

Intersecting a plane with algebraic subgroups of multiplicative groups

ENRICO BOMBIERI, DAVID MASSER AND UMBERTO ZANNIER

Abstract. Consider an arbitrary algebraic curve defined over the field of all algebraic numbers and sitting in a multiplicative commutative algebraic group. In an earlier article from 1999 bearing almost the same title, we studied the intersection of the curve and the union of all algebraic subgroups of some fixed codimension. With codimension one the resulting set has bounded height properties, and with codimension two it has finiteness properties. The main aim of the present work is to make a start on such problems in higher dimension by proving the natural analogues for a linear surface (with codimensions two and three). These are in accordance with some general conjectures that we have recently proposed elsewhere.

Mathematics Subject Classification (2000): 11G35 (primary); 11G50, 14G25, 14J20 (secondary).

1. Introduction

For $n \geq 1$ let \mathcal{X} be a variety lying in affine n -space with coordinates x_1, \dots, x_n . We suppose that none of these vanishes identically on \mathcal{X} ; thus we can think of \mathcal{X} as a quasi-affine subset of the group variety \mathbb{G}_m^n defined by the non-vanishing of the coordinates. In this paper we are interested in the intersection of \mathcal{X} with varying algebraic subgroups of \mathbb{G}_m^n restricted only by dimension. Recall that every such subgroup is defined by monomial equations $x_1^{a_1} \cdots x_n^{a_n} = 1$, and its dimension is $n - r$, where r is the rank of the subgroup of \mathbb{Z}^n generated by the exponent vectors (a_1, \dots, a_n) .

In [4] we studied such intersections when $\mathcal{X} = \mathcal{C}$ is an irreducible curve defined over the field \mathbb{Q} of algebraic numbers, and we obtained two results: the first about the boundedness of the absolute height on the set, and the second about the finiteness of the set. In this paper we make a modest start on the case of surfaces $\mathcal{X} = \mathcal{S}$ by proving the analogous results for planes $\mathcal{X} = \mathcal{P}$. These are in accordance with the Bounded Height Conjecture and the Torsion Finiteness Conjecture for general \mathcal{X} made in [7].

Received July 30, 2007; accepted December 13, 2007.

We begin by recalling the results of [4]. It is convenient to write $\mathcal{H}_d = \mathcal{H}_d^{(n)}$ for the union of all algebraic subgroups of \mathbb{G}_m^n of dimension at most d . For brevity we describe a translate of an algebraic subgroup as a coset.

We use the deprived sets \mathcal{X}^{oa} introduced in [7]. For curves, \mathcal{C}^{oa} is simply \mathcal{C} if \mathcal{C} does not lie in any coset of dimension at most $n - 1$, and otherwise \mathcal{C}^{oa} is empty. It is easy to see that $\mathcal{C}^{oa} \cap \mathcal{H}_{n-1}$ lies in the set $\mathbb{G}_m^n(\overline{\mathbb{Q}})$ of algebraic points of \mathbb{G}_m^n . We need a height function on this set; the precise choice is at the moment unimportant but for definiteness we use

$$h(P) = \max\{h(\xi_1), \dots, h(\xi_n)\}$$

for $P = (\xi_1, \dots, \xi_n)$, where $h(\xi)$ denotes the absolute (logarithmic) Weil height. This is not quite the choice made in [4] and [6].

Then [4, Theorem 1, page 1120] implies at once the following:

Theorem A. *Suppose that \mathcal{C} is defined over $\overline{\mathbb{Q}}$. Then $\mathcal{C}^{oa} \cap \mathcal{H}_{n-1}$ is a set of bounded height.*

And [4, Theorem 2, page 1121] implies similarly:

Theorem B. *Suppose that $n \geq 2$ and \mathcal{C} is defined over $\overline{\mathbb{Q}}$. Then $\mathcal{C}^{oa} \cap \mathcal{H}_{n-2}$ is a finite set.*

We also noted that the deprived set \mathcal{C}^{oa} is appropriate for Theorem A in the following sense. If \mathcal{C} does lie in a coset of dimension at most $n - 1$, then $\mathcal{C} \cap \mathcal{H}_{n-1}$ is not a set of bounded height.

And we noted that the situation for Theorem B is not so satisfying. If \mathcal{C} actually lies in a torsion coset of dimension at most $n - 1$; that is, a translate of a subgroup by a torsion point, then $\mathcal{C} \cap \mathcal{H}_{n-2}$ is not a finite set. In [6, Conjecture A, page 2248] we expressed the opinion that this type of coset provides the only obstacle to finiteness. In other words, Theorem B should remain true when \mathcal{C}^{oa} is replaced by the set \mathcal{C}^{ta} , which is \mathcal{C} if \mathcal{C} does not lie in any torsion coset of dimension at most $n - 1$, and otherwise empty. See [7] for a discussion of how this is related to conjectures of Zilber [15] and of Pink [12]. In [6, Theorem, page 2248] we proved this Conjecture A in some special cases, in particular for $n = 5$ (the cases $n = 2, 3, 4$ follow already from the work of [4]).

The proofs in [6] relied on constructing a surface \mathcal{S} in \mathbb{G}_m^3 from the curve \mathcal{C} in \mathbb{G}_m^5 and then considering $\mathcal{S} \cap \mathcal{H}_1 = \mathcal{S} \cap \mathcal{H}_1^{(3)}$ using tools developed by the first and third author in [9] and [14]. These show that most of the points on this intersection have bounded height analogous to the situation in Theorem A above. We noted that the general proof of Conjecture A would probably require appropriate generalizations to $\mathcal{S} \cap \mathcal{H}_d^{(d+2)}$ for surfaces \mathcal{S} in \mathbb{G}_m^{d+2} .

The object of the present paper is to initiate the study of such intersections $\mathcal{S} \cap \mathcal{H}_d$, not only for their applications to curves but also for their intrinsic interest. Thus it is natural first to aim at the analogues of Theorems A and B. Unfortunately the method presented here seems to be effective only for surfaces satisfying a certain

integrality condition, and for simplicity we work out the details for arbitrary planes $\mathcal{S} = \mathcal{P}$.

The nature of the deprived sets was clarified in the work [7], and we even treated \mathcal{X}^{oa} , \mathcal{X}^{ta} for varieties \mathcal{X} of arbitrary dimension. Here we restrict to surfaces \mathcal{S} , for the time being defined over the field \mathbb{C} of complex numbers.

Thus we write \mathcal{S}^{oa} for the deprived set that remains of \mathcal{S} after removing

- (i^o) \mathcal{S} itself if it lies in any coset of dimension at most $n - 1$,
- (ii^o) all irreducible curves in \mathcal{S} that lie in any coset of dimension at most $n - 2$.

We describe the curves in (ii^o) as anomalous curves for \mathcal{S} .

In [7] it is shown among other things that \mathcal{S}^{oa} is an open subset of \mathcal{S} . According to the general Bounded Height Conjecture proposed there, the set $\mathcal{S}^{oa} \cap \mathcal{H}_{n-2}$ should be a set of bounded height when \mathcal{S} is defined over $\overline{\mathbb{Q}}$. Our first result, an analogue of Theorem A, confirms this when \mathcal{S} is a plane \mathcal{P} . If $n \geq 3$ then \mathcal{P} is defined by $n - 2$ linear equations

$$\sum_{j=0}^n \alpha_{ij} x_j = 0 \quad (i = 1, \dots, n - 2) \quad (1.1)$$

with $x_0 = 1$. Again it is easy to see that $\mathcal{P}^{oa} \cap \mathcal{H}_{n-2}$ lies in $\mathbb{G}_m^n(\overline{\mathbb{Q}})$.

Theorem 1.1. *Suppose that the plane \mathcal{P} is defined over $\overline{\mathbb{Q}}$. Then $\mathcal{P}^{oa} \cap \mathcal{H}_{n-2}$ is a set of bounded height.*

We shall deduce from this the following analogue of Theorem B.

Theorem 1.2. *Suppose that $n \geq 3$ and the plane \mathcal{P} is defined over $\overline{\mathbb{Q}}$. Then $\mathcal{P}^{oa} \cap \mathcal{H}_{n-3}$ is a finite set.*

As mentioned in [7], the deprived set \mathcal{P}^{oa} seems to be appropriate for Theorem 1.1 in the same sense as for Theorem A: removals like (i^o) and (ii^o) are necessary. For the reader's convenience we reproduce the arguments of [7] for surfaces \mathcal{S} defined over $\overline{\mathbb{Q}}$.

Regarding (i^o), if \mathcal{S} lies in a coset of dimension at most $n - 1$, then $\mathcal{S} \cap \mathcal{H}_{n-2}$ is not a set of bounded height. To see this we can assume after an automorphism of \mathbb{G}_m^n that x_n is constant, say $\xi \neq 0$, on \mathcal{S} . Then after a permutation of coordinates we can assume that x_1, x_2 are algebraically independent on \mathcal{S} . The projection of \mathcal{S} to \mathbb{G}_m^2 defined by these coordinates therefore contains an open subset of \mathbb{G}_m^2 . If ξ is not a root of unity, the points (ξ^{a_1}, ξ^{a_2}) are Zariski dense as (a_1, a_2) runs over \mathbb{Z}^2 ; thus there are infinitely many (a_1, a_2) for which a point of \mathcal{S} with $x_1 = \xi^{a_1}$, $x_2 = \xi^{a_2}$ can be found. As $x_n = \xi$ these points lie in \mathcal{H}_{n-2} and clearly have unbounded height. If ξ is a root of unity a similar argument with say $(\xi_1, 2^{a_2})$ works for infinitely many roots of unity ξ_1 .

Regarding (ii^o) , the matter is a little less clear. We say that a coset is special if it lies in some algebraic subgroup of dimension at most $n-1$. Let \mathcal{C} be an anomalous curve contained in a special coset J of dimension d satisfying $1 \leq d \leq n-2$. After an automorphism we can assume that x_{d+1}, \dots, x_n are constants ξ_{d+1}, \dots, ξ_n on J ; and since J is special these constants are themselves multiplicatively dependent. We can also assume that x_1 is not constant on \mathcal{C} . If one of ξ_{d+1}, \dots, ξ_n , say ξ , is not a root of unity then we can intersect \mathcal{C} with $x_1 = \xi^a$, and as $a \rightarrow \infty$ we get points in \mathcal{H}_{n-2} with unbounded height. If ξ_{d+1}, \dots, ξ_n are all roots of unity then any point on \mathcal{C} lies in \mathcal{H}_d so in \mathcal{H}_{n-2} . So $\mathcal{C} \cap \mathcal{H}_{n-2}$ is not a set of bounded height.

Thus it might make sense to remove only the anomalous curves in (ii^o) which are contained in special cosets; these could be called special-anomalous.

But if $\tilde{\mathcal{S}}^{oa}$ denotes the set \mathcal{S} deprived as in (i^o) together with the special-anomalous curves in (ii^o) , then, as pointed out in [7], it is not always true that $\tilde{\mathcal{S}}^{oa}$ is open in \mathcal{S} . An example of this exists already for $n = 3$ given by the plane $x_3 = x_1 + x_2$. It is easy to see that the anomalous curves are defined by

$$x_1 = \alpha_1 x_3, \quad x_2 = \alpha_2 x_3 \quad (1.2)$$

with $\alpha_1 \neq 0$, $\alpha_2 \neq 0$ satisfying $\alpha_1 + \alpha_2 = 1$. They are special-anomalous exactly when α_1, α_2 are multiplicatively dependent; as mentioned in [4, page 1119], this happens infinitely often but clearly only countably so. Thus $\tilde{\mathcal{S}}^{oa}$ is not open.

In this example \mathcal{S}^{oa} is empty; an arbitrary point (ξ_1, ξ_2, ξ_3) of \mathcal{S} lies in (1.2) with $\alpha_1 = \xi_1/\xi_3$, $\alpha_2 = \xi_2/\xi_3$. Thus Theorem 1.1 tells us nothing about the points of $\mathcal{S} \cap \mathcal{H}_{n-2} = \mathcal{S} \cap \mathcal{H}_1$.

