$$

□

Proof. If v is a finite place of good reduction, then $\lambda_v(P) \geq 0$. Recall that S contains all places that divide the discriminant Δ of E. Then by definition of a canonical height through local heights we have

$$\begin{aligned}\hat{h}(P_1 - P_2) &\geq \sum_{v\in S} d_v\lambda_v(P_1 - P_2) + \sum_{v\notin S} d_v\lambda_v(P_1 - P_2)\\ &= \sum_{v\in S} d_v\lambda_v(P_1 - P_2) + \sum_{\substack{v \text{ finite}\\ v(I)>0}} d_v\lambda_v(P_1 - P_2).\end{aligned}$$

We now subdivide our set S as $S = T \cup S/T$, where T is defined in the statement of the lemma. Let us consider two differences

$$\begin{aligned}\sigma_1 &= \sum_{v\in T} d_v\lambda_v(P_1 - P_2) - (1-\varepsilon)\sum_{v\in T} d_v \min(\lambda_v(P_1), \lambda_v(P_2)),\\ \sigma_2 &= \sum_{v\in S/T} d_v\lambda_v(P_1 - P_2) - (1-2\varepsilon)\max_{j=1,2}\sum_{v\in S/T} d_v\lambda_v(P_j).\end{aligned}$$

The goal now is to show that these two quantities $\sigma_1, \sigma_2 \geq 0$. Once we are done it remains to consider only finite places v, such that $v(I) > 0$. We use the following notations $\sum^{\text{good}}, \sum^{0}, \sum^{j}$ denote that P_1, P_2 are of potentially good reduction, potentially multiplicative reduction and fall into $W_{v,0}$, potentially multiplicative reduction and fall into $W_{v,j}$ with $j > 0$ respectively. By Lemma 3.3 and Lemma 3.4 we have

$$\begin{aligned}\sigma_1 &\geq \sum_{v\in T}^{good,0} d_v \min_{j=1,2}\lambda_v(P_j) + (1-\varepsilon)\sum_{v\in T}^{j} d_v \max_{j=1,2}\lambda_v(P_j) - (1-\varepsilon)\sum_{v\in T} d_v \min_{j=1,2}\lambda_v(P_j)\\
&= \varepsilon\sum_{v\in T} d_v \min_{j=1,2}\lambda_v(P_j) - \varepsilon\sum_{v\in T}^{j} d_v \max_{j=1,2}\lambda_v(P_j) + \sum_{v\in T}^{j} d_v\big(\max_{j=1,2}\lambda_v(P_j) - \min_{j=1,2}\lambda_v(P_j)\big)\\
&\geq \varepsilon\sum_{v\in T} d_v \min_{j=1,2}\lambda_v(P_j) - \varepsilon\sum_{v\in T}^{j} d_v \max_{j=1,2}\lambda_v(P_j)\\
&= \varepsilon\sum_{v\in T}^{good,0} d_v \min_{j=1,2}\lambda_v(P_j) + \varepsilon\sum_{v\in T}^{j} d_v\big(\min_{j=1,2}\lambda_v(P_j) - \max_{j=1,2}\lambda_v(P_j)\big)\\
&= \varepsilon\sum_{v\in T}^{good,0} d_v \min_{j=1,2}\lambda_v(P_j) - \varepsilon\sum_{v\in T}^{good,0} d_v\big(\min_{j=1,2}\lambda_v(P_j) - \max_{j=1,2}\lambda_v(P_j)\big)\\
&+ \varepsilon\sum_{v\in T} d_v\big(\min_{j=1,2}\lambda_v(P_j) - \max_{j=1,2}\lambda_v(P_j)\big)\\
&= \varepsilon\sum_{v\in T}^{good,0} d_v \max_{j=1,2}\lambda_v(P_j) + \varepsilon\sum_{v\in T} d_v\big(\min_{j=1,2}\lambda_v(P_j) - \max_{j=1,2}\lambda_v(P_j)\big).\end{aligned}$$

