Heights and multiplicative relations on algebraic varieties

Inauguraldissertation
zur
Erlangung der Würde eines Doktors der Philosophie

vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Philipp Habegger

aus
Trub BE

Basel, 2007

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von Prof. Dr. D.W. Masser und Prof. Dr. S. David.

Basel, den 26. Juni 2007

Prof. Dr. Hans-Peter Hauri
Dekan

Contents

Acknowledgements

I would like to thank my adviser Prof. David W. Masser for introducing me into the active field of research revolving around the theory of heights. His large experience, his many insightful comments, and his support have helped me greatly throughout my years as a PhD student. He also carefully read this manuscript and gave many invaluable suggestions for improvement.

Many thanks go also to Prof. Sinnou David who agreed to referee my thesis, who helped me organize my six month visit the Insitut de Mathématiques de Jussieu in Paris, and with whom I had many interesting discussions there. I enjoyed my stay in Paris very much, both from a mathematical and cultural point of view.

Furthermore, I would like to thank Prof. Francesco Amoroso, Prof. Michael Nakamaye, Prof. Patrice Philippon, Prof. Gaël Rémond, Prof. Evelina Viada, and Prof. Umberto Zannier for the fruitful conversations I had over the years. I am grateful for Prof. Zannier's comments on an early version of the chapter containing the Quasi-equivalence result.

During my thesis I received financial support from a Schweizer Nationalfonds grant. I obtained financial support from the Department of Mathematics in Basel directly before and after my SNF grant. While in Paris I received financial support from the Institut de Mathémathiques de Jussieu.

In the last few years I had the pleasure to attend several conferences among them the Diophantine Geometry Research Period at the Scuola Normale Superiore di Pisa, the Diophantine Approximation and Heights Program at the ESI in Vienna, Approximation diophantienne et nombres transcendants at the CIRM in Luminy, and the Workshop Diophantische Approximationen in Oberwolfach. I would like to thank the organizers of these conferences for inviting me and for giving me the possibility to present my work.

Thanks go also to my friends, Em, Giuliana, Jonas, Martin, Patrick, Primo, Ute from inside and outside the University for keeping my spirits high and of course for the endless discussions during coffee break. I would like to especially thank Brita for the wonderful time we spent together in Basel, Berlin, and Paris.

Finally, I thank my parents, they were always there for me in the inevitable ups and downs in life. I owe this thesis to my mother who always supported my decision in life and my father who sparked my interest in mathematics.

Introduction

Points on a subvariety X of a semi-abelian variety A that are contained in a subgroup, let the subgroup be of finite rank or algebraic, are subject to severe restrictions arithmetical nature.

Finiteness results for intersections of X with subgroups of finite rank have been studied by Faltings, Hindry, Laurent, McQuillan, Raynaud, Vojta and others. More recently several authors ([**CZ00**], [**BMZ99**], [**BMZ03**], [**BMZ06a**], [**BMZ06b**], [**BMZ04**], [**Via03**], [**RV03**], [**Rém05b**], [**Rém07**], [**Pin05b**], [**Zan00**], [**Zil02**], [**Mau06**]) have considered the intersection of X with $A^{[r]}$, the set of complex points in A contained in an algebraic subgroup of codimension greater or equal to r. If H is a fixed algebraic subgroup of A with codimension strictly less than $\dim X$, then a dimension counting argument shows that $X \cap H$ is either empty or contains a curve. As we are allowing H to vary with fixed codimension, the intersection $X \cap A^{[r]}$ may be quite large if $r < \dim X$. In this thesis we are only interested in the case $r \geq \dim X$.

If not stated otherwise we will also assume throughout the introduction that all varieties are defined over $\overline{\mathbf{Q}}$, the field of algebraic numbers. One can define a height function on the set of algebraic points of A. Throughout this thesis we work only in the algebraic torus $\mathbf{G}_m^n$ or an abelian variety. So we can take the Weil height or the Néron-Tate height associated to an ample line bundle.

We will pursue two types of questions. First, for which r does the set $X'(\overline{\mathbf{Q}}) \cap A^{[r]}$ have bounded height and how do these bounds depend on X? Second, for which r is the set $X''(\overline{\mathbf{Q}}) \cap A^{[r]}$ finite? Here X' and X'' are obtained from removing from X certain subvarieties in order to to eliminate trivial counterexamples. For example if X is a proper algebraic subgroup of $\mathbf{G}_m^n$ with positive dimension, then there is no hope for a boundedness of height or finiteness result for $U(\overline{\mathbf{Q}}) \cap (\mathbf{G}_m^n)^{[r]}$ if $r \leq \dim X$ and if U is Zariski open and dense in X. In this case X' and X'' are both empty.

The simplest non-trivial example seems to be the curve defined by $x + y = 1$ in $\mathbf{G}_m^2$. Here we can take X' and X'' to equal our curve. Algebraic subgroups of $\mathbf{G}_m^2$ can be described by at most two monomial relations $x^\alpha y^\beta = 1$ with integer exponents α and β. For subgroups of dimension 1, one non-trivial relation suffices. If (x, y) is contained in such a subgroup then x and y are called multiplicatively dependent. Hence the intersection of our curve with the union of all proper algebraic subgroups of $\mathbf{G}_m^2$ can be described by the solutions of

$$x^\alpha (1 - x)^\beta = 1. \tag{0.0.1}$$

This is an equation in three unknowns x, α, and β, so one should not expect finitely many solutions. Indeed, taking $x \neq 1$ a root of unity gives infinitely many solutions.

In [**CZ00**] Cohen and Zannier showed that if H denotes the absolute non-logarithmic Weil height then (0.0.1) implies the sharp inequality $\max\{H(x), H(1-x)\} \leq 2$. In chapter 2 we start off by giving an alternative proof of Cohen and Zannier's Theorem. We even show that the possibly larger height $H(x, 1-x)$ is at most 2. In their paper, Cohen and Zannier also proved that 2 is an isolated point in the range of $\max\{H(x), H(1-x)\}$. We make this result explicit in Theorem 2.2, working instead with $H(x, 1-x)$. The proof applies Smyth's Theorem on lower bounds for heights of non-reciprocal algebraic numbers and a Theorem of Mignotte.

As was already noticed in [**CZ00**], solutions of (0.0.1) are closely linked to roots of certain trinomials whose coefficients are roots of unity. In chapter 3 Theorem 3.2 we follow this avenue by factoring such trinomials over cyclotomic fields. Having essentially a minimal polynomial in our hands, we obtain a new proof for the boundedness of $H(x, 1-x)$ with x as in (0.0.1). More importantly, in Theorem 3.1 we show that not only is 2 isolated in the range of the height function, but also that $H(x, 1-x)$ converges to an absolute constant if $[\mathbf{Q}(x) : \mathbf{Q}]$ goes to infinity. The proof determines the value of this limit: it is the Mahler measure of the two-variable polynomial $X + Y - 1$. In a certain sense this Mahler measure is the height of the curve in our problem.

In Theorem 3.3 we prove a conjecture of Masser stated in [**Mas07**]: the number of solutions of (0.0.1) with $[\mathbf{Q}(x) : \mathbf{Q}] \leq D$ is asymptotically equal to $c_0 D^3$ with $c_0 = 2.06126\ldots$ as $D \to \infty$. The constant c_0 is defined properly in chapter 3 as a converging series. This counting result is a further application of Theorem 3.2.

In chapter 4 we generalize the method from chapter 2 to bound the height of multiplicatively dependent solutions of

$$x + y = \alpha. \tag{0.0.2}$$

Here α is now any non-zero algebraic number. In [**BMZ99**] Bombieri, Masser, and Zannier prove a more general result which also implies boundedness of height in this case. Their Proposition A leads to an explicit upper bound for the height; the bound is polynomial in $H(\alpha)$. We are mainly interested in upper bounds for $H(x, y)$ which have good dependency in $H(\alpha)$. The value $H(\alpha)$ can be regarded as the height of the defining equation (0.0.2). In Theorems 4.1 and 4.2 we get the bound $H(x, y) \leq 2H(\alpha) \min\{H(\alpha), 7 \log(3H(\alpha))\}$. By Theorem 4.3 the exponent of the logarithm cannot be less than 1. But in some special cases, e.g. if α is a rational integer, we improve the upper bound to $2H(\alpha)$, see Theorem 4.4. In this theorem we also show that if α is a rational integer then $2H(\alpha)$ is attained as a height if and only if α is a power of two. Thus if α is a power of two, then our bound is sharp. For such α and if also $\alpha \geq 2$ we prove in Theorem 4.5 that $2H(\alpha)$ is isolated in the range of the height.

Starting from chapter 6 we work in an algebraic torus of arbitrary dimension. Algebraic subgroups can still be described by a finite set of monomial equations. For example $(x_1, \ldots, x_n) \in \mathbf{G}_m^n(\mathbf{C})$ is contained in a proper algebraic subgroup if and only if the x_i

satisfy a non-trivial multiplicative relation. In [**BMZ99**] Bombieri, Masser, and Zannier proved that if X is an irreducible curve which is not contained in the translate of a proper algebraic subgroup, then points on X that lie in a proper algebraic subgroup have bounded height. Moreover, they showed that this statement is false if X is contained in the translate of a proper algebraic subgroup. The authors also showed that there are only finitely many points on X that lie in an algebraic subgroup of codimension at least 2. This finiteness result was generalized by the same authors in [**BMZ03**] to algebraic curves defined over the field of complex numbers. Hence for curves it makes sense to take $X' = X$ if X is not contained in the translate of a proper algebraic subgroup and $X' = \emptyset$ else wise. But X'' is more subtle: we take $X'' = X$ if X is not contained in a proper algebraic subgroup and $X'' = \emptyset$ else wise. The point in making this distinction is that in [**BMZ06a**] the authors conjectured that X'' contains only finitely many points in an algebraic subgroup of codimension at least 2. They proved this conjecture for $n \leq 5$. Recently, in [**Mau06**] Maurin gave a proof for all n.

Let $X \subset \mathbf{G}_m^n$ be an irreducible subvariety, not necessarily a curve. In the higher dimensional case we finally need a definition of X': we get X' by removing from X all positive dimension subvarieties that show up in an improper component of the intersection of X with the translate of an algebraic subgroup. The definition of X'' is similar but we require the translates of algebraic subgroups to be algebraic subgroups. In [**BMZ06b**] Bombieri, Masser, and Zannier showed that X' is Zariski open in X.

Let h be the absolute logarithmic Weil height. Our contribution in chapter 6 is Theorem 6.1 where we give an explicit bound for the height of algebraic points p in X' that lie "uniformly close" to an algebraic subgroup of codimension strictly greater than $n - n/\dim X$. By uniformly close we mean that there exist an $\epsilon > 0$, independent of p, and an a in an algebraic subgroup of said codimension with $h(pa^{-1}) \leq \epsilon$. Actually, in Theorem 6.1 we will use a weaker notion of uniformly close. The terminology comes from the fact that the map $(p, a) \mapsto h(pa^{-1})$ has similar properties as a distance function. For example it satisfies the triangle inequality. This notion of distance was considered by several authors ([**Eve02**], [**Poo99**], [**Rém03**]) in connection with subgroups of finite rank.

Theorem 6.1 generalizes the Bounded Height Theorem for curves by Bombieri, Masser, and Zannier. We state our theorem such that it also gives an explicit version of a Theorem of Bombieri and Zannier in [**Zan00**] on the intersection of varieties with one dimensional subgroups. To do this we will need a slightly more general definition of X' which is provided in chapter 6.

The height upper bound in Theorem 6.1 involves, along with n, the degree and height of the variety X. We define these two notions in chapter 5. In simple terms, the height of X controls the heights of the coefficients of a certain set of defining equations for X whereas the degree of X controls their degrees. Just as in the second proof for height bounds on curves given in [**BMZ99**], our proof of Theorem 6.1 uses ideas from the geometry of numbers. Given $p \in X(\overline{\mathbf{Q}})$ uniformly close to an algebraic subgroup we construct a new algebraic subgroup H of codimension $\dim X$ and controlled degree, such that pH has normalized height small compared to the height of p. We then intersect

pH with X. The Arithmetic Bézout Theorem bounds the height of isolated points in this intersection leading to an explicit height bound for p.

Lehmer-type lower bounds for heights in spirit of Dobrowolski's Theorem and its generalization to higher dimension provide a method for deducing finiteness results from height bounds as given in chapter 6. This method was used together with algebraic number theory in Bombieri, Masser, and Zannier's article [**BMZ99**] to prove the finiteness of the set of points on X' in an algebraic subgroup of codimension at least 2 if X is a curve. Meanwhile, their intricate argument has been simplified in [**BMZ04**] by applying a more advanced height lower bound due to Amoroso and David [**AD04**]. In this lower bound the degree over $\mathbf{Q}$ of a point is essentially replaced by its degree over the maximal abelian extension of $\mathbf{Q}$. Using this approach we show in Corollary 6.2 that if X is a surface in $\mathbf{G}_m^5$, then there are only finitely many points on X' contained in an algebraic subgroup of codimension at least 3. Thus we have finiteness for the correct subgroup size at least in an isolated case.

Even in presence of a uniform height bound as in Theorem 6.1, the approaches in [**BMZ99**] and [**BMZ04**] cannot be used to prove the finiteness of the set of $p \in X'(\overline{\mathbf{Q}})$ with $h(pa^{-1})$ small and a contained in an algebraic subgroups of appropriate dimension: although pa^{-1} has small height, its degree cannot be controlled. In chapter 7 we pursue a new approach using a Bogomolov-type height lower bound. This bound was proved by Amoroso and David in [**AD03**]; it bounds from below the height of a generic point on a variety not equal to the translate of an algebraic subgroup. The main result of chapter 7 is Theorem 7.1: we show that for $B \in \mathbf{R}$ there exists an $\epsilon = \epsilon(X, B) > 0$ with the following property: there are only finitely many $p \in X'(\overline{\mathbf{Q}})$ with $h(pa^{-1}) \leq \epsilon$ where a is contained in an algebraic subgroup of dimension strictly less than $\mathfrak{m}(\dim X, n)$. In other words, there are only finitely many algebraic points on X' of bounded height which are uniformly close to an algebraic subgroup of dimension less than $\mathfrak{m}(\dim X, n)$. Just as was the case in Theorem 6.1 we actually use a relaxed version of uniformly close in Theorem 7.1. The somewhat unnatural function $\mathfrak{m}(\cdot, \cdot)$ is defined in (7.1.1). At least in the case of curves we have $n - 2 < \mathfrak{m}(1, n)$ and so we can take the subgroups to have the best possible dimension $n - 2$. Unfortunately this is the only interesting case where $\mathfrak{m}(r, n) > n - r - 1$.

With the height upper bound from chapter 6 we can deduce a corollary to Theorem 7.1 which proves finiteness independently of B and where the subgroup dimension is strictly less than $\min\{n/\dim X, \mathfrak{m}(\dim X, n)\}$. Let X be a curve, then this result is optimal with respect to the subgroup dimension. Let us assume that X is not contained in the translate of a proper algebraic subgroup, hence $X' = X$. Then our corollary says that there are only finitely many algebraic points on X that are close to an algebraic subgroup of codimension at least 2. Moreover, in Corollary 7.2 we use Dobrowolski's Theorem to show that if ϵ in the definition of uniformly close is small enough, then all points on X close to an algebraic subgroup of codimension at least 2 are actually contained in such a subgroup.

We now shift our focus from the algebraic torus to abelian varieties: we want to study the intersection $X'(\overline{\mathbf{Q}}) \cap A^{[r]}$ where A is an abelian variety and X is an irreducible closed subvariety of A. The definitions of X' and X'' make sense in the abelian setting and are completely analog to the multiplicative case.

Let X be a curve, then in [**Via03**] Viada proved that $X'(\overline{\mathbf{Q}}) \cap A^{[1]}$ has bounded height if A is a power of an elliptic curve. If the elliptic curve has complex multiplication she also proved that $X'(\overline{\mathbf{Q}}) \cap A^{[2]}$ is finite. Rémond in [**Rém05b**] generalized Viada's height bound to any abelian variety. In [**Rém07**] Rémond applied a generalization of Vojta's inequality which he proved in [**Rém05a**] and in Theorem 1.2 showed boundedness of height of $(X(\overline{\mathbf{Q}}) \backslash Z_X^{(r)}) \cap A^{[r]}$. Here $X \backslash Z_X^{(r)} \subset X$ is a new deprived subset which depends on r. In fact his result holds for a set larger than $A^{[r]}$ involving also the division closure of finitely generated group. If A is isogenous to a product of elliptic curves and if X is a sufficiently general surface which is not contained in the translate of a proper algebraic subgroup then $X \backslash Z_X^{(r)}$ is non-empty and Zariski open in X for $r \geq (\dim A + 3)/2$.

In [**RV03**], Rémond and Viada proved that if X is a curve then $X''(\overline{\mathbf{Q}}) \cap A^{[2]}$ is finite if A is a power of an elliptic curve E with complex multiplication. In a recent preprint, Viada [**Via07**] announced the finiteness of $X''(\overline{\mathbf{Q}}) \cap A^{[3]}$ for unrestricted E, the optimal subgroup codimension 2 is thus just missed.

We announce the following result called the Bounded Height Theorem: if $A = E^g$ is a power of an elliptic curve E and X is an irreducible closed subvariety of arbitrary dimension, then $X'(\overline{\mathbf{Q}}) \cap A^{[\dim X]}$ has bounded Néron-Tate height. Also, using a result from Kirby's Thesis [**Kir06**] and ideas from Bombieri, Masser, and Zannier's [**BMZ06b**] one can show that X' is Zariski open and give a criterion on X to decide when X' is non-empty. Using height lower bounds on abelian varieties with complex multiplication due to Ratazzi in [**Rat07**] we can use the Bounded Height Theorem to show that $X'(\overline{\mathbf{Q}}) \cap A^{[\dim X+1]}$ is finite if E has complex multiplication. For an elliptic curve without complex multiplication, finiteness of $X'(\overline{\mathbf{Q}}) \cap A^{[r]}$ can also be obtained, using for example Rémond's Theorem 2.1 from [**Rém05b**]. But r is in general sub-optimal for such elliptic curves.

The essential difference between the Bounded Height Theorem and Theorem 6.1 is that the subgroups are now allowed to have the best-possible codimension $\dim X$ for all X.

In the future we plan to publish these results.

Pink has stated a general conjecture on mixed Shimura varieties, see [**Pin05a**] and [**Pin05b**]. One special implication is his Conjecture 5.1 from [**Pin05b**]: if A is a semi-abelian variety defined over $\mathbf{C}$ and if $X \subset A$ is a subvariety also defined over $\mathbf{C}$ which is not contained in a proper algebraic subgroup of A, then $X(\mathbf{C}) \cap A^{[\dim X+1]}$ is not Zariski dense in X. Zilber's stronger Conjecture 2 in [**Zil02**] implies the same conclusion. With the Bounded Height Theorem we can prove this assertion under the following stronger hypothesis on A and X: A is a power of an elliptic curve E with complex multiplication and if $\varphi : E^g \to E^{\dim X}$ is a surjective homomorphism of algebraic groups, then the restriction $\varphi|_X : X \to E^{\dim X}$ is dominant.

The proof of the Bounded Height Theorem uses the completeness of A (and X) in an essential way as it relies on intersection theory. Nevertheless, a proof for the boundedness of height of $X'(\overline{\mathbf{Q}}) \cap A^{[\dim X]}$ for the non-complete $X \subset A = \mathbf{G}_m^n$ along the lines of the proof of the Bounded Height Theorem must not be ruled out. For instance one could compactify $\mathbf{G}_m^n \hookrightarrow \mathbf{P}^n$ and work in $\mathbf{P}^n$. Still, there seems to be no suitable Theorem of the Cube for $\mathbf{G}_m^n$. Future research could consist in finding a proof of the Bounded Height Theorem in the multiplicative case or in abelian varieties other than a power of an elliptic curve.

In the two appendices we leave the main path of the thesis. Let P be an irreducible polynomial in two variables with algebraic coefficients. Say x and y are algebraic with $P(x, y) = 0$. In appendix A, motivated by Proposition B of [**BMZ99**], we consider the problem of bounding $|\deg_X(P)h(x) - \deg_Y(P)h(y)|$ explicitly and with good dependency in $h(x)$, $h(y)$, and P.

For simple examples such as $P = X^p - Y^q$ with p and q coprime integers, the absolute value is zero. But for general and fixed P it may even be unbounded as (x, y) runs over all algebraic solutions of P. In Theorem A.1 we prove an upper bound which is of the form $c \max\{1, h_p(P)\}^{1/2} \max\{1, h(x), h(y)\}^{1/2}$ where the constant c is completely explicit and depends only on the partial degrees of P. Here $h_p(P)$ is the projective logarithmic Weil height of the coefficient vector of P. This type of height inequality is often referred to as quasi-equivalence of heights.

In appendix B we demonstrate four known results using the Quasi-equivalence Theorem from appendix A. The first application is the Theorem of Bombieri, Masser, and Zannier, already discussed above, in the case of curves in $\mathbf{G}_m^2$. We then prove a version of Runge's Theorem on the finiteness of the number of solutions of certain diophantine equations. Next we show a result of Skolem from 1929: we first generalize the greatest common divisor of pairs of integers to pairs of algebraic numbers. We then show that if x and y are coprime algebraic numbers and $P(x, y) = 0$ where P is an irreducible polynomial in $\overline{\mathbf{Q}}[X, Y]$ without constant term, then x and y have uniformly bounded height. This result has been proved independently by Abouzaid in [**Abo06**] who used it to prove a variant of the Quasi-equivalence Theorem. The fourth and final application is an explicit version of Sprindzhuk's Theorem: let P have rational coefficients, again without constant term and such that not both partial derivatives of P vanish at $(0, 0)$. Then for a sufficiently large prime l, the polynomial $P(l, Y) \in \mathbf{Q}[Y]$ is irreducible. Since the Quasi-equivalence Theorem gives explicit bounds, so do its four applications.

Chapters 1 and 5 contain no new results but serve as reference for certain theorems which we apply in the rest of the thesis. Chapter 1 introduces the Weil height and related subjects. It is used throughout the thesis. Chapter 5 contains some results from algebraic geometry and gives a definition for the height of a positive dimensional variety. These definitions and results will be used in the second part of the thesis, chapters 6 and 7.

CHAPTER 1

A review of heights

Heights play a central role in this thesis. On the one hand they are an important technical tool to control the "size" of algebraic numbers, often need to prove finiteness results. On the other hand heights have subtle properties which makes them intrinsically interesting. We dedicate this first chapter to a short review of the absolute Weil height. This particular notation of height will be used often, especially in the first part of the thesis. We present basic functional properties and also some notation which will be freely used in this work. We will then define the height and Mahler measure of polynomials and also present some results. Our main reference is Bombieri and Gubler's book [**BG06**]. This chapter should be seen mainly as a source of reference for later chapters. There is not the faintest claim that this chapter gives an complete overview of the topic of heights. In chapter 5 we will revisit heights, but from a different point of view.

1. The Weil height

Recall that an absolute value $|\cdot|$ on a field K satisfies the *triangle inequality* $|x+y| \leq |x| + |y|$ for all $x, y \in K$. The absolute value is called *ultrametric* if it satisfies the *ultrametric inequality* $|x + y| \leq \max\{|x|, |y|\}$ for all $x, y \in K$.

Let K be a number field with ring of algebraic integers $\mathcal{O}_K$. If $I \subset K$ is a fractional ideal, we set $N(I) \in \mathbf{Q}$ to be its norm. We define M_K to be the set of absolute values of K such that their restriction to $\mathbf{Q}$ is the usual p-adic absolute value or the standard complex absolute value. Elements of M_K will be called *places* of K. Let $v \in M_K$ extend $w \in M_{\mathbf{Q}}$; we will write $v \mid w$. If w is the standard complex absolute value on $\mathbf{Q}$ then v will be called an *infinite place*, or $v \mid \infty$ for short. If w is a p-adic absolute value then v will be called a *finite place*, or $v \nmid \infty$ for short. It is well-known that there are one-to-one correspondences between infinite places and embeddings $K \to \mathbf{C}$ up to complex conjugation on the one hand and between finite places and non-zero prime ideals of $\mathcal{O}_K$ on the other hand. We define the *local degree* $d_v = [K_v : \mathbf{Q}_w]$ where K_v, $\mathbf{Q}_w$ are the completions of K, $\mathbf{Q}$ with respect to the absolute values. For integers n it is sometimes useful to define $\delta_v(n) = \max\{1, |n|_v\}$.

If $\tau \in K^* = K\backslash\{0\}$, then there are at most finitely many $v \in M_K$ such that $|\tau|_v \neq 1$. We have the *product formula* (Proposition 1.4.4 [**BG06**])

$$\prod_{v \in M_K} |\tau|_v^{d_v} = 1 \text{ for any } \tau \in K^*. \tag{1.1.1}$$

Let $p = (p_1, \dots, p_n) \in K^n$, the *absolute Weil height* of p is defined as

$$H(p) = \prod_{v \in M_K} \max\{1, |p_1|_v, \dots, |p_n|_v\}^{d_v/[K:\mathbf{Q}]}. \tag{1.1.2}$$

This definition is independent of the field K containing the p_i ([**BG06**] Lemma 1.5.2 page 15) and the height function is thus defined on $\overline{\mathbf{Q}}^n$, for $\overline{\mathbf{Q}}$ the algebraic closure of $\mathbf{Q}$. For the reader's convenience we review some basic properties of the height function:

Say $\tau \in \overline{\mathbf{Q}}$ has minimal polynomial $a_d T^d + \cdots + a_0 \in \mathbf{Q}[T]$ over $\mathbf{Q}$ such that the a_i are integers having no common factor and $a_d \neq 0$. We can calculate the height of τ more directly with the formula

$$H(\tau)^d = |a_d| \prod_{i=1}^{d} \max\{1, |\tau_i|\} \tag{1.1.3}$$

where $\tau_1, \dots, \tau_d \in \mathbf{C}$ are the distinct zeros of P. The equality (1.1.3) follows from Propositions 1.6.5 and 1.6.6 in [**BG06**].

Say $p \in \overline{\mathbf{Q}}^n$ and $k \in \mathbf{N}$, then

$$H(p^k) = H(p)^k. \tag{1.1.4}$$

The proof of (1.1.4) follows directly from the definition (1.1.2). If $\tau \in \overline{\mathbf{Q}}^*$ we can say more, in fact for $k \in \mathbf{Z}$ we have

$$H(\tau^k) = H(\tau)^{|k|}. \tag{1.1.5}$$

This equality follows from (1.1.1) and (1.1.4) for negative k. If $p \in (\overline{\mathbf{Q}}^*)^n$, then in general only $H(p^{-1}) \leq H(p)^n$ holds.

Say $\tau \in \overline{\mathbf{Q}}$, then

$$H(\tau) = 1 \text{ if and only if } \tau \text{ is a root of unity or zero.}$$

This is Kronecker's Theorem, for a proof see Theorem 1.5.9 in [**BG06**].

The height is invariant under multiplication by a root of unity: if $p \in \overline{\mathbf{Q}}^n$ and ζ a root of unity, then $H(\zeta p) = H(p)$.

For $\tau, \mu \in \overline{\mathbf{Q}}$ we may bound

$$H(\tau\mu) \leq H(\tau)H(\mu).$$

This inequality follows from $\max\{1, |\tau\mu|\} \leq \max\{1, |\tau|\} \max\{1, |\mu|\}$ for any absolute $|\cdot|$ on a field containing τ and μ. An analogue bound for sums is

$$H(\tau + \mu) \leq 2H(\tau)H(\mu), \tag{1.1.6}$$

for $\tau, \mu \in \overline{\mathbf{Q}}$. Indeed (1.1.6) follows from

$$\max\{1, |\tau + \mu|\} \leq \delta \max\{1, |\tau|\} \max\{1, |\mu|\}$$

with $\delta = 1$ if $|\cdot|$ satisfies the ultrametric inequality and $\delta = 2$ else wise. If $\mu = 1$ we have the special case $H(1 + \tau) \leq 2H(\tau)$. In general the factor 2 cannot be omitted in this inequality. But in chapter 2 Lemma 2.4 we will improve 2 to $1.909\dots$ for non-zero τ of small height that are not conjugate to τ^{-1}.

If $a \in \mathbf{Z}$ and $b \in \mathbf{N}$ are coprime, then

$$H(a/b) = \max\{|a|, b\}.$$

This equality follows from (1.1.3).

Finally Northcott's Theorem says that for $C, D \in \mathbf{R}$, there are at most finitely many $\tau \in \overline{\mathbf{Q}}$ with $H(\tau) \leq C$ and $[\mathbf{Q}(\tau) : \mathbf{Q}] \leq D$. For a proof see Theorem 1.6.8 on page 25 in [**BG06**].

The product formula (1.1.1) enables us to define a height function on the algebraic points of projective space $\mathbf{P}^n$: say K is still a number field and $p = [p_0 : \cdots : p_n] \in \mathbf{P}^n(\overline{\mathbf{Q}})$ with projective coordinates $p_i \in K$, we set

$$H(p) = \prod_{v \in M_K} \max\{|p_0|_v, \ldots, |p_n|_v\}^{d_v/[K:\mathbf{Q}]}.$$

Because of (1.1.1), another choice of algebraic projective coordinates for p leads to the same height value. The product formula also implies $H(p) \geq 1$.

If $p \in \overline{\mathbf{Q}}^n$, then $H(p)$ is often called the *affine height* and if $p \in \mathbf{P}^n(\overline{\mathbf{Q}})$, then $H(p)$ is called the *projective height*.

For notational purposes it is sometimes useful to take the logarithm. We define the *absolute logarithmic Weil height* $h(p) = \log H(p)$ for $p \in \overline{\mathbf{Q}}^n$ or $p \in \mathbf{P}^n(\overline{\mathbf{Q}})$.

If $p = (p_1, \ldots, p_n) \in \overline{\mathbf{Q}}^n$, then we have the useful estimates $h(p) \leq h(p_1) + \cdots + h(p_n)$ and $\max\{h(p_1), \ldots, h(p_n)\} \leq h(p)$ which follow from local considerations.

2. Height and Mahler measure of a polynomial

Keeping track of bounds of the size of a polynomial in integer coefficients is important in transcendence theory when for example constructing auxiliary functions. "Size" could mean for example the maximum of the absolute values of the coefficients or the sum over these values. Of course more intricate definitions which anticipate a common divisor of the coefficients or which work if the coefficients are in a number fields are possible.

Let K be any field with an absolute value $|\cdot|$. If $f \in K[X_1, \ldots, X_n]$ is a polynomial with coefficients $f_{i_1 \ldots i_n} \in K$, then we set

$$|f| = \max_{i_1, \ldots, i_n} \{|f_{i_1 \ldots i_n}|\}.$$

If $|\cdot|$ satisfies the ultrametric inequality and $g \in K[X_1, \ldots, X_n]$, then by Gauss's Lemma we have $|fg| = |f||g|$. In general equality does not hold if $|\cdot|$ does not satisfy the ultrametric inequality.

Assume now that K is a number field and that $f \in K[X_1, \ldots, X_n]$ is non-zero. Then we define the *height* of f as

$$h_p(f) = \frac{1}{[K : \mathbf{Q}]} \sum_{v \in M_K} d_v \log |f|_v.$$

Therefore the height of f is just the logarithmic absolute Weil height of the point in $\mathbf{P}^n(\overline{\mathbf{Q}})$ whose projective coordinates are just the non-zero coefficients of f. By (1.1.1)

and if $\lambda \in \overline{\mathbf{Q}}^*$ then $h_p(\lambda f) = h_p(f)$. We use the subscript p to distinguish $h_p(\cdot)$ from $h(\cdot)$ as there is danger of ambiguity if the polynomial in question is constant.

Let $f \in \mathbf{C}[X_1^{\pm 1}, \ldots, X_n^{\pm 1}] \backslash \{0\}$ and say $|\cdot|$ is now the complex absolute value. The *logarithmic Mahler measure* of f is given by

$$m(f) = \int_0^1 \cdots \int_0^1 \log |f(e^{2\pi i t_1}, \ldots, e^{2\pi i t_n})| dt_1 \cdots dt_n. \tag{1.2.1}$$

It is not immediately clear that this integral converges, as the absolute value in the logarithm may vanish. For a proof of the existence of $m(f)$ we refer to Lemma 2 page 223 in [**Sch00**]. We also define $M(f) = \exp m(f)$ and $M(0) = 0$. A direct and nice consequence of (1.2.1) is the that $M(fg) = M(f)M(g)$ for any $f, g \in \mathbf{C}[X_1^{\pm 1}, \ldots, X_n^{\pm 1}]$. The Mahler measure $M(f)$ can be bounded above and below by positive multiples of $|f|$. These factors depend only on the partial degrees of f.

If $f \in \mathbf{C}[X]$ is a polynomial in one variable with $f = a_d(X - \alpha_1) \cdots (X - \alpha_d)$ and $\alpha_i \in \mathbf{C}$, then by Jensen's formula we have

$$M(f) = |a_d| \max\{1, |\alpha_1|\} \cdots \max\{1, |\alpha_d|\}. \tag{1.2.2}$$

Let $\tau \in \overline{\mathbf{Q}}$ have minimal polynomial $f \in \mathbf{Q}[X]$ with coprime integer coefficients, then by (1.1.3) and (1.2.2) we have

$$h(\tau) = \frac{1}{\deg(f)} m(f). \tag{1.2.3}$$

Hence heights and Mahler measures of integer polynomials are closed related.

In general it is non-trivial to calculate the exact value of the Mahler measure of a polynomial in more than one variable. Smyth's work [**Smy81**] contains some explicit evaluations. For example he calculates the Mahler measure of the two variable polynomial $X + Y - 1$ in terms of an L-function. We will see more of this number in chapter 3.

Even numerically approximating the integral in (1.2.1) can be tricky because of the possible singularities in the integrand. In [**Boy98**] Boyd proposed an effective method for calculating $M(f)$ up to arbitrary precision avoiding the approximation of an integral. But he states that his method is impractical from a computational point of view if f has more than one variable. Nevertheless Boyd's general approach has an important theoretical consequence. The finite dimensional $\mathbf{C}$-vector space of polynomials in $\mathbf{C}[X_1, \ldots, X_n]$ of degree bounded by some parameter carries a natural topology induced by the topology on $\mathbf{C}$. Boyd shows that for polynomials of bounded degree $M(f)$ is a continuous function of f. In the one-variable case this is a classical result going back to Mahler.

In chapter 3 we will define a family of complex polynomials depending on a real parameter and study the function which maps this real parameter to the Mahler measure of the corresponding polynomial. More precisely, we show a uniform bound on the variation of this function. Unfortunately Boyd or Mahler's results on continuity do not suffice in our situation.

We present two results on bounds for Mahler measures from Schinzel's book [**Sch00**] which will be used in chapter 3.

We begin with an upper bound for the Mahler measure of a complex polynomial f in terms of the hermitian norm of the coefficient vector of P. Concretely, if $P = f_d X^d + \cdots + f_0$ with $f_i \in \mathbf{C}$, then we set $|f|_2 = (|f_0|^2 + \cdots + |f_d|^2)^{1/2}$.

THEOREM 1.1 (Gonçalvez). *Let $f = f_d X^d + \cdots + f_0 \in \mathbf{C}[X]$ with $f_0 f_d \neq 0$. Then*

$$M(f)^2 + |f_0 f_d|^2 M(f)^{-2} \leq |f|_2$$

with equality if and only if $f(X)\overline{f}(X^{-1})$ has just three non-zero coefficients.

PROOF. This theorem is a special case of Theorem 40 in [**Sch00**]. □

The second result needed in chapter 3 is a lower bound for the Mahler measure of polynomials defined over certain number fields. We need a few definitions first.

Let $f \in \mathbf{C}[X]\backslash\{0\}$, then f is called *self-inverse* if there exists $\lambda \in \mathbf{C}^*$ with $X^{\deg f} f(X^{-1}) = \lambda \overline{f}(X)$.

Say now $f \in K[X]$ where K is a number field. The *content* of f, denoted by $\mathrm{cont}(f)$, is the fractional ideal in K generated by the coefficients of f. For example if f has coefficients in $\mathcal{O}_K$, then $\mathrm{cont}(f)$ is an ideal in $\mathcal{O}_K$.

We call a number field K *Kroneckerian* if K is a either totally real or a totally complex quadratic extension of a totally real number field.

We may now state Schinzel's Theorem.

THEOREM 1.2 (Schinzel). *Let K be a Kroneckerian field and $f \in K[X]\backslash\{0\}$ not self-inverse with $f(0) \neq 0$. Then*

$$\prod_{\sigma} M(\sigma(f)) \geq \left(\frac{1+\sqrt{5}}{2}\right)^{[K:\mathbf{Q}]/2} N(\mathrm{cont}(f)), \tag{1.2.4}$$

where the product runs over all embeddings $\sigma : K \to \mathbf{C}$.

PROOF. This result is a special case of Theorem 72 in [**Sch00**]. Schinzel proves this theorem for Laurent polynomials in any number of variables. □

The upper bound in Theorem 1.1 and the lower bound in 1.2 are so sharp that we will use them to deduce irreducibility results for certain trinomials arising in chapter 3.

We note that Theorem 1.2 in the case $K = \mathbf{Q}$ implies a lower bound as in Smyth's paper [**Smy71**] with a slightly worse constant. The important aspect of Schinzel's Theorem is the exponent $[K : \mathbf{Q}]$ in (1.2.4).

CHAPTER 2

Multiplicative dependence and isolation I

In [**CZ00**] Cohen and Zannier proved that if x is algebraic with x and $1-x$ multiplicatively dependent, then $\max\{H(x), H(1-x)\} \leq 2$; here $H(\cdot)$ is the absolute Weil height. This bound is sharp because of the exceptional values $x = -1, 1/2, 2$. Cohen and Zannier then used Bilu's Equidistribution Theorem [**Bil97**] to prove an isolation result: they showed that there exists $\epsilon > 0$ such that if x is as before but not one of the three exceptional values, then $\max\{H(x), H(1-x)\} \leq 2 - \epsilon$. In this chapter we give a concise proof of a slight strengthening of the height bound given in [**CZ00**]. We also work out an explicit ϵ. Finally we show that there exists a sequence x_n with x_n and $1 - x_n$ multiplicatively dependent and such that the height of x_n converges to the Mahler measure of the polynomial $X + Y - 1$.

The contents of this chapter will appear in the Proceedings of the Pisa research program "Diophantine Geometry" [**Hab07**].

1. Height bounds for dependent solutions of $x + y = 1$

Two elements x, y of a field are called *multiplicatively dependent* if $xy \neq 0$ and if there exist $r, s \in \mathbf{Z}$ not both zero such that $x^r y^s = 1$.

We define $\mathcal{M}$ to be the set of complex x such that x and $1 - x$ are multiplicatively dependent. Clearly the elements of $\mathcal{M}$ are algebraic. If $\zeta \neq 1$ is a root of unity, then ζ and $1 - \zeta$ are multiplicatively dependent and so $\zeta \in \mathcal{M}$. Thus $\mathcal{M}$ is infinite, a result which was made quantitative by Masser in Theorem 2 of [**Mas07**]. If y is also algebraic then $H(x, y)$ denotes the affine absolute non-logarithmic Weil height, which was defined in chapter 1. This height function corresponds to the compactification of the algebraic torus $\mathbf{G}_m^2 \hookrightarrow \mathbf{P}^2$. We have:

THEOREM 2.1. *Let $x \in \mathcal{M}$, then $H(x, 1-x) \leq 2$ with equality if and only if $x \in \{-1, 1/2, 2\}$.*

Theorem 2.1 implies Theorem 1 of [**CZ00**] since $\max\{H(x), H(y)\} \leq H(x, y)$ for algebraic x and y. We choose the particular height function $H(x, 1-x)$ because it is invariant under the maps $x \mapsto 1 - x$ and $x \mapsto x^{-1}$. Incidentally $\mathcal{M}$ is stable under these two maps. Our method of proof for Theorem 2.1 exploits this fact and relies on elementary local estimates combined with the product formula. The proof of Theorem 2.1 is a warm up for proof of Theorem 4.1 which gives a height bound for dependent x, y with $x + y = \alpha$.

Theorem 2.1 makes explicit a special case of

THEOREM ([**BMZ99**] page 1120). *Let $\mathcal{C}$ be a closed absolutely irreducible curve in $\mathbf{G}_m^n$, $n \geq 2$, defined over $\overline{\mathbf{Q}}$ and not contained in a translate of a proper subtorus of $\mathbf{G}_m^n$.*

Then the algebraic points of $\mathcal{C}$ which lie in the union of all proper algebraic subgroups of $\mathbf{G}_m^n$ form a set of bounded Weil height.

Indeed $x + y = 1$ defines a line $\mathcal{C}$ in $\mathbf{G}_m^2$ which is not contained in a translate of a proper subtorus. And any proper algebraic subgroup of $\mathbf{G}_m^2$ is contained in some set defined by $x^r y^s = 1$. Finally the Weil height used in [**BMZ99**] was the expression $H(x)H(y) \geq H(x, y)$.

To prove the isolation result with explicit ϵ mentioned above we apply a result of Mignotte from [**Mig89**] on the angular distribution of conjugates of an algebraic number of small height and big degree. Actually, large degree is guaranteed by a theorem of Smyth [**Smy71**] on lower bounds for heights of non-reciprocal algebraic numbers.

THEOREM 2.2. *If $x \in \mathcal{M} \backslash \{-1, 1/2, 2\}$ then $H(x, 1 - x) < 1.915$.*

The element of $\mathcal{M} \backslash \{-1, 1/2, 2\}$ of largest height known to the author is $1 - \zeta_3$ where ζ_3 is a primitive 3rd root of unity. In fact $H(1 - \zeta_3) = \sqrt{3}$. It would be interesting to know if $\sqrt{3}$ is already the second to largest height value obtained on $\mathcal{M}$. In chapter 3 we will develop an algorithm to decide this question.

If $\zeta \neq 1$ is a root of unity, then $1 - \zeta \in \mathcal{M}$. As the degree of ζ goes to infinity we can use Bilu's Equidistribution Theorem (Theorem 1.1, [**Bil97**]) to show that $H(1 - \zeta, \zeta) = H(1 - \zeta)$ converges to

$$\exp \int_{-1/3}^{1/3} \log |1 + \exp(2\pi i t)| dt = 1.381356..., \tag{2.1.1}$$

Let f be a polynomial in n variables with complex coefficients, the Mahler measure $M(f)$ of f was defined in chapter 1 (1.2.1). Smyth ([**Smy81**]) calculated the Mahler measure of the polynomial $X + Y - 1$ as

$$M(X + Y - 1) = \exp\left(\frac{3\sqrt{3}}{4\pi} \sum_{k \geq 1} \left(\frac{k}{3}\right) \frac{1}{k^2}\right), \tag{2.1.2}$$

here $\left(\frac{\cdot}{\cdot}\right)$ is the Legendre symbol. By Jensen's formula the Mahler measure in (2.1.2) is equal to the integral (2.1.1). We immediately obtain:

PROPOSITION 2.1. *There exists a sequence $x_n \in \mathcal{M}$ with $\lim_{n \to \infty} [\mathbf{Q}(x_n) : \mathbf{Q}] = \infty$ such that $\lim_{n \to \infty} H(x_n, 1 - x_n) = M(X + Y - 1)$.*

In chapter 3 we will prove a much stronger result. In fact $H(x_n, 1 - x_n)$ converges to $M(X + Y - 1)$ for any sequence $x_n \in \mathcal{M}$ such that $[\mathbf{Q}(x_n) : \mathbf{Q}]$ goes to infinity.

2. Proof of Theorem 2.1

We prove Theorem 2.1 via an elementary estimate which holds for any field K with any absolute value $|\cdot| : K \to \mathbf{R}$.

LEMMA 2.1. *Let $x \in K\backslash\{0,1\}$, $r,t \in \mathbf{Z}$ with $0 \neq t \geq r \geq 0$ and $x^r = (1-x)^t$. We have*

$$|1-x|^{-1}\max\{1,|x|\} \leq \delta \tag{2.2.1}$$

where $\delta = 1$ if $|\cdot|$ is ultrametric and $\delta = 2$ otherwise. Furthermore, equality in (2.2.1) implies $\delta = 1$ or $r = 0$ or $r = t$.

PROOF. Let q denote the left-hand side of (2.2.1).

First let us assume

$$|x| < \delta^{-1} \text{ or } |x| > \delta. \tag{2.2.2}$$

If $\delta = 1$, then $|x| \neq 1$, so $|1-x| = \max\{1,|x|\}$, hence $q = 1$. If $\delta = 2$ we use the triangle inequality to bound

$$|1-x| \geq \begin{cases} |x|-1 > |x|\delta^{-1} & : \text{ if } |x| > \delta, \\ 1-|x| > \delta^{-1} & : \text{ if } |x| < \delta^{-1} \end{cases}$$

which implies $q < \delta$. So in the case (2.2.2) we have $q \leq \delta$ and furthermore $q = \delta$ can only hold if $\delta = 1$.

Now let us assume $\delta^{-1} \leq |x| \leq \delta$. If $|x| < 1$, then $q = |1-x|^{-1} = |x|^{-r/t} \leq \delta^{r/t} \leq \delta$, and if $|x| \geq 1$, then $q = |x|/|1-x| = |x|^{1-r/t} \leq \delta^{1-r/t} \leq \delta$. It is clear that if we have the equalities $q = \delta = 2$, then $r = 0$ or $r = t$. □

We will see more arguments in this style in chapter 4.

LEMMA 2.2. *If $\zeta \neq 1$ is a root of unity, then $H(1+\zeta) \leq \sqrt{2\sqrt{3}} = 1.8612...$*

PROOF. Let K be a number field of degree d containing ζ. We multiply the product formula $\prod_{v\in M_K} |1-\zeta|_v^{d_v} = 1$ with the definition of the height and note that ζ is an algebraic integer to get

$$H(1+\zeta)^d \leq \min\{\prod_{v|\infty} \max\{1,|1+\zeta|_v\}^{d_v}, \prod_{v|\infty} \max\{|1-\zeta|_v, |1-\zeta^2|_v\}^{d_v}\}.$$

Let Δ_1 be the set of infinite places v with $|1-\zeta|_v \geq 1$, let Δ_2 be all other infinite places. Recall that infinite places correspond to embeddings of K into $\mathbf{C}$ up to conjugation. If $v \in \Delta_1$, then elementary geometry gives $|1+\zeta|_v \leq \sqrt{3}$; with the right-hand side replaced by 2 if we allow $v \in \Delta_2$. Similarly if $v \in \Delta_2$, then $\max\{|1-\zeta|_v, |1-\zeta^2|_v\} \leq \sqrt{3}$; and if $\sqrt{3}$ is replaced by 2, then the inequality holds for $v \in \Delta_1$. We define $\delta_i = \sum_{v\in\Delta_i} d_v/d$, then $\delta_1 + \delta_2 = 1$ and so

$$H(1+\zeta) \leq \min\{\sqrt{3}^{\delta_1}2^{\delta_2}, 2^{\delta_1}\sqrt{3}^{\delta_2}\} = \sqrt{3}(2\sqrt{3}^{-1})^{\min\{\delta_1,1-\delta_1\}} \leq \sqrt{2\sqrt{3}}.$$

□

We note that $\frac{1}{2}\log(2\sqrt{3})$ is an improvement of the trivial bound $h(1+\zeta) \leq \log 2$ which holds for any root of unity ζ. In the proof of Theorem 2.1 we need only a weak form of Lemma 2.2, namely the fact that $H(1+\zeta) < 2$ is $\zeta \neq 1$ is a root of unity.

LEMMA 2.3. *Let $x' \in \mathcal{M}$, then there exist $x \in \mathcal{M}$ and $r, t \in \mathbf{Z}$ with $0 \neq t \geq 2r \geq 0$ such that $x^r = (1-x)^t$ and $h(x', 1-x') = h(x)$. Furthermore, if $x' \notin \{-1, 1/2, 2\}$ then we can choose x such that $x \notin \{-1, 1/2, 2\}$.*

PROOF. The lemma is simple if x' is a primitive 6th root of unity, for then $1 - x'$ is also a 6th root of unity and we may take $x = x'$, $t = 6$, and $r = 0$. Hence it suffices to show the lemma for $x' \in \mathcal{M}_0$ with $\mathcal{M}_0 = \mathcal{M} \backslash \{e^{\pm 2\pi i/6}\}$. For any such x' there exists a unique $\lambda(x') = [r : t] \in \mathbf{P}^1(\mathbf{Q})$ with r and t coprime integers such that $x'^r(1-x')^{-t}$ is a root of unity. The maps $\phi_1(x) = 1/x$ and $\phi_2(x) = 1 - x$ are automorphisms of $\mathcal{M}_0$ and generate the symmetric group S_3. Thus we get an action of S_3 on $\mathcal{M}_0$ which also leaves $\{-1, 1/2, 2\}$ invariant. By the product formula the height $h(x, 1-x)$ is also invariant under this action. We check that if $\lambda(x) = [r : t]$, then $\lambda(\phi_1(x)) = [t - r : t]$ and $\lambda(\phi_2(x)) = [t : r]$. We get an action of S_3 on $\mathbf{P}^1(\mathbf{R})$. Furthermore, any element of $\mathbf{P}^1(\mathbf{R})$ lies in the orbit of some element of $\{[1 : s];\ s \geq 2\} \cup \{[0 : 1]\}$. The lemma follows since if $x^r(1-x)^{-t}$ is a root of unity with $t \geq 2r \geq 0$, then $h(x, 1-x) = h(x)$. □

Proof of Theorem 2.1: Because of Lemma 2.3 it suffices to show that if $x \in \overline{\mathbf{Q}} \backslash \{0, 1\}$ with $x \neq -1, 1/2, 2$ and $x^r = (1-x)^t$ for integers $0 \neq t \geq 2r \geq 0$, then $h(x) < \log 2$.

If $r = 0$, then $x = 1 + \zeta$ for some root of unity $\zeta \neq \pm 1$. In this case the theorem follows from Lemma 2.2.

Let us assume $r > 0$. We fix a number field K that contains x and apply the product formula (1.1.1) to $1 - x$ to deduce

$$[K : \mathbf{Q}]h(x) = \sum_v d_v \log \max\{1, |x|_v\} = \sum_v d_v \log \frac{\max\{1, |x|_v\}}{|1-x|_v},$$

where d_v are the local degrees from chapter 1. Since $0 < r < t$ we apply Lemma 2.1 to the local terms in the equality above to see that $[K : \mathbf{Q}]h(x) < \sum_{v \text{ infinite}} [K_v : \mathbf{Q}_v] \log 2$. This inequality completes the proof since the sum is just $[K : \mathbf{Q}] \log 2$. □

3. Proof of Theorem 2.2

A non-zero algebraic number α is called reciprocal if α and α^{-1} are conjugated. We apply Mignotte's equidistribution result and Smyth's Theorem ([**Smy71**]) on lower bounds for heights of non-reciprocal algebraic integers to deduce the following lemma.

LEMMA 2.4. *Let $\alpha \in \overline{\mathbf{Q}}^*$ be non-reciprocal with $h(\alpha) \leq \frac{\log 2}{3 \cdot 10^5}$, then $h(1 + \alpha) \leq 0.933 \cdot \log 2 + h(\alpha)$.*

PROOF. Let α be as in the hypothesis and $d = [\mathbf{Q}(\alpha) : \mathbf{Q}]$, furthermore let $\theta_0 > 1$ be the unique real that satisfies $\theta_0^3 - \theta_0 - 1 = 0$. If α is an algebraic integer, then $dh(\alpha) \geq \log \theta_0$ by Smyth's Theorem ([**Smy71**]). The upper bound for $h(\alpha)$ implies

$$d \geq 121700. \tag{2.3.1}$$

On the other hand, if α is not an algebraic integer, then it is well-known that $dh(\alpha) \geq \log 2$. Thus (2.3.1) holds in any case.