Anyway if we want to restrict to an open set in Theorem 1.1 then it is indeed \mathcal{P}^{oa} that we must take.

The situation for Theorem 1.2 is not so satisfying, just as for Theorem B. As in [7], for a surface \mathcal{S} defined over \mathbb{C} we write \mathcal{S}^{ta} for the deprived set that remains of \mathcal{S} after removing

- (i^t) \mathcal{S} itself if it lies in any torsion coset of dimension at most $n-1$,
- (ii^t) all irreducible curves in \mathcal{S} that lie in any torsion coset of dimension at most $n-2$.

We describe the curves in (ii^t) as torsion-anomalous curves for \mathcal{S} .

According to the general Torsion Openness Conjecture of [7], \mathcal{S}^{ta} is open in \mathcal{S} . However, as remarked there, this must lie quite deep, as it would imply Conjecture A above for curves \mathcal{C} , even those not defined over $\overline{\mathbb{Q}}$. Again for the reader's convenience, we reproduce the argument of [7]. We claim that

$$(\mathbb{G}_m \times \mathcal{C})^{ta} = \mathbb{G}_m \times \mathcal{D} \quad (1.3)$$

for any curve \mathcal{C} in \mathbb{G}_m^n defined over \mathbb{C} , where \mathcal{D} is what remains of \mathcal{C}^{ta} after removing $\mathcal{C}^{ta} \cap \mathcal{H}_{n-2}$.

Now \mathcal{C}^{ta} is empty if and only if \mathcal{C} is contained in a torsion coset of dimension at most $n - 1$, which in turn happens if and only if $\mathcal{S} = \mathbb{G}_m \times \mathcal{C}$ is contained in a torsion coset of dimension at most n . So in this case both sides of (1.3) are empty.

Otherwise, if \mathcal{C}^{ta} is non-empty, then nothing in (i^t) is removed from \mathcal{S} . And consider an anomalous curve in \mathcal{S} as in (ii^t) . The projection of this down to the last n factors in \mathbb{G}_m^n must lie in a torsion coset of dimension at most $n - 1$. It cannot be of positive dimension, else it would be \mathcal{C} , and \mathcal{C}^{ta} would be empty. So this projection is a single point P , and the anomalous curve was $\mathbb{G}_m \times \{P\}$. As this lies in a torsion coset of dimension at most $n - 2$, we see that P must lie in \mathcal{H}_{n-2} . Conversely it is clear that $\mathbb{G}_m \times \{P\}$ for such a P is anomalous; and this leads to (1.3).

However $\mathcal{C}^{ta} \cap \mathcal{H}_{n-2} \subset \mathcal{C}^{ta} \cap \mathcal{H}_{n-1}$ is at most countable. Thus the openness of (1.3) would indeed imply the finiteness of $\mathcal{C}^{ta} \cap \mathcal{H}_{n-2}$ in accordance with Conjecture A.

According to the general Torsion Finiteness Conjecture of [7], if $n \geq 3$ and \mathcal{S} is defined over $\overline{\mathbb{Q}}$, then $\mathcal{S}^{ta} \cap \mathcal{H}_{n-3}$ is a finite set. It is easy to convince oneself that the deprived set \mathcal{S}^{ta} is appropriate for this problem. The conjecture even just for planes $\mathcal{S} = \mathcal{P}$ would strengthen Theorem 1.2 above. Again the reader is referred to [7] for the relations with the conjectures of Zilber and of Pink.

Returning to planes in \mathbb{G}_m^n , we give a last result that clarifies the nature of the deprived set \mathcal{P}^{oa} in Theorems 1.1 and 1.2, even when \mathcal{P} is not defined over $\overline{\mathbb{Q}}$. Let M be the coefficient matrix in (1.1). Because \mathcal{P} has dimension 2, the matrix M has rank $n - 2$; thus at least one maximal subdeterminant is non-zero. Recall that any variety in \mathbb{G}_m^n has a degree arising from the natural embedding in projective \mathbb{P}_n .

Theorem 1.3. *Suppose that \mathcal{P} is defined over \mathbb{C} . Then the deprived set \mathcal{P}^{oa} is non-empty if and only if every maximal subdeterminant of M is non-zero. In this case $\mathcal{P} = \mathcal{P}^{oa} \cup \Gamma_1 \cup \Gamma$, where Γ_1 is a union of lines and Γ is a union of non-linear curves of degree at most $n - 2$. Further the cardinality of Γ_1 is at most $\frac{1}{8}(n - 2)(n - 1)n(n + 1)$ and the cardinality of Γ is bounded only in terms of n .*

It may be interesting to study further the anomalous curves when \mathcal{P}^{oa} is non-empty. For planes in small dimensions $n = 2, 3, 4$ there are in fact exactly $\frac{1}{8}(n - 2)(n - 1)n(n + 1)$ different lines in Γ_1 , and also Γ is empty. But if $n = 5$ then Γ_1 can have fewer lines and Γ can be non-empty. Thus not all anomalous curves on planes are lines. However, for “generic” planes in any dimension it seems likely that there are exactly $\frac{1}{8}(n - 2)(n - 1)n(n + 1)$ different lines in Γ_1 , and that Γ is empty.

We refer the reader to [7] for more detailed information in case \mathcal{P}^{oa} is empty. For example, Theorem 1 there implies that the cosets in (i^o) and (ii^o) can be defined by equations whose degrees are bounded only in terms of n .

Here is a brief description of how the proofs in this paper are arranged.

In Section 2 we define local heights h_v on $\overline{\mathbb{Q}}^m$ corresponding to valuations v on number fields, and we prove a simple but crucial inequality for these.

In Section 3 we give an estimate for points on \mathcal{P} lying on a proper coset J (that is, of dimension at most $n - 1$) defined by say

$$x_2^{b_2} x_3^{b_3} \cdots x_n^{b_n} = \theta.$$

The intersection $\mathcal{P} \cap J$ probably defines a curve, and the result may be regarded as a one-sided adelic version of Siegel's classical comparison estimate for heights on such a curve. However our dependence on the degree of J is rather bad. In the same section we sharpen the result using simple ideas from diophantine approximation. In fact this has the effect of eliminating the dependence on the degree, at least when J is a torsion coset

$$x_2^{b_2} x_3^{b_3} \cdots x_n^{b_n} = 1.$$

But we have to avoid some exceptional points.

Then in Section 4 we apply these estimates to points P of $\mathcal{P} \cap \mathcal{H}_{n-2}$. Typically we now have a second coset

$$x_1^{b_1} x_3^{b_3} \cdots x_n^{b_n} = 1.$$

The effect is to make the results of Section 3 two-sided. If $a_1 \geq 0, \dots, a_n \geq 0$ are integers and $P = (\xi_1, \dots, \xi_n)$, a consequence is that the height of $\xi_1^{a_1} \cdots \xi_n^{a_n}$ is probably nearly as large as the trivial upper bound. This shows at once the unlikelihood of a relation $\xi_1^{a_1} \cdots \xi_n^{a_n} = 1$ with $a_1 \geq 0, \dots, a_n \geq 0$ not all zero.

To include negative exponents in the picture we use an inductive argument, the details of which are supplied in Section 5.

The proof of our Theorem 1.1 can then be finished quickly in Section 6. The exceptional points lead directly to the anomalous curves.

To deduce Theorem 1.2 we will need as in [4] some lower bounds for heights in the style of the work [1] of Amoroso and David. In [6] we saw the usefulness of a one-dimensional version over abelian extensions due to Amoroso and Zannier [3]. A higher-dimensional generalization to cyclotomic fields has recently been obtained by Amoroso and David in [2], and this we recall in Section 7.

Then in Section 8 we use this to establish a general result which produces finite sets from sets of bounded height; the latter result applied to Theorem 1.1 yields at once Theorem 1.2.

Finally in Section 9 we give the proof of Theorem 1.3, and we add some remarks about what happens in small dimensions.

2. Local heights

Let K be a number field and let $V(K)$ be the set of valuations v normalized such that $[K : \mathbb{Q}]^{-1} \sum_{v \in V(K)} \log \max\{1, |\xi|_v\}$ is the absolute height $h(\xi)$ of ξ in K . Thus

$$h(\xi) = \sum_{v \in V(K)} h_v(\xi) \tag{2.1}$$

for the “local heights” $h_v(\xi) = [K : \mathbb{Q}]^{-1} \log \max\{1, |\xi|_v\}$. It is easy to define h_v on $\overline{\mathbb{Q}}$ so that (2.1) continues to hold for any ξ in $\overline{\mathbb{Q}}$, and even to do the same thing for $\overline{\mathbb{Q}}^m$ (see for example [11, page 287]). Of course

$$h_v(\xi_1, \dots, \xi_m) = [L : \mathbb{Q}]^{-1} \sum_{w \in V(L), w|v} \log \max\{1, |\xi_1|_w, \dots, |\xi_m|_w\} \quad (2.2)$$

for any ξ_1, \dots, ξ_m in $\overline{\mathbb{Q}}$ and any number field L containing $K(\xi_1, \dots, \xi_m)$; here $w|v$ means that the valuations w, v restricted to K are equivalent. Then for any number field L containing K we have

$$h_v(\xi_1, \dots, \xi_m) = \sum_{w \in V(L), w|v} h_w(\xi_1, \dots, \xi_m). \quad (2.3)$$

Also

$$h(\xi_1, \dots, \xi_m) = \sum_{v \in V(K)} h_v(\xi_1, \dots, \xi_m) \quad (2.4)$$

holds for all ξ_1, \dots, ξ_m in $\overline{\mathbb{Q}}$. This is just a technical device to allow temporary uncontrolled field extensions. It will be used for example as follows.

Lemma 2.1. *For $m \geq 1$ suppose that $\tau, \sigma_1, \dots, \sigma_m$ in $\overline{\mathbb{Q}}$ satisfy*

$$\tau^m + \sigma_1 \tau^{m-1} + \dots + \sigma_m = 0. \quad (2.5)$$

Then for any K and any v in $V(K)$ we have

$$h_v(\tau) \leq h_v(m) + h_v(\sigma_1, \sigma_2^{1/2}, \dots, \sigma_m^{1/m})$$

with any determination of the surds.

Proof. Let L be a number field containing K, τ and $\sigma_1, \sigma_2^{1/2}, \dots, \sigma_m^{1/m}$. For any w in $V(L)$ and any ξ_1, \dots, ξ_m in L we have

$$|\xi_1 + \dots + \xi_m|_w \leq \max\{1, |m|_w\} \max\{|\xi_1|_w, \dots, |\xi_m|_w\}.$$

Thus (2.5) gives

$$|\tau|_w^m \leq \max\{1, |m|_w\} \max_{1 \leq i \leq m} |\sigma_i \tau^{m-i}|_w.$$

The second maximum is attained at some $i = i(\tau, \sigma_1, \dots, \sigma_m, w)$ and we deduce

$$|\tau|_w \leq \max\{1, |m|_w\}^{1/i} |\sigma_i|_w^{1/i} \leq \max\{1, |m|_w\} \max\{1, |\sigma_1|_w, \dots, |\sigma_m^{1/m}|_w\}.$$

The far right-hand side is also an upper bound for $\max\{1, |\tau|_w\}$. Thus taking the logarithm, the sum over all $w|v$, and then dividing by $[L : \mathbb{Q}]$ yields the required result. \square

Taking the further sum over all v in $V(K)$ gives

$$h(\tau) \leq \log m + h(\sigma_1, \sigma_2^{1/2}, \dots, \sigma_m^{1/m}).$$

The height on the right-hand side could of course be estimated by

$$h(\sigma_1) + h(\sigma_2^{1/2}) + \dots + h(\sigma_m^{1/m}),$$

but in applications the number m of terms will be large, and so it is important not to do this.