Now we apply the assumption of our lemma and get

$$\begin{aligned}\sigma_1 &\geq \varepsilon\sum_{v\in T}^{good,0} d_v \max_{j=1,2}\lambda_v(P_j) - \varepsilon^2\sum_{v\in T} d_v \max_{j=1,2}\lambda_v(P_j)\\
&= (\varepsilon - \varepsilon^2)\sum_{v\in T}^{good,0} d_v \max_{j=1,2}\lambda_v(P_j) - \varepsilon^2\sum_{v\in T}^{j} d_v \max_{j=1,2}\lambda_v(P_j) \geq 0\end{aligned}$$

by choosing ε small enough. Applying the same condition again we get

$$\begin{aligned}\sum_{v\in T} d_v\lambda_v(P_1-P_2) &\geq (1-\varepsilon)\sum_{v\in T} d_v \max_{j=1,2}\lambda_v(P_j)\\ &+(1-\varepsilon)\sum_{v\in T} d_v\big(\min_{j=1,2}\lambda_v(P_j)-\max_{j=1,2}\lambda_v(P_j)\big)\\ &\geq (1-\varepsilon)\sum_{v\in T} d_v \max_{j=1,2}\lambda_v(P_j)+\sum_{v\in T} d_v\big(\min_{j=1,2}\lambda_v(P_j)-\max_{j=1,2}\lambda_v(P_j)\big)\\ &\geq (1-\varepsilon)\sum_{v\in T} d_v \max_{j=1,2}\lambda_v(P_j)-\varepsilon\sum_{v\in T} d_v \max_{j=1,2}\lambda_v(P_j)\\ &\geq (1-2\varepsilon)\sum_{v\in T} d_v \max_{j=1,2}\lambda_v(P_j).\end{aligned}$$

Similarly for σ_2

$$\sigma_2=\sum_{v\in S/T}^{good} d_v\big(\min_{j=1,2}\lambda_v(P_j)-(1-2\varepsilon)\max_{j=1,2}\lambda_v(P_j)\big)>0$$

with ε being small enough. Combining estimates for σ_1,σ_2 and using the fact that

$$\sum_{v\in T} d_v \max_{j=1,2}\lambda_v(P_j)\geq \max_{j=1,2}\sum_{v\in T} d_v\lambda_v(P_j)$$

one can see that

$$\begin{aligned}\sum_{v\in S} d_v\lambda_v(P_1-P_2) &\geq (1-2\varepsilon)\sum_{v\in T} d_v \max_{j=1,2}\lambda_v(P_j)+(1-2\varepsilon)\max_{j=1,2}\sum_{v\in S/T} d_v\lambda_v(P_j)\\ &\geq (1-2\varepsilon)\max_{j=1,2}\sum_{v\in S} d_v\lambda_v(P_j).\end{aligned}$$

Since S contains all places that do divide the discriminant, then we have

$$\lambda_v(P)=\frac{1}{2}\log^+(|x(P)|_v)=0,\ \text{for } v\notin S.$$

Then

$$\hat{h}_K(P_1-P_2)\geq (1-2\varepsilon)\max_{j=1,2}\hat{h}_K(P_j)+\sum_{\substack{v \text{ finite}\\ v(I)>0}} d_v\lambda_v(P_1-P_2).$$

It remains to consider only finite places v, such that $v(I)>0$. Let $\mathfrak{p}_v$ be the corresponding prime ideal in O_K with its multiplicity n_v in I. By reduction modulo $\mathfrak{p}_v^{n_v}$ our point P_1-P_2 becomes an origin O. Then

$$v(x(P_1-P_2))\leq -2n_v$$

and

$$\lambda_v(P_1-P_2)\geq \frac{n_v}{e_v}\log p_v,$$

where e_v is the ramification degree of K_v and p_v is the rational irreducible element under v. Thus

$$\sum_{\substack{v \text{ finite}\\ v(I)>0}} d_v\lambda_v(P_1-P_2)=\log NI.$$

■

We are going to exploit Lemma 3.5 to give an upper bound on the number of S-integral points. In order to get a good constant C, that appears in the main result of this paper we are going to apply sphere packings. We first subdivide the set of integer points on E into "good slices" and then apply sphere packing bounds to each part separately. Here we use the remarkable result of Kabatiansky and Levenstein (see, for example, [11]).

