On the torsion values for sections of
an elliptic scheme

By *Pietro Corvaja* at Udine, *Julian Demeio* at Pisa, *David Masser* at Basel and
Umberto Zannier at Pisa

Abstract. We shall consider sections of a complex elliptic scheme $\mathcal{E}$ over an affine base curve B, and study the points of B where the section takes a torsion value. In particular, we shall relate the distribution in B of these points with the canonical height of the section, proving an integral formula involving a measure on B coming from the so-called Betti map of the section. We shall show that this measure is the same one which appears in dynamical issues related to the section. This analysis will also involve the *multiplicity* with which a torsion value is attained, which is an independent problem. We shall prove finiteness theorems for the points where the multiplicity is higher than expected. Such multiplicity has also a relation with Diophantine Approximation and quasi-integral points on $\mathcal{E}$ (over the affine ring of B), and in Sections 5 and 6 of the paper we shall exploit this viewpoint, proving an effective result in the spirit of Siegel's theorem on integral points.

1. Introduction

In this paper we deal with a complex elliptic scheme $\mathcal{E} \to B$ over an affine algebraic curve B, and defined for instance by a Weierstrass equation

$$(1.1) \qquad \mathcal{E} : y^2 = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3),$$

where α_i, $i = 1, 2, 3$, lie in $\mathbb{C}(B)$, the origin being the point at infinity.

For $b \in B$, we let $\mathcal{E}_b$ be the fiber over b. We can suppose that the α_i are regular in $\mathbb{C}(B)$ such that the discriminant of the cubic polynomial on the right of (1.1) never vanishes on B (i.e., $(\alpha_i - \alpha_j)$ do not vanish for $i \neq j$), so that $\mathcal{E}_b$ will really be an elliptic curve for each b. We may, in fact, take a smooth completion $\overline{B}$ of B, and extend our scheme to a fibration above the whole $\overline{B}$, where the remaining fibers shall not necessarily be elliptic curves.

In the sequel, we shall denote by $S = \overline{B} \setminus B$ and call it the bad reduction locus (even if some points of S could lead to elliptic fibers).

The corresponding author is Julian Demeio.

We shall disregard the case when this scheme is isotrivial, which is equivalent to the j-invariant of the elliptic curve being constant. Moreover, at some points we shall also suppose that the elliptic scheme is defined over the field of algebraic numbers.

A non-zero section $\sigma : B \to \mathcal{E}$ of this scheme can be written as $\sigma = (x_\sigma, y_\sigma)$, where $x_\sigma, y_\sigma \in \mathbb{C}(B)$ satisfy the above equation.

Our main interest lies in what we call the *torsion points* of σ, namely the set of points of the base B where σ takes a value which is torsion in the appropriate fiber:

$$T_\sigma = \{b \in B : \text{there exists } n > 0 \text{ such that } n \cdot \sigma(b) = 0\}.$$

Of course, one usually refers to torsion points as points of $\mathcal{E}$, however our present terminology should not risk to create any confusion.

1.1. Generalities on T_σ.

T_σ **is large.**   We pause to remark that T_σ is always infinite. When σ is torsion as a section, of course $T_\sigma = B$, and this is a trivial case usually disregarded here, so assume that σ is non-torsion. Then the issue is maybe less obvious then it may seem. Here are some arguments.

Writing $n \cdot \sigma = (x_{n\sigma}, y_{n\sigma})$, a torsion point (in our sense) corresponds to a pole in B of some $x_{n\sigma}$. Saying that T_σ is infinite amounts to the fact that the set of these poles is infinite as n varies. Siegel's theorem over function fields implies this fact and more. Or we can use the *abc* inequality over function fields, after observing that for each $n > 0$ the functions $x_{n\sigma}$ have all poles in T_σ and the functions $x_{n\sigma} - \alpha_i$ have all zeros in T_σ. (See for instance [8] and [29] for a bit more detail. This link with Siegel's theorem will be central in Section 5).

There is also an analytic argument employing the so called Betti map, which will be central in this paper; this gives even the stronger result that the torsion points are dense for the complex topology of B.[1] (See again [8] for more on this, and [29] for a further argument using reduction modulo p.)

A recent paper of DeMarco and Mavraki [9], in the case when $\overline{\mathbb{Q}}$ is a field of definition, proves even a property of Galois equidistribution of such points. Roughly speaking this asserts that there is a probability measure on the base B such that, as the torsion order grows to infinity, the Galois conjugates of a torsion point "tend" to be equidistributed in B with respect to this measure, i.e., every prescribed sufficiently regular region contains a percentage of conjugates approximately equal to its measure. They work even with p-adic valuations, but if we stick to the archimedean case, we shall see later how the measure in this case comes naturally from the Betti map, so their results very well fit into our study.

T_σ **is small.**   In the opposite direction, one may prove that these points are *sparse* in certain meanings. For instance, when the scheme and the sections are defined over $\overline{\mathbb{Q}}$, a theorem of Silverman-Tate predicts bounded height for such points (provided the section is not torsion), hence their degree (over $\mathbb{Q}$) tends to infinity. (See [29], especially Appendix C by the third author, for a self-contained proof.) We may add to this p-adic integrality constraints of the type appearing in the theorem of Lutz and Nagell (see [20, Chapter VIII, Section 7].)

And if we seek points which are torsion simultaneously for another section which is linearly independent with the former, then one may prove their finiteness over the whole alge-

---

[1] When the section is defined over $\mathbb{R}$, this remains true over the real topology. However, as proved in [11], this is not the case for the p-adic topology.

braic closure, i.e., disregarding their degree. See e.g. the paper [6], which extends to $\mathbb{C}$ such results, previously proved over $\overline{\mathbb{Q}}$ in papers quoted therein, and see also [29, Chapter 3]. (These finiteness theorems are cases of the so-called *Pink–Zilber conjecture.*)

1.2. Our issues.

1.2.1. Multiplicities. We have just remarked that the poles of the rational functions $x_{n\sigma}$ for varying n are the elements of T_σ. So, if for instance we are interested in *counting* the elements of T_σ of a given torsion order m, and if we want to do that by comparison with the degree $\deg x_{m\sigma}$, then we must take into account the *multiplicity* with which they appear as poles of $x_{m\sigma}$.

Now, the coordinate function x for a Weierstrass equation of an elliptic curve has a pole of order 2 at infinity, hence the multiplicity for the said poles will be even. But we should not expect multiplicity higher than 2, except for special points.

Thus an issue here is to seek to *bound from above* such multiplicity, and to *discover the distribution* of torsion points on B where high multiplicity occurs.

This will be the first of the problems discussed in the paper. We shall start by describing the multiplicity in different equivalent terms. As stated in Lemma 2.7 below, this multiplicity corresponds to the local multiplicity intersection of the image of the section with a torsion curve (i.e., defined by an equation $np = 0$, $p \in \mathcal{E}$, for an integer $n > 0$).

We shall analyze the multiplicity by means of the Betti map of the section, recalled in Section 2. This leads to finiteness theorems for the points in T_σ where the multiplicity is higher than expected, and also to a bound for the maximal multiplicity. Actually, the analysis will prove a stronger result on letting the section vary through the full Mordell–Weil group of sections.[2]

We also note that a high multiplicity at a point $b \in B$, as a pole of $x_{n\sigma}$, means that $n\sigma$ is near to infinity (i.e., the origin of $\mathcal{E}$) in the valuation of $\mathbb{C}(B)$ associated to b. This fact links the issue with Diophantine Approximation over function fields. Then we shall take this viewpoint in Section 5 (see below for more).

1.2.2. Canonical height of a section, torsion points and the Betti measure. We shall also investigate the distribution of T_σ in regions of B; this will be done by counting asymptotically the points of order n with varying n.

After proving that the multiplicity at these points will be generally what is expected (i.e., 2 as a pole of $x_{n\sigma}$), the counting of points will give essentially the degree of $x_{n\sigma}$, which is asymptotically $n^2 \hat{h}(\sigma)$, where $\hat{h}$ is a canonical height associated to twice the divisor at infinity on $\mathcal{E}$.

But the counting of points can be also done through the Betti map. For given n, we have to count the number of points where the Betti map takes a rational value with denominator (dividing) n. For large n this will be essentially n^2 times the area of the base, with respect to the measure obtained locally by pulling back the Lebesgue measure on $\mathbb{R}^2$ by the Betti map.

---

[2] At the same time as the submission of this paper, D. Ulmer and G. Urzua kindly informed us about their preprint [23] in which, among other things, they proved the finiteness instance in the case of a cyclic group, using somewhat different methods, of more algebraic-geometric nature. Subsequently, some time later, they produced a follow-up paper [24], in which they were able to give very explicit estimates for the total number of multiple points. This estimate often improves on ours (which, by the way, is not explicit), although our estimates work for finitely generated groups and sometimes give other further information.

By comparing these two approaches, we shall also obtain a certain integral formula for the canonical height of the section.

It is to be noted that, beyond the control on multiplicities, we shall need control on the behavior of the Betti map near the points of bad reduction of $\mathcal{E}$ (for instance, $0, 1, \infty$ for the Legendre elliptic scheme). This will be done by using that the Betti map is *definable*, as recently proved by G. Jones and H. Schmidt [10]. (This definability will be used actually also in the previous part of the paper.) To show that the counting can be done through the usual comparison with an area, it will be necessary to use a further result by Barroero and Widmer [3] proving the required asymptotic for definable maps.

We shall also indicate other possible methods for proofs of these results.

This section will further contain an addendum, showing that the measure induced by the Betti map is the same as the one used by DeMarco and Mavraki in the above quoted context. As a byproduct of this verification, we shall obtain a certain characterization of the possible measures which could appear *a priori*.

1.2.3. Effective analysis of integral points on $\mathcal{E}$ and multiplicities again. As mentioned above, the issue of multiplicities for the poles of the rational functions $x_{n\sigma}$, for $n \in \mathbb{N}$, are related to diophantine approximation and quasi integral points in the function field of B. In Section 5 of the paper we will exploit this viewpoint.

We will first prove a version of effective Roth's theorem over function fields, by adapting a method of J. Wang [27]. Then we shall use Siegel's method for quasi-integral points on an elliptic curve, to deduce the sought approximation result.

Concerning multiplicities, this will yield results in a sense much weaker than those coming from the Betti map if applied to varying elements in a prescribed finitely generated group of sections. However these results will be stronger if applied individually to an arbitrary section. Moreover, they will be completely effective and uniform, and admit applications beyond what follows from the former methods. This analysis is accompanied by examples illustrating the conclusions in all details.

Acknowledgement. We thank Ulmer and Urzua for sending us their preprints. We thank Mavraki, Jones and Schmidt for their interest and helpful kind indications concerning their papers. Finally, we thank the anonymous referee for the careful analysis of the manuscript and for suggesting many useful changes.

2. The Betti map

After replacing y with $\frac{y}{2}$ and performing a suitable translation on x, equation (1.1) may be written in strict Weierstrass form:

$$y^2 = 4x^3 - g_2 x - g_3,$$

where g_2 and g_3 are regular functions on B. For every $b \in B(\mathbb{C})$, there is a unique lattice L_b in $\mathbb{C}$ corresponding to the specialization of this Weierstrass equation at B, i.e.,

$$y^2 = 4x^3 - g_2(b)x - g_3(b).$$

Of course, we have that $\mathcal{E}_b = \pi^{-1}(b)$ is analytically isomorphic to the quotient $\mathbb{C}/L_b$.

Let $D \subset B(\mathbb{C})$ be a simply connected domain. We may choose, for $b \in D$, a basis $\rho_1(b), \rho_2(b)$ (called periods) for L_b, where $\rho_1(b), \rho_2(b)$ are holomorphic functions on D, such that $\rho_1(b)/\rho_2(b)$ lies in the upper-half plane $\mathcal{H}$.

Also, for $P \in E_b$, we have a well-defined elliptic logarithm $\log^{\mathrm{ab}}(P)$ of P, uniquely as a point of $\mathbb{C}/L_b$, corresponding to P under the abovementioned analytic isomorphism.

Definition 2.1. For D as above, and $P \in \pi^{-1}(D)$, we call the *Betti coordinates* of P, the unique elements $\beta_1(P), \beta_2(P)$ of $\mathbb{R}/\mathbb{Z}$, such that the following equality holds:

$$\log^{\mathrm{ab}}(P) = \beta_1(P)\rho_1(\pi(P)) + \beta_2(P)\rho_2(\pi(P)).$$

We call the *Betti map* the (real analytic) map $\beta(P) := (\beta_1(P), \beta_2(P))$.

We note that the 1-forms $\mathrm{d}\beta_1$, $\mathrm{d}\beta_2$, which are going to appear commonly in this paper, are well defined on $\pi^{-1}(D)$, for any simply connected open $D \subset B$ (once the lattice bases are chosen as above). Moreover, the 2-form $\mathrm{d}\beta_1 \wedge \mathrm{d}\beta_2$ is well defined on $\mathcal{E}$, i.e., it is independent of the local choice of periods, since the determinant of a change of basis for the lattice is 1.

If we happen to deal with a specific section $\sigma : B \to \mathcal{E}$, with a slight abuse of notation (when there is no risk of confusion), we shall sometimes call, *Betti coordinates* of a point $P \in B$, the Betti coordinates of the point $\sigma(P)$ (after a specific choice of a simply connected $D \subset B$ and basis for L_b, has been done as above). It may be continued to all of $B(\mathbb{C})$, with monodromy transformation that we forget about at the moment. (See e.g. [6] for a simple description sufficient for our task here, and see [2] for much more precise information about the Betti map, in any dimension.)

For a more detailed exposition about the Betti map, see e.g. [6, Section 1.1].

Remark 2.2. The Betti map of a non-torsion section is proved to be generically submersive (see [29, Section 2.5] or [2], [6]); however it is not necessarily surjective to $(\mathbb{R}/\mathbb{Z})^2$. In fact, for every positive integer N there exists always an (algebraic) section of the Legendre scheme such that the Betti map associated to it "misses" all the rational points with denominator dividing N. One such example is any algebraic section σ_N such that $[N] \cdot \sigma_N = \sigma_0$, where, in the standard Weierstrass notation of the Legendre scheme (recalled in Section 3), $\sigma_0(\lambda) = (2, \sqrt{2(2 - \lambda)})$.

Definition 2.3. We define the multiplicity of β at a point $b \in B(\mathbb{C})$ as the minimum order of a partial derivative of β which does not vanish at b. This notion is clearly independent of the local determination for the Betti map.

For completeness, we recall from [6] that the rank of the (differential of the) Betti map (at any point) is always even, hence in our case it is 0 or 2.

We immediately prove a result concerning the multiplicity.

Proposition 2.4. *Notation being as above, assuming that σ is not torsion, the set of $b \in B(\mathbb{C})$ such that the multiplicity at b of the Betti map for σ is > 1, is a finite set.*

Proof. Let us first deal with the points in a given compact region K in $B(\mathbb{C})$, assuming that K is the closure of an open connected subset of B where a determination of the Betti map

β is well defined. If K contained infinitely many points of multiplicity > 1 for β, then K would contain a whole real analytic arc of such points, i.e., where the differential $d\beta$ vanishes. But then β would be constant along this curve. Since the fibers of β are analytic, β then would be constant on K, and hence on B. By Manin's theorem, σ would then be torsion. We conclude that K contains only finitely many of the points in question.

This suffices to cope with any portion of B in the complement of any neighborhood of the boundary. It will suffice then to deal with any small disk around a boundary point $p_0 \in \overline{B} - B$. We can cover this disk with finitely many triangular sectors, with a vertex in p_0, and such that β is well defined on each of the sectors. We now use that on each such sector β is *definable* in the structure $\mathbb{R}_{\mathrm{an, exp}}$. This has been proved in [10] for the Legendre scheme, but then it holds also in our situation, on going to a Legendre model of our scheme (after suitable base-change). Then the derivatives of β are also definable; this entails that the set of multiple points in a triangular sector is also definable, hence either is finite or contains a real-analytic arc; now the above argument again applies. $\square$

Remark 2.5.   If we deal with a scheme defined over $\overline{\mathbb{Q}}$, and if our multiple points are also *torsion points* for σ (which are those primarily of interest in this paper), then another argument is possible, avoiding the definability. Namely, one first remarks that the degree of the field of definition of these points must tend to infinity, e.g. by Silverman's bounded height theorem. Then, again by bounded height, for a given point of "large" degree, its conjugates must fall in positive percentage outside a "small" neighborhood of the boundary points. This allows the argument in the first part of the lemma to apply. (See also [29] for another similar use of bounded height.)

Now we would like to introduce a definition of multiplicity *at torsion points* of σ, associated directly to the section σ rather than to its Betti map.

Definition 2.6.   Let b be a torsion point for σ of order (dividing) n. We define the multiplicity of σ at b, denoted $m_\sigma(b)$, as the local multiplicity of the intersection at $\sigma(b) \in \mathcal{E}$ between the curves $\sigma(B)$ and the torsion divisor $\mathcal{E}[n]$.

A priori this depends as well on the n such that $n\sigma(b) = O$. However the lemma which follows shows that $m_\sigma(b)$ equals the previously considered multiplicity, and in particular does not depend on the said integer n.

Lemma 2.7.   *Let $\sigma : B \to \mathcal{E}$ be a non-torsion section and $b \in T_\sigma$ a torsion point for σ. Then the local multiplicity $m_\sigma(b)$ equals the multiplicity for the Betti map β_σ at the point b.*

Proof.   On multiplying by a non-zero integer the two multiplicities do not change (since multiplication on an elliptic curve is étale, which implies that on an elliptic surface this multiplication is étale over the set of good reduction).

So we reduce to the case when the torsion divisor intersected by σ above b is the image of the zero section.

Let then $t = \frac{x}{y}$ be a local parameter at the (image of the) zero-section (so it is a local parameter at the origin of the fibers). If the section σ has Weierstrass coordinates (x_σ, y_σ), then define $t_\sigma = \frac{x_\sigma}{y_\sigma}$.

Let now $b \in B$ be such that $\sigma(b) = 0 = 0_b$. Then the zero-multiplicity of t_σ at the point b equals the local intersection multiplicity of $\sigma(B)$ with the zero-section (by definition).

Let us denote by m this multiplicity.

