ON UNIFORM BOUNDS FOR RATIONAL POINTS ON NON-RATIONAL CURVES

J. ELLENBERG AND A. VENKATESH

Abstract. We show that the number of rational points of height $\leq H$ on a non-rational plane curve of degree d is $O_d(H^{2/d-\delta})$, for some $\delta > 0$ depending only on d. The implicit constant depends only on d. This improves a result of Heath-Brown, who proved the bound $O_\epsilon(H^{2/d+\epsilon})$. We also show that one can take $\delta = 1/450$ in the case $d = 3$.

Contents

1. Introduction

Let $P = [x_0 : \ldots : x_k]$ be a point of $\mathbb{P}^k(\mathbb{Q})$, where $x_0, \ldots, x_k$ are integers with no common factor. The *height* of P, denoted $H(P)$, is defined to be $\max_{0 \leq i \leq k} |x_i|$. The aim of this paper is to prove the following upper bound for the number of points of bounded height on a non-rational projective curve.

Theorem 1. *Let C be an irreducible curve in $\mathbb{P}^2$ of geometric genus at least 1 and degree d. Then there are constants $\delta = \delta(d)$ and $c = c(d)$ such that there are at most $c(d)H^{2/d-\delta}$ points on $C(\mathbb{Q})$ of height $\leq H$.*

Recall that the geometric genus of C is the genus of its normalization, and so a birational invariant; in particular, note that there is no requirement that C be smooth. We remark that, by a projection argument due to Heath-Brown [9], a similar bound on curves in $\mathbb{P}^m$ can be derived from Theorem 1.

The constant $\delta(d)$ is completely effective but rather tedious to compute. We will show that when $d = 3$, one can take $\delta(3) = \frac{1}{450}$. [1]

For an individual curve, this theorem is of course weaker than the results of Mordell (in genus 1) and Faltings (for higher genus). The content of the result lies in its uniformity in families of plane curves.

Theorem 1 improves, for non-rational curves, on a previous uniform result of Heath-Brown [9], which yields an exponent of $2/d$. The methods used in [9] (or the work of Bombieri-Pila, [4]) are insensitive to the genus of the curve. For instance, these methods do not distinguish between the curves $y^2 = x^d$ and $y^2 = x^d + D$, although we expect far fewer rational points on the latter. The bound $H^{2/d}$ of Heath-Brown, [9] is in fact sharp for rational curves. We improve the exponent by means of the geometry of higher-genus curves and especially their Jacobians, a tool not available for rational curves.

Let us briefly compare our result to what the usual methods can obtain. The bounds obtained by invoking the Mordell-Weil lattice of a Jacobian depend exponentially on that rank, and are also sensitive to the existence of points of small height; see e.g. [7] where uniform results are obtained under the circumstance that these two issues can be controlled. The main point of this paper is the observation that (a refinement of the) techniques of [9], together with a descent argument on the Mordell-Weil lattice, suffice to control this problem.

A very interesting problem is to give an analogue of Theorem 1 for *integral* points. Here one would aim to improve the exponent $1/d$ obtained by Bombieri-Pila [4]. The method of this paper does not apply directly, as one cannot obtain a sufficiently strong version of Prop. 1 for integral points, cf. Remark 2.

Finally, Heath-Brown has informed us that he obtained a similar result for smooth cubic curves. His methods are apparently similar to those of the present paper.

Acknowledgements: We have benefited from conversations with A. J. de Jong. We also thank Brian Conrad, Davesh Maulik and Jason Starr, who took the time to answer a great many of our algebro-geometric questions, and Tim Browning who made several comments on a draft of this paper. The second author would like to thank D.R. Heath-Brown and T. Browning for generously discussing their ideas and methods during his visit to Oxford.

2. Bounds with good dependence on $\|f\|$

2.1. **Notation.** Let $m \geq 1$, and endow $\mathbb{R}^m$ with the usual inner product in which the coordinate vectors form an orthonormal basis. Let $N \subset \mathbb{Z}^m$ be a subgroup of rank p. The inner product on $\mathbb{R}^m$ induces an inner product $\langle \cdot, \cdot \rangle_p$ on $\wedge^p \mathbb{R}^m$; on the other hand $\wedge^p N$ is a free $\mathbb{Z}$-module of rank 1. Choose η_p to be a generator. We define the height of N by the rule $\mathrm{ht}(N) := \langle \eta_p, \eta_p \rangle_p$. It is evident that $\mathrm{ht}(N)$ is an integer. It is the square of the volume of the flat torus $N \otimes \mathbb{R}/N$, where the metric on the torus arises from the restriction of the inner product on $\mathbb{R}^m$.

Put $N^\perp = \{\lambda \in \mathbb{Z}^m : \langle \lambda, N \rangle = 0\}$. Then $\mathrm{ht}(N^\perp) | \mathrm{ht}(N)$, with equality if N is "saturated," i.e. $\mathbb{R}N \cap \mathbb{Z}^m = N$.

[1] Although one can probably improve this value considerably (we suggest several methods for doing so), it seems that $\delta(d)$ will decrease very rapidly in d. The limiting factor is giving a good bound on the rank of the Jacobian.

Now, for $n \geq 1$, set $R = \mathbb{Z}[x_1, \ldots, x_{n+2}]$. Let $\mathcal{B}[l]$ be the set of monomials of degree l, and let $R[l]$ be the $\mathbb{Z}$-span of $\mathcal{B}[l]$; thus $R = \oplus_{l=0}^{\infty} R[l]$. $\mathcal{B}[l]$ is a $\mathbb{Z}$-basis for $R[l]$; in particular there is an isomorphism $\mathbb{Z}^{\mathcal{B}[l]} \to R[l]$, and a unique inner product on $R[l]$ making $\mathcal{B}[l]$ an orthonormal basis. This yields an isomorphism $\mathbb{Z}^{\mathcal{B}[l]} \to R[l]^*$ (the dual of $R[l]$), and also allows us to speak of the height of a subgroup $N \subset R[l]$. We shall denote the norm and inner product as $\| \cdot \|$ and $\langle \cdot, \cdot \rangle$, respectively.

2.2. **Bounding the number of points.** The projective scheme $\mathrm{Proj}(R)$ defined by R is just $n+1$-dimensional projective space $\mathbb{P}^{n+1}$. It carries a natural line bundle $\mathscr{O}(1)$, and the global sections $\Gamma(\mathbb{P}^{n+1}, \mathscr{O}(M))$ are naturally identified with $R[M]$.

Let $f \in R[d]$ be irreducible. Let $Z_f \subset \mathbb{P}^{n+1}$ be the n-dimensional zero-locus of f. We shall prove the following refinement of Theorems 3 and 4 of [9]. (It also implies a corresponding refinement of the results of [4]: see Remark 2.)