Lemma 3.6. [Kabatiansky-Levenstein [11]] Let $A(n,\theta)$ be the maximal number of points that can be arranged on the unit sphere of $\mathbb{R}^n$ such that the angle between P_1, O and P_2 for any two P_1, P_2 of them is no smaller than θ. Then for $0 < \theta < \frac{\pi}{2}$

$$\frac{1}{n}\log_2 A(n,\theta) \leq \frac{1+\sin\theta}{2\sin\theta}\log_2\frac{1+\sin\theta}{2\sin\theta} - \frac{1-\sin\theta}{2\sin\theta}\log_2\frac{1-\sin\theta}{2\sin\theta} + o(1),$$

where the convergence is uniform and explicit for θ within any closed subinterval of $\left(0, \frac{\pi}{2}\right)$. In particular, for $\theta = \frac{\pi}{3}$, we have

$$\frac{1}{n}\log_2 A(m,\theta) \leq 0.40141\ldots$$

□

Lemma 3.7. Let c_1, c_2 be two positive real numbers, $0 < \varepsilon < \frac{1}{2}$, n is a non-negative integer. For $\vec{X} = (X_i)_{1\leq i\leq n} \in \mathbb{F}_q^n[T]$ consider

$$S = \{\vec{X} \in \mathbb{F}_q^n[T]\ c_1 \leq |\vec{X}| \leq c_2\},$$

where $|\vec{X}| = \sum_{i=1}^n |X_i| = \sum_{i=1}^n q^{\deg X_i}$. Then there is a subset $T \subset \mathbb{F}_q^n[T]$ such that

$$\#T \leq C^n \varepsilon^{-(n+1)}\left(1 + \log\frac{c_2}{c_1}\right),$$

where the implied absolute constant C is explicit and the balls $B(\vec{Y}, \varepsilon|\vec{Y}|)$ cover all of S for $\vec{Y} \in T$. □

Proof. It is enough to show the covering by balls $B(\vec{Y}, 2\varepsilon|\vec{Y}|)$. We wish to slice S into a union of regions where $|\cdot|$ is almost constant, namely

$$T = \bigcup_{0\leq m\leq M} \frac{c_1\varepsilon(1+\varepsilon)^m}{n} T_m,$$

where

$$T_m = \{\vec{Y} \in \mathbb{F}_q^n[T] :\ \frac{n}{\varepsilon}(1-\varepsilon) \leq |\vec{Y}| \leq \frac{n}{\varepsilon}(1+\varepsilon)\} \text{ and } M = \log_{1+\varepsilon}\log\frac{c_2}{c_1}.$$

Let $\vec{X} \in S$. Consider

$$m(\vec{X}) = \left\lfloor \log_{1+\varepsilon}\frac{|\vec{X}|}{c_1}\right\rfloor \text{ and } \vec{Z}(\vec{X}) = \left\lfloor \frac{n\vec{X}}{c_1\varepsilon(1+\varepsilon)^{m(\vec{X})}}\right\rfloor,$$

where $\lfloor\cdot\rfloor$ is the floor function. Define $\vec{Y} = \frac{c_1\varepsilon(1+\varepsilon)^m}{n}\vec{Z}(\vec{X})$. Then

$$|\vec{Z}(\vec{X})| \leq \frac{n|\vec{X}|}{c_1\varepsilon(1+\varepsilon)^{m(\vec{X})}} \text{ and thus } |\vec{Y}| \leq |\vec{X}| < c_2,$$

$$|\vec{Z}(\vec{X})| \geq \frac{n|\vec{X}|}{c_1\varepsilon(1+\varepsilon)^{m(\vec{X})}} - 1 \text{ and thus } |\vec{Y}| \geq |\vec{X}| - \frac{c_1\varepsilon(1+\varepsilon)^{m(\vec{X})}}{n}$$

$$\geq c_1 - \frac{c_1\varepsilon(1+\varepsilon)^{M)}}{n} \geq c_1 - \frac{c_2\varepsilon}{n} > c_1.$$