We may now express the regular differential $\omega = \frac{\mathrm{d}x}{y}$ as a series

$$\omega = s(t)\mathrm{d}t, \quad s(t) = c_0 + c_1 t + \cdots,$$

where the c_i are functions on B regular outside the bad reduction, and $c_0 \neq 0$ outside the bad reduction, in particular at b.

An elliptic logarithm $\tilde{\sigma}(u)$ of $\sigma(u)$ may be expressed, for u in a suitable neighborhood of b, as

$$\tilde{\sigma}(u) = \int_{0_u}^{\sigma(u)} s_u(t)\,\mathrm{d}t.$$

Here the integral is on a path on the fiber E_u, where we may choose for instance the shortest path (recall that u is supposed to lie near to the given point b, where $\sigma(b) = 0_b$, hence $\sigma(u)$ is near to 0_u).

The path from 0_u to $\sigma(u)$ on E_u corresponds, via the local parameter t restricted to E_u, to a path from $0 \in \mathbb{C}$ to $t(\sigma(u)) = t_\sigma(u)$. Thus we obtain

$$\tilde{\sigma}(u) = \int_0^{t_\sigma(u)} s_u(z)\,\mathrm{d}z.$$

Therefore, since $c_0(b) \neq 0$,

$$|t_\sigma(u)| \ll_b |\tilde{\sigma}(u)| \ll_b |t_\sigma(u)|.$$

Hence,

$$|\tilde{\sigma}(u)| \asymp d(u,b)^m,$$

where $d(\cdot,\cdot)$ is a distance function in a neighborhood of b in $B(\mathbb{C})$.

On the other hand, by definition $\tilde{\sigma}(u) = \beta_1(u)\rho_1(u) + \beta_2(u)\rho_2(u)$.

Conjugating this equation, we have the vector equation

$$\begin{pmatrix} \tilde{\sigma}(u) \\ \overline{\tilde{\sigma}(u)} \end{pmatrix} = \begin{pmatrix} \rho_1(u) & \rho_2(u) \\ \overline{\rho_1(u)} & \overline{\rho_2(u)} \end{pmatrix} \begin{pmatrix} \beta_1(u) \\ \beta_2(u) \end{pmatrix}.$$

Now, the 2×2-matrix is non-singular at $u = b$ and hence is uniformly bounded together with its inverse in a whole neighborhood of b.

This shows that $\|(\beta_1(u), \beta_2(u))\| \asymp |\tilde{\sigma}(u)| \asymp d(u,b)^m$, and now the fact that β_1, β_2 are real analytic, via a Taylor expansion, proves the lemma. □

2.1. Multiplicity. The above lemma gives two equivalent ways to define the multiplicity. In fact, the lemma proves even more, and we have the following list of equivalent definitions:

(1) Local intersection between the image $\sigma(B)$ of the section and the torsion divisor $\mathcal{E}[n]$ (as in the lemma). This notion is purely algebraic.

(2) Consider an elliptic logarithm $\tilde{\sigma}$ of the section; as a function from B to $\mathbb{C}$, it is well defined in a neighborhood of b up to periods. If $b \in B$ is a torsion point for σ of order n, then the function $\tilde{\sigma} - \frac{\omega}{n}$ will vanish at b for a suitable choice of the period ω (well

defined and holomorphic in a neighborhood of b). We may define the multiplicity as the zero multiplicity of $\tilde{\sigma} - \frac{\omega}{n}$ at b. (We can also reduce to the case $n = 1$ by multiplication by n, as in the proof of the lemma.)

(3) Multiplicity of the Betti map at b, in the sense that both components are 0 up to order $\geq m$ as real analytic functions in a neighborhood of b.

(4) Multiplicity of $t_{n\sigma} = \frac{x_{n\sigma}}{y_{n\sigma}}$ at b. Note that $\frac{x}{y}$ is a local parameter at the origin of the generic fiber. This may be expressed also in terms of the valuation $|\cdot|_v$ associated to b, so that the value $-\log|t_{n\sigma}|_v$ is our multiplicity (provided v is normalized so that its value group is $\mathbb{Z}$).

Theorem 2.8. *There is a differential operator Ξ of the second order on B, acting on local holomorphic functions, such that is σ is a section and $\tilde{\sigma}$ a local determination of its logarithm, then $\Xi(\tilde{\sigma}) = 0$ if and only if σ is torsion.*

Let now σ be non-torsion. For a torsion point b for σ, $m_\sigma(b) \leq 2 + \max(0, \operatorname{ord}_b(\Xi(\tilde{\sigma})))$. In particular, this multiplicity does not exceed a certain explicitly computable function of $\hat{h}(\sigma)$.

Also, there are only finitely many torsion points for σ where $m_\sigma(b) > 1$.

Proof. We follow Manin [12] (see also [8, Section 6.3]), and work on the Legendre scheme, as we may after a base change. Let as above $\tilde{\sigma}$ denote a local determination of an elliptic logarithm of σ, and let Ξ denote the usual Legendre–Gauss differential operator

$$\Xi = 4\lambda(1 - \lambda)\frac{\mathrm{d}^2}{\mathrm{d}\lambda^2} + 4(1 - 2\lambda)\frac{\mathrm{d}}{\mathrm{d}\lambda} - 1,$$

where λ is a suitable rational function on B.

Since Ξ annihilates the periods, it follows that $\Xi(\tilde{\sigma})$ is a well-defined function on B. By easy growth estimates on the coordinates of σ we see that this function has no essential singularities on a complete model of B, and thus is a rational function on B (which can be seen also directly: see equation (6.3) below, which corrects (2) in Manin's paper).

By condition (2) above on the multiplicity, this function $\Xi(\tilde{\sigma})$ vanishes at b with multiplicity at least $-2 +$ the multiplicity of σ at b.

This proves the first assertion, while the rest is a consequence of Proposition 2.4.   □

Remark 2.9. Using Gabrielov–Hardt's theorem (see [4, Theorem 3.14]), one can give a uniform estimate for the cardinality of torsion points for σ where $m_\sigma(b) > 1$, when σ varies in a finitely generated group. However in the next theorem we shall achieve a still stronger result.

Theorem 2.10. *Let Γ be a finitely generated torsion-free group of sections $B \to \mathcal{E}$, defined over $\overline{\mathbb{Q}}$. For $\sigma \in \Gamma$, set $M_\sigma := \{b \in B : \sigma(b) \text{ is torsion}, m_\sigma(b) > 1\}$. Then the union $\bigcup_{\sigma \in \Gamma - \{0\}} M_\sigma$ is finite.*

Proof. Let $\sigma_1, \ldots, \sigma_r$ be a basis for Γ; consider the respective Betti maps $\beta_1, \ldots, \beta_r$, on some domain where they are well defined. (In those domains we make some definite choice of these maps, up to integers.)

As in the previous proofs, we may cover B with finitely many such domains (either compact or triangular regions with a vertex at a point in $\partial B = \overline{B} - B$ of bad reduction).

A (special case of a) theorem of André (see [1, Theorem 3]) ensures that the β_i are linearly independent over $\mathbb{R}$ modulo constant functions (since the sections σ_i are linearly independent).

We consider linear combinations $\sum x_i \beta_i$ with real $x_i \in \mathbb{R}$. Among these linear combinations, those with rational coefficients x_i correspond to Betti maps (up to the addition of real constants) of elements in the division group of Γ.

Let us work now in one of the above domains, call it D. We have a map from $\mathbb{R}^r \times D$ to $\mathrm{Mat}_2(\mathbb{R})$, sending $(x_1, \ldots, x_r, p) \mapsto \sum x_i \mathrm{d}\beta_i(p)$, where $\mathrm{d}\beta_i$ denotes the jacobian matrix of β_i (with respect to some chosen coordinates in D). This is a definable map, by the results that we have quoted in the above proofs. Therefore the set of its zeros forms a definable subvariety Z of the domain.

Note that if σ is a section corresponding to a rational point $\mathbf{v} = (c_1, \ldots, c_r) \in \mathbb{Q}^r - \{0\}$, the intersection of Z with the fiber above $\mathbf{v}$ in D is finite by Proposition 2.4 (since the sections are independent and $\mathbf{v} \neq 0$). However this fiber contains the set M_σ. Also, we may restrict $\mathbf{v}$ to the closed unit cube by linearity. Now, by Gabrielov–Hardt's theorem, the number of connected components of the fibers is uniformly bounded, hence $|M_\sigma|$ is uniformly bounded as well.

Now, note that for a given $\sigma \in \Gamma$ the set M_σ is stable by Galois conjugation over a number field of definition, hence by what has been proved above the degrees of the involved points are bounded independently of σ. On the other hand, these points are torsion for σ, hence of bounded height by Silverman theorem, and the conclusion follows. □

Remark 2.11. We have argued for sections defined over $\overline{\mathbb{Q}}$. However the result holds true for any ground field of characteristic 0. A proof comes from specialization, as in the paper [6], although here the specialization argument is easier than in [6]. We give here a sketch of the argument.

Note that we can always reduce to a covering of the Legendre scheme $\mathcal{E} \to B$, where $B \to \mathbb{P}_1 - \{0, 1, \infty\}$ is defined over $\overline{\mathbb{Q}}$, but we consider now a finitely generated group of sections which are defined over a transcendental extension K of $\overline{\mathbb{Q}}$. Suppose for simplicity that K has transcendence degree 1 over $\overline{\mathbb{Q}}$, so it is a finite algebraic extension of $\overline{\mathbb{Q}}(t)$. Geometrically, the field $K(\lambda)$ corresponds to a surface S defined over $\overline{\mathbb{Q}}$, endowed with a projection to B, on which λ is a rational function. The function field of S is an algebraic extension of $\overline{\mathbb{Q}}(t, \lambda)$ for some rational function $t \in \overline{\mathbb{Q}}(S)$. The elliptic scheme can be viewed as a scheme over S, and S can be taken to be affine and such that the "bad reduction" is confined to its points at infinity. A torsion point $b \in B$ for a section $\sigma : B \to \mathcal{E}$ is necessarily algebraic over K; hence it can be viewed in S by defining λ and the x-coordinate (and consequently the y-coordinate) by a certain algebraic function of t. Geometrically, this corresponds to an algebraic relation $\varphi_b(x, \lambda) = 0$, which in turn provides an algebraic curve in S; we call it a torsion curve.

The idea now is to cut the surface S with a "generic" curve Y defined over $\overline{\mathbb{Q}}$ and to consider the elliptic scheme over such a curve. The finiteness result just proved over $\overline{\mathbb{Q}}$ ought to imply the corresponding finiteness statement for our original sections, which are not defined over $\overline{\mathbb{Q}}$, but this is not obvious and further arguments are needed, as we now explain. The problem, which was the main obstacle in the specialization procedure carried out in [6], is that such a curve Y might avoid all but finitely many torsion curves; hence any finiteness result for the intersection of Y with the torsion curves of a certain type (namely those with $m_\sigma(b) > 1$) would be meaningless. As noted in [6], on any affine surface S defined over $\overline{Q}$ one can construct a sequence of algebraic curves $B_1, B_2, \ldots$ such that each algebraic curve

$Y \subset S$ defined over $\bar{\mathbb{Q}}$ avoids all but finitely many of them (in the sense that it will intersect them only at infinity). However, such a sequence of curves $B_1, B_2, \ldots$ would necessarily have a degree tending to infinity. This is the content of the following claim:

Claim. *Let $\bar{S}$ be a projective algebraic surface defined over $\bar{\mathbb{Q}}$, $S \subset \bar{S}$ an open non-empty affine subset. Let $H = \bar{S} - S$ be the divisor at infinity of S. Let $B_1, B_2, \ldots$ be a sequence of irreducible curves in S such that the intersection product $\bar{B}_i \cdot H$ of their closure with H is uniformly bounded. Then there exists a curve $Y \subset S$ defined over $\bar{\mathbb{Q}}$ such that Y intersects infinitely many of the curves B_i.*

Proof of the Claim. Let $d \geq 1$ be an upper bound for the intersection product $B_i \cdot H$; choose $d+1$ curves $Y_1, \ldots, Y_{d+1}$ in S such that for every pair (j, h) with $1 \leq j < h \leq d+1$, the complete curves $\bar{Y}_j, \bar{Y}_h$ do not intersect at infinity. Then for each i, the complete curve $\bar{B}_i$ can intersect at most d of the $\bar{Y}_j$ at infinity, so it must intersect at least one of them in S. It follows that for at least one index j, the curve Y_j intersects infinitely many of the irreducible curves B_i. □

Now, in our situation, we claim that the torsion curves in question are given by equations of the form $\varphi_b(x, \lambda) = 0$, where both partial degrees of the polynomials φ_b are bounded. The degree in x corresponds to the functional height of λ, which is bounded by a functional version of Silverman's theorem since the point is torsion. Reciprocally, the degree in λ is bounded as before by an application of Gabrielov–Hardt's theorem; this time we have to use the fact that our point is not only torsion, but satisfies $m_\sigma(b) > 1$.

It follows that the torsion curves we have to examine have bounded degree with respect to the divisors at infinity $H = \bar{S} - S$, so we can apply the previous claim and conclude that for one curve Y in S, Y intersects infinitely many torsion curves $b(x, \lambda) = 0$ with $m_\sigma(b) > 1$, and the theorem already proved gives the desired contradiction.

3. Canonical height and Betti map

Throughout this section, $\pi : \mathcal{E} \to \mathbb{P}_1 \setminus \{0, 1, \infty\}$ denotes the Legendre scheme, i.e., the elliptic surface associated to the following elliptic curve, defined over the function field $\mathbb{C}(\lambda)$ of $\mathbb{P}_1$:

$$(3.1) \qquad\qquad y^2 = x(x - 1)(x - \lambda).$$

Let $S = \{0, 1, \infty\}$ be the set of points of bad reduction for π. In the future we will also use the notation $\mathcal{S} = \pi^{-1}(S)$.

In this article, we frequently work with the abelian logarithm of an elliptic curve $X/\mathbb{C}$. We recall, as above, that, for this to be defined, a lattice $\Lambda_X \subset \mathbb{C}$, such that $\mathbb{C}/\Lambda_X \cong X(\mathbb{C})$ has to be chosen.

Remark 3.1. As remarked also above, during the definition of the Betti map, when a Weierstrass equation $y^2 = 4x^3 - ux - v$ for X is chosen, there exists a unique such lattice Λ_X that satisfies the extra condition $g_2(\Lambda_X) = u, g_3(\Lambda_X) = v$. When we work with a specific Legendre equation as (3.1), we obtain a corresponding Weierstrass equation as above after the

substitution $x \mapsto x + \frac{1}{3}(1 + \lambda)$ and $y \mapsto \frac{y}{2}$, and assume that the above canonical choice of a lattice for the abelian logarithm has been made.

3.1. Height as an integral and consequences.

The following is the main result of this section. The result was not present in the literature in the following form, though it can be deduced from the work of DeMarco and Mavraki [9] using Wirtinger's formula: we sketch this argument at the end of the following section. However, the proof presented here seems to be different and more elementary in nature.

As above, let $\pi : \mathcal{E} \to \mathbb{P}_1 \setminus \{0, 1, \infty\}$ be the Legendre scheme. Let $\overline{B}$ be a complete complex curve, $r : \overline{B} \to \mathbb{P}_1$ a finite morphism, and $B := \overline{B} \setminus r^{-1}(S)$. Let $\mathcal{E}_B := \mathcal{E} \times_{(\pi,r)} B$, and let $\pi_B : \mathcal{E}_B \to B$ be the projection on B. Let $\sigma : B \to \mathcal{E}_B := \mathcal{E} \times_{(\pi,r)} B$ be an algebraic section of $\pi_B : \mathcal{E}_B \to B$.

Theorem 3.2. *The following equality holds:*

$$(3.2) \qquad \hat{h}(\sigma) = \int_{B(\mathbb{C})} \sigma^*(\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1),$$

where (β_1, β_2) is a (local branch) of the Betti map on $\mathcal{E}_B$.

Moreover, we will see, as part of the proof of Theorem 3.2, that the set T_σ introduced at the beginning of this paper is distributed as the measure associated to the 2-form $\sigma^*(\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1)$.

Remark 3.3. Note that Theorem 3.2 implies that, when σ is not torsion, the Betti map $(\beta_1, \beta_2) : B(\mathbb{C}) \to (\mathbb{R}/\mathbb{Z})^2$ is submersive (i.e. its real Jacobian matrix is invertible) in at least one point (and, hence, on an open dense subset of $B(\mathbb{C})$, as the zero-set of the determinant of the Jacobian is a closed real-analytic subvariety).

We immediately notice the following corollary.

Corollary 3.4. *Let $\sigma : B \to \mathcal{E}_B$ denote a section of the elliptic surface $\mathcal{E}_B \to B$. Then the integral $\int_B \sigma^*(\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1)$ has a rational value.*

Proof. We have by [18, Section 11.8] that $\hat{h}(\sigma) \in \mathbb{Q}$. Hence, Corollary 3.4 is an immediate consequence of Theorem 3.2.   □

Before coming to the proof of Theorem 3.2, we give here another expression for the $(1, 1)$-form $\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1 \in \Omega^2(\mathcal{E}_B)$ that might be more suitable for calculations. To fix some notation, let Λ_λ be a local choice of a lattice in $\mathbb{C}$ corresponding to the elliptic curve

$$\mathcal{E}_\lambda = \pi^{-1}(\lambda)$$

with Weierstrass form $y^2 = x(x - 1)(x - \lambda)$ (see Remark 3.1), and let $\rho_1(\lambda), \rho_2(\lambda) \in \mathbb{C}$ be a (continuous, hence analytic) local choice of periods. Moreover, let

$$d(\lambda) := \rho_1(\lambda)\overline{\rho_2(\lambda)} - \rho_2(\lambda)\overline{\rho_1(\lambda)} = 2i V(\lambda),$$

where $V(\lambda)$ denotes the area of the fundamental domain of Λ_λ. Let $\eta_\lambda(\zeta)$ denote the linear extension to $\mathbb{C}$ of the quasi-period function η_λ (which is defined on the lattice Λ_λ as

in [21, Proposition VI.3.1]), and let $\eta_i(\lambda) = \eta_\lambda(\rho_i(\lambda))$, $i = 1, 2$. Then the following equality, which is proven in Remark 4.2 below, holds (here we are denoting, with a slight abuse of notation, by $\beta_i := \beta_i(\lambda)$ the Betti coordinates $\beta_i(\sigma(\lambda))$):

$$(3.3) \qquad \mathrm{d}\beta_2(\lambda) \wedge \mathrm{d}\beta_1(\lambda) = \frac{i}{2V(\lambda)}\omega \wedge \bar\omega,$$

where

$$\omega = \omega(\lambda) := \left[-\frac{(\eta_1\bar\rho_2 - \eta_2\bar\rho_1)z + 2\pi i\bar z}{2(\lambda^2 - \lambda)} - \frac{2\lambda - 1}{6(\lambda^2 - \lambda)}z\right]\mathrm{d}\lambda + \mathrm{d}z$$

and $z := z(\lambda) := \log_\lambda^{\mathrm{ab}}(\sigma(\lambda))$.