We split $\mathbf{C}^*$ up into three sectors

$$C_k = \{r \cdot \exp(i\phi);\ r > 0 \text{ and } \frac{2\pi}{3}(k-1) \leq \phi < \frac{2\pi}{3}k\} \text{ for } 1 \leq k \leq 3$$

and define the function

$$m(z) = \frac{\max\{1, |z+1|\}}{\max\{1, |z|\}} = \frac{\max\{1, (r^2 + 2r\cos\phi + 1)^{1/2}\}}{\max\{1, r\}}$$

for $z = r \cdot \exp(i\phi)$ with $r > 0$ and $\phi \in \mathbf{R}$. Hence

$$m(z)^2 \leq \begin{cases} \frac{\max\{1, r^2+2r+1\}}{\max\{1, r^2\}} & : \text{ if } -2\pi/3 \leq \phi \leq 2\pi/3 \\ \frac{\max\{1, r^2-r+1\}}{\max\{1, r^2\}} & : \text{ if } 2\pi/3 \leq \phi \leq 4\pi/3. \end{cases}$$

Elementary calculus now leads to

$$m|_{C_1 \cup C_3} \leq 2 \text{ and } m|_{C_2} \leq 1. \tag{2.3.2}$$

We fix an embedding of $\overline{\mathbf{Q}}$ into $\mathbf{C}$. Let $\alpha_1 \ldots, \alpha_d \in \mathbf{C}^*$ be the conjugates of α. We set $N_k = |\{i;\ \alpha_i \in C_k\}|$ for $1 \leq k \leq 3$. For any finite place v of $\mathbf{Q}(\alpha)$ we have $\max\{1, |1+\alpha|_v\} = \max\{1, |\alpha|_v\}$ by the ultrametric inequality. Since the infinite places of $\mathbf{Q}(\alpha)$ taken with multiplicities correspond to embeddings of $\mathbf{Q}(\alpha)$ into $\mathbf{C}$ and because of (2.3.2) we have

$$d(h(1+\alpha) - h(\alpha)) = \sum_{i=1}^{d} \log m(\alpha_i) \leq (N_1 + N_3)\log 2. \tag{2.3.3}$$

We set $\epsilon = (\frac{9}{2}c^2(\frac{\log(2d+1)}{d} + h(\alpha)))^{1/3}$ with $c = 2.62$. Since $\frac{\log(2d+1)}{d}$ is decreasing considered as function of $d \geq 1$, we use (2.3.1) and our hypothesis on $h(\alpha)$ to conclude $\epsilon < 0.1477$. We apply Mignotte's Theorem (ii) ([**Mig89**], page 83) to the minimal polynomial of α and to the closure of our sectors C_k to bound

$$\frac{N_k}{d} \leq \frac{1}{3} + 2.823(\frac{\log(2d+1)}{d} + h(\alpha))^{1/3}. \tag{2.3.4}$$

Our hypothesis on $h(\alpha)$ and (2.3.1) together with (2.3.4) imply $\frac{N_k}{d} < 0.4662$. This last bound applied to (2.3.3) concludes the proof. □

We note that the trivial bound $h(1+\alpha) \leq \log 2 + h(\alpha)$ holds for all algebraic α. Thus Lemma 2.4 gives a slight improvement for non-reciprocal α of small height. Instead of Smyth's lower bound for heights we could have used the lower bound by Dobrowolski which holds for any non-zero algebraic number not a root of unity. This approach leads to slightly worse numerical constants. By taking sectors with smaller angles in the proof of Lemma 2.4 the constant $0.933 \cdot \log 2$ can be replaced by any real number strictly greater than the logarithm of the number (2.1.1) if the height of α is sufficiently small but positive. But the bound given in Lemma 2.4 is apt for our application.

In [**CZ00**] Cohen and Zannier introduced a function $S : (1, \infty) \to \mathbf{R}$ relevant to our problem. We briefly recall its definition. Say $\lambda > 1$ and let $\xi, \tilde{\xi} > 1$ be the unique reals such that $\xi^\lambda = \xi + 1$ and $\tilde{\xi}^{\lambda/(\lambda-1)} = \tilde{\xi} + 1$, then

$$S(\lambda) = \frac{\log(\xi + 1)\log(\tilde{\xi} + 1)}{\log(\xi + 1) + \log(\tilde{\xi} + 1)}.$$

Lemma 1 of [**CZ00**] implies $S < \log 2$, furthermore if $x^r = (1 - x)^t$ for integers $t > r > 0$, then $h(x) \leq S(t/r)$. The proof of said lemma also shows that S increases on $[2, \infty)$.

Proof of Theorem 2.2: Because of Lemma 2.3 it suffices to show that if $x \in \mathbf{Q}\backslash\{0, 1\}$, $x \neq -1, 1/2, 2$ and $x^r = (1 - x)^t$ for integers $0 \neq t \geq 2r \geq 0$, then $h(x) < \log 1.915$. If $r = 0$, then $x = 1 + \zeta$ for a root of unity $\zeta \neq 1$. Lemma 2.2 implies $h(x) \leq \frac{1}{2}\log(2\sqrt{3}) < \log 1.915$. We now assume $r > 0$ and define $\lambda = t/r \geq 2$.

If $\lambda < 3 \cdot 10^5$, then we have $h(x) \leq S(3 \cdot 10^5)$ by the properties of $S(\cdot)$. A calculation shows that the right-hand side of the last inequality is strictly less then $\log 1.915$.

Finally we assume $\lambda \geq 3 \cdot 10^5$. Then $h(1 - x) = \lambda^{-1}h(x) \leq \frac{\log 2}{3 \cdot 10^5}$ by Theorem 2.1. Let $\alpha = x - 1$, we have

$$(-1)^t\alpha^t = (1 + \alpha)^r. \tag{2.3.5}$$

Let us assume first that α and α^{-1} are not conjugated, then $h(x) \leq 0.933 \cdot \log 2 + \frac{\log 2}{3 \cdot 10^5} < \log 1.915$ by Lemma 2.4. If α and α^{-1} are conjugated, then equality (2.3.5) must hold with α replaced by α^{-1}. Hence $1 = \alpha^{2t}(1 + \alpha^{-1})^{2r}$, or $1 = x^{2r}(1 - x)^{2(t-r)}$ in terms of x. Since $r \neq 0$ and $r \neq 2t$ this new dependency relation between x and $1 - x$ is independent of the original relation $1 = x^r(1 - x)^{-t}$. We conclude that x is a root of unity and so $h(x) = 0$. □

CHAPTER 3

Multiplicative dependence and isolation II

In this chapter we continue the study of multiplicatively dependent solutions x, y of the equation $x + y = 1$. We recall that $\mathcal{M}$ is the set of $x \in \overline{\mathbf{Q}}\backslash\{0, 1\}$ such that x and $1 - x$ are multiplicatively dependent. It was first proved by Cohen and Zannier in [**CZ00**] that if $x \in \mathcal{M}$, then $\max\{h(x), h(1 - x)\} \leq \log 2$ with equality if and only if $x \in \{-1, 1/2, 2\}$. Furthermore, they showed that $\log 2$ is an isolated point in the range of this particular height function restricted to $\mathcal{M}$. In chapter 2 we worked with the possibly larger height $h(x, 1 - x)$ and also got $\log 2$ as an upper bound. We also showed that if $x \in \mathcal{M}\backslash\{-1, 1/2, 2\}$, then $h(x, 1 - x) \leq \log 1.915$, thus making Cohen and Zannier's isolation result explicit. It is thus natural to ask for the exact value of the greatest limit point of the height. We will answer this question and even determine all limit points.

1. Introduction

Let

$$S_0 = m(X + Y - 1) = 0.3230659472194505140936\ldots$$

where $m(\cdot)$ is the logarithmic Mahler measure of a polynomial in complex coefficients, see (1.2.1). One may interpret S_0 as the logarithmic height of the curve defined by $X + Y = 1$ in $\mathbf{G}_m^2$. In [**Smy81**] Smyth expressed S_0 in terms of the value of the L-function associated to χ, the (unique) non-trivial character of conductor 3, at $s = 2$, cf. chapter 2 equation (2.1.2).

In this chapter we prove that the set of height values

$$\Sigma = \{h(x, 1 - x);\, x \in \mathcal{M}\}$$

has exactly one limit point equal to S_0. It follows that all points in Σ with at most one exception are isolated. We cannot discard of this one exception because we cannot exclude that S_0 is itself a height value. In fact it seems to be already unknown if S_0 is the logarithm of an algebraic number or not.

We now state the Limit Point Theorem:

THEOREM 3.1. *Let $\epsilon > 0$ there is $c(\epsilon)$ depending only on ϵ such that if $x \in \mathcal{M}$ and $D = [\mathbf{Q}(x) : \mathbf{Q}]$ then*

$$|h(x, 1 - x) - S_0| \leq \frac{c(\epsilon)}{D} e^{(\epsilon + \log 2)\frac{\log D}{\log\log 3D}}.$$

The constant $c(\epsilon)$ in Theorem 3.1 is effective, see Proposition 3.1 for an explicit bound of $|h(x, 1-x) - S_0|$.

Of course the theorem implies the boundedness of $h(x, 1-x)$ for $x \in \mathcal{M}$, but it does not give the optimal bound $\log 2$.

We deduce two immediate consequences of the previous theorem and Northcott's Theorem.

COROLLARY 3.1. *Let $x_n \in \mathcal{M}$ be a sequence such that each x_n occurs only finitely often, then $h(x_n, 1-x_n) \to S_0$ as $n \to \infty$.*

COROLLARY 3.2. *The set $\{h(x, 1-x);\, x \in \mathcal{M}\} \backslash \{S_0\} \subset \mathbf{R}$ is discrete.*

It seems natural to conjecture that the set $\{h(x, 1-x);\, x \in \mathcal{M}\} \subset \mathbf{R}$ is discrete.

The essential tool used in the proof of the Limit Point Theorem is the following irreducibility result on certain trinomials which may be of independent interest. We recall that a number field is called *Kroneckerian* if it is either totally real or a totally complex quadratic extension of a totally real number field.

THEOREM 3.2. *Assume K is a Kroneckerian number field. Let $n \geq m \geq 0$ be coprime integers, ξ, ζ roots of unity in K, and $P = X^n + \xi X^m + \zeta \in K[X]$. Let $P = AB$ where $A, B \in K[X]$, $A \notin K$, and such that the zeros of B are precisely the roots of unity that are zeros of P. Then $\deg B \leq 2$, furthermore A is irreducible in $K[X]$ except if $P = X^2 + \xi X - \xi^2$ and $\sqrt{5} \in K$.*

This theorem may be compared with results obtained independently by Ljunggren in [**Lju60**] and by Tverberg in [**Tve60**]. They give a complete factorization over $\mathbf{Q}$ of trinomials such as in Theorem 3.2 with $\xi, \zeta \in \{\pm 1\}$.

A result related to Theorem 3.2 was proved by Schinzel in Corollary 8 ([**Sch00**] page 417). As in his proof we prove the irreducibility of A by showing that if g is the number of irreducible factors of A, then $g < 2$. This bound for g follows from comparing upper and lower bounds for the Mahler measure of polynomials with coefficients in K. The upper bound is due to Gonçalvez (Theorem 1.1) and the lower bound is Theorem 1.2 showed by Schinzel. Our proof of Theorem 3.2 follows Schinzel's proof of Corollary 8.

Because $X^2 + \xi X - \xi^2 = (X - \xi\alpha)(X - \xi\beta)$ for $\alpha = (-1+\sqrt{5})/2$, $\beta = (-1-\sqrt{5})/2$, this polynomial is reducible over $\mathbf{Q}(\xi)$ precisely when this field contains $\sqrt{5}$. Furthermore, this situation does occur for a root of unity ξ. Indeed since $\mathbf{Q}(\sqrt{5})$ is abelian extension of $\mathbf{Q}$ it is contained in a field generated by a root of unity by the Theorem of Kronecker-Weber. Or by an elementary argument: if ξ is a primitive 5th root of unity, then it is well-known that $(\xi - \xi^2 - \xi^3 + \xi^4)^2 = 5$.

For example Theorem 3.2 implies that $X^{2007} + X + 1$ is irreducible in $\mathbf{Q}^{\mathrm{ab}}[X]$, where $\mathbf{Q}^{\mathrm{ab}}$ is the maximal abelian extension of $\mathbf{Q}$.

We give a short description of the proof of Theorem 3.1: Theorem 3.2 enables us to express the height of essentially all $x \in \mathcal{M}$ in terms of the average over the Mahler measure of the conjugates of a certain trinomial much like the one in Theorem 3.2. We use Koksma's inequality to compare this average with a 2-dimensional integral which

equals S_0. Two steps are required to get an explicit bound in Koksma's inequality. First we apply a classical result of Erdös and Turán to show that roots of unity of fixed order are sufficiently well distributed around the unit circle. Second we need a uniform bound for the total variation of a family of functions related to Mahler measures of certain trinomials. The key estimate in bounding the total variation will be Lemma 3.13 which follows from a lengthy calculation.

We remark that it is not completely evident that the set of height values Σ is an infinite set. The infinitude of Σ can be proved by an argument kindly mentioned to me by Masser and Zannier. In fact if ζ is a primitive pth root of unity and ξ is a primitive qth root of unity with p and q primes, then $h(1-\zeta) = h(1-\xi)$ if and only if $p = q$. Indeed let us assume $p \neq q$ and $H(1-\zeta) = H(1-\xi)$ where $H(\cdot) = \exp h(\cdot)$. By the definition of the height it is clear that $H(1-\zeta)^{2(p-1)(q-1)} = H(1-\xi)^{2(p-1)(q-1)} \in \mathbf{Q}(\zeta) \cap \mathbf{Q}(\xi)$. These two fields have intersection $\mathbf{Q}$, hence $H(1-\zeta)^{2(p-1)(q-1)}$ is a rational and even an integer. Since any conjugate of $1-\zeta$ generates the unique prime ideal above p in the ring integers of $\mathbf{Q}(\zeta)$ we conclude that $H(1-\zeta)^{2(p-1)(q-1)}$ is a power of p. Similarly $H(1-\zeta)^{2(p-1)(q-1)}$ is a power of q. An easy calculation shows that $1-\zeta$ is not a root of unity (cf. the remark at the beginning of section 2). So by Kronecker's Theorem $H(1-\zeta) > 0$ and we have a contradiction.

On the other hand, carefully calculating the height shows that $h(1+\zeta) < S_0 < h(1-\zeta)$ if ζ is a primitive pth root of unity for any prime p. We will not show this inequality here. Together with Theorem 3.1 it gives another proof of the fact that Σ is infinite. Moreover, we conclude that $h(x, 1-x)$ attains a second, third, etc. maximum respectively minimum as x runs over $\mathcal{M}$. Since Theorem 3.1 is effective, one could use a machine to calculate these values.

In [**Zag93**] Zagier asked the following question on the set of small height values $h(x) + h(1-x)$ where $x \in \overline{\mathbf{Q}}$ is now unrestricted: does there exist a sequence $c_1 < c_2 < \cdots$ such that $h(x) + h(1-x) = c_i$ has finitely many algebraic solutions and $h(x) + h(1-x) > \limsup c_i$ for all other $x \in \overline{\mathbf{Q}}$? In this more general setting no answer is known.

The height used by Cohen and Zannier was $\max\{h(x), h(1-x)\}$. With Theorem 3.1 we can determine the limit points of $\Sigma' = \{\max\{h(x), h(1-x)\};\, x \in \mathcal{M}\}$:

COROLLARY 3.3. *The closure of $\Sigma' \subset \mathbf{R}$ is equal to the union of $[S_0/2, S_0]$ with a discrete set.*

We complete this section by giving a further application to Theorem 3.2. The subset of elements in $\mathcal{M}$ with degree bounded by a parameter is finite by Northcott's Theorem and Theorem 2.1. In [**Mas07**] Masser proposed the problem of counting elements in $\mathcal{M}$ of bounded degree. More precisely, given a real D and

$$\mathcal{M}(D) = \{x \in \mathcal{M};\ [\mathbf{Q}(x) : \mathbf{Q}] \leq D\},$$

can one say something about the asymptotic behavior of the cardinality $|\mathcal{M}(D)|$ as D goes to infinity?

Let ϕ be Euler's totient function and $D \geq 3$. Masser proved the bounds

$$c_0 D^3 - 200 D^2 (\log D)^3 \leq |\mathcal{M}(D)| \leq 10^{18} D^3 \frac{(\log D)^9}{(\log\log D)^6} \tag{3.1.1}$$

with

$$c_0 = \frac{6}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{\phi(k)^2}.$$

Masser obtained the lower bound in (3.1.1) by carefully counting zeros of trinomials. The upper bound comes from an explicit "relative Lehmer"-type height lower bound proved by Pontreau in [**Pon05**]. The intricate definition of c_0 has some significance: in [**Mas07**] Masser conjectured

$$\lim_{D\to\infty} \frac{|\mathcal{M}(D)|}{D^3} = c_0. \tag{3.1.2}$$

With the help of Theorem 3.2 we will prove Masser's conjecture by improving the upper bound in (3.1.1). We even give an asymptotic equality.

THEOREM 3.3. *Let $D \geq 3$, then*

$$c_0 D^3 - 200 D^2 (\log D)^3 \leq |\mathcal{M}(D)| \leq c_0 D^3 + 1000 D^2 (\log D)^3. \tag{3.1.3}$$

In particular (3.1.2) holds.

2. Factorizing trinomials over Kroneckerian fields

Let $S^1 \subset \mathbf{C}$ be the unit circle around 0. We will often use the elementary fact that if $x + y = 1$ with $x, y \in S^1$, then x and y are primitive sixth roots of unity with $x = y^{-1}$. In particular there are only two possibilities for (x, y).

LEMMA 3.1. *Let $n > m > 0$ be integers, $\xi, \zeta \in S^1$, and $P = X^n + \xi X^m + \zeta \in \mathbf{C}[X]$. Then P has only simple zeros. Furthermore, if $z \in \mathbf{C}^*$ with $P(z) = \overline{P}(z^{-1}) = 0$, then $z^{n-2m} = \xi^2 \zeta^{-1}$.*

PROOF. We begin by proving the first assertion in the lemma. Let z be a zero of P with multiplicity > 1. In this case we have $z^n + \xi z^m = -\zeta$ and $nz^n + \xi m z^m = 0$. We consider the two previous equalities as linear equations in unknowns z^n and z^m. Since $n \neq m$ we can solve for z^n and z^m:

$$z^n = \frac{m}{n-m}\zeta, \quad z^m = -\frac{n}{n-m}\frac{\zeta}{\xi}.$$

Taking absolute values and setting $\mu = m/n$ gives $|z^n| = \mu/(1-\mu)$ and $|z^m| = 1/(1-\mu)$. Hence $\mu^\mu (1-\mu)^{1-\mu} = 1$, a contradiction since $0 < \mu < 1$.

The second part of the assertion is very easy. Let $z \in \mathbf{C}^*$ with

$$z^n + \xi z^m + \zeta = 0 \text{ and } z^{-n} + \overline{\xi} z^{-m} + \overline{\zeta} = 0. \tag{3.2.1}$$

So also $z^n + \zeta\overline{\xi} z^{n-m} + \zeta = 0$. We subtract this equality from the first one in (3.2.1) to see that $z^{n-2m} = \xi^2 \zeta^{-1}$. □

Proof of Theorem 3.2: If $n = m$ or $m = 0$, then the result is immediate as in this case we have $n = 1$. So let us also assume $n > m > 0$, in particular $P(0) \neq 0$.

Let $P = AB$ be a decomposition as in the statement of Theorem 3.2; we may assume that A and B are monic. Say $B(\eta) = 0$, then η is a root of unity and $P(\eta) = 0$. Hence $-\zeta^{-1}\eta^n + (-\zeta^{-1}\xi\eta^m) = 1$. We conclude that $-\zeta^{-1}\eta^n = x$, $-\zeta^{-1}\xi\eta^m = x^{-1}$ where x is a primitive 6th root of unity. Since n and m are coprime there exist $a, b \in \mathbf{Z}$ with $am + bn = 1$. Hence $x^{b-a} = (-1)^{a+b}\zeta^{-a-b}\xi^a\eta$. Since there are only 2 possibilities for x, there are at most 2 such η. By Lemma 3.1 we conclude $\deg B \leq 2$. The first part of the assertion follows.

Now let $A = A_1 \cdots A_g$ where the $A_i \in K[X]$ are irreducible and monic. By hypothesis $A \notin K$, so $g \geq 1$.

We claim that the A_i are not self-inverse, i.e. $X^{\deg A_i}\overline{A_i}(X^{-1})A_i(X)^{-1}$ is not a constant. Indeed let us assume the contrary. If $A_i(z) = 0$, then $z \neq 0$ and $P(z) = \overline{P}(z^{-1}) = 0$. By Lemma 3.1 we have $z^{n-2m} = \xi^2\zeta^{-1}$. Since z is not a root of unity we have $n = 2m$, so $n = 2$ and $m = 1$. Hence $\xi^2 = \zeta$ and so $P = X^2 + \xi X + \xi^2$. But now all zeros of P are roots of unity, contradicting our assumption $A \notin K$. Therefore none of the A_i are self-inverse.

By Gauss's Lemma we have $A_i \in \mathcal{O}_K[X]$, and since the A_i are monic we have $N(\mathrm{cont}(A_i)) = 1$. By Theorem 1.2 we get the lower bound

$$\prod_\sigma M(\sigma(A_i)) \geq \left(\frac{\sqrt{5}+1}{2}\right)^{[K:\mathbf{Q}]/2}$$

where the product runs over all embeddings $\sigma : K \to \mathbf{C}$ extended to $K[X]$ in the usual manner. The Mahler measure is multiplicative and we have $M(\sigma(B)) = 1$ by construction, therefore

$$\prod_\sigma M(\sigma(P)) = \prod_{i=1}^{g}\prod_\sigma M(\sigma(A_i)) \geq \left(\frac{\sqrt{5}+1}{2}\right)^{g[K:\mathbf{Q}]/2}. \tag{3.2.2}$$

Let us assume for the moment that $P(X)\overline{P}(X^{-1})$ has strictly more than 3 non-zero terms. Let $\sigma : K \to \mathbf{C}$ be an embedding, we will bound $M(\sigma(P))$ from above. Since $\sigma(P)$ has three coefficients of modulus 1 Theorem 1.1 in chapter 1 implies $M(\sigma(P)) < \frac{\sqrt{5}+1}{2}$. If we apply this last inequality to bound the left side of (3.2.2) we conclude $g < 2$, hence $A = A_1$ is irreducible.

But what if

$$\begin{aligned} &P(X)\overline{P}(X^{-1}) \\ &\quad = \zeta^{-1}X^n + \zeta X^{-n} + \xi^{-1}X^{n-m} + \xi X^{m-n} + \xi\zeta^{-1}X^m + \xi^{-1}\zeta X^{-m} + 3 \end{aligned}$$

has at most 3 non-zero terms? Since $n > m > 0$ are coprime we must have $n = 2$, $m = 1$, and $\xi^2 = -\zeta$. Therefore $P = X^2 + \xi X - \xi^2$, and this polynomial has no roots of unity as zeros, so $P = A$. As already pointed out in section 1 the polynomial A is

reducible K if and only if $\sqrt{5} \in K$, but this is just the case we are excluding in the assertion. □

3. A reduction

We define a norm $||\cdot||$ on $\mathbf{R}^2$ which will be used in certain places of this chapter. Let $(r,s) \in \mathbf{R}^2$, we set

$$||(r,s)|| = \begin{cases} |r|+|s| & \text{if } rs \geq 0, \\ \max\{|r|,|s|\} & \text{if } rs < 0. \end{cases}$$

The following lemma will be helpful throughout the whole chapter, it is similar to Lemma 2.3.

LEMMA 3.2. *Let $x \in \mathcal{M}$. There exist a root of unity ζ, integers k,l,m,n, and $\alpha \in \overline{\mathbf{Q}}$ such that*

$$n \geq 2m \geq 0, \qquad (n-m)k - nl = 1, \quad h(x,1-x) = nh(\alpha),$$
$$\alpha^n - \zeta^l \alpha^m + \zeta^k = 0, \quad \textit{and} \quad \mathbf{Q}(x) = \mathbf{Q}(\alpha,\zeta).$$

Furthermore, there exist coprime integers r and s such that $x^r(1-x)^s = \pm\zeta$ and $||(r,s)|| = n$.

PROOF. The lemma is simple if x and $1-x$ are both roots of unity. Indeed in this case x is a primitive 6th root of unity and $x + x^{-1} = 1$. We set $\alpha = x$, $\zeta = x^{-1}$, $n = 1$, $m = 0$, $k = 1$, $l = 0$, $r = 0$, and $s = 1$.

Let $\mathcal{M}_0 = \mathcal{M} \backslash \{e^{\pm 2\pi i/6}\}$, then for any $x \in \mathcal{M}_0$ there is a unique $\lambda(x) \in \mathbf{P}^1(\mathbf{Q})$ such that if $\lambda(x) = [r:s]$ with $r,s \in \mathbf{Z}$ coprime then $x^r(1-x)^s$ is a root of unity. Since the pair (r,s) is uniquely determined up to sign it makes sense to speak of $||\lambda(x)||$. The two maps $\phi_i : \mathcal{M}_0 \to \mathcal{M}_0$ given by $\phi_1(x) = 1/x$, $\phi_2(x) = 1-x$ generate the symmetric group S_3 which acts on $\mathcal{M}_0$. These actions leave $h(x,1-x)$ invariant by the product formula. Furthermore, we have $\mathbf{Q}(\phi_i(x)) = \mathbf{Q}(x)$. We check that if $\lambda(x) = [r:s]$ with coprime $r,s \in \mathbf{Z}$, then $\lambda(\phi_1(x)) = [r+s:-s]$ and $\lambda(\phi_2(x)) = [s:r]$. So we have an action of S_3 on $\mathbf{P}^1(\mathbf{R})$. Let $w \in \mathbf{R}\backslash\{-1,0\}$, then the orbit of $[1:w]$ equals

$$\left\{ [1:w], [1:-\frac{w}{1+w}], [1:-1-\frac{1}{w}], [1:-1-w], [1:-\frac{1}{1+w}], [1:\frac{1}{w}] \right\}.$$

It is easily verified that

$$\inf_{w \in \mathbf{R}\backslash\{-1,0\}} \max\{w, -\frac{w}{1+w}, -1-\frac{1}{w}, -1-w, -\frac{1}{1+w}, \frac{1}{w}\} \geq 1.$$

Since the orbit of $[0:1]$ contains $[1:-1]$ and $[1:0]$ we conclude that any point in $\mathbf{P}^1(\mathbf{R})$ is in the orbit of some element of $F = \{[1:w];\ w \geq 1 \text{ in } \mathbf{R}\} \cup \{[0:1]\}$. Finally it is simple to show $||\lambda(x)|| = ||\lambda(\phi_i(x))||$.

Let $x \in \mathcal{M}_0$, by letting S_3 act on x we assume $\lambda(x) = [r:s] \in F$ with $r,s \in \mathbf{Z}$ coprime and $r \geq 0$, therefore $s \geq r$ and $s > 0$. So $x^r(1-x)^s = \zeta$ for some root of unity ζ. We choose integers a,b with $ar + bs = 1$ and set $\alpha = x^b(1-x)^{-a}$. Then $\alpha^s = x\zeta^{-a}$

and $\alpha^{-r} = \zeta^{-b}(1-x)$. Clearly we have $\mathbf{Q}(x) = \mathbf{Q}(\alpha, \zeta)$ and $\alpha^{r+s} - \zeta^{-a}\alpha^r + \zeta^{b-a} = 0$. We choose $n = r+s$, $m = r$, $k = b-a$, and $l = -a$. By definition $||(r,s)|| = r+s = n$. It remains to show the height equality in the assertion. We note $h(x, 1-x) = h(x^{-1}, x^{-1}-1)$ by the product formula. We have $\zeta x^{-r-s} = (x^{-1}-1)^s$, so by local considerations $h(x^{-1}, x^{-1}-1) = h(x^{-1}-1) = (1+r/s)h(x) = (r+s)h(\alpha) = nh(\alpha)$. □

We apply Theorem 3.2 to study irreducible factors of the trinomial in the previous lemma. If ζ is a root of unity we let $\operatorname{ord}\zeta$ denote its order.

LEMMA 3.3. *Let $x \in \mathcal{M}$ and $h(x, 1-x) > 0$. In the situation of Lemma 3.2 set $K = \mathbf{Q}(\zeta)$ and $R = X^n - \zeta^l X^m + \zeta^k$. There exist $A, B \in K[X]$ with A irreducible, the zeros of B are precisely the roots of unity that are zeros of P, and $\deg B \leq 2$. Furthermore, if $\operatorname{ord}\zeta \nmid 6$ then $B \in K$.*

PROOF. Let $R = AB$ where the zeros of B are precisely the roots of unity that are zeros of R. We would like to apply Theorem 3.2. We first note that $A \notin K$ since $A \neq 0$ and $A(\alpha) = 0$ where α is no root of unity. Next we dismiss the case where R is one of the exceptions described in the statement of Theorem 3.2. What happens if $n = 2$, $m = 1$, and $\zeta^{2l} = -\zeta^k$? Since $(n-m)k - nl = 1$ we conclude $k - 2l = 1$ and so $\zeta = -1$. But then $K = \mathbf{Q}$. Theorem 3.2 applies in our situation since $\sqrt{5} \notin \mathbf{Q}$. Hence $\deg B \leq 2$ and $A \in K[X]$ is irreducible.

Let us assume that R has a zero η which is a root of unity. Then $-\zeta^{-k}\eta^n + \zeta^{l-k}\eta^m = 1$. Hence $(\zeta^{-k}\eta^n)^6 = (\zeta^{l-k}\eta^m)^6 = 1$. So

$$1 = (\zeta^{-k}\eta^n)^{-6m}(\zeta^{l-k}\eta^m)^{6n} = \zeta^{6(km+ln-kn)} = \zeta^{-6}.$$

Therefore if $\operatorname{ord}\zeta \nmid 6$, then B has no zeros, hence $B \in K$. □

4. Averaging the Mahler measure

Let $\theta \in \mathbf{R}$ and let $n \geq m \geq 0$ be integers with $n \neq 0$. We introduce the following notation

$$\begin{aligned} P(X,\theta) = P_{nm}(X,\theta) &= X^n - e^{-2\pi i\theta}X^m + 1, \\ Q(X,\theta) = Q_{nm}(X,\theta) &= (X^n - e^{-2\pi i\theta}X^m + 1)(X^{-n} - e^{2\pi i\theta}X^{-m} + 1) \\ f(\theta) = f_{nm}(\theta) &= \int_0^1 \log|P_{nm}(e^{2\pi iu}, \theta)|du. \end{aligned} \tag{3.4.1}$$

Clearly f satisfies $f(\theta+1) = f(\theta)$. This function equals $\log M(P_{nm}(X,\theta))$ and is therefore continuous by a classical result of Mahler. If $\zeta \in S^1$, then $Q_{nm}(\zeta,\theta) = |P_{nm}(\zeta,\theta)|^2 \geq 0$. So for example

$$f_{nm}(\theta) = \frac{1}{2}\int_0^1 \log Q_{nm}(e^{2\pi iu}, \theta)du. \tag{3.4.2}$$

We will often omit the indices n and m to ease notation.

LEMMA 3.4. *Let a, d be positive integers and $\zeta = e^{2\pi i a/d}$, furthermore let k, l, m, n be integers with $n \geq m \geq 0$ and $(n-m)k - nl = 1$. Then*

$$\log M(X^n - \zeta^l X^m + \zeta^k) = f_{nm}(\frac{a}{dn}).$$

PROOF. We have

$$\log M(X^n - \zeta^l X^m + \zeta^k) = \int_0^1 \log |e^{2\pi i(nu - \frac{ka}{d})} - e^{2\pi i(mu + \frac{a}{d}(l-k))} + 1| du.$$

Of course $n > 0$, so the change of variables $u' = u - \frac{ak}{dn}$ leads to

$$\log M(X^n - \zeta^l X^m + \zeta^k) = \int_{-\frac{ak}{dn}}^{1-\frac{ak}{dn}} \log |e^{2\pi i n u'} - e^{2\pi i(mu' - \frac{a}{dn})} + 1| du'.$$

Since the integrand above is invariant under $u' \mapsto u' + 1$ we may replace the integration limits by 0 and 1. This completes the proof. □

LEMMA 3.5. *Let $x \in \mathcal{M}$ and $h(x, 1-x) > 0$. In the situation of Lemma 3.2 and with $d = \operatorname{ord} \zeta$ we have*

$$h(x, 1-x) = \frac{n}{[\mathbf{Q}(\zeta, \alpha) : \mathbf{Q}]} \sum_{1 \leq a \leq d,\, (a,d)=1} f_{nm}(\frac{a}{dn}). \tag{3.4.3}$$

Furthermore, $n - 2 \leq [\mathbf{Q}(\zeta, \alpha) : \mathbf{Q}(\zeta)] \leq n$ in general and $[\mathbf{Q}(\zeta, \alpha) : \mathbf{Q}(\zeta)] = n$ if $d \nmid 6$. Finally $\phi(d) \leq [\mathbf{Q}(x) : \mathbf{Q}]$.

PROOF. By Lemma 3.2 it suffices to show that $nh(\alpha)$ is equal to the right side of (3.4.3). Let R be the polynomial given implicitly in the assertion of Lemma 3.2, i.e. $R = X^n - \zeta^l X^m + \zeta^k \in K[X]$ with $K = \mathbf{Q}(\zeta)$; we have $R(\alpha) = 0$. Also $n \geq 2m$ and so $n > m$ since n and m cannot be both zero. In particular α is an algebraic integer.

By Lemma 3.3 we may factor $R = AB$ with $A, B \in K[X]$, where A is irreducible and monic, and such that the zeros of B are precisely the roots of unity that are zeros of R. We note that B is also monic. Since α is not a root of unity we conclude $A(\alpha) = 0$. Therefore A is the minimal polynomial of α over K. So $A' = \prod_\sigma \sigma(A)$ is the minimal polynomial of α over $\mathbf{Q}$, here σ runs over all embeddings $K \to \mathbf{C}$. Since A' is monic and α is an algebraic integer we conclude $A' \in \mathbf{Z}[X]$ and therefore by (1.2.3)

$$h(\alpha) = \frac{1}{[\mathbf{Q}(x) : \mathbf{Q}]} \log M(A') = \frac{1}{[\mathbf{Q}(x) : \mathbf{Q}]} \sum_\sigma \log M(\sigma(A)). \tag{3.4.4}$$

The zeros of the monic polynomial B are roots of unity and the Mahler measure is multiplicative, so we may replace A by R in (3.4.4).

Equality (3.4.3) follows immediately from (3.4.4), Lemma 3.4, and $\mathbf{Q}(x) = \mathbf{Q}(\zeta, \alpha)$. The assertions regarding $[\mathbf{Q}(\zeta, \alpha) : \mathbf{Q}(\zeta)]$ follow from Lemma 3.3. Since $\phi(d) = [\mathbf{Q}(\zeta) : \mathbf{Q}]$ and $\mathbf{Q}(\zeta) \subset \mathbf{Q}(x)$ we conclude $\phi(d) \leq [\mathbf{Q}(x) : \mathbf{Q}]$. □

5. Bounding the variation

Throughout this section we use the notation from Lemma 3.2, i.e. $n \geq 2m \geq 0$ are coprime integers and $(n-m)k - nl = 1$ for integers k, l. These properties imply $n > 0$ so we may define $\mu = m/n \in [0, 1/2]$. Furthermore, we keep the notation introduced around (3.4.1)

We define $E = E_{nm}$ to be the set of all $\theta \in [0,1]$ such that $Q_{nm}(\zeta, \theta) = 0$ for some $\zeta \in S^1$. Equivalently, E_{nm} is the set of $\theta \in [0,1]$ with $P_{nm}(\zeta, \theta) = 0$ for some $\zeta \in S^1$.

LEMMA 3.6. *The set E_{nm} is finite.*

PROOF. We assume that $\zeta \in S^1$ satisfies $P(\zeta, \theta) = 0$. Then $-\zeta^n + e^{-2\pi i\theta}\zeta^m = 1$ and so

$$\zeta^{6n} = (e^{-2\pi i\theta}\zeta^m)^6 = 1.$$

We divide the mth power of the first expression by the nth power of the second expression and conclude $e^{12\pi i\theta n} = 1$. Since $n \neq 0$ there are only finitely many $\theta \in [0,1]$ that satisfy this last inequality. □

LEMMA 3.7. *Let $\theta \in [0,1]\backslash E_{nm}$ and say $m > 0$. We consider $Q_{nm}(X, \theta)$ as a Laurent polynomial in X. If $z \neq 0$ is a zero of $Q_{nm}(X, \theta)$, then it is a simple zero, and furthermore $\frac{\partial}{\partial\theta} Q_{nm}(z, \theta) \neq 0$.*

PROOF. For brevity we set $\xi = e^{-2\pi i\theta}$ with $\theta \in [0,1]\backslash E$. By (3.4.1) we have

$$Q(X, \theta) = P(X, \theta)P(X^{-1}, -\theta), \tag{3.5.1}$$

$$\frac{\partial}{\partial X} Q(X, \theta) = \frac{\partial P}{\partial X}(X, \theta)P(X^{-1}, -\theta) - P(X, \theta)\frac{1}{X^2}\frac{\partial P}{\partial X}(X^{-1}, -\theta). \tag{3.5.2}$$

We prove the first part by contradiction. So assume $z \neq 0$ is a zero of $Q(X, \theta)$ of multiplicity > 1, hence z is a zero of both (3.5.1) and (3.5.2). By our hypothesis on θ we have $|z| \neq 1$.

Let us assume first that $P(z, \theta) = 0$. By Lemma 3.1 $\frac{\partial}{\partial X} P(X, \theta)$ is non-zero at z. We get $P(z^{-1}, -\theta) = 0$ by (3.5.2). We apply the second statement in Lemma 3.1 and $|z| \neq 1$ to conclude that $n = 2m$ and $\xi^2 = 1$. Thus $z^{2m} - \xi z^m + \xi^2 = 0$, hence z^m/ξ is a root of unity, a contradiction.

The case $P(z^{-1}, -\theta) = 0$ is similar and also leads to a contradiction.

We turn to the second part of the lemma. We expand Q and its derivative:

$$Q = X^n + X^{-n} - \xi^{-1}X^{n-m} - \xi X^{m-n} - \xi X^m - \xi^{-1}X^{-m} + 3, \tag{3.5.3}$$

$$\frac{\partial Q}{\partial \theta} = 2\pi i(\xi X^{m-n} - \xi^{-1}X^{n-m} + \xi X^m - \xi^{-1}X^{-m}). \tag{3.5.4}$$

If $z \neq 0$ is a common zero of $Q(X, \theta)$ and $\frac{\partial}{\partial\theta} Q(X, \theta)$, we will derive a contradiction. We define $a = z^n$ and $b = -\xi z^m$. Then by (3.5.3) and (3.5.4) we have the identities $a + a^{-1} + ab^{-1} + a^{-1}b + b + b^{-1} + 3 = 0$ and $ab^{-1} - a^{-1}b - b + b^{-1} = 0$. We add the two to obtain

$$a + a^{-1} + 2ab^{-1} + 2b^{-1} + 3 = 0. \tag{3.5.5}$$

By (3.4.1) we have either $a+b+1=0$ or $a^{-1}+b^{-1}+1=0$. In both cases we can use (3.5.5) to deduce that a is a root of unity. Therefore so is z, this contradicts our choice $\theta \notin E$. □

We define

$$T_{nm}(X,\theta) = (\frac{\partial Q_{nm}}{\partial \theta}/Q_{nm})(X,\theta).$$

LEMMA 3.8. *Let $\theta \in [0,1]\backslash E_{nm}$ and say $m>0$. If $P_{nm}(z,\theta)=0$ then $z\neq 0$, $1+z^{-n}\neq 0$, $\mu-(1+z^{-n})^{-1}\neq 0$, and the residue satisfies*

$$\mathrm{res}_{X=z}\frac{T_{nm}(X,\theta)}{X} = -\frac{2\pi i}{n}\left(\mu - \frac{1}{1+z^{-n}}\right)^{-1}.$$

PROOF. Say $P(z,\theta)=0$, then $z\neq 0$ since $n>m>0$. We set $a=z^n$, $\xi = e^{-2\pi i\theta}$, and $b=-\xi z^m$, hence $a+b=-1$. The second non-vanishing statement follows since $1+z^{-n}=1+a^{-1}=-ba^{-1}$.

Lemma 3.7 implies that the non-zero poles of $T(X,\theta)/X$ are precisely the zeros of $Q(X,\theta)$ and that these poles are all simple. Thus the residue of $T(X,\theta)/X$ at z is given by

$$\mathrm{res}_{X=z}\frac{T(X,\theta)}{X} = \frac{1}{z}\frac{\frac{\partial}{\partial\theta}Q(X,\theta)}{\frac{\partial}{\partial X}Q(X,\theta)}.$$

With this equality, (3.5.3), and (3.5.4) we evaluate

$$\begin{aligned}(3.5.6)\qquad &\mathrm{res}_{X=z}\frac{T(X,\theta)}{X}\\ &= -2\pi i\frac{a^{-1}b-ab^{-1}+b-b^{-1}}{na-na^{-1}+(n-m)ab^{-1}-(n-m)a^{-1}b+mb-mb^{-1}}\\ &= -\frac{2\pi i}{n}\frac{b^2-a^2+ab^2-a}{a^2b-b+a^2-a^2\mu-b^2+b^2\mu+ab^2\mu-a\mu}\\ &= -\frac{2\pi i}{n}\frac{(a+1)(a^2+a+1)}{(\mu(a+1)-a)(a^2+a+1)}\\ &= -\frac{2\pi i}{n}\left(\mu-\frac{1}{1+a^{-1}}\right)^{-1},\end{aligned}$$

here the second to last equality follows by using $a+b=-1$. This calculation also implies the last non-vanishing statement in the assertion. □

The next Lemma is very similar to Lemma 3.8

LEMMA 3.9. *Let $\theta \in [0,1]\backslash E_{nm}$ and say $m>0$. If $P_{nm}(z,\theta)=0$ then $\overline{z}\neq 0$, $1+\overline{z}^{-n}\neq 0$, $\mu-(1+\overline{z}^{-n})^{-1}\neq 0$, and the residue satisfies*

$$\mathrm{res}_{X=\overline{z}^{-1}}\frac{T_{nm}(X,\theta)}{X} = -\frac{2\pi i}{n}\left(\mu-\frac{1}{1+\overline{z}^{-n}}\right)^{-1}.$$

PROOF. The proof follows the lines of the proof of Lemma 3.8: we set $a = \overline{z}^{-n}$ and $b = -e^{-2\pi i\theta}\overline{z}^{-m}$. This time the relation is $a+b+ab = 0$. The proof follows by applying this equation to the second to last equality in (3.5.6). □

For each $z \in \mathbf{C}$ we define a map $\sigma_z : \mathbf{C} \to \mathbf{C}$: if $|z| < 1$ we set $\sigma_z(w) = w$ for all $w \in \mathbf{C}$, if $|z| \geq 1$ we set $\sigma_z(w) = \overline{w}$ for all $w \in \mathbf{C}$.

LEMMA 3.10. *The map $\theta \mapsto f_{nm}(\theta)$ is continuous on $[0,1]$ and differentiable on $[0,1]\backslash E_{nm}$. If $\theta \in [0,1]\backslash E_{nm}$ and $m > 0$ then*

$$f'_{nm}(\theta) = -\frac{\pi i}{n} \sum_{z,\, P_{nm}(z,\theta)=0} \sigma_z\left(\mu - \frac{1}{1+z^{-n}}\right)^{-1}. \tag{3.5.7}$$

PROOF. The continuity of f was already mentioned at the beginning of section 4. Now say $\theta \in [0,1]\backslash E$. By Lemma 3.6 $Q(e^{2\pi iu}, \theta) > 0$ for all $u \in \mathbf{R}$. By compactness of S^1 we have $Q(e^{2\pi iu}, \theta') > 0$ for all θ' in some neighborhood of θ and all $u \in \mathbf{R}$. Hence f is differentiable at θ by (3.4.2).

Now let us assume $m > 0$. By the discussion in the last paragraph we may exchange integration and differentiation to obtain

$$\begin{aligned} \frac{df}{d\theta} &= \frac{1}{2}\frac{d}{d\theta}\int_0^1 \log Q(e^{2\pi iu}, \theta)du = \frac{1}{2}\int_0^1 \frac{\partial}{\partial\theta}\log Q(e^{2\pi iu}, \theta)du \\ &= \frac{1}{2}\int_0^1 T(e^{2\pi iu}, \theta)du. \end{aligned} \tag{3.5.8}$$

We can rewrite (3.5.8) as a complex integral

$$\frac{df}{d\theta} = \frac{1}{4\pi i}\int_{|z|=1} \frac{T(z,\theta)}{z}dz, \tag{3.5.9}$$

the path being understood as counterclockwise. The integrand has no pole on $|z| = 1$ because $\theta \notin E$. This meromorphic function also has no pole at $z = 0$ by (3.5.3), (3.5.4), and since $m \geq 1$.

As already stated in the proof of Lemma 3.8, the non-zero poles of $T(X,\theta)/X$ are precisely the zeros of $Q(X,\theta)$. We apply the Residue Theorem to (3.5.9) to get

$$\frac{df}{d\theta} = \frac{1}{2}\sum_{\substack{Q(z,\theta)=0 \\ 0<|z|<1}} \mathrm{res}_{X=z}\frac{T(X,\theta)}{X}. \tag{3.5.10}$$

Let us now assume $Q(z,\theta) = 0$ and $0 < |z| < 1$. By (3.4.1) either $P(z,\theta) = 0$ or $P(z^{-1}, -\theta) = 0$. These two cases are mutually exclusive since Q has only simple zeros by Lemma 3.7. We note that in the second case $P(\overline{z}^{-1}, \theta) = 0$ and $|\overline{z}^{-1}| > 1$. To determine the residue at z we apply Lemma 3.8 or 3.9 depending on whether $P(z,\theta) = 0$ or $P(\overline{z}^{-1}, \theta) = 0$. Hence each term in (3.5.10) corresponds to exactly one term in the sum (3.5.7).

On the other hand, if $z \neq 0$ is a zero of $P(X,\theta)$, then $Q(z,\theta)$ and $Q(\overline{z}^{-1}, \theta)$ both vanish. Since $\theta \notin E$ one and only one of the values $|z|$, $|\overline{z}^{-1}|$ is strictly less than 1. As

before any term in (3.5.7) corresponds to exactly one term in (3.5.10) by Lemma 3.8 or 3.9. □

LEMMA 3.11. *If $z \in \mathbf{C}^*$, $\theta \in \mathbf{R}$, and $\eta^n = 1$ with $P_{nm}(z,\theta) = P_{nm}(\eta z,\theta)$, then $\eta = 1$.*

PROOF. Say $\xi = e^{-2\pi i\theta}$, then $z^n - \xi z^m + 1 = z^n - \xi\eta^m z^m + 1$. So $\eta^m = 1$ and now $n(k-l) - mk = 1$ implies $\eta = 1$. □

Unfortunately, the absolute value of a term in the sum (3.5.7) cannot in general be bounded from above independently of n. Nevertheless we prove a vanishing result which will enable us to get such a bound for $|f'_{nm}(\theta)|$:

LEMMA 3.12. *Let $\theta \in [0,1]\backslash E_{nm}$ and $m > 0$, then*

$$\sum_{z,\, P_{nm}(z,\theta)=0} \left(\mu - \frac{1}{1+z^{-n}}\right)^{-1} = 0. \tag{3.5.11}$$

PROOF. By Lemma 3.1 the polynomial $P(X,\theta)$ has exactly n distinct zeros for fixed θ. As z runs over these zeros, the values z^n are pairwise distinct by Lemma 3.11 and thus so are

$$\left(\mu - \frac{1}{1+z^{-n}}\right)^{-1}. \tag{3.5.12}$$

For fixed $\xi = e^{-2\pi i\theta}$ we consider the non-zero polynomial

$$U = X^n - \xi^n(\mu X - 1)^m((1-\mu)X + 1)^{n-m} \in \mathbf{C}[X]$$

of degree at most n. If $P(z,\theta) = 0$, then a calculation using $z^n + 1 = \xi z^m$ shows that (3.5.12) is a zero of U. Since there are n distinct values (3.5.12), these must be precisely the zeros of U. In particular U has degree n. It follows that the sum in the assertion is a multiple of the coefficient of X^{n-1} in U. But this coefficient is

$$-\xi^n(\mu^m(1-\mu)^{n-m-1}(n-m) - m\mu^{m-1}(1-\mu)^{n-m}) = 0$$

since $\mu = m/n$. □

LEMMA 3.13. *Let $\theta \in [0,1]\backslash E_{nm}$, then $|f'_{nm}(\theta)| < 4\pi/\sqrt{3}$.*

PROOF. We eliminate the case $m = 0$ first. Indeed then automatically $n = 1$ and $P = X - e^{-2\pi i\theta} + 1$. Thus $f(\theta) = \log\max\{1, |e^{-2\pi i\theta} - 1|\}$ for all θ. The function f is differentiable on $[0,1]\backslash\{1/6, 5/6\}$. Using calculus one can show that the derivative of f is at most $\sqrt{3}\pi$ for $1/6 < \theta < 5/6$.

Let us move on to the case $m > 0$. We add $\pi i/n$ times the expression on the left of (3.5.11) to $f'(\theta)$ as in Lemma 3.10 to deduce

$$f'(\theta) = \frac{\pi i}{n} \sum_{\substack{z,\, P(z,\theta)=0 \\ |z|>1}} \left(\left(\mu - \frac{1}{1+z^{-n}}\right)^{-1} - \overline{\left(\mu - \frac{1}{1+z^{-n}}\right)^{-1}}\right). \tag{3.5.13}$$

We continue by bounding the absolute value of (3.5.13) term-wise.

Let $P(z,\theta)=0$, $|z|>1$, and let us define $\delta=\mu-\frac{1}{1+z^{-n}}$. We have $z^n=\frac{\mu-\delta}{1-\mu+\delta}$ and $z^n+1=\frac{1}{1-\mu+\delta}=\xi z^m$ where $\xi=e^{-2\pi i\theta}$. So $|1-\mu+\delta|<1$ and

$$|\mu-\delta|^m=|1-\mu+\delta|^{m-n}, \tag{3.5.14}$$

thus $|\mu-\delta|>1$. We apply the parallelogram equality $|x+y|^2+|x-y|^2=2|x|^2+2|y|^2$ $(x,y\in\mathbf{C})$ with $x=1/2$ and $y=1/2-\mu+\delta$ to obtain

$$|1-\mu+\delta|^2+|\mu-\delta|^2=\frac{1}{2}+2\left|\frac{1}{2}-\mu+\delta\right|^2=1+2|\mu-\delta|^2+2(\mathrm{Re}\delta)-2\mu. \tag{3.5.15}$$

We define $w_0=|\mu-\delta|^2>1$. From (3.5.14) we deduce $|1-\mu+\delta|^2=w_0^{-\mu/(1-\mu)}$. We insert this equality into (3.5.15) to obtain

$$2\mathrm{Re}\delta=w_0^{-\mu/(1-\mu)}-w_0+2\mu-1. \tag{3.5.16}$$

On the other hand $w_0=|\delta|^2-2\mu\mathrm{Re}\delta+\mu^2$, hence together with (3.5.16) we get $k(w_0)=0$ where

$$k(w)=\mu w^{-\mu/(1-\mu)}+(1-\mu)w-|\delta|^2+\mu^2-\mu,\quad w\in(0,\infty).$$

Elementary calculus shows that k has a unique minimum at

$$\left(\frac{\mu}{1-\mu}\right)^{2(1-\mu)}\leq 1.$$

Hence k is strictly increasing on $[1,\infty)$. Since $w_0>1$ we deduce

$$0=k(w_0)>k(1)=\mu^2-\mu+1-|\delta|^2,$$

and so $|\delta|^2>\mu^2-\mu+1$. But the right-hand side of this inequality is $\geq 3/4$ since $\mu\in(0,1/2]$. We conclude

$$|\delta|>\sqrt{3}/2. \tag{3.5.17}$$

Back to our derivative. Each term in (3.5.13) is of the form $\delta^{-1}-\overline{\delta^{-1}}$ with δ as in the last paragraph. Thus we apply the triangle inequality and (3.5.17) to (3.5.13) to complete the proof. □

Finally we have:

LEMMA 3.14. *The map $\theta\mapsto f_{nm}(\theta/n)$ is of bounded variation on $[0,1)$ with total variation bounded by $\frac{4\pi}{\sqrt{3}n}$.*

PROOF. If S is a real interval $[a,b]$ with a finite set of points removed, and F is a continuous real-valued function on $[a,b]$ and differentiable on S with $|F'|\leq M$, then it is well-known that the total variation of F is at most $M(b-a)$. So the result follows from Lemma 3.13 with $F(\theta)=f(\theta/n)$ and $M=\frac{4\pi}{\sqrt{3}n}$. □

6. Bounding the discrepancy

Recall that ϕ is Euler's totient function. Let d be a positive integer and let $\{x_k\}$ denote the sequence of numbers a/d where a runs over the $\phi(d)$ integers in $[1, d]$ that are coprime to d. We set $\mathcal{D}(d)$ to be the discrepancy of this sequence as defined in [**Har98**]. That is

$$\mathcal{D}(d) = \sup_{I \subset [0,1)} \left| \Big(\sum_{\substack{1 \le a \le d,\ (a,d)=1 \\ a/d \in I}} 1 \Big) - \phi(d)|I| \right|. \tag{3.6.1}$$

Above I runs over all open, closed, and half-open intervals of length $|I|$. Let $h_d = 1 + \frac{1}{2} + \cdots + \frac{1}{d}$ be the dth harmonic number and let $\omega(d)$ denote the number of distinct prime factors of d. We use a Theorem of Erdös and Turán to show:

Lemma 3.15. *There exists an absolute constant c such that*

$$\mathcal{D}(d) \le c \frac{h_d \phi(d)}{d} \prod_{\substack{p \mid d \\ prime}} \Big(1 + \frac{p}{p-1}\Big) \le c 2^{\omega(d)} h_d \quad \textit{for all} \quad d \ge 1. \tag{3.6.2}$$

The choice $c = 3 + 2/\pi$ is possible.