We have just shown that given an $\vec{X} \in S$ one can find a point $\vec{Y}$, that depends on $\vec{X}$ and lies in T. In addition $\vec{Y}$ has the following property

$$d(\vec{X}, \vec{Y}) = |\vec{X} - \vec{Y}| \leq 2\varepsilon|\vec{Y}|,$$

where $d(\cdot,\cdot)$ is the associated metric. It remains to estimate the size of T

$$\#T \leq \left(1 + \log_{1+\varepsilon}\frac{c_2}{c_1}\right)\#T_m \leq \left(1 + \log_{1+\varepsilon}\frac{c_2}{c_1}\right)\frac{(n(1+\frac{1}{\varepsilon})+n)^n}{n!}.$$

The result follows after application of Stirling formula. ■

We will need the following lower bound for a canonical height on E.

Lemma 3.8. Let E be an elliptic curve over K. There is an absolute constant $0 < c < 1$ such that, for every non-torsion point $P \in E(K)$ we have the bound

$$\hat{h}(P) > c^m \max\big(1, h(j(E))\big),$$

where m is the number of mulpiplicative places and $j(E)$ is as usual a j-invariant of E. □

Proof. This Lemma is an analogous result to the ones in [16] and [10]. In fact, a stronger result was proven in [10], namely: $\hat{h}(P) \geq c\sigma_E h(E)$, where σ_E is the Szpiro ratio (it gives $\hat{h}(P) \geq c_1 h(E)$ when $j(E) \in \mathbb{F}_q(T)/\mathbb{F}_q(T^p)$). ■

4 Bounding the number of S-integral points

In this section we prove the bound for the number of S-integer points on E/K of height less than h_0. Here t is a parameter to be optimized further. Then we are going to present a proof of the main result. It follows the way proposed in [7], [16], [4] and later improved in [8]. By embedding $E(K)/E(K)_{tors}$ into $E(K) \otimes_{\mathbb{Z}} \mathbb{R} \cong \mathbb{R}^{\operatorname{rank} E}$ we can take the canonical height on E to be squared Euclidean norm. The key idea consists of the fact that the points we are looking at have large distance between each other. Namely, by choosing a good division of the area into small symmetric slices we can say that any two points are separated by almost 60 degrees. Then the number of integral points on E is bounded above by $2^{\operatorname{rank} E}$ (this constant was later improved to $(1+\varepsilon)$ in [8]). It remains to apply Theorem 5 and (7) for getting the result.

Theorem 4.1. Let E be an elliptic curve over K. Let also S be a finite set of places of K, including all irreducible divisors of the discriminant of E. Then, for any $h_0 \geq 1$ and every $0 \leq t \leq 1$, the number of S-integer points P of $E(K)$ with a canonical height $\hat{h}(P) \leq h_0$ is at most

$$O\left(C^{|S|}\varepsilon^{-2(|S|+[K:L])}|S|^{[K:L]}(1+\log h_0)^2 e^{t[K:L]h_0+(\beta(t)+\varepsilon)\operatorname{rank} E}\right),$$

where C is an absolute constant and $\beta(t)$ is defined for $0 \leq t < 1$ by

$$\beta(t) = \frac{1+f(t)}{2f(t)}\log\frac{1+f(t)}{2f(t)} - \frac{1-f(t)}{2f(t)}\log\frac{1-f(t)}{2f(t)},$$
$$f(t) = \frac{\sqrt{(1+t)(3-t)}}{2}, \quad \beta(1) = 0.$$

