Lacunary polynomials

Francesco Amoroso
Appunti di Davide Lombardo

1 24.04.2018 - Overview

The basic questions we start from are

1. Let $\xi \in \mathbb{Z}$, $f \in \mathbb{Z}[x]$. Is it true that $f(\xi) = 0$?
2. Given $P \in \mathbb{Z}[x]$, does P divide f?
3. Factorization of f?
4. $\gcd(f, g)$?

For many of these questions, there are efficient algorithms for polynomials in their *dense representation*, that is, if they are given by the complete list of their coefficients. In this course we're interested in *lacunary polynomials*, which we represent as $f(t) = \sum_{i=1}^{N} c_i t^{a_i}$ with $c_i \in \mathbb{Z} \setminus \{0\}$ and $a_i \in \mathbb{N}$.

Definition 1.1. *The lacunary size of $f(t)$ is*

$$\ell(f) = \sum_{i=1}^{N} \max\{\log_2 |c_i|, \log_2 a_i\}$$

Up to constants, this is essentially $\max\{h(f), \log \deg f, N\}$*, where* $h(f) = \log \max |c_i|$.

1.1 Testing roots of polynomials in polynomial time

Let's start with the first problem. Fix $\xi \in \mathbb{Z}$: we want to decide whether $f(\xi) = 0$.

Remark 1.2. The size of $f(\xi)$ is exponential in $\ell(f)$, unless $\xi = 0, 1, -1$. Therefore evaluating the polynomial at ξ is not a good strategy to determine whether f vanishes at ξ.

Question 1.3 (Cueken-Koiran-Smale)**.** *Does there exist a polynomial-time algorithm to decide whether $f(\xi) = 0$?*

Theorem 1.4 (Lenstra)**.** *The answer is affirmative.*

Remark 1.5. The problem is nontrivial only if $\log \deg f \approx N$, for otherwise simply evaluating at ξ is efficient.

Lenstra's idea is to exploit the fact that there must be large gaps between the monomials of f. Write $f(t) = r(t) + t^u q(t)$, with $\deg r = k < u$ and $u - k$ *large*. Suppose $f(\xi) = 0$. Then, under suitable assumptions, both q and r must vanish at ξ.

sketch. Assume by contradiction $f(\xi) = 0 \neq q(\xi)$ and write. Then one has

$$|\xi|^u \leq |\xi|^u |q(\xi)| = |\xi^u q(\xi)| = |r(\xi)| \leq \|f\|_1 |\xi|^k,$$

and therefore $\|f\|_1 \geq |\xi|^{u-k} \geq 2^{u-k}$, that is,

$$u - k \leq \frac{\log \|f\|_1}{\log 2}.$$

This implies that the gap is bounded by a polynomial in the lacunary height (which is a contradiction *under suitable assumptions*, that is, there is a certain tension between this and the fact that the degree is exponentially large in the number of terms). This allows us to write

$$f(t) = \sum t^{n_j} f_j(t)$$

with $\deg f_j$ *small*, and simply evaluate the various $f_j(t)$. □

Definition 1.6. *The **Weil height** of $\xi \in \overline{\mathbb{Q}}^\times$ is*

$$h(\xi) = \frac{1}{d}\left(\log a + \sum_{i=1}^{d} \log \max\{|\xi_j|, 1\}\right),$$

where $d = \deg(\xi)$ and ξ_j are the conjugates of ξ.

A similar bound holds for algebraic numbers, namely one gets

$$u - k \leq \frac{\log \|f\|_1}{h(\xi)}.$$

Remark 1.7. $h(\xi)$ is non-negative, and is zero precisely when ξ is a root of unity. However, if one does not fix a bound on the degree of ξ, the height is not bounded away from zero: for example, $h(2^{1/d}) = \frac{\log 2}{d}$.

A conjecture of Lehmer (1933) states that there exists a universal constant $C > 0$ such that for all $\xi \in \overline{\mathbb{Q}}^\times \setminus \mu_\infty$ the inequality

$$h(\xi) \geq \frac{C}{d}$$

holds.

Theorem 1.8 (Dobrowolski 1979)**.** *For all $\varepsilon > 0$ $\exists C_\varepsilon > 0$ such that for every ξ which is not a root of unity one has*

$$h(\xi) \geq \frac{C_\varepsilon}{d^{1+\varepsilon}}$$

Putting these ingredients together, one gets:

Theorem 1.9. *There exists a polynomial-time algorithm (with respect to the lacunary height) for the research of irreducible factors of **bounded** degree.*

Example 1.10. One cannot do much better: for $f(t) = t^p - 1$, one has $\ell(f) \approx \log p$, but $f(t) = (t-1)(t^{p-1} + \ldots + 1) = (t-1)h(t)$ is such that $\ell(h) \approx p$, which is exponential in $\ell(f)$.

1.2 Multivariate polynomials

Let's consider the case of a bi-variate polynomial. One has a dictionary of sorts with the previous case:

- elements of $\overline{\mathbb{Q}}^\times$ ↔ elements of $\mathbb{G}_m(\overline{\mathbb{Q}})^2$
- $\pm 1 \leftrightarrow 1$, subtori $x^a y^b - 1 = 0$ with $a, b \in \mathbb{Z}, (a, b) = 1$.
- Roots of unity ↔ torsion subvarieties, that is, translates by a torsion point $\xi \in (\mathbb{G}_m^2)_{\text{tors}}$ of subtori

- Weil height $h(\xi) \leftrightarrow$ Weil height for points; for curves $\mathcal{C}$, one often uses the **essential minimum**:

  Definition 1.11.
  $$\mu_{ess}(\mathcal{C}) = \inf \left\{ h > 0 \,\middle|\, \exists \infty\ \alpha \in \mathcal{C} : h(\alpha) \leq h \right\}$$

Theorem 1.12 (Manin-Mumford, toric case)**.** *$\mu_{\mathrm{ess}}(\mathcal{C}) = 0$ if and only if $\mathcal{C}$ is a torsion subvariety. If nonzero, the essential minimum is bounded below by an explicit expression.*

Again, one obtains polynomial-time algorithms for the research of irreducible factors of bounded degree for lacunary polynomials in 2 variables.

1.3 Greatest common divisor

It is a fact of life that the coefficients of the GCD of two lacunary polynomials can be very large with respect to the lacunary height. For this reason, instead of working with the lacunary height introduced above, one fixes the *skeleton* of f:

$$f = f_{\underline{a}}(t) = F(t^{a_1}, \ldots, t^{a_N}),$$

where F is a fixed (Laurent) polynomial with integer coefficients.

Remark 1.13 (U. Zannier)**.** In geometric terms, F is a regular function on $\mathbb{G}_m^N$, and f is the specialization of F to a 1-parameter subgroup; the *variable* of the problem is this subgroup.

Remark 1.14. $d = \deg f_{\underline{a}} \approx \max(a_i)$

Theorem 1.15 (Plaisted)**.** *Computation of gcd with respect to the lacunary height is NP-hard (that is, any NP problem can be reduced to it in polynomial time).*

Theorem 1.16 (Filoseta, Granville, Schinzel)**.** *Let $f, g \in \mathbb{Z}[x]$ be polynomials **without cyclotomic factors**. Then one can compute (f, g) in quasi-linear time in $\log(d)$ (with constants depending on the skeletons of f and g, and where quasi-linear time means linear up to powers of $\log \log d$)*

More precisely,

Theorem 1.17 (Amoroso-Leroux-Sombra)**.** *With the same complexity and with no assumptions on f, g, one can compute $p \in \mathbb{Z}[x]$ such that $p(x) \mid \gcd(f, g)$ and $\frac{\gcd(f,g)}{p(x)}$ is a product of cyclotomic polynomials.*

Example 1.18.
$$\gcd(t^{ab} - 1, (t^a - 1)(t^b - 1)) = \frac{(t^a - 1)(t^b - 1)}{t - 1}$$
is an example of lacunary polynomials whose GCD is not lacunary.

Corollary 1.19. *One can decide whether $(f, g) = 1$ in quasi-linear time.*

Conjecture 1.20 (Schinzel)**.** *$F, G \in \mathbb{Z}[X_1, \ldots, X_N]$ relatively prime. Suppose that $\xi \neq$ root of 1 is a common root of $F(t^{a_1}, \ldots, t^{a_N})$ and $G(t^{a_1}, \ldots, t^{a_N})$. Then there exists a vector $\underline{b} \in \mathbb{Z}^N \setminus 0$ with $\|\underline{b}\| \leq C(F, G)$ such that $\underline{b} \perp \underline{a}$. The crucial point here is that $\underline{b}$ does not depend on $\underline{a}$.*

Remark 1.21. This is now a theorem of Bombieri-Zannier.

Example 1.22. The constant $C(F, G)$ must depend not just on the monomials appearing in F, G but also on their coefficients, as the following example shows.

$F(x, y) = x - 2$, $G(x, y) = y - 2^a$. Then $f(t) = F(t, t^a) = t - 2$ and $g(t) = G(t, t^a) = t^a - 2^a$ both vanish at 2, but any vector $\underline{b}$ orthogonal to $(1, a)$ must necessarily grow in norm as $a \to \infty$.

The idea behind the algorithm of Filaseta-Granville-Schinzel

- If $(F, G) \neq 1$ all the better! By specialization they will still have a common factor.
- Suppose $(F, G) = 1$ but $(f, g) \neq 1$. Then there exists ξ such that $f(\xi) = g(\xi) = 0$. By assumption, ξ is not a root of unity. By Schinzel-Bombieri-Zannier, there exists a small $\underline{b}$ orthogonal to $\underline{a}$. After a change of variables, we can assume $\underline{b} = (0, \ldots, 0, 1)$, i.e. $a_N = 0$, and now we continue by induction.

Remark 1.23. This gives more: it shows that GCDs between lacunary polynomials can be found by computing GCDs between not exactly their skeletons, but polynomials obtained from the skeletons via 'controlled' changes of variable.

Theorem 1.24 (Amoroso-Sombra-Zannier)**.** *$C(F, G)$ is bounded polynomially by the logarithmic height of the coefficients of F, G.*

Remark 1.25. This has applications to the study of multiple roots of a polynomial: one needs to compute (f, f'), and f' necessarily involves the degrees a_i. The logarithmic dependence on the coefficients ensures that this computation can be done in polynomial time in the input.