Proof. The second inequality in (3.6.2) follows from

$$\phi(d) \prod_{\substack{p \mid d \\ \text{prime}}} \Big(1 + \frac{p}{p-1}\Big) = d \prod_{\substack{p \mid d \\ \text{prime}}} \Big(1 + \frac{p-1}{p}\Big) \le d 2^{\omega(d)}.$$

We proceed to prove the first inequality in (3.6.2). The Erdös-Turán Theorem (Theorem 5.5 in [**Har98**] page 129) applied with $L = d$ and $N = \phi(d)$ gives the upper bound

$$\frac{\mathcal{D}(d)}{\phi(d)} \le \frac{1}{d+1} + 2(1 + 1/\pi) \sum_{m=1}^{d} \frac{1}{m\phi(d)} \left| \sum_{k=1}^{\phi(d)} \exp(2\pi i m x_k) \right|. \tag{3.6.3}$$

We start by handling the Ramanujan sum

$$s(m) = \sum_{k=1}^{\phi(d)} \exp(2\pi i m x_k) = \sum_{1 \le a \le d,\ (a,d)=1} \exp(2\pi i m a/d).$$

We use Theorem 272 of [**HW05**] to evaluate

$$s(m) = \frac{\mu(d/(d,m))}{\phi(d/(d,m))} \phi(d)$$

where μ is the Möbius function. We rearrange the second term on the right of (3.6.3) and obtain

$$\begin{aligned}\sum_{m=1}^{d}\frac{1}{m\phi(d)}\left|\sum_{k=1}^{\phi(d)}\exp(2\pi imx_k)\right| &= \sum_{m=1}^{d}\frac{|\mu(d/(d,m))|}{m\phi(d/(d,m))}\\ &= \sum_{e|d}\sum_{\substack{1\le m\le d\\(d,m)=e}}\frac{|\mu(d/e)|}{m\phi(d/e)}\\ &= \sum_{e|d}\frac{|\mu(d/e)|}{e\phi(d/e)}\sum_{\substack{1\le m\le d\\(d,m)=e,\ m'=m/e}}\frac{1}{m'}\\ &\le h_d\sum_{e|d}\frac{|\mu(d/e)|}{e\phi(d/e)}.\end{aligned}$$

The arithmetic function $g(a)=\sum_{e|a}\frac{|\mu(a/e)|}{e\phi(a/e)}$ is multiplicative. If p is a prime and f is a positive integer, then $g(p^f)=p^{-f}(1+p/(p-1))$. Hence

$$g(a)=\frac{1}{a}\prod_{p|a}(1+\frac{p}{p-1})\ge\frac{1}{a}.$$

The lemma now follows easily from $\mathcal{D}(d)/\phi(d)\le\frac{1}{d+1}+2(1+1/\pi)g(d)h_d$. □

It is well-known that $2^{\omega(n)}=O(n^\epsilon)$ and $n=O(\phi(n)^{1+\epsilon})$ for every $\epsilon>0$. Hence the expression $\mathcal{D}(d)/\phi(d)$ is $O(d^{-1+\epsilon})$ and especially $\mathcal{D}(d)/\phi(d)\to 0$ as $d\to\infty$.

7. Proof of Theorem 3.1 and Corollary 3.3

LEMMA 3.16. *Let m,n be coprime integers with $n\ge 1$, we set*

$$N=\begin{bmatrix} n & 0\\ m & -1/n\end{bmatrix}$$

and let $F\subset\mathbf{R}^2$ denote the image of $[0,1)^2$ under N. Then

(i) *if $x\in\mathbf{R}^2$ there exists $y\in\mathbf{Z}^2$ with $x-y\in F$,*
(ii) *if $x',x''\in F$ with $x'-x''\in\mathbf{Z}^2$ then $x'=x''$.*

PROOF. Let $x=(x_1,x_2)^{\mathrm{t}}\in\mathbf{R}^2$. Since m and n are coprime we have

$$\{\frac{mk}{n}\mod 1;\ 1\le k\le n\}=\{\frac{k}{n}\mod 1;\ 1\le k\le n\}.$$

We may choose an integer k with $1\le k\le n$ such that

$$x_2-\frac{m}{n}(x_1-[x_1])-m+\frac{mk}{n}=y_2-\frac{t}{n}\tag{3.7.1}$$

for some $t \in [0,1)$ and some $y_2 \in \mathbf{Z}$. We define $s = (x_1 - [x_1] + n - k)/n$ and note $s \in [0,1)$. The left-hand side of (3.7.1) equals $x_2 - ms$. We set $y_1 = [x_1] - n + k$ and $y = (y_1, y_2)^{\mathrm{t}} \in \mathbf{Z}^2$. Part (i) now follows since $x = y + N(s,t)^{\mathrm{t}}$.

We now prove part (ii). So let x' and x'' be as in the hypothesis, i.e. $x' - x'' = Nu \in \mathbf{Z}^2$. We may assume $u = (s,t)^{\mathrm{t}}$ with $0 \le t < 1$ and $|s| < 1$. Now ns and $ms - t/n$ are both integers. We subtract n times the second expression from m times the first and conclude $t \in \mathbf{Z}$. Hence $t = 0$ and so ms, ns are integers. There are integers a and b with $am + bn = 1$, it follows that $s = ams + bns \in \mathbf{Z}$. Hence $s = 0$; we conclude $x' = x''$. □

The map $\theta \mapsto f_{nm}(\theta/n)$ is continuous on $[0,1]$, so we may integrate.

LEMMA 3.17. *Let m, n be coprime integers with $0 \ne n \ge m \ge 0$, then*

$$\int_0^1 f_{nm}(\frac{\theta}{n})d\theta = S_0.$$

PROOF. We have

$$\int_0^1 f(\frac{\theta}{n})d\theta = \int_{[0,1]^2} \log|e^{2\pi i n u} - e^{2\pi i(mu - \theta/n)} + 1|du d\theta.$$

We may replace $[0,1]^2$ by $[0,1)^2$ in this equality, indeed the two sets differ by a set of measure zero. We apply the linear change of coordinates $u' = nu$, $\theta' = mu - \theta/n$ of determinant -1 defined in Lemma 3.16. So

$$\int_0^1 f(\frac{\theta}{n})d\theta = \int_F \log|e^{2\pi i u'} - e^{2\pi i \theta'} + 1|d\theta' du' \tag{3.7.2}$$

where $F \subset \mathbf{R}^2$ is as in Lemma 3.16. The integrand in (3.7.2) is invariant under the action of $\mathbf{Z}^2$ on $\mathbf{R}^2$. Because of Lemma 3.16 we may replace F in (3.7.2) by $[0,1)^2$ and even by $[0,1]^2$. Therefore

$$\int_0^1 f(\frac{\theta}{n})d\theta = \int_{[0,1]^2} \log|e^{2\pi i u'} - e^{2\pi i \theta'} + 1|d\theta' du'. \tag{3.7.3}$$

The translation $u' \mapsto u' + 1/2$ shows that (3.7.3) is equal to the logarithmic Mahler measure of the two variable polynomial $X + Y - 1$ which is by definition just S_0. □

PROPOSITION 3.1. *Let $x \in \mathcal{M}$ with $D = [\mathbf{Q}(x) : \mathbf{Q}]$. There exists a positive integer d with $\phi(d) \le D$ such that*

$$|h(x, 1-x) - S_0| \le c(d) D^{-1}$$

where

$$c(d) = \begin{cases} 260 & \text{if } d|6, \\ (4\pi\sqrt{3} + \frac{8}{\sqrt{3}})2^{\omega(d)} h_d & \text{otherwise.} \end{cases}$$

PROOF. Let α, ζ, m, n be as in Lemma 3.2, and define $d = \operatorname{ord}\zeta$. The proposition is easy if $h(x, 1-x) = 0$. Indeed in this case x is a root of unity by Kronecker's Theorem. By the discussion in the beginning of section 2, x is a 6th root of unity and we can take $d = 1$.

So we assume $h(x,1-x)>0$. We set $K=\mathbf{Q}(\zeta)$ and note $\phi(d)\leq D$. By Lemmas 3.5 and 3.17 the absolute value $|h(x,1-x)-S_0|$ equals

$$\frac{n}{[K(\alpha):K]}\left|\left(\frac{1}{[K:\mathbf{Q}]}\sum_{\substack{1\leq a\leq d\\(a,d)=1}}f(\frac{a}{dn})\right)-\int_0^1 f(\frac{\theta}{n})d\theta+\frac{n-[K(\alpha):K]}{n}S_0\right|,$$

so

$$|h(x,1-x)-S_0|\leq\frac{n}{[K(\alpha):K]}\left|\left(\frac{1}{\phi(d)}\sum_{\substack{1\leq a\leq d\\(a,d)=1}}f(\frac{a}{dn})\right)-\int_0^1 f(\frac{\theta}{n})d\theta\right|+\frac{|n-[K(\alpha):K]|}{[K(\alpha):K]}S_0. \tag{3.7.4}$$

Let V be the total variation of the map $\theta\mapsto f(\theta/n)$ over $[0,1)$. We apply Koksma's inequality (Theorem 5.4 [**Har98**] page 124) to the first term in the upper bound of (3.7.4) to deduce

$$|h(x,1-x)-S_0|\leq V\frac{n}{[K(\alpha):K]}\frac{\mathcal{D}(d)}{\phi(d)}+\frac{|n-[K(\alpha):K]|}{[K(\alpha):K]}S_0,$$

here $\mathcal{D}(d)$ is as defined in (3.6.1).

By Lemma 3.14 we have $V\leq\frac{4\pi}{\sqrt{3}n}$ and by Lemma 3.15 $\mathcal{D}(d)\leq(3+2/\pi)2^{\omega(d)}h_d$, therefore

$$\begin{aligned}|h(x,1-x)-S_0|&\leq\frac{4\pi}{\sqrt{3}}(3+2/\pi)\frac{2^{\omega(d)}h_d}{[K(\alpha):K]\phi(d)}+\frac{|n-[K(\alpha):K]|}{[K(\alpha):K]}S_0\\&=(4\pi\sqrt{3}+\frac{8}{\sqrt{3}})\frac{2^{\omega(d)}h_d}{D}+\frac{|n-[K(\alpha):K]|}{D}\phi(d)S_0.\end{aligned} \tag{3.7.5}$$

There are two cases.

Let us first assume that $d|6$. By Lemma 3.5 we have $|[K(\alpha):K]-n|\leq 2$. Therefore (3.7.5) implies

$$|h(x,1-x)-S_0|\leq(4\pi\sqrt{3}+\frac{8}{\sqrt{3}})\frac{2^{\omega(d)}h_d}{D}+2\frac{\phi(d)}{D}S_0$$

Since $d\in\{1,2,3,6\}$ we easily bound

$$|h(x,1-x)-S_0|<\frac{260}{D},$$

and hence the proposition holds for $d\mid 6$.

Now let us assume $d\nmid 6$. Then $[K(\alpha):K]=n$ by Lemma 3.5. Therefore the dangerously large term on the very right of (3.7.5) disappears. We have

$$|h(x,1-x)-S_0|\leq(4\pi\sqrt{3}+\frac{8}{\sqrt{3}})\frac{2^{\omega(d)}h_d}{D},$$

which completes the proof. □

The proof of Theorem 3.1 is now a consequence of the previous proposition together with standard inequalities concerning arithmetic functions.

Let $\epsilon > 0$ and say $x \in \mathcal{M}$ with $D = [\mathbf{Q}(x) : \mathbf{Q}]$. Let d be as in Proposition 3.1. By Theorem 317 and §22.13 of [**HW05**] there exists $c_1(\epsilon)$ independent of d such that $\omega(d) \leq (1+\epsilon)\frac{\log d}{\log\log(3d)} + c_1(\epsilon)$. It is well-known that there exists $c_2(\epsilon)$, also independent of d, such that $\phi(d) \geq c_2(\epsilon)d^{1-\epsilon}$. We also recall the elementary inequality $h_d \leq 1 + \log d$. Hence

$$2^{\omega(d)}h_d \leq c_3(\epsilon)e^{(\epsilon+\log 2)\frac{\log\phi(d)}{\log\log(3\phi(d))}}.$$

for some constant $c_3(\epsilon)$ and all $d \geq 1$. The theorem follows from the inequality above, Proposition 3.1, and $\phi(d) \leq D$. □

Finally we prove Corollary 3.3.

Let $x \in \mathcal{M}$ with $x^r(1-x)^s = 1$ where r, s are integers and not both zero. Then an easy local calculation gives $\max\{h(x), h(1-x)\} = \frac{\max\{|r|,|s|\}}{||(r,s)||}h(x, 1-x)$. We have the inequalities $\frac{1}{2}||(r,s)|| \leq \max\{|r|,|s|\} \leq ||(r,s)||$ which together with Theorem 3.1 imply that any height value not in $[S_0/2, S_0]$ is isolated. Furthermore, for any $\beta \in [1/2, 1]$ we may choose two sequences of positive integers $r_n \leq s_n$ with $(r_n, s_n) = 1$, s_n strictly increasing, and with $s_n/(r_n + s_n) \to \beta$. Let $x_n > 1$ be a real with $x_n^{r_n}(1-x_n)^{s_n} = 1$. Then the x_n must be pairwise different since they are not roots of unity. Because $h(x_n) \leq \log 2$ we have $[\mathbf{Q}(x_n) : \mathbf{Q}] \to \infty$. By Theorem 3.1 and our choice of r_n, s_n we have $\max\{h(x_n), h(1-x_n)\} = \frac{s_n}{r_n+s_n}h(x_n, 1-x_n)$. This sequence converges to βS_0. □

8. Counting multiplicative dependent points

We show an easy consequence of Lemmas 3.2 and 3.5.

LEMMA 3.18. *Let $D \geq 1$ and $x \in \mathcal{M}(D)$, there exist coprime integers r, s with $r \geq 0$ and a root of unity η of order d such that $x^r(1-x)^s = \eta$ and*

$$||(r,s)|| \leq \frac{D}{\phi(d)} + 2.$$

PROOF. Let $x \in \mathcal{M}(D)$. Let k, l, m, n, r, s, α, and ζ be as in Lemma 3.2. Say $R = X^n - \zeta^l X^m + \zeta^k$ and $K = \mathbf{Q}(\zeta)$. We may assume $r \geq 0$ by replacing $\pm\zeta$ with $\pm\zeta^{-1}$ if necessary.

First say $h(x, 1-x) = 0$. By Kronecker's Theorem α is a root of unity and we may take $r = 1$, $s = 0$, and $\eta = x$.

Now say $h(x, 1-x) > 0$. We set $x^r(1-x)^s = \eta$. Let d be the order of η, then $[K : \mathbf{Q}] = \phi(d)$ since $\eta = \pm\zeta^{\pm 1}$. Lemma 3.2 implies $||(r,s)|| = n$. We have $R(\alpha) = 0$ and α is not a root of unity. Hence by Lemma 3.5 we conclude $[K(\alpha) : K] \geq n - 2$. The lemma now follows from $[K(\alpha) : K] = [\mathbf{Q}(x) : K] \leq D/\phi(d)$. □

Let r, s be coprime integers and η a root of unity. In Lemma 7 of [**Mas07**] Masser proved the bound

$$|\{x \in \overline{\mathbf{Q}} \backslash \{0,1\};\ x^r(1-x)^s = \eta\}| \leq ||(r,s)||.$$

This bound together with Lemma 3.18 gives the following upper bound for the cardinality of $\mathcal{M}(D)$:

$$\begin{aligned} |\mathcal{M}(D)| &\leq \sum_{\substack{d\geq 1 \\ \phi(d)\leq D}} \sum_{\substack{r,s \text{ coprime, } r\geq 0 \\ ||(r,s)||\leq 2+\frac{D}{\phi(d)}}} \sum_{\substack{\eta \\ \text{ord } \eta = d}} ||(r,s)|| \\ &= \sum_{\substack{d\geq 1 \\ \phi(d)\leq D}} \phi(d) \sum_{\substack{r,s \text{ coprime, } r\geq 0 \\ ||(r,s)||\leq 2+\frac{D}{\phi(d)}}} ||(r,s)||. \end{aligned} \tag{3.8.1}$$

We follow [**Mas07**] and define $S(X) = \sum_{1\leq m\leq X} m\phi(m)$ for real $X \geq 1$.

LEMMA 3.19. *For $X \geq 1$ we have*

$$\sum_{\substack{r,s \text{ coprime, } r\geq 0 \\ ||(r,s)||\leq X}} ||(r,s)|| = 3S(X) + 1. \tag{3.8.2}$$

PROOF. The proof follows by splitting the sum in (3.8.2) up into three sums over $s \geq 0$, $-r \leq s < 0$, and $s < -r \leq 0$ respectively. The first sum equals $S(X) + 1$ and each of the other two is $S(X)$. □

LEMMA 3.20. *For $X \geq 1$ we have $S(X) \leq \frac{2}{\pi^2}X^3 + X^2 \log X + 2X^2$.*

PROOF. The proof is similar to the proof of Lemma 8 given in [**Mas07**] which provides an analogous lower bound for $S(X)$.

Recall that μ denotes the Möbius function. For $Y \geq 1$ we define $T(Y) = \sum_{1\leq d\leq Y} d^2$. The identity $\phi(m) = \sum_{d|m} \mu(d)m/d$ implies

$$S(X) = \sum_{1\leq d\leq X} \mu(d) d T(\frac{X}{d})$$

after changing the order of summation. Let $X \geq 1$, we use $\sum_{d=1}^{\infty} \mu(d)d^{-2} = 6/\pi^2$ to conclude

$$\begin{aligned} \left|S(X) - \frac{2}{\pi^2}X^3\right| &= \left|S(X) - \frac{X^3}{3}\sum_{d=1}^{\infty}\frac{\mu(d)}{d^2}\right| \\ &\leq \left|\sum_{1\leq d\leq X}\left(\mu(d)dT(\frac{X}{d}) - \frac{X^3}{3d^2}\mu(d)\right)\right| + \frac{X^3}{3}\left|\sum_{d>X}\frac{\mu(d)}{d^2}\right| \\ &\leq \sum_{1\leq d\leq X} d\left|T(\frac{X}{d}) - \frac{X^3}{3d^3}\right| + \frac{X^3}{3}\sum_{d>X}\frac{1}{d^2}. \end{aligned}$$

Elementary calculations give $|T(Y) - \frac{1}{3}Y^3| \leq Y^2$ for $Y \geq 1$. We use this inequality and $\sum_{d>X} d^{-2} \leq 2X^{-1}$ to bound

$$\left| S(X) - \frac{2}{\pi^2}X^3 \right| \leq \left(\sum_{1 \leq d \leq X} \frac{X^2}{d} \right) + \frac{2}{3}X^2.$$

The proof follows from the inequality $\sum_{1 \leq d \leq X} d^{-1} \leq 1 + \log X$. □

The proof of Theorem 3.3 is now simple. Certainly the lower bound in (3.1.3) follows from Theorem 2 in [**Mas07**]. We proceed to prove the upper bound.

Let $D \geq 3$. By (3.8.1) and Lemma 3.19 we have

$$|\mathcal{M}(D)| \leq \sum_{\substack{d \geq 1 \\ \phi(d) \leq D}} \phi(d)(3S(2 + \frac{D}{\phi(d)}) + 1).$$

We apply Lemma 3.20 to this last inequality to obtain

$$|\mathcal{M}(D)| \leq \sum_{\substack{d \geq 1 \\ \phi(d) \leq D}} \left(\frac{6}{\pi^2} \frac{D^3}{\phi(d)^2} + c' \frac{D^2}{\phi(d)} \log D \right), \tag{3.8.3}$$

for some absolute constant c'. We may take $c' = 200$. Lemma 9 of [**Mas07**] gives

$$\sum_{\substack{d \geq 1 \\ \phi(d) \leq D}} \frac{1}{\phi(d)} < 5(\log D)^2,$$

the fact that D may not be an integer is not important. We insert this bound into (3.8.3) to obtain

$$|\mathcal{M}(D)| \leq \frac{6}{\pi^2} D^3 \left(\sum_{\substack{d \geq 1 \\ \phi(d) \leq D}} \frac{1}{\phi(d)^2} \right) + 5c' D^2 (\log D)^3 \leq c_0 D^3 + 1000 D^2 (\log D)^3.$$

The proof follows from this last inequality. □

CHAPTER 4

Dependent solutions of $x + y = \alpha$

This chapter extends some methods used in chapter 2: we give essentially optimal upper bounds for the height of multiplicatively dependent algebraic solutions of the inhomogeneous linear equation $x + y = \alpha$ in two unknowns x and y where α is any non-zero algebraic number. Furthermore, we will study the case where $\alpha \geq 2$ is a rational power of a non-zero integer and derive a better bound for the height of the solution. We will also see that this bound is best possible in the case where $\alpha \geq 2$ is a rational power of 2 and even that the maximal height value is isolated if α is also assumed to be an integer. Moreover, for non-zero rational α we give a bound independent of α for the number of solutions of $x + y = \alpha$ if the unknowns are algebraic units in the union of all number fields that have unit group of rank 1.

The content of this chapter has been published in [**Hab05**].

1. Height bounds for dependent solutions of $x + y = \alpha$

We recall that $H(\cdot)$ denotes the non-logarithmic absolute Weil height defined in chapter 1.

Let α be a non-zero algebraic number. In Theorem 4.1 we will show that the heights of multiplicatively dependent algebraic numbers x, y satisfying

$$x + y = \alpha \tag{4.1.1}$$

are effectively bounded in terms of the height of α. More precisely:

THEOREM 4.1. *Let α be a non-zero algebraic number, and let x, y be non-zero multiplicatively dependent algebraic numbers with $x + y = \alpha$. Then*

$$H(x, y) \leq 2H(\alpha)^2, \tag{4.1.2}$$

$$\min\{H(x), H(y)\} \leq 2H(\alpha). \tag{4.1.3}$$

The general strategy in the proof of Theorem 4.1 is the following: Given non-zero algebraic numbers x, y, α as in Theorem 4.1, there are integers r, s not both zero such that $x^r y^s = 1$. Define the rational function $f = T^r(T - \alpha)^s$ and let d be the degree of f. We note that $f(x) = \pm 1$. For any algebraic τ not a pole of f we will derive a lower bound for $H(f(\tau))H^{-d}$ in terms of $H(\alpha)$ and d. Here H will be either $H(\tau)$ or $H(\tau, \tau - \alpha)$ depending on the sign of rs. By substituting $\tau = x$ we will get an upper bound for $H(x)$ independent of r, s, and d. Here it will be essential that the lower bound

has optimal dependency on d; we will achieve this through careful estimates of certain positive local minima, together with the use of the product formula to avoid zero values.

In the special case $\alpha = 1$ we get

$$H(x, y) \leq 2, \tag{4.1.4}$$

and recover the height bound from Theorem 2.1. Since $\max\{H(x), H(y)\} \leq H(x, y)$, the inequality (4.1.4) also implies Theorem 1 of Cohen and Zannier's article [**CZ00**].

The bound (4.1.4) and also the one from [**CZ00**] are easily seen to be sharp after setting $x = y = 1/2$. More generally, pick any $\phi \in \mathbf{Q}$ with $\phi \geq 0$ and set $x = y = 2^{-\phi-1}$, $\alpha = 2^{-\phi}$. Then $x + y = \alpha$. Further, by using the standard properties of the height function stated in the next section one obtains $H(x, y) = 2^{\phi+1} = 2H(\alpha)$. So an upper bound for $H(x, y)$ has to be at least linear in $H(\alpha)$ and (4.1.3) is sharp. We ask the questions: can the bound for the height in (4.1.2) be improved? If so, is the upper bound linear in $H(\alpha)$? The next theorem gives a positive answer to the first question. It is proved using simple ideas from diophantine approximation.

THEOREM 4.2. *Let α be a non-zero algebraic number, and let x, y be non-zero multiplicatively dependent algebraic numbers with $x + y = \alpha$. Then*

$$H(x, y) \leq 14H(\alpha) \log(3H(\alpha)). \tag{4.1.5}$$

Note that the estimate given in Theorem 4.2 is asymptotically much better for $H(\alpha) \to \infty$ then the one given in Theorem 4.1, even though it is worse for small $H(\alpha) < 31$.

Still the logarithm in (4.1.5) seems a bit disturbing, as it does not answer the second question posed above. In fact a linear inequality like

$$H(x, y) \leq CH(\alpha)$$

with an absolute constant C is impossible - even if x, y, and α are restricted to $\mathbf{Q}$ - as the following Theorem shows.

THEOREM 4.3. *For each pair of reals θ, Θ with $0 < \theta < 1$ and $\Theta > 1$, there are non-zero $\alpha \in \mathbf{Q}$ with $H(\alpha) > \Theta$, and non-zero multiplicatively dependent $x, y \in \mathbf{Q}$ with $x + y = \alpha$, such that*

$$\max\{H(x), H(y)\} > \theta \frac{H(\alpha) \log 3H(\alpha)}{(\log \log 3H(\alpha))^2}.$$

So the bound given in Theorem 4.2 is optimal in the sense that it cannot be replaced by $c(\epsilon)H(\alpha)(\log 3H(\alpha))^{1-\epsilon}$ for some $\epsilon > 0$ (and similarly the bound given in Theorem 4.3 cannot be replaced by $c(\epsilon)H(\alpha)(\log 3H(\alpha))^{1+\epsilon}$).

Until now we have considered any non-zero algebraic α, but if we restrict α to special values, then we are able to improve (4.1.5) to a linear bound.

THEOREM 4.4. *Let $\alpha = n^\phi$ where n is a positive integer and ϕ a positive rational, and suppose that $\alpha \geq 2$. Then for all non-zero multiplicatively dependent algebraic*

numbers x, y with $x + y = \alpha$ we have

$$H(x, y) \leq 2H(\alpha) \tag{4.1.6}$$

with equality if and only if α is a rational power of 2 and $x = 2\alpha$ or $y = 2\alpha$.

We already know that inequality (4.1.6) holds for $\alpha = 1$ by Theorem 4.1. So (4.1.6) is valid for every non-zero integer α.

Once we have a sharp upper bound, say $B(\alpha)$ for $H(x, y)$ as in Theorem 4.4, it is an interesting problem to determine if this upper bound is isolated in the sense that there exists $\epsilon(\alpha) > 0$ such that either

$$H(x, y) = B(\alpha) \quad \text{or} \quad H(x, y) \leq B(\alpha) - \epsilon(\alpha). \tag{4.1.7}$$

This kind of problem was first studied by Cohen and Zannier ([**CZ00**] Proposition 1) who proved isolation for $\alpha = 1$ with $B(\alpha) = 2$ and with $\max\{H(x), H(y)\}$ instead of $H(x, y)$. They used Bilu's Equidistribution Theorem. In chapter 2 we proved that if $\alpha = 1$, then either $H(x, y) = 2$ or $H(x, y) \leq 1.915$. We will now formulate an isolation result if $\alpha = 2^\phi$ with $\phi \in \mathbf{N}$.

THEOREM 4.5. *Let $\alpha = 2^\phi$ where ϕ is a positive integer. Then for all non-zero multiplicatively dependent algebraic numbers x, y with $x + y = \alpha$ we have either*

$$H(x, y) = 2H(\alpha) \quad \textit{or} \quad H(x, y) \leq 1.98H(\alpha).$$

Theorem 4.5 shows not only that our $\epsilon(\alpha)$ in (4.1.7) can be chosen independent of α, but that it can even be chosen such that $\epsilon(\alpha) \to \infty$ if $H(\alpha) \to \infty$.

We move on to an application. Given a subset $M \subset \overline{\mathbf{Q}}$ we define $S_M(\alpha)$ to be the set of pairs $(x, y) \in M^2$ that satisfy (4.1.1) and where x, y are also required to be units in the ring of algebraic integers. Equations of this type are called unit equations and have been studied extensively for example if $M = K$ is a number field. In this case it is a well-known result that $S_K(\alpha)$ is finite and even $|S_K(\alpha)| \leq c(K)$, i.e. the cardinality is bounded independently of α. For example Evertse showed ([**Eve84**] Theorem 1) $|S_K(\alpha)| \leq 3 \cdot 7^{2[K:\mathbf{Q}]+r}$ where r is the number of real embeddings of K. Further bounds for $|S_K(\alpha)|$ have been obtained by Beukers and Schlickewei ([**BS96**] Theorem 1.1).

Now let K be a number field with unit group of rank 1 and $\alpha \in \mathbf{Q}^*$. If x and y solve the unit equation, then they are multiplicatively dependent and so by Theorem 4.1 their height is bounded effectively in terms of $H(\alpha)$. Define $\mathscr{F}$ to be the union of all number fields with unit group of rank 1. Then the same argument just given leads to an effective height bound for x, y with $(x, y) \in S_{\mathscr{F}}(\alpha)$. In both cases Dirichlet's Unit Theorem shows that the degrees of x and y do not exceed 4. So Northcott's Theorem implies that $S_K(\alpha)$ and $S_{\mathscr{F}}(\alpha)$ are finite sets. In fact we will prove a uniform result in the spirit of Evertse.

THEOREM 4.6. *Let α be a non-zero rational number and K a number field with unit group of rank 1 and $\mathscr{F}$ as above. Then $|S_K(\alpha)| \leq 292$ and $|S_{\mathscr{F}}(\alpha)| \leq 755 \cdot 10^6$.*

The first inequality is merely a numerical improvement of a special case of Evertse's bound which leads to $|S_K(\alpha)| \leq 3 \cdot 7^8 = 17294403$. The second result is best possible

in the sense that if $\mathscr{F}$ is replaced by $\mathscr{F}'$, which is the union of all number fields with unit group of rank 2, then $S_{\mathscr{F}'}(\alpha)$ is infinite. Indeed let $a \in \mathbf{Z}$ and let x be a zero of the polynomial $T(1-T)(a-T)-1 \in \mathbf{Z}[T]$ which is easily seen to be irreducible. Then x and $1-x$ are units. Clearly there are infinitely many such x as a runs over all rational integers. This is not a contradiction to the Theorem above because for a large enough x lies in a cubic number field with unit group of rank 2.

In section 2 we introduce notation used throughout the chapter and state the required results on lower bounds for the height of values of special rational functions. In section 3 we prove the statements made in section 2. These results are then used in section 4 to prove Theorems 1 and 2. In section 5 we prove Theorem 3 by construction. Finally in sections 6, 7, and 8 we prove Theorems 4, 5, and 6 respectively.

2. Notation and auxiliary results on rational functions

Let $f = f(T)$ be a rational function with algebraic coefficients and let $\tau \in \overline{\mathbf{Q}}$. We are interested in how $H(f(\tau))$ depends on $H(\tau)$ and f (provided τ is not a pole of f). Recall that the degree of f is defined as $\max\{\deg(P), \deg(Q)\}$ for any two coprime polynomials P, Q with $f = P/Q$. Classically it is known that

$$C \leq \frac{H(f(\tau))}{H(\tau)^{\deg(f)}} \leq C' \tag{4.2.1}$$

with positive constants C, C' that are independent of τ (for example it follows easily from Theorem 1.8 page 81 of [**Lan83**]).

As pointed out in section 1 we are particularly interested in a sharp lower bound of (4.2.1) for certain rational functions. The first function we will investigate has the form $f = T^r(T-\alpha)^s$ with α a non-zero algebraic number. Its degree is $|r|+|s|$ if $rs \geq 0$, and $\max\{|r|,|s|\}$ if $rs < 0$.

PROPOSITION 4.1. *Let r, s be integers and α a non-zero algebraic number and define the rational function $f = T^r(T-\alpha)^s$. Put*

$$e = e(f) = \begin{cases} 2|s|-|r| & (rs<0, |r|<|s|), \\ |r| & (rs<0, |r|\geq|s|), \\ |r|+|s| & (rs\geq 0) \end{cases}$$

and

$$H = \begin{cases} H(\tau) & rs<0, \\ H(\tau, \tau-\alpha) & rs\geq 0. \end{cases}$$

Then for all algebraic $\tau \neq 0$, α we have

$$\frac{H(f(\tau))}{H^d} \geq \frac{1}{2^d}\frac{1}{H(\alpha)^e} \geq \frac{1}{2^d}\frac{1}{H(\alpha)^{2d}}. \tag{4.2.2}$$

This exponent $e = e(f)$ in (4.2.2) looks strange, but in fact it is best possible for each value of r and $s \neq 0$.

We will also study lower bounds for certain types of polynomials.

PROPOSITION 4.2. *Let m, n be integers with $n > m \geq 0$, let β be an algebraic number, and define the polynomial $P = T^n + \beta T^m$. Put $\theta = m/n$. Then for all algebraic τ we have*

$$\frac{H(P(\tau))}{H(\tau)^n} \geq \frac{1-\theta}{2} \frac{1}{H(\beta)^{1/(1-\theta)}} \geq \frac{1}{2n} \frac{1}{H(\beta)^n}. \tag{4.2.3}$$

Furthermore, if $n = m > 0$ and $\beta \neq -1$, then

$$\frac{H(P(\tau))}{H(\tau)^n} \geq \frac{1}{2n} \frac{1}{H(\beta)^n}.$$

If β is a root of unity, then Proposition 4.2 reduces to $H(P(\tau))H(\tau)^{-n} \geq (2n)^{-1}$. In Proposition 4.3 we use diophantine approximation to get a lower bound independent of n. This improvement comes at a price: the denominator in the left-hand side of (4.2.3) has to replaced by $\frac{X}{\log X}$ for large $X = H(\tau)^n$. So strictly speaking we are not in the situation of (4.2.1) anymore. Nevertheless Proposition 4.3 is essential for the proof of Theorem 4.2.

PROPOSITION 4.3. *Let m, n be integers with $n > m \geq 0$, let ζ be a root of unity, and define the polynomial $P = T^n + \zeta T^m$. Then for all algebraic τ we have*

$$\frac{H(P(\tau))}{H(\tau)^n} \max\{1, \log(H(\tau)^n)\} \geq \frac{1}{2e}$$

with $e = 2.71828...$

Compare this lower bound with the easy upper bound

$$H(P(\tau)) = H(\tau^m(\tau^{n-m} + \zeta)) \leq 2H(\tau^m)H(\tau^{n-m}) = 2H(\tau)^n.$$

By combining the previous inequality with the inequality from Proposition 4.3, we will deduce the following amusing consequence for the logarithmic height $h(\tau) = \log H(\tau)$.

COROLLARY 4.1. *For $\sigma \in \overline{\mathbf{Q}}$ and $\phi \in \mathbf{Q}$ with $0 \leq \phi \leq 1$ take any determination of σ^ϕ. Then*

$$|h(\sigma + \sigma^\phi) - h(\sigma)| \leq 1 + \log 2 + \log \max\{1, h(\sigma)\}.$$

3. Proofs of Propositions 4.1, 4.2, 4.3 and Corollary

Our general strategy in estimating heights $H(f(\tau))$ where f and τ are defined as in Proposition 4.1 is to consider each local factor separately. In the first step we will consider the case $rs \geq 0$. It will actually suffice to suppose $r \geq 0$ and $s \geq 0$, so that f is a polynomial. Because some results proved below remain valid in a more general context we will work with an arbitrary field K containing α and τ equipped with an absolute value $|\cdot|$.

For an absolute value $|\cdot|$ we define

$$m_f(\tau) = \frac{\max\{1, |\tau|^r |\tau - \alpha|^s\}}{\max\{1, |\tau|, |\tau - \alpha|\}^d} \tag{4.3.1}$$

with $d = \deg(f) = r + s$. The subscript f will be omitted if it is clear from the context what function is meant.

LEMMA 4.1. *Suppose $r \geq 0, s \geq 0$. Let $|\cdot|$ be an ultrametric absolute value on a field K. Then for any $\tau \in K\backslash\{0, \alpha\}$*

$$m_f(\tau) \geq \frac{1}{\max\{1, |\alpha|\}^d}.$$

PROOF. First assume $|\tau| \leq 1$, then $|\tau - \alpha| \leq \max\{1, |\alpha|\}$, so $m(\tau) \geq \max\{1, |\tau|, |\tau - \alpha|\}^{-d} \geq \max\{1, |\alpha|\}^{-d}$ and the lemma follows in this case.

Let us now assume $|\tau| \geq 1$. We split up into two cases.

The first case is $|\tau| \neq |\alpha|$. This inequality implies $|\tau - \alpha| = \max\{|\tau|, |\alpha|\}$ so

$$m(\tau) = \frac{\max\{1, |\tau|^{r+s}, |\tau|^r |\alpha|^s\}}{\max\{1, |\tau|, |\alpha|\}^d} \geq \frac{1}{\max\{1, |\alpha|\}^d}$$

since $d = r + s$ and after considering the two possibilities $|\tau| > |\alpha|$ and $|\tau| < |\alpha|$. Hence the lemma holds in this case.

The second case is $|\tau| = |\alpha|$, then

$$m(\tau) \geq \frac{\max\{1, |\alpha|^r |\tau - \alpha|^s\}}{\max\{1, |\alpha|\}^d} = \max\{|\alpha|^{-d}, \frac{|\tau - \alpha|^s}{|\alpha|^s}\} \geq |\alpha|^{-d} = \frac{1}{\max\{1, |\alpha|\}^d}.$$

□

We now treat the case that our absolute value is not ultrametric but satisfies the weaker triangle inequality. We obtain a slightly worse estimate than in Lemma 4.1.

LEMMA 4.2. *Suppose $r \geq 0$, $s \geq 0$. Let $|\cdot|$ be an absolute value on a field K. Then for any $\tau \in K\backslash\{0, \alpha\}$*

$$m_f(\tau) \geq \frac{1}{(1 + |\alpha|)^d} \geq \frac{1}{2^d \max\{1, |\alpha|\}^d}. \tag{4.3.2}$$

PROOF. The second inequality in (4.3.2) follows from

$$1 + |\alpha| \leq 2 \max\{1, |\alpha|\}.$$

We proceed by proving the first.

First say $|\tau| \leq 1$, then $|\tau - \alpha| \leq |\tau| + |\alpha| \leq 1 + |\alpha|$ so $m(\tau) \geq \max\{1, |\tau - \alpha|\}^{-d} \geq (1 + |\alpha|)^{-d}$ and (4.3.2) follows.

Now let us assume $|\tau| \geq 1$. Say first $|\tau - \alpha| \geq |\tau|$, then

$$m(\tau) = \frac{|\tau|^r}{|\tau - \alpha|^r} = \left|1 - \frac{\alpha}{\tau}\right|^{-r} \geq \left(1 + \left|\frac{\alpha}{\tau}\right|\right)^{-r} \geq \frac{1}{(1 + |\alpha|)^r} \geq \frac{1}{(1 + |\alpha|)^d},$$

which implies (4.3.2). Say now $|\tau - \alpha| \leq |\tau|$, then

$$m(\tau) = \max\{1, |\tau|^r |\tau - \alpha|^s|\}|\tau|^{-d}.$$

The lemma follows easily if $s = 0$, so let us assume $s > 0$. If $|\tau| \leq 1 + |\alpha|$; then $m(\tau) \geq |\tau|^{-d} \geq (1 + |\alpha|)^{-d}$ by the equality for $m(\tau)$ above. If on the other hand $|\tau| \geq 1 + |\alpha|$, then $|\tau - \alpha| \geq |\tau| - |\alpha| \geq 1$ and so we have

$$m(\tau)^{1/s} = \left|\frac{\tau - \alpha}{\tau}\right| = \left|1 - \frac{\alpha}{\tau}\right| \geq 1 - \left|\frac{\alpha}{\tau}\right| \geq 1 - \frac{|\alpha|}{1 + |\alpha|} = \frac{1}{1 + |\alpha|},$$

hence $m(\tau) \geq (1 + |\alpha|)^{-s} \geq (1 + |\alpha|)^{-d}$, which concludes the proof. □

We will now cover the non-polynomial case of Proposition 4.1, that is if $rs < 0$. Recalling the definition of $m_f(\tau)$ in (4.3.1) with $K = \mathbf{C}$ and with $|\cdot|$ the standard absolute value. Taking for example $r = 2$ and $s = -1$ we obtain

$$m_f(\tau) \leq \frac{|\tau|^2/|\tau - \alpha|}{|\tau|^2} = |\tau - \alpha|^{-1},$$

if $|\tau|$ is large. Therefore unfortunately $\lim_{\tau \to \infty} m_f(\tau) = 0$, so m cannot be bounded away from 0 as in Lemma 4.2.

So we redefine m_f in this case. Our motivation comes from the calculation

$$\begin{aligned} H(f(\tau))^{[K:\mathbf{Q}]} &= \prod_{v \in M_K} |\tau - \alpha|_v^{-sd_v} \prod_{v \in M_K} \max\{1, |\tau|_v^r |\tau - \alpha|_v^s\}^{d_v} \\ &= \prod_{v \in M_K} \max\{|\tau|_v^r, |\tau - \alpha|_v^{-s}\}^{d_v} \end{aligned} \tag{4.3.3}$$

where we have applied the product formula (1.1.1) for the number field K if $\tau \neq \alpha$. The d_v are the local indices defined in chapter 1.

Now if we assume $r \geq 0$ and set $t = -s \geq 0$, then the new definition

$$m_f(\tau) = \frac{\max\{|\tau|^r, |\tau - \alpha|^t\}}{\max\{1, |\tau|\}^d} \quad \text{with} \quad d = \deg(f) = \max\{r, t\} \tag{4.3.4}$$

will do the trick. We note that this time there is no extra $|\tau - \alpha|$ in the denominator of (4.3.4).

LEMMA 4.3. *Suppose $r > 0 > s = -t$ and let $\epsilon = 1 - r/t$. Let $|\cdot|$ be an ultrametric absolute value on a field K. Then for any $\tau \in K \backslash \{0, \alpha\}$*

$$m_f(\tau) \geq \begin{cases} \max\{1, |\alpha|\}^{-\epsilon d} & : \quad \textit{if} \quad r < t \quad \textit{and} \quad |\alpha| = |\tau| > 1 \\ \max\{1, |\alpha|^{-1}\}^{-d} & : \quad \textit{otherwise.} \end{cases}$$

PROOF. As in the proof of Lemma 4.1 the proof is just a study of different cases. Assume $|\tau| \neq |\alpha|$. Then one has $|\tau - \alpha| = \max\{|\tau|, |\alpha|\}$ and hence

$$m(\tau) = \frac{\max\{|\tau|^r, |\tau|^t, |\alpha|^t\}}{\max\{1, |\tau|^d\}}.$$

If $|\tau| \geq 1$ then clearly $m(\tau) \geq 1$ so let $|\tau| < 1$. Then

$$m(\tau) \geq |\alpha|^t \geq \frac{1}{\max\{1, |\alpha|^{-1}\}^t} \geq \frac{1}{\max\{1, |\alpha|^{-1}\}^d}$$

as desired.

Now assume $|\tau| = |\alpha|$. Here

$$m(\tau) \geq \frac{|\alpha|^r}{\max\{1, |\alpha|^d\}}.$$

If $|\alpha| \leq 1$ then

$$m(\tau) \geq |\alpha|^r \geq \frac{1}{\max\{1, |\alpha|^{-1}\}^d}$$

as above.

Finally if $|\alpha| > 1$ then $m(\tau) \geq |\alpha|^{r-d}$. The assertion is obvious if $r = d$, so assume $r < d = t$; then one has

$$m(\tau) \geq |\alpha|^{-\epsilon d} = \frac{1}{\max\{1, |\alpha|\}^{\epsilon d}}.$$

□

LEMMA 4.4. *Suppose $r > 0 > s = -t$ and let $\epsilon = 1 - r/t$. Let $|\cdot|$ be an absolute value on a field K. Then for any $\tau \in K \backslash \{0, \alpha\}$*

$$m_f(\tau) \geq \begin{cases} 2^{-d} \max\{1, |\alpha|^{-1}\}^{-d} \max\{1, |\alpha|\}^{-\epsilon d} & : \quad \textit{if} \quad r < t \\ 2^{-d} \max\{1, |\alpha|^{-1}\}^{-d} & : \quad \textit{otherwise.} \end{cases}$$

PROOF. We will show the slightly stronger statement

$$m(\tau) \geq \begin{cases} 2^{-d} \max\{1, |\alpha|\}^{-\epsilon d} & : \text{ if } \quad r < t, \ \frac{1}{2} \leq \left|\frac{\tau}{\alpha}\right| \leq 2, \ |\alpha| \geq 1, \\ 2^{-d} \max\{1, |\alpha|^{-1}\}^{-d} & : \text{ otherwise.} \end{cases} \tag{4.3.5}$$

First assume that $|\frac{\tau}{\alpha}| < \frac{1}{2}$ or $|\frac{\tau}{\alpha}| > 2$. Then $\frac{1}{2} \max\{|\tau|, |\alpha|\} \leq |\tau - \alpha|$, and therefore

$$m(\tau) \geq 2^{-t} \frac{\max\{|\tau|^r, |\tau|^t, |\alpha|^t\}}{\max\{1, |\tau|^d\}}.$$

If $|\tau| \geq 1$ then $m(\tau) \geq 2^{-t} \geq 2^{-d}$ and we are done. Assume that $|\tau| < 1$. Then

$$m(\tau) \geq 2^{-t} |\alpha|^t \geq 2^{-d} \max\{1, |\alpha|^{-1}\}^{-d}$$

which implies (4.3.5).

Now assume that $\frac{1}{2} \leq \left|\frac{\tau}{\alpha}\right| \leq 2$. If $|\tau| \leq |\alpha|$ then

$$m(\tau) \geq \frac{|\tau|^r}{\max\{1, |\alpha|^d\}} \geq 2^{-r} \frac{|\alpha|^r}{\max\{1, |\alpha|^d\}},$$

and if $|\tau| > |\alpha|$ then

$$m(\tau) \geq 2^{-d} \frac{|\alpha|^r}{\max\{1, |\alpha|^d\}};$$

so in both cases we have $m(\tau) \geq 2^{-d} |\alpha|^r \max\{1, |\alpha|\}^{-d}$. If $|\alpha| < 1$ then

$$m(\tau) \geq 2^{-d} |\alpha|^r \geq 2^{-d} \max\{1, |\alpha|^{-1}\}^{-d}$$

as above. Suppose $|\alpha| \geq 1$. Then $m(\tau) \geq 2^{-d}|\alpha|^{r-d}$. If $r = d$ our assertion is obvious. Finally if $r < d$ then $t = d$ and we have

$$m(\tau) \geq 2^{-d}|\alpha|^{-\epsilon d} = 2^{-d}\max\{1, |\alpha|\}^{-\epsilon d}.$$

□

Proof of Proposition 4.1: One may assume $r \geq 0$ and $s \geq 0$ if $r = 0$ (apply property 1.1.5 if necessary). Let K be any number field containing τ and α, and say $v \in M_K$.

If $rs \geq 0$, then $s \geq 0$. Lemmas 4.1 and 4.2 together imply

$$\frac{\max\{1, |\tau|_v^r |\tau - \alpha|_v^s\}}{\max\{1, |\tau|_v, |\tau - \alpha|_v\}^d} \geq \frac{1}{\delta_v(2)^d \max\{1, |\alpha|_v\}^d}.$$

Now the Proposition in the case $rs \geq 0$ is proved by raising to the d_vth power, taking the product over all elements of M_K and extracting the $[K : \mathbf{Q}]$th root.

If $rs < 0$, then $s < 0$. We set $t = -s$ and also $\epsilon = 1 - r/t$; now Lemmas 4.3 and 4.4 together imply

$$\frac{\max\{|\tau|_v^r, |\tau - \alpha|_v^t\}}{\max\{1, |\tau|_v^d\}} \geq \begin{cases} \delta_v(2)^{-d}\max\{1, |\alpha|_v\}^{-\epsilon d}\max\{1, |\alpha|_v^{-1}\}^{-d} & : r < t, \\ \delta_v(2)^{-d}\max\{1, |\alpha|_v^{-1}\}^{-d} & : r \geq t. \end{cases}$$

Recall (4.3.3); now the Proposition in the case $rs < 0$ is proved by raising to the d_vth power, taking the product over all elements of M_K and extracting the $[K : \mathbf{Q}]$th root.□

We now turn to the proof of Proposition 4.2. It follows the same idea as the proof of Proposition 4.1; that is, each factor in the height is estimated separately. Again the following two lemmas hold in a more general context where K is any field with an absolute value $|\cdot|$. For $\theta \in (0, 1)$ and $\beta, z \in K$ define

$$\hat{m}_{\theta\beta}(z) = \frac{\max\{1, |z|^{\theta/(1-\theta)}|z + \beta|\}}{\max\{1, |z|^{1/(1-\theta)}\}} \tag{4.3.6}$$

The subscripts θ and β will be omitted if the context makes it clear what is meant.

As in the previous section we shall estimate the finite part of the height function first.

LEMMA 4.5. *Let $|\cdot|$ be an ultrametric absolute value on a field K. Then for any $z \in K$*

$$\hat{m}_{\theta\beta}(z) \geq \frac{1}{\max\{1, |\beta|\}^{1/(1-\theta)}}.$$

PROOF. The case where $|z| \leq 1$ is trivial because then $\hat{m}(z) \geq 1$. If $|z| > 1$ we split into three cases. First if $|\beta| < |z|$ then $\hat{m}(z) = 1$. Next if $|\beta| = |z|$ then $\hat{m}(z) = \max\{|z|^{-1/(1-\theta)}, |1 + \beta/z|\} \geq |z|^{-1/(1-\theta)} = |\beta|^{-1/(1-\theta)}$. And finally if $|\beta| > |z|$ then $\hat{m}(z) = |\beta|/|z| > 1$.

□

Finally we will give an estimate of the factors in the infinite part of the height function in the next lemma.

LEMMA 4.6. *Let $|\cdot|$ be an absolute value on a field K. Then for any $z \in K$*

$$\hat{m}_{\theta\beta}(z) \geq \frac{1-\theta}{2} \frac{1}{\max\{1,|\beta|\}^{1/(1-\theta)}}.$$

PROOF. We split up the proof into two parts the first one being when

$$\frac{1-\theta}{2}|z|^{1/(1-\theta)} \max\{1,|\beta|\}^{-1/(1-\theta)} \leq 1. \tag{4.3.7}$$

But then one has

$$\hat{m}(z) \geq \frac{1}{\max\{1,|z|\}^{1/(1-\theta)}} \geq \frac{1-\theta}{2} \frac{1}{\max\{1,|\beta|\}^{1/(1-\theta)}}$$

which is just the assertion. Now assume that (4.3.7) does not hold. This is equivalent to saying $|z| > \phi\mu$ with

$$\phi = \left(\frac{2}{1-\theta}\right)^{1-\theta} \quad \text{and} \quad \mu = \max\{1,|\beta|\}.$$

Obviously one has $|z| > 1$ and therefore

$$\begin{aligned}\hat{m}(z) &\geq \frac{|z|^{\theta/(1-\theta)}|z+\beta|}{|z|^{1/(1-\theta)}} = \frac{|z+\beta|}{|z|} \geq 1 - \frac{|\beta|}{|z|} \geq 1 - \frac{\mu}{|z|} \\ &> 1 - \frac{1}{\phi} \geq \left(1 - \frac{1}{\phi}\right) \frac{1}{\mu^{1/(1-\theta)}}.\end{aligned}$$

The proof of the lemma is complete if the inequality

$$1 - \left(\frac{1-\theta}{2}\right)^{1-\theta} = 1 - \frac{1}{\phi} \geq \frac{1-\theta}{2} \tag{4.3.8}$$

holds. Set $\xi = 1-\theta$ then (4.3.8) is equivalent to

$$g(\xi) \leq 1 \quad \text{for all} \quad \xi \in (0,1) \quad \text{with} \quad g(\xi) = \left(\frac{\xi}{2}\right)^{\xi} + \frac{\xi}{2}. \tag{4.3.9}$$

Note that defining $g(0) = 1$ makes the map continuous on $[0,1]$ and

$$\frac{\mathrm{d}^2 g}{\mathrm{d}\xi^2} = \left(\frac{\xi}{2}\right)^{\xi} \left((\log(\xi/2)+1)^2 + \frac{1}{\xi}\right) > 0$$

for $\xi \in (0,1)$. So g is convex which implies (4.3.9) because $g(0) = g(1) = 1$. □

Note that Lemma 4.6 is in some cases an improvement of Lemma 4.2. Indeed recall the definition of f and assume $r \geq 0, s > 0$ so $d = r+s > 0$. Set $\theta = r/d, \xi = 1-\theta = s/d$. Then Lemma 4.6 with $\beta = -\alpha$, in conjuction with (4.3.1), (4.3.6) imply

$$m_f(z) = \hat{m}_{\theta\alpha}(z)^{d(1-\theta)} \geq \left(\frac{1-\theta}{2}\right)^{d(1-\theta)} \frac{1}{\max\{1,|\alpha|\}^d}$$

and the latter is larger the $(2\max\{1,|\alpha|\})^{-d}$ if and only if $(\xi/2)^{\xi} > \frac{1}{2}$, that is, $\xi < \frac{1}{2}$.

We now prove Proposition 4.2. Let K be a number field containing β, τ and let $v \in M_K$. The case $m = 0$ follows directly from property 1.1.6 of the height function so one may assume $\theta = m/n > 0$. First consider $n > m$; then

$$\frac{\max\{1, |\tau^n + \beta\tau^m|_v\}}{\max\{1, |\tau|_v^n\}} = \frac{\max\{1, |z|_v^{\theta/(1-\theta)}|z + \beta|_v\}}{\max\{1, |z|_v^{1/(1-\theta)}\}}$$

for $z = \tau^{n-m}$. We apply Lemmas 4.5 and 4.6, raise to the d_vth power, take the product over all elements of M_K and then the $[K : \mathbf{Q}]$th root to prove the first inequality of (4.2.3). The second inequality follows immediately from the fact that $1/(1 - \theta) \leq n$.