□

Proof. Briefly speaking, we subdivide S-integer points on E denoted by $E(K,S)$ into points $(\operatorname{mod} I)$ for I being a suitable ideal in O_K. Then Lemma 3.5 states that after some manipulations on this partition the points, that lie in the same class tend to be far away from each other in the Mordell-Weil lattice. Here we apply sphere packing bounds of Kabatiansky and Levenstein, namely Lemma 3.6 to each part separately. These sphere packing bounds will bring us to the term $e^{\beta(t)\operatorname{rank} E}$ on each part. Summation over all the classes gives rise to another term $e^{[K:L]h_0}$. We have only to take care of getting the right conditions to apply Lemma 3.5.

We firstly subdivide $E(K,S)$ into a very few slices to force any two points of the same slice have comparable canonical height. Consider a set

$$\{P \in E(K,S) : \hat{h}(P) \leq h_0\}.$$

We want to cover it by sets of the form

$$\{P \in E(K,S) : (1-\varepsilon)h_i \leq \hat{h}(P) \leq h_i\}.$$

By Lemma 3.8 it is enough to take $\ll \varepsilon^{-1}(\log h_0 + |S|)$ such sets. Then we are allowed to decrease the power of $(1+\log h_0)^2$ just to 1, only for the set of points

$$\{(1-\epsilon)h_0 \leq \hat{h}(P) \leq h_0\}.$$

Suppose first that $t \neq 0$. Let S' be the set of places below S. If

$$X = \max(\lceil e^{th_0}\rceil, |\bar{S}|^{1+\frac{1}{[K:L]}}),$$

then there is an irreducible polynomial f in L, such that $f \notin \bar{S}$ and $X \le |f| \le 2X$. The ideal I of $\mathcal{O}_K$ generated by f satisfies

$$\frac{\log N(I)}{[K:L]} \ge h_0 t, \quad N(I) \ll_{[K:L]} s^{[K:L]+1} e^{th_0[K:L]}.$$

The S-integer points of our curve $E(K)$ fall into no more than $O_{[K:L]}(N(I))$ classes under the reduction modulo the corresponding ideal I. Define R to be the set of all places of potentially multiplicative reduction. For any place $v \in R$ we subdivide the corresponding $E(K_v)$ into $n_v + 1$ subsets, where n_v is defined as in Lemma 3.4(we take $\frac{\varepsilon}{2}$ instead of ε). Consider arbitrary tuples of the form $(a_v)_{v\in R}$, $(b_v)_{v\in R}$, such that $0 \le a_v \le n_v$ and $b_v = 0, 1$. We define B as the set of non-torsion points $P \in E(K, S)$, such that for each $v \in R$ we have that P falls into the corresponding W-class – $P \in W_{v,a_v}$ and that $\lambda_v(p) \ge 0$ is equivalent to $b_v = 1$. Now we bound the number of elements in

$$B_{h_0} = \{P \in B : (1-\varepsilon)h_0 \le \hat{h}(P) \le h_0\}.$$

The number of such sets B is bounded above by $c_0^2 |\log \varepsilon|^{s+[K:L]_\varepsilon - 2[K:L]}$, that brings us to the desired result. Define $M = (S - R) \cup \{v \in R : b_v = 1\}$ and a map $l(P) = (d_v \lambda_v(P))_{v\in M}$. For $v \in S - M$ we know that $\lambda_v(P) < 0$, so that one can apply Lemma 3.8 and get

$$|l(P)|_1 > [K:L]\kappa^s \max(1, h(j)).$$

Using [4, Proposition 3] we get the bound

$$\sum_{v\notin M} d_v \lambda_v(P) \ge -\frac{1}{24} h_k(j) - 3[K:L].$$

On combining that we obtain

$$|l(P)|_1 \le [K:L](h_0 + 3 + h(j)/24)$$

for $P \in B_{h_0}$. By Lemma 3.4 we can cover $l(B_{h_0})$ by at most

$$O(c_1^s \varepsilon^{-(s+1)} \log(h_0 + 1))$$

balls $B(x, \frac{\varepsilon}{8}|x|_1)$ in the 1-norm. Take two points $P_1, P_2 \in B_{h_0}$ with $l(P_i) \in B(x, \frac{\varepsilon}{8}|x|_1)$ for $i = 1, 2$. We then have

$$|l(P_1) - l(P_2)|_1 \le \frac{\varepsilon}{4}|x|_1 \le \frac{\varepsilon}{2} \max_{j=1,2} |l(P_j)|_1.$$