Proposition 1. *The $\mathbb{Q}$-points on Z_f of height $\leq H$ are contained in at most $(H^{\frac{n+1}{d^{1/n}}} \cdot \|f\|^{-d^{-1-1/n}} + 1)H^{\epsilon}$ divisors of degree $O_{\epsilon,d,n}(1)$.*

We briefly give the approximate idea of the proof. One might proceed very naively as follows: if there are enough rational points of Z_f of low height, one can hope to recover f by interpolation. If there were *too many* points of low height, this interpolated form would have size smaller than $\|f\|$, a contradiction. To implement this vague idea, one needs one further idea (present in different language in [9]): it is most efficient to first choose a "cluster" of rational points that are p-adically close, for some appropriate prime p, and then use these to interpolate f. These ideas motivate the formal proof below.

Proof. In what follows, it is useful to think of M as "big." The implicit constants in the symbols $\ll, \gg$ will depend on M, n and and d. The proof follows the ideas of Heath-Brown [9], the extra ingredient being to keep track of the height of certain lattices.

Let $Z_{\mathrm{sing}} \subset Z_f$ be the singular locus of f; this being a divisor, it suffices to show the claimed result for the points of $Z_f - Z_{\mathrm{sing}}$. Let $\Lambda_M \subset R[M]$ be the subgroup consisting of functions vanishing on Z_f. Then $\Lambda_M = f \cdot R[M-d]$. Set $\gamma = \#\mathcal{B}[M]$ and $\delta = \#\mathcal{B}[M] - \#\mathcal{B}[M-d]$. It is easy to check that

$$\gamma \sim \frac{M^{n+1}}{(n+1)!}, \ \delta \sim \frac{dM^n}{n!}, \tag{1}$$

where the symbol $\sim$ should be understood as meaning that the ratio of the two quantities approaches 1 as $M \to \infty$.

A compactness argument yields

$$\mathrm{ht}(\Lambda_M) \asymp_{M,d} \|f\|^{2|\mathcal{B}[M-d]|} \tag{2}$$

(The function $f \mapsto \mathrm{ht}(f.R[M-d])$ extends to a continuous function on the real vector space $R[d] \otimes \mathbb{R}$ that is nonvanishing away from $f = 0$ and homogeneous of order $2|\mathcal{B}[M-d]|$. This yields (2).)

We say a subset $T \subset Z_f(\mathbb{C})$ is in M-general position if any polynomial $g \in R[M] \otimes \mathbb{C}$ that vanishes on T in fact vanishes on Z_f. For such T, the set of polynomials in $R[M]$ that vanish on all points on T is precisely Λ_M.

In the discussion that follows let c_1, c_2, c_3, c_4 be constants depending only on d, n and M. We claim that there is $c_1 > 0$ so that, if $X \geq c_1(1 + \log \|f\| + \log(H))$, then one may choose a set $\mathcal{P}$ of primes such that

(1) $\mathcal{P} \subset [X, 2X]$.
(2) $|\mathcal{P}| \leq c_1(1 + \log \|f\| + \log(H))$.
(3) For any point $x \in Z_f(\mathbb{Q}) - Z_{\text{sing}}(\mathbb{Q})$ of height $\leq H$, there exists $p \in \mathcal{P}$ such that $\nabla f(\tilde{x}) \not\equiv 0 \, (\text{mod}\, p)$. Here $\tilde{x}$ is a primitive integral representative for x, and $\nabla f := (\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_{n+2}})$.
(4) Any $p \in \mathcal{P}$ does not divide $\text{ht}(\Lambda_M)$.

Indeed: first note that there is is a constant c_2 so that, for any x as in (3), all coordinates of $\nabla f(\tilde{x})$ are bounded above by $c_2 \|f\| H^{d-1}$. Next, there is c_3 such that $\text{ht}(\Lambda_M) \leq c_3 \|f\|^{2(\gamma - \delta)}$. It is then easy to see that, if $\mathcal{Q}$ is a set of primes satisfying

$$\prod_{q \in \mathcal{Q}} q \geq (c_2 \|f\| H^{d-1}) \cdot c_3 \|f\|^{2(\gamma-\delta)}, \tag{3}$$

then taking $\mathcal{P}$ to be $\{q \in \mathcal{Q} : (q, \text{ht}(\Lambda_M)) = 1\}$ will satisfy the conclusions (3) and (4) above. Recalling that $\log(\prod_{X \leq p \leq 2X} p) \gg X$, we easily deduce the claimed assertion.

For each $(p \in \mathcal{P}, \overline{x} \in Z_f(\mathbb{F}_p))$ satisfying the conditions of (3) above, i.e. that $\nabla f(\overline{x}) \neq 0$, we consider the subset $S_{p,\overline{x}}$ of elements in $Z_f(\mathbb{Q})$ of height $\leq H$ that reduce to $\overline{x}$ mod p. By (1) above and the obvious bound $|Z_f(\mathbb{F}_p)| \ll_d p^n$, there are at most $O(X^n |\mathcal{P}|)$ such subsets and by (3) above the sets $S_{p,\overline{x}}$ cover all points of $Z_f - Z_{\text{sing}}$ of height $\leq H$.

Choose any such $(p, \overline{x})$ and set $S \equiv S_{p,\overline{x}}$. If S is not in M-general position, it follows that the image of S under the projective embedding of Z_f corresponding to $\mathscr{O}(M)$ lies in a proper hyperplane. The irreducibility of Z_f guarantees that this hyperplane section does not contain a component of Z_f. It follows that *either* S is contained in a proper divisor on Z_f of degree $\leq Md$, *or* S is in M-general position.

Assume then that we are in the latter case. We will eventually show that if we choose X large enough (see (6)) this cannot happen. In any case, choosing a primitive integral representative for each point in S, we observe that each point in S gives a linear functional (evaluation at the chosen representative) on $R[M]$, so we can regard S as a subset of the dual lattice $R[M]^*$. Let $\mathbb{Z}.S$ be the lattice spanned by S inside $R[M]^*$. In view of the fact that S is a set of points in M-general position, it follows that

$$(\mathbb{Z}.S)^{\perp} := \{\lambda \in R[M] : s(\lambda) = 0 \ \forall \ s \in S\} = \Lambda_M.$$

Thus $\text{rank}(\mathbb{Z}.S) = \delta$; moreover, it is easy to deduce from the facts about height stated earlier that $\text{ht}(\Lambda_M) | \text{ht}(\mathbb{Z}.S)$.

It will now be our aim to show that – if X is large enough – $\text{ht}((\mathbb{Z}.S)^{\perp})$ is "small", contradicting $(\mathbb{Z}.S)^{\perp} = \Lambda_M$. The intuition is as follows: to say that $\text{ht}((\mathbb{Z}.S)^{\perp})$ is "small" is to say there are "many" functions that vanish on S. On the other hand, one expects that it is "easier" for a function to vanish on a set of points if those points are close together. The points of S are, by choice, p-adically close; it is therefore reasonable to expect that $\text{ht}((\mathbb{Z}.S)^{\perp})$ is unusually small. What follows formalizes this vague idea.