Finally if $n = m$ and $\beta \neq -1$ then standard height properties imply

$$\frac{H(P(\tau))}{H(\tau)^n} \geq \frac{1}{H(1 + \beta)} \geq \frac{1}{2H(\beta)}.$$

□

To prove Proposition 4.3 we will apply diophantine approximation to Proposition 4.2. This idea came up in a private correspondence Bombieri-Masser-Zannier June 2002.

Set $\theta = m/n$. The case $m = 0$ can be dismissed as trivial so we may assume $\theta \in (0, 1)$. We would like to apply diophantine approximation to θ to create a new polynomial with small degree.

Recall that for any $Q > 1$ there exists a pair $p, q \in \mathbf{Z}$ such that

$$|\theta q - p| \leq \frac{1}{Q} \quad \text{and} \quad 0 < q < Q.$$

For a reference see [**Cas57**] page 1. Because $\theta \in (0, 1)$ one has

$$\theta q - p \leq \frac{1}{Q} \quad \text{so} \quad p \geq \theta q - \frac{1}{Q} > \theta q - 1 > -1$$

thus $p \geq 0$; and furthermore

$$p - q\theta \leq \frac{1}{Q} \quad \text{so} \quad p \leq \frac{1}{Q} + q\theta < 1 + q\theta < 1 + q$$

which implies $p \leq q$.

We choose any $u \in \overline{\mathbf{Q}}^*$ with $u^q = \tau^n$. We set $k = mq - np$ and choose a $\beta \in \overline{\mathbf{Q}}^*$ with $\beta^q = \zeta\tau^k$. Note that $(\beta u^p)^q = \beta^q u^{pq} = \zeta\tau^{k+np} = \zeta\tau^{mq}$, so $\tau^m = \eta\beta u^p$ for some root of unity η. Now

$$P(\tau) = \tau^n + \zeta\tau^m = u^q + \xi\beta u^p$$

for a root of unity $\xi = \zeta\eta$. If $p = q$ and $\xi\beta = -1$ then $P(\tau) = 0$ so $\tau^{n-m} = -\zeta$ so τ is a root of unity and the required result follows trivially. If $p < q$ or $\xi\beta \neq -1$ we can apply Proposition 4.2 to get

$$\begin{aligned}\frac{H(P(\tau))}{H(\tau)^n} = \frac{H(u^q + \xi\beta u^p)}{H(u)^q} &\geq \frac{1}{2qH(\xi\beta)^q} \geq \frac{1}{2QH(\beta)^q} \\ &= \frac{1}{2QH(\tau)^{|k|}} \geq \frac{1}{2QH(\tau)^{n/Q}}\end{aligned}$$

because $|k| = n|\theta q - p|$. Therefore one has

$$H(P(\tau)) \geq \frac{H(\tau^n)^{1-1/Q}}{2Q} \tag{4.3.10}$$

for each $Q > 1$. Because the lower bound is continuous $Q = 1$ is also allowable. Optimization of the right hand side leads to the choice

$$Q_0 = \max\{1, \log H(\tau^n)\} \geq 1.$$

Inserting Q_0 into (4.3.10) gives the bound

$$\begin{aligned} H(P(\tau)) &\geq \begin{cases} H(\tau^n)(2e \log H(\tau^n))^{-1} & : \quad \text{if} \quad H(\tau^n) > e \\ 1/2 & : \quad \text{otherwise} \end{cases} \\ &\geq \frac{H(\tau^n)}{2e \max\{1, \log H(\tau^n)\}} \end{aligned}$$

as desired. □

To prove the Corollary we note that the bound $h(\sigma + \sigma^\phi) - h(\sigma) \leq \log 2$ has already been shown in section 2. For the corresponding lower bound we note that the case $\phi = 1$ is covered by elementary height properties; so assume $\phi < 1$. Choose $m, n \in \mathbf{Z}$ with $\phi = m/n$; then there is $\tau \in \overline{\mathbf{Q}}^*$ and ζ a root of unity with $\sigma = \tau^n$ and $\sigma^\phi = \zeta\tau^m$. We apply Proposition 4.3 to get

$$h(\sigma + \sigma^\phi) - h(\sigma) \geq -1 - \log 2 - \log \max\{1, h(\sigma)\},$$

thus concluding the proof. □

4. Proofs of Theorems 4.1 and 4.2

Before we prove Theorems 4.1 and 4.2 we need a simple lemma.

LEMMA 4.7. *Say $x, y \in \overline{\mathbf{Q}}^*$ with $H(x) \geq H(y)$ and $x^r y^s = 1$ for integers r and s such that $rs < 0$, then $H(x, y) = H(x)$. Furthermore, if $H(y) > 1$, then $|s| \geq |r|$.*

PROOF. Let K be a number field containing x and y. If $H(y) = 1$, then $|y|_v = 1$ for all $v \in M_K$ and the lemma follows. From now on we assume $H(y) > 1$. By inverting $x^r y^s = 1$ we may also assume $r > 0$. Therefore $s < 0$ and $-r/s > 0$. By functional properties of the height function we have $H(x) = H(y)^{-s/r} \geq H(y)$ and so $-s \geq r$ because $H(y) > 1$. The second assertion of the lemma follows. To prove the first we note that for any $v \in M_K$ we have $\max\{1, |x|_v, |y|_v\} = \max\{1, |x|_v, |x|_v^{-r/s}\} = \max\{1, |x|_v\}$ since $-r/s \in (0, 1]$. These local equalities imply $H(x, y) = H(x)$. □

The proof of Theorem 4.1 is a simple task with the help of Proposition 4.1. Let for example $x^r y^s = 1$ with $r, s \in \mathbf{Z}$ not both zero and say $H(x) \geq H(y)$. We apply Proposition 4.1 to $\pm 1 = x^r(x - \alpha)^s$ using $H(\pm 1) = 1$. If $rs \geq 0$ we have $H(y) \leq H(x, y) \leq 2H(\alpha)$, and the theorem follows. So say $rs < 0$, then $H(x, y) = H(x)$ by Lemma 4.7 and furthermore $H(x) \leq 2H(\alpha)^2$ by Proposition 4.1. It remains to show that $H(y) \leq 2H(\alpha)$, and for this we may clearly assume $H(y) > 1$. By Lemma 4.7 we

conclude $|s| \geq |r|$. Finally we apply Proposition 4.1 to $y^s(y-\alpha)^r = \pm 1$ and conclude $H(y) \leq 2H(\alpha)$. □

We now prove Theorem 4.2. Let for example $x^r y^s = 1$ with $r, s \in \mathbf{Z}$ not both zero and assume $r \geq 0$ and $H(x) \geq H(y)$. If $rs \geq 0$ the same argument as in the proof of Theorem 4.1 gives $H(x,y) \leq 2H(\alpha)$, and of course this estimate is better than our assertion. Hence let us assume $rs < 0$, which implies $H(x,y) = H(x)$. Hence it suffices to show (4.1.5) with $H(x,y)$ replaced by $H(x)$.

If $H(y) = 1$, then $H(x) = H(\alpha - y) \leq 2H(\alpha)$ by elementary height inequalities. So say $H(y) > 1$, then Lemma 4.7 implies $|s| \geq |r|$. We may even assume $|s| > |r|$, for if equality were to hold, then $H(x) = H(y) \leq 2H(\alpha)$ by Theorem 4.1. We set $n = |s|$ and $m = r$; then there exists $\tau \in \overline{\mathbf{Q}}^*$ with $x = \tau^n$ and $y = \zeta\tau^m$ and for some root of unity ζ. Hence we have $P(\tau) = \alpha$ with $P = T^n + \zeta T^m$. Now Proposition 4.3 implies

$$H(\alpha) \geq \varphi(H(x)) \quad \text{with} \quad \varphi(z) = \frac{z}{2e\max\{1, \log z\}}, \tag{4.4.1}$$

φ being understood as a continuous map on $[1, \infty)$. Note that φ is increasing which can be easily verified by restricting it to $[1, e]$ and $[e, \infty)$. Theorem 4.2 follows if

$$\varphi(z_0) \geq H(\alpha) \quad \text{with} \quad z_0 = 14H(\alpha)\log(3H(\alpha)) \tag{4.4.2}$$

holds. Indeed (4.4.1) combined with (4.4.2) leads to $\varphi(z_0) \geq \varphi(H(x))$ and further to $z_0 \geq H(x)$ because φ is increasing.

Because $z_0 > e$ the inequality (4.4.2) is equivalent to

$$\frac{7}{e}\log(3w) - \log(14w) - \log\log(3w) \geq 0 \quad \text{with} \quad w = H(\alpha) \tag{4.4.3}$$

which certainly holds for $w = 1$. The derivative of the left-hand side of (4.4.3) with respect to w is

$$\frac{1}{w}\left(\frac{7}{e} - 1 - \frac{1}{\log(3w)}\right) \geq \frac{1}{w}\left(\frac{7}{e} - 2\right)$$

if $w \geq 1$. The right-hand side is positive for every $w \geq 1$ so we may conclude that (4.4.3) holds for every α thus completing the proof. □

5. Dependent solutions with large height

Choose a large integer q, and define

$$x = \left(\frac{q}{q-1}\right)^n, \quad y = -\left(\frac{q}{q-1}\right)^{n-1} \quad \text{with} \quad n = [q\log q],$$

so that $H(x) = q^n > q^{n-1} = H(y)$. Then $x + y = \alpha$ with $\alpha = q^{n-1}/(q-1)^n$. Now $H(\alpha) = \max\{q^{n-1}, (q-1)^n\}$ so

$$\lim_{q\to\infty} \frac{\log H(\alpha)}{n\log q} = \lim_{q\to\infty} \max\left\{\frac{n-1}{n}, \frac{\log(q-1)}{\log q}\right\} = 1 \tag{4.5.1}$$

and one clearly has

$$\lim_{q\to\infty} \frac{n\log q}{q(\log q)^2} = 1. \tag{4.5.2}$$

We multiply (4.5.1) and (4.5.2) to get

$$\log H(\alpha) = q(\log q)^2\kappa(q) \text{ with } \lim_{q\to\infty} \kappa(q) = 1.$$

Taking the logarithm gives $\log\log H(\alpha) = \log q + 2\log\log q + \log\kappa(q)$ which leads us to

$$\lim_{q\to\infty} \frac{\log\log H(\alpha)}{\log q} = 1. \tag{4.5.3}$$

We want to show that the quotient $H(x)/H(\alpha)$ is large, so we will evaluate the limit of

$$q^{-1}\frac{H(x)}{H(\alpha)} = \min\{1, q^{-1}(1-q^{-1})^{-n}\} \tag{4.5.4}$$

for $q\to\infty$. Now

$$\begin{aligned}
&\log(q^{-1}(1-q^{-1})^{-n}) \\
&\quad = -\log q - n\log\left(1-q^{-1}\right) = -\log q + n\left(q^{-1}+O(q^{-2})\right) \\
&\quad = -\log q + (q\log q + O(1))\left(q^{-1}+O(q^{-2})\right) = O\left(\frac{\log q}{q}\right)
\end{aligned}$$

for $q\to\infty$. By inserting this expression into (4.5.4) one obtains

$$\lim_{q\to\infty} q^{-1}\frac{H(x)}{H(\alpha)} = 1. \tag{4.5.5}$$

Finally, by combining (4.5.1), (4.5.2), (4.5.3), and (4.5.5) we conclude

$$\lim_{q\to\infty} \frac{H(x)}{\frac{H(\alpha)\log H(\alpha)}{(\log\log H(\alpha))^2}} = 1.$$

□

6. Proof of Theorem 4.4

The next lemma will take care of some special cases of Theorem 4.4.

LEMMA 4.8. *Let $\alpha = n^\phi \geq 2$ with $n\in\mathbf{N}$, ϕ a positive rational and $x,y\in\overline{\mathbf{Q}}^*$ and $x+y=\alpha$.*

(i) If $r,s\in\mathbf{Z}$ are not both zero with $rs\geq 0$ and $x^ry^s=1$, then

$$H(x,y) \leq \frac{3}{2}H(\alpha).$$

(ii) If $y=\zeta x$ for some root of unity ζ, then

$$H(x,y) \leq \sqrt{2\sqrt{3}}H(\alpha).$$

PROOF. We can and will assume $H(x) \geq H(y)$. Let K be a number field containing x and y. For part (i) we assume $r \geq 0$, and if $r = 0$ also $s > 0$. Lemmas 4.1 and 4.2 applied to $f = T^r(T-\alpha)^s$ and $\tau = x$ give

$$\prod_{v\in M_K} \frac{1}{\max\{1,|x|_v,|x-\alpha|\}^{d_v d}} \geq \prod_{v|\infty} \frac{1}{(1+|\alpha|_v)^{d_v d}} \prod_{v\nmid\infty} \frac{1}{\max\{1,|\alpha|_v\}^{d_v d}}$$

with $d = r + s$. For a finite place v we have $|\alpha|_v \leq 1$ and therefore $H(x,y) \leq 1 + \alpha \leq \frac{3}{2}\alpha \leq \frac{3}{2}H(\alpha)$.

Now to part (ii): We note that $H(x,y) = H(x)$ by Lemma 4.7. We have $x(1+\zeta) = \alpha$ and by hypothesis $\zeta \neq -1$; elementary height properties lead to $H(x) = H((1+\zeta)^{-1}\alpha) \leq H((1+\zeta)^{-1})H(\alpha) = H(1+\zeta)H(\alpha) \leq \sqrt{2\sqrt{3}}H(\alpha)$ by Lemma 2.2 in chapter 2 if $\zeta \neq 1$. So now assume $\zeta = 1$, then $x = y = \alpha/2$ and so

$$\begin{aligned} H(x,y)^{[K:\mathbf{Q}]} = H(x)^{[K:\mathbf{Q}]} &= \prod_{v\in M_K} \max\{1,|\alpha/2|_v^{d_v}\} \\ &= \prod_{v|\infty} |\alpha/2|_v^{d_v} \prod_{v\nmid\infty} \max\{1,|\alpha/2|_v^{d_v}\} \leq (\alpha/2)^{[K:\mathbf{Q}]} \prod_{v\nmid\infty} |2|_v^{-d_v}, \end{aligned}$$

which implies $H(x,y) \leq H(\alpha)$. □

LEMMA 4.9. *Let $\alpha = n^\phi \geq 2$ with $n \in \mathbf{N}$, ϕ a positive rational and let $x, y \in \overline{\mathbf{Q}}^*$ with $x+y = \alpha$ and $x^r = y^t$ where $0 < r < t$ are rational integers. Define $\lambda = t/r$ and let K be a number field containing x, y. Then x is an algebraic integer and furthermore:*

(i) If v is a finite place of K with $|x|_v < 1$, then $|x|_v = |y|_v^\lambda = |\alpha|_v^\lambda$.

(ii) One has

$$H(x) = \prod_{\substack{v\nmid\infty \\ |x|_v<1}} \max\{1,|\alpha|_v^{-1}\}^{d_v\lambda/[K:\mathbf{Q}]} \leq \alpha^\lambda. \tag{4.6.1}$$

(iii) Let $\epsilon \in [0,1)$ and $\delta \in (1,2]$ be such that

$$\lambda \geq 1 + \frac{\log 2}{\log\alpha}(1+\epsilon) \quad \text{and} \quad (\delta-1)(1-\delta^{-1})^{\frac{\log\alpha}{(1+\epsilon)\log 2}}\alpha \geq 1. \tag{4.6.2}$$

Then

$$|x|_v \leq \delta|\alpha|_v \text{ for all } v \mid \infty \quad \text{and} \quad H(x) \leq \delta H(\alpha). \tag{4.6.3}$$

PROOF. Note that x is an algebraic integer. Indeed α is an algebraic integer and x is a zero of the monic polynomial $(-1)^t(\alpha - T)^t - (-1)^t T^r \in \mathcal{O}_K[T]$.

Let v be as in part (i); then $x^r = y^t$ implies $|y|_v > |x|_v$ and so because of $x + y = \alpha$ we have $|\alpha|_v = |y|_v = |x|_v^{r/t} < 1$.

For part (ii) recall the definition (4.3.4) of $m = m_f$ with f defined as in Proposition 4.1 with $s = -t$. For any finite place $v \in M_K$ one has

$$m(x) = \max\{|x|_v^r, |x-\alpha|_v^t\} = |x|_v^r \leq 1. \tag{4.6.4}$$

If on the other hand v is an infinite place one has $|\alpha|_v = |n|_v^{\phi} = \alpha \geq 2$ so clearly $|x|_v \geq 1$ and therefore

$$m(x) = |x|_v^{r-t}. \tag{4.6.5}$$

Using (4.6.4), (4.6.5) and applying the product formula, we conclude

$$\prod_{v \in M_K} m(x)^{d_v} = \prod_{v \nmid \infty} |x|_v^{d_v r} \prod_{v \mid \infty} |x|_v^{d_v(r-t)} = \prod_{v \nmid \infty} |x|_v^{d_v t} = \prod_{\substack{v \nmid \infty \\ |x|_v < 1}} |x|_v^{d_v t}. \tag{4.6.6}$$

Then by applying part (i) to (4.6.6) and using (4.3.3) with $f(x) = \pm 1$ we get

$$\begin{aligned} \frac{1}{H(x)^{[K:\mathbf{Q}]t}} = \prod_{v \in M_K} m(x)^{d_v} &= \prod_{\substack{v \nmid \infty \\ |x|_v < 1}} |\alpha|_v^{d_v \lambda t} = \prod_{\substack{v \nmid \infty \\ |x|_v < 1}} \max\{1, |\alpha|_v^{-1}\}^{-d_v \lambda t} \\ &\geq H(\alpha^{-1})^{-[K:\mathbf{Q}]\lambda t}, \end{aligned}$$

from which (ii) follows at once.

To prove part (iii) assume (4.6.2) holds. Note that the statement on the left-hand side of (4.6.3) implies the inequality on the right-hand side because x is an algebraic integer. We will prove the left-hand side of (4.6.3) by contradiction. Assume $|x|_v > \delta |\alpha|_v$ for some $v \mid \infty$. Then $|x - \alpha|_v > |x|_v(1 - \delta^{-1})$, and so $|x|_v^r = |x - \alpha|_v^t > (|x|_v(1 - \delta^{-1}))^t$. The first inequality in (4.6.2) implies

$$\delta\alpha < |x|_v < (1 - \delta^{-1})^{-\frac{\lambda}{\lambda - 1}} \leq (1 - \delta^{-1})^{-(1 + \frac{\log \alpha}{(1+\epsilon)\log 2})}$$

which contradicts the second inequality in (4.6.2). □

We can now prove Theorem 4.4. Let for example $x^r y^s = 1$ with $r, s \in \mathbf{Z}$ not both zero and assume $r \geq 0, H(x) \geq H(y)$. Fix a number field K with $x, y \in K$. We will begin by proving the inequality (4.1.6). If $rs \geq 0$ or $r = -s$ or y is a root of unity, then Lemma 4.8 implies (4.1.6). So assume $rs < 0$ and $-s \neq r$ and y not a root of unity. Then $H(x, y) = H(x)$ and $-s > r$ by Lemma 4.7. Hence it suffices to prove $H(x) \leq 2H(\alpha)$.

For brevity we will set $t = -s$ and $\lambda = t/r$. If $\lambda \leq 1 + \log 2 / \log \alpha$ then (4.6.1) in Lemma 4.9 gives

$$H(x) \leq \alpha^{\lambda} \leq \alpha^{1 + \log 2/\log \alpha} = 2H(\alpha). \tag{4.6.7}$$

On the other hand if $\lambda > 1 + \log 2 / \log \alpha$, there exists an $\epsilon > 0$ such that the first inequality in (4.6.2) holds. Now the second inequality in (4.6.2) holds strictly when $\delta = 2$, and so it must continue to hold for some $\delta < 2$. Hence the left-hand side of (4.6.3) holds with some $\delta < 2$ and therefore $H(x) < 2H(\alpha)$

Finally we prove that $H(x, y) = 2H(\alpha)$ holds if and only if α is a rational power of 2 and $x = 2\alpha$ or $y = 2\alpha$. The "if" part is trivial. For the "only if" part let us assume $H(x, y) = 2H(\alpha)$ and $H(x) \geq H(y)$. As above we use Lemma 4.8 to reduce to the case $rs < 0$ and $t = -s > r$. Let again $\lambda = t/r$. By Lemma 4.7 we have $H(x, y) = H(x) = 2H(\alpha)$. We have already showed that $\lambda > 1 + \log 2 / \log \alpha$ implies

$H(x) < 2H(\alpha)$. But if $\lambda < 1 + \log 2/\log \alpha$ then (4.6.7) also implies $H(x) < 2H(\alpha)$. So we must have

$$\lambda = 1 + \log 2/\log \alpha \tag{4.6.8}$$

and thus α is a rational power of 2. Because of (4.6.8) the choice $\epsilon = 0$, $\delta = 2$ satisfies the hypothesis of Lemma 4.9(iii). We conclude $|x|_v \le 2|\alpha|_v$ for infinite places v. As α, x are algebraic integers we even have $|x|_v = |2\alpha|_v$ for all infinite places v. Note that (4.6.8) implies

$$\alpha^\lambda = 2\alpha. \tag{4.6.9}$$

If v is a finite place with $|2|_v < 1$ then $|x|_v < 1$; indeed we must have equality in (4.6.1). So Lemma 4.9(i) gives $|x|_v = |\alpha|_v^\lambda$, and because of (4.6.9) we conclude $|x|_v = |2\alpha|_v$. But this last equation holds for any finite place: indeed if $|2|_v = 1$ then $|x|_v = |\alpha|_v = 1$ because of Lemma 4.9(i). Hence $|x|_v = |2\alpha|_v$ for all places finite or infinite: therefore $x = 2\alpha\xi$ for a root of unity ξ. Let v be a infinite place. Then the equality $|x|_v^r = |x-\alpha|_v^t$ and (4.6.9) imply $|2\xi - 1|_v = 1$. For any $z, w \in K$ we have the equality $|z+w|_v^2 + |z-w|_v^2 = 2|z|_v^2 + 2|w|_v^2$; We take $z = \xi - 1$ and $w = \xi$ to conclude $\xi = 1$. Thus $x = 2\alpha$. □

7. Proof of Theorem 4.5

We will start with a lemma concerning an elementary estimate.

LEMMA 4.10. *Let $\phi \in \mathbf{N}$ and $1 + \frac{2}{3}\phi^{-1} \le \lambda \le 1 + \frac{4}{3}\phi^{-1}$, then*

$$\frac{1}{2}\max\{1, \lambda(1+\phi^{-1})^{-1}\}\log(2^{2\phi+1} + 2\max\{1, (\lambda/2)^{\frac{1}{1-\lambda}}\}) \le \log(1.98 \cdot 2^\phi). \tag{4.7.1}$$

PROOF. Let $g(\lambda)$ denote the left-hand side of (4.7.1). The map $\lambda \to (\lambda/2)^{1/(1-\lambda)}$ decreases for $1 < \lambda < 4$. If $\phi = 1$ the lemma follows easily by considering the cases $\lambda \le 2$ and $\lambda > 2$. We will therefore assume $\phi \ge 2$. Since $(1 + 1/w)^w$ increases for $w \ge 1$ we conclude

$$(\lambda/2)^{\frac{1}{1-\lambda}} \le \frac{2^{3\phi/2}}{(1+\frac{2}{3\phi})^{3\phi/2}} \le \frac{3^3}{2^6} \cdot 2^{3\phi/2}.$$

If x and y are positive then $\log(x+y) \le \frac{x}{y} + \log y$, thus

$$\begin{aligned} g(\lambda) &\le \frac{1}{2}\max\{1, \frac{\lambda}{1+\phi^{-1}}\}\log(\frac{3^3}{2^6}2^{3\phi/2+1} + 2^{2\phi+1}) \\ &\le \max\{1, \frac{\lambda}{1+\phi^{-1}}\}(\frac{3^3}{2^8} + \frac{1}{2}\log 2^{2\phi+1}). \end{aligned}$$

If $\lambda \le 1 + \phi^{-1}$, then

$$g(\lambda) \le \frac{3^3}{2^8} + \frac{1}{2}\log 2 + \log 2^\phi$$

and we are done. On the other hand if $\lambda > 1 + \phi^{-1}$, then

$$g(\lambda) \leq \frac{\phi + \frac{4}{3}}{\phi + 1}(\frac{3^3}{2^8} + \frac{1}{2}\log 2^{2\phi+1}) = \frac{\phi(\frac{5}{6}\log 2 + \frac{3^3}{2^8}) + \frac{3^2}{2^6} + \frac{2}{3}\log 2}{\phi + 1} + \log 2^{\phi}$$
$$< \frac{5}{6}\log 2 + \frac{3^3}{2^8} + \log 2^{\phi}.$$

□

We proceed as follows: let x and $y = \alpha - x$ be multiplicative dependent with $H(y) \leq H(x) < 2H(\alpha)$, and let $P \in \mathbf{Z}[T]$ be the minimal polynomial of x. Then we will show $P(2\alpha) \neq 0$. With the help of the finite places lying above 2 we will even find a lower bound for $|P(2\alpha)|$ in terms of $H(x)$. More precisely:

LEMMA 4.11. *Let $\alpha = 2^{\phi}$ for $\phi \in \mathbf{N}$ and let $x, y \in \overline{\mathbf{Q}}^*$ with $x + y = \alpha$ and $x^r = y^t$ where $0 < r < t$ are rational integers. Define $\lambda = t/r$ and $k(w) = 2\alpha^2 + 2w^{2/\lambda} - w^2$. If $H(x) < 2H(\alpha)$, then*

$$\log H(x) \leq \frac{1}{2}\max\{1, \lambda(1 + \phi^{-1})^{-1}\}\log\sup_{w \geq 1} k(w).$$

PROOF. Fix a finite galois extension $K/\mathbf{Q}$ with galois group G such that $x \in K$. Let $P \in \mathbf{Z}[T]$ be the minimal polynomial of x; then because x is an algebraic integer we have

$$P^{[K:\mathbf{Q}(x)]} = \prod_{\sigma \in G}(T - \sigma x). \tag{4.7.2}$$

Let v be any finite place of K extending the 2-adic absolute value i.e. $v \mid 2$, then

$$|P(2\alpha)|_v^{[K:\mathbf{Q}(x)]} = \prod_{\sigma \in G}|2\alpha - \sigma x|_v \leq \prod_{\sigma \in G}\max\{|2|_v^{1+\phi}, |\sigma x|_v\}.$$

Recall from Lemma 4.9(i) that if $v' \nmid \infty$ then $|x|_{v'} < 1$ implies $|x|_{v'} = |\alpha|_{v'}^{\lambda}$. Let g be the number of finite places of K lying over 2 and g' be the number of finite places v' with $|2|_{v'}, |x|_{v'} < 1$. Now G acts transitively on the set of all finite places lying over 2 (Proposition 11 page 12 of [**Lan94**]) therefore each stabilizer has cardinality $[K : \mathbf{Q}]/g$, so

$$|P(2\alpha)|_v^{[K:\mathbf{Q}(x)]} \leq \prod_{\sigma \in G}\max\{|2|_{\sigma^{-1}v}^{1+\phi}, |x|_{\sigma^{-1}v}\} = \prod_{v' \mid 2}\max\{|2|_{v'}^{1+\phi}, |x|_{v'}\}^{[K:\mathbf{Q}]/g}$$
$$= \prod_{\substack{v' \mid 2 \\ |x|_{v'} < 1}}\max\{|2|_{v'}^{1+\phi}, |\alpha|_{v'}^{\lambda}\}^{[K:\mathbf{Q}]/g} = |2|_v^{[K:\mathbf{Q}]\frac{g'}{g}\min\{1+\phi, \phi\lambda\}}.$$

Recall that $\alpha \in \mathbf{Q}$, so if $P(2\alpha) = 0$ then $x = 2\alpha$. In this case $H(x) = 2H(\alpha)$, which contradicts our hypothesis. We conclude $P(2\alpha) \neq 0$. Now $P(2\alpha) \in \mathbf{Z}$ and the product

formula imply

$$|P(2\alpha)|^{[K:\mathbf{Q}]} = \prod_{v|\infty} |P(2\alpha)|_v^{d_v} = \prod_{v\nmid\infty} |P(2\alpha)|_v^{-d_v} \geq \prod_{v|2} |P(2\alpha)|_v^{-d_v} \tag{4.7.3}$$
$$\geq 2^{[K:\mathbf{Q}]\frac{g'}{g}\min\{1+\phi,\phi\lambda\}\deg(P)}.$$

Because $K/\mathbf{Q}$ is galois we have $d_v g = [K:\mathbf{Q}]$ for any $v \mid 2$. So Lemma 4.9(ii) yields

$$H(x) = \prod_{\substack{v\nmid\infty \\ |x|_v<1,|2|_v<1}} \max\{1, |\alpha|_v^{-1}\}^{d_v\lambda/[K:\mathbf{Q}]} = 2^{\frac{g'}{g}\phi\lambda}.$$

This equation inserted into (4.7.3) gives

$$|P(2\alpha)|^{[K:\mathbf{Q}]} \geq H(x)^{[K:\mathbf{Q}]\min\{1,\lambda^{-1}(1+\phi^{-1})\}\deg(P)}. \tag{4.7.4}$$

We continue by bounding the left hand side of (4.7.4) from above. Let v be an infinite place. Recall that for $z_1, z_2 \in K$ we have $|z_1+z_2|_v^2+|z_1-z_2|_v^2 = 2|z_1|_v^2+2|z_2|_v^2$. Let $\sigma \in G$, take $z_1 = \sigma\alpha$, $z_2 = \sigma\alpha - \sigma x$ in the previous equation and recall $|\sigma\alpha - \sigma x|_v = |\sigma x|_v^{1/\lambda}$ to conclude

$$|2\alpha - \sigma x|_v^2 = |2\sigma\alpha - \sigma x|_v^2 = 2|\sigma\alpha|_v^2 + 2|\sigma x|_v^{2/\lambda} - |\sigma x|_v^2 = k(|\sigma x|_v). \tag{4.7.5}$$

Note that $|\sigma x|_v \geq 1$; indeed if $|\sigma x|_v < 1$ then $|\sigma y|_v = |\sigma x|_v^{1/\lambda} < 1$ and so $|\alpha|_v < 2$, a contradiction. Apply (4.7.5) to (4.7.2) to get

$$|P(2\alpha)|^{[K:\mathbf{Q}(x)]} = \prod_{\sigma\in G} |2\alpha - \sigma x|_v = \prod_{\sigma\in G} k(|\sigma x|_v)^{1/2} \leq (\sup_{w\geq 1} k(w))^{[K:\mathbf{Q}]/2}.$$

We combine the previous upper bound with the lower bound in (4.7.4) to conclude the proof. □

We can now prove Theorem 4.5. Assume $H(x) \geq H(y)$ and $H(x,y) < 2H(\alpha)$, furthermore let $r, s \in \mathbf{Z}$ not both zero with $r \geq 0$ such that $x^r y^s = 1$. With the help of Lemma 4.8 we reduce to the case $t = -s > r > 0$ and $H(x,y) = H(x)$ as we did in the beginning of the proof of Theorem 4.4. It suffices to show $H(x) \leq 1.98H(\alpha)$. Define $\lambda = t/r$; we will split up into cases.

The two first cases $\lambda < 1 + \frac{2}{3}\phi^{-1}$ and $\lambda > 1 + \frac{4}{3}\phi^{-1}$ are effectively covered in Lemma 4.9. Indeed part (ii) applied to the first case gives $H(x) \leq \alpha^\lambda \leq 2^{2/3}H(\alpha)$. For the second case set $\epsilon = 1/3$ and $\delta = 1.9$. This choice clearly satisfies the left-hand side of (4.6.2). Because

$$(\delta-1)(1-\delta^{-1})^{\frac{\phi}{1+\epsilon}}\alpha = \frac{9}{10}\left(2(\frac{9}{19})^{3/4}\right)^\phi \geq 1$$

the right-hand side of (4.6.2) holds as well. Thus $H(x) \leq 1.9H(\alpha)$.

Now assume $1 + \frac{2}{3}\phi^{-1} \leq \lambda \leq 1 + \frac{4}{3}\phi^{-1}$. Let k be the function defined in Lemma 4.11. Elementary calculus yields

$$\sup_{w\geq 1} k(w) = k(w_0) \quad \text{with} \quad w_0 = \max\{1, (\lambda/2)^{\frac{\lambda}{2(1-\lambda)}}\}.$$

Apply the previous inequality to Lemma 4.11 and obtain the bound

$$\begin{aligned}\log H(x) &\leq \frac{1}{2}\max\{1, \lambda(1+\phi^{-1})^{-1}\}\log k(w_0)\\ &\leq \frac{1}{2}\max\{1, \lambda(1+\phi^{-1})^{-1}\}\log(2\alpha^2 + 2\max\{1, (\lambda/2)^{\frac{1}{1-\lambda}}\}).\end{aligned}$$

Now Lemma 4.10 implies $H(x) \leq 1.98H(\alpha)$. □

8. Proof of Theorem 4.6

The next lemma proves the first inequality in Theorem 4.6.

LEMMA 4.12. *Let K be a number field with* $\operatorname{rank}\mathcal{O}_K^* = 1$ *and $\alpha \in \mathbf{Q}^*$; then $|S_K(\alpha)| \leq 292$.*

PROOF. If $\alpha \notin \mathbf{Z}$ then $S_K(\alpha)$ is empty so assume $\alpha \in \mathbf{Z}\setminus\{0\}$. Let ω be the number of roots of unity in K, η a fundamental unit, R the regulator, and D the degree of K.

If $(x,y) \in S_K(\alpha)$ there exist $r, s \in \mathbf{Z}$ not both zero such that $x^r y^s = 1$; we shall furthermore assume that $r \geq 0$ and $H(x) \geq H(y)$ (so we will have to multiply the number of solutions under this hypothesis by 2 to get a bound for the total number of solutions). If $H(y) = 1$ then y is a root of unity, hence in this case we can choose $r = 0$, $s > 0$. So in all cases one may assume $|s| \geq r$.

First assume $\alpha \neq \pm 1$. Let Δ denote the set of all infinite places v for which $|x|_v \geq 1$ and let $\delta \in [0,1]$ with $\delta D = \sum_{v\in\Delta} d_v$. Note that in the case $s < 0$ one has $\delta = 1$ because $|\alpha| \geq 2$. Because x is a unit the height is given by

$$H(x) = \prod_{v\in\Delta} |x|_v^{d_v/D}. \tag{4.8.1}$$

Let $v \in \Delta$; if $|x|_v < |\alpha|_v/2$, then $|x|_v^{-r/s} = |\alpha - x|_v \geq |\alpha|_v - |x|_v > |x|_v$ which leads to $r/s < -1$ so $|r/s| > 1$ contradicting $|r| \leq |s|$. We conclude $|x|_v \geq |\alpha|_v/2 = \max\{1, |\alpha|_v\}/2$ for all $v \in \Delta$. So by (4.8.1)

$$H(x) \geq 2^{-\delta} H(\alpha)^{\delta}. \tag{4.8.2}$$

We now deduce a corresponding upper bound for $H(x)$. If $s \geq 0$, then Lemma 4.2 with $f = T^r(T-\alpha)^s$ and $\tau = x$ applied to (4.8.1) leads to

$$H(x) \leq 2^{\delta} H(\alpha)^{\delta}. \tag{4.8.3}$$

If on the other hand $s < 0$, then (4.8.3) also holds because of Theorem 4.4 and $\delta = 1$.

With the bounds (4.8.2) and (4.8.3) we can apply a gap principle. There exists a unique $a \in \mathbf{Z}$ and a root of unity ζ such that $x = \zeta\eta^a$. We apply height functional properties (cf. chapter 1) and the bounds to see that $|a|$ lies in an interval of length $\frac{\delta \log 4}{\log H(\eta)}$. Hence there are at most $2(\log 4/\log H(\eta) + 1)$ possibilities for a. Clearly

this estimate remains valid for $\alpha = \pm 1$ because in this case Theorem 4.1 implies $0 \leq \log H(x) \leq \log 2$. We also note that $R = D \log H(\eta)$ and therefore

$$|S_K(\alpha)| \leq 4\omega \left(\frac{D \log 4}{R} + 1 \right).$$

Elementary considerations lead to $D \leq 4$ and $\omega \leq 12$. Now a result of Friedman ([**Fri89**] Theorem B) which states $R/\omega \geq 0.09058$ completes the proof. □

To prove Theorem 4.6 let $(x, y) \in S_{\mathscr{F}}(\alpha)$. The proof splits up into two cases:

(i) There exist $n \in \mathbf{Z}$ and $\zeta \in \mathscr{F}$ a root of unity such that $x = \zeta y^n$ with $-2 \leq n \leq 2$ or $y = \zeta x^n$ with $n = 0, \pm 2$.
(ii) Otherwise.

First assume case (i). Elementary arguments show that there are 24 roots of unity ζ in $\mathscr{F}$. For each such ζ, substituting $y = \alpha - x$ in (i) gives eight polynomial equations in x of degree at most 3; thus the number of x is at most $24 \cdot 8 \cdot 3 = 576$.

Now assume case (i) does not hold. Set $K = \mathbf{Q}(x) = \mathbf{Q}(x, y)$ and $D = [K : \mathbf{Q}]$. Because x and y are not roots of unity we have $\operatorname{rank} \mathcal{O}_K^* = 1$. Let η be a fundamental unit of K. We claim that $H(\eta)^D \leq 4$ will complete the proof. Indeed assuming this inequality a well-known argument bounds the number of units with degree d and height at most $4^{1/d}$ by $2d \prod_{k=1}^{d-1} (2 \cdot \binom{d}{k} \cdot 4 + 1)$. Take the sum over this expression for $2 \leq d \leq 4$; thus there are at most 430706 possibilities for η. There are 6 roots of unity in $\mathscr{F}$ such that one of these generates the group of roots of unity in K. Let ζ be such a root of unity, then $K = \mathbf{Q}(\eta, \zeta)$. Now the Theorem follows from Lemma 4.12 applied to the field K.

We will now show $H(\eta)^D \leq 4$. There are $a, b \in \mathbf{Z}$ and roots of unity ζ, ξ such that $x = \zeta \eta^a$ and $y = \xi \eta^b$. Let σ_1, σ_2 be two distinct non-conjugate embeddings of K into $\mathbf{C}$. These correspond to the two infinite places of K. Define $d_i = 1$ if $\sigma_i(K) \subset \mathbf{R}$ and $d_i = 2$ otherwise. We may assume $d_1 \geq d_2$. Now $|\sigma_i(\eta)| \neq 1$ for $i = 1, 2$ by Kronecker's Theorem, so by replacing η with η^{-1} if necessary we may assume $|\sigma_1(\eta)| > 1$. Let l_i be a logarithm of $\sigma_i(\eta)$. Note that $D \log H(\eta) = d_1 \log |\sigma_1(\eta)| = d_1 \mathrm{Re}(l_1)$, hence it suffices to show $\mathrm{Re}(l_1) \leq \log 2$. The equality $|\sigma_1(\eta)|^{d_1} |\sigma_2(\eta)|^{d_2} = 1$ implies

$$d_1 \mathrm{Re}(l_1) + d_2 \mathrm{Re}(l_2) = 0. \tag{4.8.4}$$

Because $\alpha \in \mathbf{Q}$ one has $\sigma_1(x) - \sigma_2(x) = \sigma_2(y) - \sigma_1(y)$. Apply (4.8.4) to get

$$|e^{q(a)|a|\mathrm{Re}(l_1)+\gamma_1 i} - e^{-q(-a)|a|\mathrm{Re}(l_1)+\gamma_2 i}| = |e^{q(b)|b|\mathrm{Re}(l_1)+\gamma_3 i} - e^{-q(-b)|b|\mathrm{Re}(l_1)+\gamma_4 i}| \tag{4.8.5}$$

where the γ_i are real numbers and

$$q(w) = \begin{cases} 1 & : w \geq 0, \\ \frac{d_1}{d_2} & : w < 0 \end{cases}$$

is an integer. Now define $k = q(b)|b| - q(a)|a| \in \mathbf{Z}$, then $k \neq 0$ because otherwise we would be in case (i). So first assume $k \geq 1$. Apply the triangle inequality to (4.8.5) and

use $\mathrm{Re}(l_1) \geq 0$ to conclude

$$e^{q(a)|a|\mathrm{Re}(l_1)} + 1 \geq e^{(q(a)|a|+k)\mathrm{Re}(l_1)} - 1. \tag{4.8.6}$$

Note that we have $|a| \geq 1$, or else we would be back in case (i). Now (4.8.6) and $q(a) \geq 1$ imply

$$e^{\mathrm{Re}(l_1)} - 2e^{-\mathrm{Re}(l_1)} - 1 \leq 0. \tag{4.8.7}$$

The left-hand side of (4.8.7) increases in $\mathrm{Re}(l_1)$. Substitute $\log 2$ for $\mathrm{Re}(l_1)$ to get the bound $\mathrm{Re}(l_1) \leq \log 2$. Now assume $k \leq -1$; then similar arguments as above involving the triangle inequality and this time $|b| \geq 1$ also lead to (4.8.7) and thus again $\mathrm{Re}(l_1) \leq \log 2$. □

CHAPTER 5

More on heights and some algebraic geometry

The purpose of this chapter is to review some facts from algebraic geometry and their application to heights. Furthermore, we give a glimpse of some results concerning heights which will be used in the second part of the thesis. The two subsequent chapters will make use of the concepts introduced here. Therefore we do not strive for utmost generality; we present the material in a form useful for our applications. Also, we give no details of the constructions and almost no proofs.

1. Some preliminaries

By a *variety* we mean a Zariski open subset of a possibly reducible projective algebraic variety defined over a field K. For simplicity we assume throughout the whole chapter that K is algebraically closed and of characteristic zero unless stated otherwise.

Throughout this section X and Y will denote irreducible varieties defined over K. All morphisms are considered to be defined over this field. By abuse of notation we will often identify varieties with their set of K-rational points.

We start off with the Fibre Dimension Theorem, a result we will often apply in the next chapters.

THEOREM 5.1. *Let $f : X \to Y$ be a morphism of varieties. Then for $p \in X$ all irreducible components of $f^{-1}(f(p))$ have dimension greater or equal to $\dim X - \dim Y$. Furthermore, there exists a Zariski open, dense subset $V \subset Y$ such that if $q \in V$ then all irreducible components of $f^{-1}(q)$ have dimension $\dim X - \dim Y$.*

PROOF. Follows easily from the first theorem on page 228 of [**Dan94**]. □

The next results is the Theorem on Semi-continuity of Fibre Dimension; it will be used in chapter 6.

THEOREM 5.2. *Let $f : X \to Y$ be a morphism of varieties and say $k \in \mathbf{Z}$. The set*

$$\{p \in X;\ \dim_p f^{-1}(f(p)) \geq k\}$$

is Zariski closed in X.

PROOF. This is the second theorem on page 228 of [**Dan94**]. □

The following result will also be useful:

THEOREM 5.3. *Let $f : X \to Y$ be a dominant morphism of varieties. Then $f(X)$ contains an Zariski open, dense subset of Y.*

PROOF. Follows from the theorem on page 219 of [**Dan94**]. □

2. Chow forms and higher dimensional heights

We continue using the notation of the previous section but with the exception that we now assume $X \subset \mathbf{P}^n$ to be an irreducible closed projective variety defined over K.

The goal of this section is to define the height of a variety defined over $\overline{\mathbf{Q}}$. The most naive approach would be to define this height somehow in terms of the height of defining equations for X. A *hypersurface*, i.e. a subvariety of codimension 1, is the zero set of one homogeneous polynomial unique up to multiplication by a non-zero scalar. Our naive approach works in this situation because of the product formula. Unfortunately if $0 \leq \dim X \leq n-2$ there is no unique system of polynomials which define X. The solution is to take a suitable height of the Chow form of X. We follow Philippon's definition in [**Phi95**]. We begin by defining the Chow form of X.

Let $\dim X = r$. For $0 \leq i \leq r$, let $U_i = (U_{i0}, \ldots, U_{in})$ where the U_{ij} are independent variables. By Proposition 2.2 on page 99 and the discussion on pages 100-101 of [**GKZ94**] there exist $d \in \mathbf{N}$ and $\mathcal{F}_X \in K[U_{ij}; 0 \leq i \leq r, 0 \leq j \leq n]$ with $\mathcal{F}_X$ multi-homogeneous of degree d in each U_i such that

$$\mathcal{F}_X(u_0, \ldots, u_r) = 0 \quad \text{if and only if} \quad \{p \in X;\ u_0(p) = \cdots = u_r(p) = 0\} \neq \emptyset.$$

Here $u_i \in K^{n+1}$ are identified with homogeneous polynomials of degree 1. The form $\mathcal{F}_X$ is called *Chow form* or sometimes also *Cayley form* or *Elimination form*. The integer d is just the geometric degree $\deg(X)$ which equals the cardinality of the intersection of X with a generic linear subvariety of $\mathbf{P}^n$ of dimension $n-r$. The Chow form $\mathcal{F}_X$ is determined uniquely up to multiplication by a non-zero scalar. A converse statement holds: the Chow form $\mathcal{F}_X$ determines a set of homogeneous polynomials whose set of common zeros is precisely X. Indeed, for $0 \leq i \leq r$ let $s^{(i)} = (s^{(i)}_{jk})$ denote skew-symmetric $(n+1) \times (n+1)$ matrices with coefficients $s^{(i)}_{jk}$ $(0 \leq j, k \leq n)$ which are algebraically independent for $j < k$. Then

$$\mathcal{F}_X((X_0, \ldots, X_n)s^{(0)}, \ldots, (X_0, \ldots, X_n)s^{(r)}) \tag{5.2.1}$$

considered as a polynomial in the variables $X_0, \ldots, X_n$ and coefficients in the field $K(s^{(i)}_{jk})$ vanishes precisely on X. Then (5.2.1) provides us with a finite number of homogeneous polynomials of degree $d(r+1)$ whose set of common zeros in $\mathbf{P}^n$ is X.

For the rest of the section we assume that $K = \overline{\mathbf{Q}}$. The coefficients of $\mathcal{F}_X$ are contained in a number field F. The height of X is essentially the height of the polynomial $\mathcal{F}_X$ but with a different norm at the infinite places. This deviation is motivated by Arakelov theory and the definition of the so-called Faltings height in (3.1.2.2) of [**BGS94**]. The norms are chosen such that the resulting height equals Philippon's height in [**Phi95**].

More precisely, let $v \in M_F$. If v is a finite place, then we set $||\mathcal{F}_X||_v = |\mathcal{F}_X|_v$, where $|\cdot|_v$ is defined as in section 2, chapter 1. If v is an infinite place and f is the image of

$\mathcal{F}_X$ under an embedding $F \to \mathbf{C}$ corresponding to v, we define

$$\log ||\mathcal{F}_X||_v = \int_{S^{r+1}} \log |f(u)| \sigma^{r+1} + (r+1)\deg(X) \sum_{i=1}^{n} \frac{1}{2i},$$

where $S \subset \mathbf{C}^{n+1}$ is the unit sphere and σ is the rotation invariant measure on S of total measure 1. We define the *height* of X as

$$h_V(X) = \frac{1}{[F:\mathbf{Q}]} \sum_{v \in M_F} d_v \log ||\mathcal{F}_X||_v,$$

where d_v are the local degrees defined in chapter 1 section 1. This quantity does not depend on the number field F containing the coefficients of $\mathcal{F}_X$.

At this point we must warn the reader about a possible ambiguity: a point $p \in \mathbf{P}^n(\overline{\mathbf{Q}})$ already has a well-defined height from chapter 1 which does not necessarily equal the height $h_V(\{p\})$ of the corresponding variety. For this reason, we use two different symbols to distinguish the two heights.

There is a natural embedding $\iota : \mathbf{G}_m^n \hookrightarrow \mathbf{P}^n$ given by taking a point $(p_1, \ldots, p_n)$ to $[1 : p_1 : \cdots : p_n]$. This embedding will be used throughout the rest of the thesis. Having defined the degree and height of a subvariety of $\mathbf{P}^n$ it makes sense to speak of the degree and height of a subvariety of $\mathbf{G}_m^n$ using this embedding. Concretely, if $Z \subset \mathbf{G}_m^n$ is Zariski dense in an irreducible subvariety defined over $\overline{\mathbf{Q}}$, we define $h_V(Z) = h_V(\overline{\iota(Z)})$ and $\deg(Z) = \deg(\overline{\iota(Z)})$.

Say $Y \subset \mathbf{P}^n$ is an irreducible projective variety. Bézout's Theorem provides an upper bound for the degrees of the irreducible components of the set theoretic intersection $X \cap Y$.

THEOREM 5.4. *Let $Z_1, \ldots, Z_g$ be the distinct irreducible components of $X \cap Y$, then*

$$\sum_{i=1}^{g} \deg(Z_i) \leq \deg(X)\deg(Y).$$

PROOF. See [**Ful84**] example 8.4.6 on page 148 or Main Theorem of [**Vog84**] on page 60. □

From the point of view of Arakelov theory the height of a variety should be understood as an arithmetic version of the degree. The arithmetic analog of Bézout's Theorem is then naturally called the Arithmetic Bézout Theorem. If Y is defined over $\overline{\mathbf{Q}}$ it bounds the height of the irreducible components in $X \cap Y$ from above in terms of heights and degrees of X and Y:

THEOREM 5.5. *Let $Z_1, \ldots, Z_g$ be the distinct irreducible components of $X \cap Y$, then*

$$\sum_{i=1}^{g} h_V(Z_i) \leq \deg(X)h_V(Y) + \deg(Y)h_V(X) + c\deg(X)\deg(Y),$$

where c is a positive effective constant which depends only on n.

PROOF. See [**Phi95**] Theorem 3. □

It is clear that Theorems 5.4 and 5.5 hold if X and Y are subvarieties of $\mathbf{G}_m^n$. We will apply Bézout's Theorem in chapter 7 and the Arithmetic Bézout Theorem in chapter 6.

3. Normalized height and essential minimum of a variety

All varieties in this section are defined over $\overline{\mathbf{Q}}$. Let $X \subset \mathbf{G}_m^n$ denote an irreducible closed subvariety. We use h to denote the absolute logarithmic Weil height restricted from $\mathbf{P}^n(\overline{\mathbf{Q}})$ to $\mathbf{G}_m^n(\overline{\mathbf{Q}})$.

In [**Phi95**] Philippon described how to normalize the height of a subvariety of an abelian variety. This construction is a generalization to higher dimension of Tate's normalization of the height of a point on an abelian variety. In [**DP99**] David and Philippon did the same for subvarieties of $\mathbf{G}_m^n$. For an integer $m \in \mathbf{Z}$, let $[m] : \mathbf{G}_m^n \to \mathbf{G}_m^n$ denote the homomorphism which takes p to p^m. In [**DP99**] the authors showed that the limit

$$\hat{h}(X) = \lim_{m\to\infty} \frac{h_V([m]X)\deg(X)}{m\deg([m]X)} \tag{5.3.1}$$

exists. Its value $\hat{h}(X)$ is called the *normalized height* of X. In Proposition 2.1 of [**DP99**] it is showed that

$$|\hat{h}(X) - h_V(X)| \leq \frac{7}{2}(\dim X + 1)\deg(X)\log(n+1). \tag{5.3.2}$$

Furthermore, if $p \in \mathbf{G}_m^n(\overline{\mathbf{Q}})$, then $\hat{h}(\{p\}) = h(p)$ where the left-hand side is the absolute logarithmic Weil height. In view of (5.3.2) we conclude $|h_V(\{p\}) - h(p)| \leq \frac{7}{2}\log(n+1)$. If X is the translate of an algebraic subgroup by a torsion point, then by Proposition 2.1 of [**DP99**] we have $\hat{h}(X) = 0$. We will see a converse statement further down.

We define the *essential minimum* of X as

$$\hat{\mu}^{\text{ess}}(X) = \sup_{Z\subsetneq X} \inf\{h(p);\ p \in (X\backslash Z)(\overline{\mathbf{Q}})\},$$

here Z runs over all proper Zariski closed subsets of X. Equivalently one could take the supremum over all Zariski closed subsets of X with codimension 1 in X. Another equivalent definition is

$$\hat{\mu}^{\text{ess}}(X) = \inf\{\delta \in \mathbf{R};\ \{p \in X(\overline{\mathbf{Q}});\ h(p) \leq \delta\} \text{ is Zariski dense in } X\}.$$

For example if $X = \{p\}$ is a point, then $\hat{\mu}^{\text{ess}}(X) = h(p)$. The case where X is a curve is more interesting. Then for any $\delta < \hat{\mu}^{\text{ess}}(X)$ there are at most finitely many points $p \in X(\overline{\mathbf{Q}})$ with $h(p) \leq \delta$.