Example 3.5. Let us do an explicit computation of the terms of (3.2), in the case of a specific section, for instance:

$$\sigma(\lambda) = (2, \sqrt{2(2 - \lambda)}).$$

This section is not defined over the base curve $\mathbb{P}_1$ (because of the sign choice). It is, however, well defined as a section of the elliptic surface obtained by base-change along the quadratic cover defined by the equation $t^2 = 2(2 - \lambda)$. We let $r : \bar B \to \mathbb{P}_1$ be the (smooth, complete) cover of $\mathbb{P}_1$ that corresponds to this equation, and $B := \bar B \setminus r^{-1}(S)$, consistently with our notation for this section.

One can compute $\hat h(\sigma)$ explicitly, by using the intersection product on a complete regular minimal model $\overline{\mathcal{E}_B} \to \bar B$ of $\mathcal{E}_B$ (see [18, Section 11.8][3] and [21, Section III.8, especially p. 244]). A simple application of Tate's algorithm reveals that $\overline{\mathcal{E}_B} \to \bar B$ has five singular fibers, four of which are of type I_2 (the ones over $\lambda = 0, 1$) and one of type I_4 (the one over $\lambda = \infty$). Looking at the intersection of the section σ with the singular fibers reveals that the canonical height $\hat h(\sigma) = \frac{1}{2}$.

To make the calculation explicit, we will think of $\bar B$ as being a union of 2 projective plane charts (which we denote by $\bar B_+, \bar B_-$) that are united through the segment $[2, \infty]$, with the topology that switches chart when "we cross the segment". We use the rational function λ as a parameter on these charts.

Let us note that the $(1, 1)$-form $\sigma^*(\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1) \in \Omega^2(B)$ is invariant by the involution $\iota : B \to B$, $t \mapsto -t$ (since $\sigma \circ \iota = -\sigma$). Hence, the form $\sigma^*(\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1) \in \Omega^2(B)$ is actually a pullback along r of a form, which we denote by $(\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1)(\lambda)$, defined over $\mathbb{P}_1 \setminus S$. The $(1, 1)$-form $(\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1)(\lambda) \in \Omega^2(\mathbb{P}_1 \setminus S)$ is equal to (3.3) above, with

$$z = z(\lambda) = \int_\infty^2 \mathrm{d}x/(\pm\sqrt{x(x-1)(x-\lambda)})$$

(the determination chosen for the square root is irrelevant for our purposes).

Hence, we deduce the following integral identity from Theorem 3.2:

$$\frac{1}{4} = \frac{1}{2} \cdot \hat h(\sigma) = \frac{1}{2} \cdot \int_B \mathrm{d}(\beta_2 \circ \sigma) \wedge \mathrm{d}(\beta_1 \circ \sigma) = \int_{\lambda \in \mathbb{C}} \mathrm{d}\beta_2(\lambda) \wedge \mathrm{d}\beta_1(\lambda).$$

---

[3] Our normalization of the height function is $\frac{1}{2}$ with respect to that of [18].

3.2. Proof of Theorem 3.2. We start with a premise to the reader. In this proof, we are going to use concepts of o-minimality. To explain why we are doing so, let us first recall that the main idea of the proof will be to study the quantity of torsion points of σ of order n, and show that, as $n \to \infty$, this number, divided by n^2, compares to both the left-hand side and the right-hand side of (3.2). The reason we decided to rely here on the machinery of o-minimality is because studying the asymptotics for torsion points near the singular fibers turns out to be quite tricky (on the other hand, studying the asymptotics away from the bad reduction locus is quite easy). Indeed, *a priori*, the Betti map could behave quite nastily on the bad fibers. Luckily, this map is definable in an o-minimal structure [10], and combining this with its boundedness (proved by Jones and Schmidt in [10]), provides the necessary tools to still obtain quite easily the sought asymptotic. The reader that is not familiar with o-minimality concepts (such as definability) may just think of them as a "black box", indeed we always mention explicitly what the consequences that we need of this theory are. Moreover, in Remark 3.8, we describe a second proof that avoids entirely the use of o-minimal theory (and uses only the easier-to-prove part of the asymptotics described above, i.e., the one away from the bad reduction), and in Section 4.1, we show, as mentioned above, how our result can be obtained also a consequence of a theorem of DeMarco and Mavraki [9].

We recall the following lemma, which follows directly from [3, Theorem 1.3], and that will be used in the proof:

Lemma 3.6. *Let $X \subset \mathbb{R}^k$ be a bounded subset of $\mathbb{R}^k$ definable in an o-minimal structure (note that this implies that X is Lebesgue-measurable). Let, for each $n > 0$,*

$$\alpha_n := \#\left\{ \left(\frac{a_1}{n}, \ldots, \frac{a_k}{n} \right) \in X : a_1, \ldots, a_k \in \mathbb{Z} \right\}.$$

Then the limit $\lim_{n\to\infty} n^{-k}\alpha_n$ exists and is equal to $\mu(X)$, the Lebesgue measure of X.[4]

Proof of Theorem 3.2. Let us first consider the case where σ is torsion. Since the Betti map of a torsion section is constant, in this case, left- and right-hand side of (3.2) are both equal to 0.

We restrict now to the case where σ is not torsion, and consider, for each $n \geq 1$ and subset $U \subseteq B(\mathbb{C})$ (which will usually be open), the following quantity:

$$A_n(U) := \#\{t \in U : \sigma(t) \text{ is } n\text{-torsion in } \mathcal{E}_t\}.$$

We set $A_n := A_n(B(\mathbb{C}))$. We notice that, since σ is not torsion, this quantity is finite for each $n \geq 1$. We claim the following (which obviously implies the thesis):

(a) The limit $\lim_{n\to\infty} \frac{A_n}{n^2}$ exists, is finite, and $\lim_{n\to\infty} \frac{A_n}{n^2} = \hat{h}(\sigma)$.

(b) The limit $\lim_{n\to\infty} \frac{A_n}{n^2} = \int_B \sigma^*(d\beta_2 \wedge d\beta_1)$.

Let us first prove (a). Let $\overline{\pi_B} : \overline{\mathcal{E}_B} \to \bar{B}$ be a regular complete minimal model of $\mathcal{E}_B \to B$. We denote by $\bar{\sigma} : \bar{B} \to \overline{\mathcal{E}_B}$ the completion of the section σ. We know that (see e.g. [21, Section III.9] or [19, Chapters 2 and 3])

$$\hat{h}(\sigma) = \lim_{n\to\infty} \frac{\langle n\bar{\sigma}, O \rangle}{n^2},$$

---

[4] Actually, [3, Theorem 1.3] also provides the following error bound: $\alpha_n = n^k \mu(X) + O(n^{k-1})$.

where $n\bar{\sigma}$ denotes, with a slight abuse of notation, the graph of the section $n\bar{\sigma}$, O denotes the zero section of $\overline{\pi_B}$, and $\langle n\bar{\sigma}, O\rangle$ denotes the intersection product in $\overline{\mathcal{E}_B}$. We write

$$\langle n\bar{\sigma}, O\rangle = A_n + \delta_n + s_n,$$

where s_n denotes the intersection of $n\bar{\sigma}$ and O on the bad fibers of $\overline{\pi_B} : \overline{\mathcal{E}_B} \to \bar{B}$ (those lying over $r^{-1}(S)$), and δ_n is a correction term that keeps track of the intersection that happens with multiplicity greater than 1. That is,

$$\delta_n = \sum_{t \in B}(\mathrm{ord}_t(n\sigma) - 1),$$

where $\mathrm{ord}_t(n\sigma)$ denotes the multiplicity of intersection of $n\sigma$ and O at t. We will now prove that $\delta_n + s_n = O(1)$ as $n \to \infty$. Note that this will conclude the proof of (a).

We prove first that $s_n = O(1)$. To do so, it suffices to show that, for a point $s \in r^{-1}(S)$, $\langle n\bar{\sigma}, O\rangle_s$ (i.e., the local multiplicity of intersection) is bounded in $n > 0$. Let $s \in r^{-1}(S)$, and let m be the least positive integer such that $(m\sigma)(s) = 0$ on the singular fiber $\mathcal{E}_s := \overline{\pi_B}^{-1}(s)$. Since $\langle n\bar{\sigma}, O\rangle_s$ can be positive only when n is a multiple of m, we may assume without loss of generality (replacing, if necessary, $\bar{\sigma}$ by $m\bar{\sigma}$), that $m = 1$, i.e., that $\bar{\sigma}(s) = 0$.

Let now λ denote a local parameter for $s \in \bar{B}$. Let K denote the completed field of Laurent series $\mathbb{C}\{\{\lambda\}\}$, and let O_K denote the ring of integers of K. Let

$$\mathcal{E}_1(K) := \{P \in \mathcal{E}(K) : P(s) = 0\}.$$

We recall that $\mathcal{E}_1(K)$ has the structure of a Lie group over the local field K, given by the restriction of the sum operation on the elliptic curve $\mathcal{E}/K$. The Lie group $\mathcal{E}_1(K)$ is isomorphic to a Lie group $(\lambda O_K, \tilde{+})$, in a way that we briefly recall now.

We recall from [20, Chapter IV] that the formal group on the elliptic curve $\mathcal{E}/K$ is the unique formal power series

$$F(+)(X, Y) \in K[[z_1, z_2]]$$

such that

$$F(+)(z(P), z(Q)) = z(P + Q), \quad P, Q \in \mathcal{E},$$

as *formal* power series, where $z(P) := \frac{x(P)}{y(P)}$ denotes the z-coordinate in a given minimal Weierstrass model for $\mathcal{E}/K$.

We recall, moreover, that the following holds:

$$F(+)(z_1, z_2) = z_1 + z_2 + O(z^2).$$

One can verify that, when an integral Weierstrass model is chosen, $F(+) \in O_K[[z_1, z_2]]$. Hence, $F(+)(X, Y)$ converges when $X, Y \in \lambda O_K$. One can then see that the Lie group $\mathcal{E}_1(K)$ is isomorphic to the Lie group $(\lambda O_K, \tilde{+})$, where

$$\tilde{+}(z_1, z_2) := F(+)(z_1, z_2)$$

(see e.g. [20, Proposition VII.2.2]).

It follows that for $P \in \mathcal{E}(K)$, denoting by $z(P)$ the z-coordinate in a minimal Weierstrass model, one has that

$$(3.4) \qquad\qquad z(n\sigma) = nz(\sigma) + O(z^2) \quad \text{for all } n \in \mathbb{Z}.$$

Moreover, for any $P \in \mathcal{E}_1(K)$, we have $\mathrm{ord}_s(P) = \mathrm{ord}_\lambda(z(P))$ (since z is a local parameter for the 0-section O). Hence, $\mathrm{ord}_s(n\bar{\sigma}) = \mathrm{ord}_\lambda(z(n\bar{\sigma})) = \mathrm{ord}_\lambda(z(\bar{\sigma})) = \mathrm{ord}_s(\bar{\sigma})$, where the middle equality follows from (3.4). Thus, we have $s_n = O(1)$.

We prove now that δ_n is uniformly bounded (for any $n > 0$). We notice that Lemma 2.7 implies that, in order for the intersection of $n\sigma$ and O to be of order greater than 1 at a certain point $t \in B$, the differential of (a local branch of) the Betti map $(\beta_1, \beta_2) \circ \sigma$ would have to be 0 at the point t. Because of Proposition 2.4, this happens only for finitely many base points t in the base. We denote them by $t_1, \ldots, t_k \in B$. Moreover, by the discussion of Section 2.1 following Lemma 2.7, we have that, for every $n > 0$ such that $n\sigma$ and O intersect over t, $\langle n\sigma, O \rangle_t = \mathrm{ord}_t(x_{n\sigma}/y_{n\sigma})$ is equal to the multiplicity of the Betti map of σ at t (see Definition 2.3), and, in particular, does not depend on n (see also the first paragraph of the proof of Lemma 2.7). This concludes the proof of (a).

We prove now point (b). As mentioned above, let us first present a straightforward (and elementary) proof for the asymptotic away from the singular fibers:

$$(3.5) \qquad \lim_{n \to \infty} \frac{A_n(B \setminus r^{-1}(S_\varepsilon))}{n^2} = \int_{B(\mathbb{C}) \setminus r^{-1}(S_\varepsilon)} \sigma^*(\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1),$$

where S_ε is an open neighborhood of radius[5] ε of S.

Indeed, let $B(\mathbb{C}) \setminus r^{-1}(S_\varepsilon) = \overline{D_1} \cup \cdots \cup \overline{D_k}$ be a decomposition where the sets D_i are open (the bar represents their closure) whose borders are piece-wise real analytic closed simple curves such that $D_i \cap D_j = \emptyset$ when $i \neq j$, and such that on each of the D_i there is a well-defined (lift to $\mathbb{R}^2$ of a) local branch of the Betti map of σ. Note that, for $t \in B(\mathbb{C})$, $\sigma(t) \in \mathcal{E}_t$ is n-torsion if and only if the Betti coordinates $(\beta_1(\sigma(t)), \beta_2(\sigma(t)))$ are rational with denominator dividing n. Hence the identity (3.5) follows at once from an application of the following lemma to the maps

$$\beta' := \beta \circ \sigma|_{D_i} : D_i \to \mathbb{R}^2.$$

Lemma 3.7. *Let $U \subseteq \mathbb{R}^2$ be a bounded open subset, and let λ be the Lebesgue measure on $\mathbb{R}^2$. Let $b : \overline{U} \to \mathbb{R}^2$ be a function which is the restriction of a C^1 map from a neighborhood V of $\overline{U}$. Assume that $\lambda(\partial U) = 0$. Then*

$$\int_U b^*(\mathrm{d}x \wedge \mathrm{d}y) = (b^*\lambda)(U) = \lim_{n \to \infty} \sum_{\frac{a}{n}, a \in \mathbb{Z}^2} \#b^{-1}\left(\frac{a}{n}\right)$$

$$= \lim_{n \to \infty} \sum_{\frac{a}{n}, a \in \mathbb{Z}^2} \#\left\{b^{-1}\left(\frac{a}{n}\right) \cap U\right\}.$$

Proof. The lemma easily follows by applying the Lebesgue–Vitali Theorem[6] to the function $\mathbb{R}^2 \ni x \mapsto \#b^{-1}(x) \in \mathbb{N}$ (the first equality is standard, the second equality follows by applying the theorem on the domain $\overline{U}$, while the third follows by applying it on U). Indeed, let b' the map $V \to \mathbb{R}^2$ of which b is the restriction, and let $R \subseteq V$ be the finite set

---

[5] With respect to some choice of a metric, which is irrelevant for our purposes. For instance, one may choose the Fubini-Study metric on $\mathbb{P}_1(\mathbb{C})$.

[6] That is, a bounded function $\mathbb{R}^k \to \mathbb{R}$, with bounded support, is Riemann integrable if and only if it continuous outside a set of Lebesgue measure 0.

of points where $\mathrm{d}b'$ is not invertible. Note that the function $x \mapsto \#b^{-1}(x)$ is locally constant, hence continuous, outside $b(\partial(U) \cup R)$, and that this set has Lebesgue measure 0, because the C^1-hypothesis implies that b' is Lipschitz (say with constant L) on $\overline{U}$, which in turn implies that $\lambda(b(\partial(U) \cup R)) = \lambda(b(\partial(U))) \leq L^2 \lambda(\partial(U)) = 0$ is 0. $\square$

To deal with the neighborhood of $r^{-1}(S)$, we use the fact that the Betti map is (locally) definable and bounded (by definable, we will always mean definable in $(\mathbb{R}_{\mathrm{an}, \exp})$).

Namely, it is proven in [10, Section 10] that there exists a (definable) partition

$$r^{-1}(S_\varepsilon \setminus S) = Y_1 \cup \cdots \cup Y_l,$$

and on each Y_j there exists a well-defined branch of (a lift to $\mathbb{R}^2$ of) the Betti map, which we will denote by $B^j = (\beta_1^j, \beta_2^j) \circ \sigma : Y_j \to \mathbb{R}^2$, such that B^j is definable, and in [10, Proposition 4] that it is bounded.

Let us consider now, for each $j = 1, \ldots, l$, the following definable set in $Y_j \times \mathbb{R}^2$:

$$X_j := \{(y, t) \in Y_j \times \mathbb{R}^2 : y \in Y_j \text{ and } t = B^j(y)\},$$

i.e., X_j is the graph of B^j. Applying Gabrielov–Hardt's theorem (as formulated in [25, Theorem 9.1.2]) to the second projection $X_j \to \mathbb{R}^2$, we know that there exists a finite partition of $\mathbb{R}^2$, say

$$A_j^1 \cup \cdots \cup A_j^{d(j)} = \mathbb{R}^2,$$

such that X_j is definably trivial over each A_j^m. Recall that this means that, for each $m \leq d(j)$, there exist a definable set F_j^m and a (definable) isomorphism

$$h_{A_j^m} : X_j \cap (Y_j \times A_j^m) \to F_j^m \times A_j^m$$

that commutes with the projection to A_j^m. Moreover, since the Betti map is bounded on each Y_j, we may assume that, for each $m \leq d(j)$, either A_j^m is bounded, or $X_j \cap (A_j^m \times \mathbb{R}^2) = \emptyset$.