We may express $\text{ht}(\mathbb{Z}.S)$ concretely as follows. Regarding S as a subset of $R[M]^*$, we can pair any element $s \in S$ with any $e \in R[M]$. We thereby obtain a $|S| \times \gamma$ matrix $\langle s, e \rangle_{s \in S, e \in \mathcal{B}[M]}$, and we can compute $\text{ht}(\mathbb{Z}.S)$ in terms of the minors of this matrix. More specifically, we have:

$$\mathrm{ht}(\mathbb{Z}.S) = \mathrm{gcd}_{T \subset S, |T| = \delta} \left(\sum_{E \subset \mathcal{B}[M], |E| = \delta} \det(\langle e, s \rangle_{e \in E, s \in T})^2 \right) \tag{4}$$

From (4), and the fact that all points of $S_{p,\overline{x}}$ have height $\leq H$, one deduces that $\mathrm{ht}(\mathbb{Z}.S)$ is $\ll H^{2.M.\delta}$.

On the other hand, it is proven in [9] (see esp. equation (3.7) and (3.11); cf. also Elkies [8]) that there is a positive integer $c(M,d)$, depending only on M and d and satisfying

$$c(M,d) \sim \frac{dM^n}{n!} \cdot \frac{d^{1/n} M}{(1+n^{-1})} \quad (M \to \infty),$$

so that each determinant appearing in (4) is divisible by $p^{c(M,d)}$.

(Although we will not repeat Heath-Brown's computation in this paper, we note that $c(M,d)$ has an intrinsic interpretation as the order of vanishing of a certain section of a certain line bundle on Z_f^{δ} along the diagonally embedded Z_f. We refer the reader to Remark 1 for the construction of this section, which also highlights the connection with traditional methods in Diophantine approximation.)

Since p does not divide $\mathrm{ht}(\Lambda_M)$, and $\mathrm{ht}(\Lambda_M) | \mathrm{ht}(\mathbb{Z}.S)$, the remarks above show that $\mathrm{ht}(\Lambda_M) \ll H^{2M\delta} p^{-2c(M,d)}$. It now follows from (2) that

$$p \ll \|f\|^{-\frac{|\mathcal{B}[M-d]|}{c(M,d)}} H^{\frac{M\delta}{c(M,d)}}.$$

Taking $M \to \infty$ and apply (1) to obtain the inequality:

$$p \leq c(\epsilon) \|f\|^{-d^{-1-1/n} n^{-1} + \epsilon} H^{(1+n^{-1}) d^{-1/n} + \epsilon}, \tag{5}$$

where $\epsilon \to 0$ as $M \to \infty$, and the constant $c(\epsilon)$ depends on ϵ, d and n.

Now suppose we choose X such that

$$X \geq \max(c_1(1 + \log H + \log \|f\|), c(\epsilon) \|f\|^{-d^{-1-1/n} n^{-1} + \epsilon} H^{(1+n^{-1}) d^{-1/n} + \epsilon}), \tag{6}$$

where c_1 is as discussed in the choice of $\mathcal{P}$. (A caution: this c_1 depends on M, and so implicitly on ϵ also).

With this choice of X, (5) cannot hold; our previous discussion now shows that, for any admissible pair $(p, \overline{x})$, the set $S_{p,\overline{x}}$ is contained in a divisor of degree $O_\epsilon(1)$. The number of pairs $(p, \overline{x})$ is, as previously remarked, $O(|\mathcal{P}| X^n)$. Taking into account condition (2) in the choice of $\mathcal{P}$ yields that all points on $Z_f(\mathbb{Q})$ of height $\leq H$ are contained in at most $(H^{\frac{n+1}{d^{1/n}}} \cdot \|f\|^{-d^{-1-1/n}} + 1) H^\epsilon \|f\|^\epsilon$ divisors.

It remains to remove the factor $\|f\|^\epsilon$, which follows the idea behind [9, Theorem 4]. Indeed, we may assume that all points of height $\leq H$ lie on a unique hypersurface of degree d; if not, all the points on $Z_f(\mathbb{Q})$ are contained in a single divisor and the assertion is trivially true. However, in this case, one can recover f by interpolation, and we see in particular that $\|f\| \ll H^{Ad^{n+2}}$, for some constant A depending only on n. Thus the factor $\|f\|^\epsilon$ is subsumed in H^ϵ. □

The case $n = 1$ is nontrivial: it gives the bound $(H^{2/d} \|f\|^{-1/d^2} + 1) H^\epsilon$ for the number of points of height less than H on the plane curve cut out by f. In the next section, we show that, by treating the cases where $\|f\|$ small and large w.r.t H separately, one can obtain a uniform improvement on $2/d$ for genus ≥ 1. Let us remark that such a result (not just uniform in $\|f\|$, but improving as $\|f\|$ grows) is by no means surprising. For example the improvements in the bounds of [9] as one passes to skew-shaped boxes reflect the same general principle.

Remark 1. As remarked above, the quantity $c(M,d)$ is the order of vanishing, along the diagonally embedded Z_f, of a certain section of a line bundle on Z_f^δ. Since the construction of this section is the key part of Heath-Brown's original approach in [9] we give it here for completeness. In essence, one can regard the proof of Prop. 1 as giving an "interpretation" of this section in terms of the height of certain lattices. (See also [5, Appendix] for another interpretation of $c(M,d)$).

For $1 \le i \le \delta$ let $\pi_i : (\mathbb{P}^{n+1})^\delta \to \mathbb{P}^{n+1}$ be projection onto the ith factor. Then $\oplus_{i=1}^{\delta} \pi_i^* \mathscr{O}(M)$ defines a vector bundle of rank δ on $(\mathbb{P}^{n+1})^\delta$. Any subset $E \subset \mathcal{B}[M]$ with $|E| = \delta$ gives rise to a certain section of the line bundle $\wedge^\delta \left(\oplus_i \pi_i^* \mathscr{O}(M)\right)$ on $(\mathbb{P}^{n+1})^\delta$. Namely, fix an ordering $e_1, \ldots, e_\delta$ for E, and take the section defined in local coordinates by the rule

$$(P_1, \ldots, P_\delta) \mapsto \det(e_i(P_j))_{1 \le i,j \le \delta}, \qquad P_j \in \mathbb{P}^{n+1}. \tag{7}$$

Let us make some very crude comments which might suggest the intuition behind the construction of this section. Intrinsically, one aims to construct, by analogy with arguments from Diophantine approximation, a section vanishing to high order along the diagonal $Z_f \hookrightarrow Z_f^\delta \to (\mathbb{P}^{n+1})^\delta$; the existence of such a section turns out to present an obstacle to the clustering of rational points on Z_f. Now, an appropriate choice of E will define an embedding $\iota_E : Z_f \to \mathbb{P}^{\delta-1}$ by the rule

$$\iota_P(E) = (e_1(P) : \ldots : e_\delta(P)).$$

The value of the section defined above on $(P_1, \ldots, P_\delta)$ should be (crudely) thought of as measuring the "volume" of the simplex $(\iota_E(P_1), \ldots, \iota_E(P_\delta))$. This description suggests that if $P_1, \ldots, P_\delta$ are "close" in a real or p-adic topology, then the corresponding simplex would be "small", and therefore one expects that the value of this section to be "small" in the real or p-adic valuation. If this effect is strong enough compared to the height of the value of the section, we can show that the value of the section is actually equal to 0. (The protoypical example: if a rational number x of height H reduces mod p to 0, and $p > H$, then $x = 0$.)