If X is the translate of an algebraic subgroup of $\mathbf{G}_m^n$ by a torsion point then $\hat{\mu}^{\text{ess}}(X) = 0$. Hence both normalized height and essential minimum of such an X vanish. The following theorem of Zhang explains this observation:

THEOREM 5.6 (Zhang). *We have the inequality*

$$\hat{\mu}^{\mathrm{ess}}(X) \leq \frac{\hat{h}(X)}{\deg(X)} \leq (1+\dim X)\hat{\mu}^{\mathrm{ess}}(X). \tag{5.3.3}$$

PROOF. Follows easily from Theorem 5.2 in [**Zha95a**] and (5.3.1). Compare also Theorem 1.10 of [**Zha95b**]. □

In fact Zhang proved a stronger lower bound: the left side of (5.3.3) can be replaced by the sum of $\hat{\mu}^{\mathrm{ess}}(X)$ with the higher successive minima. The higher successive minima are non-negative; we will not define or use them here though.

It is an important problem to determine all X with $\hat{\mu}^{\mathrm{ess}}(X) = 0$ or equivalently $\hat{h}(X) = 0$. The Bogomolov Conjecture for $\mathbf{G}_m^n$ states that if $\hat{\mu}^{\mathrm{ess}}(X) = 0$ then X is the translate of an algebraic subgroup by a torsion point. Bogomolov originally conjectured that if C is a non-singular projective curve of genus at least 2 embedded in its Jacobian, then there exists an $\epsilon > 0$ such that there are only finitely many points in $C(\overline{\mathbf{Q}})$ with Néron-Tate height at most ϵ. This conjecture was proved by Ullmo in [**Ull98**] and a more general version was proved by Zhang in [**Zha98**].

In Theorem 6.2 of [**Zha95a**] Zhang proved Bogomolov's Conjecture in the algebraic torus. Zhang's Theorem 6.2 in [**Zha95a**] together with Theorem 5.6 above can be seen as a higher dimension generalization of Kronecker's Theorem. Zhang also showed that if X^* is X deprived of all its subvarieties that are translates of algebraic subgroups by torsion points, then $X^* \subset X$ is Zariski open and $\inf_{p \in X^*(\overline{\mathbf{Q}})} h(p) > 0$. Bombieri and Zannier gave an elementary and effective proof of these two statements in [**BZ95**]. They used a slightly different height in their paper which can easily be compared with our height. Finding lower bounds for the infimum (if X^* is non-empty) is often called the Lehmer Problem. Lower bounds involve the field of definition of X, $\deg(X)$, and n. See for example Corollary 1.3 in [**AD01**].

If we assume that X is not equal to the translate of a proper algebraic subgroup of $\mathbf{G}_m^n$, then in [**AD03**] Amoroso and David proved a lower bound for $\hat{\mu}^{\mathrm{ess}}(X)$ which only depends on $\deg(X)$ and n but not on the field of definition of X. The dependence in the degree is best-possible up to a logarithmic factor. Actually, instead of the degree Amoroso and David used the more subtle obstruction index. If V is a proper subvariety of $\mathbf{P}^n$ the *obstruction index* is defined as

$$\omega(V) = \inf_{V \subset Z \subsetneq \mathbf{P}^n} \{\deg(Z)\},$$

here Z runs over all Zariski closed subsets of $\mathbf{P}^n$ that contain V and whose irreducible components have dimension $n-1$. As Amoroso and David pointed out in their article, a result of Chardin implies

$$\omega(V) \leq n \deg(V)^{1/\operatorname{codim} V}. \tag{5.3.4}$$

Using the embedding $\iota : \mathbf{G}_m^n \hookrightarrow \mathbf{P}^n$ we may consider the obstruction index of subvarieties $X \subset \mathbf{G}_m^n$.

THEOREM 5.7 (Amoroso, David). *Assume $X \subsetneq \mathbf{G}_m^n$ is a proper irreducible closed subvariety of codimension k defined over $\overline{\mathbf{Q}}$. If X is not contained in the translate of a proper algebraic subgroup of $\mathbf{G}_m^n$, then*

$$\hat{\mu}^{\mathrm{ess}}(X) \geq \frac{c}{\omega(X)}(\log(3\omega(X)))^{-\lambda(k)}$$

where $c > 0$ depends only on n and $\lambda(k) = (9(3k)^{k+1})^k$.

PROOF. This is Theorem 1.4 of [**AD03**]. □

If X is as in Theorem 5.7 then (5.3.4) implies

$$\hat{\mu}^{\mathrm{ess}}(X) \geq \frac{c'}{\deg(X)^{1/\operatorname{codim} X}}(\log(3\deg(X)))^{-\lambda(k)}, \tag{5.3.5}$$

where $c' > 0$ depends only on n.

Theorem 5.7 will be applied in chapter 7 to obtain the following result: if $X \subset \mathbf{G}_m^n$ is an irreducible curve not contained in the translate of an algebraic subgroup then there are at most finitely many points in $X(\overline{\mathbf{Q}})$ that lie in an algebraic subgroup of dimension $n-2$ "up to an ϵ". In the proof of this result it will be essential that the lower bound in (5.3.5) is best possible up to a power of a logarithm.

CHAPTER 6

Intersecting varieties with small subgroups

Let $X \subset \mathbf{G}_m^n$ be an irreducible closed subvariety. In this chapter we will bound the height of algebraic points on X that are in a certain sense close to the union of all algebraic subgroup of $\mathbf{G}_m^n$ of dimension $m < n/\dim X$. These results give a qualitative generalization of results from previous chapters. The bounds we obtain are effective and will be expressed as functions of the height of X, the degree of X, and n. We will use these height bounds to derive finiteness results if the points in question actually lie on the union of all algebraic subgroup of dimension m. In chapter 7 we will prove finiteness results for points close to such a subgroup.

1. Introduction

Unless stated otherwise, all varieties in this chapter are assume to be defined over $\overline{\mathbf{Q}}$. Fix an irreducible closed subvariety $X \subset \mathbf{G}_m^n$. In this chapter and the next a *coset* is the translate of an algebraic subgroup of $\mathbf{G}_m^n$. We do not require a coset to be irreducible.

Let H be a subset of $\mathbf{G}_m^n(\overline{\mathbf{Q}})$ and let $\epsilon \geq 0$, we define the "cone" around H:

$$C(H, \epsilon) = \{ab;\ a \in H,\, b \in \mathbf{G}_m^n(\overline{\mathbf{Q}}),\, h(b) \leq \epsilon(1 + h(a))\},$$

here $h(\cdot)$ is the absolute logarithmic Weil height on $\mathbf{G}_m^n(\overline{\mathbf{Q}})$ defined in chapter 1.

Let X^{o} be the complement in X of the union of all positive dimensional cosets contained in X. The work of Bombieri and Zannier (Theorem 1, [**BZ95**]) shows that X^{o} is Zariski open in X. Let Γ be a finitely generated subgroup of $\mathbf{G}_m^n(\overline{\mathbf{Q}})$ and let $\overline{\Gamma}$ be the group of $p \in \mathbf{G}_m^n(\overline{\mathbf{Q}})$ such that $p^k \in \Gamma$ for a positive integer k. Evertse showed in Theorem 1.7 of [**Eve02**] that $X^{\mathrm{o}} \cap C(\overline{\Gamma}, \epsilon)$ is finite for some positive ϵ. The definition of X^{o} makes sense if $\mathbf{G}_m^n$ is replaced by any semi-abelian variety A. One can define a reasonable height on $A(\overline{\mathbf{Q}})$ if A is defined over $\overline{\mathbf{Q}}$. In this context Poonen ([**Poo99**]) proved that if A is isogenous to the product of an abelian variety with $\mathbf{G}_m^n$ ($n = 0$ is allowed) then there exists a positive ϵ with the following property: there are at most finitely many points $p = ab \in X^{\mathrm{o}}$ where $a \in \overline{\Gamma}$ and where b has height at most ϵ. In fact Poonen's Corollary 9 is slightly stronger. In [**Rém03**] Theorem 1.2 Rémond proves the finiteness of $X^{\mathrm{o}} \cap C(\overline{\Gamma}, \epsilon)$ for any semi-abelian variety A defined over $\overline{\mathbf{Q}}$ and for some $\epsilon > 0$.

For $\epsilon = 0$ we describe a result by Bombieri, Masser, and Zannier in [**BMZ99**] which concerns the intersection of X with the union of all algebraic subgroups of fixed dimension. When $0 \leq m \leq n$ is an integer, we define $\mathcal{H}_m^n$ to be the union of all algebraic subgroups of $\mathbf{G}_m^n$ with dimension less or equal to m. If not stated otherwise, we usually

identify $\mathcal{H}_m^n$ with the set of its algebraic points. To avoid trivialities we set $\mathcal{H}_m^n = \emptyset$ for negative m. In [**BMZ99**] the authors showed that if $X \subset \mathbf{G}_m^n$ is an irreducible algebraic curve not contained in a proper coset, then $X \cap \mathcal{H}_m^n$ is a set of bounded height for $m = n-1$ and finite for $m = n-2$.

Motivated by the results described in the previous two paragraphs we pose the problem of finding boundedness of height results for the intersection $X^{\mathrm{oa}} \cap C(\mathcal{H}_m^n, \epsilon)$ for a small $\epsilon > 0$ and appropriate m. Here X is not necessarily a curve and $X^{\mathrm{oa}} \subset X$ will be defined further down.

In some instances height bounds for $X^{\mathrm{oa}} \cap C(\mathcal{H}_m^n, \epsilon)$ imply the finiteness of this set as we will see later on. We will discuss finiteness results with $\epsilon = 0$ here and with $\epsilon > 0$ in chapter 7. By Kronecker's Theorem the set $C(\mathcal{H}_m^n, 0)$ equals $\mathcal{H}_m^n$. Hence Bombieri, Masser, and Zannier's finiteness result in [**BMZ99**] covers the case $\epsilon = 0$ for curves not contained in a proper coset.

We start off with a short review of what else is known about intersections $X \cap C(\mathcal{H}_m^n, \epsilon)$ from the point of view of boundedness of height and finiteness.

If we have $m = 0$ then $\mathcal{H}_0^n$ is the set of torsion points of $\mathbf{G}_m^n$ by Kronecker's Theorem. Certainly the height is bounded on $X \cap \mathcal{H}_0^n$, but finiteness questions are already interesting: one is lead to the study of torsion points contained in X. For example, a special case of a result of Laurent in [**Lau84**] shows that the intersection $X \cap \mathcal{H}_0^n$ is contained in a finite union of translates of algebraic subgroups by torsion points. Furthermore, these translates are contained in X.

Next, if we allow ϵ to be positive but still with $m = 0$, then $C(\mathcal{H}_0^n, \epsilon)$ is just the set of algebraic points $\mathbf{G}_m^n$ with height at most ϵ. So again bounding the height on $X \cap C(\mathcal{H}_0^n, \epsilon)$ is trivial. Finiteness (and non-density) questions of this set are related to the Bogomolov Conjecture already discussed to some extent in section 3 of chapter 5.

Up to now most work concerning the intersection of varieties with $C(\mathcal{H}_m^n, \epsilon)$ where positive dimensional algebraic subgroups are involved assumed $\epsilon = 0$. Hence we continue our review considering only the case $\epsilon = 0$.

For X of any dimension, Bombieri and Zannier showed in Theorem 1 ([**Zan00**] page 524) that $X^{\mathrm{o}} \cap \mathcal{H}_1^n$ is a set of bounded height. The subgroup dimension in this theorem is not believed to be best-possible for $\dim X \leq n-2$. This belief is supported by [**BMZ99**] where the subgroups have dimension $n-1$, i.e. equal to the codimension of the variety in question. For $2 \leq \dim X \leq n-2$ only one boundedness of height result is known where $m > 1$: if X is a plane, then in the preprint [**BMZ04**] Bombieri, Masser, and Zannier proved that $X^{\mathrm{oa}} \cap \mathcal{H}_{n-2}^n$ has bounded height and $X^{\mathrm{oa}} \cap \mathcal{H}_{n-3}^n$ is finite. Furthermore, they proved that for planes, the still undefined set $X^{\mathrm{oa}} \subset X$ is Zariski open.

In [**BMZ06b**] Bombieri, Masser, and Zannier proved that $X^{\mathrm{ta}} \cap \mathcal{H}_1^n$ is finite if $\dim X = n-2$ and in this case X^{ta} is Zariski open in X. The definition of X^{ta} will also be given below.

As already mentioned, if X is a curve which is not contained in a proper coset, then $X \cap \mathcal{H}_{n-2}^n$ is finite. Say now that X is contained in a proper coset, then Bombieri, Masser, and Zannier showed that the height on $X \cap \mathcal{H}_{n-1}^n$ is necessarily unbounded. Nevertheless it is conjectured in [**BMZ99**] that $X \cap \mathcal{H}_{n-2}^n$ is finite if X is allowed to be

contained in a proper coset but not in a proper algebraic subgroup. Recently, Maurin has announced a proof of this conjecture in [**Mau06**].

We will prove the following simple proposition:

PROPOSITION 6.1. *Let $X \subset \mathbf{G}_m^n$ be an irreducible closed subvariety defined over $\overline{\mathbf{Q}}$ of dimension r with $1 \leq r \leq n$ and let $U \subset X$ be Zariski open and dense. Then the set $U \cap \mathcal{H}_{n-r+1}^n$ has unbounded height and $U \cap \mathcal{H}_{n-r}^n$ is infinite.*

Therefore to prove a reasonable boundedness of height result on intersections of X with $\mathcal{H}_m^n$ one needs to assume $m \leq n - \dim X$. And to prove finiteness results one needs $m \leq n - \dim X - 1$. There is a series of conjectures stated by Bombieri, Masser, and Zannier which state that one can take m to be equal to these upper bounds and expect boundedness of height or finiteness respectively, at least when restricting to geometrically motivated subsets X^{oa} of X which we now finally define.

An irreducible closed subvariety $Y \subset X$ is called *X-anomalous* if there exists a coset $H \subset \mathbf{G}_m^n$ such that $Y \subset H$ and

$$\dim Y \geq \max\{1, \dim X + \dim H - n + 1\}. \tag{6.1.1}$$

If $Y \subset X$ is X-anomalous and there can be no confusion regarding X we will simply call Y *anomalous*. We define X^{oa} to be X deprived of all its anomalous subvarieties. If H is an algebraic subgroup we call Y *X-torsion anomalous*, or just *torsion anomalous*. We define X^{ta} to be X deprived of all its torsion anomalous subvarieties. These definitions continue to make sense when all varieties in question are defined over $\mathbf{C}$.

Informally speaking, a positive dimensional $Y \subset X$ is anomalous if it is contained in a component of the intersection of X with a coset whose dimension is smaller than expected.

We present some special cases where X^{oa} and X^{ta} can easily be determined: if X is a curve, then $X^{\mathrm{oa}} = X$ if X is not contained in a proper coset, and $X^{\mathrm{oa}} = \emptyset$ otherwise. Similarly, $X^{\mathrm{ta}} = X$ if X is not contained in a proper algebraic subgroup, and $X^{\mathrm{ta}} = \emptyset$ otherwise. If X is a hypersurface, i.e. if $\dim X = n - 1$, then it is easy to see that $X^{\mathrm{oa}} = X^{\mathrm{o}}$.

CONJECTURE 6.1. *Let $X \subset \mathbf{G}_m^n$ be an irreducible closed subvariety defined over $\mathbf{C}$. Then $X^{\mathrm{oa}} \subset X$ is Zariski open.*

In the work [**BMZ06b**], Bombieri, Masser, and Zannier proved Conjecture 6.1 and even a "Structure Theorem" for anomalous subvarieties.

The conjectures of Bombieri, Masser, and Zannier are:

CONJECTURE 6.2. *(Bombieri, Masser, Zannier) Let $X \subset \mathbf{G}_m^n$ be an irreducible closed subvariety defined over $\overline{\mathbf{Q}}$. Then $X^{\mathrm{oa}} \cap \mathcal{H}_{n-\dim X}^n$ is a set of bounded height.*

CONJECTURE 6.3. *(Bombieri, Masser, Zannier) Let $X \subset \mathbf{G}_m^n$ be an irreducible closed subvariety defined over $\mathbf{C}$. Then $X^{\mathrm{ta}} \subset X$ is Zariski open and furthermore $X^{\mathrm{ta}} \cap \mathcal{H}_{n-\dim X-1}^n$ is finite.*

In Conjecture 6.3 one identifies $\mathcal{H}_m^n$ with the set of its complex points.

By the survey of known results above, Conjecture 6.2 has been proved for curves, hypersurfaces, and planes.

Pink and Zilber independently stated conjectures concerning general semi-abelian varieties which are related to the conjectures above. In [**Pin05b**], Pink conjectured:

CONJECTURE 6.4 (Pink, Zilber). *Let A be a semi-abelian variety defined over $\mathbf{C}$ and let $X \subset A$ be an irreducible closed subvariety also defined over $\mathbf{C}$ which is not contained in a proper algebraic subgroup of A. Then the set of points in $X(\mathbf{C})$ that are contained in an algebraic subgroup of A of codimension greater or equal to $\dim X + 1$ is not Zariski dense in X.*

In [**Zil02**] Zilber stated a conjecture which implies Conjecture 6.4.

Theorem 6.1, the main result of this chapter, goes in the direction of Conjecture 6.2 but with positive ϵ. The drawback is that the subgroups involved have dimension strictly less than $n/\dim X$. This strong restriction on the subgroup dimension allows us to prove Conjecture 6.2 only in some special cases.

To formulate Theorem 6.1 we need a refined version of X^{oa}: let t be an integer with $0 \leq t \leq n$, we define $X^{\mathrm{oa},t}$ to be X deprived of all its closed irreducible subvarieties Y that are contained in a coset H satisfying (6.1.1) and $\dim H \leq t$. Clearly we have

$$X = X^{\mathrm{oa},0} \supset X^{\mathrm{oa},1} \supset \cdots \supset X^{\mathrm{oa},n} = X^{\mathrm{oa}}. \tag{6.1.2}$$

And even

$$X^{\mathrm{oa},n-\dim X} = X^{\mathrm{oa}}. \tag{6.1.3}$$

Indeed by (6.1.2) it suffices to prove that if $p \in Y \subset pH$ where $Y \subset X$ and H is an algebraic subgroup satisfying (6.1.1), then $p \notin X^{\mathrm{oa},n-\dim X}$. If $\dim H \leq n - \dim X$, then we are done; hence say $\dim H > n - \dim X$. We will see in Proposition 6.2 that there exist linearly independent $u_1, \ldots, u_{n-\dim H} \in \mathbf{Z}^n$ such that $H = \{x \in \mathbf{G}_m^n,\ x^{u_1} = \cdots = x^{u_{n-\dim H}} = 1\}$. Here notation introduced in section 2 is used. We may choose $v_1, \ldots, v_{\dim H + \dim X - n} \in \mathbf{Z}^n$ such that the u_i, v_i are linearly independent. By Proposition 6.2 the v_i define an algebraic subgroup $H' \subset \mathbf{G}_m^n$ with $\dim H' = 2n - \dim H - \dim X$ and $\dim H \cap H' = n - \dim X$. Let Y' be an irreducible component of $Y \cap pH'$ containing p. By [**Ful84**] (§8.2, page 137) we have $\dim Y' \geq \dim Y + \dim H' - n$ and we use (6.1.1) to conclude $\dim Y' \geq 1$. Furthermore, $Y' \subset p(H \cap H')$ where the right side is a coset of dimension $n - \dim X$. We conclude $p \notin X^{\mathrm{oa},n-\dim X}$. Therefore (6.1.3) is established.

We are now ready to state Theorem 6.1.

THEOREM 6.1. *Let $X \subset \mathbf{G}_m^n$ be an irreducible closed subvariety defined over $\overline{\mathbf{Q}}$. Let s be an integer with $\dim X \leq s \leq n$ and let m be an integer with $m \cdot s < n$. Then there exists $\epsilon > 0$ which depends only on $\deg(X)$ and n such that if $p \in X^{\mathrm{oa},n-s} \cap C(\mathcal{H}_m^n, \epsilon)$ then*

$$h(p) \leq c(n) \cdot \deg(X)^{\frac{ms}{n-ms}} (\deg(X) + h_V(X)).$$

The constant $c(n)$ is effective and depends only on n.

We recall that degrees and heights of subvarieties of $\mathbf{G}_m^n$ are defined in chapter 5.

It is tempting to take $s = \dim X$ in Theorem 6.1, since then one can take m as large as possible and also $X^{\mathrm{oa},n-\dim X} = X^{\mathrm{oa}}$. If X is a curve or a hypersurface, then we may choose $m = n-1$ or $m = 1$ respectively. In these two cases the theorem implies Conjecture 6.2 "with an ϵ" and with explicit height bounds. Although Conjecture 6.2 has already been proved for such X, no explicit height bounds have appeared in literature.

The proof of Theorem 6.1 does not use the now proven fact (cf. [**BMZ06b**]) that $X^{\mathrm{oa}} \subset X$ is Zariski open. In fact if $s \geq \dim X$ the proof shows that the height is bounded above on $U \cap C(\mathcal{H}_m^n, \epsilon)$ for some non-empty Zariski open set $U \subset X$ with $X^{\mathrm{oa},n-s} \subset U$ as soon as the following hypothesis on X is satisfied.

(6.1.4) For any surjective homomorphism of algebraic groups $\varphi : \mathbf{G}_m^n \to \mathbf{G}_m^s$ one has $\dim \varphi(X) = \dim X$.

In Lemma 6.4 we will show that (6.1.4) follows from $X^{\mathrm{oa},n-s} \neq \emptyset$.

If $X \neq \mathbf{G}_m^n$ then taking $s = n-1$ also has interesting consequences. Indeed it is not difficult to see that $X^{\mathrm{oa},1} = X^{\mathrm{o}}$. Now Theorem 6.1 with $m = 1$ gives an explicit height bound for the set $X^{\mathrm{o}} \cap C(\mathcal{H}_1^n, \epsilon)$. We have therefore recovered an explicit version of Bombieri and Zannier's Theorem ([**Zan00**] Theorem 1, page 523) mentioned above. Their theorem also holds for $X = \mathbf{G}_m^n$, but then of course $X^{\mathrm{o}} = \emptyset$.

In [**BMZ04**] Bombieri, Masser, and Zannier show the following theorem:

THEOREM 6.2. *(Bombieri, Masser, Zannier) Let $X \subset \mathbf{G}_m^n$ be an irreducible closed subvariety defined over $\overline{\mathbf{Q}}$ of dimension r. If $B \in \mathbf{R}$, then*

$$\{p \in X^{\mathrm{ta}} \cap \mathcal{H}_{n-r-1}^n;\ h(p) \leq B\}$$

is finite.

If we combine Bombieri, Masser, and Zannier's result from [**BMZ04**] with Theorem 6.1 we get a finiteness result with $\epsilon = 0$.

COROLLARY 6.1. *Let $X \subset \mathbf{G}_m^n$ be an irreducible closed subvariety defined over $\overline{\mathbf{Q}}$ of dimension $r \geq 1$ and let m be an integer with $m < \min\{n/r, n-r\}$. Then $X^{\mathrm{oa}} \cap \mathcal{H}_m^n$ is finite.*

The proof is an immediate consequence of Theorems 6.1 (with $s = \dim X$) and 6.2, and the fact that $X^{\mathrm{oa}} \subset X^{\mathrm{ta}}$.

For example, the previous corollary implies the finiteness statement of Conjecture 6.3 in the cases of curves ($r = 1$), hypersurfaces ($r = n-1$), and also $r = n-2$ but always with X^{ta} replaced by X^{oa} and if everything is defined over $\overline{\mathbf{Q}}$.

There is one further situation where Corollary 6.1 implies finiteness with the correct subgroup dimension: namely surfaces in $\mathbf{G}_m^5$.

COROLLARY 6.2. *Let $X \subset \mathbf{G}_m^5$ be an irreducible closed algebraic surface defined over $\overline{\mathbf{Q}}$. Then $X^{\mathrm{oa}} \cap \mathcal{H}_2^5$ is finite.*

Of course this corollary follows from Corollary 6.1 by taking $n = 5$, $r = 2$, and $m = 2$.

We remark that we are showing the finiteness of $X^{\mathrm{oa}} \cap \mathcal{H}_2^5$ and not of the potentially larger set $X^{\mathrm{ta}} \cap \mathcal{H}_2^5$ which appears in Conjecture 6.3.

As we will see, proving a version of Theorem 6.1 with $\epsilon = 0$ is not essentially simpler than the proof of Theorem 6.1 itself. Hence in the problem of bounding the height we get the ϵ for free. Unfortunately the same cannot be said about the problem of proving finiteness in the style of Theorem 6.2. In fact we reserve the whole next chapter to prove a version of Theorem 6.2 "with an ϵ". Although the subgroup dimension m will often be less than $n - r - 1$.

2. Algebraic subgroups of $\mathbf{G}_m^n$

We let $c_1, \ldots, c_6$ denote positive constants which depend only on n.

As all our work is done in the torus $\mathbf{G}_m^n$ we introduce some notation which eases work in this and the next chapter. Let K be a field. Let $p = (p_1, \ldots, p_n) \in \mathbf{G}_m^n(K)$ and $u = (\alpha_1, \ldots, \alpha_n) \in \mathbf{Z}^n$, we set $p^u = p_1^{\alpha_1} \cdots p_n^{\alpha_n}$. Say $u_1, \ldots, u_m \in \mathbf{Z}^n$. If U is an $n \times m$ matrix with columns $u_1, \ldots, u_m \in \mathbf{Z}^n$ then we set $p^U = (p^{u_1}, \ldots, p^{u_m}) \in \mathbf{G}_m^m(K)$. For $q \in \mathbf{G}_m^n(K)$ we have $(pq)^U = p^U q^U$. If V is a matrix with m rows and integer coefficients, then $(p^U)^V = p^{UV}$. If $K = \mathbf{R}$ and all p_i are positive we will also allow exponent vectors or matrices with rational numbers as entries.

Let $p = (p_1, \ldots, p_n) \in \mathbf{G}_m^n(\overline{\mathbf{Q}})$ and say U is as above. By elementary height inequalities we have

$$h(p^U) \leq m \sum_{j=1}^{n} x_j h(p_j), \tag{6.2.1}$$

where x_j is the maximum of the absolute values of the elements of the jth row of U.

Finally we define the morphism of algebraic groups $\varphi_{(u_1,\ldots,u_m)} : \mathbf{G}_m^n \to \mathbf{G}_m^m$ by sending p to $(p^{u_1}, \ldots, p^{u_m})$.

For $u \in \mathbf{R}^n$ $|u|$ will always denote the euclidean norm of u. Furthermore, if $L \in \mathbf{R}[X_1, \ldots, X_n]$ is a linear form, then $|L|$ will denote the euclidean norm of the coefficient vector of L.

Let $\Lambda \subset \mathbf{Z}^n$ be a subgroup. Then we define

$$\mathcal{H}(\Lambda) = \{p \in \mathbf{G}_m^n;\ p^u = 1 \text{ for all } u \in \Lambda\}.$$

It is clear that $\mathcal{H}(\Lambda)$ is an algebraic subgroup of $\mathbf{G}_m^n$. We define $\det \Lambda$ to be the determinant of the subgroup Λ considered as a lattice in $\mathbf{R}^n$.

Algebraic subgroups of $\mathbf{G}_m^n$ and subgroups of $\mathbf{Z}^n$ are closely related:

PROPOSITION 6.2. *Let $\Lambda \subset \mathbf{Z}^n$ be a subgroup of rank r, then $\dim \mathcal{H}(\Lambda) = n - r$ and*

$$\deg(\mathcal{H}(\Lambda)) \leq c_1 \det \Lambda. \tag{6.2.2}$$

Furthermore, for any algebraic subgroup $H \subset \mathbf{G}_m^n$ there exists a subgroup $\Lambda \subset \mathbf{Z}^n$ with $H = \mathcal{H}(\Lambda)$.

PROOF. The equality $\dim \mathcal{H}(\Lambda) = n - r$ follows from Proposition 3.2.7 in [**BG06**]. The last statement of the assertion is Corollary 3.2.15 of [**BG06**].

Inequality (6.2.2) can be proved as follows: by Minkowski's Second Theorem there exist linearly independent $u_1, \dots, u_r \in \Lambda$ with $|u_1| \cdots |u_r| \leq c \det \Lambda$, here c depends only on n. Say the u_i generate the subgroup $\Lambda' \subset \mathbf{Z}^n$. Then by the first part of the proposition $\mathcal{H}(\Lambda)$ and $\mathcal{H}(\Lambda')$ have equal dimension. Since $\mathcal{H}(\Lambda) \subset \mathcal{H}(\Lambda')$ the irreducible components of $\mathcal{H}(\Lambda)$ are irreducible components of $\mathcal{H}(\Lambda')$. Now $x^{u_1} = \cdots = x^{u_r} = 1$ are defining equations for $\mathcal{H}(\Lambda')$. By multiplying these with suitable monomials we get polynomial equations. Inequality (6.2.2) then follows from Bézout's Theorem (Theorem 5.4). □

3. Some geometry of numbers

We start off with a result very much in the spirit of an argument given in the alternative proof of Theorem 1 in [**BMZ99**].

LEMMA 6.1. *Let $1 \leq m \leq n$ and let $a \in \mathcal{H}_m^n$. Then there exist linear forms $L_1, \dots, L_m \in \mathbf{R}[X_1, \dots, X_n]$ such that $|L_j| \leq 1$ and*

$$h(a^u) \leq c_2 \max_{1 \leq j \leq m} \{|L_j(u)|\} h(a)$$

for all $u \in \mathbf{Z}^n$.

PROOF. Let $a = (a_1, \dots, a_n) \in \mathcal{H}_m^n$; then by Proposition 6.2 we may assume that the a_i lie in a finitely generated subgroup of $\overline{\mathbf{Q}}^*$ of rank m. By a result of Schlickewei (Theorem 1.1 of [**Sch97**]) there exist multiplicatively independent elements $g_1, \dots, g_m \in \overline{\mathbf{Q}}^*$ and roots of unity $\zeta_1, \dots, \zeta_n$ such that

$$a_i = \zeta_i g_1^{v_{i1}} \cdots g_m^{v_{im}} \quad \text{for some} \quad v_{ij} \in \mathbf{Z} \tag{6.3.1}$$

and

$$h(g_1^{b_1} \cdots g_m^{b_m}) \geq c(|b_1| h(g_1) + \cdots + |b_m| h(g_m)) \tag{6.3.2}$$

for all $(b_1, \dots, b_m) \in \mathbf{Z}^m$; here $c > 0$ depends only on n.

We define $A = \max_{i,j}\{|v_{ij}| h(g_j)\}$. The assertion of the lemma obviously holds if all a_i are roots of unity. So let us assume that at least one v_{ij} is non-zero; then $A > 0$. For $1 \leq j \leq m$ we define the linear forms

$$\tilde{L}_j = v_{1j} X_1 + \cdots + v_{nj} X_n \quad \text{and} \quad L_j = \frac{h(g_j)}{\sqrt{n} A} \tilde{L}_j. \tag{6.3.3}$$

Clearly we have $|L_j| \leq 1$.

Let $u \in \mathbf{Z}^n$, then (6.3.1) and (6.3.3) imply

$$a^u = \xi g_1^{\tilde{L}_1(u)} \cdots g_m^{\tilde{L}_m(u)}$$

for some root of unity ξ. We apply standard height properties and (6.3.3) to conclude

$$\begin{aligned} h(a^u) &= h(g_1^{\tilde{L}_1(u)} \cdots g_m^{\tilde{L}_m(u)}) \\ &\le |\tilde{L}_1(u)|h(g_1) + \cdots + |\tilde{L}_m(u)|h(g_m) \\ &= \sqrt{n}A(|L_1(u)| + \cdots + |L_m(u)|) \\ &\le m\sqrt{n} \max_{1\le j\le m}\{|L_j(u)|\}A. \end{aligned} \tag{6.3.4}$$

We choose i_0 and j_0 such that $A = |v_{i_0 j_0}|h(g_{j_0})$. Then (6.3.1) and (6.3.2) imply

$$h(a) \ge h(a_{i_0}) \ge c|v_{i_0 j_0}|h(g_{j_0}) = cA.$$

We insert this inequality into (6.3.4) to complete the proof. □

Now we approximate zeros of linear forms with real coefficients by integral vectors with controlled coefficients. The next lemma will be used in this chapter and also in chapter 7. The proof uses Minkowski's Second Theorem.

LEMMA 6.2. *Let $1 \le m \le n$ and let $L_1, \dots, L_m \in \mathbf{R}[X_1, \dots, X_n]$ be linear forms with $|L_j| \le 1$. If $\rho \ge 1$, there exist $\lambda_1, \dots, \lambda_n$ with $0 < \lambda_1 \le \cdots \le \lambda_n$ and linearly independent $u_1, \dots, u_n \in \mathbf{Z}^n$ such that*

$$|u_k| \le \lambda_k, \quad |L_j(u_k)| \le \rho^{-1}\lambda_k, \text{ and } \lambda_1 \cdots \lambda_n \le c_3\rho^m. \tag{6.3.5}$$

PROOF. We set $\Lambda \subset \mathbf{R}^{m+n}$ to be the rank n lattice generated by the columns of the $(m+n) \times n$-matrix

$$A = \begin{bmatrix} \rho L_1 \\ \vdots \\ \rho L_m \\ E \end{bmatrix}$$

where E is the $n \times n$ unit matrix and the L_i are identified with the coefficient vectors in $\mathbf{R}^n$. By the Cauchy-Binet formula we have $\det \Lambda = (\det A^{\mathrm{t}}A)^{1/2} = (\sum_{A'}(\det A')^2)^{1/2}$ where the sum ranges over all $n \times n$ minors A' of A. By Hadamard's inequality we deduce $|\det A'| \le \rho^m$ and so

$$\det \Lambda \le \binom{n+m}{n}^{1/2} \rho^m.$$

Let $Q = \{x \in \mathbf{R}^{m+n};\ |x| \le 1\}$ be the unit ball and $V = \Lambda \otimes \mathbf{R} \subset \mathbf{R}^{m+n}$. Then V is an n-dimensional vector space and $\mathrm{vol}(Q \cap V) = \pi^{n/2}/\Gamma(1 + n/2)$. Here vol is the n-dimensional Lebesgue volume on V and Γ is the usual Gamma function. In particular $\mathrm{vol}(Q \cap V)$ depends only on n and not on V.

Let $\lambda_1, \dots, \lambda_n$ be the successive minima of Λ with respect to the convex, symmetric, and compact set $Q \cap V$. Minkowski's Second Theorem implies

$$\lambda_1 \cdots \lambda_n \le 2^n \frac{\det \Lambda}{\mathrm{vol}(Q \cap V)} \le c_3\rho^m \tag{6.3.6}$$

with $c_3 = 2^n \binom{n+m}{n}^{1/2} \Gamma(1+n/2)/\pi^{n/2}$. By definition there exist

$$v_k \in (\lambda_k Q) \cap \Lambda \text{ with } v_1, \ldots, v_n \text{ linearly independent.} \tag{6.3.7}$$

For $1 \leq k \leq n$ there are $u_k \in \mathbf{Z}^n$ with

$$v_k = (\rho L_1(u_k), \ldots, \rho L_m(u_k), u_k). \tag{6.3.8}$$

Clearly the $u_1, \ldots, u_n$ are also linearly independent. The first two inequalities in (6.3.5) follow from (6.3.7) and (6.3.8). The last one is just (6.3.6). □

LEMMA 6.3. *Let $1 \leq m \leq n$ and let $L_1, \ldots, L_m \in \mathbf{R}[X_1, \ldots, X_n]$ be linear forms with $|L_j| \leq 1$. If $T \geq 1$, then for any integer s with $1 \leq s \leq n$ there exist $u_1, \ldots, u_s \in \mathbf{Z}^n$ linearly independent such that $|u_1| \cdots |u_s| \leq T$ and*

$$|u_1| \cdots |u_s| \frac{|L_j(u_k)|}{|u_k|} \leq c_4 T^{1-\frac{n}{ms}} \quad \textit{for} \quad 1 \leq j \leq m \textit{ and } 1 \leq k \leq s.$$

PROOF. Say c_3 is the constant from Lemma 6.2. If $T < c_3^{s/n}$, then $c_3 \geq 1$ and

$$T^{1-\frac{n}{sm}} \geq T^{1-\frac{n}{s}} \geq T^{-\frac{n}{s}} > c_3^{-1}.$$

In this case it suffices to take for $u_1, \ldots, u_s$ any distinct standard basis elements of $\mathbf{R}^n$ and $c_4 = c_3$.

So let us assume that $T \geq c_3^{s/n}$. We set $\rho = c_3^{-1/m} T^{n/(ms)} \geq 1$. Applying Lemma 6.2, we get λ_k and u_k as in (6.3.5). Then

$$|u_1| \cdots |u_s| \leq \lambda_1 \cdots \lambda_s \leq (\lambda_1 \cdots \lambda_n)^{s/n} \leq c_3^{s/n} \rho^{ms/n} = T.$$

Furthermore, by (6.3.5) and the inequality above we may estimate

$$|u_1| \cdots |u_s| \frac{|L_j(u_k)|}{|u_k|} \leq \rho^{-1} \lambda_1 \cdots \lambda_s \leq \rho^{-1} T = c_3^{1/m} T^{1-\frac{n}{ms}}$$

for $1 \leq j \leq m$ and $1 \leq k \leq s$ as desired. □

4. Proof of Theorem 6.1

We start this section by pointing out a simple consequence of the assumption $X^{\mathrm{oa}} \neq \emptyset$. This result will not be used in this chapter.

LEMMA 6.4. *Let $X \subset \mathbf{G}_m^n$ be an irreducible closed subvariety defined over $\mathbf{C}$ with $1 \leq \dim X \leq s$. Let us assume $X^{\mathrm{oa},n-s} \neq \emptyset$. Then for any surjective homomorphism of algebraic groups $\varphi : \mathbf{G}_m^n \to \mathbf{G}_m^s$ one has $\dim \varphi(X) = \dim X$.*

PROOF. Say $\dim \varphi(X) < \dim X$, and let $p \in X(\mathbf{C})$. The Fibre Dimension Theorem implies that each irreducible component of $\varphi|_X^{-1}(\varphi|_X(p))$ has positive dimension. Each such component is anomalous since it is contained in a coset of dimension $n - s \leq n - \dim X$. One of these components contains p, therefore $p \notin X^{\mathrm{oa},n-s}$. Hence $X^{\mathrm{oa},n-s} = \emptyset$. □

We will now show Proposition 6.1.

Proof of Proposition 6.1: Let $x_1, \dots, x_n$ denote the coordinate functions on X. Since X has dimension r we may assume without loss of generality that $x_1, \dots, x_r$ are algebraically independent over $\overline{\mathbf{Q}}$. Let $\pi : \mathbf{G}_m^n \to \mathbf{G}_m^r$ denote the projection onto the first r coordinates. Then $\pi(X)$ is Zariski dense in $\mathbf{G}_m^r$, and so is $\pi(U)$. By Theorem 5.3 $\pi(U)$ contains a Zariski open dense subset $V \subset \mathbf{G}_m^r$. The torsion points of $\mathbf{G}_m^r$ lie Zariski dense and hence have infinite intersection with V. Now any point $p \in U(\overline{\mathbf{Q}})$ such that $\pi(p)$ is a torsion point already lies in $\mathcal{H}_{n-r}^n$. Therefore $U \cap \mathcal{H}_{n-r}^n$ is infinite.

Let $\pi' : \mathbf{G}_m^n \to \mathbf{G}_m^{r-1}$ denote the projection onto the first $r-1$ coordinates. Again $\pi'(U)$ lies Zariski dense and thus contains $V' \neq \emptyset$ which is Zariski open in $\mathbf{G}_m^{r-1}$. Now V' must contain a torsion point of $\mathbf{G}_m^{r-1}$. By the Fibre Dimension Theorem, $\pi'|_U$ has positive dimensional fibres. By taking the fibre above a torsion point in V' we see that $U \cap \mathcal{H}_{n-r+1}^n$ has unbounded height. □

Before giving the main argument for the proof of Theorem 6.1 we make some preliminary observations.

Assume for the moment that m is an integer with $0 \leq m \leq n$. Say $p \in C(\mathcal{H}_m^n, \epsilon)$ and $\epsilon \leq (2n)^{-1}$. Then $p = ab$ with $a \in \mathcal{H}_m^n$ and $h(b) \leq \epsilon(1 + h(a))$. By elementary height properties described in chapter 1 we have $h(a) = h(pb^{-1}) \leq h(p) + h(b^{-1}) \leq h(p) + nh(b)$. So $h(a) \leq h(p) + n\epsilon(1 + h(a)) \leq h(p) + \frac{1}{2}(1 + h(a))$. We conclude that

$$h(a) \leq 1 + 2h(p) \quad \text{and} \quad h(b) \leq 2\epsilon(1 + h(p)). \tag{6.4.1}$$

For $T \in \mathbf{R}$ and an integer s with $1 \leq s \leq n$ we define

$$\Phi_s(T) = \{\varphi_{(u_1,\dots,u_s)} : \mathbf{G}_m^n \to \mathbf{G}_m^s; \\ u_1, \dots, u_s \in \mathbf{Z}^n \text{ linearly independent and } |u_1| \cdots |u_s| \leq T\}.$$

Clearly $\Phi_s(T)$ is a finite set.

LEMMA 6.5. *Let $1 \leq m \leq n$ be an integer, let $p \in C(\mathcal{H}_m^n, \epsilon)$ with $\epsilon \leq (2n)^{-1}$, and let $T \geq 1$. Say s is an integer with $1 \leq s \leq n-1$; then there exists a map $\varphi \in \Phi_s(T)$ such that if H is the irreducible component of $\ker \varphi$ containing 1, then the normalized height satisfies*

$$\hat{h}(pH) \leq c_5(T^{1-\frac{n}{ms}} + T\epsilon)(h(p) + 1) \quad \textit{and} \quad \deg(pH) \leq c_6 T. \tag{6.4.2}$$

PROOF. In this proof $c_1', \dots, c_{10}'$ will denote constants which depend only on n.

We have $p = ab$ with $a \in \mathcal{H}_m^n$ and $h(b) \leq \epsilon(1 + h(a))$. By Lemmas 6.1 and 6.3 there exist linearly independent $u_1, \dots, u_s \in \mathbf{Z}^n$ such that $|u_1| \cdots |u_s| \leq T$ and

$$\frac{|u_1| \cdots |u_s|}{|u_k|} h(a^{u_k}) \leq c_1' T^{1-\frac{n}{ms}} h(a) \leq 2c_1' T^{1-\frac{n}{ms}}(1 + h(p)) \quad \text{for} \quad 1 \leq k \leq s. \tag{6.4.3}$$

The second inequality used the first part of (6.4.1). We define $\varphi = \varphi_{(u_1,\dots,u_s)} \in \Phi_s(T)$. By elementary height inequalities (cf. chapter 1) we have $h(b^{u_k}) \leq c_2'|u_k|h(b)$, hence

using the second part of (6.4.1) we get

$$\frac{|u_1|\cdots|u_s|}{|u_k|}h(b^{u_k}) \le c_2'|u_1|\cdots|u_s|h(b) \le c_2'Th(b) \le 2c_2'\epsilon T(1+h(p)) \tag{6.4.4}$$

for $1 \le k \le s$. By (6.4.3) and (6.4.4), and since $h(p^{u_k}) \le h(a^{u_k}) + h(b^{u_k})$ we conclude

$$\frac{|u_1|\cdots|u_s|}{|u_k|}h(p^{u_k}) \le c_3'(T^{1-\frac{n}{ms}} + \epsilon T)(1+h(p)) \quad \text{for} \quad 1 \le k \le s. \tag{6.4.5}$$

Let U be the $n \times s$ matrix with columns $u_1, \dots, u_s$. If we write $u_k = (u_{1k}, \dots, u_{nk})$, we may assume that

$$U' = \begin{bmatrix} u_{11} & \cdots & u_{1s} \\ \vdots & & \vdots \\ u_{s1} & \cdots & u_{ss} \end{bmatrix}$$

is an $s \times s$ minor of U with maximal absolute determinant. We set $\Delta = \det U'$ and let $\Lambda \subset \mathbf{Z}^n$ be the rank s subgroup generated by $u_1, \dots, u_s$. The Cauchy-Binet formula says that $(\det \Lambda)^2$ is equal to the sum over the squares of the determinants of all $s \times s$-minors of U. Hence we may bound

$$|\Delta| \ge c_4' \det \Lambda \tag{6.4.6}$$

for some positive c_4'. The lattice Λ defines an algebraic subgroup $\ker \varphi = \mathcal{H}(\Lambda) \subset \mathbf{G}_m^n$ of dimension $n-s$ by Proposition 6.2. We let $H \subset \mathcal{H}(\Lambda)$ be the irreducible component of $\mathcal{H}(\Lambda)$ that contains the unit element. If $\pi : \mathbf{G}_m^n \to \mathbf{G}_m^{n-s}$ is the projection onto the last $n-s$ coordinates, then $\pi(pH) = \mathbf{G}_m^{n-s}$ because $\Delta \neq 0$. Since the torsion points of $\mathbf{G}_m^{n-s}$ lie Zariski dense we conclude that the set $V = \pi|_{pH}^{-1}(\text{torsion points of } \mathbf{G}_m^{n-s})$ lies Zariski dense in pH.

Let $w \in V \subset pH$, then $w = ph$ for some $h \in H(\overline{\mathbf{Q}})$. We define the $(n-s) \times s$-matrix U'' by

$$U = \begin{bmatrix} U' \\ U'' \end{bmatrix}.$$

Furthermore, let us set $w = (w', w'')$ with $w' \in \mathbf{G}_m^s(\overline{\mathbf{Q}})$ and $w'' \in \mathbf{G}_m^{n-s}(\overline{\mathbf{Q}})$. Then $w'^{U'}w''^{U''} = w^U = p^U h^U = p^U$. So $w'^{\Delta} = w'^{U'(\#U')} = p^{U(\#U')}w''^{-U''(\#U')}$ where $\#U'$ is the adjoint matrix of U'. Note that the coordinates of w'' are roots of unity. Elementary height properties and (6.2.1) provide

$$|\Delta|h(w) = |\Delta|h(w') \le sh(w'^{\Delta}) = sh(p^{U(\#U')}) \le c_5' \sum_{k=1}^{s} x_k h(p^{u_k}).$$

Here x_k denotes the maximum of the absolute values of the elements of the kth row of $\#U'$. Linear algebra implies $x_k \le c_6' \frac{|u_1|\cdots|u_s|}{|u_k|}$. Hence with (6.4.5) we get

$$h(w) \le \frac{c_7'}{|\Delta|}(T^{1-\frac{n}{ms}} + \epsilon T)(1+h(p)), \tag{6.4.7}$$

and this inequality holds for all $w \in V$.

Since $V \subset pH$ lies Zariski dense, the right-hand side of (6.4.7) is an upper bound for $\hat{\mu}^{\mathrm{ess}}(pH)$. Zhang's Theorem (Theorem 5.6) implies

$$\hat{h}(pH) \leq \deg(pH)(1+\dim pH)\hat{\mu}^{\mathrm{ess}}(pH) \leq c_8' \frac{\deg(pH)}{|\Delta|}(T^{1-\frac{n}{ms}} + \epsilon T)(1+h(p)). \tag{6.4.8}$$

We have

$$\deg(pH) = \deg(H) \leq \deg(\mathcal{H}(\Lambda)) \leq c_9' \det \Lambda, \tag{6.4.9}$$

where the last inequality follows from the bound in (6.2.2). We recall (6.4.6) to deduce $|\Delta|^{-1}\deg(pH) \leq c_{10}'$ and so (6.4.8) implies the height bound in (6.4.2). The degree bound in (6.4.2) follows from (6.4.9) and $\det \Lambda \leq |u_1| \cdots |u_s| \leq T$. □

We can now prove Theorem 6.1.

Clearly we may assume $X^{\mathrm{oa},n-s} \neq \emptyset$. To avoid trivialities say $\dim X \geq 1$ and $m \geq 1$.

We will assume that $T \geq 1$ and $\epsilon > 0$ are fixed and depend only on $\deg(X)$ and n. At the end of the proof we will see how to choose them appropriately. In this proof $c_1', \ldots, c_5'$ will denote constants which depend only on n.

Say $\varphi : \mathbf{G}_m^n \to \mathbf{G}_m^s$ is a surjective homomorphism of algebraic groups. We set

$$Z_\varphi = \{p \in X;\ p \text{ is not an isolated point of } \varphi|_X^{-1}(\varphi|_X(p))\}$$

Theorem 5.2 with $k = 1$ implies that $Z_\varphi \subset X$ is Zariski closed. We claim $Z_\varphi \subset X \backslash X^{\mathrm{oa},n-s}$. Indeed say $p \in Z_\varphi$, then p is contained in an irreducible $Y \subset X$ of positive dimension with $\varphi(Y) = \varphi(p)$. So Y is contained in the coset $p\ker\varphi$ of dimension $n-s$. Since $\dim X \leq s$ we conclude $p \notin X^{\mathrm{oa},n-s}$. Thus indeed $Z_\varphi \subset X \backslash X^{\mathrm{oa},n-s}$. In particular $Z_\varphi \neq X$ by our assumption.

We define Z as the union

$$Z = \bigcup_{\varphi \in \Phi_s(T)} Z_\varphi.$$

Clearly $Z \subsetneq X$ is Zariski closed because the union is taken over a finite set. Furthermore, $X^{\mathrm{oa},n-s} \subset X \backslash Z$.

We proceed to show that we have bounded height on $(X \backslash Z) \cap C(\mathcal{H}_m^n, \epsilon)$. Hence let us assume $p \in (X \backslash Z) \cap C(\mathcal{H}_m^n, \epsilon)$ with $m \cdot s < n$. We may assume $\epsilon \leq (2n)^{-1}$. By Lemma 6.5 there exists $\varphi \in \Phi_s(T)$ such that if H is the irreducible component of $\ker\varphi$ containing 1, then (6.4.2) holds. Since $p \notin Z_\varphi$ we conclude that $\{p\}$ is an irreducible component of the intersection $X \cap pH$. The Arithmetic Bézout Theorem (Theorem 5.5) implies

$$h_V(\{p\}) \leq \deg(X)h_V(pH) + \deg(pH)h_V(X) + c_1' \deg(X)\deg(pH). \tag{6.4.10}$$

We have the bound $\deg(pH) \leq c_6 T$. To bound $h_V(pH)$ we note that by (5.3.2) in chapter 5 the inequality $|\hat{h}(pH) - h_V(pH)| \leq c_2' \deg(pH) \leq c_3' T$ holds. By (6.4.2) we conclude

$$h_V(pH) \leq c_5(T^{1-\frac{n}{ms}} + \epsilon T)(h(p)+1) + c_3' T.$$

We insert this inequality and the bound for $\deg(pH)$ into (6.4.10) to get

$$h_V(\{p\}) \leq c_5 \deg(X)(T^{1-\frac{n}{ms}} + \epsilon T)(h(p)+1) + c_6 T h_V(X) + c_4' T \deg(X). \tag{6.4.11}$$

The estimate (5.3.2) lets us replace $h_V(\{p\})$ by $h(p)$ at the cost of replacing c_4' by c_5' in (6.4.11).

Without loss of generality $c_5 \geq \max\{1, 2n/3\}$. As $\frac{n}{ms} > 1$ we choose $T \geq 1$ such that

$$c_5 T^{1-\frac{n}{ms}} \deg(X) = \frac{1}{3}$$

and then $\epsilon > 0$ such that

$$c_5 T \epsilon \deg(X) = \frac{1}{3}.$$

This choice implies $\epsilon \leq (2n)^{-1}$. Inequality (6.4.11) concludes the proof after a short calculation. □

CHAPTER 7

A Bogomolov property modulo algebraic subgroups

Let $X \subset \mathbf{G}_m^n$ be an irreducible closed subvariety defined over $\overline{\mathbf{Q}}$ which is not contained in the translate of a proper algebraic subgroup of $\mathbf{G}_m^n$. Given a real B we show that for an appropriate integer $m = m(\dim X, n)$ there exists a positive ϵ with the following property: the set of $p \in X \cap C(\mathcal{H}_m^n, \epsilon)$ with $h(p) \leq B$ is not Zariski dense in X. A related statement was proved with $\epsilon = 0$ and optimal $m = n - \dim X - 1$ by Bombieri, Masser, and Zannier in [**BMZ04**] (cf. Theorem 6.2 in the previous chapter). If X is a curve, we will show that $m = n - 2$ suffices. This particular value of m is best possible for curves. We then continue by applying results from chapter 6 to deduce finiteness results independent of B.

1. Introduction

In this chapter all varieties are considered to be defined over $\overline{\mathbf{Q}}$. We will work with much the same notation as in chapter 6, i.e. the same notion of height and the same definition for the set $\mathcal{H}_m^n$. We also refer to chapter 6 for the motivation of the problem considered here.

Let $1 \leq r \leq n$ be real numbers, we define

$$\mathfrak{m}(r,n) = n - 2r + 2^{-d}(r(d+2) - n) \text{ with } d = \left[\frac{n-1}{r}\right], \tag{7.1.1}$$

here $[x]$ denotes the greatest integer less or equal to x.

We recall that the degree of a subvariety of $\mathbf{G}_m^n$ is defined using the embedding $\iota : \mathbf{G}_m^n \hookrightarrow \mathbf{P}^n$ introduced in chapter 5. Recall also that a coset is the translate of an algebraic subgroup of $\mathbf{G}_m^n$. The precise definition of the deprived set X^{oa} can be found in chapter 6.