We define now, for each $j = 1, \ldots, l$, $m \leq d(j)$,

$$f_j^m := \#F_j^m.$$

We notice that f_j^m is always finite. In fact, this is a direct consequence of the fact that the fibers of (any local branch of) the Betti map of σ are isolated points (see [6, Proposition 1.1]), and that these fibers are definable (since so is the Betti map), and definable sets made of isolated points are finite. Then the following equality holds:

$$(3.6) \qquad \sum_{m \leq d(j)} f_j^m \mu(A_j^m) = \int_{r^{-1}(S_\varepsilon \setminus S)} \sigma^*(\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1).$$

In fact, notice that both left- and right-hand side of (3.6) are equal to the measure of the graph of the section $\sigma|_{r^{-1}(S_\varepsilon \setminus S)} : r^{-1}(S_\varepsilon \setminus S) \to \mathcal{E}_B$, given by the integration of the restriction of the $(1, 1)$-form $\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1 \in \Omega^2 \mathcal{E}_B$.

Remember that, if $t \in B(\mathbb{C})$, then $\sigma(t) \in \mathcal{E}_t$ is n-torsion if and only if the Betti coordinates $(\beta_1(\sigma(t)), \beta_2(\sigma(t)))$ are rational with denominator dividing n. Hence, the following equality holds, for each $n > 0$:

$$(3.7) \quad A_n(r^{-1}(S_\varepsilon \setminus S)) = \sum_{m \leq d(j)} f_j^m \alpha_{n,j}^m, \quad \alpha_{n,j}^m := \left\{ \left(\frac{a}{n}, \frac{b}{n} \right) \in A_j^m : a, b \in \mathbb{Z} \right\}.$$

Hence

$$\frac{A_n(r^{-1}(S_\varepsilon \setminus S))}{n^2} = \sum_{m \leq d(j)} f_j^m \frac{\alpha_{n,j}^m}{n^2}.$$

Letting $n \to \infty$, and combining (3.6), (3.7) with Lemma 3.6, we obtain

$$(3.8) \qquad \lim_{n \to \infty} \frac{A_n(r^{-1}(S_\varepsilon \setminus S))}{n^2} = \int_{r^{-1}(S_\varepsilon \setminus S)} \sigma^*(\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1).$$

Combining this with (3.5),[7] concludes the proof of (b), and hence the theorem. □

Remark 3.8. To prove Theorem 3.2, one may also use a different argument. Namely, one first proves equality (3.2) when $r : \bar{B} \to \mathbb{P}_1$ is defined over $\bar{\mathbb{Q}}$ using a height argument, and then uses a continuity argument to prove it for every r. We give here a sketch of these two steps.

Assume that $r : \bar{B} \to \mathbb{P}_1$ is defined over a number field K. As presented in the proof of Theorem 3.2, one can prove in a straightforward way that

$$(3.9) \qquad \lim_{n \to \infty} \frac{A_n^\varepsilon}{n^2} = \int_{B(\mathbb{C}) \setminus r^{-1}(S_\varepsilon)} \sigma^*(\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1),$$

where S_ε is a neighborhood of radius ε of S, and

$$A_n^\varepsilon := A_n(B(\mathbb{C}) \setminus r^{-1}(S_\varepsilon)) = \#\{t \in B(\mathbb{C}) \setminus r^{-1}(S_\varepsilon) : \sigma(t) \text{ is } n\text{-torsion in } E_t\}.$$

Now, one proves that

$$(3.10) \qquad \left| \frac{A_n^\varepsilon}{A_n} - 1 \right| \leq \frac{C}{|\log \varepsilon|}, \quad \text{where } C \in \mathbb{R}_+.$$

Partitioning A_n in Galois orbits, this becomes a consequence of the fact that the points $t \in \bigcup_{n \in \mathbb{N}} A_n$ are of bounded height in $\mathbb{P}_1(\bar{\mathbb{Q}})$,[8] and the fact[9] that points $x \in \mathbb{P}_1(\bar{\mathbb{Q}})$ of height $\leq C'$ have at most $\frac{C'}{|\log \varepsilon|}[K(x) : K][K : \mathbb{Q}]$ conjugates of absolute value $< \varepsilon$.

By letting $\varepsilon \to 0$ in equation (3.9) and using (3.10) to employ a uniform convergence argument, one gets the sought equality (3.2) for r.

When r is not defined over a number field, let $K_0 \subset \mathbb{C}$ denote its minimal field of definition, and let $K = K_0 \bar{\mathbb{Q}} \subset \mathbb{C}$. The field K has finite transcendence degree, hence there exists an integral algebraic variety $X/\bar{\mathbb{Q}}$ such that $\bar{\mathbb{Q}}(X) \cong K$. By construction, up to restricting X to a Zariski-open subset, we may assume that there exists an algebraic family of morphisms

$$r_t : \bar{B}_t \to \mathbb{P}_1, \quad t \in X,$$

---

[7] Let us remark that the proof of (3.8) could be repeated *ad litteram* replacing $r^{-1}(S_\varepsilon)$ with $B(\mathbb{C})$. This would immediately conclude the proof of the theorem, without making use of (3.5). However, we decided to present things in a way that makes clear how the argument is much more elementary away from the singularities.

[8] This is a direct consequence of the well-known result of Tate [22] that, if $\pi : E \to B$ is an elliptic fibration, and $P : B \to E$ is a non-torsion section, then the function $t \mapsto \hat{h}_{E_t}(P_t)$, is, up to a bounded constant, a Weil height on B.

[9] One way to prove this fact is to use the following easy result, which can be found, for instance, in [28, Remark 3.10 (ii)]. Let $\xi \in L$, where $L \subset \mathbb{C}$ is a Galois number field of degree d over $\mathbb{Q}$; then

$$\sum_{g \in \mathrm{Gal}(L/\mathbb{Q})} |\log |\xi^g|| \leq 2d\,\hat{h}(\xi).$$

and sections

$$\sigma_t : \bar{B}_t \to \mathcal{E} \times_{(\pi, r_t)} (\bar{B}_t \setminus r_t^{-1}(S)), \quad t \in X,$$

such that there exists $t_0 \in X(\mathbb{C})$ such that $B_{t_0} = B$, $r_{t_0} = r$ and $\sigma_t = \sigma_{t_0}$.

We want to show now that, for any $t \in X(\mathbb{C})$ (and hence, in particular, for $t = t_0$), the following equality holds:

$$(3.11) \qquad \hat{h}(\sigma_t) = \int_{B_t \setminus r_t^{-1}(S)} \sigma^*(d\beta_2^t \wedge d\beta_1^t),$$

where, for a point $P \in B_t$, β_1^t, β_2^t denote the Betti coordinates of $\sigma_t(P)$. Now, we know that, for each $t \in X(\bar{\mathbb{Q}})$, (3.11) is true by the argument presented above in this Remark. Since $X(\bar{\mathbb{Q}})$ is dense (in the Euclidean topology) in $X(\mathbb{C})$, it is hence sufficient to show that both right- and left-hand side of (3.11) are continuous in $t \in X(\mathbb{C})$.

For the left-hand side, this is an immediate consequence of the fact that the height may be expressed through an explicit intersection formula on a complete regular minimal model (see [18, Section 11.8]), hence a standard flatness argument tells us that it is constant (hence, continuous) for t in a non-empty Zariski-open subset of X (that automatically contains t_0 since this is, by construction, a generic point).

For the right-hand side, this follows from a dominated convergence argument, which proceeds as follows. Since the $(1, 1)$-form $d\beta_2 \wedge d\beta_1$ diverges only when λ approaches $0, 1$ or ∞, it is sufficient to give a bound for $d\beta_2 \wedge d\beta_1$ in small disks around these three points, locally uniformly in $t \in X(\mathbb{C})$. Without loss of generality, one may do so only for a small disk around 0. Here, one may use results of Jones and Schmidt [10][10] to obtain the following bound in a circle $|\lambda| < \varepsilon$, locally uniformly in $t \in X(\mathbb{C})$:

$$(3.12) \qquad d\beta_2^t \wedge d\beta_1^t = \Omega\left(\frac{1}{|\lambda|^2 |\log \lambda|^3}\right) d\lambda \wedge d\bar{\lambda}.$$

Since the right-hand side of (3.12) is an integrable 2-form in the circle $|\lambda| < \varepsilon$, this allows a dominated convergence argument to be employed to show that the right-hand side of (3.11) is continuous, hence reproving Theorem 3.2.

Remark 3.9. Note that to carry out the proof of Theorem 3.2, a weaker estimate (than the finiteness granted by Proposition 2.4) for the number of torsion points of σ with $m_\sigma(b) > 1$ would have sufficed. Indeed (we keep here the notation of the proof of Theorem 3.2) if we had the estimate $\delta_n = o(n^2)$ (recall that what we obtain with Proposition 2.4 is $\delta_n = O(1)$), the proof would carry out analogously.[11] However, our approach for the formula (using the stronger bound $\delta_n = O(1)$) has also the advantage of giving some quite precise estimates for the number of torsion points of σ. We give here two such estimates.

---

[10] Of their paper one uses, most importantly, the lower bound $\det A(\lambda) \gg |\log |\lambda||$ of [10], and then the handy formulas, numbered (22)–(23), (27) in [10]. The complete calculation to obtain the bound below is quite long (but straightforward, given the just mentioned ingredients) and we prefer to omit it here, leaving it as an exercise to the interested reader.

[11] In fact, one may also deduce, using the argument for the proof of Theorem 3.2, that *a priori*, i.e., not using Proposition 2.4, the integral height formula (3.2) *is equivalent* to the bound $\delta_n = o(n^2)$. So, for instance, since, as explained later in Section 4.1, one can prove the height formula using the work [9] of DeMarco and Mavraki (without using Proposition 2.4), one can actually deduce this bound already from their work.

The first is the following:

$$A_n = n^2 \hat{h}(\sigma) + O(1) = n^2 \int_{B(\mathbb{C})} \sigma^*(\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1) + O(1).$$

Indeed, as shown in the proof of Theorem 3.2, we have $A_n = \langle n\bar{\sigma}, \bar{O}\rangle + O(1)$ (recall we used $\delta_n = O(1)$ to prove this); on the other hand, $\langle n\bar{\sigma}, \bar{O}\rangle = n^2 \hat{h}(\sigma) + O(1)$ by [18, Section 11.8].

The second is the following: let $U \subseteq B(\mathbb{C})$ be a *definable* domain. We then have that

$$A_n(U) = n^2 \int_U \sigma^*(\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1) + O(n).$$

Indeed, this may be proven following the same proof pattern as for point (b) in Theorem 3.2,[12] and using the error bound of Lemma 3.6 (see the footnote there).

Remark 3.10.   We want to address here the following question: What happens to the integral $\int_{B(\mathbb{C})} \sigma^*(\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1)$ when σ is a transcendental section?

In this case, the situation is well studied in the work [7] of Corvaja, Noguchi and Zannier, where they prove that the integral $\int_{B(\mathbb{C})} \sigma^*(\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1)$ is in fact *divergent* when the section σ is holomorphic and not rational (by this we mean that σ is a holomorphic section $B(\mathbb{C}) \to \mathcal{E}(\mathbb{C})$ of $\mathcal{E}(\mathbb{C}) \to B(\mathbb{C})$ that does not arise from an algebraic section, or, equivalently, by well-known GAGA results, such that it does not extend to a holomorphic section of the whole complete model $\overline{\mathcal{E}}(\mathbb{C}) \to \overline{B}(\mathbb{C})$). See, in particular, [7, Section 8 (iii)] (and refer also to [7, equation (4.8)]), where the situation is analyzed in detail, using the notion of "height characteristic function" (which, in this case, serves as a substitute for the height and analyses the behavior of the integral $\int_{B(r)} \sigma^*(\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1)$, where r varies in $\mathbb{R}_{>0}$ and $B(r)$ is an open subset of $B(\mathbb{C})$ that avoids a neighborhood, whose size depends on r, of the singularities, see [7] for the precise definition of this open subset and of the mentioned height function).

4. Addendum: Comparison with a measure coming from dynamics

In this addendum we compare the measure $\sigma^*(\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1)$ with a measure appearing in the following theorem of DeMarco and Mavraki [9, Section 3]:

Theorem 4.1 (DeMarco and Mavraki).   *Let* $\pi : \mathcal{E} \to B$ *be an elliptic surface and let* $P : B \to \mathcal{E}$ *be a non-torsion section, both defined over* $\overline{\mathbb{Q}}$. *Let* $\mathcal{S} \subset \mathcal{E}$ *be the union of the finitely many singular fibers in* E. *There is a positive, closed* $(1, 1)$-*current* T *on* $\mathcal{E}(\mathbb{C}) \setminus \mathcal{S}$ *with locally continuous potentials such that* $T|_{\mathcal{E}_t}$ *is the Haar measure on each smooth fiber, and* P^*T *is equal to a measure* μ_P, *that satisfies the following property. For any infinite non-repeating sequence of* $t_n \in B(\overline{\mathbb{Q}})$ *such that* $\hat{h}_{\mathcal{E}_{t_n}}(P_{t_n}) \to 0$ *as* $n \to \infty$, *the discrete measures*

$$\frac{1}{\#\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})t_n} \sum_{t \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})t_n} \delta_t$$

converge weakly on $B(\mathbb{C})$ *to* μ_P.[13]

---

[12] Note that, to carry out the proof here, one needs at the very least the bound $\delta_n = O(n)$, and a bound of the form $\delta_n = o(n^2)$ would no longer be sufficient.

[13] DeMarco and Mavraki [9] proved this result also in the non-archimedean setting.

In particular, we will prove that the current T is equal to the $(1,1)$-current defined by the $(1,1)$-form $\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1$ on $\mathcal{E} \setminus \mathcal{S}$ (we warn the reader that, whereas in the last section the form $\mathrm{d}\beta_2(\lambda) \wedge \mathrm{d}\beta_1(\lambda)$ we worked with was defined on the basis of the elliptic fibration B, the $(1,1)$-form $\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1$ we are working with now is define on the whole fibration space $\mathcal{E} \setminus \mathcal{S}$). A direct consequence of this is that the pullback measure $\sigma^*(\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1)$ is characterized by the arithmetic condition appearing in Theorem 4.1.

The restriction of the current T to the open set $\mathcal{E} \setminus (\mathcal{S} \cup O)$ (where we are denoting by O, with a slight abuse of notation, the image of the zero section of $\pi : \mathcal{E} \to B$), can be written down explicitly as follows (see [9, Proposition 3.1] and [9, proof of Proposition 3.6]):

$$T = \mathrm{dd}^c H_N,$$

where H_N denotes the Néron local (archimedean) height function. We recall that the following formula holds [21, p. 466]:

$$H_N(b, z) = -\log |e^{-\frac{1}{2} z \eta_b(z)} \sigma_b(z) \Delta(\Lambda_b)^{\frac{1}{12}}|.$$

Here $b \in B(\mathbb{C})$, Λ_b denotes a lattice in $\mathbb{C}$ (chosen to vary holomorphically in b) such that $\mathbb{C}/\Lambda_b \cong E_b := \pi^{-1}(b)$, z denotes the complex variable of $\mathbb{C}/\Lambda_b$ and η_b and σ_b indicate the semiperiod function η and, resp., the function σ (as defined in [21, Proposition VI.3.1, Proposition I.5.4]) associated to the lattice Λ_b. Since $\sigma_b(z) \Delta(\Lambda_b)^{\frac{1}{12}}$ is a holomorphic function in both variables b and z, this gives the following expression for T (on the open set $\mathcal{E} \setminus (\mathcal{S} \cup O)$):

$$(4.1) \qquad T = \frac{1}{2} \mathrm{dd}^c (\Re(z \eta_b(z))).$$

Using [9, Proposition 3.1], one shows that the above equality actually holds over the whole $\mathcal{E} \setminus \mathcal{S}$ (i.e., T does not have any part that is supported on the graph of the section O).

We shall prove that the current T matches the 2-form $\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1$ in two different ways: through a direct calculation (of which we give a sketch in Remark 4.2) and through a dynamical argument (in Corollary 4.6).

We notice that both T and $\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1$ restrict to the Haar measure on the fibers, normalized in such a way that the area of each fiber is 1.

Remark 4.2. Let us give a sketch of a calculation that shows that $T = \mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1$. This proves also (3.3). With a standard pullback argument, we may assume that $B = \mathbb{P}_1$ and that the scheme $\mathcal{E} \to B$ is (a regular complete minimal model of) the Legendre scheme. One first checks (using (4.1)) that

$$(4.2) \qquad T = \frac{i}{4\pi} \partial \bar{\partial} \left[\begin{pmatrix} z \\ \bar{z} \end{pmatrix}^\mathsf{T} \begin{pmatrix} \eta_1 & \eta_2 \\ \bar{\eta}_1 & \bar{\eta}_2 \end{pmatrix} \begin{pmatrix} \rho_1 & \rho_2 \\ \bar{\rho}_1 & \bar{\rho}_2 \end{pmatrix}^{-1} \begin{pmatrix} z \\ \bar{z} \end{pmatrix} \right],$$

where $\rho_i, \eta_i, i = 1, 2$ denote, respectively, the periods and the quasi-periods of E_λ, and

$$(4.3) \qquad \begin{pmatrix} \mathrm{d}\beta_1 \\ \mathrm{d}\beta_2 \end{pmatrix} = A^{-1} \left(\mathrm{d}(A) A^{-1} \begin{pmatrix} z \\ \bar{z} \end{pmatrix} + \begin{pmatrix} \mathrm{d}z \\ \mathrm{d}\bar{z} \end{pmatrix} \right), \quad \text{where } A := \begin{pmatrix} \rho_1 & \rho_2 \\ \bar{\rho}_1 & \bar{\rho}_2 \end{pmatrix}.$$

Using the relations

$$\frac{\mathrm{d}\rho_i}{\mathrm{d}\lambda} = f(\lambda)\eta_i + g(\lambda)\rho_i, \quad i = 1, 2$$

(with $f(\lambda) = -\frac{1}{2(\lambda^2 - \lambda)}$, $g(\lambda) = -\frac{2\lambda - 1}{6(\lambda^2 - \lambda)}$, see e.g. [16, p. 18], although the specific expres-

sions for these holomorphic functions will be irrelevant in this calculation), and remembering that ρ_i, η_i are holomorphic functions in λ, we see that

$$
d(A)A^{-1} = \begin{pmatrix} f(\lambda)d\lambda & 0 \\ 0 & \bar{f}(\lambda)d\bar{\lambda} \end{pmatrix} \begin{pmatrix} \eta_1 & \eta_2 \\ \bar{\eta}_1 & \bar{\eta}_2 \end{pmatrix} \begin{pmatrix} \rho_1 & \rho_2 \\ \bar{\rho}_1 & \bar{\rho}_2 \end{pmatrix}^{-1} + \begin{pmatrix} g(\lambda)d\lambda & 0 \\ 0 & \bar{g}(\lambda)d\bar{\lambda} \end{pmatrix}.
$$

To simplify the notation, we define[14]

$$
C := \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} := \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} := \begin{pmatrix} \eta_1 & \eta_2 \\ \bar{\eta}_1 & \bar{\eta}_2 \end{pmatrix} \begin{pmatrix} \rho_1 & \rho_2 \\ \bar{\rho}_1 & \bar{\rho}_2 \end{pmatrix}^{-1}
$$

$$
= \frac{1}{\det(A)} \begin{pmatrix} \eta_1\bar{\rho}_2 - \eta_2\bar{\rho}_1 & 2\pi i \\ 2\pi i & -\bar{\eta}_1\rho_2 + \bar{\eta}_2\rho_1 \end{pmatrix}.
$$

We wish to express now both $d\beta_2 \wedge d\beta_1$ and T as linear combinations (with coefficients in $\mathcal{C}^{(\infty)}(\pi^{-1}(U))$, $U \subset \mathbb{P}_1 \setminus \{0, 1, \infty\}$ being an open simply connected domain) of the forms $d\lambda \wedge d\bar{\lambda}$, $d\lambda \wedge d\bar{z}$, $dz \wedge d\bar{\lambda}$, $dz \wedge d\bar{z}$. We note that a (smooth) $(1, 1)$-form in $\Omega^{1,1}(\pi^{-1}(U))$ can be written uniquely as a linear combination of these forms, since the two 1-forms $d\lambda$ and dz span the cotangent bundle over each point in $\pi^{-1}(U)$. Hence, if we verify that the two expressions for $d\beta_2 \wedge d\beta_1$ and T are the same, we are done.