Remark 2. The result also implies a result for integral points. Let $f \in \mathbb{Z}[x_1, \ldots, x_{n+1}]$ be an irreducible polynomial of degree d, and let f_k be the degree k component of f. Define $F \in \mathbb{Z}[x_1, \ldots, x_{n+1}, x_{n+2}]$ by the rule $F = \sum_{i=0}^{d} H^i x_{n+2}^{d-i} f_i$. Then F is homogeneous of degree d and irreducible, and it is clear that $\|F\| \gg H^d \|f_d\|$.

Now each integral solution $(x_1, x_2, \ldots, x_{n+1})$ gives rise to a point $(x_1, \ldots, x_{n+1}, x_{n+2} = H)$ on the projective n-fold defined by $F = 0$. It follows from Prop. 1 that for each $\epsilon > 0$, there exists a constant $c = c(n, \epsilon)$ such that the set of integral solutions to $f(x_1, \ldots, x_n, x_{n+1}) = 0$ with $|x_i| \le H$ can be covered by $H^{\frac{n}{d^{1/n}}+\epsilon} \|f_d\|^{-d^{-1-1/n}} + H^\epsilon$ proper subvarieties, each of degree $O_{\epsilon,d,n}(1)$.

Taking $n = 1$, this shows that the number of integral points on an irreducible plane curve $f(x,y) = 0$ is $\ll_\epsilon H^{1/d+\epsilon} \|f_d\|^{-1/d^2} + H^\epsilon$. This represents a very slight refinement of a result of Bombieri and Pila [4].

In connection with obtaining uniform improvements of [4] in genus ≥ 1, it would be desirable to obtain better dependence on f; in particular, it would be nice to replace $\|f_d\|$ by a different norm that would also take into account the size of lower degree terms of f. However, as the example $y = (x-N)^2$ shows, such a result must take a different form to the rational case. Indeed, the curve $f_N(x,y) = y - (x-N)^2$

has its largest coefficient of size N^2. On the other hand, the number of solutions to $f_N = 0$ in the box $\{|x|, |y| \leq 2N\}$ is $\gg N^{1/2}$. Thus one cannot naively replace $\|f_d\|$ by $\|f\|$ in the foregoing remarks.

Remark 3. The proof may be applied without essential modification to subvarieties $Z \subset \mathbb{P}^m$ of arbitrary codimension. In this context $\|f\|$ should be replaced by a *subvariety height* of Z (in the sense of Bost-Gillet-Soulé [2]). Indeed, $\|f\|$ enters the argument only through the computation of the height of the lattice Λ_M. In the general setting, the computation of the height of the (analogous) lattice is given by the arithmetic Hilbert-Samuel formula [15], [2, 3.2.5]. We have not written it in this generality largely because of the following technical difficulty: the Hilbert-Samuel formula as stated in the literature is not *a priori* sufficiently uniform in Z. It is seems extremely likely that a uniform version should exist, but we do not know of a reference.

This generalization should also encompass the situation of counting points in "non-square boxes", by altering the archimedean metric on $\mathscr{O}(1)$ appropriately.

We remark that such a result and its variants should be useful in bounding the number of rational points on higher-dimensional varieties. (The idea will be to keep track at all stages not only of the number of points of bounded height, but also of the number of cycles – of all intermediate dimensions – of bounded height.)

3. Uniform improvement on $2/d$ for non-rational curves.

Throughout this section we will use capital letters H, H_0 to denote exponential heights, and lower-case h, h_0 for logarithmic heights.

3.1. Introduction.

We now turn to the proof of Theorem 1. Suppose we want to prove a bound better than $H_0^{2/d}$ for the number of $\mathbb{Q}$-points of height $\leq H_0$ on a plane curve given by a single equation $f(x, y, z) = 0$.

The idea is that one may restrict (by Prop. 1) to the case where $\|f\|$ is very small compared to H_0, i.e. $\|f\| \asymp H_0^{\delta}$ for some δ small. In this range we can hope to profitably pass to the Jacobian and make the proof of Néron's bound $O_\epsilon(H_0^\epsilon)$ effective. While this fairly accurately describes the plan of attack, there are some technicalities involved in carrying it out. We shall proceed in the following steps.

1. Reduction to a problem about smooth curves, allowing us to talk about Jacobians more easily. We carry out this part in Section 3.2, essentially by normalization.
2. To control points on the curve of low height (say $\leq H_0^\alpha$ for some α much less than 1) it will suffice to apply trivial bounds.
3. To control points on the curve of intermediate height (between H_0^α and H_0) we use descent. If there were many points of this intermediate height, one could construct (by an appropriate version of descent; there is more than one way to proceed!) yet more points of small height ($\leq H_0^\alpha$). Then again the trivial bounds yield the desired conclusion. This is carried out in Sec. 3.3.

Let us remark that a naive approach to the third step would be as follows: bound the rank of the Mordell-Weil lattice by descent, and then use trivial bounds for the number of points in a lattice inside a large ball. Unfortunately, when one implements this most directly, the resulting bound is not even polynomial in H_0.

This gives rise to the need for the slightly more elaborate approach used in the third step.

There are various alternate ways of dealing unconditionally with the "intermediate case" of height between H_0^α and H_0, and it seems likely that these might lead to better bounds. We have only sketched them here, in Sec. 4.2.1 and Sec. 4.2.2. One is based on Vojta's work and works only in genus ≥ 2. The other uses Mumford's orthogonality relation together with an interesting feature of the geometry of $\mathbb{R}^N$: one cannot pack too many vectors at mutual angles near $\pi/2$. This idea was used in [11] to bound 3-torsion in class groups. A striking feature of this method is that the exponent $2/d$ emerges naturally from the geometry of the Mordell-Weil lattice and the Picard group!

3.2. **Bounds for rational points on families of non-rational curves.** In this section we state a theorem about upper bounds for the density of rational points on smooth genus g curves which are *uniform* as the curves vary in families; this theorem, in combination with a normalization argument, will prove Theorem 1.