Remark 1.26. Schinzel ('65) conjectures analogue results for the (complete) factorisation of $f_{\underline{a}} = F(t^{a_1}, \ldots, t^{a_N})$. More precisely, there should exist a matrix $A \in M_{N\times N}(\mathbb{Z})$ in a finite set (depending only on F) such that $\underline{a} = A\underline{a}'$, and the factorisation of $f_{\underline{a}}$ comes, up to cyclotomic factors, from the factorisation of $F(\underline{y}^{\underline{A}}$ by specialisation $\underline{y} \mapsto t^{\underline{a}'}$.

Remark 1.27. This conjecture seems to be out of reach with the current methods. However, the function field analogue (that is, replace $\mathbb{Z}[t]$ with $\mathbb{C}(z)[t]$) is (almost) a direct consequence of work of Dvornicich-Zannier and Zannier.

2 26.04.2018 – Valuations, height...

2.1 Absolute values

Let K be a field. An **absolute value** on K is a map $|\cdot| : K \to \mathbb{R}^+$ that satisfies:

- $x \in K$, $|x| = 0$ if and only if $x = 0$
- $|xy| = |x| \cdot |y|$
- $|x + y| \leq |x| + |y|$; if furthermore $|x + y| \leq \max\{|x|, |y|\}$, then $|\cdot|$ is said to be **non-archimedean**

Remark 2.1.
- $|1| = 1$
- More generally, $|\omega| = 1$ for every root of unity $\omega \in K$
- if $A \subseteq K$ is a subring and K is the fraction field of A and $|\cdot|$ is defined on A, then $|\cdot|$ can be extended to K

Remark 2.2. The **trivial absolute value** is $|x| = 1$ for all $x \in K^\times$ and $|0| = 0$.

2.2 Valuations

A **valuation** on K is a function $v : K \to \mathbb{R} \cup \{\infty\}$ such that:

- $v(x) = \infty$ if and only if $= \infty$
- $\forall x, y \in K \quad v(xy) = v(x) + v(y)$
- $\forall x, y \in K \quad v(x + y) \geq \min\{v(x), v(y)\}$

If v is a valuation on K and $a \in \mathbb{R}$ is > 1, then $|x| := a^{-v(x)}$ is a non-archimedean absolute value (with the convention $a^{-\infty} = 0$). Conversely, given a non-archimedean absolute value, the function $v(x) = -b \log |x|$ is a valuation.

2.2.1 Valuation ring

Let v be a valuation on K. The set

$$A = A_v = \{x \in K : v(x) \geq 0\}$$

is a valuation ring (that is, for all $x \in K$ either x or x^{-1} belongs to A). It is a local ring with maximal ideal

$$M = \{x \in K : v(x) > 0\}$$

and group of units

$$A^{\times} = \{x \in K : v(x) = 0\}.$$

Remark 2.3. When $|x|_v = a^{-v(x)}$ is induced by a valuation v, the ring $A_v = \{x \in K : |x|_v \leq 1\}$ and the ideal $M_v = \{x \in K : |x| < 1\}$ are independent of a.

Definition 2.4. *Two absolute values $|\cdot|_1$, $|\cdot|_2$ are* ***equivalent*** *if they define the same topology.*

Proposition 2.5. *Two absolute values $|\cdot|_1$, $|\cdot|_2$ are equivalent if and only if the corresponding maximal ideals are the same, that is, if and only if*

$$\{x \in K : |x|_1 < 1\} = \{x \in K : |x|_2 < 1\}.$$

In turn, this is equivalent to the fact that there exists $a > 0$ such that $|\cdot|_1 = |\cdot|_2^a$.

Definition 2.6. *Two valuations v_1, v_2 of K are equivalent if the corresponding absolute values are equivalent.*

Proposition 2.7. *Two valuations are equivalent if and only if they are proportional (with the proportionality constant being positive).*

2.3 Valuations on $\mathbb{Q}$

The following are absolute values on $\mathbb{Q}$:

- the trivial absolute value
- the standard (real) absolute value $|x|_\infty = \begin{cases} x, & \text{if } x \geq 0 \\ -x, & \text{if } x < 0 \end{cases}$
- for every prime number p, we can write each nonzero rational number x as $p^{v_p(x)}\frac{a}{b}$, where $a, b \in \mathbb{Z}$ are prime to p and $v_p(x) \in \mathbb{Z}$. In the equivalence class of absolute values induced by v_p we take $|x|_p = p^{-v_p(x)}$, so that $|p| = \frac{1}{p}$.

Theorem 2.8 (Ostrowski)**.** *The previous list of absolute values on $\mathbb{Q}$ is complete.*

Definition 2.9. *We denote by $M_{\mathbb{Q}}$ the set of nontrivial absolute values on $\mathbb{Q}$, normalised as in the previous list.*

2.4 Some diophantine inequalities

We saw the other day that the following two (trivial) inequalities were essential in studying the problem of determining whether $f(\xi)$ vanishes or not:

- $\forall x \in \mathbb{Z}, x \neq 0$, we have $|x| \geq 1$
- $\forall x \in \mathbb{Z}, x \neq 0, \pm 1$, we have $|x| \geq 2$

These inequalities generalise to $\mathbb{Q}$ in the following way:

Theorem 2.10 (Product formula, or the fundamental theorem of arithmetic)**.** *For every $x \in \mathbb{Q}^{\times}$ we have*

$$\prod_{v \in M_{\mathbb{Q}}} |x|_v = 1.$$

2.5 Valuations on number fields

Let $K/\mathbb{Q}$ be a number field of degree $d = r+2s$ (where as usual r is the number of real embeddings and $2s$ is the number of complex embeddings). For every $\sigma : K \hookrightarrow \mathbb{C}$ we obtain an archimedean absolute value by setting $|x|_\sigma = |\sigma(x)|$. Then:

Proposition 2.11. *K admits exactly $r+s$ archimedean (equivalence classes of) absolute values; a set of representative is given by $|\cdot|_\sigma$, where σ varies among the r real embeddings and s non-conjugate complex embeddings.*

As for the non-archimedean valuations, recall from algebraic number theory the notion of rings of integers, prime ideals, ramification index e and inertia degree f and the fact that e and f are multiplicative in towers of extensions.

Theorem 2.12. *Let L/K be an extension of number fields and let $\mathcal{P}$ be a prime of $\mathcal{O}_K$. Writing $\mathcal{P}\mathcal{O}_L = \mathcal{Q}_1^{e_1} \cdots \mathcal{Q}_k^{e_k}$ we have*

$$\sum_{i=1}^{k} e(\mathcal{Q}_i \mid \mathcal{P}) f(\mathcal{Q}_i \mid \mathcal{P}) = [L : K].$$

Definition 2.13. *Let $\mathcal{P}$ be a prime ideal of $\mathcal{O}_K$. For $x \in K^\times$, write $(x) = \mathcal{P}^\lambda \cdot \prod_{\mathcal{Q} \neq \mathcal{P}} \mathcal{Q}^{e_\mathcal{Q}}$ with $\lambda \in \mathbb{Z}$ and set*

$$v_\mathcal{P}(x) = \frac{\lambda}{e(\mathcal{P} \mid p)},$$

where $(p) = \mathcal{P} \cap \mathbb{Z}$. We normalize the absolute value by $|x|_\mathcal{P} = p^{-v_\mathcal{P}(x)}$; this normalization ensures that the product formula holds.

Theorem 2.14. *The previous list of nontrivial absolute values on K is complete.*

Definition 2.15. *We denote by M_K the set of nontrivial absolute values on K, normalised as in the previous constructions. We shall confuse an equivalence class of absolute values with its representative constructed above.*

Given a *place* v (that is, an equivalence class of absolute values and/or of absolute values) we write

$$n_v = [K_v : \mathbb{Q}_v] = \begin{cases} 1, \text{ if } v \mid \infty, v = \sigma \text{ with } \sigma(K) \subseteq \mathbb{R} \\ 2, \text{ if } v \mid \infty, v = \sigma \text{ with } \sigma(K) \not\subseteq \mathbb{R} \\ e(\mathcal{P} \mid p) f(\mathcal{P} \mid p), \text{ if } v = \mathcal{P} \mid p \end{cases}$$

Remark 2.16. The algebraic closure of $\mathbb{Q}_p$ is *not* complete; its completion $\widehat{\overline{\mathbb{Q}_p}}$ is often denoted by $\mathbb{C}_p$. With the language of completions, one recovers a certain symmetry between archimedean and non-archimedean valuations: the former are given by the embeddings of K in $\mathbb{C}$, the latter by embeddings of K in $\mathbb{C}_p$ (indeed, given $\sigma : K_v \hookrightarrow \mathbb{C}_p$ we get an absolute value by setting $|x|_v = |\sigma(x)|_p$).

Theorem 2.17 (Product formula)**.** *One has: $\forall x \in K^\times$,*

$$\prod_{v \in M_K} |x|_v^{n_v} = 1.$$

Definition 2.18. *Let $K \subseteq L$ be an extension of number fields, $v \in M_K, w \in M_L$. We say that $w \mid v$ if $\forall \alpha \in K$ we have $|\alpha|_w = |\alpha|_v$.*

Remark 2.19. For $w, v \mid \infty$, $w = [\sigma]$, $v = [\tau]$, $w \mid v$ is equivalent to $\sigma|_K = \tau$.

For w, v finite, $w = [\mathcal{Q}], v = [\mathcal{P}]$ $w \mid v$ is equivalent to $\mathcal{Q} \mid \mathcal{P}$.

Proposition 2.20. *For $v \in M_K$ we have $\sum_{w \in M_L, w \mid v} \frac{n_w}{n_v} = [L : K]$*

2.6 Weil height

Given $\alpha \in \overline{\mathbb{Q}}^{\times}$, fix a number field K that contains α.

Definition 2.21. *The (logarithmic absolute)* ***Weil height*** *of α is*

$$h(\alpha) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M_K} n_v \log^+ |\alpha|_v,$$

where $\log^+(x) = \max\{\log x, 0\}$ *for* $x > 0$.