3. Comparison of local heights

If $n \geq 3$ then our plane \mathcal{P} in \mathbb{G}_m^n is defined by (1.1) with $x_0 = 1$ and α_{ij} the entries of a matrix M . Of course here it is the Grassmann coordinates, or maximal subdeterminants, that have particular significance. We will say that \mathcal{P} is non-degenerate if every such maximal subdeterminant is non-zero. This implies that any two x_i, x_j ($1 \leq i < j \leq n$) in (1.1) are independent on \mathcal{P} and that x_1, \dots, x_n are linear polynomials in them. In particular for $i = 1, j = 2$ it will be convenient to write

$$x_k = \beta_k x_1 + \alpha_k x_2 + \gamma_k \quad (k = 3, \dots, n) \quad (3.1)$$

for $\alpha_k, \beta_k, \gamma_k$ in \mathbb{C} . Further these are all non-zero; more generally every maximal subdeterminant of the matrix obtained by adjoining the rows $(\beta_k, \alpha_k, \gamma_k)$ ($k = 3, \dots, n$) to the identity matrix is non-zero. As is usual, when \mathcal{P} is defined over $\overline{\mathbb{Q}}$ we write $\mathcal{P}(\overline{\mathbb{Q}})$ for the set of points of \mathcal{P} defined over $\overline{\mathbb{Q}}$.

Lemma 3.1. *For $n \geq 3$ let \mathcal{P} be non-degenerate and defined over $\overline{\mathbb{Q}}$. Given an integer B there is a constant $C_{\mathcal{P}}(B)$, depending only on \mathcal{P} and B , with the following property. Let $\theta \neq 0$ be in $\overline{\mathbb{Q}}$ and let $P = (\xi_1, \dots, \xi_n)$ be in $\mathcal{P}(\overline{\mathbb{Q}})$ satisfying*

$$\xi_2^{b_2} \dots \xi_n^{b_n} = \theta \quad (3.2)$$

for integers b_2, \dots, b_n with

$$B = |b_2| + \dots + |b_n|. \quad (3.3)$$

Then for any number field K we have

$$\sum_{v \in V(K)} \max\{0, h_v(\xi_2) - h_v(\xi_1)\} \leq 2h(\theta) + C_{\mathcal{P}}(B) \quad (3.4)$$

provided

$$\theta \neq \alpha_3^{b_3} \dots \alpha_n^{b_n} \quad (*)$$

for $\alpha_3, \dots, \alpha_n$ in (3.1).

Proof. What is the meaning of (3.4)? If one forgets the zeroes on the left-hand side, then one sees that $h(\xi_2) \leq h(\xi_1) + 2h(\theta) + C_{\mathcal{P}}(B)$; the height of ξ_2 is not much larger than the height of ξ_1 . This is a one-sided version of the sort of heights comparison estimate that Siegel might have proved. But (3.4) says more; for example each local height $h_v(\xi_2)$ is not much larger than $h_v(\xi_1)$. And this even for the sum over all v or indeed over any set of v whatsoever. Furthermore the error term is bounded, which even Néron did not prove. We will see later how to make everything two-sided.

We start the proof by noting that $(*)$ and (3.2) imply $B \geq 1$. We do the special case $b_2 \geq 0, \dots, b_n \geq 0$ first. We write (3.2) using (3.1) as

$$\xi^{b_2}(\beta_3\eta + \alpha_3\xi + \gamma_3)^{b_3} \cdots (\beta_n\eta + \alpha_n\xi + \gamma_n)^{b_n} = \theta$$

with $\xi = \xi_2, \eta = \xi_1$. We then expand in powers of ξ to get

$$\omega_0\xi^m + \omega_1\xi^{m-1} + \cdots + \omega_m = 0 \quad (3.5)$$

with $m = B$ and

$$\omega_i = U_i(\eta) + \theta V_i(\eta) \quad (i = 0, \dots, m) \quad (3.6)$$

for polynomials U_i, V_i ($i = 0, \dots, m$). The crucial point here is that U_i, V_i have degree at most i ($i = 0, \dots, m$) and their coefficients depend only on \mathcal{P} and b_2, \dots, b_n . Now $\omega_0 = \alpha_3^{b_3} \cdots \alpha_n^{b_n} \neq 0$. So Lemma 2.1 gives for any K and any v in $V(K)$

$$h_v(\xi) \leq h_v(m) + h_v(\omega_1/\omega_0, (\omega_2/\omega_0)^{1/2}, \dots, (\omega_m/\omega_0)^{1/m}). \quad (3.7)$$

The second local height on the right-hand side is at most

$$h_v(\omega_0^{-1}, \omega_0^{-1/2}, \dots, \omega_0^{-1/m}) + h_v(\omega_1, \omega_2^{1/2}, \dots, \omega_m^{1/m}). \quad (3.8)$$

The first term in (3.8) is just $h_v(\omega_0^{-1})$. We estimate the second by taking the field K so large to include $\omega_1, \omega_2^{1/2}, \dots, \omega_m^{1/m}$ as well as the coefficients of U_i, V_i and ξ, η . Now the local height is formed with a single valuation, and by (3.6)

$$\max\{1, |\omega_i|_v\} \leq \max\{1, |2i + 2|_v\} \max\{1, |\theta|_v\} |U_i|_v |V_i|_v \max\{1, |\eta|_v^i\},$$

where $|U_i|_v, |V_i|_v$ are the local heights of the vectors formed with 1 and the coefficients of U_i, V_i respectively. Taking the i -th root and then the maximum over i we get

$$h_v(\omega_1, \omega_2^{1/2}, \dots, \omega_m^{1/m}) \leq h_v(2m + 2) + h_v(\theta) + h_v(\eta) + l_v$$

with say $l_v = \log \prod_{i=1}^m |U_i|_v |V_i|_v$. So from (3.7) and (3.8) we find

$$h_v(\xi) \leq h_v(m) + h_v(\omega_0^{-1}) + h_v(2m + 2) + h_v(\theta) + h_v(\eta) + l_v.$$

The sum $\sum_{v \in V(K)} l_v = \sum_{i=1}^m (h(U_i) + h(V_i))$ for the absolute heights of the polynomials U_i, V_i ($i = 1, \dots, m$), which is at most some C_1 depending only on \mathcal{P} and B . We conclude that

$$\sum_{v \in V(K)} \max\{0, h_v(\xi) - h_v(\eta)\} \leq h(\omega_0) + h(\theta) + C_2$$

for some C_2 also depending only on \mathcal{P} and B .

Finally $\omega_0 = U_0(\eta) + \theta V_0(\eta)$ in this case happens to be just $\alpha_3^{b_3} \cdots \alpha_n^{b_n}$, and so its height is bounded in terms only of \mathcal{P} and B . So we obtain the required result (3.4) with $h(\theta)$ in place of $2h(\theta)$.

If some of the exponents b_2, \dots, b_n are negative, then this has no great effect on the proof. We take the corresponding terms in (3.2) over to the θ . The new left-hand side has degree $\sum_{b_i \geq 0} b_i$ in ξ and the new right-hand side degree $-\sum_{b_i < 0} b_i$ in ξ . If these degrees are equal then there is a danger of cancellation; but this is excluded by (*), which says that then the leading coefficients are unequal. So (3.5) continues to hold, now with $m \leq B$ but still with $m \geq 1$. And $\omega_0 = U_0(\eta) + \theta V_0(\eta)$ may now involve $h(\theta)$ in its height (but still no $h(\eta)$), leading to the extra $2h(\theta)$ in (3.4). This completes the proof. It may seem that we have on the way assumed the field K to be “large”, but in fact there is no restriction because we can immediately descend to arbitrary subfields using (2.3). \square

The condition (*) cannot be omitted in Lemma 3.1. For example, if $n = 3$ and $x_3 = x_1 + x_2 + 1$ in (3.1) then any algebraic $P = (-1, \xi, \xi)$ satisfies $\xi_2^{-1} \xi_3 = 1$ in (3.2), but $h(\xi_2) = h(\xi)$ is not bounded above in terms of $h(\xi_1) = 0$. And indeed it is this (*) which will eventually lead to the anomalous curves.

The next result eliminates the term $C_{\mathcal{P}}(B)$, which may depend badly on B , from (3.4). It suffices for our purposes to treat the case $\theta = 1$ in (3.2), corresponding to a coset that is a subgroup.

Lemma 3.2. *For $n \geq 3$ let \mathcal{P} be non-degenerate and defined over $\overline{\mathbb{Q}}$. Given $\epsilon > 0$ there is a constant $C_{\mathcal{P}}(\epsilon)$, and a finite collection $\mathcal{J}_{\mathcal{P}}(\epsilon)$ of proper cosets of \mathbb{G}_m^n defined over $\overline{\mathbb{Q}}$, all depending only on \mathcal{P} and ϵ , with the following property. Let $P = (\xi_1, \dots, \xi_n)$ be in $\mathcal{P}(\overline{\mathbb{Q}})$ such that ξ_2, \dots, ξ_n are multiplicatively dependent. Then for any number field K we have*

$$\sum_{v \in V(K)} \max\{0, h_v(\xi_2) - h_v(\xi_1)\} \leq \epsilon h(P) + C_{\mathcal{P}}(\epsilon). \quad (3.9)$$

provided P is not in $\mathcal{J}_{\mathcal{P}}(\epsilon)$.

Proof. We have a relation

$$\xi_2^{b_2} \cdots \xi_n^{b_n} = 1 \quad (3.10)$$

for integers b_2, \dots, b_n not all zero. With

$$B = |b_2| + \cdots + |b_n| \geq 1$$

and a real parameter $Q > 1$ we can find integers r_2, \dots, r_n, q such that

$$1 \leq q \leq Q, \quad \left| \frac{b_i}{B} - \frac{r_i}{q} \right| < q^{-1} Q^{-\frac{1}{n-1}} \quad (i = 2, \dots, n); \quad (3.11)$$

see for example [10, Theorem VI, page 13]. After raising (3.10) to the power q , we can rewrite it as

$$\xi_2^{r_2} \dots \xi_n^{r_n} = \theta \quad (3.12)$$

with some θ satisfying

$$\theta^B = \xi_2^{r_2 B - b_2 q} \dots \xi_n^{r_n B - b_n q}. \quad (3.13)$$

As

$$\begin{aligned} \sum_{i=2}^n \frac{|r_i|}{q} &> -(n-1)q^{-1}Q^{-\frac{1}{n-1}} + \sum_{i=2}^n \frac{|b_i|}{B} \\ &= -(n-1)q^{-1}Q^{-\frac{1}{n-1}} + 1 \geq -(n-1)Q^{-\frac{1}{n-1}} + 1 \end{aligned}$$

we see that r_2, \dots, r_n are not all zero provided $Q \geq (n-1)^{n-1}$. In this case

$$R = |r_2| + \dots + |r_n| \geq 1.$$

Applying Lemma 3.1 to (3.12) we get

$$\sum_{v \in V(K)} \max\{0, h_v(\xi_2) - h_v(\xi_1)\} \leq 2h(\theta) + C_{\mathcal{P}}(R) \quad (3.14)$$

subject to (*), which for the time being we assume. From (3.11) and (3.13) we get $h(\theta) \leq (n-1)Q^{-\frac{1}{n-1}}h(P)$ which is at most $\frac{1}{2}\epsilon h(P)$ if further $Q \geq (\frac{2(n-1)}{\epsilon})^{n-1}$. So fixing for example

$$Q = (n-1)^{n-1} \max \left\{ 1, \left(\frac{2}{\epsilon} \right)^{n-1} \right\}$$

we find $2h(\theta) \leq \epsilon h(P)$ in (3.14). Also

$$\frac{R}{q} = \sum_{i=2}^n \frac{|r_i|}{q} < (n-1)q^{-1}Q^{-\frac{1}{n-1}} + \sum_{i=2}^n \frac{|b_i|}{B} \leq 2$$

so $R < 2q \leq 2Q$ is bounded by a function of only n and ϵ . Thus $C_{\mathcal{P}}(R)$ in (3.14) becomes $C_{\mathcal{P}}(\epsilon)$ and we have arrived at (3.9).