The main result of this chapter is:

THEOREM 7.1. *Let $X \subset \mathbf{G}_m^n$ be an irreducible closed subvariety of dimension $r \geq 1$ defined over $\overline{\mathbf{Q}}$. Let $B \in \mathbf{R}$ and let m be an integer with $m < \mathfrak{m}(r,n)$.*

(i) *If X is not contained in a proper coset, then there exists an $\epsilon > 0$ depending only on B, $\deg(X)$, and n such that*

$$\{p \in X \cap C(\mathcal{H}_m^n, \epsilon);\ h(p) \leq B\}$$

is not Zariski dense in X.

(ii) *For unrestricted X there exists an $\epsilon > 0$ such that*

$$\{p \in X^{\mathrm{oa}} \cap C(\mathcal{H}_m^n, \epsilon);\ h(p) \leq B\}$$

is finite.

In Proposition 7.4 we provide a more precise version of part (i) of this theorem.

For integers $1 \leq r \leq n$ we always have

$$\mathfrak{m}(r, n) > n - 2r,$$

hence the choice $m = n - 2r$ is always possible. In particular if $r = 1$, then we may take $m = n - 2$ in Theorem 7.1. This value of m is best possible for curves. Unfortunately if $r > n/2$, then $n - 2r$ is negative, thus uninteresting as a choice for m. But if we already have $(n-1)/2 < r \leq n - 1$, then $d = 1$ in (7.1.1). So

$$\mathfrak{m}(r, n) = \frac{n - r}{2} \quad \text{if} \quad \frac{n-1}{2} < r \leq n.$$

In this case any integer $m < (n - r)/2$ is admissible in Theorem 7.1.

Under what restriction on $r = \dim X$ can we derive a non-density statement from Theorem 7.1(i) if the points in question lie in $X \cap C(\mathcal{H}_1^n, \epsilon)$, have bounded height, and $\epsilon > 0$ is small? Say $n \geq 4$ and $1 \leq r \leq n - 3$, then one can always take $m = 1$ in Theorem 7.1. Indeed by Lemma 7.8 we have $\mathfrak{m}(r, n) \geq \mathfrak{m}(n - 3, n)$. If $n \geq 6$ then $\mathfrak{m}(n - 3, n) = 3/2$ follows directly from the definition (7.1.1); moreover $\mathfrak{m}(2, 5) = 7/4$ and $\mathfrak{m}(1, 4) = 17/8$. But as $\mathfrak{m}(n - 2, n) = 1$ for all $n \geq 4$ we cannot apply Theorem 7.1 in the critical case $r = n - 2$ and $m = 1$.

From the definition of $\mathfrak{m}$ it is not difficult to deduce

$$\mathfrak{m}(r, n) \leq n - r - 1 \quad \text{if} \quad 2 \leq r \leq n - 2.$$

This inequality shows that if $2 \leq r \leq n - 2$ one can never apply Theorem 7.1 with the critical subgroup dimension $n - r - 1$.

We recall that Kronecker's Theorem implies $C(\mathcal{H}_m^n, 0) = \mathcal{H}_m^n$; Theorem 7.1(ii) is therefore already known to hold with the best-possible $m = n - r - 1$ and even with X^{oa} replaced by X^{ta} if we allow $\epsilon = 0$, cf. Theorem 6.2. This result follows from the work of Bombieri, Masser, and Zannier. Their approach uses a higher dimension Lehmer-type lower bound for heights relative to the maximal abelian extension of $\mathbf{Q}$ due to Amoroso and David. Unfortunately, no way is known to directly reduce Theorem 7.1 to the result of Bombieri, Masser, and Zannier. The proof of Theorem 7.1 seems to be a new approach to this sort of problem since a version of this theorem in the setting of abelian varieties implies new results. The end of the present contains a discussion of this line of thought. We will deduce Theorem 7.1 using a Bogomolov type lower bound for heights given in the work [**AD03**] of Amoroso and David.

By Theorem 6.1 we know that $X^{\mathrm{oa}} \cap C(\mathcal{H}_m^n, \epsilon)$ has bounded height for small ϵ if $m < n/r$. Hence we have the following corollary:

COROLLARY 7.1. *Let $X \subset \mathbf{G}_m^n$ be an irreducible closed subvariety defined over $\overline{\mathbf{Q}}$.*

(i) *If X is a curve then there exists $\epsilon > 0$ depending only on $h_V(X)$, $\deg(X)$, and n such that $X^{\mathrm{oa}} \cap C(\mathcal{H}^n_{n-2}, \epsilon)$ is finite.*
(ii) *If $1 \leq r = \dim X$ and if $m < \min\{n/r, \mathfrak{m}(r,n)\}$ is an integer then there exists $\epsilon > 0$ such that $X^{\mathrm{oa}} \cap C(\mathcal{H}^n_m, \epsilon)$ is finite.*
(iii) *If $\dim X \leq n-3$ and $n \geq 4$ then there exists $\epsilon > 0$ such that $X^{\mathrm{oa}} \cap C(\mathcal{H}^n_1, \epsilon)$ is finite.*

In Corollary 6.2 we showed the finiteness of $X^{\mathrm{oa}} \cap \mathcal{H}^5_2$ for a surface X in $\mathbf{G}^5_m$. In view of the fact that the height is bounded on $X^{\mathrm{oa}} \cap C(\mathcal{H}^5_2, \epsilon)$ for some $\epsilon > 0$ by Theorem 6.1, it is tempting to try to prove finiteness of this set with a (possibly smaller) positive ϵ. Unfortunately as $\mathfrak{m}(2,5) = 7/4$ we cannot take $m = 2$ in Theorem 7.1(ii). Proving the finiteness of $X^{\mathrm{oa}} \cap C(\mathcal{H}^5_2, \epsilon)$ for some positive ϵ remains an open problem.

For curves that are not contained in a proper coset Corollary 7.1(i) is a generalization of Theorem 2 of Bombieri, Masser, and Zannier in [**BMZ99**], but also a generalization of the Bogomolov property. Indeed for $n \geq 2$ the set $C(\mathcal{H}^n_{n-2}, \epsilon)$ contains all algebraic points of $\mathbf{G}^n_m$ of height $\leq \epsilon$. (The Bogomolov property holds for the more general class of curves which are not the translate of an algebraic subgroups by a torsion point.) So Corollary 7.1(i) can be viewed as a sort of Bogomolov property for curves modulo subgroups of dimension $n-2$. If $n = 2$ this corollary follows from the Bogomolov property since $C(\mathcal{H}^2_0, \epsilon)$ is precisely the set of points with height $\leq \epsilon$; we get nothing new here.

Let $H \subset \mathbf{G}^n_m(\overline{\mathbf{Q}})$ be any subset and $\epsilon \geq 0$. In chapter 6 we defined the "cone" $C(H, \epsilon)$ around H; we now define the "tube" around H:

$$T(H, \epsilon) = \{ab;\ a \in H,\ b \in \mathbf{G}^n_m(\overline{\mathbf{Q}}),\ h(b) \leq \epsilon\}.$$

Corollary 7.1(i) motivates the following definition: let $n \geq 2$ and let $X \subset \mathbf{G}^n_m$ be an irreducible algebraic curve, we define

$$\hat{\mu}^{\mathrm{ess}}_C(X) = \sup\{\epsilon \geq 0;\ X \cap C(\mathcal{H}^n_{n-2}, \epsilon) \text{ finite}\}$$

where by convention $\sup \emptyset = -\infty$. We also define

$$\hat{\mu}^{\mathrm{ess}}_T(X) = \sup\{\epsilon \geq 0;\ X \cap T(\mathcal{H}^n_{n-2}, \epsilon) \text{ finite}\}.$$

Then clearly $\hat{\mu}^{\mathrm{ess}}_C(X) \leq \hat{\mu}^{\mathrm{ess}}_T(X)$. We are interested in lower bounds for $\hat{\mu}^{\mathrm{ess}}_C(X)$ or $\hat{\mu}^{\mathrm{ess}}_T(X)$.

In the special situation $n = 2$ we have $\hat{\mu}^{\mathrm{ess}}_T(X) = \hat{\mu}^{\mathrm{ess}}_C(X) = \hat{\mu}^{\mathrm{ess}}(X)$ where $\hat{\mu}^{\mathrm{ess}}(X)$ is the usual essential minimum defined in chapter 5. For example in [**Zag93**] Zagier gave an explicit positive lower bound for $\hat{\mu}^{\mathrm{ess}}(X)$ if $X \subset \mathbf{G}^2_m$ is the line defined by $x + y = 1$. Although he worked with the slightly different height $h(x) + h(y) \leq 2h(x,y)$.

Let us study $\hat{\mu}^{\mathrm{ess}}_C(X)$ in the so-called geometric case, i.e. if the curve X is not contained in a proper coset. Corollary 7.1(i) says that $\hat{\mu}^{\mathrm{ess}}_C(X)$ can be bounded below in terms of $\deg(X)$, $h_V(X)$, and n only. If we consider again the case $n = 2$, the results of Amoroso and David from [**AD03**] imply that $\hat{\mu}^{\mathrm{ess}}_C(X) = \hat{\mu}^{\mathrm{ess}}(X)$ is bounded below only in terms of $\deg(X)$ and n. We pose the following question.

QUESTION 1. *Let $X \subset \mathbf{G}_m^n$ be an irreducible curve not contained in a proper coset, is there a positive lower for $\hat{\mu}_C^{\mathrm{ess}}(X)$ in terms of* $\deg(X)$ *and n only?*

What happens in the arithmetic case, i.e. if we allow the curve X to be contained in a proper coset but not in a proper algebraic subgroup? As was stated in chapter 6, $X \cap \mathcal{H}_{n-1}^n$ does not necessarily have bounded height. Conjecture A of Bombieri, Masser, and Zannier's work [**BMZ06a**] expects that $X \cap \mathcal{H}_{n-2}^n$ is finite, or in other words $\hat{\mu}_C^{\mathrm{ess}}(X) \geq 0$. This conjecture was proved by Maurin in [**Mau06**]. We consider an example:

Let $X \subset \mathbf{G}_m^n$ be contained in a coset of codimension 2. After an automorphism of $\mathbf{G}_m^n$ we may assume $X = \{(\gamma_1, \gamma_2)\} \times X'$ where $X' \subset \mathbf{G}_m^{n-2}$ is a curve. Let ϵ be a positive real, then any $p \in X(\overline{\mathbf{Q}})$ can be written as $p = ab$ with $a = (1, 1, p') \in \mathcal{H}_{n-2}^n$ and $b = (\gamma_1, \gamma_2, 1, \ldots, 1)$. Hence if $h(p)$ is large with respect to ϵ and $h(b)$ then $h(p') = h(a)$ will be large with respect to ϵ and $h(b)$. Therefore $p \in C(\mathcal{H}_{n-2}^n, \epsilon)$ if $h(p)$ is large. So $X \cap C(\mathcal{H}_{n-2}^n, \epsilon)$ is infinite for all positive ϵ. Hence $\hat{\mu}_C^{\mathrm{ess}}(X) \leq 0$ and so $\hat{\mu}_C^{\mathrm{ess}}(X) = 0$ by Maurin's Theorem.

We note that this example does not imply that $\hat{\mu}_T^{\mathrm{ess}}(X) \leq 0$ if we also assume that X is not contained in a proper algebraic subgroup. Indeed under this extra hypothesis γ_1 and γ_2 cannot be roots of unity. By Kronecker's Theorem our b above satisfies $h(b) = h(\gamma_1, \gamma_2) > 0$. We pose a further question.

QUESTION 2. *Let $X \subset \mathbf{G}_m^n$ be an irreducible curve not contained in a proper algebraic subgroup, does $\hat{\mu}_T^{\mathrm{ess}}(X) > 0$ hold?*

We use Dobrowolski's classical theorem to prove a not completely immediate corollary to Theorem 7.1:

COROLLARY 7.2. *Let $X \subset \mathbf{G}_m^n$ be an irreducible closed subvariety defined over $\overline{\mathbf{Q}}$.*

(i) *If X is a curve then there exists $\epsilon > 0$ such that $X^{\mathrm{oa}} \cap C(\mathcal{H}_{n-2}^n, \epsilon)$ is finite and equal to $X^{\mathrm{oa}} \cap \mathcal{H}_{n-2}^n$.*

(ii) *If $1 \leq r = \dim X$ and if $m < \min\{n/r, \mathfrak{m}(r, n)\}$ is an integer then there exists $\epsilon > 0$ such that $X^{\mathrm{oa}} \cap C(\mathcal{H}_m^n, \epsilon)$ is finite and equal to $X^{\mathrm{oa}} \cap \mathcal{H}_1^n$.*

(iii) *If $\dim X \leq n - 3$ and $n \geq 4$ then there exists $\epsilon > 0$ such that $X^{\mathrm{oa}} \cap C(\mathcal{H}_1^n, \epsilon)$ is finite and equal to $X^{\mathrm{oa}} \cap \mathcal{H}_1^n$.*

So for small ϵ there are no points on a curve X that are close to a subgroup of dimension $n - 2$ without already lying on a subgroup of dimension $n - 2$ (if $X^{\mathrm{oa}} = X$). In Corollary 7.2(i) the quantity ϵ may also depend on the field of definition of X, even when C is replaced by T, as is illustrated by the following simple example.

Let X be the curve defined by $x + y = 2^{1/k} + 1$ with $k \in \mathbf{N}$. Then $\deg(X) = 1$ and $h_V(X)$ is bounded independently of k. But X contains the non-torsion point $(2^{1/k}, 1)$ which lies in $T(\mathcal{H}_0^n, \epsilon)$ with $\epsilon = (\log 2)/k$.

We conclude the introduction by discussing the abelian situation. More specifically we replace $\mathbf{G}_m^n$ by E^n where E is an elliptic curve. The set $\mathcal{H}_m^n$ also makes sense in this setting, as do $T(\cdot, \cdot)$ and $C(\cdot, \cdot)$ when using for example the Néron-Tate height

associated to an ample symmetric line bundle. Let $X \subset E^n$ be an irreducible curve. Intersections of X with $\mathcal{H}_m^n$ have been studied by Viada in [**Via03**] and Rémond and Viada in [**RV03**].

Say X is not contained in the translate of a proper algebraic subgroup of E^n. If E does not have complex multiplication, then Viada proves the finiteness of $X \cap \mathcal{H}_m^n$ for $m \leq n/2 - 2$. If E has complex multiplication, Viada shows that one has finiteness for the optimal $m = n - 2$. In her proof she uses a height upper bound analog to Theorem 6.1 and also a lower bound for the Néron-Tate height on powers of elliptic curves. The reason for the apparently non-optimal $n/2 - 2$ in the non-complex multiplication case comes from the fact that up to now no one can prove essentially best-possible Lehmer-type height lower bounds here. The problem can be traced back to the fact that on an arbitrary elliptic curves the Frobenius automorphism cannot in general be lifted from the reduction modulo a suitable prime.

The work of Amoroso and David in [**AD03**] on Bogomolov-type lower bounds for heights in the toric case uses methods from transcendence theory. Their extrapolation step does not make use of the Frobenius automorphism. Thus there is hope that lower bounds of the same quality as in [**AD03**] hold in the case where the torus is replaced by a power of an elliptic curve. In fact Galateau has recently announced such a result in the more general case of a product of elliptic curves in [**Gal07**].

In a preprint [**Via07**], Viada improved on her result obtained together with Rémond. She proved that if X is a curve not contained in a proper algebraic subgroup, then $X \cap \mathcal{H}_{n-3}^n$ is finite regardless of C.M. type of the elliptic curve E. The subgroup dimension $n - 3$ is already close to the best-possible value $n - 2$. Her approach, obtained independently from ours, is based on Galateau's Bogomolov-type height lower bound.

Let X be a curve in E^n not contained in the translate of a proper algebraic subgroup. Galateau's result together with the methods presented in this chapter and the previous one could provide a proof for the finiteness of $X \cap C(\mathcal{H}_{n-2}^n, \epsilon)$ for some $\epsilon > 0$ regardless if E has complex multiplication or not. Of course such a result would imply the finiteness of $X \cap \mathcal{H}_{n-2}^n$, which is still unknown.

2. Auxiliary results

We reuse notation introduced in section 2 of chapter 6. For example $|\cdot|$ denotes the euclidean norm on $\mathbf{R}^n$ and also the euclidean norm of the coefficient vector of a linear form. Unless otherwise stated the symbols $c_1, \ldots, c_{14}$ denote constants which depend only on n.

We recall two lemmas from chapter 6 section 3 which will also be used in this chapter.

LEMMA 7.1. *Let $1 \leq m \leq n$ and let $a \in \mathcal{H}_m^n$. Then there exist linear forms $L_1, \ldots, L_m \in \mathbf{R}[X_1, \ldots, X_n]$ such that $|L_j| \leq 1$ and*

$$h(a^u) \leq c_1 \max_{1 \leq j \leq m} \{|L_j(u)|\} h(a)$$

for all $u \in \mathbf{Z}^n$.

The second lemma wraps up all the geometry of numbers we will use.

LEMMA 7.2. *Let $1 \leq m \leq n$ and let $L_1, \dots, L_m \in \mathbf{R}[X_1, \dots, X_n]$ be linear forms with $|L_j| \leq 1$. If $\rho \geq 1$, there exist $\lambda_1, \dots, \lambda_n$ with $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ and linearly independent $u_1, \dots, u_n \in \mathbf{Z}^n$ such that*

$$|u_k| \leq \lambda_k, \quad |L_j(u_k)| \leq \rho^{-1}\lambda_k, \text{ and } \quad \lambda_1 \cdots \lambda_n \leq c_2\rho^m.$$

We recall Dobrowolski's Theorem: if $\alpha \in \overline{\mathbf{Q}}^*$ is not a root of unity and $D = [\mathbf{Q}(\alpha) : \mathbf{Q}]$, then

$$h(\alpha) \geq c_3 \frac{1}{D} \left(\frac{\log\log 3D}{\log 2D} \right)^3 \tag{7.2.1}$$

where $c_3 > 0$ is an absolute constant.

The geometry of numbers machinery and Dobrowolski's Theorem give the following lemma which is the main ingredient in the proof of Corollary 7.2. It can be regarded as a Dobrowolski type result modulo subgroups.

LEMMA 7.3. *Let $1 \leq m \leq n$ be an integer, let $\delta > 0$, and let $p \in C(\mathcal{H}_m^n, \epsilon)$ with $h(p) \leq B$ and $[\mathbf{Q}(p) : \mathbf{Q}] \leq D$. If*

$$\epsilon^{-1} \geq c_4(1+B)^{m+1}D^{m+1+\delta}. \tag{7.2.2}$$

then $p \in \mathcal{H}_m^n$. Here $c_4 > 0$ depends only on n and δ.

PROOF. The symbols c_1', c_2', c_3' denote constants which depend only on n and δ. We may assume $c_4 \geq 2n$; we will see how to choose c_4 appropriately further down. We define $\rho \geq 1$ to be the right-hand side of (7.2.2).

We write $p = ab$ with $a \in \mathcal{H}_m^n$ and $h(b) \leq \epsilon(1 + h(a))$. By (6.4.1) we have $h(a) \leq 1 + 2B$ and $h(b) \leq 2\epsilon(1+B)$. Let $L_1, \dots, L_m$ be the linear forms from Lemma 7.1 and let λ_k and u_k be from Lemma 7.2 applied to the L_j. We deduce

$$h(a^{u_k}) \leq c_1\lambda_k\rho^{-1}h(a) \leq 2c_1\lambda_k\rho^{-1}(1+B). \tag{7.2.3}$$

Furthermore, by elementary height inequalities we have $h(b^{u_k}) \leq \sqrt{n}|u_k|h(b) \leq 2\sqrt{n}\lambda_k\epsilon(1+B)$. We combine this inequality with (7.2.3) and use $\epsilon \leq \rho^{-1}$ to get

$$h(p^{u_k}) \leq h(a^{u_k}) + h(b^{u_k}) \leq c_1'\lambda_k(\rho^{-1} + \epsilon)(1+B) \leq 2c_1'\lambda_k\rho^{-1}(1+B). \tag{7.2.4}$$

Say $1 \leq k \leq n-m$. By Lemma 7.2 we have $\lambda_k \geq |u_k| \geq 1$ and

$$\lambda_k \leq (\lambda_{n-m} \cdots \lambda_n)^{\frac{1}{m+1}} \leq (\lambda_1 \cdots \lambda_n)^{\frac{1}{m+1}} \leq c_2'\rho^{\frac{m}{m+1}}.$$

We apply this inequality to (7.2.4) and use the definition of ρ to get

$$h(p^{u_k}) \leq c_3'\rho^{-\frac{1}{m+1}}(1+B) = c_3'c_4^{-\frac{1}{m+1}}D^{-1-\frac{\delta}{m+1}},$$

this inequality holds for all $1 \leq k \leq n-m$.

Now $[\mathbf{Q}(p^{u_k}) : \mathbf{Q}] \leq D$, so if c_4 is large enough with respect to c_3' and δ we use Dobrowolski's Theorem (7.2.1) to deduce that $p^{u_1}, \dots, p^{u_{n-m}}$ are roots of unity. Since $u_1, \dots, u_{n-m}$ are linearly independent we conclude that $p \in \mathcal{H}_m^n$ by Proposition 6.2. □

Say $X \subset \mathbf{G}_m^n$ is an algebraic curve not contained in a proper coset. Then by Theorem 6.1 chapter 6 there is a B and $\epsilon > 0$ such that if $p \in X \cap C(\mathcal{H}_{n-2}^n, \epsilon)$ then $h(p) \leq B$. By the lemma above we conclude that p is already contained in a algebraic subgroup of dimension $n-2$ if ϵ is small with respect to B, $\deg(X)$, $[\mathbf{Q}(p) : \mathbf{Q}]$, and n. In this case p is contained in the set $X \cap \mathcal{H}_{n-2}^n$, which is finite by the result of Bombieri, Masser, and Zannier mentioned in the introduction. Of course we cannot conclude the finiteness of

$$\{p \in X \cap C(\mathcal{H}_{n-2}^n, \epsilon);\ h(p) \leq B\}$$

for some fixed positive ϵ since as of yet we know nothing about the degree $[\mathbf{Q}(p) : \mathbf{Q}]$. Also, if we already knew that the degrees $[\mathbf{Q}(p) : \mathbf{Q}]$ were uniformly bounded, then finiteness would follow from boundedness of height and Northcott's Theorem.

3. Push-forwards and pull-backs

In this section we prove two lemmas on bounds for degrees of push-forwards and pull-backs of varieties by a homomorphism of algebraic groups.

Let $X \subset \mathbf{G}_m^n$ be an irreducible subvariety throughout this section.

LEMMA 7.4. *Let $u_1, \ldots, u_t \in \mathbf{Z}^n$ with $0 < |u_1| \leq \cdots \leq |u_t|$ and $\varphi = \varphi_{(u_1,\ldots,u_t)} : \mathbf{G}_m^n \to \mathbf{G}_m^t$. If $q = \dim \varphi(X)$, then*

$$\deg(\overline{\varphi(X)}) \leq c_5 |u_{t-q+1}| \cdots |u_t| \deg(X).$$

PROOF. For brevity we set $Y = \overline{\varphi(X)}$. By Theorem 5.3 there exists $U \subset Y$ Zariski open and dense with $U \subset \varphi(X)$. By a standard Bertini type argument we may find polynomials $l_i \in \overline{\mathbf{Q}}[X_1, \ldots, X_t]$ $(1 \leq i \leq q)$ of degree 1 such that $l_i - X_{t-q+i} \in \overline{\mathbf{Q}}[X_1, \ldots, X_{t-q}]$ with the following properties: the set

$$S = \{y \in Y;\ l_1(y) = \cdots = l_q(y) = 0\}$$

is finite of cardinality $\deg(Y)$ and contained in U. We define $R = \varphi|_X^{-1}(S)$. Then $R \subset X$ is Zariski closed and has at least $\deg(Y)$ irreducible components. On the other hand we have $R = \{x \in X;\ l_1(\varphi(x)) = \cdots = l_q(\varphi(x)) = 0, \}$. The exponent vectors in $l_i \circ \varphi$ have norm bounded by $|u_{t-q+i}|$. Hence Bézout's Theorem 5.4 implies that the number of irreducible components of R is bounded above by $c_5 |u_{t-q+1}| \cdots |u_t| \deg(X)$. This completes the proof. □

If $\varphi = \pi : \mathbf{G}_m^n \to \mathbf{G}_m^t$ is the projection onto any set of t coordinates, then proof above can easily be modified to give $\deg(\overline{\pi(X)}) \leq \deg(X)$.

LEMMA 7.5. *Let $u_1, \ldots, u_n \in \mathbf{Z}^n$ be linearly independent and $\varphi = \varphi_{(u_1,\ldots,u_n)} : \mathbf{G}_m^n \to \mathbf{G}_m^n$. Then there exists an irreducible component $W \subset \varphi^{-1}(X)$ such that $\varphi(W) \subset X$ is Zariski dense,*

$$\dim W = \dim X, \text{ and } \deg(W) \leq c_6 |u_1| \cdots |u_n| \deg(X).$$

PROOF. Let $W_1, \dots, W_l$ be the irreducible components of $\varphi^{-1}(X)$. Since φ is surjective $\varphi(W_{i_0}) \subset X$ is Zariski dense for some i_0. We set $W = W_{i_0}$. Because φ has finite fibres we conclude that $\dim W = \dim X$ by Theorem 5.1.

It remains to prove the upper bound for $\deg(W)$. By Bézout's Theorem (Theorem 5.4) the irreducible components of

$$\{(p,q) \in \mathbf{G}_m^n \times \mathbf{G}_m^n;\ \varphi(p) = q\} \cap (\mathbf{G}_m^n \times X)$$

have degree bounded by $c_6|u_1| \cdots |u_n| \deg(X)$. Now W is the projection of such an irreducible component onto the first factor of $\mathbf{G}_m^n \times \mathbf{G}_m^n$. The lemma follows from the remark after the proof of Lemma 7.4. □

4. A lower bound for the product of heights

In [**AD03**] Amoroso and David deduced a positive lower bound for the height of a point on an open subset of certain subvarieties of $\mathbf{G}_m^n$, cf. Theorem 5.7. Their bound depends only on the degree of the variety and on n. In this section we derive a corollary of this result by deducing a lower bound for the product over heights of some coordinates of a generic point on a variety. The lower bounds depends only on the degree of the variety and n.

Let us consider for the moment the curve X in $\mathbf{G}_m^3$ defined by $p_1 + p_2 = 1$ and $p_2 + p_3 = 2$ (we have $X^{\mathrm{oa}} = X$). Because of examples like $(2^{1/k}, 1 - 2^{1/k}, 1 + 2^{1/k}) \in X$ with $k \in \mathbf{N}$ the product

$$h(p_1)h(p_2)h(p_3) \tag{7.4.1}$$

cannot be bounded below by some positive constant on any Zariski open dense subset of X. The trick is to forget about the coordinate in (7.4.1) with minimal height, in this case p_1 if k is large. In Proposition 7.1 below we will show that for all p on a open dense subset X the quantity

$$\max\{h(p_1)h(p_2), h(p_1)h(p_3), h(p_2)h(p_3)\}$$

is bounded below by a positive constant.

First we need a preparatory lemma which encapsulates the consequence (5.3.5) of Amoroso and David's Theorem (Theorem 5.7).

LEMMA 7.6. *Let $X \subsetneq \mathbf{G}_m^n$ be an irreducible closed subvariety that is not contained in a proper coset. Let $u_1, \dots, u_n \in \mathbf{Z}^n$ be linearly independent and $\varphi = \varphi_{(u_1,\dots,u_n)} : \mathbf{G}_m^n \to \mathbf{G}_m^n$. We set $\Pi = |u_1| \cdots |u_n|$. There exists a Zariski open dense $U \subset X$ such that if $q \in \mathbf{G}_m^n$ with $\varphi(q) \in U$, then*

$$h(q) \geq \frac{c_7}{(\Pi \deg(X))^{1/\operatorname{codim} X} (\log(3\Pi \deg(X)))^{c_8}}$$

where c_7 and c_8 are positive and depend only on n.

PROOF. Let $W \subset \varphi^{-1}(X)$ be an irreducible component as in Lemma 7.5. Clearly W is not contained in a proper coset of $\mathbf{G}_m^n$. By (5.3.5), the consequence of Amoroso and David's Theorem, we have

$$\hat{\mu}^{\mathrm{ess}}(W) \geq \frac{c}{(\Pi \deg(X))^{1/\operatorname{codim} X}(\log(3\Pi \deg(X)))^{c'}} > 0,$$

where c and c' depend only on n. By definition of the essential minimum there exists $V \subset W$ Zariski open and dense such that if $q \in V(\overline{\mathbf{Q}})$ then $h(q) \geq \frac{1}{2}\hat{\mu}^{\mathrm{ess}}(W)$. Finally by Theorem 5.3 there exists $U \subset \varphi(V)$ such that $U \subset X$ is Zariski open and dense. The lemma follows because if $\varphi(q) \in U$ then $q = q'\zeta$ for some $q' \in V$ and some torsion point ζ; thus $h(q) = h(q')$. □

The following proposition is a geometric analogue of Theorem 1.6 in [**AD99**].

PROPOSITION 7.1. *Let $X \subset \mathbf{G}_m^n$ be an irreducible closed subvariety that is not contained in a proper coset. There exists $U \subset X$ Zariski open and dense such that for each $p = (p_1, \ldots, p_n) \in U(\overline{\mathbf{Q}})$ there is a subset $\Sigma \subset \{1, \ldots, n\}$ with $|\Sigma| \geq n - \dim X$ and*

$$\prod_{k \in \Sigma} h(p_k) \geq \frac{c_9}{\deg(X)(\log(3\deg(X)))^{c_{10}}} \tag{7.4.2}$$

with $c_9 > 0$.

PROOF. Throughout the proof $c_1', \ldots, c_6'$ will denote constants which depend only on n. For brevity let $r = \dim X$. Clearly we may assume that $X \neq \mathbf{G}_m^n$. We prove the proposition by contradiction. To do this we assume that there exists $V \subset X(\overline{\mathbf{Q}})$ Zariski dense such that for all $p = (p_1, \ldots, p_n) \in V$ and all $\Sigma \subset \{1, \ldots, n\}$ with $|\Sigma| \geq n - r$ we have

$$\prod_{j \in \Sigma} h(p_j) < \frac{C}{\deg(X)(\log(3\deg(X)))^{C'}} \tag{7.4.3}$$

where $C, C' > 0$ are fixed and depend only on n. We will see how to choose these constants properly later on.

First we do some reduction steps. By replacing V by a smaller, yet still Zariski dense subset of $X(\overline{\mathbf{Q}})$ and after permuting coordinates we may assume that

$$h(p_1) \leq h(p_2) \leq \cdots \leq h(p_n)$$

for all $(p_1, \ldots, p_n) \in V$. These inequalities will be used freely throughout the proof.

Let us temporarily assume that the set of $(p_1, \ldots, p_n) \in V$ with $h(p_n) > 1$ is dense in X. Let $X' \subset \mathbf{G}_m^{n-1}$ be the Zariski closure of the projection of X onto the first $n-1$ coordinates. Now $\dim X' \leq r$ and X' has degree at most $\deg(X)$ by the remark after Lemma 7.4. Also, X' is not contained in a proper coset. By induction on n and using (7.4.3) with $\Sigma = \{r+1, \ldots, n\}$ we conclude a contradiction if C is small enough and C' is big enough. Hence we may suppose that

$$\max_j \{h(p_j)\} = h(p_n) \leq 1 \quad \text{for} \quad (p_1, \ldots, p_n) \in V, \tag{7.4.4}$$

where V is still Zariski dense in X.

Let $p = (p_1, \dots, p_n) \in V$ be arbitrary for the moment. Let $\pi : \mathbf{G}_m^n \to \mathbf{G}_m^{r+1}$ be the projection onto the first $r+1$ coordinates. The variety $Y = \overline{\pi(X)}$ is not contained in a proper coset since X itself is not. We note that $\deg(Y) \le \deg(X)$ by the comment after Lemma 7.4 and $\operatorname{codim} Y \ge 1$. Thus (5.3.5) gives a lower bound for $\hat{\mu}^{\mathrm{ess}}(Y)$. So after replacing V by a smaller, dense subset of $X(\overline{\mathbf{Q}})$ we may assume that

$$(r+1)h(p_{r+1}) = (r+1)\max\{h(p_1), \dots, h(p_{r+1})\} \ge h(\pi(p)) \ge \ge \frac{c_1'}{\deg(X)(\log(3\deg(X)))^{c_2'}}. \tag{7.4.5}$$

In particular we have $h(p_{r+1}) > 0$. For $r+1 \le j \le n$ we define

$$k_1 = k_2 = \cdots = k_r = 1 \text{ and } k_j = \left[\frac{h(p_j)}{h(p_{r+1})}\right] \ge 1. \tag{7.4.6}$$

We note $k_{r+1} = 1$. The k_j depend on the point p. For $r+1 \le j \le n$ we have

$$\frac{1}{2}\frac{h(p_j)}{h(p_{r+1})} \le k_j \le \frac{h(p_j)}{h(p_{r+1})} \tag{7.4.7}$$

Next we would like to bound the k_j from above. For $r+1 \le j \le n$ we apply (7.4.4) and (7.4.5) to deduce

$$1 \le k_j \le \frac{1}{h(p_{r+1})} \le c_3' \deg(X)(\log(3\deg(X)))^{c_2'}, \tag{7.4.8}$$

with $c_3' \ge 1$. In fact (7.4.8) holds for all $1 \le j \le n$. As an important consequence we see that the k_j are bounded independently of $p \in V$.

Still considering $p \in V$ we define a homomorphism $\varphi : \mathbf{G}_m^n \to \mathbf{G}_m^n$ by $\varphi(x_1, \dots, x_n) = (x_1^{k_1}, \dots, x_n^{k_n})$ where the k_j are as in (7.4.6). Of course φ depends on p. Each such φ gives rise to an open dense $U_\varphi \subset X$ as in Lemma 7.6. Since the set of possible φ, i.e. with k_j satisfying (7.4.8), is finite and independent of p there are only finitely many U_φ. Because $V \subset X$ is Zariski dense we may pick once and for all a $p \in V$ with $p \in U_\varphi$ for all φ as defined above with k_j satisfying (7.4.8).

Let $q = (q_1, \dots, q_n) \in \mathbf{G}_m^n(\overline{\mathbf{Q}})$ with $\varphi(q) = p$. By our choice of p and Lemma 7.6 we have the lower bound

$$h(q) \ge \frac{c_7}{(k_1 \cdots k_n \deg(X))^{1/(n-r)}(\log(3k_1 \cdots k_n \deg(X)))^{c_8}}. \tag{7.4.9}$$

We have

$$h(p_{r+1}) \cdots h(p_n) = k_{r+1} \cdots k_n h(q_{r+1}) \cdots h(q_n). \tag{7.4.10}$$

From (7.4.7) we derive

$$1 \le \frac{h(q_j)}{h(p_{r+1})} \le 2 \quad \text{for} \quad r+1 \le j \le n. \tag{7.4.11}$$

We apply the lower bound in (7.4.11) to the right side of (7.4.10) and get

$$h(p_{r+1})\cdots h(p_n) \geq k_{r+1}\cdots k_n h(p_{r+1})^{n-r}. \tag{7.4.12}$$

Say $1 \leq j' \leq n$ with $h(q_{j'}) = \max\{h(q_1),\dots,h(q_n)\}$. Now $h(q_{r+1}) = h(p_{r+1}) \geq h(p_j) = h(q_j)$ for $1 \leq j \leq r$, hence we may assume $j' \geq r+1$. We insert the upper bound from (7.4.11) with $j = j'$ into (7.4.12) and use $h(q_{j'}) \geq \frac{1}{n}h(q)$ to derive

$$h(p_{r+1})\cdots h(p_n) \geq 2^{-(n-r)}k_{r+1}\cdots k_n h(q_{j'})^{n-r} \geq c_4' k_{r+1}\cdots k_n h(q)^{n-r}.$$

We use the lower bound for $h(q)$ in (7.4.9) to get

$$\begin{aligned} h(p_{r+1})\cdots h(p_n) &\geq \frac{c_5' k_{r+1}\cdots k_n}{k_1\cdots k_n \deg(X)(\log(3k_1\cdots k_n\deg(X)))^{(n-r)c_8}} \\ &= \frac{c_5'}{\deg(X)(\log(3k_1\cdots k_n\deg(X)))^{(n-r)c_8}}. \end{aligned}$$

We bound the remaining k_j in the logarithm from above with the help of (7.4.8) and obtain

$$h(p_{r+1})\cdots h(p_n) \geq \frac{c_6'}{\deg(X)(\log(3\deg(X)))^{(n-r)c_8}}. \tag{7.4.13}$$

If we choose $C = c_6'$ and $C' = (n-r)c_8$, then (7.4.3) and (7.4.13) contradict. □

It is not too difficult to construct examples of X of arbitrary dimension, as at the beginning of this section, where one really needs to omit $\dim X$ factors in the product (7.4.2) to obtain a positive lower bound on an open dense set. Furthermore, the dependency of (7.4.2) on $\deg(X)$ is optimal up to the usual log power.

We will prove a variant of Proposition 7.1 where the variety is allowed to be contained in a proper coset. Let $X \subset \mathbf{G}_m^n$, we define

$$\mathrm{s}^{\mathrm{o}}(X) = \inf\{\dim H;\ H \subset \mathbf{G}_m^n \text{ a coset with } X \subset H\}.$$

Proposition 7.1 only holds for varieties $X \subset \mathbf{G}_m^n$ with $\mathrm{s}^{\mathrm{o}}(X) = n$. A simple projection argument shows that in general we have:

Proposition 7.2. *Let $X \subset \mathbf{G}_m^n$ be an irreducible closed subvariety. There exists $U \subset X$ Zariski open and dense such that for each $p = (p_1,\dots,p_n) \in U(\overline{\mathbf{Q}})$ there is a subset $\Sigma \subset \{1,\dots,n\}$ with $|\Sigma| \geq \mathrm{s}^{\mathrm{o}}(X) - \dim X$ and*

$$\prod_{k\in\Sigma} h(p_k) \geq \frac{c_{11}}{\deg(X)(\log(3\deg(X)))^{c_{12}}}$$

with $c_{11} > 0$.

Proof. Let H be a coset with $X \subset H$ and $\mathrm{s}^{\mathrm{o}}(X) = \dim H = n-h$. By Proposition 6.2 there are linearly independent $u_1,\dots,u_h \in \mathbf{Z}^n$ such that x^{u_i} is constant on H. We may assume that the $h \times h$ matrix whose ith row consists of the first h entries of u_i is non-singular. In this case the projection $\pi : \mathbf{G}_m^n \to \mathbf{G}_m^{n-h}$ onto the last $n-h$ coordinates has finite fibres when restricted to H. Therefore $\pi|_X$ has finite fibres too.

By Theorem 5.1 we have $\dim \overline{\pi(X)} = \dim X$. The comment after Lemma 7.4 implies $\deg(\overline{\pi(X)}) \leq \deg(X)$. Furthermore, $\overline{\pi(X)}$ is not contained in a proper coset, indeed otherwise X would be contained in a coset of dimension strictly less than $n-h$. The proposition now follows from Proposition 7.1 applied to $\overline{\pi(X)}$. □

If for example X is itself a coset, then $\mathrm{s}^{\mathrm{o}}(X) = \dim X$; in this case Proposition 7.2 is an empty statement.

5. Proof of Theorem 7.1 and corollaries

We start off with an auxiliary lemma on linear programming. We recall that $\mathfrak{m}(r,s)$ was defined in (7.1.1).

LEMMA 7.7. *Let $1 \leq r < s \leq n$ be integers, let $M = (m_{ij})$ be the $n \times (s-r)$ matrix defined by*

$$m_{ij} = \begin{cases} 1 & \text{if} \quad i+j \leq n-r+1, \\ 2 & \text{if} \quad n-r+2 \leq i+j \leq n+1, \\ 0 & \text{else,} \end{cases}$$

and let $w = (w_1, \dots, w_{s-r})^{\mathrm{t}} \in \mathbf{R}^{s-r}$ be the column vector with

$$w_j = \begin{cases} 2^{-(\frac{j-1}{r}+1)} & \text{if} \quad r|(j-1), \\ 0 & \text{else.} \end{cases}$$

Then $v = (v_1, \dots, v_n)^{\mathrm{t}} = Mw$ satisfies $v_i \leq 1$ and furthermore

$$\sum_{j=1}^{s-r} (s-r-j+1) w_j = \mathfrak{m}(r,s). \tag{7.5.1}$$

PROOF. The vector w looks like

$$w = (\frac{1}{2}, 0, \dots, 0, \frac{1}{4}, 0, \dots, 0, \frac{1}{8}, 0, \dots)^{\mathrm{t}} \tag{7.5.2}$$

with $r-1$ zeros between consecutive negative powers of 2 (there are no zeros if $r = 1$). When $1 \leq i \leq n$, the ith row of M starts off with a certain number (possibly zero) of consecutive ones followed by say N consecutive twos and finally consecutive zeros. By definition we have $N \leq r$, hence by (7.5.2) there is at most one j with $m_{ij} = 2$ and $w_j \neq 0$. Let N' be the number of j with $m_{ij} = 1$ and $w_j \neq 0$, then

$$v_i = \sum_{j,\, m_{ij}=1} w_j + 2 \sum_{j,\, m_{ij}=2} w_j \leq (\frac{1}{2} + \cdots + \frac{1}{2^{N'}}) + 2\frac{1}{2^{N'+1}} = 1.$$

This inequality proves the first part of the lemma. The second part, the equality (7.5.1), follows from an elementary calculation. □

The following proposition will imply part (i) of Theorem 7.1.

PROPOSITION 7.3. *Let $X \subset \mathbf{G}_m^n$ be an irreducible closed subvariety of dimension $r \geq 1$ and assume that $s = \mathrm{s}^{\mathrm{o}}(X) \geq r+1$. Let m be an integer with $0 \leq m < \mathfrak{m}(r,s)$, let $B \geq 1$, and let $\delta > 0$. Then there exists $c_{13} > 0$ which depends only on n and δ such that if*

$$\epsilon \leq c_{13}(B^{m+\delta}(\deg(X))^{1+\delta})^{-\frac{1}{\mathfrak{m}(r,s)-m}} \tag{7.5.3}$$

the set $\{p \in X \cap T(\mathcal{H}_m^n, \epsilon);\ h(p) \leq B\}$ is not Zariski dense in X.

PROOF. Throughout the proof $c_1', \ldots, c_{12}'$ will denote constants which only depend on n and δ. We consider $c_{13} \leq 1$ as fixed and depending only on n and δ; we will see how to choose it later on. We also define $\rho \in \mathbf{R}$ such that $B\rho^{-1}$ is equal to the right-hand side of (7.5.3). We note that $\rho \geq B \geq 1$.

Now say $p \in X \cap T(\mathcal{H}_m^n, \epsilon)$ and $h(p) \leq B$. We will show that p is contained in some proper Zariski closed subset of X independent of p.

By definition we may write $p = ab$ with $a \in \mathcal{H}_m^n$ and $h(b) \leq \epsilon$. Elementary height properties imply $h(a) \leq h(p) + h(b^{-1}) \leq h(p) + nh(b) \leq h(p) + n\epsilon \leq 2nB$. Let us assume for the moment that $m \geq 1$. By Lemmas 7.1 and 7.2 there exist $\lambda_1, \ldots, \lambda_n$ with $0 < \lambda_1 \leq \cdots \leq \lambda_n$ and linearly independent $u_1, \ldots, u_n \in \mathbf{Z}^n$ such that for $1 \leq k \leq n$

$$|u_k| \leq \lambda_k, \quad h(a^{u_k}) \leq c_1' h(a)\rho^{-1}\lambda_k \leq c_2' B\rho^{-1}\lambda_k, \text{ and } \lambda_1 \cdots \lambda_n \leq c_3'\rho^m. \tag{7.5.4}$$

In the case $m = 0$ the statements in (7.5.4) also hold if we take $\lambda_k = 1$ and u_k the standard basis elements of $\mathbf{R}^n$. Indeed if $m = 0$, then a is a torsion point and thus has height 0. Elementary height inequalities give $h(b^{u_k}) \leq \sqrt{n}h(b)|u_k| \leq \sqrt{n}\epsilon\lambda_k$, hence

$$h(p^{u_k}) \leq h(a^{u_k}) + h(b^{u_k}) \leq c_4'(B\rho^{-1} + \epsilon)\lambda_k \leq 2c_4' B\rho^{-1}\lambda_k, \tag{7.5.5}$$

here we used the bound $\epsilon \leq B\rho^{-1}$ in the last inequality.

Let $t_0 = n - s + r + 1$, so that $r+1 \leq t_0 \leq n$; also let t be an integer with $t_0 \leq t \leq n$. We define the surjective homomorphism of algebraic groups $\varphi_t = \varphi_{(u_1,\ldots,u_t)} : \mathbf{G}_m^n \to \mathbf{G}_m^t$ and the irreducible variety $X_t = \overline{\varphi_t(X)} \subset \mathbf{G}_m^t$. Both φ_t and variety X_t depend on p. On the other hand the quantity $c_3'\rho^m$ depends only on X, B, and δ but not on the point p. Since $|u_1| \cdots |u_t| \leq c_3'\rho^m$ by (7.5.4) there are only finitely many possibilities for φ_t. These morphisms give a finite set of X_t which does not depend on p.

We proceed by bounding $\mathrm{s}^{\mathrm{o}}(X_t)$ from below. To do this let $H \subset \mathbf{G}_m^t$ be a coset of dimension $\mathrm{s}^{\mathrm{o}}(X_t)$ that contains X_t. Then $\dim \varphi_t^{-1}(H) = \dim H + (n-t)$ and $\varphi_t^{-1}(H)$ is a coset containing X. Hence

$$\mathrm{s}^{\mathrm{o}}(X_t) = \dim \varphi_t^{-1}(H) - (n-t) \geq s + t - n \geq s + t_0 - n = r + 1. \tag{7.5.6}$$

In particular X_t has positive dimension.

Lemma 7.4 and (7.5.4) give the upper bound

$$\deg(X_t) \leq c_5'\lambda_{t-\dim X_t+1} \cdots \lambda_t \deg(X)$$

since the λ_k are in ascending order. Furthermore, $\dim X_t \leq r$ and $\lambda_k \geq 1$, so

$$\deg(X_t) \leq c_5'\lambda_{t-r+1} \cdots \lambda_t \deg(X) \quad \text{for all} \quad t_0 \leq t \leq n. \tag{7.5.7}$$

Let $c_{11} > 0$ and c_{12} be the constants from Proposition 7.2. Let us assume for the moment that there exists an integer t with $t_0 \leq t \leq n$ such that for all $\Sigma \subset \{1, \ldots, t\}$ with $|\Sigma| \geq \mathrm{s}^{\mathrm{o}}(X_t) - \dim X_t$ we have the inequality

$$c_{11}^{-1} \left(\prod_{k \in \Sigma} h(p^{u_k}) \right) \deg(X_t) (\log(3 \deg(X_t)))^{c_{12}} < 1. \tag{7.5.8}$$

The product in (7.5.8) is actually a product over heights of certain coordinates of the point $\varphi_t(p) \in X_t$. From (7.5.8) and Proposition 7.2 we conclude that $\varphi_t(p)$ is contained in $Z_{\varphi,t}$, a proper and Zariski closed subset of X_t. As mentioned above, the set of possible φ_t and X_t is finite. Finally $\varphi_t|_X : X \to X_t$ is by definition a dominant morphism. The point p is therefore contained in the proper, Zariski closed subset $\bigcup_{\varphi,t} \varphi_t|_X^{-1}(Z_{\varphi,t})$, where the union is taken over a finite set independent of p. The proposition follows in this case.

What if p does not satisfy the property described around (7.5.8)? Then we will derive a contradiction. Let $t_0 \leq t \leq n$ and let $\Sigma \subset \{1, \ldots, t\}$, we define $f_{n-t+1}(\Sigma) \in \mathbf{R}$ to be the expression on the left-hand side of (7.5.8). We are assuming that for all $t_0 \leq t \leq n$ there exists a subset $\Sigma(t) \subset \{1, \ldots, t\}$ with $|\Sigma(t)| \geq \mathrm{s}^{\mathrm{o}}(X_t) - \dim X_t$ and

$$f_{n-t+1}(\Sigma(t)) \geq 1. \tag{7.5.9}$$

For brevity we set $f_{n-t+1} = f_{n-t+1}(\Sigma(t))$.

We continue by bounding f_{n-t+1} from above. To do this we apply (7.5.5) to the definition of f_{n-t+1} and get

$$f_{n-t+1} \leq c_{11}^{-1} \left(\prod_{k \in \Sigma(t)} (c_6' B \rho^{-1} \lambda_k) \right) \deg(X_t) (\log(3 \deg(X_t)))^{c_{12}}.$$

Next we bound $\deg(X_t)$ for above using (7.5.7):

$$f_{n-t+1} \leq c_7' \left(\prod_{k \in \Sigma(t)} \lambda_k \right) \lambda_{t-r+1} \cdots \lambda_t (B\rho^{-1})^{|\Sigma(t)|} \deg(X) (\log(3\lambda_t \deg(X)))^{c_{12}}.$$

By (7.5.6) we have $|\Sigma(t)| \geq \mathrm{s}^{\mathrm{o}}(X_t) - \dim X_t \geq s+t-n-r \geq 1$ where we used $\dim X_t \leq r$. Since $\lambda_k \geq 1$, $\rho \geq B$, and $\prod_{k \in \Sigma(t)} \lambda_k \leq \lambda_1 \cdots \lambda_t$ we deduce

$$\begin{aligned} f_{n-t+1} \leq \\ c_7' \lambda_1 \cdots \lambda_{t-r} (\lambda_{t-r+1} \cdots \lambda_t)^2 (B\rho^{-1})^{s+t-n-r} \deg(X) (\log(3\lambda_t \deg(X)))^{c_{12}}. \end{aligned} \tag{7.5.10}$$

And this inequality holds for all $t_0 \leq t \leq n$.

Let M, v, w be the matrix respectively vectors from Lemma 7.7. Using notation from chapter 6, section 2 we define

$$\begin{aligned} \Lambda &= (\lambda_1, \ldots, \lambda_n)^M \\ &= (\lambda_1 \cdots \lambda_{n-r} (\lambda_{n-r+1} \cdots \lambda_n)^2, \ldots, \lambda_1 \cdots \lambda_{n-s+1} (\lambda_{n-s+2} \cdots \lambda_{n+r-s+1})^2). \end{aligned}$$

That is, the jth entry of Λ is the main contribution of the λ_k's to the bound for f_j in (7.5.10). By Lemma 7.7 and $\lambda_k \geq 1$ we have

$$\Lambda^w = (\lambda_1, \dots, \lambda_n)^{Mw} = \lambda_1^{v_1} \cdots \lambda_n^{v_n} \leq \lambda_1 \cdots \lambda_n. \tag{7.5.11}$$

We define the product

$$f = (f_1, \dots, f_{s-r})^w = f_1^{w_1} \cdots f_{s-r}^{w_{s-r}}.$$

By (7.5.9) and since $w_j \geq 0$ we conclude

$$f \geq 1. \tag{7.5.12}$$

We bound f from above with the help of (7.5.10) and (7.5.11)

$$\begin{aligned} f &\leq c_8' \Lambda^w (B\rho^{-1})^{\sum_{j=1}^{s-r}(s-r-j+1)w_j} (\deg(X)(\log(3\lambda_n \deg(X)))^{c_{12}})^{w_1+\cdots+w_{s-r}} \\ &\leq c_8' \lambda_1 \cdots \lambda_n (B\rho^{-1})^{\sum_{j=1}^{s-r}(s-r-j+1)w_j} \deg(X)(\log(3\lambda_n \deg(X)))^{c_{12}}. \end{aligned}$$

We also used $w_1 + \cdots + w_{s-r} < 1$, which follows easily from the definition of the w_j in Lemma 7.7, cf. (7.5.2). By the same lemma the exponent of $B\rho^{-1}$ equals $\mathfrak{m}(r,s)$. We recall (7.5.4) and observe $\lambda_n \leq \lambda_1 \cdots \lambda_n \leq c_3' \rho^m$ to bound

$$f \leq c_9' \rho^{m-\mathfrak{m}(r,s)} B^{\mathfrak{m}(r,s)} \deg(X)(\log(3\rho \deg(X)))^{c_{12}}. \tag{7.5.13}$$

By the definition of ρ we have

$$\rho^{m-\mathfrak{m}(r,s)} B^{\mathfrak{m}(r,s)+\delta} \deg(X)^{1+\delta} = c_{13}^{\mathfrak{m}(r,s)-m} \tag{7.5.14}$$

Together with (7.5.13) we get

$$f \leq c_9' c_{13}^{\mathfrak{m}(r,s)-m} \frac{(\log(3\rho \deg(X)))^{c_{12}}}{B^\delta \deg(X)^\delta}.$$

By (7.5.14) the ρ in the logarithm is bounded above by $(Bc_{13}^{-1} \deg(X))^{c_{10}'}$. Together with elementary inequalities we get

$$\begin{aligned} f &\leq c_{11}' c_{13}^{\mathfrak{m}(r,s)-m} \frac{(\log(3Bc_{13}^{-1} \deg(X)))^{c_{12}}}{B^\delta \deg(X)^\delta} \\ &\leq c_{12}' c_{13}^{\mathfrak{m}(r,s)-m} \frac{(Bc_{13}^{-1} \deg(X))^{\min\{\frac{\mathfrak{m}(r,s)-m}{2}, \delta\}}}{B^\delta \deg(X)^\delta} \\ &\leq c_{12}' c_{13}^{\frac{\mathfrak{m}(r,s)-m}{2}}. \end{aligned} \tag{7.5.15}$$

Thus we may choose $c_{13} \in (0,1]$ depending only on n and δ such that (7.5.15) implies $f < 1$. But this inequality contradicts (7.5.12). □

For $B \geq 1$ we have the following inclusions

$$\begin{aligned} \{p \in T(\mathcal{H}_m^n, \epsilon);\ h(p) \leq B\} &\subset \{p \in C(\mathcal{H}_m^n, \epsilon);\ h(p) \leq B\} \\ &\subset \{p \in T(\mathcal{H}_m^n, 4\epsilon B);\ h(p) \leq B\}, \end{aligned} \tag{7.5.16}$$

if $\epsilon \leq \frac{1}{2n}$. The first inclusion is trivial and holds for unrestricted ϵ, the second one follows easily from (6.4.1). Therefore Proposition 7.3 can be reformulated with $T(\cdot,\cdot)$ replaced by $C(\cdot,\cdot)$ and after choosing a possibly smaller ϵ.