2.6.1 Properties of the Weil height

1. $h(\alpha) \geq 0$, with equality iff α is a root of unity (proof: $h(\alpha) = 0$ iff all summands are 0. This is only possible if α is integral *and* all its archimedean absolute values are 1. Br Kronecker, this implies that α is a root of unity)
2. $h(\alpha\beta) \leq h(\alpha) + h(\beta)$ (proof: it follows immediately from $\max\{1, xy\} \leq \max\{1, x\} \max\{1, y\}$ for $x, y > 0$)
3. $\forall n \in \mathbb{Z}$, $h(n\alpha) = |n| h(\alpha)$ (proof: $h(\alpha^{-1}) = h(\alpha)$ by the product formula; the case of positive n is obvious)
4. $\forall \sigma : \mathbb{Q}(\alpha) \hookrightarrow \overline{\mathbb{Q}}$, $h(\sigma(\alpha)) = h(\alpha)$ (proof: σ permutes the absolute values on $\mathbb{Q}(\alpha)$).

2.6.2 Relation with the elementary definition

Let $P \in \mathbb{Z}[x]$, $P(x) = a(x - \alpha_1) \cdots (x - \alpha_d)$ with $\alpha_j \in \overline{\mathbb{Q}}$ and $a \in \mathbb{Z}$.

Definition 2.22. *The* ***Mahler measure*** *of P is* $M(P) := |a| \prod_{i=1}^{d} \max\{1, |\alpha_j|\}$.

Proposition 2.23.

$$M(P) = \exp\left(\frac{1}{2\pi} \int_0^{2\pi} \log |P(e^{it})| dt \right)$$

Let $\alpha \in \overline{\mathbb{Q}}^{\times}$ with minimal polynomial $f \in \mathbb{Z}[x]$. Let $K = \mathbb{Q}(\alpha)$. Then

$$\sum_{v \in M_K, v \nmid \infty} n_v \log^+ |\alpha|_v = \log(a),$$

which implies $h(\alpha) = \frac{1}{d} \log M(P)$.

3 02.05.2018 – Lower bounds on the height, Dobrowolski's theorem

3.1 Finding irreducible factors of a lacunary polynomial $f(t)$

Write

$$f(t) = (t) + t^u q(t), \quad \deg r = k < u$$

Now let $\xi \in \overline{\mathbb{Q}}$, $\deg \xi = d$. We want to show that if $u - k$ is large, then $f(\xi) = 0$ implies $r(\xi) = q(\xi) = 0$. By contradiction, assume $x := \xi^{-u} r(\xi) = -q(\xi) \neq 0$.

To simplify the notation, given a place v write $\delta_v = 0$ if $v \nmid \infty$ and $\delta_v = 1$ if $v \mid \infty$. For any given place v we have both

$$|x|_v = |\xi^{-u} r(\xi)|_v \leq \|f\|_1^{\delta_v} |\xi|_v^{-u} \max\{1, |\xi|_v\}^k$$

and

$$|q(\xi)| \leq \|f\|_1^{\delta_v} \max\{1, |\xi|_v\}^{\deg q};$$

optimizing (ie choosing the first inequality if $|\xi| > 1$ and the second otherwise) we obtain

$$|x|_v \leq \|f\|^{\delta_v} \max\{1, |\xi|_v\}^{k-u},$$

where $k - u$ is negative. Using the product formula, we get

$$1 = \prod_v |x|_v^{n_v/d} \leq \|f\|_1^{\sum_{v|\infty} \frac{n_v}{d}} \left(\prod_v \max\{1, |\xi|_v\}^{n_v/d} \right)^{-(u-k)} = \|f\|_1 H(\xi)^{-(u-k)}.$$

If ξ is *not* a root of unity, this implies

$$u - k \leq \frac{\log \|f\|_1}{h(\xi)},$$

which is a contradiction if $u - k$ is large enough. For our applications it is enough to know any nontrivial lower bound on $h(\xi)$ (when ξ is not a root of unity); however, for completeness, we now discuss finer lower bounds on the height.

3.2 Lower bounds on the height I: algebraic numbers

Let $\xi \in \overline{\mathbb{Q}}^{\times}$ be an algebraic number which is not a root of unity.

Conjecture 3.1. *('Optimal' conjecture; Lehmer's conjecture) There exists an absolute constant $C > 0$ such that, for every algebraic ξ of degree d which is not a root of unity, the inequality $h(\alpha) \geq \frac{C}{d}$ holds.*

Theorem 3.2. *(Dobrowolski '77) There exists $C > 0$ such that, for every ξ of degree $d \geq 3$ (ξ not a root of unity), the following inequality holds:*

$$h(\xi) \geq \frac{C}{d} \left(\frac{\log \log d}{\log d} \right)^3$$

Theorem 3.3. *(Smyth) Lehmer's conjecture is true if one restricts to algebraic numbers which are not reciprocal (that is, such that α^{-1} is not a conjugate of α).*

Definition 3.4. *(Bombieri-Zannier) Let $K \subseteq \overline{\mathbb{Q}}$ be an infinite algebraic extension of $\mathbb{Q}$. We say that K has property (B) if $\exists C > 0$ such that $h(\alpha) > C$ for every $\alpha \in K^{\times}$ which is not a root of unity.*

Remark 3.5. (B) stands for Bogomolov.

Example 3.6. Here are some fields for which property (B) holds:

- (Schinzel) $\mathbb{Q}^{tr}$, the composiitum of all totally real extensions of $\mathbb{Q}$. More precisely, for every $\alpha \in \mathbb{Q}^{tr}$, $\alpha \neq 0, 1, -1$, one has $H(\alpha) \geq \left(\frac{1+\sqrt{5}}{2}\right)^{1/2}$. One can prove this result by applying Bilu's equidistribution theorem (see theorem 3.7 below, which shows that $h(\alpha)$ cannot go to zero if α ranges over totally real algebraic numbers, because the conjugates of α are certainly not equidistributed near the unit circle)

- (Bombieri-Zannier) p prime, $L \subseteq$ finite extension of $\mathbb{Q}_p$. Then L has property (B). For example, given $r \in \mathbb{N}$, the field $\mathbb{Q}^{(r)}$ (compositum of all extensions of $\mathbb{Q}$ of degree $\leq r$) has local degrees which are bounded uniformly in p.

- (Dvornicich-Amoroso) $\mathbb{Q}^{\text{ab}}$. Notice that by Kronecker-Weber $\mathbb{Q}^{\text{ab}} = \mathbb{G}_m(\overline{\mathbb{Q}})_{\text{tors}}$.

- More generally, K^{ab} (K number field) satisfies property (B)
- (Habegger) Let $E/\mathbb{Q}$ be an elliptic curve. Then $\mathbb{Q}(E_{\mathrm{tors}})$ has property (B). Notice that if E has CM then $\mathbb{Q}(E_{\mathrm{tors}})$ is an abelian extension of a quadratic extension of $\mathbb{Q}$.

Theorem 3.7 (Bilu's equidistribution theorem)**.** *Let $(\alpha_n) \subseteq \overline{\mathbb{Q}}$ be a sequence of pairwise distinct algebraic numbers such that $h(\alpha_n) \to 0$ as $n \to \infty$. Then $\delta_{\alpha_n} \to \mu_{\mathbb{T}}$, where*

$$\delta_\alpha = \frac{1}{[\mathbb{Q}(\alpha) : \mathbb{Q}]} \sum_\sigma \delta_{\sigma(\alpha)},$$

the measure $\mu_{\mathbb{T}}$ is the uniform measure on the torus $\mathbb{T} = S^1$, and the convergence is in the weak- topology.*

Observe that $\alpha \in \mathbb{Q}^\times, \alpha \neq \pm 1$ implies $h(\alpha) \geq \log 2$: this is obvious. We now prove a weaker statement than this (in a more complicated way) using techniques that appear in the proof of Dobrowolski and of the fact that $\mathbb{Q}^{\mathrm{ab}}$ has property (B).

Proof. Fix a prime p. Fermat's little theorem gives $a^p - a \equiv 0 \pmod p$ for all $a \in \mathbb{Z}$, which can be restated as $|a^p - a|_p \leq \frac{1}{p}$. By continuity, this is true for every $a \in \mathbb{Z}_p$. If, on the other hand, a does not belong to $\mathbb{Z}_p$, then a^{-1} is in $\mathbb{Z}_p$ and in the same way one obtains

$$|a^p - a| \leq \frac{1}{p}|a|_p^{p+1}.$$

In particular, the inequality $|a^p - a|_p \leq \frac{1}{p}|a|_p^{p+1}$ holds for every a. Applying the product formula to $x := a^p - a$ we obtain

$$1 = \prod_{v \in M_{\mathbb{Q}}} |x|_v = |x| \cdot |x|_p \cdot \prod_{\ell \neq p} |x|_\ell.$$

Notice that for $\ell \neq p$ we also have

$$|a^p - a|_\ell \leq \max\{1, |a^p|_\ell\} \max\{1, |a|_\ell\} = \max\{1, |a|_\ell\}^{p+1},$$

and

$$|a^p - a|_\infty \leq 2 \max\{1, |a|\}^{p+1}.$$

Multiply all these inequalities together to get

$$1 \leq \frac{2}{p} H(a)^{p+1} \Rightarrow h(\alpha) \geq \frac{\log(p/2)}{p+1}.$$

Choosing for example $p = 3$ one gets an absolute lower bound on $h(\alpha)$. □

3.3 Proof of Dobrowolski's theorem (sketch)

Remark 3.8. We can assume α to be integral. Indeed, if α is *not* integral, then there exists $v \nmid \infty$ such that $|\alpha|_v > 1$, and therefore

$$h(\alpha) \geq \frac{n_v \log |\alpha|_v}{d},$$

where $n_v \log |\alpha|_v \in \mathbb{Z} \cdot \log p$. Therefore $h(\alpha) \geq \frac{\log 2}{d}$.

Hence, from now on, we assume that α is an algebraic *integer*.