We write the explicit expression for $d\beta_2 \wedge d\beta_1$. Namely, using (4.3), one deduces that

$$
(4.4) \qquad d\beta_2 \wedge d\beta_1 = \frac{-1}{\det(A)}[(fC_{11}z + fC_{12}\bar{z} + gz)d\lambda + dz]
$$

$$
\wedge [(\bar{f}C_{21}z + \bar{f}C_{22}\bar{z} + \bar{g}\bar{z})d\bar{\lambda} + d\bar{z}],
$$

where $f = f(\lambda) = -\frac{1}{2(\lambda^2 - \lambda)}$ and $g = g(\lambda) = -\frac{2\lambda - 1}{6(\lambda^2 - \lambda)}$.

We turn now to T. One can verify that

$$
(4.5) \qquad \partial C_2 = 2\pi i \frac{fC_1 + (g \quad 0)}{\det(A)}d\lambda, \quad \bar{\partial}C_1 = 2\pi i \frac{\bar{f}C_2 + (0 \quad \bar{g})}{\det(A)}d\bar{\lambda},
$$

$$
(4.6) \qquad \partial\bar{\partial}C = \frac{2\pi i}{\det(A)} \begin{pmatrix} 2\bar{f}C_{12}(fC_{11}+g) & (fC_{11}+g)(\bar{f}C_{22}+\bar{g})+\bar{f}fC_{12}^2 \\ (fC_{11}+g)(\bar{f}C_{22}+\bar{g})+\bar{f}fC_{12}^2 & 2fC_{12}(\bar{f}C_{22}+\bar{g}) \end{pmatrix}.
$$

Using the Leibniz rule in (4.2), one gets

$$
(4.7) \quad T = \frac{i}{4\pi}\left[\frac{4\pi i}{\det A}dz \wedge d\bar{z} + 2\begin{pmatrix} dz \\ d\bar{z} \end{pmatrix}^{\mathsf{T}} \wedge \begin{pmatrix} -\bar{\partial}C_1 \\ \partial C_2 \end{pmatrix}\begin{pmatrix} z \\ \bar{z} \end{pmatrix} + 2\pi i \begin{pmatrix} z \\ \bar{z} \end{pmatrix}^{\mathsf{T}}(\partial\bar{\partial}C)\begin{pmatrix} z \\ \bar{z} \end{pmatrix}\right].
$$

Using formulas (4.5) and (4.6), an easy term-by-term comparison of (4.7) and (4.4) reveals that these last two expressions are the same.

We focus now on giving a more conceptual proof of the equality $T = d\beta_2 \wedge d\beta_1$. With a standard pullback argument, we may assume that $B = \mathbb{P}_1$ and that the scheme $\mathcal{E} \to B$ is (a complete minimal model of) the Legendre scheme.

---

[14] We have used here the Legendre relation $\rho_1\eta_2 - \rho_2\eta_1 = 2\pi i$ (see e.g. [21, Proposition I.5.2 (d)]).

Proposition 4.3.   *Let $D \subset B(\mathbb{C})$ be an open simply connected domain. The $(1,1)$-form $\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1$ is, up to scalar multiplication, the unique closed 2-form in $\Omega_0^2(\pi^{-1}(D))$[15] that satisfies $[2]^*\omega = 4\omega$, where $[2] : \pi^{-1}(D) \to \pi^{-1}(D)$ denotes the endomorphism of multiplication by 2 on the fibers.*

Proof.   Let $\omega \in \Omega_0^2(\pi^{-1}(D))$ be a 2-form such that $[2]^*\omega = 4\omega$. Let

$$\omega = \sum_{i,j} \alpha_{ij} \gamma_i \wedge \gamma_j, \quad \text{where } i, j \in \{1, 2, \lambda, \bar{\lambda}\}, \quad \begin{cases} \gamma_i = \mathrm{d}\beta_i, & i = 1, 2, \\ \gamma_i = \mathrm{d}\lambda, & i = \lambda, \\ \gamma_i = \mathrm{d}\bar{\lambda}, & i = \bar{\lambda}, \end{cases}$$

and $\alpha_{ij} \in \mathcal{C}^0(\pi^{-1}(D))$. We notice that

$$\frac{1}{4}[2]^*\omega = \sum_{ij} \alpha_{ij}([2]P) 2^{\delta i + \delta j - 2} \gamma_i \wedge \gamma_j,$$

where $\delta_i = 1$ when $i = 1, 2$, and $\delta_i = 0$ otherwise. We know that $\frac{1}{4}[2]^*\omega = \omega$ and hence

$$(4.8) \quad \alpha_{ij}(P) = 2^{\delta i + \delta j - 2} \alpha_{ij}([2]P) \quad \text{for all } i, j \in \{1, 2, \lambda, \bar{\lambda}\} \text{ and } P \in \pi^{-1}(D').$$

However, for each $x \in D$, restricting both hand sides of equation (4.8) to the fiber $\pi^{-1}(x)$, and then taking the ∞-norm yields

$$\max_{P \in \pi^{-1}(x)} |\alpha_{ij}(P)| = 2^{\delta i + \delta j - 2} \max_{P \in \pi^{-1}(x)} |\alpha_{ij}(P)|$$

$$\leq 2^{-1} \max_{P \in \pi^{-1}(x)} |\alpha_{ij}(P)| \quad \text{if } \{i, j\} \neq \{1, 2\}.$$

Hence, for $\{i, j\} \neq \{1, 2\}$, $\alpha_{ij} \equiv 0$. Therefore, we have

$$\omega = \alpha_{12} \, \mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1$$

and

$$\alpha_{12}([2]P) = \alpha_{12}(P) \quad \text{for all } P \in \pi^{-1}(D').$$

Since the function α_{12} is continuous, this implies that α_{12} is constant on the fibers of π, and hence it depends just on λ and $\bar{\lambda}$. Since ω is closed by hypothesis, it follows that α_{12} is constant in λ and $\bar{\lambda}$ as well. Hence $\omega = c\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1$, with $c \in \mathbb{C}$. $\square$

Remark 4.4.   In Proposition 4.3, one can replace the *closed* hypothesis with the hypothesis that the restriction of ω to the fibers of π is the normalized Haar measure. We note that the two different hypothesis do not *a priori* imply each other. On the other hand, the proof that was presented here works in both cases. In fact, the only point in which we used the closed assumption was to deduce that the function α_{12} is constant. However, this is automatic if ω restricts to the normalized Haar measure on the fibers.

Remark 4.5.   In Proposition 4.3, the hypothesis that ω is continuous on $\pi^{-1}(D)$ is crucial. The $(1, 1)$-form $\mathrm{d}z \wedge \mathrm{d}\bar{z}$, which is not a constant multiple of $\mathrm{d}\beta_2 \wedge \mathrm{d}\beta_1$, satisfies all the hypothesis of the proposition, except that it is not a well-defined continuous 2-form on $\pi^{-1}(D)$. This last fact may be easily seen by noticing that adding a period $\rho(\lambda)$ to z changes the 2-form $\mathrm{d}z \wedge \mathrm{d}\bar{z}$ by a (non-zero) term $\rho'(\lambda)\overline{\rho'(\lambda)}\mathrm{d}\lambda \wedge \mathrm{d}\bar{\lambda} + \overline{\rho'(\lambda)}\mathrm{d}z \wedge \mathrm{d}\bar{\lambda} - \rho'(\lambda)\mathrm{d}\bar{z} \wedge \mathrm{d}\lambda$.

---

[15] We are using the notation Ω_0^2 to denote continuous 2-forms.

Corollary 4.6. *The restriction to $\mathcal{E} \setminus \mathcal{S}$ of the $(1,1)$-current T is equal to the current associated to the $(1,1)$-form $d\beta_2 \wedge d\beta_1$.*

Proof. Let $D \subset \mathbb{P}_1(\mathbb{C}) \setminus S$ be an open relatively compact simply connected domain. Recall from Theorem 4.1 that T is closed. Moreover, we have

$$([2]^*T)(P) = \frac{1}{2} \, dd^c (\Re(z(2P)\eta_{\pi(P)}(z(2P)))) \quad \text{for all } P \in \pi^{-1}(D),$$

where $z(P)$ denotes the complex parameter of the complex torus $\mathbb{C}/\Lambda_{\pi(P)}$. We have the following:

$$(\Re(z(2P)\eta_b(z(2P)))) = 4(\Re(z(P)\eta_{\pi(P)}(z(P)))),$$

and, hence $[2]^*T = 4T$ on $\pi^{-1}(D)$. Covering $\mathbb{P}_1 \setminus S$ with such D proves that $[2]^*T = 4T$ on the whole $\mathcal{E} \setminus \mathcal{S}$. Moreover, note that expression (4.1) shows that T is smooth.

Hence, by Proposition 4.3[16], $T = c \, d\beta_2 \wedge d\beta_1$, where $c \in \mathbb{C}$ is a constant. Since, as remarked at the beginning of this subsection, both T and $d\beta_2 \wedge d\beta_1$ restrict to the normalized Haar measure on the fibers of π, $c = 1$, as we wanted to prove.   □

4.1. An alternative proof of Theorem 3.2. As mentioned in the beginning of Section 3.1, we sketch here another argument, using the work of DeMarco and Mavraki, that reproves Theorem 3.2.

Let

$$D_E(\sigma) := \sum_{\gamma \in B(\bar{K})} \hat{\lambda}_{E,\mathrm{ord}_\gamma}(\sigma) \cdot (\gamma),$$

where $\hat{\lambda}_{E,\mathrm{ord}_\gamma}(\sigma)$ is the local canonical height of the point corresponding to σ on the elliptic curve E over $k = \mathbb{C}(B)$ at the place ord_γ (i.e., the place corresponding to the point $\gamma \in B$), for each $\gamma \in B(\mathbb{C})$. The degree of $D_E(\sigma)$ is equal to $\hat{h}_E(\sigma)$. Then DeMarco and Mavraki [9, Sections 3 and 4] prove that (keeping the notation above):

$$c_1(D_E(\sigma)) = \sigma^*T = \sigma^* d\beta_2 \wedge d\beta_1,$$

where the last equality is Corollary 4.6, i.e., the comparison result proven in this section. Hence, Theorem 3.2 becomes a consequence of Wirtinger's formula applied to the divisor $D_E(\sigma)$.

Remark 4.7. In higher dimension (i.e., for a fibration $\mathcal{A} \to S$ in abelian varieties, where $\dim \mathcal{A} = 2 \cdot \dim(S) = 2g > 1$) the situation is slightly different. For instance, one could not expect a formula as (3.2) because the function $\hat{h}([n]\sigma)$ is always quadratic in $n \in \mathbb{Z}$, while $\int_B ([n]\sigma)^*(d\beta_1 \wedge \cdots \wedge d\beta_{2g})$ has degree $2g > 2$ in $n \in \mathbb{Z}$.

5. Quasi-integral points

We start by interpreting Theorem 2.8 in terms of points on the elliptic curve that are not quite integral, but quasi-integral or "almost-integral" (see the explanation following (5.7) below), in the sense that they have a small "denominator" (an exact meaning will be implicit in what follows). But from a certain point of view the result is not optimal, and the main purpose of this section is to remedy this defect.

---

[16] Here, instead of using Proposition 4.3, we could use its modified version, as in Remark 4.4.

5.1. Introduction. Theorem 2.8 may be reformulated in terms of heights on $\mathbb{C}(B)$. We may identify the section σ with a point (ξ, η) on $E(\mathbb{C}(B))$. The arguments in Section 6 will make it clear that $m_\sigma(b) = -\frac{1}{2}v(\xi)$, at least if $-v(\xi) \geq 0$ is large enough (indeed, recall anyway that the multiplicity is defined only at points where the section attains the value which is the origin of the elliptic curve), where v is the valuation corresponding to b (so that in particular $\sigma(b)$ is the origin of E).

For the Legendre model we will also see that

$$(5.1) \qquad\qquad h(\Xi(\sigma)) \leq 4h(\xi) + c_0$$

for the natural height $h(\xi) = \sum_v \max\{0, -v(\xi)\}$ on $\mathbb{C}(B)$ and some c_0 (here absolute). So for this model we deduce from Theorem 2.8 that $|\xi|_v = \exp(-v(\xi))$ satisfies

$$(5.2) \qquad\qquad |\xi|_v \leq e^{2c_0} H(\xi)^8$$

with

$$(5.3) \qquad\qquad H(\xi) = \exp(h(\xi)) = \prod_v \max\{1, |\xi|_v\}.$$

But of course (5.2) is worse than the trivial

$$(5.4) \qquad\qquad |\xi|_v \leq H(\xi).$$

And we will also see that the 4 in (5.1) cannot be avoided.

Further, we will see (for general models) that for sections $\sigma = n\sigma_0$ with σ_0 fixed, then thanks to $m_\sigma(b) = m_{\sigma_0}(b)$, Theorem 2.8 gives bounds for $m_\sigma(b)$ independent of n; and in fact similar remarks hold for linear combinations like $n\sigma_0 + m\tau_0$. Of course, $E(\mathbb{C}(B))$ is finitely generated so every (ξ, η) corresponds to such a linear combination; but as finding a basis for the Mordell–Weil group remains an ineffective procedure, this does not provide an effective non-trivial improvement of (5.2) in general.

The main object of the present section is to provide such an improvement. We shall drop the geometric terminology and replace $\mathbb{C}(B)$ by a function field $\mathcal{K}$ of transcendence degree 1 over an algebraically closed field k of zero characteristic.

Thus let E be an elliptic curve over $\mathcal{K}$, defined by say

$$y^2 = x^3 + ax^2 + bx + c$$

(now without danger of confusing a coefficient with a point of some B). As in Section 1, we assume (just for convenience) that the cubic factorizes completely over $\mathcal{K}$.

Now an argument of Manin [12] (see also Voloch [26]) using formal groups shows that (5.4) can be replaced by

$$(5.5) \qquad\qquad |\xi|_v \leq C_1,$$

where C_1 depends only on E and v. However (5.5) is again not effective, and not just because of Mordell–Weil.

In fact, it is already implicit in the literature that an effective bound

$$(5.6) \qquad\qquad |\xi|_v \leq C_2 H(\xi)^\theta$$

holds for some absolute $\theta < 1$ with C_2 depending only on E and $\mathcal{K}$. Take for simplicity $\mathcal{K}$ as $\mathbb{C}(t)$ and suppose we have absolute constants γ, δ such that

$$(5.7) \qquad \deg X \leq \gamma \max\{\deg A, \deg B, \deg C\} + \delta$$

for all "integral" X, Y in $\mathbb{C}[t]$ with $Y^2 = X^3 + AX^2 + BX + C$. Writing $\xi = \frac{X}{Z^2}, \eta = \frac{Y}{Z^3}$ for our point (ξ, η) of $E(\mathbb{C}(t))$, we deduce easily (5.6) for $\theta = 1 - \frac{1}{3\gamma}$.

The X above has of course $|X|_v = H(X)$ for v over ∞. A quasi-integral point would have $|\xi|_v$ "almost as large as" $H(\xi)$, so (5.6) usually implies their finiteness, too.

Now (5.7) has been known for some time, and for example [13, Theorem 6, p. 30] leads to $\theta = \frac{155}{156}$ (for general $\mathcal{K}$). But this procedure is wasteful, and a more detailed direct analysis yields $\theta = \frac{7}{8}$. Here the arguments use the *abc* inequality for function fields, whereas our arguments use a refined version (due to Wang) involving carefully chosen *abcd* . . . .

Here is our improvement of (5.4) and (5.6). To highlight the effectivity we give a completely explicit bound depending for example on a height

$$H(E) = \prod_v \max\{1, |a|_v^6, |b|_v^3, |c|_v^2\}.$$

Crucial for certain applications will be the fact that it no longer depends on the valuation v in (5.4). From now on all valuations will be supposed to have value group $\mathbf{Z}$ when written additively.

Theorem 5.1. *Given an elliptic curve E over a function field $\mathcal{K}$ as above, and any $\varepsilon > 0$, there is an effective constant $C = C(E, \mathcal{K}, \varepsilon)$, depending only on $E, \mathcal{K}, \varepsilon$, such that*

$$|\xi|_v \leq CH(\xi)^\varepsilon$$

for any (ξ, η) on $E(\mathcal{K})$ and any valuation v on $\mathcal{K}$. In fact, if $\varepsilon \leq \frac{1}{16}$, then we may take $C = (e^g H(E))^\rho$, where g is the genus of $\mathcal{K}$, and $\rho = 2^{10000/\varepsilon^2}$.