PROPOSITION 7.4. *Let X, r, s, m, n, B, and δ be as in Proposition 7.3. Then there exists $c_{14} > 0$ which depends only on n and δ such that if*

$$\epsilon \leq c_{14}(B^{\mathfrak{m}(r,s)+\delta}(\deg(X))^{1+\delta})^{-\frac{1}{\mathfrak{m}(r,s)-m}}$$

the set $\{p \in X \cap C(\mathcal{H}^n_m, \epsilon);\ h(p) \leq B\}$ is not Zariski dense in X.

PROOF. The proof follows immediately from the second inclusion in (7.5.16) and Proposition 7.3. □

LEMMA 7.8. *Let n, n', r, r' be integers with $1 \leq r \leq n-1$, $1 \leq r' \leq n'-1$, $r' \leq r$, and $n - r \leq n' - r'$, then $\mathfrak{m}(r,n) \leq \mathfrak{m}(r',n')$.*

PROOF. From (7.5.1) and taking $j = kr + 1$ we get

$$\mathfrak{m}(r,n) = \sum_{k=0}^{\infty} \max\{0, \frac{n-(k+1)r}{2^{k+1}}\}.$$

As

$$n-(k+1)r = (n-r) - kr \leq (n'-r') - kr' = n' - (k+1)r'$$

for non-negative k, the result follows at once. □

LEMMA 7.9. *Let $X \subsetneq \mathbf{G}^n_m$ be an irreducible closed subvariety of dimension $r \geq 1$. Let m be an integer with $0 \leq m < \mathfrak{m}(r,n)$ and let $B \geq 1$. Say $Z \subset X$ is a closed irreducible subvariety of positive dimension with $Z \cap X^{\mathrm{oa}} \neq \emptyset$. Then there exists $\epsilon > 0$ such that*

$$\{p \in Z \cap C(\mathcal{H}^n_m, \epsilon);\ h(p) \leq B\}$$

is not Zariski dense in Z.

PROOF. The hypothesis $Z \cap X^{\mathrm{oa}} \neq \emptyset$ implies that Z is not an anomalous subvariety of X. Say $H \subset \mathbf{G}^n_m$ is a coset containing Z with $\mathrm{s}^{\mathrm{o}}(Z) = \dim H$. By hypothesis we have $\dim Z \leq r + \dim H - n$ or

$$\mathrm{s}^{\mathrm{o}}(Z) - \dim Z \geq n - r \tag{7.5.17}$$

and in particular $\mathrm{s}^{\mathrm{o}}(Z) \geq \dim Z + 1$ since X is proper. Inequalities (7.5.17), $\dim Z \leq \dim X$, and Lemma 7.8 imply $\mathfrak{m}(\dim Z, \mathrm{s}^{\mathrm{o}}(Z)) \geq \mathfrak{m}(r,n) > m$. Therefore the lemma follows from Proposition 7.4 applied to Z. □

Before proving Theorem 7.1 we prove Corollaries 7.1 and 7.2.

We start with Corollary 7.1(i). Let X be as in the hypothesis of part (i). We may assume $X^{\mathrm{oa}} = X$, i.e. X is not contained in a proper coset. By Theorem 6.1 with $s = \dim X = 1$ the set $X \cap C(\mathcal{H}^n_{n-1}, \epsilon)$ has height bounded by B for an $\epsilon > 0$ such that B depends only on $h_V(X)$, $\deg(X)$, and n. Moreover, ϵ depends only on $\deg(X)$ and n. Since $\mathfrak{m}(1,n) > n-2$ we conclude finiteness of $X \cap C(\mathcal{H}^n_{n-2}, \epsilon)$ from Theorem 7.1(i)

after choosing a possibly smaller $\epsilon > 0$ which now depends only on $h_V(X)$, $\deg(X)$, and n.

The proof of part (ii) is similar.

For part (iii) we may assume $r = \dim X \geq 1$. In section 1 we saw that $1 \leq r \leq n-3$ implies $\mathfrak{m}(r,n) > 1$. Hence the proof follows from taking $m = 1$ in part (ii). □

To prove Corollary 7.2(i) we note that by Corollary 7.1(i) we may choose $\epsilon > 0$ such that $X^{\mathrm{oa}} \cap C(\mathcal{H}^n_{n-2}, \epsilon)$ is finite. Hence the points of this set have height bounded by some fixed B and degree bounded by some fixed D. The proof now follows easily from Lemma 7.3 after adjusting ϵ if necessary.

The proofs of parts (ii) and (iii) are similar. □

Proof of Theorem 7.1: We may assume $m \geq 0$, thus $X \neq \mathbf{G}^n_m$ since $\mathfrak{m}(n,n) = 0$. Part (i) follows from Proposition 7.4 and $\mathrm{s}^{\mathrm{o}}(X) = n$, so it remains to prove part (ii). We will prove the following statement:

Let $Z \subset X$ be an irreducible closed subvariety, then there exists $\epsilon > 0$ such that

$$\{p \in X^{\mathrm{oa}} \cap Z \cap C(\mathcal{H}^n_m, \epsilon);\ h(p) \leq B\} \tag{7.5.18}$$

is finite.

Of course the theorem follows by taking $Z = X$ in the statement above. We prove the statement by induction on $\dim Z$. The case $\dim Z = 0$ being trivial we assume $\dim Z \geq 1$ and also $X^{\mathrm{oa}} \cap Z \neq \emptyset$. By Lemma 7.9 there exists an $\epsilon > 0$ such that

$$\{p \in X^{\mathrm{oa}} \cap Z \cap C(\mathcal{H}^n_m, \epsilon);\ h(p) \leq B\} \subset Y_1 \cup \cdots \cup Y_l \tag{7.5.19}$$

where the $Y_i \subsetneq Z$ are proper, Zariski closed, and irreducible. As $\dim Y_i \leq \dim Z - 1$ we reduce ϵ if necessary and apply the induction hypothesis to conclude that $\{p \in X^{\mathrm{oa}} \cap Y_i \cap C(\mathcal{H}^n_m, \epsilon);\ h(p) \leq B\}$ is finite for each i. The statement around (7.5.18) now follows from (7.5.19). □

APPENDIX A

Quasi-equivalence of heights

Say two algebraic number x and y satisfy a simple equation such as $x^p + y^q = 0$ with p and q positive integers. Then the elementary height properties described in chapter 1 imply that the heights are related as $ph(x) = qh(y)$. Now say $P(x, y) = 0$ where P is an irreducible polynomial in two variables with algebraic coefficients and $p = \deg_X P$, $q = \deg_Y P$, then how are $h(x)$ and $h(y)$ related? In this generality the expression $|ph(x) - qh(y)|$ need not be bounded independently of x and y. Nevertheless in this appendix we prove an upper bound in terms of $\max\{h(x), h(y)\}$, $h_p(P)$, p, and q.

1. Introduction

Let C be a smooth projective curve defined over $\overline{\mathbf{Q}}$. We assume that C is embedded into projective space $\mathbf{P}^n$. Let $f \in \overline{\mathbf{Q}}(C)$ be a non-constant rational function on C. We consider f as a rational map $C \to \mathbf{A}^1$. Then f has a well-defined degree $\deg(f)$, the cardinality of $f^{-1}(q)$ for generic q. Let $g \in \overline{\mathbf{Q}}(C)$ be a further non-constant rational function. We would like to compare the heights $h(f(p))$ and $h(g(p))$ as p runs over the algebraic points of C that are not poles of f or g. It was known to Siegel that the quantities $h(f(p))/\deg(f)$ and $h(g(p))/\deg(g)$ are asymptotically equal. One can even show

$$\left| \frac{h(f(p))}{\deg(f)} - \frac{h(g(p))}{\deg(g)} \right| = O(\max\{1, h(f(p)), h(g(p))\}^{1/2}). \tag{A.1.1}$$

The implied constant may depend on C, f, and g. The inequality (A.1.1) can for example be proved using the Néron-Tate height on the Jacobian of C if our curve has positive genus.

As f and g are rational functions on the curve C they are contained in a field of transcendence degree 1 over $\overline{\mathbf{Q}}$. Hence f and g satisfy some algebraic relation. That is, there exists a polynomial $P \in \overline{\mathbf{Q}}[X, Y]$ with $P(f, g) = 0$ on C. We may assume that P is irreducible over $\overline{\mathbf{Q}}$ and in this case we have

$$\frac{\deg_X P}{\deg_Y P} = \frac{\deg g}{\deg f}.$$

Hence inequality (A.1.1) follows from the following statement:

Let $P \in \overline{\mathbf{Q}}[X, Y]$ be irreducible over $\overline{\mathbf{Q}}$ with $p = \deg_X P > 0$ and $q = \deg_Y P > 0$. There exists a constant $c_1 = c_1(P)$ such that if x and y are algebraic numbers with

$P(x,y)=0$, then

$$\left|\frac{h(x)}{q}-\frac{h(y)}{p}\right|\leq c_1\max\{1,h(x),h(y)\}^{1/2}. \tag{A.1.2}$$

The purpose of this appendix is to give a proof of (A.1.2) and to determine a constant c_1 which is completely explicit in terms of P. We will strive for a good dependency in the height of P and the partial degrees p and q. For example if we assume that $h_p(P)$ is large compared with $\max\{p,q\}$, then we will show that c_1 depends only on $h_p(P)$.

Property (A.1.2) is often referred to as quasi-equivalence of heights. We will now state the Quasi-equivalence Theorem:

THEOREM A.1. *Let $P\in\overline{\mathbf{Q}}[X,Y]$ be irreducible with $p=\deg_X P>0$ and $q=\deg_Y P>0$. If $P(x,y)=0$ with $x,y\in\overline{\mathbf{Q}}$, then*

$$\left|\frac{h(x)}{q}-\frac{h(y)}{p}\right|\leq 51\max\{p,q,h_p(P)\}^{1/2}\max\{1,h(x),h(y)\}^{1/2}.$$

In [**Abo06**] Abouzaid proved a variant of the Quasi-equivalence Theorem. The dependency in $h_p(P)$ and $\max\{h(x),h(y)\}$ is essentially the same in both version. Though in Abouzaid's bound, the dependence on the partial degrees is slightly worse.

It is essential that P is irreducible. For example the partial degrees of $(X^2-Y)(X-Y^2)$ are equal to 3, but clearly (A.1.2) cannot hold for this polynomial. Although it is possible to formulate a version of Theorem A.1 with $P\in K[X,Y]$ where K is a number field and P is irreducible over K. In this case P is up to a scalar factor the product of polynomials which are irreducible in $\overline{\mathbf{Q}}[X,Y]$ and conjugated over K. Thus all factors have equal partial degrees.

Sometimes it is useful to bound $h(y)$ uniformly in terms of $h(x)$ if the height of x is large:

COROLLARY A.1. *Let P be as in Theorem A.1. Say $P(x,y)=0$ with $x,y\in\overline{\mathbf{Q}}$ and*

$$h(x)\geq 2.10^4(\deg P)^2\max\{p,q,h_p(P)\}, \tag{A.1.3}$$

then $h(y)\leq 2\frac{p}{q}h(x)$.

The proof of Theorem A.1 depends on the Absolute Siegel Lemma by Roy and Thunder and some elementary theory of algebraic functions. We note that a proof along the lines of the proof of Theorem A.1 is possible using Bombieri and Vaaler's classical version of Siegel's Lemma. But this comes at the cost of introducing a dependency on the field of definition of P, more precisely the discriminant of this field, in c_1 of (A.1.2).

We give a short sketch of the proof of Theorem A.1. Let m and n be large integers with n/m approximately equal to p/q. In section 2 we use the Absolute Siegel Lemma to construct polynomials A and B in two variables with algebraic coefficients of bounded height such that P divides AY^m-B as a polynomial. Furthermore, we require that $\deg_X A,\deg_X B\leq n$ and $\deg_Y A,\deg_Y B<\deg_Y P$. If $P(x,y)=0$, then $A(x,y)y^m=B(x,y)$. If we assume for the moment $A(x,y)\neq 0$, then we may bound the height of y in terms of the height of x by using the product formula and the fact that m is much

larger than $\deg_Y P$. In section 3 we show that if $A(x,y)=0$ then the order of vanishing cannot be too large if m and n are well chosen. We then differentiate appropriately and replace A, B by new polynomials A', B' with controlled height and degree such that $A'(x,y)y^m = B'(x,y)$ and $A'(x,y) \neq 0$. Thus again we get a height bound of y. By interchanging x and y we can bound the height of x by the height of y and this completes the proof. We note that the proof does not depend on an explicit version of Eisenstein's Theorem.

2. Construction via Siegel's Lemma

If K is a number field with an absolute value $|\cdot| = |\cdot|_v$ ($v \in M_K$) and P is a polynomial with coefficients in K in any number of variables, then, just as in chapter 1, we define $|P|$ to be the maximum of the absolute values of the coefficients of P. We also recall that the height $h_p(P)$ was defined in chapter 1. If Q is a further polynomial with algebraic coefficients it will sometimes be useful to define $h_p(P,Q)$ as the projective height of the vector consisting of the coefficients of P and Q. Finally δ_v and d_v are defined in chapter 1.

We start off by proving a standard lemma concerning simple properties of absolute values and heights.

LEMMA A.1. *Let K be a field and $A, B \in K[X,Y]$.*

(i) *If K is a number field and $v \in M_K$, then*

$$|A+B|_v \leq \delta_v(2)\max\{|A|_v, |B|_v\},$$
$$|AB|_v \leq \delta_v(1+\min\{\deg_X A, \deg_Y A, \deg_X B, \deg_Y B\})|A|_v|B|_v.$$

(ii) *If $K = \overline{\mathbf{Q}}$ and $x, y \in \overline{\mathbf{Q}}$ with $A(x,y) \neq 0$, then*

$$h(B(x,y)/A(x,y)) \leq h_p(A,B) + \max\{\deg_X A, \deg_X B\}h(x) + \max\{\deg_Y A, \deg_Y B\}h(y) + \log\max\{(1+\deg_X A)(1+\deg_Y A), (1+\deg_X B)(1+\deg_Y B)\}$$

(iii) *If $K = \overline{\mathbf{Q}}$, $x, y \in \overline{\mathbf{Q}}$ with $A(x,y) = 0$ and assume furthermore that A is not divisible in $\overline{\mathbf{Q}}[X,Y]$ by $X - \alpha$ for some $\alpha \in \overline{\mathbf{Q}}$, then*

$$h(y) \leq (\deg_X A)h(x) + \log((1+\deg_X A)\deg_Y A) + h_p(A). \tag{A.2.1}$$

PROOF. The first inequality in (i) follows directly from the triangle inequality. To prove the second inequality we write $A = \sum_{i,j} a_{ij}X^iY^j$ and $B = \sum_{i,j} b_{ij}X^iY^j$. Then $AB = \sum_{i,j} c_{ij}X^iY^j$ with

$$c_{ij} = \sum_{i'+i''=i, j'+j''=j} a_{i'j'}b_{i''j''}.$$

The sum above involves at most $1 + \min\{\deg_X A, \deg_Y A, \deg_X B, \deg_Y B\}$ non-zero terms. Hence the desired inequality follows from the triangle inequality.

Now to part (ii). We note that the product formula (1.1.1) implies

$$h(B(x,y)/A(x,y)) = [F:\mathbf{Q}]^{-1} \sum_{v\in M_F} d_v \log \max\{|A(x,y)|_v, |B(x,y)|_v\} \tag{A.2.2}$$

where F is a number field containing x, y and the coefficients of A and B. Note that the polynomial A involves at most $(1+\deg_X A)(1+\deg_Y A)$ non-zero coefficients, hence the triangle inequality gives

$$|A(x,y)|_v \leq \\ \delta_v((1+\deg_X A)(1+\deg_Y A))|A|_v \max\{1,|x|_v\}^{\deg_X A} \max\{1,|y|_v\}^{\deg_Y A}$$

for each $v \in M_F$. Of course a similar inequality holds for $|B(x,y)|_v$. These inequalities inserted into (A.2.2) conclude this part of the lemma.

Our proof for Part (iii) follows the lines of Proposition 5 in [**BM06**]: say $A = a_q Y^q + \cdots + a_0$ with $a_i \in \overline{\mathbf{Q}}[X]$ and $a_q \neq 0$. Because of hypothesis there exists $q' \geq 1$ maximal such that $a_{q'}(x) \neq 0$. Let F be a number field that contains x, y, and the coefficients of A. If $|\cdot| = |\cdot|_v$ is an absolute value on F, then

$$|a_{q'}(x)y^{q'}| \leq \delta_v(q') \max_{0\leq k\leq q'-1}\{|a_k(x)|\} \max\{1,|y|\}^{q'-1}$$

and so

$$\max\{1,|y|\} \leq \delta_v(q') \max_{0\leq k\leq q'}\{|a_k(x)|/|a_{q'}(x)|\}.$$

We use the last inequality and the product formula to show

$$h(y) \leq \log q + [F:\mathbf{Q}]^{-1} \sum_{v\in M_F} d_v \log \max_{0\leq k\leq q'}\{|a_k(x)|_v\}. \tag{A.2.3}$$

By the triangle inequality we have

$$|a_k(x)|_v \leq \delta_v(1+\deg_X P)\max\{1,|x|_v\}^{\deg_X P}|A|_v$$

for any absolute value on F. We apply this inequality to (A.2.3) to complete the proof. □

We will often apply property (iii) of the previous lemma to a non-zero $A \in \overline{\mathbf{Q}}[Y]$ and $x = 0$. Inequality (A.2.1) then reduces to $h(y) \leq \log \deg_Y A + h_p(A)$.

We need to introduce a crude notion of a sparsity: if $\mathcal{A} = (a_{ij})$ is an $M \times N$ matrix, then we set

$$S(\mathcal{A}) = \max_{1\leq i\leq M} |\{j;\ a_{ij} \neq 0\}|.$$

If $\mathcal{A}$ has algebraic coefficients and is non-zero we define the height $h_p(\mathcal{A})$ as the height of $[a_{ij}; 1 \leq i \leq M, 1 \leq j \leq N] \in \mathbf{P}^{MN-1}(\overline{\mathbf{Q}})$.

Next we adapt Roy and Thunder's Absolute Siegel Lemma from [**RT96**] to our notation. But first we remark that a version due to David and Philippon (Lemma 4.7 in [**DP99**]) which uses Zhang's Theorem 5.2 from [**Zha95a**] would also suffice.

LEMMA A.2. *Let* $\mathcal{A} = (a_{ij}) \in Mat_{MN}(\overline{\mathbf{Q}})$, $\operatorname{rank}\mathcal{A} = M < N$. *Then there exists* $v \in \mathbf{P}^{N-1}(\overline{\mathbf{Q}})$ *such that* $\mathcal{A}v = 0$ *and*

$$h(v) \leq \frac{\log 2}{2}\frac{N}{N-M} + \frac{M}{N-M}(\log S(\mathcal{A}) + h_p(\mathcal{A})) + \frac{N-M}{4}.$$

PROOF. We apply Theorem 1 of [**RT99**] (cf. [**RT96**]) which states that there exists $v \in \mathbf{P}^{N-1}(\overline{\mathbf{Q}})$ with $\mathcal{A}v = 0$ and

$$h(v) \leq \frac{1}{N-M}\log H(V) + \frac{N-M}{4} \tag{A.2.4}$$

where $V \subset \overline{\mathbf{Q}}^N$ is the kernel of $\mathcal{A}$ and $H(V)$ is the height of the vector space V as defined in [**RT96**]. We note that the (logarithm of the) height used in [**RT99**] and [**RT96**] uses the Euclidean norm at infinite places and is hence at least as big as our notion of height which uses the infinity norm at all places. By Theorem 1.1 in [**RT96**] $H(V)$ is equal to the height of the vector space $W \subset \overline{\mathbf{Q}}^N$ spanned by the rows of $\mathcal{A}$. Because the rank of $\mathcal{A}$ is M, the rows of $\mathcal{A}$ are a basis of W. The height $H(W)$ is then just the height of the projective vector composed of the determinants of all $M \times M$ minors of $\mathcal{A}$ with the Euclidean norm taken at the infinite places and maximum norm at the finite places. A simple application of the triangle inequality and the fact that the expansion of each determinant in question involves at most $S(\mathcal{A})^M$ non-zero terms gives

$$\log H(V) = \log H(W) \leq \frac{1}{2}\log\binom{N}{M} + M(\log S(\mathcal{A}) + h_p(\mathcal{A})).$$

The proof follows from this inequality, (A.2.4), and $\binom{N}{M} \leq 2^N$. □

LEMMA A.3. *Let* $P \in \overline{\mathbf{Q}}[X,Y]$ *with* $p = \deg_X P > 0$ *and* $q = \deg_Y P > 0$. *Furthermore, let m,n be integers with* $m \geq q$,$n \geq p$ *and define* $t = q(n+1) - mp$. *If* $t \geq 1$, *then there exist* $A, B \in \overline{\mathbf{Q}}[X,Y]$ *with*

$$AY^m - B \in P \cdot \overline{\mathbf{Q}}[X,Y]\backslash\{0\},\ \deg_X A, \deg_X B \leq n,\ \deg_Y A, \deg_Y B \leq q-1 \tag{A.2.5}$$

and

$$h_p(A,B) \leq \frac{mn}{t}(\log(\sqrt{8}p) + h_p(P)) + \frac{nq}{2}.$$

PROOF. Let $\mathfrak{Q} = \sum_{k,l} q_{kl}X^kY^l \in \mathbf{Z}[X,Y,q_{kl}]$ with $\deg_X \mathfrak{Q} = n-p$ and $\deg_Y \mathfrak{Q} = m-1$. Define the linear forms $f_{ij} \in \overline{\mathbf{Q}}[q_{kj}]$ $(0 \leq i \leq n,\ 0 \leq j \leq m+q-1)$ by

$$P\mathfrak{Q} = \sum_{i,j} f_{ij}X^iY^j.$$

We note that f_{ij} is a linear form in the q_{kj} such that its non-zero coefficients are coefficients of P. So

$$f_{ij} = 0, \quad \text{for} \quad 0 \leq i \leq n, \quad q \leq j \leq m-1 \tag{A.2.6}$$

is a system of linear equations with rank M in the

$$N = m(n-p+1)$$

variables q_{ij}. Clearly one has

$$M \leq (n+1)(m-q) = N - t. \tag{A.2.7}$$

Because $N - M \geq t \geq 1$ there is a non-trivial solution. Any such solution gives rise to a non-zero polynomial $Q \in \overline{\mathbf{Q}}[X,Y]$ such that the coefficients of PQ satisfy (A.2.6) and hence $PQ = AY^m - B$ for unique polynomials $A, B \in \overline{\mathbf{Q}}[X,Y]$ with $\deg_X A, \deg_X B \leq n$ and $\deg_Y A, \deg_Y B \leq q-1$.

A peculiarity of Roy and Thunder's version of Siegel's Lemma is that the second term in upper bound (A.2.4) works against us if the rank M is small. We work out a lower bound for M: a non-trivial linear combination of X^iY^jP where $0 \leq i \leq n-p$ and $q \leq j \leq m-q-1$ is not of the form $AY^m - B$ with A and B satisfying the degree bounds in (A.2.5). Therefore we get a lower bound $M \geq (n-p+1)\max\{0, m-2q\}$, so

$$N - M \leq (n-p+1)(m - \max\{0, m-2q\}) \leq 2nq. \tag{A.2.8}$$

We will apply Siegel's lemma to find a solution Q with small height. If $M = 0$, nothing has to be done as $Q = 1$ is possible. So say $M \geq 1$. We choose a subset of the linear forms f_{ij} ($0 \leq i \leq n$, $q \leq j < m$) with rank M. We use the coefficients of each such linear form to define a row in the $M \times N$ matrix $\mathcal{A}$. It was noted that the non-zero entries of $\mathcal{A}$ are coefficients of P, hence $h_p(\mathcal{A}) \leq h_p(P)$. Furthermore, it is clear from the definition that each f_{ij} involves at most $1+\min\{p,q\}$ non-zero coefficients and hence $S(\mathcal{A}) \leq 2p$. By Lemma A.2 and our discussion above there exists a non-zero solution $Q \in \overline{\mathbf{Q}}[X,Y]$ of (A.2.6) that satisfies

$$h_p(Q) \leq \frac{\log 2}{2}\frac{N}{N-M} + \frac{M}{N-M}(\log(2p) + h_p(P)) + \frac{N-M}{4}.$$

Lemma A.1(i) implies $h_p(PQ) \leq \log(2p) + h_p(P) + h_p(Q)$. Furthermore, we use the inequalities (A.2.7) and (A.2.8) to conclude

$$\begin{aligned} h_p(PQ) &\leq \log(2p) + h_p(P) + \frac{\log 2}{2}\frac{N}{t} + \frac{M}{t}(\log(2p) + h_p(P)) + \frac{nq}{2} \\ &\leq \frac{M+t}{t}\log(2p) + \frac{\log 2}{2}\frac{N}{t} + \frac{M+t}{t}h_p(P) + \frac{nq}{2} \\ &\leq \frac{N}{t}(\log(\sqrt{8}p) + h_p(P)) + \frac{nq}{2}. \end{aligned}$$

This inequality completes the proof because $N \leq mn$. □

Let A, B, and P be as in the lemma above, then P cannot divide A. Indeed assuming the contrary, then P also divides B. Because $\deg_Y A, \deg_Y B < \deg_Y P$ we have $A = B = 0$, a contradiction to $AY^m - B \neq 0$. For similar reasons and since $\deg_X P > 0$, P cannot divide B.

3. Zero bounds

We need only the most basic facts about function fields which we recall here for the reader's convenience.

Let F be a field which is an extension of an algebraically closed field L. Assume that there exists an element $t \in F$ transcendental over L such that F is a finite field extension of $L(t)$. Then F is a function field over L. We define M_F to be the set of the maximal ideals of all the proper valuation rings of F containing L. This set is the function field analogue of M_K for a number field K. We will identify an element of M_F with the valuation function it induces. Hence the elements of M_F are surjective maps $v : F^* \to \mathbf{Z}$ with $v(ab) = v(a) + v(b)$ if $a, b \in F^*$, $v(a+b) \geq \min\{v(a), v(b)\}$ if $a, b, a+b \in F^*$, and $v(\lambda) = 0$ if $\lambda \in L^*$.

We have the following properties: if $a \in F^*$ we have $v(a) = 0$ for all but finitely many $v \in M_F$ and

$$\sum_{v \in M_F} v(a) = 0.$$

Furthermore, if $a \in F \backslash L$ then

$$\sum_{v \in M_F} \max\{0, v(a)\} = [F : L(a)].$$

For the rest of this section $P \in \overline{\mathbf{Q}}[X, Y]$ will, unless stated otherwise, be a fixed irreducible polynomial and F will denote the quotient field of the domain $\overline{\mathbf{Q}}[X, Y]/(P)$; then F is a function field over $L = \overline{\mathbf{Q}}$. By abuse of notation we shall consider polynomials in $\overline{\mathbf{Q}}[X, Y]$ as elements of F via the natural map. Note that any polynomial in $\overline{\mathbf{Q}}[X, Y]$ that is not divisible by P maps to F^*.

Let $\pi = (x, y) \in \overline{\mathbf{Q}}^2$ with $P(\pi) = 0$ such that $\frac{\partial P}{\partial X}, \frac{\partial P}{\partial Y}$ do not both vanish at π, then we call π a *regular zero* of P. Let us assume for the moment $\pi = (x, y)$ and $\frac{\partial P}{\partial Y}(\pi) \neq 0$, then there exist a unique $v_\pi \in M_F$ with $v_\pi(X - x) = 1$ and $v_\pi(Y - y) > 0$. There exists E in $\overline{\mathbf{Q}}[[T]]$, the ring of formal power series, such that $E(0) = 0$ and $P(x+T, y+E) = 0$. For any $A \in \overline{\mathbf{Q}}[X, Y]$ not divisible by P we have

$$\operatorname{ord} A(x + T, y + E) = v_\pi(A)$$

where ord is the usual valuation on $\overline{\mathbf{Q}}[[T]]$. Therefore $v_\pi(A) = 0$ if and only if $A(x, y) \neq 0$.

Of course if $\frac{\partial P}{\partial X}(\pi) \neq 0$ then the results in the last paragraph hold with the roles of X and Y reversed.

Proofs for these statements can be found in [**Che51**].

Let $A \in \overline{\mathbf{Q}}[X, Y]$, we define

$$D(A) = \frac{\partial P}{\partial Y}\frac{\partial A}{\partial X} - \frac{\partial P}{\partial X}\frac{\partial A}{\partial Y} \in \overline{\mathbf{Q}}[X, Y].$$

We also set $D^0(A) = A$ and inductively $D^s(A) = D(D^{s-1}(A))$ for a positive integer s. A formal verification gives $D(AB) = D(A)B + AD(B)$ if B is in $\overline{\mathbf{Q}}[X, Y]$. Thus we have

Leibniz's rule:

$$D^s(AB) = \sum_{k=0}^{s} \binom{s}{k} D^k(A) D^{s-k}(B).$$

LEMMA A.4. *Let K be a number field and $P \in K[X,Y]$ with $p = \deg_X P > 0$ and $q = \deg_Y P > 0$. Furthermore, assume $v \in M_K$ and $A \in K[X,Y]$ with $\deg_X A \le n$, $\deg_Y A \le q$. Then for any non-negative $s \in \mathbf{Z}$*

$$\deg_X D^s(A) \le s(p-1) + n, \quad \deg_Y D^s(A) \le s(q-1) + q, \tag{A.3.1}$$

and if $r = \max\{p,q\}$ then

$$|D^s(A)|_v \le \delta_v (4pq(sr+n))^s |P|_v^s |A|_v. \tag{A.3.2}$$

PROOF. We note $\deg_X D(A) \le p - 1 + \deg_X A$ and so the first inequality in (A.3.1) follows by induction on s. The second inequality is proved similarly. We now prove (A.3.2) by induction on s. The case $s = 0$ being trivial we assume $s \ge 1$. For brevity set $|\cdot| = |\cdot|_v$. We apply Lemma A.1(i) to show

$$|D^s(A)| \le \delta_v(2(r'+1)) \max\{\left|\frac{\partial P}{\partial Y}\right| \left|\frac{\partial D^{s-1}(A)}{\partial X}\right|, \left|\frac{\partial P}{\partial X}\right| \left|\frac{\partial D^{s-1}(A)}{\partial Y}\right|\}$$

where $r' = \min\{p,q\}$. By bounding the partial derivatives of the polynomials in a usual way we get

$$|D^s(A)| \le \delta_v(4r'r)|P||D^{s-1}(A)| \max\{\delta_v(\deg_X D^{s-1}(A)), \delta_v(\deg_Y D^{s-1}(A))\}.$$

We apply $r'r = pq$ and the inequality (A.3.1) to get

$$|D^s(A)| \le \delta_v(4pq \max\{(s-1)(p-1)+n, (s-1)(q-1)+q\})|P||D^{s-1}(A)|.$$

The expressions inside "max" are bounded above by $sr + n$. Applying the induction hypothesis completes the proof. □

LEMMA A.5. *Let $\pi = (x,y) \in \overline{\mathbf{Q}}^2$ be a regular zero of P and let $v = v_\pi \in M_F$ be the valuation described above. If $A \in \overline{\mathbf{Q}}[X,Y]$ is not divisible by P and $A(\pi) = 0$, then $D(A)$ is not divisible by P and*

$$v(D(A)) = v(A) - 1.$$

PROOF. We shall assume $\frac{\partial P}{\partial Y}(\pi) \ne 0$, the case $\frac{\partial P}{\partial X}(\pi) \ne 0$ is similar. There exists $E \in T\overline{\mathbf{Q}}[[T]]$ such that $P(x+T, y+E) = 0$ and $v(A) = \operatorname{ord} A(x+T, y+E) \ge 1$. By the chain rule we have

$$0 = \frac{d}{dT} P(x+T, y+E) = \frac{\partial P}{\partial X}(x+T, y+E) + \frac{dE}{dT} \frac{\partial P}{\partial Y}(x+T, y+E). \tag{A.3.3}$$

We use the definition of D and (A.3.3) to obtain

$$\begin{aligned}\operatorname{ord} D(A)(x+T,y+E) &= \operatorname{ord}\left(\left(\frac{\partial P}{\partial Y}\frac{\partial A}{\partial X}-\frac{\partial P}{\partial X}\frac{\partial A}{\partial Y}\right)(x+T,y+E)\right)\\ &= \operatorname{ord}\frac{\partial P}{\partial Y}(x+T,y+E)+\operatorname{ord}\left(\frac{\partial A}{\partial X}(x+T,y+E)+\frac{dE}{dT}\frac{\partial A}{\partial Y}(x+T,y+E)\right).\end{aligned}$$

Note that by hypothesis we have

$$\operatorname{ord}\frac{\partial P}{\partial Y}(x+T,y+E)=0. \tag{A.3.4}$$

We insert (A.3.4) into the equality above and use the chain law to get

$$\operatorname{ord} D(A)(x+T,y+E)=\operatorname{ord}\frac{d}{dT}A(x+T,y+E)=\operatorname{ord} A(x+T,y+E)-1$$

hence $v(D(A))=v(A)-1$. In particular P does not divide $D(A)$. □

LEMMA A.6. *Let A,B,P,m,p,q,t be as in Lemma A.3. Furthermore, assume P is irreducible and $\deg P=p+q$. Let $\pi\in\overline{\mathbf{Q}}^2$ be a regular zero of P, then there exists an integer s with $0\le s\le t+pq-p-q$ such that $D^s(A)(\pi)\neq 0$ and $D^k(A)(\pi)=0$ for $0\le k<s$.*

PROOF. For brevity set $v=v_\pi$. Clearly $X,Y\in F\backslash\overline{\mathbf{Q}}$ since p and q are both positive. Furthermore, $A,B\neq 0$ in F by the comment after the proof of Lemma A.3. Finally $v(X),v(Y)\ge 0$.

We first claim that for any $v'\in M_F$ at least one of the two $v'(X),v'(Y)$ is non-negative. Indeed we argue by contradiction so let us assume $v'(X)<0$ and $v'(Y)<0$. Then for any integers i,j with $0\le i\le p$, $0\le j\le q$ and $i+j<p+q$ we have

$$iv'(X)+jv'(Y)>pv'(X)+qv'(Y). \tag{A.3.5}$$

Now by hypothesis $P=\alpha X^pY^q+\tilde{P}$ with $\alpha\neq 0$ and $\deg\tilde{P}<p+q$. We apply the ultrametric inequality and (A.3.5) to get

$$pv'(X)+qv'(Y)\ge\min_{\substack{0\le i\le p,0\le j\le q\\ i+j<p+q}}\{iv'(X)+jv'(Y)\}>pv'(X)+qv'(Y),$$

a contradiction.

Now assume $v'\in M_F$ such that $v'(Y)<0$. Then $v'(X)\ge 0$ by the discussion above and

$$v'(A)=v'(Y^{-m}B)=-mv'(Y)+v'(B). \tag{A.3.6}$$

Now $\deg_Y B\le q-1$ as a polynomial so we apply the ultrametric inequality to get

$$v'(B)\ge(q-1)v'(Y).$$

We insert this last inequality into (A.3.6) to find

$$v'(A)\ge(q-m-1)v'(Y)\ge -v'(Y)>0 \tag{A.3.7}$$

because $q \le m$. Hence

$$\sum_{v' \in M_F} \max\{0, v'(A)\} \ge \max\{0, v(A)\} + \sum_{v' \in M_F, v'(Y)<0} \max\{0, v'(A)\}$$

$$\ge v(A) + (m+1-q) \sum_{v' \in M_F, v'(Y)<0} \max\{0, -v'(Y)\} \tag{A.3.8}$$

where the last inequality follows from (A.3.7). Next we insert the equality

$$\sum_{v' \in M_F, v'(Y)<0} \max\{0, v'(Y^{-1})\} = [F : \overline{\mathbf{Q}}(Y^{-1})] = [F : \overline{\mathbf{Q}}(Y)] = p$$

into (A.3.8) to see

$$\sum_{v' \in M_F} \max\{0, v'(A)\} \ge v(A) + (m+1-q)p > 0. \tag{A.3.9}$$

In particular $A \notin \overline{\mathbf{Q}}$.

We continue by bounding the left-hand side of (A.3.9) from above. If $v' \in M_F$ with $v'(X) \ge 0$ and $v'(Y) \ge 0$ then $v'(A) \ge 0$ because A is a polynomial in X and Y. Hence

$$\sum_{v' \in M_F} \max\{0, -v'(A)\} \le \sum_{v' \in M_F, v'(X)<0} \max\{0, -v'(A)\} + \sum_{v' \in M_F, v'(Y)<0} \max\{0, -v'(A)\}. \tag{A.3.10}$$

Actually equality holds above because at most one $v'(X)$, $v'(Y)$ can be negative, but this is not important here. Around (A.3.7) we showed that if $v'(Y) < 0$ then $v'(A) > 0$, hence the second term on the right-hand side of (A.3.10) is zero. Now recall that $\deg_X A \le n$; if $v'(X) < 0$ then $v'(Y) \ge 0$ and the ultrametric inequality leads us to $v'(A) \ge n v'(X)$. If we insert this inequality into (A.3.10) we get

$$\sum_{v' \in M_F} \max\{0, -v'(A)\} \le n \sum_{\substack{v' \in M_F \\ v'(X)<0}} \max\{0, -v'(X)\} = n[F : \overline{\mathbf{Q}}(X^{-1})] = n[F : \overline{\mathbf{Q}}(X)] = nq. \tag{A.3.11}$$

The left-hand sides of (A.3.9) and (A.3.11) are both equal to $[F : \overline{\mathbf{Q}}(A)]$, so we get

$$v(A) \le nq - mp + pq - p = t + pq - p - q.$$

If we set $s = v(A)$, then Lemma A.5 and induction give $v(D^k(A)) = v(A) - k$ for $0 \le k \le s$. Hence $D^s(A)(x, y) \ne 0$ and $D^k(A)(x, y) = 0$ for $0 \le k < s$. □

4. Completion of proof

LEMMA A.7. *Let $P \in \overline{\mathbf{Q}}[X,Y]$ be irreducible with $p = \deg_X P > 0$ and $q = \deg_Y P$. Then there is a root of unity ξ such that the polynomial $\tilde{P} = X^p P(X^{-1}+\xi, Y)$ has total degree $p+q$, is irreducible in $\overline{\mathbf{Q}}[X,Y]$, and satisfies*

$$\deg_X \tilde{P} = p, \quad \deg_Y \tilde{P} = q, \quad h_p(\tilde{P}) \le h_p(P) + p\log 2.$$

PROOF. We may write $P = \sum_j a_j Y^j$ with $a_j = \sum_i a_{ij} X^i \in \overline{\mathbf{Q}}[X]$. By hypothesis we have $a_q \neq 0$ and a_q has degree at most p as a polynomial in X. We may choose a root of unity ξ such that $a_q(\xi) \neq 0$. A direct computation shows that $\tilde{P}$ is irreducible. And by construction $\deg_X \tilde{P} \le p$, $\deg_Y \tilde{P} \le q$, therefore $\deg \tilde{P} \le p+q$.

Say $0 \le i \le p$ and $0 \le j \le q$, then the coefficient of $X^i Y^j$ in $\tilde{P}$ equals

$$\sum_{k=p-i}^{p} a_{kj} \binom{k}{i-p+k} \xi^{i-p+k}. \tag{A.4.1}$$

So if $i = p$ and $j = q$ we see that the coefficient of $X^p Y^q$ is non-zero and conclude that $\deg_X \tilde{P} = p$, $\deg_Y \tilde{P} = q$, and $\deg \tilde{P} = p+q$.

Now say K is a number field that contains ξ and the coefficients of P. If $v \in M_K$, then by (A.4.1) and standard facts on binomial coefficients we get

$$|\tilde{P}|_v \le \delta_v \left(\sum_{k=p-i}^{p} \binom{k}{i-p+k} \right) |P|_v = \delta_v \left(\binom{p+1}{i} \right) |P|_v \le \delta_v(2^p)|P|_v.$$

These local bounds complete the proof. □

LEMMA A.8. *Let $P \in \overline{\mathbf{Q}}[X,Y]$ be irreducible with $p = \deg_X P > 0$ and $q = \deg_Y P > 0$. If $(x,y) \in \overline{\mathbf{Q}}^2$ with $P(x,y) = 0$ is not a regular zero of P then*

$$h(y) \le 2p h_p(P) + 5p\log(2pq).$$

PROOF. Let $D \in \overline{\mathbf{Q}}[Y]$ be the resultant of the two polynomials $P, \frac{\partial P}{\partial X} \in \overline{\mathbf{Q}}(Y)[X]$ (cf. [**Lan02**] page 200). Then $D \neq 0$ because P is irreducible in $\overline{\mathbf{Q}}(Y)[X]$. The resultant D is the determinant of a $(2p-1)\times(2p-1)$ matrix whose entries, denoted here by m_{ij}, are polynomials in Y with degrees bounded by q. Hence $\deg D \le (2p-1)q \le 2pq$. If K is a number field containing the coefficients of the m_{ij} and $v \in M_K$, then by Lemma A.1(i)

$$|D|_v \le \delta_v((2p-1)!) \max_\sigma \{|m_{1,\sigma(1)} \cdots m_{2p-1,\sigma(2p-1)}|_v\} \tag{A.4.2}$$

where σ runs over all permutations of the first $2p-1$ positive integers. We apply Lemma A.1(i) to bound

$$|m_{1,\sigma(1)} \cdots m_{2p-1,\sigma(2p-1)}|_v \le \delta_v(2q)^{2p-2} |m_{1,\sigma}|_v \cdots |m_{2p-1,\sigma(2p-1)}|_v.$$

This inequality and $|m_{ij}|_v \le \delta_v(p)|P|_v$ inserted into (A.4.2) gives

$$|D|_v \le \delta_v((2p-1)!(2pq)^{2p})|P|_v^{2p-1} \le \delta_v(4p^2q)^{2p}|P|_v^{2p-1} \le \delta_v(2pq)^{4p}|P|_v^{2p-1}.$$

Hence $h_p(D) \le 2ph_p(P) + 4p\log(2pq)$.

Now if $(x,y) \in \overline{\mathbf{Q}}^2$ with $P(x,y) = 0$ is not a regular zero of P, then $D(y) = 0$. By Lemma A.1(iii) we can bound $h(y) \le h_p(D) + \log\deg D \le h_p(D) + \log(2pq)$. The proof follows from this inequality together with the bound for $h_p(D)$. □

LEMMA A.9. *Let $P \in \overline{\mathbf{Q}}[X,Y]$ be irreducible with $p = \deg_X P > 0$, $q = \deg_Y P > 0$, and $\deg P = p+q$. If $(x,y) \in \overline{\mathbf{Q}}^2$ is a regular zero of P and $h(x) \ge h_p(P)$, then*

$$h(y) \le \frac{p}{q}h(x) + 16p\max\{\log(\sqrt{8}p)^2, h_p(P)\}^{1/2}\max\{1,h(x)\}^{1/2} + 17p\log(3q).$$

PROOF. For brevity we set

$$h = \max\{1,h_p(P)\}^{1/2}, \quad k = \max\{1,h(x)\}^{1/2}.$$

We note $k \ge h$ and define

$$m = pq^2\left[\frac{k}{h}\right], \quad n = m\frac{p}{q} + p - 1. \tag{A.4.3}$$

Then m and n are integers with $m \ge q$, $n \ge p$. Let t be as in Lemma A.3, we have $t = pq \ge 1$. Now let $A, B \in \overline{\mathbf{Q}}[X,Y]$ be as in Lemma A.3. Because of Lemma A.6 there exists an integer s with $0 \le s \le t + pq - p - q \le 2pq - 1$ such that $D^s(A)(x,y) \ne 0$ and $D^k(A)(x,y) = 0$ for all $0 \le k < s$. We apply Leibniz's rule to $D^s(AY^m - B)$ and the use the fact that $AY^m - B$ is divisible by P to conclude

$$y^m = \frac{D^s(B)(x,y)}{D^s(A)(x,y)}.$$

An application of Lemma A.1(ii) gives

$$\begin{aligned} mh(y) \le{}& h_p(D^s(A), D^s(B)) + \max\{\deg_X D^s(A), \deg_X D^s(B)\}h(x) \\ &+ \max\{\deg_Y D^s(A), \deg_Y D^s(B)\}h(y) \\ &+ \log\max\{(1+\deg_X D^s(A))(1+\deg_Y D^s(A)), \\ &\qquad (1+\deg_X D^s(B))(1+\deg_Y D^s(B))\}. \end{aligned} \tag{A.4.4}$$

Lemma A.4 applied to A and B implies

$$\begin{aligned} &\deg_X D^s(A), \deg_X D^s(B) \le sp + n, \\ &\deg_Y D^s(A), \deg_Y D^s(B) \le q(s+1), \text{ and} \\ &\qquad h_p(D^s(A), D^s(B)) \le h_p(A,B) + sh_p(P) + s\log(4pq(n+sr)). \end{aligned}$$

The last line follows from summing up the local bounds in (A.3.2). We insert these bounds in (A.4.4) to see

$$\begin{aligned} mh(y) \le{}& h_p(A,B) + sh_p(P) + s\log(4pq(n+sr)) + (sp+n)h(x) + q(s+1)h(y) \\ &+ \log(sp+n+1)(sq+q+1). \end{aligned}$$

Next we use the bound given for $h_p(A,B)$ in Lemma A.3 to get

$$\begin{aligned} mh(y) \le \frac{mn}{t}(\log(\sqrt{8}p) + h_p(P)) + \frac{nq}{2} + sh_p(P) + s\log(4pq(n+sr)) \\ + (sp+n)h(x) + q(s+1)h(y) + \log(sp+n+1)(sq+q+1). \end{aligned} \tag{A.4.5}$$

To control $h(y)$ on the right side of (A.4.5) we apply Lemma A.1(iii) to P; we obtain

$$h(y) \le ph(x) + h_p(P) + \log(2pq).$$

We insert the inequality above into (A.4.5) and see

$$\begin{aligned} mh(y) \le &\frac{mn}{t}(\log(\sqrt{8}p) + h_p(P)) + \frac{nq}{2} + sh_p(P) + s\log(4pq(n+sr)) \\ &+ (sp+n+pq(s+1))h(x) + q(s+1)h_p(P) + q(s+1)\log(2pq) \\ &+ \log(sp+n+1)(sq+q+1). \end{aligned}$$

We recall $h(x) \le k^2$ and $h_p(P) \le h^2$. By collecting the $h_p(P)$'s, $h(x)$'s and applying $\frac{n}{m} \le \frac{p}{q} + \frac{p}{m}$, which holds by (A.4.3), we get

$$\begin{aligned} h(y) \le &\left(\frac{n}{t} + \frac{s+q(s+1)}{m}\right)h^2 + \frac{p(s+1)+pq(s+1)}{m}k^2 + \frac{p}{q}h(x) + \frac{n}{t}\log(\sqrt{8}p) \\ &+ \frac{nq}{2m} + \frac{s}{m}\log(4pq(n+sr)) + \frac{q(s+1)}{m}\log(2pq) \\ &+ \frac{1}{m}\log(sp+n+1)(sq+q+1). \end{aligned} \tag{A.4.6}$$

We now continue by bounding each term in the right-hand side of inequality (A.4.6). The elementary inequalities $\frac{1}{2}pq^2\frac{k}{h} \le m \le pq^2\frac{k}{h}$ and $m \ge pq^2$, which hold because of our assumption $k \ge h$, will be used throughout the process.

Our definition in (A.4.3) and $t = pq$ imply

$$\frac{n}{t} \le \frac{m}{q^2} + \frac{1}{q}. \tag{A.4.7}$$

Furthermore, we use $s \le 2pq - 1$ to obtain a bound for the first term on the right of (A.4.6)

$$\left(\frac{n}{t} + \frac{s+q(s+1)}{m}\right)h^2 \le \left(\frac{m}{q^2} + \frac{1}{q} + \frac{4pq^2}{m}\right)h^2 \le (p\frac{k}{h} + \frac{1}{q} + 4)h^2 \le 6phk. \tag{A.4.8}$$

The second term in (A.4.6) can bounded as follows

$$\frac{p(s+1)+pq(s+1)}{m}k^2 \le \frac{4p^2q^2}{m}k^2 \le 8phk. \tag{A.4.9}$$

We skip the third term, which is the main contribution to the height of y, and bound the fourth with the help of (A.4.7):

$$\frac{n}{t}\log(\sqrt{8}p) \le (\frac{m}{q^2} + \frac{1}{q})\log(\sqrt{8}p) \le (p\frac{k}{h} + \frac{1}{q})\log(\sqrt{8}p) \le 2p\frac{k}{h}\log(\sqrt{8}p). \tag{A.4.10}$$

We use $\frac{n}{m} \le \frac{p}{q} + \frac{p}{m}$ again and also $m \ge q$ to bound the fifth term:

$$\frac{nq}{2m} \le \frac{q}{2}(\frac{p}{q} + \frac{p}{m}) = \frac{p}{2} + \frac{pq}{2m} \le p. \tag{A.4.11}$$

For the sixth term we use $n + sr \le n + 2pqr \le 3p^2q^2n$ to get

$$\begin{aligned} \frac{s}{m}\log(4pq(n+sr)) &\le 2\frac{pq}{m}\log(12p^3q^3n) = 2\frac{pq}{m}\log(12p^3q^3) + 2pq\frac{\log n}{m} \\ &\le 8\frac{\log(2pq)}{q} + 2pq\frac{\log n}{m}. \end{aligned} \tag{A.4.12}$$

For positive α and β we have $\log(\alpha+\beta) \le \frac{\alpha}{\beta} + \log\beta$. So $\log n \le \log(m\frac{p}{q}+p) \le \frac{m}{q} + \log p$ with $\alpha = \frac{mp}{q}$ and $\beta = p$. Therefore $pq\frac{\log n}{m} \le p + \log p$. We apply this inequality and $\frac{\log q}{q} \le \frac{1}{e}$, here $e = 2.71828\ldots$, to (A.4.12) and get

$$\frac{s}{m}\log(4pq(n+sr)) \le 8\log(2p) + \frac{8}{e} + 2p + 2\log p. \tag{A.4.13}$$

The second to last term in (A.4.6) can be bounded as

$$\frac{q(s+1)}{m}\log(2pq) \le \frac{2pq^2}{m}\log(2pq) \le 2\log(2pq). \tag{A.4.14}$$

Finally, to tackle the last term we use $n \le 2pm$ and $p \le m$ which follow from (A.4.3), so $\frac{\log n}{m} \le \frac{\log(2pm)}{m} \le \frac{\log(2p)}{p} + \frac{\log m}{m} \le \frac{2}{e} + \frac{1}{e}$ by elementary calculus. Hence with $s \le 2pq$

$$\begin{aligned} \frac{1}{m}\log(sp+n+1)(sq+q+1) &\le \frac{1}{m}\log(2p^2q+n+1)(2pq^2+q+1) \\ &\le \frac{1}{m}\log(16p^3q^3n) \le \log 16 + 3\frac{\log p}{p} + 3\frac{\log q}{q^2} + \frac{\log n}{m} \\ &\le \log 16 + \frac{3}{e} + \frac{3}{2e} + \frac{3}{e} < \frac{28}{5}. \end{aligned} \tag{A.4.15}$$

In the second to last inequality we used $\frac{\log q}{q^2} \le \frac{1}{2e}$.

The sum of the right-hand sides in (A.4.11), (A.4.13), (A.4.14), and (A.4.15) can be bounded with elementary calculus and the fact that $p \in \mathbf{N}$:

$$p + 8\log(2p) + \frac{8}{e} + 2p + 2\log p + 2\log(2pq) + \frac{28}{5} \le 13p + 2\log q + \frac{28}{5}.$$

It is easy to show $13p + 2\log q + 28/5 \le 17p\log(3q)$ by considering the two cases $p = 1$ and $p \ge 2$ separately.