We denote by $H_{\mathbb{P}^m}, h_{\mathbb{P}^m}$ the usual exponential and logarithmic heights on $\mathbb{P}^m(\overline{\mathbb{Q}})$: in particular, for $(x_1, \ldots, x_{m+1}) \in \mathbb{Z}^{m+1}$ we have $h_{\mathbb{P}^m}([x_1 : x_2 : \cdots : x_{m+1}]) = \log(\max_i |x_i|)$. Let $B = \mathbb{P}^{\frac{(d+1)(d+2)}{2}-1}$ be the projective space parametrizing plane curves of degree d, and let $P \to \mathbb{P}^2_B$ be the universal flat family of plane curves.

Proposition 2. *Let $W \subset B$ be an irreducible subvariety. Let $\mathcal{C} \to W$ be a morphism whose generic fiber is geometrically connected and smooth of genus at least 1. Suppose given a finite morphism $\mathcal{C} \to P$ over W.*

For each point $p \in \mathcal{C}(\mathbb{Q})$, we write $H_{\mathbb{P}^2}(p)$ for the height of the image of p in $\mathbb{P}^2(\mathbb{Q})$ (under the map $\mathcal{C} \to P \to \mathbb{P}^2$.)

Then there exists a Zariski open set $U \subset W$, and constants $c, \delta > 0$ depending only on $\mathcal{C} \to W$, such that for any $u \in U(\mathbb{Q})$ and any $H_0 \geq 1$,

$$\#\{p \in \mathcal{C}_u(\mathbb{Q}) : H_{\mathbb{P}^2}(p) \leq H_0\} \leq cH_0^{2/d-\delta}.$$

Proposition 2 will be proven in Sec. 3.3. We now explain how to deduce Theorem 1 from Proposition 2.

Let $V \hookrightarrow B$ be a locally closed irreducible subvariety of B and let $P_V \hookrightarrow \mathbb{P}^2_V$ be the universal plane curve over V; we suppose that the generic fiber of $P_V \to V$ is a geometrically reduced and irreducible plane curve of geometric genus at least 1. Now let $\tilde{P}_V \to P_V$ be the normalization of P. By Serre's criterion, the singular locus of $\tilde{P}_V$ has codimension at least 2. Since normalization commutes with localization, the generic fiber of $\tilde{P}_V \to V$ is in fact a smooth geometrically connected curve of genus at least 1. The map $\tilde{P}_V \to P_V$ is an isomorphism away from some proper closed subscheme $Z \subset P_V$; let V_1 be an open dense subset of V such that, for all $v \in V_1$, the fiber P_v is not contained in Z. Then the size of $P_v \cap Z(\mathbb{Q})$ is uniformly bounded above for all $v \in V_1(\mathbb{Q})$.

Now we can apply Prop. 2 with $\mathcal{C} = \tilde{P}_V \times V_1, W = V_1$ and conclude that there is an open dense subset $V_2 \subset V_1$ and a constant c_V such that, for every $v \in V_2(\mathbb{Q})$ and every $H_0 \geq 1$,

$$\#\{p \in \tilde{P}_v(\mathbb{Q}) : H_{\mathbb{P}^2}(p) \leq H_0\} \leq c_V H_0^{2/d-\delta}.$$

Moreover, for every $v \in V_2(\mathbb{Q})$ the map $\tilde{P}_v \to P_v$ is an isomorphism on the complement of the proper closed subset $Z \cap P_v$. It follows that all the rational points of P_v of height at most H_0 lift to rational points of $\tilde{P}_v$, with the possible exception of those points in $(Z \cap P_v)(\mathbb{Q})$. The cardinality of $(Z \cap P_v)(\mathbb{Q})$ is bounded as v varies over $V_2(\mathbb{Q})$. So there exists a constant c'_V such that

$$\#\{p \in P_v(\mathbb{Q}) : H_{\mathbb{P}^2}(p) \le H_0\} \le c'_V H_0^{2/d-\delta} \tag{8}$$

for all $v \in V_2(\mathbb{Q})$ and every $H_0 \ge 1$.

Let $B' \subset B$ be the constructible subset of B consisting of all $b \in B$ such that P_b is geometrically reduced and irreducible. (See [3, (9.7.7)] for the constructibility.) We may write B' as a finite union $\cup_i V_i$ of locally closed irreducible subvarieties. For each i there is an open dense subset $V_{i,2} \subset V_i$ such that (8) holds for all $v \in V_{i,2}(\mathbb{Q})$, for some constant c'_{V_i}. Let B_1 be the complement of $\cup_i V_{i,2}$ in B'; then the dimension of B_1 (the largest dimension of any Zariski point of B_1) is strictly less than that of B'. Proceeding by induction, we obtain a decomposition of B' into a finite union of locally closed subvarieties W_j such that (8) holds for all j and all $v \in W_j(\mathbb{Q})$, for some constants c'_{W_j}. In particular, (8) holds for all $v \in B'(\mathbb{Q})$, taking the constant $c'_{B'}$ to be the maximum of all constants c'_{W_j}. This is precisely the conclusion of Theorem 1.

3.3. **Proof of Prop. 2.**

Proof. (of Prop. 2). Let $\mathcal{C}$ and W be as in the statement of Prop. 2. We may assume W is defined over $\mathbb{Q}$. Let T be the closure of W in B, and let $h_T : T(\overline{\mathbb{Q}}) \to \mathbb{R}$ be a (logarithmic) Weil height function associated to the natural projective embedding $T \hookrightarrow B$; by translating if necessary, we may assume that h_T takes only non-negative values.

Replacing W with an open dense subset if necessary, we may assume that $\mathcal{C} \to W$ has smooth connected fibers, and that the Jacobian $\mathscr{J}/W$ of $\mathcal{C}/W$ is an abelian scheme projective over W. Fix a projective embedding $\mathscr{J} \xrightarrow{\iota} \mathbb{P}^M_W$. We choose this embedding to be symmetric; that is, for each fiber $\mathscr{J}_w$ the divisor class of $\iota_w^*(\mathscr{O}(1))$ is invariant under $P \mapsto -P$.

For each $w \in W$ and $P, Q \in \mathcal{C}_w$ the expression $(P) - (Q)$ defines a divisor of degree 0; this yields a morphism of W-schemes $\phi : \mathcal{C} \times \mathcal{C} \to \mathscr{J}$.

We shall need the following standard results. All constants are understood to depend on $\mathcal{C}, W$, and the choice of projective embeddings of W and $\mathscr{J}$.

(1) Let $w \in W(\overline{\mathbb{Q}})$. On each fiber $\mathscr{J}_w$ of the Jacobian, the height function $h_{\mathbb{P}^M} \circ \iota$, differs from the canonical height $\hat{h}_w$ by a bounded amount. Indeed, for $w \in W(\overline{\mathbb{Q}})$, there exists an constant a_1 such that $|h_{\mathbb{P}^m} \circ \iota - \hat{h}_w| \le a_1(1 + h_T(w))$ on $X_w(\overline{\mathbb{Q}})$. (See [13, Thm. A]. The assumption of symmetry enters to assure that $\hat{h}_w$ is genuinely quadratic, i.e. $\hat{h}_w(-P) = \hat{h}_w(P)$.)