Remark 3.9. If $F \in \mathbb{Z}[x]$ which vanishes at α with multiplicity $\geq T$, then

$$|F(\alpha^p)|_v \leq p^{-T} \quad \forall p \, \forall v \mid p$$

Proof. Let $f \in \mathbb{Z}[x]$ be the minimal polynomial of α (it is monic since α is integral). From Fermat's little theorem we get

$$f(x)^p \equiv f(x^p) \pmod{p\mathbb{Z}[x]},$$

which (evaluating at α) gives $f(\alpha^p) \equiv 0 \pmod{)}Z[\alpha]$, and therefore $|f(\alpha^p)|_v \leq p^{-1}$. The assumption is that $f^T \mid F$, and therefore $|F(\alpha^p)|_v \leq p^{-T}$. □

We work with an extra parameter L (a bound for the degree of F). Suppose therefore that $\deg F \leq L$ **and** $F(\alpha^p) \neq 0$. Then by the product formula

$$1 = \prod_v |F(\alpha^p)|^{n_v/d} \leq \prod_{v|\infty} |F(\alpha^p)|^{n_v/d} \prod_{v|p} |F(\alpha^p)|^{n_v/d}$$

Furthermore, if $v \mid \infty$, we have the obvious upper bound

$$|F(\alpha^p)|_v \leq \|F\|_1 \max\left\{1, |\alpha|_v\right\}^{pL},$$

and for $v \mid p$ we have already proven $|F(\alpha^p)|_v \leq p^{-T}$. Combining all this with $\sum_{v|\infty} \frac{n_v}{d} = \sum_{v|p} \frac{n_v}{d} = 1$ we obtain

$$1 \leq \|F\|_1 H(\alpha)^{pL} p^{-T},$$

which gives a lower bound on $h(\alpha)$ of the form

$$h(\alpha) \geq \frac{T \log p - \log \|F\|_1}{pL}.$$

If we try to take $F = f$ and $T = 1$ we're not too happy: 'trival' inequalities give $\|f\|_1 \leq 2^d M(f) = 2^d H(\alpha)^d$. One can assume that $H(\alpha)$ is small (because if it's big we're already done!), but $\log \|F\|$ is still linear in d, which forces us to take p exponential in d, and therefore gives a very weak lower bound on $h(\alpha)$. Let's do something better (ie the only thing we know how to do: Siegel's lemma); for $L > 2dT$ (and, for technical reasons, $\log L \ll \log d$), it gives

$$\log \|F\|_1 \leq \left(1 + \frac{dT^2}{L}\right) \log L + dTh(\alpha),$$

where one should think that $dh(\alpha)$ is small (because $h(\alpha)$ is supposed to be very small). Because of the $1 + \frac{\cdots}{L}$, one might as well choose $L \approx dT^2$. Suppose by contradiction that $h(\alpha) \ll \frac{\log d}{dT}$: then we can find an F (of degree at most L, vanishing at order at least T at α) such that

$$\log \|F\|_1 \ll \log d.$$

We have an extra problem: F might vanish at α^p. Let's now do things (almost) properly. We work with an extra parameter N and consider primes p with $\log d \leq p \leq N$.

- if by contradiction $F(\alpha^p) \neq 0$, then the product formula implies $1 \leq p^{-T} \|F\|_1 H(\alpha)^{pL}$, that is, $T \log p \ll \log d + pLh(\alpha) \Rightarrow T \log\log d \ll \log d + NLh(\alpha)$. If we suppose $h(\alpha) < c\frac{\log d}{NL}$, it suffices to show T large enough to obtain a contradiction ($T = C\frac{\log d}{\log\log d}$ suffices), so **for this choice of parameters** $F(\alpha^p) = 0$. Now T is fixed (hence $L \approx dT^2$ is also fixed) and we can only choose N. Notice that the technical assumption $\log L \ll \log d$ is satisfied.

- Now we know that F vanishes at β^p for all primes between $\log d$ and N and for all β conjugate of α. Up to a technical detail ($[\mathbb{Q}(\alpha^p) : \mathbb{Q}] = [\mathbb{Q}(\alpha) : \mathbb{Q}]$ – one can reduce to this case[1]) we obtain

$$L \geq \#\{\text{zeroes of } F\} \geq \sum_{\log d \leq p \leq N} d \gg_{PNT} d\frac{N}{\log N}.$$

[1] if I understand correctly, this reduction needs an assumption like $\log d = o(N)$)

It is now enough to choose $\frac{dN}{\log N} \gg d\left(\frac{\log d}{\log\log d}\right)^2$, which leads to choosing $N \approx d\frac{\log d}{(\log\log d)^2}$. This choice of parameters (N, L, T) gives a contradiction, hence one of the assumptions on $h(\alpha)$ must fail. One then obtains

$$h(\alpha) \gg \frac{\log d}{NL} = \frac{1}{d}\left(\frac{\log\log d}{\log d}\right)^3.$$

4 03.05.2018 – More lower bounds on the height; geometry of $\mathbb{G}_m^N$

4.1 Lower bounds on the height II: abelian extensions

Theorem 4.1 (Amoroso-Dvornicich 2000). *$\forall\alpha \in (\mathbb{Q}^{ab})^\times$, α not a root of unity,*

$$h(\alpha) \geq \frac{\log 5}{12}.$$

Remark 4.2. The constant might be non-optimal, but there exists $\alpha \in (\mathbb{Q}^{ab})^\times$ such that $h(\alpha) = \frac{\log 7}{12}$. To construct such an α, we work in the field $\mathbb{Q}(\zeta_{21})$. In the ring of integers of this field we have $(7) = (\mathcal{P}\overline{\mathcal{P}})^6$; furthermore, $\mathbb{Q}(\zeta_{21})$ has class number 1, so $\mathcal{P} = (\gamma)$. Set $\alpha = \gamma/\overline{\gamma}$. Then α and all its conjugates have modulus 1, and therefore

$$h(\alpha) = \frac{\log N(\gamma)}{[\mathbb{Q}(\zeta_{21}) : \mathbb{Q}]} = \frac{\log 7}{12}.$$

Remark 4.3. One can also invert the argument to prove *lower* bounds on the class number based on lower bounds on the height.

Generalizing the previous remark, one obtains

Theorem 4.4 (Amoroso-Dvornicich 2003). *The exponent of the ideal class group of a CM field K goes to infinity as* $\operatorname{disc}(K) \to \infty$.

Remark 4.5. There is no uniform lower bound on the height for algebraic numbers lying in CM fields. Indeed, a CM field K can be generated by a single element α with $|\alpha| = 1$ (and α not a root of unity; notice that the condition $|\alpha| = 1$ is independent of the choice of absolute value because the field is CM). The field $\mathbb{Q}(\alpha^{1/n})$ is then also CM, and the height of $\alpha^{1/n}$ goes to zero as $n \to \infty$.

4.1.1 Sketch of proof of Theorem 4.1

By Kronecker-Weber, every abelian extension is contained in a cyclotomic one, hence we might work with a cyclotomic field $L = \mathbb{Q}(\zeta_n)$ and with a (generic, but integral) element $\alpha \in L^\times$. Fix p prime (which will be chosen at the end of the proof to be 'small').

We have an extension

$$\begin{tikzcd}
\mathbb{Q}(\zeta_n) \arrow[d, dash, "\text{ab}"'] & (\mathcal{P}_1 \dots \mathcal{P}_r)^e \arrow[d, dash] \\
\mathbb{Q} & p
\end{tikzcd}$$

We consider two cases:

- $e = 1$, or equivalently $p \nmid n$. Let σ be the Frobenius of $\mathcal{P}_j$ over p (this is independent of j, because $L/\mathbb{Q}$ is abelian). Then if α is integral ($\alpha \in \mathcal{O}_L$) we have $\alpha^p \equiv \sigma(\alpha) \pmod{\mathcal{P}_j}$ for every j, and therefore $\alpha^p \equiv \sigma(\alpha) \pmod{p\mathcal{O}_L}$. Concretely, if one doesn't want to talk about Frobenius automorphisms, σ is the unique automorphism of L that sends ζ_n to ζ_n^p.

Now let $x = \alpha^p - \sigma(\alpha) \neq 0$ (if $\alpha^p = \sigma(\alpha)$, then $ph(\alpha) = h(\alpha^p) = h(\sigma(\alpha)) = h(\alpha)$, so $h(\alpha) = 0$ and α is a root of unity). For any α (not necessarily integral) set

$$\Delta_v = \Delta_{v,p,\alpha} := \max\{1, |\alpha|_v\}^p \max\{1, |\sigma\alpha|\}.$$

As in yesterday's lecture, one obtains

$$|x|_v \leq \begin{cases} 2\Delta_v & v \mid \infty \\ \Delta_v & v \nmid \infty \\ \frac{1}{p}\Delta_v & v \mid p \end{cases}$$

To prove the inequality for $v \mid p$: if $\alpha \in \mathcal{O}_L$ we are done, and otherwise one uses the fact that $|\alpha^{-1}|_v < 1$ and $|\alpha^{-p} - \sigma(\alpha^{-1})|_v < 1$.

Applying the product formula,

$$\begin{aligned} 1 &= \prod_{v|\infty} |x|_v^{n_v/d} \cdot \prod_{v|p} |x|_v^{n_v/d} \cdot \prod_{\substack{v\nmid p \\ v\nmid\infty}} |x|_v^{n_v/d} \\ &\leq 2^{\sum_{v|\infty} n_v/d} p^{-\sum_{v|p} n_v/d} \prod_v \Delta_v \\ &\leq 2p^{-1} H(\alpha^p) H(\sigma\alpha) \\ &= \frac{2}{p} H(\alpha)^{p+1}, \end{aligned}$$

which gives $h(\alpha) \geq \frac{\log(p/2)}{p+1}$.

- $p \mid n$, that is, $e > 1$. The idea is to choose *another* automorphism of L which still allows us to obtain lower bounds on the height. Observe that

$$\operatorname{Gal}\left(\mathbb{Q}(\zeta_n)/\mathbb{Q}(\zeta_{n/p})\right) = \langle\sigma\rangle,$$

where $\sigma(\zeta_n) = \zeta_n\zeta_p$ (for a fixed primitive p-th root ζ_p of 1). If $\alpha \in \mathcal{O}_L$ we have $\alpha^p \equiv \sigma(\alpha^p) \pmod{p\mathcal{O}_L}$: this can be proven immediately by just writing $\alpha = f(\zeta_n)$ with $f \in \mathbb{Z}[x]$ and applying Fermat's little theorem:

$$\alpha^p = f(\zeta_n)^p \equiv f(\zeta_n^p) = f((\zeta_n\zeta_p)^p) = f((\sigma\zeta_n)^p) \equiv f(\sigma\zeta_n)^p = \sigma(f(\zeta_n)^p) = \sigma(\alpha^p).$$

Set $x = \alpha^p - \sigma\alpha^p$.