Note that the same bound is of course valid for $\max\{1, |\xi|_v\}$. Our methods will almost certainly establish

$$\prod_{v \in S} \max\{1, |\xi|_v\} \leq (e^{g-1+|S|} H(E))^\rho H(\xi)^\varepsilon$$

for any finite set S of valuations of $\mathcal{K}$. This would not follow simply by multiplying the individual bounds in the theorem.

Our proof essentially follows the classical strategy of Siegel. If ξ is v-adically large, then $P = (\xi, \eta)$ is close to the origin O. Regarding ε as fixed, we fix a positive integer m (later to be related to ε) and note that for every w over v of a suitable field extension some submultiple $\frac{P}{m}$ is w-adically close to some torsion point $T = \frac{O}{m}$. Here the point is that dividing by m does not essentially worsen the closeness. If $\frac{P}{m} = (\xi_m, \eta_m)$ and $T = (\tau, \omega)$, then the fixed quantity τ algebraic over $\mathcal{K}$ is well approximated by the varying quantity ξ_m also algebraic over $\mathcal{K}$, and we are set up for Roth's theorem (which of course Siegel did not have), taking all the available w. As in Siegel's strategy, we win because the logarithmic height $h(\xi_m)$ gets a lot smaller, about $\frac{1}{m^2} h(\xi)$. This means that we do not need the arbitrary Roth exponent $\kappa > 2$, and for example $\kappa = 3$ would suffice. And it is slightly simpler technically first to add a fixed point of order 2 to P and work with a finite target; this rules out the possibility $T = O$.

Actually Siegel argued slightly differently, using the Mordell–Weil Theorem that $E(\mathcal{K})$ is finitely generated to write $P = mQ + R$ with Q in $E(\mathcal{K})$ and a remainder R; then the fixed algebraic $-\frac{R}{m}$ is approximated by Q which needs no field extension. We can avoid this because the function field versions of Roth are much more uniform, so we do not suffer from the Mordell–Weil non-effectivity.

That Roth's theorem for function fields is effective seems to have been proved first by Osgood [15], at least for a single valuation; and Wang [27] treated several valuations, on the way considerably simplifying the proof. In Section 5.2 we shall give some extra minor simplifications in the proof of her Main Theorem and its applications to Roth's theorem.

Then in Section 5.3 we highlight the effectivity by giving an example in more traditional form for a particular algebraic function of degree 4 over $\mathbb{C}(t)$. We even calculate all the implied constants effectively.

In Section 5.4 we record some observations preliminary to the proof of our Theorem 5.1, which then follows in Section 5.5.

5.2. Wang's version of Roth's theorem.   We stay with a function field $\mathcal{K}$ of transcendence degree 1 over an algebraically closed field k of zero characteristic. It has a genus $g \geq 0$. We normalize the valuations v on $\mathcal{K}$ such that $v(f) = -\log |f|_v$ has value group $\mathbf{Z}$. Then the logarithmic height

$$h(f) = -\sum_v \min\{0, v(f)\},$$

corresponding to (5.3), is also given by

(5.8) $h(f) = [\mathcal{K} : k(f)].$

We fix a finite set S of these valuations and define

$$\chi = 2g - 2 + |S|$$

for the cardinality of S.

Now we take a finite set A^* consisting of 0 together with a non-empty set of S-units.

For an integer $r \geq 1$ denote by $L(r)$ the vector space over k spanned by monomials of degree r in the elements of A^*, and write $l(r) \geq 1$ for its dimension. Define also $L(0) = k$, so $l(0) = 1$. Note that

$$a^* L(r) \subseteq L(r + 1)$$

for every a^* in A^*. Here is a version of the Main Theorem of [27, p. 1226].

Lemma 5.2.   *For each v in S choose some a_v^* in A^*. Suppose $f \neq 0$ is in $\mathcal{K}$ and $r \geq 0$ are such that*

(5.9) $f L(r) \cap L(r + 1) = \{0\}.$

Then f is not in A^* and we have

$$\sum_{v \in S} \max\{0, v(f - a_v^*)\} + \sum_{v \notin S} \max_{a^* \in A^*} \max\{0, v(f - a^*) - (m + n - 1)\}$$

$$\leq \frac{m + n}{n} h(f) + \frac{(m + n)(m + n - 1)}{2n} \chi$$

for $n = l(r)$ and $m = l(r + 1)$.

In fact, the (non-negative) sum over $v \notin S$ does not appear in [27]; here it allows our result to be considered as an analogue of Nevanlinna's Second Main Theorem with ramification. We have also eliminated some extra heights appearing in [27].

Proof.   We employ the standard sort of Wronskian argument, but using suitable bases as in [27].

If f were in A^*, then $fL(r)$ would lie in $L(r+1)$, forcing $fL(r) = \{0\}$ by (5.9), a contradiction.

Choose basis elements $\beta_1, \ldots, \beta_n$ of $L(r)$ and basis elements $b_1, \ldots, b_m$ of $L(r+1)$. Write

$$\mu_v = \max_{a^* \in A^*} \max\{0, v(f - a^*) - (m + n - 1)\}$$

as in the lemma. As in [27], we let t be a non-constant element of $\mathcal{K}$, we let t_v be a local parameter at v, and we consider ω, ω_v defined as the Wronskians with respect to t, t_v respectively of $f\beta_1, \ldots, f\beta_n, b_1, \ldots, b_m$. Our assumption (5.9) together with $f \neq 0$ and $n \geq 1$ imply that these latter are linearly independent over k. So $\omega \neq 0, \omega_v \neq 0$.

We let S_1 be the subset of S made up of those v such that $v(f - a_v^*) > 0$; clearly the sum on the far left can be restricted to S_1.

We estimate $v(\omega_v)$ according to five disjoint cases for v.

Case (i): $v \notin S, v(f) \geq 0.$ If $\mu_v = 0$, we use just

(5.10) $$v(\omega_v) \geq 0.$$

If $\mu_v > 0$, then there is a^* in A^* with $v(f - a^*) = m + n - 1 + \mu_v$. We note as in [27] that ω_v is also the Wronskian of $(f - a^*)\beta_1, \ldots, (f - a^*)\beta_n, b_1, \ldots, b_m$ (with respect to t_v). This is because $a^*\beta_1, \ldots, a^*\beta_n$ are in $L(r+1)$ and so we may use column operations.

Now by looking at the first n columns we see that

$$v(\omega_v) \geq \mu_v + (\mu_v + 1) + \cdots + (\mu_v + n - 1) \geq n\mu_v;$$

thus from (5.10) we have

$$v(\omega_v) \geq n\mu_v$$

for all v in Case (i).

Case (ii): $v \notin S, v(f) < 0.$ Standard identities show that $\frac{\omega_v}{f^{m+n}}$ is the Wronskian of $\beta_1, \ldots, \beta_n, \frac{b_1}{f}, \ldots, \frac{b_m}{f}$ (with respect to t_v). These functions are all regular at v, hence

$$v(\omega_v) \geq (m + n)v(f).$$

Case (iii): $v \in S \setminus S_1, v(f) < 0.$ We use the same formula as in Case (ii). Now taking into account possible poles of the functions $\beta_1, \ldots, \beta_n, b_1, \ldots, b_m$, the usual computation yields

$$v(\omega_v) \geq (m + n)v(f) + \sum_{i=1}^{n} v(\beta_i) + \sum_{j=1}^{m} v(b_j) - B.$$

for the binomial coefficient $B = \binom{m+n}{2}$.

Case (iv): $v \in S \setminus S_1$, $v(f) \geq 0$. The usual computation on the original Wronskian yields

$$v(\omega_v) \geq \sum_{i=1}^{n} v(\beta_i) + \sum_{j=1}^{m} v(b_j) - B.$$

Case (v): $v \in S_1$, that is, $v(f - a_v^*) > 0$. As above in Case (i) we can replace f in the Wronskian by $f - a_v^*$. Then as in Case (iv), but now taking into account $f - a_v^*$ multiplying $\beta_1, \ldots, \beta_n$, we get

$$v(\omega_v) \geq nv(f - a_v^*) + \sum_{i=1}^{n} v(\beta_i) + \sum_{j=1}^{m} v(b_j) - B.$$

This completes the analysis, as we have covered all possible cases.

Let us now sum over all v on $\mathcal{K}$, and use the five inequalities so obtained, noting that the respective sets give a partition of all v. We obtain as lower bound for $\sum_v v(\omega_v)$

$$n \sum_{v \notin S, \, v(f) \geq 0} \mu_v + n \sum_{v \in S_1} v(f - a_v^*) + (m + n) \sum_{v \notin S_1, \, v(f) < 0} v(f)$$
$$+ \sum_{i=1}^{n} \sum_{v \in S} v(\beta_i) + \sum_{j=1}^{m} \sum_{v \in S} v(b_j) - B|S|.$$

Finally, in the sum over μ_v we may omit $v(f) \geq 0$, since $\mu_v = 0$ when $v(f) < 0$. Also the sum $\sum_{v \notin S_1, v(f) < 0} v(f)$ is at least $\sum_{v(f) < 0} v(f)$, which in turn is the number of poles of f counted with (negative) multiplicity, hence equals $-h(f)$; the two subsequent double sums vanish because β_i, b_j are S-units. Also, by a well-known formula for Wronskians, we have

$$\omega_v = \omega \left(\frac{\mathrm{d}t}{\mathrm{d}t_v} \right)^B.$$

Therefore, since $\sum_v v(\omega) = 0$, we see that $\sum_v v(\omega_v) = B(2g - 2)$ (by the Hurwitz formula).

This immediately leads to the stated inequality and completes the proof. □

We turn now to the application to Roth's theorem for function fields, which as in [27] involves the elimination of (5.9). But here we also drop all references to S-units.

Thus we take a finite set A consisting of 0 together with a non-empty set of l non-zero elements of $\mathcal{K}$.

Proposition 5.3. *For each v in S choose some a_v in A. Then for $f \neq 0$ in $\mathcal{K}$ and any positive $\varepsilon \leq \frac{1}{16}$ we have either*

$$h(f) \leq \frac{6l}{\varepsilon} \left(\log \frac{1}{\varepsilon} \right) \sum_{a \in A} h(a)$$

or f is not in A and

$$\sum_{v \in S} \max\{0, v(f - a_v)\} \leq (2 + \varepsilon)h(f) + 3\left(\frac{1}{\varepsilon} \right)^l \left(\chi + 2 \sum_{a \in A} h(a) \right).$$

Proof. We start by enlarging the set S to make the non-zero elements a of A into S-units. So for each such a we must throw in the v with $v(a) \neq 0$. Their number is at most

$$(5.11) \qquad \sum_{v(a)>0} v(a) - \sum_{v(a)<0} v(a) \leq 2h(a).$$

Now we are set up to apply Lemma 5.2.

We have to find an r with $\frac{l(r+1)}{l(r)} \leq 1 + \varepsilon$. If this fails for say $r = 0, 1, \ldots, R-1$, then we get

$$(1 + \varepsilon)^R < l(R) \leq \binom{R + l - 1}{l - 1}.$$

Write $\delta = \sqrt{1 + \varepsilon} - 1$. We have

$$(5.12) \qquad \binom{R + l - 1}{l - 1} \leq (1 + \delta)^{R + l - 1} \delta^{1-l},$$

hence $(1 + \varepsilon)^{\frac{R}{2}} < (1 + \varepsilon)^{\frac{l-1}{2}} \delta^{1-l}$, which yields

$$R \leq l - 1 + 2(l - 1) \frac{\log(\frac{1}{\delta})}{\log(1 + \varepsilon)} \leq R'$$

for $R' = 3(l - 1)\frac{1}{\varepsilon} \log \frac{1}{\varepsilon}$.

So if we choose

$$R = [R'] + 1 \leq \frac{3l}{\varepsilon} \log \frac{1}{\varepsilon},$$

we can find $r \leq R - 1$ as above, that is, with $l(r + 1) \leq (1 + \varepsilon)l(r)$. We then apply Lemma 5.2 with $m = l(r + 1), n = l(r)$. If (5.9) is not satisfied, then with the bases $\beta_1, \ldots, \beta_n, b_1, \ldots, b_m$ as before, we see that $f\beta_1, \ldots, f\beta_n, b_1, \ldots, b_m$ must be linearly dependent over k. So by [27, Lemma 5, p. 1232], we have

$$h(f) \leq (2r + 1) \sum_{a \in A} h(a) \leq 2R \sum_{a \in A} h(a) \leq \frac{6l}{\varepsilon} \left(\log \frac{1}{\varepsilon} \right) \sum_{a \in A} h(a),$$

which is the first of the two alternative conclusions of the present proposition.

Therefore we may indeed assume that the conclusion of Lemma 5.2 holds.

Now by (5.12),

$$\log \binom{R + l - 1}{l - 1} \leq (R + l - 1) \log(1 + \delta) + (l - 1) \log \frac{1}{\delta}$$

which is at most

$$l \left(\frac{4}{\varepsilon} \left(\log \frac{1}{\varepsilon} \right) \log(1 + \delta) + \log \frac{1}{\delta} \right) < l \log \frac{1}{\varepsilon}.$$

As

$$n = l(r) \leq \binom{r + l - 1}{l - 1} \leq \binom{R + l - 1}{l - 1},$$

we deduce $n \leq (\frac{1}{\varepsilon})^l$. Also

$$\frac{m + n}{n} \leq 2 + \varepsilon \leq \frac{33}{16},$$

so

$$\frac{(m+n)(m+n-1)}{2n} \leq \frac{33}{16}\frac{m+n-1}{2} < 3n.$$

We had enlarged the size of S by at most $2\sum_{a\in A} h(a)$ from (5.11), and this completes the proof of Proposition 5.3, as the left-hand side of the second alternative conclusion only gets bigger. $\qquad\square$

5.3. An example. It is clear that by optimizing ε in Proposition 5.3 we obtain something of the shape

$$2h(f) + O(h(f)^{\frac{l}{l+1}})$$

and therefore we obtain (effective) versions of Roth's theorem that are stronger than the analogues over finite extensions of $\mathbf{Q}$. Similar results, even with $O(h(f)^\theta)$ for any $\theta > \frac{2}{3}$, were found by Osgood [15] – see [15, Theorem VIII, p. 382] with $M = 2$. Since [15] is written from a rather more general point of view, and also the constants are not always calculated, we feel it may be of interest to work out a completely explicit example.

So now we take α in a finite extension of say $\mathbb{C}(t)$, and for simplicity we want to bound the traditional $|\alpha - \frac{p}{q}|$ from below, where p and $q \neq 0$ are in $\mathbb{C}[t]$ and the valuation extends that on $\mathbb{C}(t)$ defined by $v(\frac{1}{t}) = 1$.

If α has degree 2 over $\mathbb{C}(t)$, then of course we can avoid the ε altogether. This is true also if α has degree 3 over $\mathbb{C}(t)$; it was noted first also by Osgood [14], even in effective form, and Schmidt [17] worked these out in detail. An example is for

$$\alpha = -\frac{1}{t} - \frac{1}{t^3} - \frac{3}{t^5} - \frac{12}{t^7} - \frac{55}{t^9} - \cdots$$

satisfying $\alpha^3 - \alpha = \frac{1}{t}$; then there is an obvious extension of v to $\mathbb{C}(t,\alpha)$ with $v(\alpha) = 1$, and [17, Theorem 1 (i), p. 2] implies that

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{e^{-6}}{|q|^2}.$$

Actually, we know of no obstacle to the conjecture that this can be done for α of any degree $d \geq 2$ over $\mathbb{C}(t)$. But already for $d = 4$ the methods of [17], based on the use of differential equations, yield only $|q|^3$.

We will work out a corresponding result for the example

$$(5.13) \qquad \alpha = -\frac{1}{t} + \frac{1}{t^4} - \frac{4}{t^7} + \frac{22}{t^{10}} - \frac{140}{t^{13}} + \cdots$$

with

$$\alpha^4 - \alpha = \frac{1}{t}.$$

Example. We have

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{e^{-10^9}}{|q|^2} \exp\{-3(\log|q|)^{\frac{4}{5}}\}$$

for all p and $q \neq 0$ in $\mathbb{C}[t]$.

Verification. We take $\mathcal{K}$ as the Galois closure of $\mathbb{C}(t, \alpha)$, with $[\mathcal{K} : \mathbb{C}(t)] = 24$. We check that the genus $g = 4$. In fact, the above v extends easily to $\mathcal{K}$, the other conjugates of α being

$$(5.14) \qquad \alpha' = 1 + \frac{\frac{1}{3}}{t} - \frac{\frac{2}{9}}{t^2} + \cdots, \qquad \alpha'' = \omega + \frac{\frac{1}{3}}{t} + \cdots, \qquad \alpha''' = \omega^2 + \frac{\frac{1}{3}}{t} + \cdots$$

(for $\omega = e^{2\pi \frac{i}{3}}$) still in $\mathbb{C}((\frac{1}{t}))$. We get 24 valuations v_σ on $\mathcal{K}$ given by $v_\sigma(x) = v(x^\sigma)$ for σ in the Galois group S_4 of $\mathcal{K}/\mathbb{C}(t)$. These will make up our set S. So $\chi = 30$. Now with $f = \frac{p}{q}$

$$24v(f - \alpha) = 24v_\sigma(f - \alpha^{\sigma^{-1}}) = \sum_{\sigma \in S_4} v_\sigma(f - \alpha^{\sigma^{-1}})$$

and so our set $A = \{0, \alpha, \alpha', \alpha'', \alpha'''\}$ with $l = 4$. By (5.8) we have

$$h(\alpha) = [\mathcal{K} : \mathbb{C}(\alpha)] = [\mathcal{K} : \mathbb{C}(t, \alpha)] = 6$$

and the same for the other conjugates. So we deduce from Proposition 5.3 that either

$$(5.15) \qquad h \leq \frac{576}{\varepsilon} \log \frac{1}{\varepsilon}$$

or

$$(5.16) \qquad 24v(f - \alpha) \leq 2h + \varepsilon h + 234\varepsilon^{-4}$$

for $h = h(f)$ and any positive $\varepsilon \leq \frac{1}{16}$.