(2) There are constants a_2, a_3 such that for $w \in W(\overline{\mathbb{Q}}), P, Q \in \mathcal{C}_w(\overline{\mathbb{Q}})$ we have:

$$\hat{h}_w(\phi(P,Q)) \le \frac{a_2}{2}(h_{\mathbb{P}^2}(P) + h_{\mathbb{P}^2}(Q)) + a_3(1 + h_T(w)). \tag{9}$$

By (1) it suffices to prove the claim when the left-hand side is replaced by $h_{\mathbb{P}^M} \circ \iota(\phi(P,Q))$. The argument is routine; we shall just state the point in words. For $w \in W(\overline{\mathbb{Q}})$ and $P, Q \in \mathcal{C}_w(\overline{\mathbb{Q}})$, the coordinates of $\iota \circ \phi(P,Q)$ are algebraic over the coordinates of P, Q considered in $\mathbb{P}^2$, and the coefficients of this algebraic relation are some algebraic functions of w. (9) immediately follows.

(3) There exists a constant a_4 such that, for all $w \in W(\mathbb{Q})$, we have $\operatorname{rank} \mathscr{J}_w(\mathbb{Q}) \leq a_4(1 + h_T(w))$. This follows by descent (implicitly using the trivial upper bound on class numbers).

(4) There is a constant a_5 such that the unit ball in $\mathbb{R}^n$ can be covered by $\exp(a_5 n)$ balls of radius $1/2$.

Now choose $w \in W(\mathbb{Q})$, and set $h_0 = \log(H_0)$. Let $N_w(h_0)$ be the number of points $P \in \mathcal{C}_w(\mathbb{Q})$ with $h_{\mathbb{P}^2}(P) \leq h_0$. In view of Claim (2) there are at least $N_w(h_0)$ points $J \in \mathscr{J}_w(\mathbb{Q})$ with canonical height $\hat{h}_w(J) \leq R = a_2 h_0 + a_3(1 + h_T(w))$. (Indeed, fix Q with $h_{\mathbb{P}^2}(Q) < h_0$ and consider points of the form $\phi(P, Q)$. These are mutually inequivalent since the genus of $\mathcal{C}_w$ is nonzero.)

Let $j \in \mathbb{N}$. We can cover any $\sqrt{R}$-ball in the Mordell-Weil lattice of $\mathscr{J}_w(\mathbb{Q})$ by $\exp(j a_5 \operatorname{rank}(\mathscr{J}_w))$ balls of radius $\sqrt{R}/2^j$. The idea of the argument is now as follows: if $N_w(h_0)$ is very big, at least one of these balls must contain "many" points by the pigeonhole principle; then the $\sqrt{R}.2^{1-j}$ ball around the origin also contains many points, i.e. there are many points with canonical height $\leq 4^{1-j}R$.

In fact, using Claims (3) and (4) we see that there exist at least

$$\exp(-j a_4 a_5 (1 + h_T(w))) N_w(h_0) \tag{10}$$

points of $\mathscr{J}_w(\mathbb{Q})$ with canonical height $\leq 4^{1-j}(a_2 h_0 + a_3(1 + h_T(w)))$. By Claim (1), all these points are sent by ι to points in $\mathbb{P}^M$ of height at most $4^{1-j}(a_2 h_0 + a_3(1 + h_T(w)) + a_1(1 + h_T(w)))$.

Now the number of points $P \in \mathbb{P}^M(\mathbb{Q})$ of height $\operatorname{ht}_{\mathbb{P}^M}(P) \leq h'$ is $\ll \exp((M + 1)h')$. Applying this trivial bound to (10), we deduce:

$$N_w(h_0) \leq \exp(C(1 + h_T(w))) \cdot H_0^{4^{1-j} a_2 (M+1)} \tag{11}$$

where C is a constant depending on $a_1, \ldots, a_5, M$ and j. Taking $j \to \infty$, we have shown that one may obtain bounds for $N_w(h_0)$ involving an arbitrarily small power of H_0, at the price of a large, but in principle explicit, power of $h_T(w)$.

On the other hand, the image of $\mathcal{C}_w$ in $\mathbb{P}^2_w$ is a curve Z_{f_w}, for some $f_w \in \mathbb{Q}[x_0, x_1, x_2]$. By definition of the height h_T, we have $\|f\| \asymp \exp(1 + h_T(w))$. Moreover, the map $\mathcal{C}_w \to Z_{f_w}$ is finite, with degree bounded as w varies. It now follows from Proposition 1 that there is a constant $C' > 0$ so that

$$N_w(h_0) \leq C_{\epsilon,d} \left(\exp(-C'(1 + h_T(w))) H_0^{2/d+\epsilon} + H_0^{\epsilon} \right). \tag{12}$$

Choosing j sufficiently large, and combining (11) and (12), we obtain the conclusion of Prop. 2. □

4. Complements.

4.1. **Explicit constants for cubics – l-adic descent.** We shall now give an explicit version of Thm. 1 for smooth cubic curves, justifying the assertion $\delta(3) > \frac{1}{450}$ stated in the introduction.

Although the constants involved in the proof just presented of Prop. 2 may be easily found in the literature, we shall use a slightly different approach, mainly for illustrative purposes. Indeed, the proof of Prop. 2 partitions the real vector space $J(\mathbb{Q}) \otimes_{\mathbb{Z}} \mathbb{R}$, where J is the Mordell-Weil lattice of the Jacobian of the curve, into small archimedean regions. Similarly, the argument below operates by partitioning $J(\mathbb{Q}) \otimes_{\mathbb{Z}} \mathbb{Z}_\ell$ into small ℓ-adic regions, for suitably chosen ℓ.

Let $C/\mathbb{Q}$ be a smooth cubic curve in $\mathbb{P}^2$ with equation $f = 0$, where f is an irreducible cubic polynomial with integral coefficients:

$$\begin{aligned}(13)\quad & f(x,y,z)\\ &= a_1x^3+a_2y^3+a_3z^3+a_4x^2y+a_5xy^2+a_6x^2z+a_7xz^2+a_8y^2z+a_9y^2z+a_{10}xyz.\end{aligned}$$

Put $|C| = \log(\max_i |a_i|)$. Let $N_C(h)$ be the number of rational points $P \in C(\mathbb{Q})$ whose logarithmic height $h_{\mathbb{P}^2}(P)$ is at most h.

Proposition 3. *For every positive integer m and every $\epsilon > 0$, we have:*

$$\log N_C(h) \le (1+\epsilon)\left(\frac{32}{3m^2}h + \left(\frac{32}{m^2} + \frac{6}{\log 2}\log m + 6\right)|C|\right) + B(m,\epsilon). \tag{14}$$

for some constant $B(m,\epsilon)$ depending only on m and ϵ.