If these points have the same reduction modulo I, then we apply Lemma 3.5 and get that

$$\hat{h}(P_1 - P_2) \ge (1-\varepsilon) \max_{j=1,2} \hat{h}(P_j) + \frac{\log N(I)}{[K:L]} \ge (1 + t - \varepsilon) \max_{j=1,2} \hat{h}(P_j).$$

Now we embed the Mordell-Weil lattice modulo torsion into $\mathbb{R}^{\operatorname{rank} E}$ by taking $\hat{h}$ to be the square of the Euclidean height. Since all $\hat{h}(P_1)$, $\hat{h}(P_2)$, $\hat{h}(P_1 - P_2)$ are positive, then the images of P_1, P_2, say, $Q_1, Q_2 \in \mathbb{R}^{\operatorname{rank} E}$ aree different from each other and from the origin, so that the angle between them is at least $\arccos \frac{1-t+O(\varepsilon)}{2}$. We now apply Lemma 3.6 and get that there are at most $e^{r(\beta(t)+O(\varepsilon))} O_{[K:L]}(1)$ points of B_{h_0} with an image in a given ball and with a prescribed reduction modulo I. Now we combine these results with the number of variants for I, the number of possible sets B and the number of balls to get the theorem. Notice, that in the case $t = 0$ one simply proceeds without I. ■

The case $t = 0$ is the pure application of sphere-packing results of Lemma 3.6, while the case $t = 1$ is related to the corresponding result of Bombieri-Pila type.

Corollary 4.2. Let E be an elliptic curve over K defined by a Weierstrass equation with integer coefficients. Let S be a finite set of places of K, including all places dividing the discriminant of E. Then for every sufficiently small ε the number of S integral points on E/K is at most

$$O_\varepsilon\left(C^s \varepsilon^{-2(s+1)} (\log|\Delta| + \log p)^2 e^{\operatorname{rank} E(\beta(0)+\varepsilon)}\right).$$

□

We need as well upper bound for the canonical height. Here we adapt the result of Pacheco [14]. There are known bounds over $\mathbb{Q}$, see, for example [5]. Also one finds good bounds in [10], but they work only in characteristic 0.

Lemma 4.3. Let E be an elliptic curve over K defined by a Weierstrass equation $y^2 = f(x)$. Let $\mathcal{O}_S$ be the ring of S-inetegers and $\mathcal{O}_S^*$ be the ring of S-units. Suppose that $f(X) \in \mathcal{O}_S$ and the discriminant $\Delta \in \mathcal{O}_S^*$, $p > 2$. Define a set Ξ in the following way. Let $f(X) = (X - x_1)(X - x_2)(X - x_3)$ be the factorization of $f(X)$ in $\bar{K}[X]$. Let $P = (x_P, y_P) \in \mathcal{O}_S$. Define $\xi_i^2 = X - x_i$, $i = 1, 2, 3$. Let $L = K(x_1, x_2, x_3, \xi_1, \xi_2, \xi_3)$. For any permutation $\{i, l, m\}$ of $\{1, 2, 3\}$ define

$$\Xi = \left\{ \frac{(\xi_i - \xi_l)}{(\xi_i - \xi_m)}, \frac{(\xi_i - \xi_l)}{(\xi_i + \xi_m)}, \frac{(\xi_i + \xi_l)}{(\xi_i - \xi_m)}, \frac{(\xi_i + \xi_l)}{(\xi_i + \xi_m)} \right\}.$$

Then for any $\eta \in \Xi$ we have

$$\hat{h}_L(\eta) \leq 2p^e(2g_L - 2 + |S_L|),$$

where S_L is the set of places of L lying over S and g_L is the genus of L. Moreover, if $p > 3$, then for any $P = (x_P, y_P) \in \mathcal{O}_S$ we have $\hat{h}_K(y_P^4/\Delta) \leq 48p^e(2g - 2 + |S|)$. □

We are now ready to give a version of Theorem 4.1 with an optimized parameter t.