Optimizing ε in (5.16) gives $\varepsilon = (\frac{936}{h})^{\frac{1}{5}}$ so we need

$$(5.17) \qquad h \geq 936.16^5.$$

If we temporarily assume this, then (5.15) is impossible, and we conclude

$$(5.18) \qquad 24v(f - \alpha) \leq 2h + \frac{5}{4} 936^{\frac{1}{5}} h^{\frac{4}{5}}.$$

Also $h(t) = [\mathcal{K} : \mathbb{C}(t)] = 24$ so

$$h = h\left(\frac{p}{q}\right) = 24 \max\{\deg p, \deg q\}.$$

The further assumption $v(f - \alpha) > 0$ is harmless and then by (5.13) we see that $\deg p < \deg q$ and so $h = 24 \deg q$. Now dividing (5.18) by 24 and then exponentiating gives

$$\left| \alpha - \frac{p}{q} \right| \geq \frac{1}{|q|^2} \exp\left\{ -\frac{5}{4} \left(\frac{936}{24} \right)^{\frac{1}{5}} (\log |q|)^{\frac{4}{5}} \right\} \geq \frac{1}{|q|^2} \exp\{ -3(\log |q|)^{\frac{4}{5}} \}.$$

When (5.17) fails, we fall back on Liouville's argument; this amounts to

$$|(\alpha - f)(\alpha' - f)(\alpha'' - f)(\alpha''' - f)| = \left| f^4 - f - \frac{1}{t} \right| = \left| \frac{tp^4 - tpq^3 - q^4}{tq^4} \right| \geq \frac{e^{-1}}{|q|^4}.$$

Again the harmless $|f - \alpha| < 1$ implies $|\alpha' - f| = |\alpha'' - f| = |\alpha''' - f| = 1$ by (5.14), so we deduce

$$\left| \alpha - \frac{p}{q} \right| \geq \frac{e^{-1}}{|q|^4} = e^{-1} e^{-\frac{h}{6}} \geq \exp(-10^9),$$

accounting for the extra factor in the example above. This completes the verification.

Here the exponential factor beats that in the analogous Cugiani–Mahler–Bombieri–van der Poorten result [5] for algebraic numbers, which is

$$\exp\left\{-c \log |q| \left(\frac{\log \log \log |q|}{\log \log |q|}\right)^{\frac{1}{4}}\right\}$$

(for ineffective c) – and furthermore there is a condition of slowly growing denominators.

Here we may go beyond [15] by treating several valuations. This leads to new results of the type associated with Ridout. For example, with $v'(t) = 1$ and q a power of t it is easy to show that immediately after a term $\frac{c}{t^d}$ in (5.13) the number of consecutive zero coefficients is $O(d^{\frac{4}{5}})$.

5.4. Preliminaries. With the proof of our Theorem 5.1 in mind, we return now to our elliptic curve E over $\mathcal{K}$ with equation $y^2 = x^3 + ax^2 + bx + c$ for a, b, c in $\mathcal{K}$ and logarithmic height

$$h(E) = \log H(E) = \sum_v \max\{0, -6v(a), -3v(b), -2v(c)\}.$$

We will often use Zimmer's inequality

$$(5.19) \qquad \left|h(P) - \frac{3}{2}\hat{h}(P)\right| \leq \frac{1}{2}h(E)$$

for $P = (\xi, \eta)$ in $E(\mathcal{K})$ with

$$h(P) = \sum_v \max\{0, -v(\xi), -v(\eta)\},$$

and $\hat{h}$ as before is with respect to twice the origin O. See [30, Proposition 11.1, p. 484].

The h is with respect to $3O$, and we will also need the easy (one-sided)

$$(5.20) \qquad h(P) \leq \frac{3}{2}h(\xi) + \frac{1}{4}h(E).$$

On the other side $h(P) \geq \frac{3}{2}h(\xi) - \frac{1}{2}h(E)$ could almost as easily be checked using the identity

$$(x^2 - ax + a^2 - b)(x^3 + ax^2 + bx + c) - (a^3 - 2ab + c)x^2$$
$$- (a^2b - ac - b^2)x - (a^2c - bc) = x^5$$

but this we will not need.

In our Theorem 5.1 we are implicitly considering $|\xi|_v$ to be large. Then $P = (\xi, \eta)$ is near O, so $P + Q_0$ is near Q_0 for any fixed Q_0. Thus the abscissae of $P + Q_0, Q_0$ are close. This is expressed precisely in the following result, where we choose Q_0 as a point of order 2 (recall these are defined over $\mathcal{K}$), and revert to the additive notation.

Lemma 5.4. *For* $Q_0 = (\mu_0, 0)$ *and any* $O \neq P = (\xi, \eta) \neq Q_0$ *in* $E(\mathcal{K})$ *we have for* $P + Q_0 = (\mu, \nu)$ *the inequality*

$$\max\{0, v(\mu - \mu_0)\} \geq -v(\xi) - 2h(E).$$

Proof. Factor $x^3 + ax^2 + bx + c = (x - \mu_0)B(x)$. We have $\xi \neq \mu_0$ and we check $\mu - \mu_0 = \frac{B(\mu_0)}{\xi - \mu_0}$. Hence

$$v(\mu - \mu_0) = v(B(\mu_0)) - v(\xi - \mu_0).$$

Now

$$h(\mu_0) \leq h(Q_0) = h(Q_0) - \frac{3}{2}\hat{h}(Q_0) \leq \frac{1}{2}h(E)$$

by (5.19). Hence

$$v(B(\mu_0)) \geq -h(B(\mu_0)) = -h(3\mu_0^2 + 2a\mu_0 + b) \geq -2h(\mu_0) - \frac{1}{3}h(E) \geq -2h(E).$$

Also, $v(\xi - \mu_0) = v(\xi)$ if $v(\xi) < v(\mu_0)$. When this holds, we then have

$$v(\mu - \mu_0) \geq -v(\xi) - 2h(E),$$

proving the stated inequality. When this does not hold, then

$$-v(\xi) \leq -v(\mu_0) \leq h(\mu_0) \leq \frac{h(E)}{2}.$$

Hence $-v(\xi) - 2h(E) \leq 0$, thus completing the proof of the lemma.   □

We are now implicitly considering $Q = P + Q_0$ to be near a point of order 2. When we divide by a positive integer m, there are several possibilities for the quotients $\frac{Q}{m}$, and each of them should be almost as close (with a loss essentially independent of m) to a point of $E[2m]$, that is, a point of order dividing $2m$. This is expressed in the next key result, again in terms of abscissae.

Lemma 5.5. *With $Q_0 = (\mu_0, 0)$ suppose that for some odd integer $m \geq 3$ the points in $E[2m]$ are in $E(\mathcal{K})$. Then for any $O \neq R = (\zeta, \theta)$ in $E(\mathcal{K})$ with $O \neq Q = mR = (\mu, \nu) \neq Q_0$, we can find $O \neq T = (\tau, \omega)$ in $E[2m]$ with*

$$\max\{0, v(\zeta - \tau)\} \geq \frac{1}{2}\max\{0, v(\mu - \mu_0)\} - m^2 h(E).$$

Proof. Let φ be the rational function representing multiplication by m on abscissae; the numerator has degree $n = m^2$ and is monic, while the denominator has degree $n - 1$ with leading coefficient n. We may therefore write

$$\varphi(x) - \mu_0 = \frac{1}{n}\frac{(x - \tau_1)\cdots(x - \tau_n)}{(x - \sigma_1)\cdots(x - \sigma_{n-1})}.$$

Here $\tau_1, \ldots, \tau_n$ correspond to $\varphi(x) = \mu_0$, that is, $mT = \pm Q_0 = Q_0$; thus they are the abscissae of all such T, each in $E[2m]$. And $\sigma_1, \ldots, \sigma_{n-1}$ correspond to $\varphi(x) = \infty$, so are the abscissae of all non-zero points of $E[m]$. Thus putting $x = \zeta$ gives

$$\mu - \mu_0 = \frac{1}{n}\frac{(\zeta - \tau_1)\cdots(\zeta - \tau_n)}{(\zeta - \sigma_1)\cdots(\zeta - \sigma_{n-1})}.$$

We may assume $\tau = \tau_n$ satisfies

$$(5.21) \qquad\qquad v(\zeta - \tau) \geq v(\zeta - \tau_j) \quad (j = 1, \ldots, n).$$

We note that $h(\sigma_i) \leq \frac{h(E)}{2}$ and similarly

$$(5.22) \qquad\qquad h(\tau - \sigma_i) \leq h(\tau) + h(\sigma_i) \leq h(E) \quad (i = 1, \ldots, n - 1)$$

and

$$(5.23) \qquad\qquad h(\tau - \tau_j) \leq h(E) \quad (j = 1, \ldots, n).$$

We have

$$(5.24) \qquad v(\mu - \mu_0) = \sum_{j=1}^{n} v(\zeta - \tau_j) - \sum_{i=1}^{n-1} v(\zeta - \sigma_i)$$

and so by (5.21) and the ultrametric inequality

$$(5.25) \qquad v(\mu - \mu_0) \le nv(\zeta - \tau) - \sum_{i=1}^{n-1} \min\{v(\zeta - \tau), v(\tau - \sigma_i)\}.$$

We will see that the factor n can be reduced to 2 (giving the desired independence of m).

We treat three possibilities for $v(\zeta - \tau)$.

First suppose

$$(5.26) \qquad v(\zeta - \tau) < -h(E).$$

Now

$$(5.27) \qquad v(\tau - \sigma_i) \ge -h(\tau - \sigma_i) \ge -h(E)$$

by (5.22), so by (5.26) the minima in (5.25) are all $v(\zeta - \tau)$. So we obtain

$$(5.28) \qquad v(\mu - \mu_0) \le v(\zeta - \tau)$$

rather stronger than needed.

Second suppose

$$(5.29) \qquad -h(E) \le v(\zeta - \tau) \le h(E).$$

Using the left-hand inequality in (5.29) together with (5.27), we find that all the minima in (5.25) are at least $-h(E)$, so the right-hand inequality in (5.29) gives

$$v(\mu - \mu_0) \le nh(E) + (n-1)h(E) = 2nh(E) - h(E).$$

Thus again the left part of (5.29) gives

$$v(\mu - \mu_0) \le v(\zeta - \tau) + 2nh(E)$$

a little worse than (5.28) but still better than needed.

Our last possibility is

$$(5.30) \qquad v(\zeta - \tau) > h(E).$$

Now in (5.24) we have

$$(5.31) \qquad v(\zeta - \tau_j) \ge \min\{v(\tau - \tau_j), v(\tau - \zeta)\}.$$

Here $v(\tau - \tau_j) \le h(\tau - \tau_j) \le h(E)$ by (5.23), so by (5.30) the two values in the minimum in (5.31) are distinct. Thus we actually have equality in (5.31), and

$$(5.32) \qquad v(\zeta - \tau_j) = v(\tau - \tau_j) \le h(E).$$

Now $\tau_1, \ldots, \tau_n$ are not all different; in fact the T with $mT = Q_0$ come in pairs $\{T, -T\}$ together with $T = Q_0$. Thus our $\tau - \tau_n$ occurs e times, where $e = 1$ or $e = 2$.

Using (5.32) in (5.24)–(5.25) for the $\tau_j \neq \tau$, we get

$$(5.33) \qquad v(\mu - \mu_0) \leq ev(\zeta - \tau) + (n - e)h(E) - \sum_{i=1}^{n-1} \min\{v(\zeta - \tau), v(\tau - \sigma_i)\}.$$

Now in the minima we have $v(\tau - \sigma_i) \geq -h(\tau - \sigma_i) \geq -h(E)$ by (5.22), and using (5.30), we see that each minimum is at least $-h(E)$. Thus (5.33) gives

$$v(\mu - \mu_0) \leq ev(\zeta - \tau) + (n - e)h(E) + (n - 1)h(E) \leq 2v(\zeta - \tau) + 2nh(E).$$

The lemma now follows. □

5.5. Proof of Theorem 5.1. Take any $O \neq P = (\xi, \eta)$ in $E(\mathcal{K})$ and pick $Q_0 = (\mu_0, 0)$ in $E(\mathcal{K})$. If $P = Q_0$, then

$$\log |\xi|_v \leq h(\xi) \leq \frac{h(E)}{2},$$

so we get a much stronger bound.

So we may assume $P \neq Q_0$. By Lemma 5.4 we get for $Q = P + Q_0 = (\mu, v) \neq O, Q_0$ the inequality

$$(5.34) \qquad \max\{0, v(\mu - \mu_0)\} \geq \log |\xi|_v - 2h(E).$$

Next we pick an odd integer $m \geq 3$ and try to apply Lemma 5.5. Now it may not be true that $E[2m]$ lies in $E(\mathcal{K})$. But certainly it lies in $E(\mathcal{L})$ for some $\mathcal{L}$ with $[\mathcal{L} : K] \leq (2m)^4$. And it may not be true that there is R in $E(\mathcal{K})$ with $mR = Q$. So we fix some R with $mR = Q$ and then R lies in $E(\mathcal{F})$ for some $\mathcal{F}$ with $[\mathcal{F} : L] \leq m^2$. As $Q \neq O$, also $O \neq R = (\zeta, \theta)$.

We take any w of $\mathcal{F}$ over v (as always with value group $\mathbf{Z}$). By Lemma 5.5 with $\mathcal{F}$ instead of $\mathcal{K}$ we get $O \neq T_w = (\tau_w, \omega_w)$ in $E(\mathcal{F})$ with

$$\max\{0, w(\zeta - \tau_w)\} \geq \frac{1}{2} w(\mu - \mu_0) - m^2 h_{\mathcal{F}}(E)$$

for the height $h_{\mathcal{F}}$ with respect to $\mathcal{F}$. Using $h_{\mathcal{F}}(E) = nh(E)$ with $n = [\mathcal{F} : K] \leq 16m^6$ and summing over all w dividing v, we deduce for

$$s = \sum_{w|v} \max\{0, w(\zeta - \tau_w)\}$$

the lower bound

$$s \geq \frac{1}{2} nv(\mu - \mu_0) - 16m^8 nh(E)$$

(in which the exponent 8 will eventually play hardly any role). So by (5.34) we get

$$(5.35) \qquad s \geq \frac{1}{2} n \log |\xi|_v - 32m^8 nh(E).$$

Using Proposition 5.3, with S as the set of w dividing v, A as 0 together with the non-zero τ_w, $\varepsilon = \frac{1}{16}$ and $\mathcal{F}$ in place of $\mathcal{K}$, yields two alternatives: either

$$h_{\mathcal{F}}(\zeta) \leq (384 \log 2)l \sum_{w|v} h_{\mathcal{F}}(\tau_w) \leq 300l \sum_{w|v} h_{\mathcal{F}}(\tau_w)$$

or

$$s \leq 3h_{\mathcal{F}}(\zeta) + 3(16)^l \left(\chi_{\mathcal{F}} + 2 \sum_{w|v} h_{\mathcal{F}}(\tau_w)\right)$$

for the characteristic $\chi_{\mathcal{F}}$ with respect to $\mathcal{F}$.

Now there may be up to n different w, but as T_w lies in $E[2m]$, we have $l \leq (2m)^4$; further $h_{\mathcal{F}}(\tau_w) \leq 2h_{\mathcal{F}}(E)$.

Next

$$\hat{h}_{\mathcal{F}}(R) \leq \frac{2}{3}h_{\mathcal{F}}(R) + \frac{1}{3}h_{\mathcal{F}}(E)$$

which by (5.20) is at most $h_{\mathcal{F}}(\zeta) + \frac{1}{2}h_{\mathcal{F}}(E)$. Also $\hat{h}_{\mathcal{F}}(P) = \hat{h}_{\mathcal{F}}(Q) = m^2\hat{h}_{\mathcal{F}}(R)$, whence

$$h_{\mathcal{F}}(\xi) \leq h_{\mathcal{F}}(P) \leq \frac{3}{2}\hat{h}_{\mathcal{F}}(P) + \frac{1}{2}h_{\mathcal{F}}(E) \leq \frac{3}{2}m^2h_{\mathcal{F}}(\zeta) + m^2h_{\mathcal{F}}(E).$$

Hence in the first alternative we get

$$h(\xi) \leq 500m^2 ln h(E) \leq 128000m^{12}h(E).$$

Let us now deal with the second alternative. To estimate $\chi_{\mathcal{F}}$, we use Hurwitz in the form

$$2g_{\mathcal{F}} - 2 = [\mathcal{F} : K](2g - 2) + \sum_w (e_w - 1)$$

(over all w of $\mathcal{F}$) for the ramification indices. Here $\mathcal{F}/\mathcal{K}$ is unramified for w over the places v of $\mathcal{K}$ outside the set S_E of bad reduction. Thus

$$\sum_w (e_w - 1) = \sum_{v \in S_E} \sum_{w|v} (e_w - 1) \leq [\mathcal{F} : K] \sum_{v \in S_E} 1.$$

Also for the discriminant $\Delta = -4a^3c + a^2b^2 + 18abc - 4b^3 - 27c^2$ we have

$$\sum_{v \in S_E} 1 \leq 2h(\Delta) \leq 2h(E)$$

as in (5.11). So we get

$$2g_{\mathcal{F}} - 2 \leq (2g - 2)n + 2nh(E).$$

Also $|S| \leq n$, so $\chi_{\mathcal{F}} \leq (2g - 1 + 2h(E))n$. Thus taking into account the $h_{\mathcal{F}}(\tau_w)$, we find

$$s \leq 3h_{\mathcal{F}}(\zeta) + 3(16)^l(2gn + 4n^2h(E)).$$

Here also using (5.20), we have

$$h_{\mathcal{F}}(\zeta) \leq \frac{3}{2}\hat{h}_{\mathcal{F}}(R) + \frac{1}{2}h_{\mathcal{F}}(E) = \frac{3n}{2m^2}\hat{h}(P) + \frac{n}{2}h(E) \leq \frac{3n}{2m^2}h(\xi) + nh(E)$$

so we end up with

$$s \leq \frac{9n}{2m^2}h(\xi) + 12(16)^l n^2(g + h(E)).$$

Comparing with (5.35), recalling $l \leq 16m^4$ and dividing by $\frac{n}{2}$, we find that

$$\log|\xi|_v \leq \frac{9}{m^2}h(\xi) + 2^{66m^4}(g + h(E)).$$

Finally, to deduce Theorem 5.1 take m odd minimal with $\frac{9}{m^2} \leq \varepsilon$. Then indeed $m \geq 3$ so $\frac{9}{(m-2)^2} > \varepsilon$ and

$$m < 2 + 3\sqrt{\frac{1}{\varepsilon}} \leq \frac{\frac{7}{2}}{\sqrt{\varepsilon}}.$$

Thus $66m^4 < \frac{10000}{\varepsilon^2}$ and we are done (even without a little square).