Proof. The computations in [1] give equations for a degree 9 map $\psi : C \to E$, where E is the Jacobian of C placed in Weierstrass form. One has

$$\psi(x:y:z) = (\Theta(x,y,z)H(x,y,z) : J(x,y,z) : H(x,y,z)^3)$$

where H, Θ, J are forms in x,y,z of degree $3,6,9$ respectively. The coefficients of x,y,z in H (resp. Θ, J) are polynomials in the a_i of degree 2 (resp. $6,9$). It follows that for each $P \in C(\mathbb{Q})$, one has $h_{\mathbb{P}^2}(\psi(P)) \le 9h_{\mathbb{P}^2}(P) + 9|C| + O(1)$. (Every $O(1)$ in this discussion refers to an *absolute* constant.)

Let $\hat{h}$ be the canonical height on E. We know (see, e.g. [16]) there are constants a, A such that

$$-a|C| + O(1) < \hat{h}(\psi(P)) - (1/3)h_{\mathbb{P}^2}(\psi(P)) < A|C| + O(1) \tag{15}$$

So we obtain in fact $\hat{h}(\psi(P)) < 3h_{\mathbb{P}^2}(P) + (3+A)|C| + O(1)$.

Next, we note that the image of $\psi(C(\mathbb{Q}))$ in $E(\mathbb{Q})$ is a coset of $3E(\mathbb{Q})$. Fix some point $P_0 \in C(\mathbb{Q})$ of height at most h. Then for each $P \in C(\mathbb{Q})$ satisfying $h_{\mathbb{P}^2}(P) \le h$, we have that $\hat{h}(\psi(P) - \psi(P_0)) < 12h_{\mathbb{P}^2}(P) + (12+4A)|C| + O(1)$. Since $\psi(P) - \psi(P_0) = 3(P - P_0)$, we obtain

$$\hat{h}(P - P_0) < \frac{4}{3}h_{\mathbb{P}^2}(P) + \left(\frac{4}{3} + \frac{4A}{9}\right)|C| + O(1). \tag{16}$$

We have now reduced our problem to bounding the number of points of bounded canonical height on an elliptic curve in Weierstrass form. Write $\hat{N}_E(h)$ for the number of points on E of canonical height at most h.

Now let $m > 1$ be an integer; we are going to count the points in $E(\mathbb{Q})$ by counting each coset of $mE(\mathbb{Q})$ separately. First of all, the number of points in $mE(\mathbb{Q})$ with canonical height at most h is just $\hat{N}_E(h/m^2)$. On the other hand, if $mE(\mathbb{Q}) + Q$ is a coset of $mE(\mathbb{Q})$, and there are N points in $m(E(\mathbb{Q})) + Q$ of canonical height at most h, then by taking differences we see that there are N points in $m(E(\mathbb{Q}))$ of canonical height at most $4h$. (Note the similarity with the proof of Prop. 2!)

We conclude that $\hat{N}_E(h) \le |E(\mathbb{Q})/mE(\mathbb{Q})|\hat{N}_E(4h/m^2)$. Clearly $\log|E(\mathbb{Q})/mE(\mathbb{Q})|$ is bounded above by $(r+2)\log m$, so

$$\log \hat{N}_E(h) \le (r+2)\log m + \log \hat{N}_E(4h/m^2). \tag{17}$$

Let $N_E(h)$ be the number of points on $E(\mathbb{Q})$ with $h_{\mathbb{P}^2} \leq h$. By the theorem of Heath-Brown [9], $N_E(h) \leq \exp((\frac{2}{3}+\epsilon)h + O_\epsilon(1))$. So by (15) we find that

$$\log \hat{N}_E(\frac{4h}{m^2}) < \log N_E(\frac{12h}{m^2} + 3a|C| + O(1)) < \frac{(8+\epsilon)h}{m^2} + (2+\epsilon)a|C| + O_\epsilon(1),$$

which in combination with (17) shows that

$$\log \hat{N}_E(h) < \frac{(8+\epsilon)h}{m^2} + (r+2)\log m + (2+\epsilon)a|C| + O_\epsilon(1).$$

By executing a 2-descent we see, as in [6], that for any c greater than $\frac{1}{2\log(2)}$ we have the rank bound $r < c\log(|\Delta_E|) + O_\epsilon(1)$, where Δ_E is the discriminant of the Weierstrauss cubic defining E. Moreover, from [1, eq. (8) and (9)] one deduces that $\log(\Delta_E) - 12\log|C| = O(1)$.

It follows that $r + 2 < (\frac{6}{\log 2} + \epsilon)|C| + O_\epsilon(1)$. We conclude that

$$\log \hat{N}_E(h) < \frac{(8+\epsilon)h}{m^2} + \left((\frac{6}{\log 2} + \epsilon)\log m + (2+\epsilon)a\right)|C| + O_{m,\epsilon}(1). \tag{18}$$

Now, by (16), we note that $N_C(h) \leq \hat{N}_E(\frac{4}{3}h + (\frac{4}{3} + \frac{4A}{9})|C|)$. Applying (18) and using the result of Zimmer [16] – which shows that a and A can be taken to be 3 and 6 respectively – we obtain (14). □

We have already shown in Proposition 1 that

$$\log N_C(h) < \max((2/3+\epsilon)h - (1/9)|C| + O_\epsilon(1), \epsilon h)$$

Combining this inequality with Proposition 3 yields an upper bound on $N_C(h)$ which is uniform in $|C|$; we observe that our two upper bounds on $\log N_C(h)$ are equal when (up to a constant factor between $1-\epsilon$ and $1+\epsilon$)

$$(\frac{2}{3} - \frac{32}{3m^2})h = (\frac{32}{m^2} + \frac{6}{\log 2}\log m + 6 + 1/9)|C|,$$

so we have in general that

$$\log N_C(h) < \left(\frac{2}{3} + \epsilon - \frac{1}{9}\frac{2/3 - 32/3m^2}{32/m^2 + (6/\log 2)\log m + 55/9}\right)h + O_{m,\epsilon}(1). \tag{19}$$

Now we choose m so as to minimize the coefficient of h in (19). An easy calculation shows this is minimized at $m = 11$. We obtain $\log N_C(h) < (2/3 - \delta)h + O(1)$, where δ can be taken to be, for instance, $1/450$, and the implicit constant in $O(1)$ is absolute.

4.2. **Two alternate proofs of Prop. 2.**

4.2.1. *Approach via Vojta orthogonality.* Using results of de Diego [7], which give uniform versions of Vojta's orthogonality results, we now sketch a proof of Prop. 2 when the genus of the curve $\mathcal{C} \to W$ is ≥ 2. The idea is that de Diego's results show that, on any fiber $\mathcal{C}_w$, there are very few points of large height, and to count points of low height is easy using Prop. 1.

Let notations be as in the statement of Prop. 2. Let T be the closure of W in B, as above, and let $T' \subset W$ be a nonempty open subset on which $\mathcal{C} \to T'$ is smooth.

The following Lemma is a consequence of the main result of [7].