Corollary 4.4. Let E be an elliptic curve over a field K. Let S be a finite set of places of K, that contains all places dividing the discriminant of E. Let $\alpha(x) = \min(xt + \beta(t), 0 \leq t \leq 1)$, where β is as in Theorem 4.1. Let also $R = \max(1, \operatorname{rank} E(K))$. Then for every $h_0 \geq 1$ and for every sufficiently small ε, the number of S-integral points on E over K, that have canonical height less or equal to h_0 is at most

$$O_{\varepsilon,[K:L]}\left(C^{\#S}\varepsilon^{-2(\#S+[K:L])}\#S^{[K:L]}(1 + \log h_0)^2 e^{R\alpha([K:L]h_0/R)+\varepsilon R}\right),$$

where C is an absolute constant. □

We derive some quantitative bounds on the height of integral points on elliptic curve. We follow exactly the way proposed in [8]. A combintation of a bound of Hajdu-Herendi [5] together with our previous results gives the following.

Corollary 4.5. Let E be an elliptic curve over a field K. Let S be a finite set of places of K, that contains all places dividing the discriminant of E. Then the number of S-integral points on E is at most

$$O_\varepsilon\left(C^{\#S}\varepsilon^{-2(\#S+1)}(\log|f| + \log|\Delta|)^2 e^{(\beta(0)+\varepsilon)\operatorname{rank} E}\right),$$

where C is a constant, f is the largest in norm element of S, Δ is the disciminant of E. The calculation gives $\beta(0) = 0.2782...$ □

Furthermore, in the same manner as in [8] we obtain the next corollary.

Corollary 4.6. Let $\varepsilon > 0$, E be an elliptic curve over a field K. Then the number of integral points on E is at most

$$O_\varepsilon\left(|\Delta|^{c+\varepsilon}\right),$$

where Δ is the disciminant of E and the constant $c = \frac{\beta(0)}{\log 2} = 0.20070...$ □

5 Bounds on an algebraic rank

Here we get the desired bound for an algebraic rank and give a bound for the number of S integral points on E in terms of its conductor. Due to the results of the previous section we have

$$\begin{aligned}\#E(K) \ll c^{\operatorname{rank} E+m} &\leq c^{\operatorname{rank}_{an} E} \\ &\leq \exp\left(\log c\left(\frac{(\deg N - 8)\log q}{2\log\deg N} + O\left(\frac{\deg N \log^2 q}{\sqrt{q}\log^2\deg N}\right)\right)\right),\end{aligned}$$

where we used the fact that $\operatorname{rank} E \leq \operatorname{rank}_{an} E$ as well as the explicit formula given in Theorem 6. We see that the term in $O(\cdot)$ is smaller than the main term, so we can simply rewrite

$$\#E(K) \ll c^{\operatorname{rank} E+m} \leq \exp\left(c\frac{\deg N \log q}{\log\deg N}\right).$$

5.1 Comparison to Bombieri-Pila type bound

Let S be the set of all points of bad reduction of an elliptic curve E/K. Consider $h_0 > c\max(\deg\Delta, h(j))$, where Δ is the discriminant and j is the j-invariant of E/K for some constant c. The main contribution to Theorem 4.1 and, respectively, Corollary 4.4 is given by $e^{R\alpha(h_0/R)}$. The minimum in α is attained to the left of $t = 1$. Since $h_0 > c\deg\Delta$, then $\alpha(h_0/R) < (1-\delta_0)h_0/R$, where δ_0 positive and depending only on c. Thus for any $\delta_1 \leq \delta_0$ we obtain a bound

$$\#E(K,S) \ll e^{(1-\delta_1)h_0},$$

while Bombieri-Pila type result brings us to e^{h_0}, thus this method gives an improvement in the exponent and also improves the corresponding results from [9].