Remark 4.6. As in the proof of Dobrowolski's theorem, there is no guarantee that x is nonzero. However, if x is zero, there exists a root of unity η such that $\eta\alpha \in \mathbb{Q}(\zeta_{n/p})$. Here we use a fundamental property of the Weil height – namely $h(\eta\alpha) = h(\alpha)$ if η is a root of unity – to proceed by induction on the degree of the cyclotomic extension.

We can then assume $x \neq 0$. Set

$$\Delta_v = \Delta_{v,p,\alpha} = \max\{1, |\alpha|_v\}^p \max\{1, |\sigma\alpha|_v\}^p;$$

the same kind of computation we did before yields

$$h(\alpha) \geq \frac{\log(p/2)}{2p}.$$

Choosing $p = 3$ yields a universal lower bound $h(\alpha) \geq \frac{\log(3/2)}{6}$, which however is not as good as $\frac{\log 5}{12}$. To obtain this improved lower bound one needs to work harder (but we won't do it).

4.1.2 Abelian extensions of number fields

$$\begin{tikzcd}
L \arrow[d, dash, "\mathrm{ab}"'] & (\mathcal{Q}_1 \dots \mathcal{Q}_r)^e \arrow[d, dash] \\
K \arrow[d, dash, "d"'] & \mathcal{P} \arrow[d, dash] \\
\mathbb{Q} & p
\end{tikzcd}$$

1. if $e = 1$ one uses Frobenius
2. if $e \neq 1$ some ramification theory is needed. Let $G = \mathrm{Gal}(L/K)$ and
$$G_j = \{\sigma \in G : \forall \gamma \in \mathcal{O}_L, \sigma\gamma \equiv \gamma \pmod{\mathcal{Q}^{j+1}}\}.$$
These groups are independent of the prime $\mathcal{Q}$ above $\mathcal{P}$ because the extension is abelian. $G_0 \supset G_1 \supset G_2 \supset \cdots G_{k-1} \supsetneq G_k = \{0\}$, by using Hasse-Arf one obtains $kq \geq e$, where $q = N(\mathcal{P})$. Let now $\sigma \in G_{k-1}$, $\sigma \neq \mathrm{id}$. Then $\forall \gamma \in \mathcal{O}_L$ we have
$$(\sigma\gamma - \gamma)^q \in \mathcal{Q}^{kq} \subseteq \mathcal{Q}^e,$$
and since this holds for every $\mathcal{Q}$ over $\mathcal{P}$ this implies $(\sigma\gamma - \gamma)^q \equiv 0 \pmod{\mathcal{P}\mathcal{O}_L}$, which (by Fermat again) gives
$$\sigma\gamma^q \equiv \gamma^q \pmod{\mathcal{P}\mathcal{O}_L}.$$
We now proceed as before with $x = \sigma\gamma^q - \gamma^q$; the difficulty is now that of dealing with the case $x = 0$ (the induction is more complicated, but works).

Putting everything together, one obtains lower bounds on the height for points in L that depend only on d.

4.2 Lower bounds on the height III: $\mathbb{Q}(E_{\mathrm{tors}})$

Let E be an elliptic curve over $\mathbb{Q}$ (the field of definition is important). The proof of the fact that $\mathbb{Q}(E_{\mathrm{tors}})$ has property (B) is similar in structure, but more complicated, to the proofs we've seen so far. If E has CM, then $\mathbb{Q}(E_{\mathrm{tors}})$ is an abelian extension of a quadratic extension of $\mathbb{Q}$ and the result follows from what we've already seen.

By Elkies, there are infinitely many supersingular primes; the corresponding Frobenius has the property that the square is central in the Galois group. If there is no (or little) ramification an approach similar to the above works; otherwise, one needs some more ramification theory (Lubin-Tate).

The role of the cyclotomic extensions $\mathbb{Q}(\zeta_n)$ is now played by the torsion extensions $\mathbb{Q}(E[N])$.

4.3 Kronecker's theorem in higher dimensions

Lang's problem. Characterise irreducible polynomials $f \in \overline{\mathbb{Q}}[x, y]$ such that there exist infinitely many pairs $(\alpha, \beta) \in \mu_\infty \times \mu_\infty$ with $f(\alpha, \beta) = 0$. Geometrically, characterise curves in $\mathbb{G}_m^2$ that contain infinitely many torsion points.

Example 4.7. The polynomials $f(x, y) = x^m y^n - 1$ (with $(m, n) = 1$) provide examples of this behaviour. More generally, 1 can be replaced by a root of unity.

Theorem 4.8 (Liardet)**.** *f has this property if and only if it is of the form $f(x, y) = x^n y^m - a$ or $x^n - a y^m$ for some root of unity a.*

Remark 4.9. More geometrically: the curves in question are the translates of subgroups by torsion points. Geometrically one would also allow Laurent polynomials, in which case $x^n y^m - a$ and $x^n - a y^m$ are essentially the same.

Proof. Let K be a number field containing the coefficients of f. Let $(\alpha, \beta) \in \mu_\infty^2$ be a zero of f; write $(\alpha, \beta) = (\zeta_N^r, \zeta_N^s)$ with $\gcd(r, s, N) = 1$. We look for an upper bound on N; notice that we have $K(\alpha, \beta) = K(\zeta_N)$. Consider the group homomorphism

$$\begin{array}{rccc} \lambda: & \mathbb{Z}^2 & \to & \mathbb{C}^\times \\ & (t, u) & \mapsto & \alpha^t \beta^u = \zeta_N^{rt+su}; \end{array}$$

its kernel is $\Lambda = \{(t, u) \in \mathbb{Z}^2 : rt + su \equiv 0 \pmod{N}\}$. By the pigeonhole principle, there exist $(a_1, b_1), (a_2, b_2)$ with $0 \leq a_i, b_j \leq \sqrt{N}$ such that $(a_1, b_1) \neq (a_2, b_2)$ and $(a_1 - a_2, b_1 - b_2) \in \Lambda \setminus \{0\}$.

Interpolation: construction of an auxiliary function. Let $g(x, y) = x^t y^u - 1 \in \mathbb{Q}[x^{\pm 1}, y^{\pm 1}]$: it vanishes at (α, β). If we want to work with polynomials, we can just multiply by $x^{\max\{-t,0\}} y^{\max\{-u,0\}}$ to obtain $g_1(x, y) = x^{\max\{-t,0\}} y^{\max\{-u,0\}} g(x, y) \in \mathbb{Z}[x, y]$.

Extrapolation: for arithmetical reasons, the auxiliary function must vanish at many more points than just those we imposed. Indeed, for every $\sigma \in \mathrm{Gal}(K(\zeta_n)/K) = \mathrm{Gal}(K(\alpha, \beta)/K)$ the point $(\sigma\alpha, \sigma\beta)$ is a common zero of f and g_1. Therefore

$$\#\{f = 0\} \cap \{g_1 = 0\} \geq [K(\zeta_n) : K] \geq \frac{\varphi(N)}{[K : \mathbb{Q}]} \gg \frac{N}{d \log\log N},$$

where $d = [K : \mathbb{Q}]$.

Zeroes lemma. If f divides $g_1(x, y)$ we are done, because all factors of $g_1(x, y)$ are of the desired form. Otherwise f, g_1 are relatively prime (recall that f is irreducible), and by Bézout they meet in $(\deg f)(\deg g_1) \leq 2\sqrt{N} \deg(f)$ points.

Conclusion. Putting the inequalities together we obtain

$$\frac{N}{d \log\log N} \ll 2\sqrt{N} \deg(f) \Rightarrow N \text{ is bounded}$$

Notice that the bound on N is completely explicit. □

Definition 4.10. *A curve $\mathcal{C}$ in $\mathbb{G}_m^2$ is a* ***torsion curve*** *if it is the translate of an algebraic subgroup by a torsion point. Notice that $\mathcal{C}$ is not required to be irreducible.*

Remark 4.11. We shall see in the next lecture that the (1-dimensional) algebraic subgroups of $\mathbb{G}_m^2$ are all of the form

$$\{(x, y) \in \mathbb{G}_m^2(\overline{\mathbb{Q}}) : x^n y^m - 1 = 0\}$$

Corollary 4.12. *A curve $\mathcal{C} \subset \mathbb{G}_m^2$ has the property that $\mathcal{C}_{tors} := (\mathbb{G}_m^2)_{tors} \cap \mathcal{C}$ is Zariski-dense in $\mathcal{C}$ if and only if it is a torsion curve.*

5 08.05.2018 – Geometry of $\mathbb{G}_m^N$

Last time we proved:

Theorem 5.1. *Let $f \in \overline{\mathbb{Q}}[x, y]$ be an irreducible polynomial. There exist infinitely many pairs $(\alpha, \beta) \in \mu_\infty \times \mu_\infty$ with $f(\alpha, \beta) = 0$ if and only if $f(x, y) = ax^n y^m - 1$ or $f(x, y) = x^n y^m - 1$ for some root of unity a.*

We consider $\mathbb{G}_m^n := \mathbb{G}_m^n(\overline{\mathbb{Q}}) = (\overline{\mathbb{Q}}^\times)^n$ and its natural embedding in $\mathbb{P}^n$ given by

$$\begin{array}{ccc} \mathbb{G}_m(\overline{\mathbb{Q}})^n & \to & \mathbb{P}^n(\overline{\mathbb{Q}}) \\ \underline{x} & \mapsto & (1 : x_1 : \cdots : x_n) \end{array}$$

There is also another natural compactification given by

$$\begin{array}{ccc} \mathbb{G}_m^n & \hookrightarrow & \mathbb{P}^1 \times \cdots \times \mathbb{P}^1 \\ \underline{x} & \mapsto & (1 : x_1), (1 : x_2), \ldots, (1 : x_n) \end{array}$$

Remark 5.2.
- it's an algebraic group
- given $\underline{\lambda} \in \mathbb{Z}^n$ and $\underline{x} \in \mathbb{G}_m^n$, we set $\underline{x}^{\underline{\lambda}} := x_1^{\lambda_1} \cdots x_n^{\lambda_n}$
- the ring of regular functions on $\mathbb{G}_m^n$ is $\overline{\mathbb{Q}}[x_1^{\pm 1}, \ldots, x_n^{\pm 1}]$
- $\underline{\alpha} \in \mathbb{G}_m^n$ is a **torsion point** if $\exists k \in \mathbb{N}$ such that $\underline{\alpha}^k = 1$, i.e. iff the coordinates of $\underline{\alpha}$ are roots of unity.
- $H \subseteq \mathbb{G}_m^n$ is an algebraic subgroup iff H is a closed subgroup of $\mathbb{G}_m^n$

Definition 5.3. *$H \subseteq \mathbb{G}_m^n$ algebraic subgroup is a **torus** iff it is irreducible.*

5.1 Lattices

Let $\Lambda \subseteq \mathbb{Z}^n$ be a subgroup of finite rank. Then it is torsion-free, and therefore free. It is then a lattice in $\mathbb{R}\Lambda := \Lambda \otimes_{\mathbb{Z}} \mathbb{R}$. The **rank** of Λ is the dimension of the $\mathbb{R}$-vector space $\Lambda \otimes_{\mathbb{Z}} \mathbb{R}$.