But what if (*) fails for (3.12)? Then

$$\xi_2^{r_2} \dots \xi_n^{r_n} = \alpha_3^{r_3} \dots \alpha_n^{r_n},$$

which says that P lies in a proper coset depending only on \mathcal{P} and ϵ . Defining the collection $\mathcal{J}_{\mathcal{P}}(\epsilon)$ to consist of such cosets, we see that the present lemma is proved. \square

4. Heights of monomials

The results of the preceding sections, notably Lemma 3.2, deal with points P on a particular subgroup of \mathbb{G}_m^n of dimension $n - 1$. We now go down to dimension $n - 2$ and the set \mathcal{H}_{n-2} of such subgroups.

Lemma 4.1. *For $n \geq 3$ let \mathcal{P} be non-degenerate and defined over $\overline{\mathbb{Q}}$. Given $\epsilon > 0$ there is a constant $C_{\mathcal{P}}(\epsilon)$, and a finite collection $\mathcal{J}_{\mathcal{P}}(\epsilon)$ of proper cosets of \mathbb{G}_m^n defined over $\overline{\mathbb{Q}}$, all depending only on \mathcal{P} and ϵ , with the following property. Let $P = (\xi_1, \dots, \xi_n)$ be in $\mathcal{P}(\overline{\mathbb{Q}}) \cap \mathcal{H}_{n-2}$. Then for any integers $b_1 \geq 0, \dots, b_n \geq 0$ with $B = b_1 + \dots + b_n$ we have*

$$|h(\xi_1^{b_1} \dots \xi_n^{b_n}) - Bh(P)| \leq B \left(12n^2 \epsilon h(P) + C_{\mathcal{P}}(\epsilon) \right),$$

provided P is not in $\mathcal{J}_{\mathcal{P}}(\epsilon)$.

Proof. Because $\xi_2, \xi_3, \dots, \xi_n$ are multiplicatively dependent, Lemma 3.2 implies

$$\sum_{v \in V(K)} \max\{0, h_v(\xi_2) - h_v(\xi_1)\} \leq \epsilon h(P) + C_1 \quad (4.1)$$

for any number field K , where C_1 (and subsequently C_2, \dots) are constants depending only on \mathcal{P} and ϵ . But equally $\xi_1, \xi_3, \dots, \xi_n$ are multiplicatively dependent, and this leads to the opposite bound

$$\sum_{v \in V(K)} \max\{0, h_v(\xi_1) - h_v(\xi_2)\} \leq \epsilon h(P) + C_2,$$

possibly after enlarging the collection $\mathcal{J}_{\mathcal{P}}(\epsilon)$. Adding this to (4.1) gives

$$\sum_{v \in V(K)} |h_v(\xi_1) - h_v(\xi_2)| \leq 2\epsilon h(P) + C_3,$$

which is the two-sided inequality promised in Section 3. It signifies not only that the global heights of ξ_1 and ξ_2 are roughly the same, but also each of the local heights in an adelic sense.

We can apply the same argument to any pair ξ_i, ξ_j ($i \neq j$). Therefore

$$\sum_{v \in V(K)} |h_v(\xi_i) - h_v(\xi_j)| \leq 2\epsilon h(P) + C_4, \quad (4.2)$$

and this holds even if $i = j$. Thus the local heights of all the ξ_1, \dots, ξ_n are adelically roughly the same; a very strong assertion which implies that there is almost no cancellation when we calculate the height of $\xi_1^{b_1} \dots \xi_n^{b_n}$.

The details of this calculation are as follows. Define

$$c_v = c_v(P) = \max_{i,j} |h_v(\xi_i) - h_v(\xi_j)| \geq 0. \quad (4.3)$$

Then (4.2) gives

$$\sum_{v \in V(K)} c_v \leq 2n^2 \epsilon h(P) + C_0 \quad (4.4)$$

for C_0 also depending only on \mathcal{P} and ϵ . Now

$$h_v(\xi_i) \leq h_v(\xi_j) + c_v \quad (i, j = 1, \dots, n), \quad (4.5)$$

and summing over v using (4.4), we get

$$h(\xi_i) \leq h(\xi_j) + 2n^2 \epsilon h(P) + C_0 \quad (i, j = 1, \dots, n).$$

Taking the maximum over i gives

$$h(P) \leq h(\xi_j) + 2n^2 \epsilon h(P) + C_0 \quad (j = 1, \dots, n).$$

By definition $h(\xi_j) \leq h(P)$, and so

$$|h(\xi_j) - h(P)| \leq 2n^2 \epsilon h(P) + C_0 \quad (j = 1, \dots, n). \quad (4.6)$$

This is the special case of Lemma 4.1 if $b_j = 1$ and all the other $b_i = 0$.

To deduce the general case for $\eta = \xi_1^{b_1} \dots \xi_n^{b_n}$ we take K so large to contain ξ_1, \dots, ξ_n and we note that then

$$h(\eta) = \sum_{v \in V(K)} h_v(\eta) = [K : \mathbb{Q}]^{-1} \sum_{v \in V(K)} \log \max\{1, |\eta|_v\}.$$

With $C_v = \exp(c_v [K : \mathbb{Q}]) \geq 1$ this is at least

$$\begin{aligned} & [K : \mathbb{Q}]^{-1} \sum_{v \in V(K)} \log \left(C_v^{-B} \max\{C_v^B, |\eta|_v\} \right) \\ &= -Bc + [K : \mathbb{Q}]^{-1} \sum_{v \in V(K)} \log \max\{C_v^B, |\eta|_v\} \end{aligned}$$

for $c = \sum_{v \in V(K)} c_v$. So

$$\begin{aligned} Bc + h(\eta) &\geq [K : \mathbb{Q}]^{-1} \sum_{v \in V(K)} \log \max\{C_v^B, |\eta|_v\} \\ &\geq [K : \mathbb{Q}]^{-1} \sum_{v \in V(K), |\eta|_v > C_v^B} \log |\eta|_v. \end{aligned} \quad (4.7)$$

Thus we can focus only on those v with $|\eta|_v > C_v^B$. Now if $|\eta|_v$ is large then some $|\xi_i|_v$ is large, and then the adelic cohesion forces all the other $|\xi_j|_v$ to be large. More precisely

$$\prod_{i=1}^n |\xi_i|_v^{b_i} = |\eta|_v > C_v^B = \prod_{i=1}^n C_v^{b_i},$$

and so there is $i = i(v)$ with $|\xi_i|_v > C_v$. Now (4.5) reads

$$\max\{1, |\xi_i|_v\} \leq C_v \max\{1, |\xi_j|_v\}.$$

It follows that $|\xi_j|_v > 1$ and even

$$|\xi_j|_v > C_v^{-1} \max\{1, |\xi_i|_v\} \quad (j = 1, \dots, n).$$

Therefore

$$|\eta|_v = \prod_{j=1}^n |\xi_j|_v^{b_j} \geq C_v^{-B} \max\{1, |\xi_i|_v\}^B \quad (4.8)$$

in (4.7), or

$$\log |\eta|_v \geq [K : \mathbb{Q}](-Bc_v + Bh_v(\xi_i)).$$

We cannot yet sum over i as in (4.7), because $i = i(v)$. But using (4.5) (with subscripts interchanged) we get

$$\log |\eta|_v \geq [K : \mathbb{Q}](-2Bc_v + Bh_v(\xi_j)) \quad (j = 1, \dots, n).$$

Now (4.7) leads to

$$Bc + h(\eta) \geq -2Bc + B \sum_{v \in V(K), |\eta|_v > C_v^B} h_v(\xi_j) \quad (j = 1, \dots, n). \quad (4.9)$$

The right-hand sum in (4.9) is

$$h(\xi_j) - \sum_{v \in V(K), |\eta|_v \leq C_v^B} h_v(\xi_j) \quad (4.10)$$

and the sum here is small. For if $|\eta|_v$ is not large then no $|\xi_j|_v$ is large else the adelic cohesion would again make $|\eta|_v$ large. More precisely, if $|\xi_j|_v > C_v^2$ for a v as in (4.10), then (4.5) (with i replaced by j , and j replaced by k) would give $|\xi_k|_v > C_v$ ($k = 1, \dots, n$), and so $|\eta|_v > C_v^B$ a contradiction.

So $|\xi_j|_v \leq C_v^2$ in (4.10) and the sum in (4.10) is at most

$$[K : \mathbb{Q}]^{-1} \sum_{v \in V(K)} \log(C_v^2) = 2c.$$

Together with (4.9) all this leads to

$$Bc + h(\eta) \geq -2Bc + Bh(\xi_j) - 2Bc \quad (j = 1, \dots, n).$$

Finally recalling (4.6) we get

$$h(\eta) \geq -5Bc + Bh(\xi_j) \geq -5Bc + Bh(P) - B(2n^2\epsilon h(P) + C_0).$$

As $c = \sum_{v \in V(K)} c_v$ satisfies (4.4) we end up with

$$h(\eta) \geq Bh(P) - 6B \left(2n^2 \epsilon h(P) + C_0 \right),$$

and this is one half (the less easy half) of the present lemma.

The other half

$$h(\eta) \leq \sum_{j=1}^n b_j h(\xi_j) \leq Bh(P)$$

is indeed more easy. This completes the proof. A variation would be first to do it for $b_1 = \dots = b_n = 1$ and then use the positivity trick in [4, bottom of page 1128]. \square

5. Inductive step

Lemma 4.1 makes it unlikely that a point $P = (\xi_1, \dots, \xi_n)$ in $\mathcal{P} \cap \mathcal{H}_{n-2}$ exists satisfying a relation $\xi_1^{b_1} \dots \xi_n^{b_n} = 1$ with exponents $b_1 \geq 0, \dots, b_n \geq 0$. For if P is not in $\mathcal{J}_{\mathcal{P}}(\epsilon)$ then we would conclude

$$Bh(P) \leq B \left(12n^2 \epsilon h(P) + C_{\mathcal{P}}(\epsilon) \right).$$

Choosing $\epsilon = \frac{1}{24n^2}$ gives $h(P) \leq 2C_{\mathcal{P}}(\epsilon)$ and the boundedness of the height.

The exceptional cosets $\mathcal{J}_{\mathcal{P}}(\epsilon)$ here cause no trouble (see later). We can reduce the number of negative exponents by an inductive argument. For example if

$$\xi_1^{b_1} \dots \xi_{n-1}^{b_{n-1}} = \xi_n^{b_n} \tag{5.1}$$

for $b_1 \geq 0, \dots, b_{n-1} \geq 0, b_n \geq 0$ then Lemma 4.1 shows that the heights of the two sides are “asymptotically”

$$(b_1 + \dots + b_{n-1})h(P), \quad b_n h(P)$$

respectively. So $b_1 + \dots + b_{n-1}$ is asymptotic to b_n . Now dividing (5.1) by $\xi_n^{b_1 + \dots + b_{n-1}}$ gives

$$(\xi_1/\xi_n)^{b_1} \dots (\xi_{n-1}/\xi_n)^{b_{n-1}} = \xi_n^{b_n - (b_1 + \dots + b_{n-1})}, \tag{5.2}$$

which we can regard as an “approximate relation” between $\xi_1/\xi_n, \dots, \xi_{n-1}/\xi_n$. The key observation here is that the coordinates $x_1/x_n, \dots, x_{n-1}/x_n$ with $1/x_n$ also describe a plane. In fact the map $f(x_1, \dots, x_n) = (x_1/x_n, \dots, x_{n-1}/x_n, 1/x_n)$ defines an automorphism of \mathbb{G}_m^n taking planes to planes and preserving non-degeneracy. Together with all the affine coordinate permutations it generates a finite group Φ (isomorphic to the symmetric group on the $n+1$ projective coordinates in (1.1)).