Another possible way to get this sort of bounds is using the work of Bhargava et al. on bounding the size of 2-torsion group, see [1]. The authors of [1] proved the first nontrivial bounds on the sizes of 2-torsion subgroups of the class groups of cubic and higher degree number fields. This is also an improvement on the bounds on the number of integral points given in [9]. They also gave a result for the function fields, see [1, Theorem 7.1].

Acknowledgements

The author thanks Marc Hindry for useful remarks on the previous version of this paper. The author would also like to thank Göttingen University for its hospitality while completing presented work.

References

[1] Manjul Bhargava, Arul Shankar, T Taniguchi, F Thorne, J Tsimerman, and Y Zhao. Bounds on 2-torsion in class groups of number fields and integral points on elliptic curves. *preprint arXiv:1701.02458v1*, pages 1–12, 2017.

[2] Enrico Bombieri and Jonathan Pila. The number of integral points on arcs and ovals. *Duke Mathematical Journal*, 3(59):337–357, 1926.

[3] Armand Brumer. The average rank of elliptic curves I. *Inventiones Mathematicae*, 472(1):445–472, 1992.

[4] Robert Gross and Joseph H Silverman. S-integer points on elliptic curves. *Pacific journal of mathematics*, 167(2):263–288, 1995.

[5] L Hajdu and T Herendi. Explicit Bounds for the Solutions of Elliptic Equations with Rational Coefficients. *Journal of Symbolic Computation*, 25:361–366, 1998.

[6] D. R Heath-Brown. The Density of Rational Points on Curves and Surfaces. *Annals of Mathematics*, 155(2):553–595, 2002.

[7] Harald Helfgott. On the square-free sieve. *Acta Arithmetica*, 115(4):349–402, 2004.

[8] Harald Helfgott and Akshay Venkatesh. Integral points on elliptic curves and 3-torsion in class groups. *J. Amer. Math. Soc.*, 19(3):527–550, 2006.

[9] Harald Helfgott and Akshay Venkatesh. How small must ill-distributed sets be? *Analytic number theory*, 2:224–234, 2009.

[10] Marc Hindry and Joseph H Silverman. The canonical height and integral points on elliptic curves. *Inventiones Mathematicae*, 93(2):419–450, 1988.

[11] G Kabatiansky and V Levenstein. On bounds for packings on a sphere and in soace. *Problemy Peredachi Informacii*, 14(1):3–25, 1978.

[12] James Milne. On a conjecture of Artin and Tate. *Annals of Mathematics*, 102:517–533, 1975.

[13] James Milne. *Elliptic Curves*. BookSurge Publishers, 2006.

[14] Amilcar Pacheco. Integral points on elliptic curves over function fields of positive characteristic. *Bull. Austral. Math. Soc.*, 58:353–357, 1998.

[15] Alisa Sedunova. On the Bombieri-Pila Method Over Function Fields. *Arxiv preprint arXiv:1506.08757*, 2015.

[16] Joseph H Silverman. Lower bounds for height functions. *Duke Mathematical Journal*, 51(2):395–403, 1984.

[17] John Tate. On the conjectures of Birch and Swinnerton- Dyer and a geometric analog. *Séminaire N. Bourbaki*, 306:415–440.

[18] Douglas Ulmer. *Elliptic curves over function fields.* Arithmetic of L-functions; 211-280, IAS/Park City Math. Ser., 18, Amer. Math. Soc., Providence, RI, 2011.

[19] Felipe Voloch. Explicit p-descent for elliptic curves in characteristic p. *Compositio Math*, 74(3):247–258, 1990.