The **saturation** of Λ is

$$\overline{\Lambda} = \mathbb{Q}\Lambda \cap \mathbb{Z}^n = \{\lambda \in \mathbb{Z}^n \mid \exists k \in \mathbb{N} k\lambda \in \Lambda\}.$$

Definition 5.4.

$$\rho(\Lambda) := [\overline{\Lambda} : \Lambda]$$

Remark 5.5. If Λ is contained in Λ', $\operatorname{rk}\Lambda = \operatorname{rk}\Lambda'$ and $\rho(\Lambda) = \rho(\Lambda')$, then $\Lambda' = \Lambda$.

Definition 5.6. *Λ is **primitive** if $\Lambda = \overline{\Lambda}$.*

Remark 5.7. Every lattice Λ admits a $\mathbb{Z}$-basis. Let $\underline{a}_1, \ldots, \underline{a}_r$ be a basis: then Λ is primitive if and only if the gcd of the $r \times r$ minors of (a_{ij}) is ± 1, or equivalently that $\underline{a}_1, \ldots, \underline{a}_r$ can be completed to a basis of $\mathbb{Z}^n$.

5.1.1 Connection with the algebraic subgroups of $\mathbb{G}_m^n$

There is a natural map

$$\begin{array}{ccc} \{\text{lattices } \subseteq \mathbb{Z}^n\} & \to & \{\text{algebraic subgroups of } \mathbb{G}_m^n\} \\ \Lambda & \mapsto & H_\Lambda = \{\underline{\alpha} \in \mathbb{G}_m^n \mid \forall \underline{\lambda} \in \Lambda, \underline{\alpha}^{\underline{\lambda}} = 1\} \end{array}$$

Remark 5.8. This map is contravariant: given $\Lambda \subseteq \Lambda'$, we have $H_{\Lambda'} \subseteq H_\Lambda$.

The aim of today's lecture is to show that this map is a bijection.

5.1.2 Morphisms $\mathbb{G}_m^k \to \mathbb{G}_m^n$

Let A be an element of $\mathrm{Mat}_{n \times k}(\mathbb{Z})$. With A we can associate a morphism

$$\varphi_A : \mathbb{G}_m^k \to \mathbb{G}_m^n$$
$$\underline{\alpha} \mapsto \left(\underline{\alpha}^{A_1}, \ldots, \underline{\alpha}^{A_n}\right),$$

where A_j is the j-th row of A. Product of matrices corresponds to composition of morphisms:

$$\varphi_A \circ \varphi_B = \varphi_{AB}$$

We define $\underline{\alpha}^A := \varphi_A(\alpha)$. Notice that $\underline{\alpha}^{AB} = (\underline{\alpha}^B)^A$.

Remark 5.9. Let $A \in \mathrm{Mat}_{n \times n}(\mathbb{Z})$. Then $\varphi_A : \mathbb{G}_m^n \to \mathbb{G}_m^n$ is a finite morphism (ie $|\ker \varphi_A| < \infty$) iff $\det A \neq 0$; more precisely, the order of $\ker \varphi_A$ is $|\det A|$. In particular, φ_A is an isomorphism iff $\det A = \pm 1$.

Theorem 5.10. *$\Lambda \subseteq \mathbb{Z}^n$ of rank r. Then H_Λ is isomorphic to $F \times H_{\overline{\Lambda}}$, where F is finite, of order $[\overline{\Lambda} : \Lambda]$. Moreover, $H_{\overline{\Lambda}} \cong \mathbb{G}_m^{n-r}$ (and is therefore irreducible). In particular, H is the finite union of translates of $H_{\overline{\Lambda}}$ by torsion points.*

Sketch of proof. Let $\underline{a}_1, \dots, \underline{a}_r$ be a basis of $\overline{\Lambda}$ and complete it to a basis $\underline{a}_1, \dots, \underline{a}_n$ of $\mathbb{Z}^n$. The $\underline{a}_i$ define a matrix $A \in \mathrm{GL}_n(\mathbb{Z})$ such that

$$A\overline{\Lambda} = \mathbb{Z}^r \times \{0\}.$$

Therefore

$$\varphi_A(H_{\overline{\Lambda}}) = H_{A\overline{\Lambda}} = \{1\} \times \mathbb{G}_m^{n-r},$$

hence (up to a change of coordinates and a projection) we may assume $r = n, \overline{\Lambda} = \mathbb{Z}^r$ and $H_{\overline{\Lambda}} = \{1\}$. Notice that we have already proven that $H_{\overline{\Lambda}} \cong \mathbb{G}_m^{n-r}$ is irreducible.

By the elementary divisors theorem, there exist unique $\rho_1, \dots, \rho_r \in \mathbb{N}$ with $\rho_r \mid \rho_{r-1} \mid \cdots \mid \rho_1$ and a basis $\underline{a}_1, \dots, \underline{a}_r$ of $\mathbb{Z}^r$ such that $\rho_1\underline{a}_1, \dots, \rho_r\underline{a}_r$ is a basis of Λ. Up to a further change of coordinates we can then assume that $\underline{a}_i = \underline{e}_i$ is the canonical basis and $\Lambda = \langle(\rho_1, 0, \dots, 0), \dots, (0, \dots, 0, \rho_r)\rangle$. But then

$$H_\Lambda = \{\alpha \in \mathbb{G}_m^n : \forall i, \alpha_i^{\rho_i} = 1\} = \ker[\rho_1] \times \cdots \times \ker[\rho_r]$$

that as claimed has order $H_\Lambda = \prod \rho_i = [\mathbb{Z}^r : \Lambda]$. □

Corollary 5.11. *H_Λ is a torus if and only if $H_\Lambda \cong \mathbb{G}_m^{n-r}$ (where $r = \operatorname{rank} \Lambda$), if and only if Λ is primitive.*

Example 5.12. $r = 1, \Lambda = \langle\underline{\lambda}\rangle$, $\underline{\lambda} = (\lambda_1, \dots, \lambda_n)$. Let $d = \mathrm{GCD}(\lambda_1, \dots, \lambda_n)$ and $\underline{\mu} = \frac{1}{d}\underline{\lambda} \in \mathbb{Z}^n$ primitive. We have

$$H = H_\Lambda = \{\underline{\alpha} \in \mathbb{G}_m^n \mid \alpha_1^{\lambda_1} \cdots \alpha_n^{\lambda_n} = 1\}$$

and

$$T = H_{\overline{\Lambda}} = \{\underline{\alpha} \in \mathbb{G}_m^n \mid \alpha_1^{\mu_1} \cdots \alpha_n^{\mu_n} = 1\};$$

one checks that $H = \bigcup_{\underline{\zeta}^d = 1} T\underline{\zeta}$.

5.1.3 Decomposition in $\mathbb{Q}$-irreducible components

$$H = \bigcup_{k|d} \{\underline{\alpha} \in \mathbb{G}_m^n \mid \Phi_k(\underline{\alpha}^\mu) = 0\},$$

where $\underline{x} \mapsto \Phi_k(\underline{x}^{\underline{\mu}})$ is called a *generalized cyclotomic polynomial.*

5.1.4 Every algebraic subgroup of $\mathbb{G}_m^n$ is of the form H_Λ

Theorem 5.13. *Let*

- *$V \subseteq \mathbb{G}_m^n$ be defined by equations $f_l(\underline{x}) = \sum_{\underline{\lambda} \in I} a_{l,\underline{\lambda}} \underline{x}^{\underline{\lambda}} = 0$*
- *H a maximal algebraic subgroup contained in V*

Then there exists a lattice $\Lambda \subseteq \mathbb{Z}^n$, generated by vectors in the set

$$D(I) = \{\underline{\lambda} - \underline{\mu} \mid \underline{\lambda}, \underline{\mu} \in I\},$$

such that $H = H_\Lambda$.

Corollary 5.14. *$\Lambda \mapsto H_\Lambda$ is surjective (onto the set of algebraic subgroups of $\mathbb{G}_m^n$).*

Proof. Let $\underline{\lambda} \in \mathbb{Z}^n$ and let $\varphi_{\underline{\lambda}}$ be the corresponding morphism $\mathbb{G}_m^n \to \mathbb{G}_m$ $(x \mapsto x^{\underline{\lambda}})$. By definition, $\chi_{\underline{\lambda}} := \varphi_{\underline{\lambda}}|_H$ is a character of H. Let us consider the partition of I given by the sets

$$I_\chi := \{\underline{\lambda} \in I \mid \chi_{\underline{\lambda}} = \chi\}$$

for χ ranging over $\hat{H}$ (the character group of H).

Remark 5.15. $\forall \chi \in \hat{H}$, $\forall \underline{\lambda}, \underline{\mu} \in I_\chi$, $\chi_{\underline{\lambda}-\underline{\mu}} = 1$.

Let

$$\Lambda = \left\langle \underline{\lambda} - \underline{\mu} \mid \underline{\lambda}, \underline{\mu} \in I_\chi, \chi \in \hat{H} \right\rangle.$$

By Remark 5.15 we have $H \subseteq H_\Lambda$. By maximality of H, it is now enough to show that H_Λ is contained in V.