To make the above concept of approximate relation precise, we say for $\delta \geq 0$ that $P = (\xi_1, \dots, \xi_n)$ in $\mathbb{G}_m^n(\overline{\mathbb{Q}})$ is δ -dependent if there are a_1, \dots, a_n in \mathbb{Z} , not all zero, with

$$h(\xi_1^{a_1} \dots \xi_n^{a_n}) \leq \delta Ah(P) \quad (5.3)$$

for $A = |a_1| + \dots + |a_n|$. This property is the same as multiplicative dependence when $\delta = 0$, but in view of our choice of height function, it holds for any P as soon as $\delta \geq 1$.

To express the fact that (5.2) involves fewer coordinates than (5.1) we say that P is (δ, m) -dependent if at most m among a_1, \dots, a_n in (5.3) are non-zero. Here $1 \leq m \leq n$.

Lemma 5.1. *For $n \geq 3$ let \mathcal{P} be non-degenerate and defined over $\overline{\mathbb{Q}}$, and let $\epsilon > 0$. Let $\delta \geq 0$ and m satisfy*

$$\delta + 12n^2\epsilon \leq \frac{1}{2}, \quad 1 \leq m \leq n, \quad (5.4)$$

and suppose P in $\mathcal{P}(\overline{\mathbb{Q}}) \cap \mathcal{H}_{n-2}$ is (δ, m) -dependent not in $\mathcal{J}_{\mathcal{P}}(\epsilon)$. Then either

(a) *we have*

$$h(P) \leq \frac{2C_{\mathcal{P}}(\epsilon)}{\min\{1, 4\delta + 24n^2\epsilon\}}$$

or

(b) *if $m \geq 2$ there is f in Φ such that $P^* = f(P)$ is $(\delta^*, m-1)$ -dependent on $f(\mathcal{P})$, where*

$$\delta^* = 16\delta + 96n^2\epsilon.$$

Proof. Say $\theta = \xi_1^{a_1} \dots \xi_n^{a_n}$ in (5.3) satisfies

$$h(\theta) \leq \delta Ah(P) \quad (5.5)$$

for $A = |a_1| + \dots + |a_n| \geq 1$. If $m < n$ then after a permutation we can assume that $a_{m+1} = \dots = a_n = 0$. If the remaining a_1, \dots, a_m are all the same sign, we can assume them all non-negative. In this case we reach (a) as follows. By Lemma 4.1

$$h(\theta) \geq Ah(P) - A \left(12n^2\epsilon h(P) + C_{\mathcal{P}}(\epsilon) \right).$$

Comparing with (5.5) and dividing by A gives

$$h(P)(1 - \delta - 12n^2\epsilon) \leq C_{\mathcal{P}}(\epsilon),$$

from which (a) is clear in view of (5.4).

If a_1, \dots, a_m are not all the same sign (that is, some $a_i > 0$ and some $a_j < 0$), then $m \geq 2$ and after another permutation we can assume that

$$b_1 = a_1 \geq 0, \dots, b_r = a_r \geq 0, b_{r+1} = -a_{r+1} \geq 0, \dots, b_m = -a_m \geq 0$$

for some r with $1 \leq r \leq m-1$ and b_1, \dots, b_r not all zero. We could similarly assume, also after changing all the signs if necessary, that $\min_{1 \leq i \leq m} |a_i| = |a_m| = -a_m$. This implies

$$b_1 + \dots + b_{m-1} \geq \frac{m-1}{m}(b_1 + \dots + b_m). \quad (5.6)$$

Now

$$\xi_1^{b_1} \dots \xi_r^{b_r} = \theta \xi_{r+1}^{b_{r+1}} \dots \xi_m^{b_m} = \eta, \quad (5.7)$$

say. Here Lemma 4.1 gives

$$\begin{aligned} |h(\eta) - Bh(P)| &\leq B \left(12n^2 \epsilon h(P) + C_{\mathcal{P}}(\epsilon) \right) \\ |h(\eta) - \tilde{B}h(P)| &\leq h(\theta) + \tilde{B} \left(12n^2 \epsilon h(P) + C_{\mathcal{P}}(\epsilon) \right) \end{aligned}$$

for

$$B = b_1 + \dots + b_r, \quad \tilde{B} = b_{r+1} + \dots + b_m.$$

Eliminating $h(\eta)$, we find

$$|B - \tilde{B}|h(P) \leq h(\theta) + (B + \tilde{B}) \left(12n^2 \epsilon h(P) + C_{\mathcal{P}}(\epsilon) \right),$$

which expresses the notion that B, \tilde{B} are asymptotic. Here

$$h(\theta) \leq \delta Ah(P) = \delta(B + \tilde{B})h(P)$$

so we get

$$|B - \tilde{B}|h(P) \leq A \left(\delta h(P) + 12n^2 \epsilon h(P) + C_{\mathcal{P}}(\epsilon) \right). \quad (5.8)$$

Next rewrite (5.7) as

$$(\xi_1/\xi_m)^{b_1} \dots (\xi_r/\xi_m)^{b_r} = \theta (\xi_{r+1}/\xi_m)^{b_{r+1}} \dots (\xi_{m-1}/\xi_m)^{b_{m-1}} \xi_m^{-(B-\tilde{B})}, \quad (5.9)$$

or

$$\xi_1^{*b_1} \dots \xi_r^{*b_r} \xi_{r+1}^{*-b_{r+1}} \dots \xi_{m-1}^{*-b_{m-1}} = \theta^*, \quad (5.10)$$

with

$$\xi_i^* = \xi_i/\xi_m \quad (i = 1, \dots, m-1) \quad (5.11)$$

and

$$\theta^* = \theta \xi_m^{-(B-\tilde{B})}. \quad (5.12)$$

Now (5.10) suggests the $(\delta^*, m-1)$ -dependence of

$$P^* = (\xi_1/\xi_m, \dots, 1/\xi_m, \dots, \xi_n/\xi_m) = f(P)$$

for f in Φ and some δ^* . Certainly the exponents in (5.10) are not all zero. And in view of (5.8) and (5.12)

$$h(\theta^*) \leq h(\theta) + A \left(\delta h(P) + 12n^2 \epsilon h(P) + C_{\mathcal{P}}(\epsilon) \right),$$

which by (5.5) gives

$$h(\theta^*) \leq A \left(2\delta h(P) + 12n^2 \epsilon h(P) + C_{\mathcal{P}}(\epsilon) \right). \quad (5.13)$$

So δ^* looks like $2\delta + 12n^2 \epsilon$; but we have to take into account the new exponents in (5.10) as well as the extra $C_{\mathcal{P}}(\epsilon)$ in (5.13).

Well, the exponents in (5.10) satisfy

$$A^* = b_1 + \cdots + b_r + b_{r+1} + \cdots + b_{m-1} \geq \frac{1}{2}A$$

by (5.6). So (5.13) gives

$$h(\theta^*) \leq A^* \left(4\delta h(P) + 24n^2 \epsilon h(P) + 2C_{\mathcal{P}}(\epsilon) \right). \quad (5.14)$$

Next $P = f^{-1}(P^*)$ and it follows easily that

$$h(P) \leq 2h(P^*). \quad (5.15)$$

So (5.14) gives in turn

$$h(\theta^*) \leq A^* \left(8\delta h(P^*) + 48n^2 \epsilon h(P^*) + 2C_{\mathcal{P}}(\epsilon) \right).$$

This implies

$$h(\theta^*) \leq A^* \left(16\delta h(P^*) + 96n^2 \epsilon h(P^*) \right) \quad (5.16)$$

provided

$$8\delta h(P^*) + 48n^2 \epsilon h(P^*) \geq 2C_{\mathcal{P}}(\epsilon). \quad (5.17)$$

If (5.17) fails, then using (5.15) we get

$$h(P) < \frac{2C_{\mathcal{P}}(\epsilon)}{4\delta + 24n^2 \epsilon}$$

leading back to (a). So we can assume (5.17); and so (5.16) does follow.

Therefore (5.10) does indeed say that $f(P)$ is $(\delta^*, m-1)$ -dependent on $f(\mathcal{P})$, with $\delta^* = 16\delta + 96n^2 \epsilon$ as required in (b). This proves the present lemma. \square

6. Proof of Theorem 1.1

This is almost immediate after repeated applications of Lemma 5.1, at least for non-degenerate \mathcal{P} . We just have to dispose of exceptional cosets by means of the following simple observation.

Lemma 6.1. *Let \mathcal{S} be an irreducible surface in \mathbb{G}_m^n defined over $\overline{\mathbb{Q}}$, and let J be a proper coset defined over $\overline{\mathbb{Q}}$. Then $\mathcal{S}^{oa} \cap J \cap \mathcal{H}_{n-2}$ is a set of bounded height.*

Proof. If \mathcal{S} lies in J then according to (i^o) of the definition \mathcal{S}^{oa} is empty so there is nothing to do. Otherwise $\mathcal{S} \cap J$ is contained in a finite union of irreducible curves \mathcal{C} defined over $\overline{\mathbb{Q}}$. So

$$\mathcal{S}^{oa} \cap J \cap \mathcal{H}_{n-2} = (\mathcal{S}^{oa} \cap \mathcal{S}) \cap J \cap \mathcal{H}_{n-2} = \mathcal{S}^{oa} \cap (\mathcal{S} \cap J) \cap \mathcal{H}_{n-2}$$

is contained in the finite union of

$$\mathcal{S}^{oa} \cap \mathcal{C} \cap \mathcal{H}_{n-2}. \quad (6.1)$$

If one of the $\mathcal{C} \cap \mathcal{H}_{n-2}$ is a set of bounded height, then its contribution to (6.1) is as we want in the lemma. Otherwise $\mathcal{C} \cap \mathcal{H}_{n-2}$ is a set of unbounded height, and then Theorem A above implies that \mathcal{C} lies in a coset of dimension at most $n-1$. After the usual automorphism we can assume that $x_n = \xi_n$ is constant on \mathcal{C} . Let $\pi(\mathcal{C})$ in \mathbb{G}_m^{n-1} be the projection to the first $n-1$ coordinates. Then $\pi(\mathcal{C}) \cap \mathcal{H}_{n-2} = \pi(\mathcal{C}) \cap \mathcal{H}_{n-2}^{(n-1)}$ is also a set of unbounded height, so $\pi(\mathcal{C})$ lies in a coset of dimension $n-2$. We can assume $x_{n-1} = \xi_{n-1}$ is constant on $\pi(\mathcal{C})$. Now $x_{n-1} = \xi_{n-1}$, $x_n = \xi_n$ on \mathcal{C} show that \mathcal{C} is anomalous according to (ii^o) of the definition, so $\mathcal{S}^{oa} \cap \mathcal{C}$ is empty and once again (6.1) is fine. This proves the lemma. \square

Now to prove Theorem 1.1 we can assume $n \geq 3$, because torsion points have zero height. Pick δ and ϵ sufficiently small as functions of n . Take P in $\mathcal{P} \cap \mathcal{H}_{n-2}$. Certainly P is $(0, n)$ -dependent. So if P does not lie in any of the exceptional cosets of $\mathcal{J}_{\mathcal{P}}(\epsilon)$, we can apply Lemma 5.1. The case (a) gives at once the boundedness of $h(P)$. The case (b) gives $P^* = f(P)$ which is $(\delta^*, n-1)$ -dependent on $\mathcal{P}^* = f(\mathcal{P})$, of course still in \mathcal{H}_{n-2} . If P^* does not lie in any of the exceptional cosets of $\mathcal{J}_{\mathcal{P}^*}(\epsilon)$ corresponding to \mathcal{P}^* , we try to apply Lemma 5.1 again. If δ and ϵ were sufficiently small, then δ^* and ϵ will continue to satisfy (5.4). Case (a) is fine; case (b) leads to $(\delta^{**}, n-2)$ -dependence. And so on. After finitely many steps we must reach (a) and so the required boundedness of the height. If δ and ϵ were sufficiently small then $\delta, \delta^*, \delta^{**}, \dots$ and ϵ satisfy (5.4) at each step. We can always ignore the excluded cosets thanks to Lemma 6.1.