We use this estimate together with (A.4.8), (A.4.9), and (A.4.10) to bound (A.4.6) from above as follows:

$$h(y) \le \frac{p}{q}h(x) + 14phk + 2p\frac{k}{h}\log(\sqrt{8}p) + 17p\log(3q).$$

If we abbreviate $h' = \max\{(\log(\sqrt{8}p))^2, h_p(P)\}^{1/2} \geq h$, then $\frac{k}{h}\log(\sqrt{8}p) \leq h'k$ and so

$$h(y) \leq \frac{p}{q}h(x) + 16ph'k + 17p\log(3q)$$

as required. □

PROPOSITION A.1. *Let $P \in \overline{\mathbf{Q}}[X,Y]$ be irreducible with $p = \deg_X P > 0$ and $q = \deg_Y P > 0$. If $(x,y) \in \overline{\mathbf{Q}}^2$ with $P(x,y) = 0$, then*

$$h(y) \leq \frac{p}{q}h(x) + 28p\max\{p, h_p(P)\}^{1/2}\max\{1, h_p(P), h(x)\}^{1/2} + 18p\log(3q).$$

PROOF. If $h(x) < h_p(P) + 2p\log 2$, then Lemma A.1(iii) applied to P gives

$$\begin{aligned} h(y) &\leq ph(x) + \log(2pq) + h_p(P) \\ &\leq p(h_p(P) + 2p\log 2)^{1/2}h(x)^{1/2} + \log(2pq) + h_p(P). \end{aligned}$$

This inequality is clearly stronger than the assertion.

So we will assume

$$h(x) \geq h_p(P) + 2p\log 2. \tag{A.4.16}$$

Now if both partial derivatives of P vanish at (x,y), then Lemma A.8 and $\log p \leq p^{1/2}$ give

$$h(y) \leq 2ph_p(P) + 5p\log(2pq) \leq 2ph_p(P) + 5p^{3/2} + 5p\log(2q),$$

which is also stronger than our assertion. From now on we assume that (x,y) is a regular zero of P.

There exist $\xi \in \overline{\mathbf{Q}}$ and $\tilde{P}$ as in Lemma A.7. If $x = \xi$ or $x = 0$, then $h(x) = 0$, but this contradicts (A.4.16). Hence $x \neq 0, \xi$ and $((x-\xi)^{-1}, y)$ is a zero of $\tilde{P}$. A short calculation shows that this zero is regular. The height inequalities

$$h((x-\xi)^{-1}) = h(x-\xi) \geq h(x) - h(\xi) - \log 2 = h(x) - \log 2,$$

$$h_p(\tilde{P}) \leq h_p(P) + p\log 2, \tag{A.4.17}$$

and (A.4.16) imply $h((x-\xi)^{-1}) \geq h_p(\tilde{P})$. We may thus apply Lemma A.9 to the point $((x-\xi)^{-1}, y)$. We use $h((x-\xi)^{-1}) \leq h(x) + \log 2$ and again (A.4.17) to show

$$\begin{aligned} h(y) \leq& \frac{p}{q}h(x) + \frac{p}{q}\log 2 \\ &+ 16p\max\{(\log(\sqrt{8}p))^2, h_p(P) + p\log 2\}^{1/2}\max\{1, h(x) + \log 2\}^{1/2} \\ &+ 17p\log(3q). \end{aligned} \tag{A.4.18}$$

We note $(\log(\sqrt{8}p))^2 \leq 1.6p \leq (1+\log 2)p$ since $p \geq 1$ and so

$$\max\{(\log(\sqrt{8}p))^2, h_p(P) + p\log 2\} \leq (1+\log 2)\max\{p, h_p(P)\}.$$

Furthermore, we have $\max\{1, h(x) + \log 2\} \leq (1+\log 2)\max\{1, h(x)\}$. The proposition follows from these last two inequalities, (A.4.18), and $\frac{p}{q}\log 2 \leq p\log(3q)$.

□

Proof of Theorem A.1: For brevity say $r = \max\{p, q\}$. If $\max\{h(x), h(y)\} < h_p(P)$, then clearly

$$|ph(x) - qh(y)| \leq (p+q)\max\{h(x), h(y)\} \leq 2pqh_p(P)^{1/2}\max\{h(x), h(y)\}^{1/2}$$

and the theorem follows in this case.

So let us assume $\max\{h(x), h(y)\} \geq h_p(P)$. We note that by symmetry Proposition A.1 implies

$$|ph(x) - qh(y)| \leq 28pq\max\{r, h_p(P)\}^{1/2}\max\{1, h(x), h(y)\}^{1/2} + 18pq\log(3r).$$

The Theorem clearly follows from this inequality and $18\log(3r) \leq 23r^{1/2}$. □

Corollary A.1 is easy to prove with the Quasi-equivalence Theorem.

Say $P(x, y) = 0$ with $x, y \in \overline{\mathbf{Q}}$, $h(y) > 2\frac{p}{q}h(x)$, and $h(x) \geq 1$. If $h(x) \leq h(y)$, then by Theorem A.1

$$\frac{1}{2}h(y) < h(y) - \frac{p}{q}h(x) \leq 51p\max\{p, q, h_p(P)\}^{1/2}h(y)^{1/2}.$$

Hence the height of y is less then the right of (A.1.3) and therefore so is the height of x. On the other hand if $h(y) \leq h(x)$ we get

$$\frac{p}{q}h(x) < h(y) - \frac{p}{q}h(x) \leq 51p\max\{p, q, h_p(P)\}^{1/2}h(x)^{1/2}.$$

The resulting bound for $h(x)$ is less than the right side of (A.1.3). □

APPENDIX B

Applications of the Quasi-equivalence Theorem

In this second appendix we show how to use the Quasi-equivalence Theorem from the previous appendix to deduce explicit versions of four number theoretic results: Theorems of Bombieri, Masser, and Zannier, of Runge, of Skolem, and of Sprindzhuk. The goal is not to get optimal dependency in the various quantities, but rather to show how Theorem A.1 is connected to these four theorems. All four theorems mentioned are proved using the same auxiliary function constructed in section 2 of appendix A.

1. On a Theorem of Bombieri, Masser, and Zannier

In this first section we give a completely explicit proof of Theorem 1 of [**BMZ99**] in the two dimensional case. The proof uses Theorem A.1 and closely mimics Bombieri, Masser, and Zannier's "alternative proof" given in §3 of [**BMZ99**]. In fact the occurrence of a quasi-equivalence of heights statement in [**BMZ99**] was the main drive behind proving Theorem A.1.

Bombieri, Masser, and Zannier's Theorem has already been discussed in chapter 6.

THEOREM B.1. *Let $P \in \overline{\mathbf{Q}}[X,Y]$ be irreducible with $p = \deg_X P$, $q = \deg_Y P$, and not of the form $\alpha X^p Y^q - \beta$ or $\alpha X^p - \beta Y^q$. If $P(x,y) = 0$ with $x, y \in \overline{\mathbf{Q}}^*$ multiplicatively dependent, then*

$$\max\{h(x), h(y)\} \leq 3.10^5 (\deg P)^3 \max\{pq, h_p(P)\}.$$

We recall that x and y are multiplicatively dependent if and only if (x,y) is contained in a proper algebraic subgroup of $\mathbf{G}_m^n$.

We need a preparatory lemma.

LEMMA B.1. *Let $P \in \overline{\mathbf{Q}}[X,Y]$ be irreducible with $p = \deg_X P > 0$ and $q = \deg_Y P > 0$. If $b \in \mathbf{N}$, there exists $Q \in \overline{\mathbf{Q}}[X,Y]$ irreducible with*

$$1 \leq \deg_X Q \leq bp,\ 1 \leq \deg_Y Q \leq q, \text{ and } h_p(Q) \leq b h_p(P) + 2b(p+q)\log 2 \tag{B.1.1}$$

such that P divides $Q(X, Y^b)$.

Before we prove Lemma B.1 we recall two height inequalities for polynomials: let $f_1, \ldots, f_n \in \overline{\mathbf{Q}}[X,Y]$ be non-zero with $f = f_1 \cdots f_n$ and $d = \deg_X f + \deg_Y f$, then

$$-d\log 2 + \sum_{i=1}^{n} h_p(f_i) \leq h_p(f) \leq d\log 2 + \sum_{i=1}^{n} h_p(f_i). \tag{B.1.2}$$

A proof not restricted to the two variable case is given in [**BG06**] (Theorem 1.6.13 page 28).

Proof of Lemma B.1: Let ζ be a primitive bth root of unity and set $A = \prod_{i=1}^{b} P(X, \zeta^i Y)$. Since $A(X,Y) = A(X,\zeta Y)$ we have $A(X,Y) = B(X,Y^b)$ for some $B \in \overline{\mathbf{Q}}[X,Y]$. Clearly $\deg_X B = bp$, $\deg_Y B = q$, and $h_p(B) = h_p(A)$. With (B.1.2) and $h_p(P) = h_p(P(X,\zeta^i Y))$ we may bound

$$h_p(B) \leq b h_p(P) + b(p+q)\log 2. \tag{B.1.3}$$

Also P divides $B(X,Y^b)$ by construction. But our B need not be irreducible, nevertheless B has an irreducible factor Q such that P divides $Q(X,Y^b)$. The degree upper bounds in (B.1.1) hold and so do the lower bounds since $pq > 0$. Furthermore, the height bound in (B.1.1) follows from (B.1.2) and (B.1.3). □

Proof of Theorem B.1: Let P be as in the hypothesis, we have $p, q > 0$. By symmetry we may assume $h(x) \geq h(y)$ and without loss of generality also $h(x) \geq 1$. There exist integers r, s not both zero with $x^r y^s = 1$. We will often use the fact that $|r|h(x) = |s|h(y)$ which follows from chapter 1. If $s = 0$, then $r \neq 0$ and so $h(x) = 0$, contradicting our assumption. Hence $s \neq 0$ and we may find $t \in \overline{\mathbf{Q}}^*$ with $x = t^s$, $y = \zeta t^{-r}$ for some root of unity ζ. We define $\theta = r/s$, then $|\theta| = h(y)/h(x) \leq 1$.

Let $B > 1$, by [**Cas57**] page 1, there exist integers a, b with $1 \leq b < B$ and $|\theta b - a| \leq B^{-1}$. We take $B = 2q$ and define $\gamma = x^a y^b = \zeta^b t^{as-br}$. Therefore

$$h(\gamma) = |as - br|h(t) = |\theta b - a|h(x) \leq \frac{h(x)}{B} < h(x). \tag{B.1.4}$$

Let Q be the polynomial from Lemma B.1. We set $R = Q(X, X^{-a}Y)X^w$ where w is an integer such that R is a polynomial not divisible by X. We have $R(x,\gamma) = Q(x,y^b) = 0$ since P divides $Q(X,Y^b)$.

We continue by comparing $h(x)$ and $h(\gamma)$ with Theorem A.1. It is easy to see that R is irreducible. We define $m = \deg_X R$ and $n = \deg_Y R$. Bounding n is not difficult: as $n = \deg_Y Q$ we obtain

$$1 \leq n \leq q.$$

Since $|\theta b - a| \leq B^{-1}$ we have $|a| \leq B^{-1} + |\theta| b \leq 1 + b < 1 + B = 1 + 2q$, so $|a| \leq 2q$. With the bound for $\deg_X Q$ and $\deg_Y Q$ in Lemma B.1 we have

$$m \leq 2\max\{1, |a|\}\deg Q \leq 4q(bp+q) \leq 4q^2(2p+1) \leq 12pq^2.$$

Finally we show $m \geq 1$. Indeed assume $m = 0$, then R has degree 1 in Y by irreducibility. Therefore P divides a polynomial of the form $\alpha X^u Y^v - \beta$ or $\alpha X^u - \beta Y^v$. But then P is also of this form, contradicting the hypothesis. Since R and Q have the same coefficients we get

$$h_p(R) = h_p(Q) \leq b h_p(P) + 2b(p+q)\log 2 \leq 2q h_p(P) + 4q(p+q)\log 2 \tag{B.1.5}$$

by Lemma B.1.

Having bounded height and partial degrees of R and since $R(x,\gamma)=0$ we apply Theorem A.1 to deduce

$$\begin{aligned} h(x) &\le \frac{n}{m}h(\gamma) + 51n\max\{m,n,h_p(R)\}^{1/2}h(x)^{1/2} \\ &\le q\frac{h(x)}{B} + 51q\max\{12pq^2, 2qh_p(P)+4q(p+q)\log 2\}^{1/2}h(x)^{1/2}, \end{aligned}$$

in the second inequality we used (B.1.4) and (B.1.5). Since $\frac{q}{B}=\frac{1}{2}$ we conclude

$$\frac{1}{2}h(x)^{1/2} \le 51\sqrt{12}q^{3/2}\max\{pq, h_p(P)+pq\}^{1/2} \le 250q^{3/2}\max\{pq,h_p(P)\}^{1/2}.$$

This inequality completes the proof. □

2. On a Theorem of Runge

In [**Run87**] Runge proved that $P(x,y)=0$ admits only finitely many solutions $x,y\in\mathbf{Z}$ if $P\in\mathbf{Q}[X,Y]$ is irreducible and satisfies the following condition: we have $\deg_X P=\deg_Y P=\deg P$ and the homogeneous part of P with maximal degree is not a scalar times the power of an irreducible polynomial in $\mathbf{Q}[X,Y]$. In fact Runge proved a finiteness result under a weaker hypothesis on P. Runge's method is effective and several explicit bounds for $\max\{|x|,|y|\}$ have been determined: for example Hilliker and Straus in [**HS83**] or Walsh in [**Wal92**]. Many approaches use Eisenstein's Theorem on bounding the coefficients in power series of algebraic functions.

In this section we will prove a simple, explicit version of Runge's Theorem using the Quasi-equivalence Theorem proved in appendix A. We state the main result of the section:

THEOREM B.2. *Let $P\in\mathbf{Z}[X,Y]$ be irreducible in $\overline{\mathbf{Q}}[X,Y]$ and assume $d=\deg_X P=\deg_Y P=\deg P$. Furthermore, assume that the homogeneous part of P with degree d is not a scalar times the power of an irreducible polynomial in $\mathbf{Q}[X,Y]$. If $P(x,y)=0$ with $x,y\in\mathbf{Z}$, then*

$$\log\max\{1,|x|,|y|\} \le 10^5 d^5\max\{d,h_p(P)\}. \tag{B.2.1}$$

A more sophisticated version of Runge's Theorem, similar to the one of Bombieri ([**Bom83**] page 304), can also be proved using our Quasi-equivalence Theorem.

For the rest of this section let $P=\sum_{i,j}p_{ij}X^iY^j\in\mathbf{Z}[X,Y]$ be irreducible in $\overline{\mathbf{Q}}[X,Y]$ with $d=\deg P=\deg_X P=\deg_Y P$. Furthermore, since Theorem B.2 applies only to polynomials of degree at least 2, we assume $d\ge 2$. We use $P_d=\sum_{i+j=d}p_{ij}X^iY^j$ to denote the homogeneous part of degree d. Finally we can factor P_d as $p_{d0}\prod_{s=1}^{d}(X-t_sY)$ with $t_s\in\overline{\mathbf{Q}}^*$.

For any s, the function $X-t_sY$ on the curve defined by P has degree strictly less than d, the degree of the function X:

LEMMA B.2. *Say $t=t_s$. We have $h(t)\le h_p(P)+\log d$. Furthermore, if $R_t(X,Z)=P(X,t^{-1}(X-Z))\in\overline{\mathbf{Q}}[X,Z]$, then $\deg_Z R=d$, $\deg_X R\le d-1$, and $h_p(R_t)\le d(2h_p(P)+\log(2d))$.*

PROOF. The bound for $h(t)$ follows from Lemma A.1(iii); indeed t is a zero of $P_d(X,1)$.

Let $i,j \in \mathbf{Z}$, the coefficients of X^iZ^j in R_t is

$$(-1)^j \sum_{k=0}^{d-j} \binom{j+k}{k} p_{i-k,j+k} t^{-j-k}.$$

The coefficient of Z^d is nonzero and that of X^iZ^j is zero provided $i \geq d$. The lemma now follows from elementary inequalities as in the proof of Lemma A.7. □

By the previous lemma and the Quasi-equivalence Theorem, the height $h(x-ty)$ is small compared to $h(x)$. We can say even more:

LEMMA B.3. *Let $P(x,y) = 0$ with $x,y \in \mathbf{Z}$ and*

$$h(x) \geq 2.10^4 d^2 \max\{d, h_p(P)\}, \tag{B.2.2}$$

then $h(y) \leq 2h(x)$. Furthermore, if $t = t_s$ for some $1 \leq s \leq d$, then

$$\log \max\{1, |x - \sigma(t)y|\} \leq \frac{d-1}{d} \log|x| + 177 d^{3/2} \max\{d, h_p(P)\}^{1/2} \sqrt{\log|x|}$$

for some embedding $\sigma : \mathbf{Q}(t) \to \mathbf{C}$.

PROOF. The inequality $h(y) \leq 2h(x)$ follows from Corollary A.1 in appendix A.

Let $R = R_t$ be as in Lemma B.2, then R is irreducible. Now $\deg_X R \geq 1$, indeed if $\deg_X R = 0$, then by construction P has degree 1, which contradicts our hypothesis $d \geq 2$. Let $z = x - ty$, so $R(x,z) = 0$. Theorem A.1, $\deg_X R \leq d-1$, and $\deg_Z R = d$ imply

$$h(x-ty) = h(z) \leq \frac{d-1}{d} h(x) + 51 d \max\{d, h_p(R)\}^{1/2} \max\{1, h(x), h(z)\}^{1/2}. \tag{B.2.3}$$

We have the elementary bound $h(z) \leq \log 2 + h(x) + h(y) + h(t)$ which yields $h(z) \leq \log(2d) + h_p(P) + h(x) + h(y) \leq \log(2d) + h_p(P) + 3h(x)$ by Lemma B.2 and $h(y) \leq 2h(x)$. Therefore $h(z) \leq 4h(x)$, in view of the lower bound (B.2.2). By Lemma B.2 we also conclude $\max\{d, h_p(R)\} \leq 3d \max\{d, h_p(P)\}$. With (B.2.3) we obtain

$$h(x-ty) \leq \frac{d-1}{d} h(x) + 177 d^{3/2} \max\{d, h_p(P)\}^{1/2} h(x)^{1/2}.$$

By the definition of the height (cf. chapter 1) there exists an embedding $\sigma : \mathbf{Q}(t) \to \mathbf{C}$ with $\log \max\{1, |x - \sigma(t)y|\} \leq h(x-ty)$. The lemma follows from this statement and $h(x) = \log|x|$. □

We now prove Theorem B.2:

Let P be as in the hypothesis. By assumption there are non-conjugated $t', t'' \in \overline{\mathbf{Q}}^*$ that are zeros of $P_d(X,1)$. Say x and y are as in the hypothesis and $|x| \geq |y|$. We may assume that x satisfies (B.2.2).

Let σ', σ'' be the embeddings given by Lemma B.3 applied to t', t'' respectively, then $\sigma'(t') \neq \sigma''(t'')$. For brevity we define $\xi = x - \sigma'(t')y$ and $\eta = x - \sigma''(t'')y$; we eliminate y to get

$$x = \frac{\xi\sigma''(t'') - \eta\sigma'(t')}{\sigma''(t'') - \sigma'(t')}.$$

So

$$|x| \leq 2\max\{|\xi|, |\eta|\} \max\{|\sigma'(t')|, |\sigma''(t'')|\}|\sigma'(t') - \sigma''(t'')|^{-1} \tag{B.2.4}$$

for the complex absolute value $|\cdot|$.

We will bound $|\xi|$ and $|\eta|$ using Lemma B.3, but first we bound the remaining absolute values in (B.2.4) using height inequalities: if $\alpha \in \overline{\mathbf{Q}}$, then $\log\max\{1, |\alpha|_v\} \leq [\mathbf{Q}(\alpha) : \mathbf{Q}]h(\alpha)$ for any $v \in M_{\mathbf{Q}(\alpha)}$. This inequality follows immediately from the definition of the height. Since for example $[\mathbf{Q}(t') : \mathbf{Q}] \leq d$ and $[\mathbf{Q}(t', t'') : \mathbf{Q}] \leq d^2$ we get a bound for (B.2.4):

$$\log|x| \leq \log 2 + d\max\{h(t'), h(t'')\} + d^2 h(\sigma'(t') - \sigma''(t'')) + \log\max\{1, |\xi|, |\eta|\}. \tag{B.2.5}$$

Let $h = \max\{d, h_p(P)\}$. Lemma B.2 implies the upper bounds

$$\max\{h(t'), h(t'')\} \leq h_p(P) + \log d.$$

Next we apply Lemma B.3 to bound $|\xi|$ and $|\eta|$. Together with (B.2.5) and standard height inequalities we have

$$\begin{aligned}\log|x| \leq& \log 2 + d(\log d + h_p(P)) + d^2(\log 2 + 2\log d + 2h_p(P)) + \frac{d-1}{d}\log|x| \\ &+ 177d^{3/2}h^{1/2}\sqrt{\log|x|} \\ \leq& \frac{d-1}{d}\log|x| + 8d^2h + 177d^{3/2}h^{1/2}\sqrt{\log|x|}.\end{aligned}$$

Hence

$$\log|x| \leq 8d^3h + 177d^{5/2}h^{1/2}\sqrt{\log|x|}.$$

If $B, C, T \geq 0$ with $T \leq B\sqrt{T} + C$, then

$$T \leq (1+\sqrt{2})^2\max\{(B/2)^2, C\} \leq (1+\sqrt{2})^2((B/2)^2 + C). \tag{B.2.6}$$

We use (B.2.6) with $B = 177d^{5/2}h^{1/2}$, $C = 8d^3h$, and $T = \log|x|$ to conclude that $\log|x|$ is at most half the right side of (B.2.1). The upper bound for $\log|y|$ follows from Lemma B.3 which implies $\log\max\{1, |y|\} \leq 2\log|x|$.

3. On a Theorem of Skolem

In 1929 Skolem showed that if $P \in \mathbf{Z}[X, Y]$ is irreducible in $\mathbf{Q}[X, Y]$ and $P(0, 0) = 0$, then there are only finitely many coprime $x, y \in \mathbf{Z}$ with $P(x, y) = 0$. Of course Siegel's Theorem on the finiteness of the integral points on a curve of genus at least 1 supersedes this result in many cases. Contrary to Siegel's Theorem, which remains ineffective, Skolem's Theorem is effective. For example in [**Wal92**] Walsh proved an effective and explicit version of Skolem's Theorem. In [**Pou04**] Poulakis improved on Walsh's bound

under some additional restrictions to P. In [**Abo06**] Abouzaid extended Skolem's Theorem in two ways: first, he considered polynomials P with algebraic coefficients. Second, he generalized the definition of the greatest common divisor, or gcd for short, from pairs of integers to pairs of algebraic numbers. Say $P \in \overline{\mathbf{Q}}[X,Y]$ is irreducible with $\deg_Y P > 0$ such that $(0,0)$ is a non-singular point on the curve defined by P. If $P(x,y) = 0$ with x, y algebraic and not both zero, then Abouzaid proved that $\log \gcd(x,y)$ is asymptotically equal to $\frac{1}{\deg_Y P} h(x)$. He also explicitly bounded the difference of these two values. In Theorem B.3 we give an independent proof of Abouzaid's result based on the work in appendix A on quasi-equivalence of heights.

Let $x, y \in K$ where K is a number field and assume that x and y are not both zero. We follow Abouzaid and define

$$\operatorname{lgcd}(x,y) = \frac{1}{[K:\mathbf{Q}]} \sum_{v \in M_K} [K_v : \mathbf{Q}_v] \log \min\{\max\{1, |x|_v^{-1}\}, \max\{1, |y|_v^{-1}\}\}.$$

Here $\operatorname{lgcd}(x,y) = h(y)$ if $x = 0$ and $y \neq 0$. Just like the height, $\operatorname{lgcd}(x,y)$ does not depend on the field K containing x and y. For $a, b \geq 0$ we have the identity

$$\max\{\min\{1,a\}, \min\{1,b\}\} = \frac{\max\{a,b\}}{\max\{1,a,b\}},$$

and so we may rewrite $\operatorname{lgcd}(x,y)$ as

$$\operatorname{lgcd}(x,y) = h(x,y) - h([x:y]) \geq 0. \tag{B.3.1}$$

Here $(x,y) \in \overline{\mathbf{Q}}^2$ and $[x:y] \in \mathbf{P}^1(\overline{\mathbf{Q}})$. We recall our notion of height defined in chapter 1 corresponding to these two cases. Finally it is readily checked that $\operatorname{lgcd}(x,y) = \log \gcd(x,y)$ if x and y are integers.

For any non-zero polynomial $P(X,Y)$ in two variables we define

$$\mathrm{e}(P) = \max\{f \in \mathbf{Z};\ X^f | P(X, XY)\} \in [0, \deg P].$$

If $P = \sum_{i,j} p_{ij} X^i Y^j$, then $\mathrm{e}(P) = \min\{i + j;\ p_{ij} \neq 0\}$. Clearly $\mathrm{e}(P)$ is positive if and only if $P(0,0) = 0$. Furthermore, $\mathrm{e}(P) = 1$ precisely when $\frac{\partial P}{\partial X}(0,0) \neq 0$ or $\frac{\partial P}{\partial Y}(0,0) \neq 0$. That is when $(0,0)$ is a non-singular point on the curve defined by P.

We will use Theorem A.1 to give an independent proof of Abouzaid's Theorem with slightly improved dependency on the degree of P and without the non-singularity hypothesis. Abouzaid proved his theorem without referring to a quasi-equivalence of heights result, in fact he uses it to prove variant of our Theorem A.1.

THEOREM B.3. *Let $P \in \overline{\mathbf{Q}}[X,Y]$ be irreducible with $p = \deg_X P > 0$, $q = \deg_Y P > 0$, and $d = \deg P$. If $P(x,y) = 0$ where x and y are non-zero algebraic numbers, then*

$$\begin{aligned} \max\{ \left| \operatorname{lgcd}(x,y) - \frac{\mathrm{e}(P)}{q} h(x) \right|, \left| \operatorname{lgcd}(x,y) - \frac{\mathrm{e}(P)}{p} h(y) \right| \} \\ \leq 183 d \max\{d, h_p(P)\}^{1/2} \max\{1, h(x), h(y)\}^{1/2}. \end{aligned} \tag{B.3.2}$$

Deriving a slightly weaker version of Theorem A.1 from inequality (B.3.2) is simple if $P(0,0) = 0$. In this case the difference $|\frac{h(x)}{q} - \frac{h(y)}{p}|$ is at most twice the right-hand side of (B.3.2). The case $P(0,0) \neq 0$ can be handled by replacing $P(X,Y)$ with $P(X+\xi, Y+\eta)$ where $P(\xi,\eta) = 0$

As a corollary we get a bound for the heights of x and y in terms of $\operatorname{lgcd}(x,y)$ as soon as P has no constant term.

COROLLARY B.1. *Let $P \in \overline{\mathbf{Q}}[X,Y]$ and d be as in Theorem B.3. We assume $\mathrm{e}(P) \geq 1$, that is $P(0,0) = 0$. If $P(x,y) = 0$ where x and y are non-zero algebraic numbers, then*

$$\max\{h(x), h(y)\} \leq 6\frac{d}{\mathrm{e}(P)}\operatorname{lgcd}(x,y) + 5.10^4\frac{d^4}{\mathrm{e}(P)^2}\max\{d, h_p(P)\}. \tag{B.3.3}$$

In particular $h(x)$ and $h(y)$ are bounded if $\operatorname{lgcd}(x,y) = 0$.

We prove Theorem B.3 in a series of elementary lemmas and then apply Theorem A.1.

Throughout the remainder of this section $P \in \overline{\mathbf{Q}}[X,Y]$ will be irreducible with $p = \deg_X P > 0$, $q = \deg_Y P > 0$, and $d = \deg P$.

LEMMA B.4. *There exists a root of unity λ such that $Q(X,Z) = P(X, Z - \lambda X)$ is irreducible and satisfies $\deg Q = \deg_X Q = d$, $\deg_Z Q = q$, and $h_p(Q) \leq h_p(P) + q\log 2$.*

PROOF. The proof is very similar to the proof of Lemma A.7, so we omit it. □

LEMMA B.5. *Let $x, y \in \overline{\mathbf{Q}}$ with $P(x,y) = 0$, then*

$$|qh(x,y) - dh(x)| \leq 110dq\max\{d, h_p(P)\}^{1/2}\max\{1, h(x), h(y)\}^{1/2}.$$

PROOF. Let λ and Q be as in Lemma B.4. For brevity we define $h = \max\{d, h_p(P)\}$ and $k = \max\{1, h(x), h(y)\}$. We set $z = \lambda x + y$ and note $Q(x,z) = 0$. Let K be a number field containing the coefficients of Q and the algebraic numbers λ, x, and y. Let $v \in M_K$, the inequality

$$\delta_v(2)^{-1} \leq \frac{\max\{1, |x|_v, |y|_v\}}{\max\{1, |x|_v, |z|_v\}} \leq \delta_v(2)$$

implies

$$|h(x,y) - h(x,z)| \leq \log 2. \tag{B.3.4}$$

We continue by comparing $h(x,z)$ with $\frac{d}{q}h(z)$. Say $Q = \sum_{i,j} q_{ij}X^iY^j$, then $q_{d0} \neq 0$. We claim

$$|x|_v \leq \delta_v(3d^2)\frac{|Q|_v}{|q_{d0}|_v}\max\{1, |z|_v\}. \tag{B.3.5}$$

Indeed, let us assume (B.3.5) is false. In particular $|x|_v > 1$. Let $i + j \leq d$ with i, j non-negative integers such that $(i,j) \neq (d,0)$ and $q_{ij} \neq 0$. If $j = 0$, then $d \geq i + 1$ and so $|q_{d0}x^d|_v \geq |q_{d0}x^i|_v|x|_v > \delta_v(3d^2)|Q|_v|x^i|_v$ by the negation of (B.3.5). If $j > 0$, then

$|q_{d0}x^d|_v \geq |q_{d0}|_v|x|_v^i|x|_v^j > \delta_v(3d^2)|Q|_v|x|_v^i|z|_v^j$, again using the negation of (B.3.5). In any case we have

$$|q_{d0}x^d|_v > \delta_v(3d^2)|q_{ij}x^iz^j|_v \quad \text{if} \quad (i,j) \neq (d,0).$$

But since $Q(x,z) = 0$ and because Q has no more than $\frac{1}{2}(d+1)(d+2) \leq 3d^2$ non-zero coefficients we have a contradiction.

Therefore B.3.5 holds for all $v \in M_K$, and even with $|x|_v$ replaced by the possibly larger $\max\{1, |x|_v, |z|_v\}$. Passing over to heights we have

$$0 \leq h(x,z) - h(z) \leq \log(3d^2) + h_p(Q),$$

here the lower bounds is obvious. We combine this inequality with (B.3.4) and the bound for $h_p(Q)$ from Lemma B.4 to obtain

$$|h(x,y) - h(z)| \leq \log(6d^2 2^q) + h_p(P) \leq \log(6d^2 2^d) + h_p(P) \leq 4h. \tag{B.3.6}$$

The last inequality used $\log(6d^2 2^d) \leq 3d$.

We have $\max\{d, h_p(Q)\} \leq (1+\log 2)h$ since $h_p(Q) \leq h_p(P) + q\log 2$. Furthermore, since $h(z) \leq h(x)+h(y)+\log 2$ by (1.1.6) in chapter 1 we conclude $\max\{1, h(x), h(z)\} \leq (2+\log 2)k$. By Theorem A.1 applied to Q we get $|dh(x) - qh(z)| \leq 109dqh^{1/2}k^{1/2}$. This inequality and (B.3.6) together imply

$$|qh(x,y) - dh(x)| \leq 109dq(hk)^{1/2} + 4qh.$$

The lemma now follows easily if $k \geq 16h$. If $k < 16h$, then

$$|qh(x,y) - dh(x)| \leq 3dk \leq 12d(hk)^{1/2}$$

completes the proof. □

LEMMA B.6. *Let $P(x,y) = 0$ with $x, y \in \overline{\mathbf{Q}}$ not both zero, then*

$$|(d - \mathrm{e}(P))h(x) - qh([x:y])| \leq 73dq \max\{d, h_p(P)\}^{1/2} \max\{1, h(x), h(y)\}^{1/2}.$$

PROOF. We may assume $x \neq 0$ and in this case we have $h([x:y]) = h(y/x)$ by the product formula. Let $e = \mathrm{e}(P) \leq d$. If $e = d$, then by irreducibility P must be homogeneous of degree 1. The lemma holds in this case as then $h_p(P) = h([x:y]) \leq h(x) + h(y)$. So let us assume $e < d$.

We define $F(U,V) = U^{-e}P(U,UV) \in \mathbf{Q}[U,V]$, then F is irreducible with $\deg_U F = d-e$, $\deg_V F = q$, and $h_p(F) = h_p(P)$. We have $F(x, y/x) = 0$ and so by Theorem A.1:

$$|(d-e)h(x) - qh(y/x)| \leq 51dq \max\{d, h_p(P)\}^{1/2} \max\{1, h(x), h(y/x)\}^{1/2}.$$

The lemma now follows from $h(y/x) \leq h(x) + h(y)$. □

Proof of Theorem B.3: Let $P(x,y) = 0$ with $x, y \in \overline{\mathbf{Q}}$ not both zero. As an immediate consequence of Lemmas B.5, B.6, and (B.3.1) we get

$$|\mathrm{lgcd}(x,y) - \frac{\mathrm{e}(P)}{q}h(x)| \leq 183d \max\{d, h_p(P)\}^{1/2} \max\{1, h(x), h(y)\}^{1/2}.$$

This inequality bounds the first quantity on the left in (B.3.2). The bound for the second quantity follows by considering $Q(X,Y) = P(Y,X)$; indeed we have $\mathrm{e}(P) = \mathrm{e}(Q)$. □

Proof of Corollary B.1: By symmetry we may assume $h(x) \geq h(y)$ and by (B.3.3) also $h(x) \geq 1$. Theorem B.3 and $q \leq d$ imply

$$h(x) \leq \frac{d}{\mathrm{e}(P)}\mathrm{lgcd}(x,y) + 183\frac{d^2}{\mathrm{e}(P)}\max\{d, h_p(P)\}^{1/2}h(x)^{1/2}.$$

The upper bound for $h(x)$ in (B.3.3) follows from (B.2.6) with $T = h(x)$, $B = 183\frac{d^2}{\mathrm{e}(P)}\max\{d, h_p(P)\}^{1/2}$, and $C = \frac{d}{\mathrm{e}(P)}\mathrm{lgcd}(x,y)$. The same upper bound holds for $h(y)$ since $h(y) \leq h(x)$. □

4. On a Theorem of Sprindzhuk

Let $P \in \mathbf{Q}[X,Y]$ be an irreducible polynomial with $\deg_Y P > 0$. A famous theorem of Hilbert states that $P(x,Y) \in \mathbf{Q}[Y]$ is irreducible for infinitely many integers x.

In [**Spr79b**] and [**Spr79a**] Sprindzhuk proved that if $P(0,0) = 0$ and $\frac{\partial P}{\partial Y}(0,0) \neq 0$ then $P(l,Y) \in \mathbf{Q}[Y]$ is irreducible for a sufficiently large prime l. His Theorem is effective and even holds if l is allowed to be a prime power. In [**BM06**] Bilu and Masser gave a quick proof of a more advanced version of this result, originally also due to Sprindzhuk.

In this section we prove a quantitative version of Sprindzhuk's result using our Quasi-equivalence Theorem and its consequence Theorem B.3.

THEOREM B.4. *Let $P \in \mathbf{Q}[X,Y]$ be irreducible with $\deg_Y P > 0$ and $\mathrm{e}(P) = 1$. If $d = \deg P$ and $l \in \mathbf{N}$ is a prime with*

$$\log l \geq 7.10^4 d^6 \max\{d, h_p(P)\}, \tag{B.4.1}$$

then $P(l,Y) \in \mathbf{Q}[Y]$ is irreducible.

The condition $\mathrm{e}(P) = 1$ is equivalent to $P(0,0) = 0$ and $(\frac{\partial P}{\partial X}, \frac{\partial P}{\partial Y})(0,0) \neq 0$. Our hypothesis is thus slightly weaker than Sprindzhuk's: he assumes the non-vanishing of $\frac{\partial P}{\partial Y}$ at $(0,0)$.

In the proof of Theorem B.4 we need the following simple remark: let $l \in \mathbf{N}$ be a prime and $y \in \overline{\mathbf{Q}}$ such that $\mathrm{lgcd}(l,y) > 0$, then

$$[K : \mathbf{Q}]\mathrm{lgcd}(l,y) \geq \log l, \tag{B.4.2}$$

with $K = \mathbf{Q}(y)$. Indeed, by definition we have

$$[K : \mathbf{Q}]\mathrm{lgcd}(l,y) = \sum_{v \in M_K} [K_v : \mathbf{Q}_v] \log \frac{\max\{1, |l|_v, |y|_v\}}{\max\{|l|_v, |y|_v\}}. \tag{B.4.3}$$

One of the terms on the right side of the expression above must be positive. Therefore $|l|_v < 1$ and $|y|_v < 1$ for some $v \in M_K$. In particular this v is a finite place. Since $[K_v : \mathbf{Q}_v]$ is the product of ramification index and rest-class residue of the prime ideal corresponding to v we have $-[K_v : \mathbf{Q}_v] \log \max\{|l|_v, |y|_v\} \geq \log l$. So (B.4.2) follows since all terms of the sum in (B.4.3) are non-negative.

Proof of Theorem B.4: The assumption $\mathrm{e}(P) = 1$ and $P \in \mathbf{Q}[X,Y]$ irreducible imply that P is irreducible in $\overline{\mathbf{Q}}[X,Y]$. Indeed otherwise P would be the product of 2 polynomials both vanishing at $(0,0)$, therefore $\frac{\partial P}{\partial X}$ and $\frac{\partial P}{\partial Y}$ would both vanish there too.

Say for brevity $p = \deg_X P$ and $q = \deg_Y P$. Clearly we may assume $p > 0$. Let l be as in the hypothesis, we note $\log l \geq 1$. We first show that $P(l,Y) \notin \mathbf{Q}$. Indeed this follows easily for example by Lemma A.1(iii). We fix $y \in \overline{\mathbf{Q}}$ with $P(l,y) = 0$. We must show $D = [\mathbf{Q}(y) : \mathbf{Q}] = q$. The inequality $D \leq q$ holds trivially so it suffices to show $D \geq q$.

By Corollary B.1 and the hypothesis (B.4.1) we conclude $\mathrm{lgcd}(l,y) > 0$ and so

$$D\mathrm{lgcd}(l,y) \geq \log l \tag{B.4.4}$$

by (B.4.2).

By Theorem B.3 and $\mathrm{e}(P) = 1$ we obtain

$$\mathrm{lgcd}(l,y) - \frac{1}{q}h(l) \leq 183d \max\{d, h_p(P)\}^{1/2} \max\{1, h(l), h(y)\}^{1/2}. \tag{B.4.5}$$

Furthermore, $h(l) = \log l$. In (B.4.5) we bound $\mathrm{lgcd}(l,y)$ from below with (B.4.4) and then multiply with $Dq \leq q^2$ to get

$$(q-D)\log l \leq 183dq^2 \max\{d, h_p(P)\}^{1/2} \max\{1, \log l, h(y)\}^{1/2}. \tag{B.4.6}$$

The proof follows by splitting up into two cases:

First say $\log l \geq h(y)$. Then (B.4.6) implies

$$(q-D)^2 \log l \leq 183^2 d^2 q^4 \max\{d, h_p(P)\}.$$

This inequality contradicts (B.4.1) if $q - D \geq 1$. Hence $q \leq D$ and we are done in this case.

Second say $\log l \leq h(y)$. By Corollary A.1 and (B.4.1) we have $h(y) \leq 2\frac{p}{q}\log l$. Therefore $\max\{1, \log l, h(y)\} \leq 2\frac{p}{q}\log l$. We insert this inequality in (B.4.6) to conclude

$$(q-D)^2 \log l \leq 7.10^4 d^2 pq^3 \max\{d, h_p(P)\}. \tag{B.4.7}$$

As before, if $q - D \geq 1$, then (B.4.7) is incompatible with (B.4.1). We conclude $q \leq D$.□

Bibliography

[Abo06] Mourad Abouzaid, *Heights and logarithmic gcd on algebraic curves*, Preprint (Oct. 3rd 2006).

[AD99] Francesco Amoroso and Sinnou David, *Le problème de Lehmer en dimension supérieure*, J. Reine Angew. Math. **513** (1999), 145–179.

[AD01] ———, *Densité des points à coordonnées multiplicativement indépendantes*, Ramanujan J. **5** (2001), 237–246.

[AD03] ———, *Minoration de la hauteur normalisée dans un tore*, J. Inst. Math. Jussieu **2** (2003), no. 3, 335–381.

[AD04] ———, *Distribution des points de petite hauteur dans les groupes multiplicatifs*, Ann. Scuola Norm. Sup. Pisa **3** (2004), 325–348.

[BG06] Enrico Bombieri and Walter Gubler, *Heights in Diophantine Geometry*, Cambridge University Press, 2006.

[BGS94] J.-B. Bost, H. Gillet, and C. Soulé, *Heights of projective varieties and positive Green forms*, J. Amer. Math. Soc. **7** (1994), no. 2, 903–1027.

[Bil97] Y. Bilu, *Limit distribution of small points on algebraic tori*, Duke Math. J. **89** (1997), no. 3, 465–476.

[BM06] Y.F. Bilu and D. Masser, *A quick proof of Sprindzhuk's decomposition theorem*, Bolyai Soc. Math. Stud. **15** (2006), 25–32.

[BMZ99] E. Bombieri, D. Masser, and U. Zannier, *Intersecting a curve with algebraic subgroups of multiplicative groups*, Int. Math. Res. Not. **20** (1999), 1119–1140.

[BMZ03] ———, *Finiteness results for multiplicatively dependent points on complex curves*, Michigan Math. J. **51** (2003), no. 3, 451–466.

[BMZ04] ———, *Intersecting a plane with algebraic subgroups of multiplicative groups*, Preprint (Dec. 29th 2004).

[BMZ06a] ———, *Intersecting curves and algebraic subgroups: conjectures and more results*, Trans. Amer. Math. Soc. **358** (2006), no. 5, 2247–2257.

[BMZ06b] ———, *Anomalous subvarieties - structure theorems and applications*, Preprint (July 19th 2006).

[Bom83] Enrico Bombieri, *On Weil's "théorème de décomposition"*, Amer. J. Math. **105** (1983), no. 2, 295–308.

[Boy98] David W. Boyd, *Uniform approximation to mahler's measure in several variables*, Canad. Math. Bull. **41** (1998), no. 1, 125–128.

[BS96] F. Beukers and H.P. Schlickewei, *The equation $x + y = 1$ in finitely generated groups*, Acta Arith. **78** (1996), 189–199.

[BZ95] E. Bombieri and U. Zannier, *Algebraic Points on Subvarieties of $\mathbf{G}_m^n$*, Int. Math. Res. Not. **7** (1995), 333–347.

[Cas57] J.W.S. Cassels, *An Introduction to Diophantine Approximation*, Cambridge University Press, 1957.

[Che51] C. Chevalley, *Introduction to the theory of algebraic functions of one variable*, American Mathematical Society, 1951.

[CZ00] Paula B. Cohen and Umberto Zannier, *Multiplicative dependence and bounded height, an example*, Proceedings of the Algebraic Number Theory and Diophantine Approximation Conference, Graz, 1998 (Walter de Gruyter, 2000), 93–101.

[Dan94] V.I. Danilov, *Algebraic varieties and schemes*, Algebraic geometry I (I.R. Shafarevich, ed.), Encyclopaedia of Mathematical Sciences **23**, Springer, 1994.

[DP99] Sinnou David and Patrice Philippon, *Minorations des hauteurs normalisées des sous-variétés des tores*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **28** (1999), no. 3, 489–543.

[Eve84] J.H. Evertse, *On equations in S-units and the Thue-Mahler equation*, Invent. Math. **75** (1984), 561–584.

[Eve02] Jan-Hendrik Evertse, *Points on subvarieties of tori*, pp. 214–230, Cambridge Univ. Press, 2002.

[Fri89] Eduardo Friedman, *Analytic formulas for the regulator of a number field*, Invent. Math. **98** (1989), 599–622.

[Ful84] William Fulton, *Intersection theory*, Springer, 1984.

[Gal07] Aurélien Galateau, *Minoration de la hauteur normalisée dans un produit de courbes elliptiques*, Preprint (Jan. 2007).

[GKZ94] I.M. Gel'fand, M.M. Kapranov, and A.V. Zelevinsky, *Discriminants, Resultants, and Multidimensional Determinants*, Birkäuser, 1994.

[Hab05] P. Habegger, *The equation $x + y = \alpha$ in multiplicatively dependent unknowns*, Acta Arith. **119** (2005), no. 4, 349–372.

[Hab07] ———, *Multiplicative dependence and isolation I*, In Diophantine Geometry, Scuola Normale Superiore. Serie CRM. Vol. 4 (2007), 189–196.

[Har98] Glyn Harman, *Metric number theory*, Oxford University Press, 1998.

[HS83] David Lee Hilliker and E.G. Strauss, *Determination of bounds for the solutions to those binary Diophantine equations that satisfy the hypotheses of Runge's theorem*, Trans. Amer. Math. Soc. **280** (1983), no. 2, 637–657.

[HW05] G.H. Hardy and E.M. Wright, *An Introduction to the Theory of Numbers*, Oxford University Press, 2005.

[Kir06] J. Kirby, *The Theory of Exponential Differential Equations*, Ph.D. thesis, 2006.

[Lan83] Serge Lang, *Fundamentals of Diophantine Geometry*, Springer, 1983.

[Lan94] ———, *Algebraic Number Theory*, Springer, 1994.

[Lan02] ———, *Algebra*, Springer, 2002.

[Lau84] Michel Laurent, *Équations diophantiennes exponentielles*, Invent. Math. **78** (1984), no. 2, 299–327.

[Lju60] Wilhelm Ljunggren, *On the irreducibility of certain trinomials and quadrinomials*, Math. Scand. **8** (1960), 65–70.

[Mas07] D.W. Masser, *Counting points with multiplicatively dependent coordinates on a curve*, In Diophantine Geometry, Scuola Normale Superiore. Serie CRM. Vol. 4 (2007), 221–236.

[Mau06] Guillaume Maurin, *Conjecture de Zilber-Pink pour les courbes tracées sur des tores*, Preprint (Nov. 10th 2006).

[Mig89] M. Mignotte, *Sur un théorème de M. Langevin*, Acta Arith. **54** (1989), 81–86.

[Phi95] Patrice Philippon, *Sur des hauteurs alternatives III*, J. Math. Pures Appl. **74** (1995), 345–365.

[Pin05a] Richard Pink, *A combination of the conjectures of Mordell-Lang and André-Oort*, Geometric Methods in Algebra and Number Theory, Progr. Math. **235**, Birkhäuser Boston, 2005, pp. 251–282.

[Pin05b] ———, *A Common Generalization of the Conjectures of André-Oort, Manin-Mumford, and Mordell-Lang*, Preprint (Apr. 17th 2005).

[Pon05] Corentin Pontreau, *Minoration effective de la hauteur des points d'une courbe de $\mathbb{G}_m^2$ définie sur $\mathbb{Q}$*, Acta Arith. **120** (2005), no. 1, 1–26.

[Poo99] Bjorn Poonen, *Mordell-Lang plus Bogomolov*, Invent. Math. **137** (1999), no. 2, 413–425.

[Pou04] Dimitrios Poulakis, *Integer points on rational curves with fixed gcd*, Publ. Math. Debrecen **64** (2004), 369–379.

[Rat07] N. Ratazzi, *Intersection de courbes et de sous-groupes, et problèmes de minoration de hauteur dans les variétés abéliennes C.M.*, Preprint (Jan. 30th 2007).

[Rém03] Gaël Rémond, *Approximation diophantienne sur les variétés semi-abéliennes*, Ann. Scuola Norm. Sup. Pisa **36** (2003), no. 2, 191–212.

[Rém05a] ———, *Inégalité de Vojta généralisée*, Bull. Soc. Math. France **133** (2005), 459–495.

[Rém05b] ———, *Intersection de sous-groupes et de sous-variétés I*, Math. Ann. **333** (2005), 525–548.

[Rém07] ———, *Intersection de sous-groupes et de sous-variétés II*, J. Inst. Math. Jussieu **6** (2007), no. 2, 317–348.

[RT96] Damien Roy and Jeffrey Lin Thunder, *An absolute Siegel's lemma*, J. Reine Angew. Math. **476** (1996), 1–26.

[RT99] ———, *Addendum and erratum to: "An absolute Siegel's lemma"*, J. Reine Angew. Math. **508** (1999), 47–51.

[Run87] C. Runge, *Ueber ganzzahlige Lösungen von Gleichungen zwischen zwei Veränderlichen*, J. Reine Angew. Math. **100** (1887), 425–435.

[RV03] Gaël Rémond and Evelina Viada, *Problème de Mordell-Lang modulo certaines sous-variétés abéliennes*, Int. Math. Res. Not. **35** (2003), 1915–1931.

[Sch97] H.P. Schlickewei, *Lower bounds for heights on finitely generated groups*, Monatsh. Math. **123** (1997), no. 2, 171–178.

[Sch00] A. Schinzel, *Polynomials with special regard to reducibility. With an appendix by Umberto Zannier*, Encyclopedia of Mathematics and its Applications, vol. 77, Cambridge University Press, 2000.

[Smy71] C.J. Smyth, *On the product of the conjugates outside the unit circle of an algebraic integer*, Bull. London Math. Soc. **3** (1971), 169–175.

[Smy81] ———, *On measures of polynomials in several variables*, Bull. Austral. Math. Soc. **23** (1981), 49–63.

[Spr79a] V.G. Sprindzhuk, *Hilbert's irreducibility theorem and rational points on algebraic curves. (English)*, Soviet Math. Dokl. **20** (1979), no. 4, 701–705.

[Spr79b] ———, *Hilbert's irreducibility theorem and rational points on algebraic curves. (Russian)*, Dokl. Akad. Nauk SSSR **247** (1979), no. 2, 285–289.

[Tve60] Helge Tverberg, *On the irreducibility of the trinomials* $x^n \pm x^m \pm 1$, Math. Scand. **8** (1960), 121–126.

[Ull98] Emmanuel Ullmo, *Positivité et discrétion des points algébriques des courbes*, Ann. of Math. **147** (1998), 167–179.

[Via03] Evelina Viada, *The Intersection of a Curve with Algebraic Subgroups in a Product of Elliptic Curves*, Ann. Scuola Norm. Sup. Pisa **2** (2003), 47–753.

[Via07] ———, *Finiteness statements for algebraic points on a curve embedded in a product of elliptic curves*, Preprint (2007).

[Vog84] W. Vogel, *Lectures on results on Bezout's theorem*, Tata Institute of Fundamental Research Lectures on Mathematics and Physics, 74, Springer, 1984.

[Wal92] P.G. Walsh, *A quantitative version of Runge's theorem on diophantine equations*, Acta Arith. **62** (1992), 157–172.

[Zag93] D. Zagier, *Algebraic numbers close to both* 0 *and* 1, Math. Comp. **61** (1993), no. 203, 485–491.

[Zan00] Umberto Zannier, *Appendix by Umberto Zannier in* [**Sch00**] *(pp. 517-539)*, 2000.

[Zha95a] Shouwu Zhang, *Positive line bundles on arithmetic varieties*, J. Amer. Math. Soc. **8** (1995), no. 1, 187–221.

[Zha95b] ———, *Small points and adelic metrics*, J. Algebraic Geom. **4** (1995), no. 2, 281–300.

[Zha98] ———, *Equidistribution of small points on abelian varieties*, Ann. of Math. **147** (1998), 159–165.

[Zil02] Boris Zilber, *Exponential sums equations and the Schanuel conjecture*, J. London Math. Soc. (2) **65** (2002), no. 1, 27–44.

Curriculum Vitae

I was born on July 23rd 1978 in Schlieren, Zurich. I am citizen of Trub, Bern. My parents are Erich and Annemarie Habegger.

Education

Oct.2003	-	Jun.2007	PhD in Mathematics at the University of Basel, Switzerland. Adviser: Prof. D. W. Masser.
Sep.2005	-	Feb.2006	Research visit at the Insitut de Mathématiques de Jussieu, Paris, France.
Oct.2000	-	Apr.2003	Advanced studies in Mathematics at the University of Basel. Degree: Diplom.
Oct.1998	-	Oct.2000	Basic studies in Mathematics at the University of Basel. Degree: Vordiplom.
Aug.1994	-	Dec.1997	Gymnasium Oberwil, Baselland. Degree: Matura.

I have visited lectures and seminars of Prof. C. Bandle, Prof. Y. Bilu, Prof. H. Kraft, Dr. J. Lieberum, Prof. D. W.Masser, Dr. S. Mohrdieck, Prof. W. Reichel, Prof. B. Scarpellini, Prof. G. Tammann. Prof. F. Thielemann, Prof. H. Thomas, and Prof. D. Trautmann.