We now observe that

$$f_l|_H = \sum_{\chi \in \hat{H}} \left(\sum_{\underline{\lambda} \in I_\chi} a_{l,\underline{\lambda}} \right) \chi \equiv 0,$$

simply because these are the defining equations of $V \supseteq H$. By independence of characters (Artin's theorem), this implies that $\forall \chi \in \hat{H}, \sum_{\underline{\lambda} \in I_\chi} a_{l,\underline{\lambda}} = 0$. Now notice that, given $\underline{\alpha} \in H_{\underline{\lambda}}$, the map

$$\begin{array}{ccc} I & \to & \mathbb{G}_m \\ \underline{\lambda} & \mapsto & \underline{\alpha}^{\underline{\lambda}} \end{array}$$

is constant on I_χ. Therefore for $\underline{\alpha} \in H_\Lambda$ we have

$$f_l(\alpha) = \sum_{\chi \in \hat{H}} \sum_{\underline{\lambda} \in I_\chi} a_{l,\underline{\lambda}} \underline{\alpha}^{\underline{\lambda}} = \sum_{\chi \in \hat{H}} \left(\sum_{\underline{\lambda} \in I_\chi} a_{l,\underline{\lambda}} \right) \underline{\alpha}^{\underline{\lambda}} = 0,$$

so $H_\Lambda \subseteq V$ as desired. □

Corollary 5.16. *Every morphism $\phi : \mathbb{G}_m^n \to \mathbb{G}_m^r$ is of the form φ_A for a suitable matrix A.*

Sketch of proof. One is immediately reduced to the case $r = 1$. Consider the graph of φ,

$$\Gamma = \left\{ (\underline{x}, \varphi(\underline{x})) \mid \underline{x} \in \mathbb{G}_m^n \right\}$$

Γ is an n-dimensional torus, and therefore $\Gamma = H_{\underline{\lambda}}$. Hence we that $y = \varphi(\underline{x})$ iff $x_1^{\lambda_1} \cdots x_n^{\lambda_n} y^{\lambda_{n+1}} = 1$. As φ is a morphism, λ_{n+1} is forced to be ± 1, and we are done. □

Theorem 5.17. *1. $\Lambda \mapsto H_\Lambda$ is bijective*

2. $H_\Lambda H_{\Lambda'} = H_{\Lambda \cap \Lambda'}$ and $H_\Lambda \cap H_{\Lambda'} = H_{\Lambda + \Lambda'}$.

Proof.

1. We've already shown surjectivity, so it remains to show that injectivity. Suppose we have two lattices Λ, Λ' which correspond to the same $H = H_\Lambda = H_{\Lambda'}$. Then $\forall x \in H, \forall \underline{\lambda} \in \Lambda, \forall \underline{\lambda}' \in \Lambda'$ we have $\underline{x}^{\underline{\lambda}} = \underline{x}^{\underline{\lambda}'} = 1$. Hence $H \subseteq H_{\Lambda+\Lambda'}$, and clearly $H_{\Lambda+\Lambda'} \subseteq H_\Lambda = H$. It follows that $H_\Lambda = H = H_{\Lambda+\Lambda'}$. By a previous theorem, $\mathrm{rk}(\Lambda+\Lambda') = \mathrm{rk}(\Lambda)$, and furthermore $\rho(\Lambda + \Lambda') = \rho(\Lambda)$. This implies $\Lambda = \Lambda + \Lambda'$, hence $\Lambda' \subseteq \Lambda$. By symmetry, $\Lambda = \Lambda'$.

2. $H_\Lambda H_{\Lambda'}$ is an algebraic subgroup of $\mathbb{G}_m^n$ (proof/exercise: a dominant morphism $\varphi : \mathbb{G}_m^n \to \mathbb{G}_m^r$ is surjective). In particular, it is the smallest algebraic subgroup containing both H_Λ and $H_{\Lambda'}$. Now $\Lambda \cap \Lambda'$ is the largest lattice contained in both Λ and Λ'; since $\Lambda \mapsto H_\Lambda$ reverses the inclusions, this implies that $H_{\Lambda \cap \Lambda'}$ is the smallest algebraic subgroup containing H_Λ and $H_{\Lambda'}$, whence $H_\Lambda H_{\Lambda'} = H_{\Lambda \cap \Lambda'}$.

   Similarly, $H_\Lambda \cap H_{\Lambda'}$ is the largest algebraic subgroup contained in both H_Λ and $H_{\Lambda'}$, and one deduces $H_\Lambda \cap H_{\Lambda'} = H_{\Lambda + \Lambda'}$.

□

Definition 5.18. *$V \subseteq \mathbb{G}_m^n$ is a* ***torsion subvariety*** *if V is of the form $H\underline{\zeta}$, where H is a torus and $\underline{\zeta} \in (\mathbb{G}_m^n)_{tors}$.*

Theorem 5.19 (Toric case of the Manin-Mumford conjecture)**.** *THe following hold:*

- *$V \subseteq \mathbb{G}_m^n$ irreducible is a torsion subvariety if and only if the Zariski closure of $\overline{V_{\text{tors}}}$ is equal to V.*
- *$V \subset \mathbb{G}_m^n$, not necessarily irreducible. Then $\overline{V_{\text{tors}}}$ is a finite union of torsion subvarieties; equivalently, it is the (finite) union of the maximal torsion subvarieties contained in V.*

Remark 5.20. Notice that part (i) in the case of curves is Liardet's theorem.

Proof (of (1) $\to$ (2)*).* Write $\overline{V_{\text{tors}}} = V_1 \cap \cdots \cap V_n$ for the decomposition in irreducible components. Then

$$\overline{V_{\text{tors}}} = \overline{\left(\overline{V_{\text{tors}}}\right)_{\text{tors}}} = \overline{(V_1)_{\text{tors}}} \cup \cdots \cup \overline{(V_n)_{\text{tors}}}$$

By uniqueness of the decomposition in irreducible components, $\overline{(V_i)_{\text{tors}}} = V_i$, hence V_i is a torsion subvariety. □

5.2 References

- Bombieri-Gubler. *Heights in diophantine geometry*, chapter 3
- Zannier. *Lecture notes on diophantine analysis*

6 09.05.2018 – Heights and Lenstra's algorithm in more than one variable

6.1 Kronecker in dimension > 1?

There are (at least) two natural notions of height on $\mathbb{G}_m^n$, corresponding to the two compactifications $\mathbb{P}_1^n$ and $\mathbb{P}_n$:

- $\hat{h}_1(\underline{\alpha}) = \sum_{i=1}^n h(\alpha_i)$, if we consider $\mathbb{G}_m^n \hookrightarrow \mathbb{P}_1^n$;
- $\hat{h}_2(\underline{\alpha}) = h((1 : \alpha_1 : \cdots : \alpha_n))$, if we consider $\mathbb{G}_m^n \hookrightarrow \mathbb{P}_n$.

Definition 6.1. *Recall the usual Weil height on $\mathbb{P}_n$: given a point $\beta = (\beta_0 : \cdots : \beta_n) \in \mathbb{P}_n(\overline{K})$, take a number field K that contains all the β_i and define*

$$h(\beta) = \frac{1}{[K : \mathbb{Q}]} \sum_{v \in M_K} [K_v : \mathbb{Q}_v] \log \max\{|\beta_0|_v, \ldots, |\beta_n|_v\}.$$

It depends neither on the choice of K nor on the choice of the representatives for the coordinates of β.

Remark 6.2.

$$\hat{h}_1 \leq \hat{h}_2 \leq n\hat{h}_1$$

Remark 6.3. The following properties hold:

1. $\hat{h}(\underline{\alpha}) = 0$ if and only if $\underline{\alpha}$ is torsion
2. $\hat{h}(\underline{\alpha}\underline{\beta}) \leq \hat{h}(\underline{\alpha}) + \hat{h}(\underline{\beta})$
3. $\hat{h}(\alpha^n) = n\hat{h}(\alpha)$ for $n \in \mathbb{N}$, and also for $n \in \mathbb{Z}$ if $\mathbb{G}_m^n \hookrightarrow \mathbb{P}_1^n$.
4. Northcott: the set of points of bounded height and degree is finite.

Definition 6.4. *Let $V \subseteq \mathbb{G}_m^n$ and $\vartheta > 0$. Set $V(\vartheta) = \{\underline{\alpha} \in V \mid \hat{h}(\underline{\alpha}) \leq \vartheta\}$. The* ***essential minimum*** *of a subvariety is*

$$\mu_{\text{ess}}(V) = \inf\{\vartheta > 0 \mid \overline{V(\vartheta)} = V\}$$

Theorem 6.5 (Toric case of Bogomolov's conjecture, Zhang/Bilu)**.** *$V \subseteq \mathbb{G}_m^n$, V/K (K a number field). Suppose that V is K-irreducible. Then the following are equivalent:*

1. *V is a union of torsion subvarieties*
2. *$\mu_{\text{ess}}(V) = 0$*

Remark 6.6. The implication $1 \Rightarrow 2$ is obvious.

More precisely:

Theorem 6.7 (Schmidt, Bombieri-Zannier)**.** *If V is not a union of torsion subvarieties, then*

$$\mu_{\text{ess}} \geq C([K:\mathbb{Q}], \deg V) > 0$$

for some function $C(d,m)$.

Theorem 6.8 (Bombieri-Zannier)**.** *In fact, $\mu_{\text{ess}}(V) \geq C(\deg V) > 0$ provided that V is not a union of translates of tori (translation by arbitrary points of $\mathbb{G}_m^n$).*

Remark 6.9. To see that the exception is necessary consider the case of a single point (the height can go to zero as $[K:\mathbb{Q}] \to \infty$).

6.2 Lenstra in more than one variable

We describe an approach due to Averdaro-Krick-Sombra. The problem is the following: given $f \in \overline{\mathbb{Q}}[x,y]$, suppose that $\deg_y f > 0$. What is the analogue of the gap principle we saw for the case of univariate polynomials?

Write $f = r + y^u q$ with $q, r \in \overline{\mathbb{Q}}[x,y]$. Let $k = \deg_y r$. Assume that $u - k$ is *large* (in a sense to be specified later). Let now $P \in K[x,y]$ (K a number field) be an irreducible polynomial such that $\deg_y P > 0$ and the curve $C = \{P = 0\}$ is *not* a union of torsion curves.

Suppose (by contradiction[2]) that $p \mid f$ and $p \nmid r$ (hence $p \nmid q$); we want to show that the gap is small. By Bogomolov, the essential minimum of C is positive. Take $h_0 \in (0, \mu_{\text{ess}}(C))$: then

$$\big\{(\omega,\xi) \in C(\overline{\mathbb{Q}}) : \omega \in \mu_\infty, h(\xi) \leq h_0\big\}$$

is **finite** (for otherwise it would be Zariski-dense[3]). On the other hand,

$$\big\{(\omega,\xi) \in C \mid \omega \in \mu_\infty\big\}$$

is infinite, and moreover $C \cap \{q = 0\}$ is finite ($p \nmid q$). Therefore there exists $\xi \in \overline{\mathbb{Q}}^\times, h(\xi) > h_0$ and a root of unity ω such that $P(\omega,\xi) = 0$ and $q(\omega,\xi) \neq 0$.