Finally what happens if \mathcal{P} is degenerate?

Lemma 6.2. *If \mathcal{P} is degenerate and defined over \mathbb{C} then \mathcal{P}^{oa} is empty.*

Proof. Applying an automorphism of Φ we can assume that x_{n-1}, x_n are dependent on \mathcal{P} . This means that one of the defining equations (1.1) can be taken as $\alpha + \beta x_{n-1} + \gamma x_n = 0$, with say $\beta \neq 0$. Now an arbitrary $P = (\xi_1, \dots, \xi_n)$ in \mathcal{P} can be removed using (i^o) or (ii^o) in the definition of \mathcal{P}^{oa} as follows. If the hyperplane coset J defined by $x_n = \xi_n$ contains \mathcal{P} , then it is removed (with P) according to (i^o) . Otherwise $\mathcal{P} \cap J$ is a line \mathcal{L} . The above equation now shows that $x_{n-1} = \xi_{n-1}$ on \mathcal{L} . So \mathcal{L} is anomalous, and it is removed (also with P) according to (ii^o) . This proves the lemma. \square

The proof of Theorem 1.1 is thereby complete.

7. Height lower bounds

We start by quoting a recent result of Amoroso and David, in which the choice of height function on \mathbb{G}_m^r is no longer so unimportant: critical is the property $h(\xi^a) = |a|h(\xi)$ for the individual coordinates.

Theorem C. *Given $r \geq 1$ there are positive constants $c(r)$ and $\kappa(r)$, depending only on r , with the following property. Let K be a cyclotomic extension of degree d over \mathbb{Q} and suppose P in $\mathbb{G}_m^r(\overline{\mathbb{Q}})$ has multiplicatively independent coordinates. Then*

$$h(P) \geq \frac{c(r)}{\omega (\log(3d\omega))^{\kappa(r)}}, \quad (7.1)$$

where $\omega = \omega_K(P)$ is the smallest degree of any non-zero polynomial in $K[t_1, \dots, t_r]$ that vanishes at P .

Proof. This is an immediate consequence of [2, Théorème 1.5, page 329], using the inequality

$$rh(P) = r \max\{h(\xi_1), \dots, h(\xi_r)\} \geq h(\xi_1) + \dots + h(\xi_r) \geq h(\xi_1, \dots, \xi_r)$$

for $P = (\xi_1, \dots, \xi_r)$ and the height used there. Note that the “obstruction subgroup” B there (which must be proper) does not exist here, in view of the multiplicative independence of the coordinates. \square

Here is a consequence for heights of numbers.

Lemma 7.1. *Given $r \geq 1$ there are positive constants $\tilde{c}(r)$ and $\kappa(r)$ with the following property. Let K be a cyclotomic extension of degree d over \mathbb{Q} and let η_1, \dots, η_r be multiplicatively independent non-zero algebraic numbers with $[K(\eta_1, \dots, \eta_r) : K] = \tilde{d}$. Then*

$$h(\eta_1) \cdots h(\eta_r) \geq \frac{\tilde{c}(r)}{\tilde{d}(\log(3d\tilde{d}))^{r\kappa(r)}}.$$

Proof. This is a fairly routine argument. As $[K(\eta_i) : \mathbb{Q}] \leq d\tilde{d}$ ($i = 1, \dots, r$) we can use crude lower bounds to find a positive integer

$$k \leq C(d\tilde{d})^2 \quad (7.2)$$

with C absolute, such that

$$h_i = h(\eta_i) \geq \frac{1}{k} \quad (i = 1, \dots, r).$$

Write

$$k_i = [kh_i] \geq 1 \quad (i = 1, \dots, r)$$

for the greatest integer parts, and define non-zero $\xi_i = \eta_i^{1/k_i}$ ($i = 1, \dots, r$) in $\overline{\mathbb{Q}}$. These remain independent and so Theorem C gives (7.1) for $P = (\xi_1, \dots, \xi_r)$. On the other hand

$$h(P) = \max \left\{ \frac{h_1}{k_1}, \dots, \frac{h_r}{k_r} \right\} \leq \frac{2}{k}.$$

We deduce

$$k \leq c_1 \omega (\log(3d\omega))^\kappa, \quad (7.3)$$

for $c_1 = c_1(r)$.

Now the usual linear algebra shows that $\omega \leq r[K(\xi_1, \dots, \xi_r) : K]^{1/r}$, and clearly

$$[K(\xi_1, \dots, \xi_r) : K] \leq \tilde{d}k_1 \cdots k_r \leq \tilde{d}k^r h_1 \cdots h_r.$$

Thus $\omega \leq rk(\tilde{d}h_1 \cdots h_r)^{1/r}$. We use this bound for the first ω in (7.3). As for the second ω , we can assume $h_1 \cdots h_r \leq \frac{1}{\tilde{d}}$ otherwise the lemma is trivially true. So we can use $\omega \leq rk$ in the second ω . We obtain

$$h_1 \cdots h_r \geq \frac{1}{c_2 \tilde{d} (\log(3drk))^{r\kappa(r)}}$$

for $c_2 = c_2(r)$ after cancelling k . The present lemma follows on recalling (7.2). \square

8. Proof of Theorem 1.2

This will be an immediate consequence of the following result, which generalizes [6, Lemma 1, page 2249]. As in [7], for any variety \mathcal{X} of dimension $m \geq 1$ in \mathbb{G}_m^n , we denote by \mathcal{X}^{ta} what remains of \mathcal{X} after removing, for each $s = 1, \dots, m$, all irreducible varieties \mathcal{Y} of dimension s contained in \mathcal{X} that lie in any torsion coset of dimension at most $n - 1 - m + s$. Just as for surfaces, we do not know if \mathcal{X}^{ta} is open.

Lemma 8.1. *Let \mathcal{X} be a variety in \mathbb{G}_m^n of dimension $m \leq n - 1$ defined over $\overline{\mathbb{Q}}$. Then for any $B \geq 0$ there are at most finitely many points P in $\mathcal{X}^{ta}(\overline{\mathbb{Q}}) \cap \mathcal{H}_{n-m-1}$ with $h(P) \leq B$.*

Proof. Let $P = (\xi_1, \dots, \xi_n)$ be in $\mathcal{X}^{ta}(\overline{\mathbb{Q}}) \cap \mathcal{H}_{n-m-1}$ with $h(P) \leq B$, and let r be the multiplicative rank of its coordinates. Then $r \leq n - m - 1$. Assume for the moment that $r \geq 1$. Applying [4, Lemma 2, page 1130] to the group generated by the coordinates, we obtain η_1, \dots, η_r in the field $\mathbb{Q}(P) = \mathbb{Q}(\xi_1, \dots, \xi_n)$ and multiplicatively independent, together with roots of unity ζ_1, \dots, ζ_n in $\mathbb{Q}(P)$, such that

$$\xi_i = \zeta_i \eta_1^{a_{i1}} \cdots \eta_r^{a_{ir}} \quad (i = 1, \dots, n) \quad (8.1)$$

for integers a_{ij} ; further

$$|a_{ij}|h(\eta_j) \leq c_1 h(P) \leq c_1 B \quad (i = 1, \dots, n; j = 1, \dots, r)$$

where c_1 (and subsequently c_2, \dots) are positive constants depending only on \mathcal{X} . So for the vectors $\mathbf{v}_j = (a_{1j}, \dots, a_{nj})$ with Euclidean lengths $|\mathbf{v}_j|$ we deduce

$$|\mathbf{v}_j| h(\eta_j) \leq c_2 B \quad (j = 1, \dots, r). \quad (8.2)$$

Write also $\mathbb{Q}(\zeta_1, \dots, \zeta_n) = \mathbb{Q}(\zeta)$ for a primitive N th root of unity ζ in $\mathbb{Q}(P)$, and $\zeta_i = \zeta^{l_i}$ for integers l_i with $0 \leq l_i < N$ ($i = 1, \dots, n$). Consider the $r + 1$ linear forms

$$-NX_0 + \sum_{i=1}^n l_i X_i, \quad \sum_{i=1}^n a_{ij} X_i \quad (j = 1, \dots, r) \quad (8.3)$$

in the $n + 1$ variables X_0, X_1, \dots, X_n . From (8.1) and the definition of r , we see that these forms are linearly independent. The height of the associated matrix is at most $c_3 N \Pi$ for $\Pi = |\mathbf{v}_1| \cdots |\mathbf{v}_r|$. So by Siegel's lemma (see for example [8, Theorem 9, page 27]) we can make these forms vanish at independent points $\mathbf{b}_1, \dots, \mathbf{b}_{n-r}$ in \mathbb{Z}^{n+1} whose lengths satisfy $|\mathbf{b}_1| \cdots |\mathbf{b}_{n-r}| \leq c_4 N \Pi$. With the ordering $|\mathbf{b}_1| \leq \dots \leq |\mathbf{b}_{n-r}|$, we deduce for the first $m \leq n - r - 1$ points

$$|\mathbf{b}_1| \cdots |\mathbf{b}_m| \leq (c_4 N \Pi)^{\frac{m}{n-r}} \leq (c_4 N \Pi)^{\frac{m}{m+1}}. \quad (8.4)$$

Writing $\mathbf{b}_k = (b_{0k}, b_{1k}, \dots, b_{nk})$ we get from (8.1) the relations

$$\xi_1^{b_{1k}} \cdots \xi_n^{b_{nk}} = 1 \quad (k = 1, \dots, m).$$

These say that $P = (\xi_1, \dots, \xi_n)$ lies in a certain algebraic subgroup H of \mathbb{G}_m^n . By (8.3) also (b_{1k}, \dots, b_{nk}) ($k = 1, \dots, m$) are independent, and so the dimension of H is at most $n - m$.

We now use standard degree theory in \mathbb{G}_m^n inside projective \mathbb{P}_n . It is well-known that (8.4) implies that the degree of H is at most $c_5 (N \Pi)^{\frac{m}{m+1}}$.

Consider the intersection $\mathcal{X} \cap H$. The component \mathcal{Y} passing through P cannot have dimension $s \geq 1$ otherwise it would have been removed to make \mathcal{X}^{ta} . Thus P is an isolated component of $\mathcal{X} \cap H$. By Bézout's theorem the number of such P does not exceed the product of the degrees of \mathcal{X} and H ; that is, at most $c_6 (N \Pi)^{\frac{m}{m+1}}$.

On the other hand this intersection contains P and all its conjugates over the field of definition of \mathcal{X} . Thus $D = [\mathbb{Q}(P) : \mathbb{Q}]$ satisfies

$$D \leq c_7(N\Pi)^{\frac{m}{m+1}}. \quad (8.5)$$

On yet another hand, η_1, \dots, η_r are multiplicatively independent in $\mathbb{Q}(P)$ containing $K = \mathbb{Q}(\zeta)$, and

$$\begin{aligned} \tilde{d} = [K(\eta_1, \dots, \eta_r) : K] &= \frac{[K(\eta_1, \dots, \eta_r) : \mathbb{Q}]}{\phi(N)} \\ &\leq \frac{D}{\phi(N)} \leq \frac{c_8 D(\log \log 3N)}{N} \end{aligned}$$

with Euler's function. And $d = [\mathbb{Q}(\zeta) : \mathbb{Q}] = \phi(N)$, so Lemma 7.1 implies

$$h(\eta_1) \cdots h(\eta_r) \geq \frac{N}{c_9 D(\log \log 3N)(\log 3D)^{r\kappa}} \geq \frac{N}{c_{10} D(\log 3D)^\lambda}$$

for κ and λ depending only on n .

From (8.2) we deduce that $N\Pi \leq c_{11} D(\log 3D)^\lambda B^r$. Finally (8.5) gives $D \leq c_{12} (D(\log 3D)^\lambda B^r)^{\frac{m}{m+1}}$, which implies that the degree D of P is bounded. As $h(P) \leq B$, its height is also bounded, and the present lemma follows by a standard application of Northcott's theorem, at least when $r \geq 1$.