Lemma 1. *There exist constants A, B such that, for any $w \in T'(\mathbb{Q})$, there exists at most* $\exp(A(1 + h_T(w))$ *points $p \in \mathcal{C}_w(\mathbb{Q})$ satisfying $h_{\mathbb{P}^2}(p) \geq B(1 + h_T(w))$.*

We now sketch a second proof of Prop. 2 in the case where relative genus ≥ 2. Set $h_0 = \log(H_0)$, where H_0 is as in the statement of Prop. 2. Fix a very small $\alpha > 0$. We are interested in bounding the number of points $P \in \mathcal{C}_w(\mathbb{Q})$ with $h_{\mathbb{P}^2}(P) \leq h_0$; divide into two cases according to whether $h_T(w) \leq \alpha h_0$ or $h_T(w) \geq \alpha h_0$.

In the first case, apply Lem. 1 to count points $P \in \mathcal{C}_w(\mathbb{Q})$ with height $\geq B(1 + h_T(w))$, and apply the trivial bound to count points $P \in \mathcal{C}_w(\mathbb{Q})$ with height $\leq B(1 + h_T(w))$. In the latter case, apply Prop. 1 directly to count points on the image of $\mathcal{C}_w$ in $\mathbb{P}^2_w$, and use the fact that $\mathcal{C}$ maps to this image with bounded finite degree.

Choosing α sufficiently small proves Prop. 2. This proof is shorter than the one we give in Section 3.3, but has the disadvantage that it does not apply to the genus 1 case, and it uses much more difficult machinery.

4.2.2. *Proof via sphere packing bounds.* Vojta's techniques show that, in genus ≥ 2, the points of a curve repel each other when embedded in its Jacobian. Evidently such a phenomenon does not occur in genus 1, so the technique of Section 4.2.1 fails. We may use a technique from [11] instead which "creates" repulsion by partitioning the points on a curve according to their reductions mod p. This method in fact works for curves of all genera ≥ 1. However, we expect that this method may allow the best estimate for δ. The article [11] deals primarily with integral points on elliptic curves, and in the final section of that paper it is indicated how the results extend to rational points and to curves of genus ≥ 1. We will just reprise the main idea, referring to [11] for a further discussion of the manner in which the exponent $2/d$ arises "naturally" in the method.

Again, take T and h_T as in Sec. 3.3.

Let $w \in W(\mathbb{Q})$. We wish to count points $P \in \mathcal{C}_w(\mathbb{Q})$ with $h_{\mathbb{P}^2}(P) \leq h_0$. Fix a very small α. As in Sec. 4.2.1, we split into the two cases where $h_T(w) \leq \alpha h_0$ and $h_T(w) \geq \alpha h_0$. As in that section, one may deal with the latter case directly by applying Prop. 1.

The crucial case is where $h_T(w) \leq \alpha h_0$. Here the key idea is now to choose an auxiliary prime p and partition $\mathcal{C}_w(\mathbb{Q})$ into the fibers of the reduction map mod p, i.e. the fibers of $\mathcal{C}_w(\mathbb{Q}) \to \mathcal{C}_w(\mathbb{F}_p)$.

Let $\mathscr{J}_w$ be the Jacobian of $\mathcal{C}_w$, $\hat{h}_w$ a canonical height on $\mathscr{J}_w(\overline{\mathbb{Q}})$. Let $j_w : \mathcal{C}_w \to \mathscr{J}_w$ the morphism that associates to $P \in \mathcal{C}_w$ the divisor class $dP - D_w$, where D_w is the pullback of $\mathscr{O}(1)$ under $\mathcal{C}_w \to P_w \to \mathbb{P}^2$ and $d = \deg(D_w)$.

Let $\Delta \subset \mathcal{C}_w \times \mathcal{C}_w$ be the diagonal. The point is that if P, Q belong to the same part of the partition defined above, the theory of local heights forces the Weil height $h_\Delta(P, Q)$ to be unusually large; Mumford's gap principle now shows that that $\hat{h}(j_w(P) - j_w(Q))$ is large. Thus $j_w(P), j_w(Q)$ are well-separated in the Mordell-Weil lattice of $\mathscr{J}_w$. It turns out that by optimizing p the gains from this "well-separatedness" can overwhelm the losses from the further partitioning.

We now apply the Kabatiansky-Levenshtein sphere packings bounds (and optimize p) to control the size of each part of the partition. Piecing together the resulting bounds yields (eventually) another proof of Prop. 2.

References

[1] Sang Yook An, Seog Young Kim, David C. Marshall, Susan H. Marshall, William G. McCallum, and Alexander R. Perlis. Jacobians of genus one curves. *J. Number Theory*, 90(2):304–315, 2001.

[2] J.-B. Bost, H. Gillet and C. Soulé. Heights of projective varieties and positive Green forms. *J. Amer. Math. Soc*, 7:903–1027, 1994.

[3] A. Grothendieck. Elements de geometrie algebrique, IV. *Publ. Math. IHES*, 28, 1966.

[4] E. Bombieri and J. Pila. The number of integral points on arcs and ovals. *Duke Math. J.*, 59(2):337–357, 1989.

[5] T. D. Browning, D.R. Heath-Brown and P. Salberger. Counting rational points on algebraic varieties. Preprint available at `http://maths.ox.ac.uk/ntg/preprints/browning/pila.pdf`

[6] Armand Brumer and Kenneth Kramer. The rank of elliptic curves. *Duke Math. J.*, 44(4):715–743, 1977.

[7] Teresa de Diego. Points rationnels sur les familles de courbes de genre au moins 2. *J. Number Theory*, 67(1):85–114, 1997.

[8] Noam Elkies. Rational points near curves and small nonzero $|x^3 - y^2|$ via lattice reduction. *Lecture Notes in Computer Science* 1838 (proceedings of ANTS-4, 2000; W.Bosma, ed.), 33-63, 2000.

[9] D. R. Heath-Brown. The density of rational points on curves and surfaces. *Ann. of Math. (2)*, 155(2):553–595, 2002.

[10] M. Hindry and J. Silverman. Diophantine geometry, *Graduate texts in mathematics* 201, Springer-Verlag, New York, 2000.

[11] H. Helfgott and A. Venkatesh. Integral points on elliptic curves and 3-torsion in class groups. *Preprint.*

[12] J. Kollar. Shafarevich maps and automorphic forms. Princeton University Press, Princeton, 1995.

[13] Joseph H. Silverman. Heights and the specialization map for families of abelian varieties. *J. Reine Angew. Math.*, 342:197–211, 1983.

[14] Joseph H. Silverman. The difference between the Weil height and the canonical height on elliptic curves. *Math. Comp.*, 55(192):723–743, 1990.

[15] Shouwu Zhang. Positive line bundles on arithmetic varieties. *J. Amer. Math Soc.*, 8:187–221, 1995.

[16] H. Zimmer. On the difference of the Weil height and the Néron-Tate height. *Math. Z.* 174:35–51, 1976.