6. Comparison of multiplicity bounds

In this section we work out some examples of Theorem 2.8 and Theorem 5.1. The comparison seems to run best with the Legendre curve. At first we will use the point

$$P = (2, \sqrt{4 - 2\lambda})$$

on $y^2 = x(x - 1)(x - \lambda)$. For each positive integer n there are coprime polynomials A_n, B_n in $\mathbf{Z}[\lambda]$ such that $\frac{A_n(\lambda)}{B_n(\lambda)}$ is the abscissa of $n(2, \sqrt{4 - 2\lambda})$. For example with $n = 1, 2, 3, 4$ we get

$$2, \quad \frac{(\lambda - 4)^2}{8(\lambda - 2)}, \quad \frac{2(5\lambda^2 - 16\lambda + 16)^2}{(\lambda^2 + 8\lambda - 16)^2}, \quad \frac{(\lambda^4 - 80\lambda^3 + 352\lambda^2 - 512\lambda + 256)^2}{32(\lambda - 2)\lambda^2(3\lambda^2 - 16\lambda + 16)^2}.$$

With a slightly laborious induction on the standard formulas relating the abscissae of $P_1 + P_2$, $P_1 - P_2$, P_1, P_2, one can check that the degree of A_n is $\frac{1}{2}(n^2 - 1)$ (odd n) and $\frac{n^2}{2}$ (even n) and the degree of B_n is $\frac{1}{2}(n^2 - 1)$ (odd n) and $\frac{1}{2}(n^2 - 2)$ (even n). Making $n \to \infty$, we see that this confirms the calculation $\hat{h}(P) = \frac{1}{2}$ of Section 3. One can also check that

$$(6.1) \qquad B_n(\lambda) = \begin{cases} b_n C_n(\lambda)^2 & \text{if } n \text{ is odd,} \\ b_n(\lambda - 2)C_n(\lambda)^2 & \text{if } n \text{ is even,} \end{cases}$$

for b_n in $\mathbf{Z}$ and C_n in $\mathbf{Z}[\lambda]$.

For every $\lambda_0 \neq 0, 1$ we will estimate from above

$$w_n(\lambda_0) = \mathrm{ord}_{\lambda=\lambda_0} B_n(\lambda)$$

assuming it is positive.

First we use Theorem 2.8. The curve B is given by $\mu^2 = 4 - 2\lambda$, with typical point $b = (\lambda, \mu)$ (say), and the section σ is given by $\sigma(b) = (2, \mu)$. We are considering a particular point $b_0 = (\lambda_0, \mu_0)$ for $\mu_0 = \sqrt{4 - 2\lambda_0}$. By the discussion just before Theorem 2.8 we have $m_\sigma(b_0) = v(\frac{x_{n\sigma}}{y_{n\sigma}})$ for the valuation v on $\mathbb{C}(B) = \mathbb{C}(\lambda, \mu)$ corresponding to b_0. Here coprimality gives

$$(6.2) \qquad v(x_{n\sigma}) = v\left(\frac{A_n(\lambda)}{B_n(\lambda)}\right) = -w_n(\lambda_0)$$

at least if $\lambda_0 \neq 2$; while if $\lambda_0 = 2$, it is $-2w_n(2)$. So $v(x_{n\sigma}) < 0 = \min\{0, v(\lambda)\}$, and it follows easily that $v(y_{n\sigma}) = \frac{3}{2}v(x_{n\sigma})$. We conclude

$$m_\sigma(b_0) = \frac{1}{2}w_n(\lambda_0)$$

if $\lambda_0 \neq 2$; and if $\lambda_0 = 2$, it is $w_n(2)$.

If we fix $dz = \frac{dx}{y}$ as a basis of the cotangent space, then the operator Ξ is given by

$$(6.3) \quad \Xi(x, y) = 4\lambda(1 - \lambda)\left(D\left(\frac{Dx}{y}\right) + \frac{Dx}{2(x - \lambda)y}\right) + 4(1 - 2\lambda)\frac{Dx}{y} + \frac{2x(x - 1)}{(x - \lambda)y}$$

with $D = \frac{d}{d\lambda}$ (even for x, y in the algebraic closure of $\mathbb{C}(\lambda)$; and the last term could be taken as $\frac{2y}{(x-\lambda)^2}$ as well). This corrects (2) of Manin [12, p. 1397]. We find

$$(6.4) \qquad \Xi(\sigma) = \frac{2\mu}{(2 - \lambda)^2}.$$

So $v(\Xi(\sigma)) = 0$ if $\lambda_0 \neq 2$ and $v(\Xi(\sigma)) = -3$ if $\lambda_0 = 2$. We deduce $m_\sigma(b_0) \leq 2$ for all b_0; and that there are at most finitely many b_0 with $m_\sigma(b_0) \geq 2$ (that is, $m_\sigma(b_0) = 2$).

It follows that

(6.5) $w_n(\lambda_0) \leq 4$

if $\lambda_0 \neq 2$, and $w_n(2) \leq 2$; and that there are at most finitely many λ_0 with $w_m(\lambda_0) \geq 4$ (that is, $w_n(\lambda_0) = 4$).

In turn this implies that the C_n in (6.1) are squarefree apart from at most finitely many squared factors $(\lambda - \lambda_0)^2$. We found no such λ_0 for $n = 1, 2, \ldots, 20$.

Next we use Theorem 5.1 for $\mathcal{K} = \mathbb{C}(B) = \mathbb{C}(\lambda, \mu)$ and $P = (\xi, \eta) = n(2, \mu)$, with v corresponding to (λ_0, μ_0) as above. We find that

$$v(\xi) = -w_n(\lambda_0)$$

if $\lambda_0 \neq 2$ as in (6.2). Also the logarithmic height of ξ with respect to $\mathbb{C}(\lambda)$ is at most $\frac{n^2}{2}$, so with respect to $\mathbb{C}(\lambda, \mu)$ we get $h(\xi) \leq n^2$.

As the genus $g = 0$ and $h(E) = 12$, we obtain

$$w_n(\lambda_0) \leq \varepsilon n^2 + 12(2^{\frac{10000}{\varepsilon^2}})$$

whenever $0 < \varepsilon \leq \frac{1}{16}$. Choosing say

$$\varepsilon = \left(\frac{10000 \log 2}{\log n} \right)^{\frac{1}{2}}$$

for $n \geq 2^{2560000}$, we get

$$w_n(\lambda_0) \leq 84 \frac{n^2}{(\log n)^{\frac{1}{2}}} + 12n \leq 100 \frac{n^2}{(\log n)^{\frac{1}{2}}}.$$

This is a lot worse than (6.5). However we looked at only the points nP on $E(\mathcal{K})$, whereas this group contains $\mathbf{Z}P + E[2]$. It is also conceivable that the rank exceeds 1, say

$$E(\mathcal{K}) = \mathbf{Z}P + \mathbf{Z}Q + \cdots,$$

and similar arguments would apply to mQ or even $nP + mQ$, giving an explicit $o(n^2 + m^2)$. But as mentioned, the determination of $E(\mathcal{K})$ is not yet an effective procedure.

We can allow ourselves an extra Q simply by increasing the field to get

$$P = (2, \sqrt{4 - 2\lambda}), \quad Q = (3, \sqrt{18 - 6\lambda}).$$

Now B is given by $\mu^2 = 4 - 2\lambda$, $\nu^2 = 18 - 6\lambda$, and we have a second section τ, and we apply Theorem 2.8 to $n\sigma + m\tau$. Here $\Xi(\tau) = \frac{2\nu}{(3-\lambda)^2}$ as in (6.4), and so $\Xi(n\sigma + m\tau) = \beta$ for

$$\beta = \frac{2n\mu}{(2 - \lambda)^2} + \frac{2m\nu}{(3 - \lambda)^2}.$$

It is not hard to show that the order at any point b of B is bounded independently of b, n, m. In fact, one finds for "generic" n, m that $h(\beta) = 12$ so $h(\beta) \leq 12$ for all n, m. Thus $v(\beta) \leq 12$ too, and the corresponding order of the abscissa (now a rational function of λ, μ, ν) is at most 14.

Theorem 5.1 gives as above an explicit estimate $o(n^2 + m^2)$; but again it can be applied to the full $E(\mathcal{K}) = \mathbf{Z}P + \mathbf{Z}Q + \cdots$, whose generators may now be rather difficult to find.

We finish by proving (5.1) and that the 4 is best possible. From

$$\Xi(x, -y) = \Xi(-(x, y)) = -\Xi(x, y)$$

(or (6.3) directly) we see that $\Xi(x, y)$ is an odd function, and we calculate it as $\frac{\Upsilon}{y^3}$, where

(6.6) $$\Upsilon = f_0 D^2 x + f(Dx)^2 + f_1 Dx + f_2$$

and f_0, f, f_1, f_2 are polynomials in x of degree $3, 2, 3, 4$, respectively, with coefficients in $\mathbf{Z}[\lambda]$. Now (5.1) is clear (and the right-hand side of (6.6) provides a differential equation vanishing at the abscissae of all points of finite order at least 3).

To see that the factor 4 is best possible, we write

$$\Xi(x, y)^2 = \frac{\Upsilon^2}{x^3(x-1)^3(x-\lambda)^3}$$

and we calculate this for abscissae $x = \xi = \lambda^d + 6\lambda + 70$ $(d \geq 2)$ and corresponding ordinate $y = \eta$. We find

(6.7) $$\frac{P_d(\lambda)}{(\lambda^d + 6\lambda + 70)^3(\lambda^d + 6\lambda + 69)^3(\lambda^d + 5\lambda + 70)^3},$$

where P_d is in $\mathbf{Z}[\lambda]$ with leading term $4(d-1)^4\lambda^{8d}$. Thus a lower bound

$$h(\Xi(\xi, \eta)) \geq 4dh(\lambda) - O(1) \geq 4h(\xi) - O(1)$$

will follow as soon as the numerator and denominator in (6.7) have no common zeros.

Now we find that the numerator

$$P_d(\lambda) = \frac{P(d, \lambda, \lambda^d)}{\lambda^2}$$

for a fixed polynomial P. If this had a zero in common with, say, $\lambda^d + 6\lambda + 70$ in the denominator, then so would $P(d, \lambda, -6\lambda - 70)$. This is in $\mathbf{Z}[\lambda]$ with leading term

$$5184(d-1)^4\lambda^8.$$

So the zero λ would have to be of degree at most 8 over $\mathbf{Q}$. However by Eisenstein with prime 2 we see that it has degree d. Therefore when $d \geq 9$ there is no such common zero.

Similar arguments work with the other factors $\lambda^d + 6\lambda + 69$ and $\lambda^d + 5\lambda + 70$ in the denominator, for which the Eisenstein primes 3 and 5 suffice. We get leading terms

$$5184(d-1)^4\lambda^8, \quad 2500(d-1)^4\lambda^{10}$$

but in the second case there is a factor λ^2. So again when $d \geq 9$ there is no common zero.

7. Analogues over number fields

For simplicity we confine ourselves to E over $\mathbf{Q}$ and points (ξ, η) of $E(\mathbf{Q})$, with $\xi = \frac{p}{q}$ for integers $p, q > 0$ and $H(\xi) = \max\{|p|, q\}$ for the conventional absolute value $|\xi|$. The bound $|\xi| \leq CH(\xi)^\varepsilon$ in our Theorem 5.1 holds, but now C is not effective due to Roth.

The case $q = 1$ corresponds to Siegel's theorem, which was made effective using the theory of linear forms in logarithms. With the latest versions this leads to

$$|\xi| \le C \frac{H(\xi)}{(\log H(\xi))^\delta}$$

for some absolute (not too small) $\delta > 0$ and effective $C = C(E)$.

Furthermore, linear forms in elliptic logarithms lead to

$$|\xi| \le C(\log H(\xi))^\kappa$$

but this is no longer effective, this time due to Mordell–Weil (here κ depends on E).

See [20, Chapter 9] for some more details.

We do not know an analogue of the latter method for function fields.

Also we do not know if, for example, there are at most finitely many (ξ, η) in $E(\mathbf{Q})$ such that the numerator of ξ is a prime number or a prime power.

References

[1] *Y. André*, Mumford–Tate groups of mixed Hodge structures and the theorem of the fixed part, Compos. Math. **82** (1992), no. 1, 1–24.

[2] *Y. André*, *P. Corvaja* and *U. Zannier*, The Betti map associated to a section of an abelian scheme, Invent. Math. **222** (2020), no. 1, 161–202.

[3] *F. Barroero* and *M. Widmer*, Counting lattice points and O-minimal structures, Int. Math. Res. Not. IMRN **2014** (2014), no. 18, 4932–4957.

[4] *E. Bierstone* and *P. D. Milman*, Semianalytic and subanalytic sets, Publ. Math. Inst. Hautes Études Sci. **67** (1988), 5–42.

[5] *E. Bombieri* and *A. J. van der Poorten*, Some quantitative results related to Roth's theorem, J. Aust. Math. Soc. Ser. A **45** (1988), no. 2, 233–248.

[6] *P. Corvaja*, *D. Masser* and *U. Zannier*, Torsion hypersurfaces on abelian schemes and Betti coordinates, Math. Ann. **371** (2018), no. 3–4, 1013–1045.

[7] *P. Corvaja*, *J. Noguchi* and *U. Zannier*, Analytic and rational sections of relative semi-abelian varieties, preprint 2020, `https://arxiv.org/abs/2005.06349`.

[8] *P. Corvaja* and *U. Zannier*, Poncelet games and elliptic schemes, preprint.

[9] *L. DeMarco* and *N. M. Mavraki*, Variation of canonical height and equidistribution, Amer. J. Math. **142** (2020), no. 2, 443–473.

[10] *G. Jones* and *H. Schmidt*, Pfaffian definitions of Weierstrass elliptic functions, Math. Ann. **379** (2021), no. 1–2, 825–864.

[11] *B. Lawrence* and *U. Zannier*, On the p-adic distribution of torsion values for a section of an abelian scheme, Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl. **31** (2020), no. 2, 285–295.

[12] *J. I. Manin*, Rational points on algebraic curves over function fields, Izv. Akad. Nauk SSSR Ser. Mat. **27** (1963), 1395–1440.

[13] *R. C. Mason*, Diophantine equations over function fields, London Math. Soc. Lecture Note Ser. **96**, Cambridge University, Cambridge 1984.

[14] *C. F. Osgood*, Effective bounds on the "Diophantine approximation" of algebraic functions over fields of arbitrary characteristic and applications to differential equations, Nederl. Akad. Wetensch. Proc. Ser. A **37** (1975), 105–119.

[15] *C. F. Osgood*, Sometimes effective Thue–Siegel–Roth–Schmidt–Nevanlinna bounds, or better, J. Number Theory **21** (1985), no. 3, 347–389.

[16] *H. Schmidt*, Relative Manin–Mumford in additive extensions, Trans. Amer. Math. Soc. **371** (2019), no. 9, 6463–6486.

[17] *W. M. Schmidt*, On Osgood's effective Thue theorem for algebraic functions, Comm. Pure Appl. Math. **29** (1976), no. 6, 749–763.

[18] *M. Schütt* and *T. Shioda*, Elliptic surfaces, in: Algebraic geometry in East Asia—Seoul 2008, Adv. Stud. Pure Math. **60**, The Mathematical Society of Japan, Tokyo (2010), 51–160.

[19] *J.-P. Serre*, Lectures on the Mordell–Weil theorem, Aspects Math. **E15**, Friedrich Vieweg & Sohn, Braunschweig 1989.

[20] *J. H. Silverman*, The arithmetic of elliptic curves, Grad. Texts in Math. **106**, Springer, New York 1986.

[21] *J. H. Silverman*, Advanced topics in the arithmetic of elliptic curves, Grad. Texts in Math. **151**, Springer, New York 1994.

[22] *J. Tate*, Variation of the canonical height of a point depending on a parameter, Amer. J. Math. **105** (1983), no. 1, 287–294.

[23] *D. Ulmer* and *G. Urzua*, Tansversality of sections on elliptic surfaces with applications to elliptic divisibility sequences and geography of surfaces, preprint 2019, `https://arxiv.org/abs/1908.02208`.

[24] *D. Ulmer* and *G. Urzúa*, Bounding tangencies of sections on elliptic surfaces, Int. Math. Res. Not. IMRN **2021** (2021), no. 6, 4768–4802.

[25] *L. van den Dries*, Tame topology and o-minimal structures, London Math. Soc. Lecture Note Ser. **248**, Cambridge University, Cambridge 1998.

[26] *J. F. Voloch*, Siegel's theorem for complex function fields, Proc. Amer. Math. Soc. **121** (1994), no. 4, 1307–1308.

[27] *J. T.-Y. Wang*, An effective Roth's theorem for function fields, Rocky Mountain J. Math. **26** (1996), no. 3, 1225–1234.

[28] *U. Zannier*, Lecture notes on Diophantine analysis, Appunti. Sc. Norm. Super. Pisa (N. S.) **8**, Edizioni della Normale, Pisa 2009.

[29] *U. Zannier*, Some problems of unlikely intersections in arithmetic and geometry, Ann. of Math. Stud. **181**, Princeton University, Princeton 2012.

[30] *H. G. Zimmer*, Die Neron–Tate'schen quadratischen Formen auf der rationalen Punktgruppe einer elliptischen Kurve, J. Number Theory **2** (1970), 459–499.

———————————

Pietro Corvaja, Dipartimento di Scienze Matematiche, Informatiche e Fisiche (DMIF),
Università degli Studi di Udine, Via delle Scienze 206, 33100 Udine, Italy
e-mail: pietro.corvaja@uniud.it

Julian Demeio, Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy;
and Faculté des Sciences d'Orsay, Université Paris-Saclay, Bâtiment 307,
rue Michel Magat, F-91405 Orsay Cedex, France
e-mail: julian.demeio@sns.it

David Masser, Department of Mathematics and Computer Science, University of Basel,
Spiegelgasse 1, 4051 Basel, Switzerland
e-mail: david.masser@unibas.ch

Umberto Zannier, Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy
e-mail: umberto.zannier@sns.it