When $r = 0$, the above argument remains valid after trivial adjustments. We omit η_1, \dots, η_r in (8.1), ignore (8.2), use a single form in (8.3), write $\Pi = 1$ and finally note that now $D = \phi(N)$ in (8.5). Of course P is now a torsion point on \mathcal{X} . This means that we could also appeal to more classical results (see for example the discussion after the statement of [7, Theorem 2]). \square

It is now easy to deduce Theorem 1.2. By Theorem 1.1 the points of $\mathcal{P}^{oa} \cap \mathcal{H}_{n-2}$ have height at most B depending only on \mathcal{P} . So this also holds for the points of $\mathcal{P}^{oa} \cap \mathcal{H}_{n-3}$. These in turn lie in $\mathcal{P}^{ta} \cap \mathcal{H}_{n-3}$ and then application of Lemma 8.1 with $m = 2$ completes the proof.

9. Proof of Theorem 1.3

Most of the arguments of this section are simplified explicit versions of similar arguments in [7]. For the reader's convenience we give all the details. We emphasize again that the plane \mathcal{P} is defined over \mathbb{C} but not necessarily over $\overline{\mathbb{Q}}$.

Lemma 9.1. *For $n \geq 3$ let \mathcal{P} be non-degenerate. Then the functions $\frac{1}{x_i x_j}$ ($1 \leq i < j \leq n$) are linearly independent over \mathbb{C} on \mathcal{P} .*

Proof. Suppose that some linear combination $y = \sum_{1 \leq i < j \leq n} \frac{\beta_{ij}}{x_i x_j}$ vanishes on \mathcal{P} . As in (3.1), we can express the x_k as linear combinations of x_1, x_2 and $x_0 = 1$, with non-zero coefficients γ_k ($k = 3, \dots, n$) of x_0 . Substituting into $x_1 x_2 \cdots x_n y$ we obtain a polynomial in x_1 and x_2 , with constant term $\beta_{12} \gamma_3 \cdots \gamma_n$, that also vanishes on \mathcal{P} . But x_1 and x_2 are algebraically independent on \mathcal{P} , which forces $\beta_{12} = 0$. In a similar way we see that $\beta_{ij} = 0$ for all $i < j$, and this proves the present lemma. \square

Recall that the degree of a variety in \mathbb{G}_m^n means the degree of its closure in \mathbb{P}_n . This is invariant under group translation.

Lemma 9.2. *For $n \geq 3$ let \mathcal{P} be non-degenerate.*

- (a) *Every anomalous curve \mathcal{C} of \mathcal{P} has degree at most $n - 2$.*
- (b) *Given a positive integer Δ , there are at most finitely many anomalous curves of \mathcal{P} lying in any translate of any algebraic subgroup of dimension $n - 2$ and degree at most Δ .*

Proof. On \mathcal{C} there are relations

$$\prod_{i=1}^n x_i^{a_i} = \alpha, \quad \prod_{j=1}^n x_j^{b_j} = \beta \quad (9.1)$$

with $(a_1, \dots, a_n), (b_1, \dots, b_n)$ in \mathbb{Z}^n linearly independent. Choose any parameter t on \mathcal{C} and differentiate the first in (9.1) with respect to t to give $\sum_{i=1}^n \frac{a_i \dot{x}_i}{x_i} = 0$. As the differentials dx_1, \dots, dx_n on \mathcal{P} span a vector space of dimension 2 we can write

$$\dot{x}_i = \lambda_i u + \mu_i v \quad (i = 1, \dots, n) \quad (9.2)$$

for basis elements u, v (for example $u = \dot{x}_1, v = \dot{x}_2$). Thus

$$u \sum_{i=1}^n \frac{\lambda_i a_i}{x_i} + v \sum_{i=1}^n \frac{\mu_i a_i}{x_i} = 0.$$

Similarly the second in (9.1) gives

$$u \sum_{j=1}^n \frac{\lambda_j b_j}{x_j} + v \sum_{j=1}^n \frac{\mu_j b_j}{x_j} = 0.$$

Elimination of u and v now yields an equation $\sum_{1 \leq i < j \leq n} \frac{\beta_{ij}}{x_i x_j} = 0$ holding on the curve \mathcal{C} . Here

$$\begin{aligned} \beta_{ij} &= \lambda_i a_i \mu_j b_j + \lambda_j a_j \mu_i b_i - \mu_i a_i \lambda_j b_j - \mu_j a_j \lambda_i b_i \\ &= (\lambda_i \mu_j - \lambda_j \mu_i)(a_i b_j - a_j b_i). \end{aligned}$$

Now $\lambda_1\mu_2 - \lambda_2\mu_1 \neq 0$ otherwise \dot{x}_1 and \dot{x}_2 would be dependent in (9.2). Similarly $\lambda_i\mu_j - \lambda_j\mu_i \neq 0$ for any $i < j$. As (a_1, \dots, a_n) and (b_1, \dots, b_n) are independent it follows that the β_{ij} are not all zero. So by Lemma 9.1 the equation $\sum_{1 \leq i < j \leq n} \beta_{ij}(x_i x_j)^{-1} = 0$ defines a proper subvariety of \mathcal{P} , of which \mathcal{C} is a component. Visibly this equation has degree at most $n - 2$, and this proves part (a) of the present lemma.

Part (b) follows from the same calculation, because the exponents in (9.1) can easily be estimated in terms of the degree of the coset defined by (9.1), and so there are at most finitely many possibilities for the β_{ij} . In fact such estimates are not needed; one can choose the exponents such that the subdeterminants $d_{ij} = a_i b_j - a_j b_i$ are coprime, and then this degree has the same order of magnitude as $\max_{1 \leq i < j \leq n} |d_{ij}|$. \square

Lemma 9.3. *For $n \geq 3$ let \mathcal{P} be non-degenerate. Then there is a constant $c(n)$, depending only on n , such that every anomalous curve on \mathcal{P} lies in a translate of some algebraic subgroup H of dimension $n - 2$ with the degree of H not exceeding $c(n)$.*

Proof. Let \mathcal{C} be anomalous, and choose x among x_1, \dots, x_n not constant on \mathcal{C} . Then $F = \mathbb{C}(x)$ is a “field with a proper set of absolute values satisfying a product formula” in the sense of [5, page 452]. We regard the function field $\mathbb{C}(\mathcal{C})$ as lying in the algebraic closure \bar{F} , so that there is an absolute height function on these fields, which we denote also by h . We will apply the “multiplicative dependence estimate” [5, Lemma 2.2, page 457] to the x_1, \dots, x_n . As \mathcal{C} is anomalous, the rank r of the subgroup generated by these coordinates over the “zero height group” Z , in this case the group of non-zero complex numbers, satisfies $1 \leq r \leq n - 2$.

Now for any non-constant y in $\mathbb{C}(\mathcal{C})$ we have

$$h(y) = \frac{[\mathbb{C}(x, y) : \mathbb{C}(y)]}{[\mathbb{C}(x, y) : \mathbb{C}(x)]} \quad (9.3)$$

(see for example [5, (4.4) page 461]). Here $[\mathbb{C}(x, y) : \mathbb{C}(x)] \leq [\mathbb{C}(\mathcal{C}) : \mathbb{C}(x)]$ is at most the degree of \mathcal{C} , which by Lemma 9.2(a) is at most $n - 2$. Thus

$$h(y) \geq \frac{1}{n - 2} \quad (9.4)$$

for any non-constant y .

Similarly (9.3) gives $h(x_i) \leq n - 2$ ($i = 1, \dots, n$). So [5, Lemma 2.2] supplies for any positive integer T a non-zero (a_1, \dots, a_n) in \mathbb{Z}^n with $|a_i| \leq T$ ($i = 1, \dots, n$) and $h(y) \leq c'(n)T^{-\frac{2}{n-2}}(n - 2)$ for $y = x_1^{a_1} \cdots x_n^{a_n}$, where $c'(n)$ depends only on n . Choosing T minimally to contradict (9.4), we deduce that the resulting $y = \alpha$ is constant on \mathcal{C} .

This gives one of the relations (9.1), with exponents a_1, \dots, a_n bounded in terms only of n . The second can be obtained similarly; for example we can assume $a_n \neq 0$, and then [5, Lemma 2.2] applied to just x_1, \dots, x_{n-1} gives bounded b_1, \dots, b_{n-1} with $x_1^{b_1} \cdots x_{n-1}^{b_{n-1}} = \beta$. Now $(a_1, \dots, a_{n-1}, a_n), (b_1, \dots, b_{n-1}, 0)$ are independent as in (9.1). \square

In the above proof we used a multiplicative dependence estimate for a function field of transcendence degree 1. The next rather simple observation can be regarded as a very special case of a generalization to arbitrary transcendence degree. But it is also very sharp; for example the resulting exponents are bounded in absolute value by 1.

Lemma 9.4. *For $n \geq 2$ let L_1, \dots, L_n be complex non-constant linear polynomials in several variables. Then they are multiplicatively dependent modulo constants if and only if there are i, j with $1 \leq i < j \leq n$ such that L_i, L_j are linearly dependent.*

Proof. For any field K a non-constant linear polynomial in the unique factorization domain $K[t_1, \dots, t_m]$ is prime, and two such primes are associate if and only if they are proportional. The result follows at once. Alternatively one can specialize the variables $t_1 = \sigma_1 t + \tau_1, \dots, t_m = \sigma_m t + \tau_m$ to involve a single variable t , when the result is obvious. \square

We can now finish the proof of Theorem 1.3. We claim that it suffices to verify for a non-degenerate \mathcal{P} that

- (i) \mathcal{P} is not contained in a proper coset, and
- (ii) the number of anomalous lines is at most $\frac{1}{8}(n-2)(n-1)n(n+1)$.

For if \mathcal{P}^{oa} is non-empty we deduce from Lemma 6.2 that \mathcal{P} is non-degenerate. Conversely suppose that \mathcal{P} is non-degenerate. Then (i) above would show that nothing in (i^o) of the definition of \mathcal{P}^{oa} is removed. As for a curve \mathcal{C} removed in (ii^o) of the definition, Lemma 9.3 implies that \mathcal{C} lies in a translate of some algebraic subgroup H of dimension at most $n-2$ with degree at most $\Delta = c(n)$. Then Lemma 9.2(b) shows that the number of possibilities for \mathcal{C} is bounded above by a function only of n . In particular \mathcal{P}^{oa} is non-empty. And Lemma 9.2(a) implies that \mathcal{C} has degree at most $n-2$, and (ii) above would then complete the proof of Theorem 1.3.

Regarding (i), if this were false then the linear polynomials

$$t_1, t_2, \beta_k t_1 + \alpha_k t_2 + \gamma_k \quad (k = 3, \dots, n)$$

in (3.1) would be multiplicatively dependent modulo constants. Because $\beta_k, \alpha_k, \gamma_k$ are all non-zero, it is easy to see from Lemma 9.4 that this implies $n \geq 4$ and the linear dependence of two of the $\beta_k t_1 + \alpha_k t_2 + \gamma_k$, say for $k = 3, 4$. But then x_3 and x_4 would be linearly dependent, contradicting their algebraic independence. This gives (i).

Regarding (ii), assume for the moment $n \geq 4$ and consider the space \mathcal{V} defined by equations

$$\frac{x_1}{x_2} = \lambda, \quad \frac{x_3}{x_4} = \mu.$$

It is readily checked that $\mathcal{V} \cap \mathcal{P}$ is empty or a single point unless $\gamma_3 = \mu\gamma_4$ and $\beta_3\lambda + \alpha_3 = \mu(\beta_4\lambda + \alpha_4)$. These determine μ as $\frac{\gamma_3}{\gamma_4}$ and then λ as $-\frac{\alpha_3\gamma_4 - \alpha_4\gamma_3}{\beta_3\gamma_4 - \beta_4\gamma_3}$; here the non-degeneracy of \mathcal{P} implies that both numerator and denominator of λ are non-zero. Thus in this case we get a line \mathcal{L}_{1234} , which is clearly anomalous. And associated with this are two other similarly defined lines \mathcal{L}_{1324} and \mathcal{L}_{1423} .

More generally each of the $\binom{n+1}{4}$ four-element subsets of $\{0, 1, \dots, n\}$ determines three anomalous lines \mathcal{L}_{ijkl} , giving at most $3\binom{n+1}{4} = \frac{1}{8}(n-2)(n-1)n(n+1)$.

In fact every anomalous line \mathcal{L} arises in this way. To see this, just parametrize \mathcal{L} by linear polynomials L_1, \dots, L_n in $\mathbb{C}[t]$. If they are all non-constant, then by Lemma 9.4 at least two are proportional, leading to say $\frac{x_1}{x_2} = \lambda$ on \mathcal{L} . But we can apply Lemma 9.4 also to the last $n-1$ coordinates, leading to a second such relation. This relation cannot involve x_2 , else some $\frac{x_k}{x_2}$ ($k \geq 3$) would be constant; impossible because $\frac{x_1}{x_2}$ and $\frac{x_k}{x_2}$ are coordinates on $f(\mathcal{P})$ for some f in Φ and so algebraically independent. Thus the second relation looks like $\frac{x_3}{x_4} = \mu$, giving \mathcal{L}_{1234} .

And if say L_1 is constant, then a similar argument leads to lines like \mathcal{L}_{0123} . So (ii) is proved for $n \geq 4$. The case $n = 3$ can now be left to the reader, and the case $n = 2$ is trivial.

This completes the proof of Theorem 1.3. We finish our paper with some remarks about what happens in small dimensions.

In dimension $n = 3$ it is also easy to check for a non-degenerate plane that Γ_1 has exactly $\frac{1}{8}(n-2)(n-1)n(n+1) = 3$ different lines, and that Γ is empty. We claim the same is true even in dimension $n = 4$; i.e. there are exactly 15 anomalous curves, all of which are lines.

To begin with, we note that an anomalous line like $\mathcal{L} = \mathcal{L}_{1234}$ cannot be the same as another line, say \mathcal{L}_{1324} , formed with the same four-element subset; for then $\frac{x_1}{x_2}$ and $\frac{x_3}{x_2} = \frac{x_1}{x_2} / \frac{x_1}{x_3}$, already algebraically independent, would be constant on \mathcal{L} . But also \mathcal{L} cannot be the same as a line, say \mathcal{L}_{0123} , formed with a different four-element subset, for then $\frac{x_1}{x_2}$ and $\frac{x_1}{x_0} = x_1$ so also x_2 would be constant on \mathcal{L} . Thus all 15 anomalous lines are different.

There is more than one way of seeing the non-existence of other anomalous curves, but each seems to be slightly troublesome. One can imitate the arguments of Sections 2, 3, 4, 5 and 6 for the function field analogue. Another proof due to Habegger, avoiding Lemmas 9.1, 9.2 and 9.3, uses valuations and the ultrametric inequality. Here is an alternative in the style of Lemma 9.4.

Let \mathcal{C} be an anomalous curve. By Lemma 9.2(a) it is defined by an equation of degree at most 2. So it is rational, and indeed it can be homogeneously parametrized by polynomials Q_0, Q_1, Q_2, Q_3, Q_4 in $\mathbb{C}[t]$ of degree at most 2, naturally without a common factor. If each of these is a square, then Lemma 9.4 on the non-constant

square roots leads quickly to the constancy of some $\frac{x_i}{x_j}$ ($i \neq j$) on \mathcal{C} . Therefore in this case \mathcal{C} is a line.

So we can assume that say Q_0 is not a square. If Q_0 is quadratic then by means of a fractional linear transformation we can move one of its zeros to 0 and the other to ∞ . By the homogeneity this amounts to the assumption $Q_0 = t$; and of course this can be achieved even more simply if Q_0 was not quadratic.

In general one would hope that none of Q_1, Q_2, Q_3, Q_4 are divisible by t or constant. If this is so, then we can argue as follows. A first non-trivial multiplicative relation involves say $x_1 = \frac{Q_1}{t}$. So a zero of x_1 must be balanced by a zero of say $x_2 = \frac{Q_2}{t}$. But there is also a second multiplicative relation involving only x_1, x_3, x_4 . If it involves x_1 then this zero of x_1 also occurs in say x_3 . Now Q_1, Q_2, Q_3 have a common factor and so must be linearly dependent. Thus x_1, x_2, x_3 are linearly dependent on \mathcal{C} . As \mathcal{P} is non-degenerate this too implies that \mathcal{C} is a line. If the second multiplicative relation does not involve x_1 , then consideration of the factor t shows that $\frac{x_3}{x_4}$ is constant on \mathcal{C} and so again we get a line.

This argument still works if none of Q_1, Q_2, Q_3, Q_4 are divisible by t but some are constant. For then we can suppose that only Q_4 is constant (else for example $\frac{x_3}{x_4}$ would be constant on \mathcal{C}), and that the first multiplicative relation still involves x_1 .

Finally what if some of Q_1, Q_2, Q_3, Q_4 are divisible by t ?

If only one is, say Q_1 , then Q_1 must be quadratic (else $\frac{x_1}{x_0}$ would be constant on \mathcal{C}) and the argument can be made to work provided Q_1 is not divisible by t^2 . For if no multiplicative relation involves x_1 , then we can argue with the two relations on the projection to the last three coordinates. But also if Q_1 is divisible by t^2 , then provided none of Q_2, Q_3, Q_4 are constant we can deduce in the same way that these are linearly dependent and so x_2, x_3, x_4 are linearly dependent on \mathcal{C} . However if say Q_2 is constant, then any non-trivial multiplicative relation among x_1, x_2, x_3, x_4 implies that Q_3, Q_4 have the same zeroes. So either they are proportional, in which case $\frac{x_3}{x_4}$ is constant on \mathcal{C} , or say Q_3 is linear. But then x_0, x_2, x_3 would be linearly dependent on \mathcal{C} .

If more than one of Q_1, Q_2, Q_3, Q_4 is divisible by t then, as no three can be, we may suppose that Q_1 and Q_2 , but no others, are divisible by t . If neither is divisible by t^2 , then a non-trivial multiplicative relation between x_1, x_2, x_3 cannot involve x_3 and so it forces $\frac{x_1}{x_2}$ to be constant. And if Q_1 but not Q_2 is divisible by t^2 , then looking at the multiplicative relations not involving x_3 and not involving x_4 shows that Q_2, Q_3, Q_4 have a common factor. Finally Q_1 and Q_2 cannot both be divisible by t^2 else $\frac{x_1}{x_2}$ is again constant.

Thus indeed in dimension $n = 4$ all the possible anomalous lines are different and there are no other anomalous curves.

In dimension $n = 5$ both of these can fail. For example each of the systems

$$\begin{aligned} x_3 &= x_1 + 2x_2 + 1, & x_4 &= x_1 + 3x_2 + 2, & x_5 &= x_1 + x_2 + 4, \\ x_3 &= x_1 + 2x_2 + 1, & x_4 &= x_1 + 3x_2 + 2, & x_5 &= x_1 + 4x_2 + 4 \end{aligned}$$

defines a plane. Each is non-degenerate because all the maximal subdeterminants of each of

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 2 & 1 \\ 1 & 3 & 2 \\ 1 & 1 & 4 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 2 & 1 \\ 1 & 3 & 2 \\ 1 & 4 & 4 \end{pmatrix}$$

are non-zero. On the first plane the anomalous lines \mathcal{L}_{1234} and \mathcal{L}_{0512} are the same, being defined parametrically by

$$(x_1, x_2, x_3, x_4, x_5) = (-t, t, t + 1, 2t + 2, 4)$$

(and so lies in a coset of dimension $2 = n - 3$). The second plane contains the curve \mathcal{C} parametrized by

$$(x_1, x_2, x_3, x_4, x_5) = (t^2, t, t^2 + 2t + 1, t^2 + 3t + 2, t^2 + 4t + 4).$$

This curve is clearly quadratic irreducible and it lies in the subgroup defined by

$$x_1 = x_2^2, \quad x_3 x_5 = x_4^2.$$

So it is anomalous.

From the point of view of enumerative geometry it may be interesting to take the study of anomalous curves on non-degenerate planes further. What is the correct upper bound for their degree? Is their number bounded polynomially in terms of the dimension n ? The use of [5, Lemma 2.2] leads only to bounds like n^{cn^2} or worse. The arguments of [11, Corollary 3.2, page 281] might yield generalizations of Lemma 9.4 with polynomial estimates for the exponents. Do there exist anomalous curves with arbitrarily large genus? And so on. And finally, as we said in Section 1, for “generic” planes it seems likely that there are exactly $\frac{1}{8}(n-2)(n-1)n(n+1)$ anomalous curves, all of which are lines.

References

- [1] F. AMOROSO and S. DAVID, *Le problème de Lehmer en dimension supérieure*, J. Reine Angew. Math. **513** (1999), 145–179.
- [2] F. AMOROSO and S. DAVID, *Distribution des points de petite hauteur dans les groupes multiplicatifs*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (5) **3** (2004), 325–348.
- [3] F. AMOROSO and U. ZANNIER, *A relative Dobrowolski lower bound over abelian extensions*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **29** (2000), 711–727.
- [4] E. BOMBIERI, D. MASSER and U. ZANNIER, *Intersecting a curve with algebraic subgroups of multiplicative groups*, Internat. Math. Res. Notices **20** (1999), 1119–1140.
- [5] E. BOMBIERI, D. MASSER and U. ZANNIER, *Finiteness results for multiplicatively dependent points on complex curves*, Michigan Math. J. **51** (2003), 451–466.

- [6] E. BOMBIERI, D. MASSER and U. ZANNIER, *Intersecting curves and algebraic subgroups: conjectures and more results*, Trans. Amer. Math. Soc. **358** (2006), 2247–2257.
- [7] E. BOMBIERI, D. MASSER and U. ZANNIER, *Anomalous subvarieties - structure theorems and applications*, Int. Math. Res. Not. IMRN **19** (2007), 33 pages.
- [8] E. BOMBIERI and J. VAALER, *On Siegel's Lemma*, Invent. Math. **73** (1983), 11–32.
- [9] E. BOMBIERI and U. ZANNIER, *Algebraic points on subvarieties of \mathbb{G}_m^n* , Internat. Math. Res. Notices **7** (1995), 333–347.
- [10] J. W. S. CASSELS, “An Introduction to Diophantine Approximation”, Cambridge Tracts in Mathematics and Mathematical Physics, Vol. 45, Cambridge, 1965.
- [11] T. LOHER and D. MASSER, *Uniformly counting points of bounded height*, Acta Arith. **111** (2004), 277–297.
- [12] R. PINK, *A common generalization of the conjectures of André-Oort, Manin-Mumford, and Mordell-Lang*, manuscript dated 17th April 2005.
- [13] A. SCHINZEL, “Polynomials with Special Regard to Reducibility”, Encyclopaedia of Mathematics and its Applications, Vol. 77, Cambridge, 2000.
- [14] U. ZANNIER, *Proof of Conjecture 1*, Appendix to [13], 517–539.
- [15] B. ZILBER, *Exponential sums equations and the Schanuel conjecture*, J. London Math. Soc. **65** (2002), 27–44.

School of Mathematics
 Institute for Advanced Study
 Princeton, New Jersey 08540, USA
 eb@math.ias.edu

Mathematisches Institut
 Universität Basel
 Rheinsprung 21
 4051 Basel, Switzerland
 David.Masser@unibas.ch

Scuola Normale Superiore
 Piazza dei Cavalieri 7
 56126 Pisa, Italy
 u.zannier@sns.it