Definition 6.10. *$g \in \overline{\mathbb{Q}}[x,y]$, $g = \sum g_{ij} x^i y^j$. Define*

$$|g|_v := \begin{cases} \|\sigma g\|_1, & \text{if } \sigma : K \hookrightarrow \mathbb{C} \\ \max |g_{ij}|_v, & \text{if } v \nmid \infty \end{cases}$$

Finally set

$$h(g) = \sum_{v \in M_K} [K_v : \mathbb{Q}_v] \log |g|_v$$

[2] well, morally by contradiction

[3] invariance of the Weil height under the action of Galois shows that the essential minimum is the same on all irreducible components, and we have assumed that C is K-irreducible, hence the $\overline{\mathbb{Q}}$-irreducible components are permuted transitively by Galois

Remark 6.11. $h(g)$ is independent of the choice of the number field K containing the coefficients of g and is invariant under multiplication by scalars in $\overline{\mathbb{Q}}^\times$.

Let now $\alpha = (\omega, \xi)$, $z := \xi^{-u} r(\alpha) = -q(\alpha) \neq 0$. We have

$$|\xi^{-u} r(\alpha)|_v \leq |f|_v |\xi|_v^{-u} \max\{1, |\xi|_v\}^k$$

and

$$|q(\alpha)|_v \leq |f|_v \max\{1, |\xi|_v\}^{\deg q};$$

this implies $|z|_v \leq |f|_v \max\{1, |\xi|_v\}^{-(u-k)}$, and by using the product formula we obtain

$$1 \leq H(f) H(\xi)^{-(u-k)},$$

hence

$$u - k \leq \frac{h(f)}{\mu_{\mathrm{ess}}(C)} \leq \begin{cases} h(f)/C([K:\mathbb{Q}], \deg P) \\ h(f)/C'(\deg P) \text{ if } C \text{ is not a union of translates} \end{cases}$$

6.3 Heights of subvarieties

Reference: appendix to Zannier's book *Lectures on diophantine analysis*

Let V be a $\overline{\mathbb{Q}}$-irreducible variety.

Fact. Consider the morphism

$$\begin{array}{cccc} [\ell]: & \mathbb{G}_m^n & \to & \mathbb{G}_m^n \\ & \underline{\alpha} & \mapsto & \left(\alpha_1^\ell, \cdots, \alpha_n^\ell\right) \end{array}$$

Then $\deg([\ell]^{-1} V) = \ell^{\operatorname{codim} V} \deg V$.

Example 6.12. If V is the hypersurface defined by $V = \{F = 0\}$, then $[\ell]^{-1} V = \{F(x_1^\ell, \ldots, x_n^\ell) = 0\}$ has degree $\ell \deg F = \ell \deg V$.

On the other hand,

$$\deg([\ell]V) = \frac{\ell^{\dim V} \deg V}{|\ker[\ell] \cap \operatorname{Stab}(V)|},$$

where $\operatorname{Stab}(V) = \{\alpha \in \mathbb{G}_m^n : \alpha V = V\} = \bigcap_{\underline{x} \in V} \underline{x}^{-1} V$. Let $h : \mathbb{P}_n \to \mathbb{R}_{\geq 0}$ be an (almost) arbitrary height. Then one can define a height *à la Néron-Tate*, namely

$$\hat{h}(V) = \lim_{\ell \to \infty} \frac{h([\ell]V) \deg V}{\ell \deg([\ell]V)}.$$

With this definition, one has

$$\hat{h}([\ell]^{-1} V) = \ell^{\operatorname{codim} V - 1} \hat{h}(V)$$

and

$$\hat{h}([\ell]V) = \frac{\ell^{\dim V + 1} \hat{h}(V)}{|\ker[\ell] \cap \operatorname{Stab}(V)|}$$

Remark 6.13. The height depends on the choice of compactification of $\mathbb{G}_m^n$. If we work with $\mathbb{P}_n$ and if $V = \{F = 0\}$, then

$$\hat{h}(V) = \log \mathcal{M}(F) = \int_0^1 \cdots \int_0^1 \log |F(e^{2\pi i \theta_1, \ldots, 2\pi i \theta_n})| \, d\theta_1 \cdots d\theta_n$$

Theorem 6.14 (Zhang). $\mu_{\mathrm{ess}}(V) \leq \frac{\hat{h}(V)}{\deg V} \leq (1 + \dim V) \mu_{\mathrm{ess}}(V)$

6.3.1 Quantitative results

Let $V/\mathbb{Q}$ be an irreducible subvariety of $\mathbb{G}_m^n$. Suppose that V is not contained in an algebraic subgroup of $\mathbb{G}_m^n$. Then for all $\varepsilon > 0$ there exists $C_\varepsilon(n)$ such that

$$\mu_{\text{ess}}(V) \geq C_\varepsilon(n)\omega(V)^{-1-\varepsilon},$$

where $C_\varepsilon(n) > 0$ and

$$\omega(V) = \min\{\deg(Z) : Z \text{ hypersurface defined over } \mathbb{Q} \text{ containing } V\}.$$

This is a result of Amoroso-David.

Example 6.15. $\underline{\alpha} = (\alpha_1, \dots, \alpha_n) \in \overline{\mathbb{Q}}^n$ multiplicatively independent. Let $V = \{\sigma\underline{\alpha} \mid \sigma : \overline{\mathbb{Q}} \hookrightarrow \mathbb{C}\}$. Then $\forall\varepsilon, h(\underline{\alpha}) \geq C_\varepsilon(n)\omega^{-1-\varepsilon}$. Moreover, by linear algebra one obtains

$$\omega = \omega(V) \leq n[K : \mathbb{Q}]^{1/n},$$

where $K = \mathbb{Q}(\alpha_1, \dots, \alpha_n)$.

Definition 6.16. *Let*

$$V^u = \textit{(finite) union of the maximal torsion subvarieties of } V, \quad V^* = V \setminus V^u,$$

and

$$\delta(V) = \{\delta : V \textit{ is the intersection of hypersurfaces of degree at most } \delta \textit{ defined over } \mathbb{Q}\}.$$

Then:

Theorem 6.17 (Amoroso-Viada 2013)**.** *$\forall\underline{\alpha} \in V^*$ one has*

$$h(\underline{\alpha}) \geq \delta(V)^{-1}(935n^5 \log(n^2\delta(V)))^{-(n+1)^2 \dim V}.$$

Definition 6.18. *We denote by V^a the (not necessarily finite) union of the positive-dimension translates contained in V.*

Example 6.19. To see that the union is not necessarily finite, consider

$$V = \{x_1 + x_2 + x_3 = 0\} \subseteq \mathbb{G}_m^3, \quad H = \{(t, t, t) \mid t \in \mathbb{G}_m\}$$

Then $\forall\underline{\alpha} \in V$, $H\underline{\alpha} \subseteq V$.

Remark 6.20. Even though it's not obvious, V^a is closed.

Set $V^0 = V \setminus V^a$.

Theorem 6.21 (Bombieri-Zannier)**.** *For all but finitely many $\underline{\alpha} \in V^0$ the inequality $h(\underline{\alpha}) \geq C(\deg V) > 0$ holds.*

Let now $V \subseteq \mathbb{G}_m^n$ be defined over K and K-irreducible. Suppose that V is *not* contained in a finite union of translates. Analogously to the case of varieties defined over $\mathbb{Q}$, set

$$\delta(V) = \min\{\delta : V \text{ is intersection of hypersurfaces of degree } \leq \delta\}$$

and

$$\theta = \theta(V) = \delta(V)(200n^5 \log(n^2\delta(V)))^{n(n-1)\deg V}$$

Theorem 6.22 (Amoroso-Viada '09)**.**

$$\overline{V(\theta^{-1})} = B_1 \cup \ldots \cup B_k$$

where every B_i is a translate and $\sum_{i=1}^k \theta^{\dim B_i} \deg B_i \leq \theta^n$.

Corollary 6.23. *The maximal torsion subvarieties B_j of V satisfy*

$$\sum_{i=1}^{k} \theta^{\dim B_i} B_i \leq \theta^n.$$

Example 6.24. $n = 2, V = \{F(x,y) = 0\}$ $\overline{\mathbb{Q}}$-irreducible and not torsion. Then the number of torsion points on V satisfies

$$\# V_{\text{tors}} \leq C \deg(F)^2$$

and in fact

$$\# V_{\text{tors}} \leq \operatorname{vol}(\Delta_F),$$

where Δ_F is the Newton polytope of F. These are results of Ruppert and Aliev-Smyth, generalized to hypersurfaces of $\mathbb{G}_m^n$ by César Martinez.

7 29.05.2018

Recall the following notions:

1. Weil height in $\mathbb{G}_m^n \hookrightarrow \mathbb{P}_n$: see sections 2.6 and 5.
2. height of a polynomial
3. Mahler measure: see definition 2.22
4. normalised height of a hypersurface, or more generally of a subvariety:
$$\hat{h}(V) = \lim_{\ell \to \infty} \frac{h([\ell]V) \deg V}{\ell \deg([\ell]V)}.$$
5. non-normalised heights of varieties. Let $V \subseteq \mathbb{P}_n$ be a projective variety of dimension d, and let F be the Chow form (see section 13.1) of V. Then one can define $h(V)$ as the height of the hypersurface $\{F = 0\}$ of $\mathbb{G}_m^{(d+1)n}$.

We begin with two exercises:

Exercise 7.1. *Let $P = \sum_{\underline{\lambda}} c_{\underline{\lambda}} x^{\underline{\lambda}}$ be a polynomial in n variables $x_1, \dots, x_n$.*

1. *For $n = 1$ show that*
$$M(P) \leq \|P\|_1 \leq 2^{\deg P} M(P).$$
2. *Show by induction on $n > 1$ that in general we have*
$$M(P) \leq \|P\|_1 \leq 2^{\sum_j \deg_{x_j} P} M(P) \leq 2^{n \deg P} M(P)$$
3. *Show that in fact $\|P\|_1 \leq (n+1)^{\deg P} M(P)$ by using the inequality*
$$|c_{\underline{\lambda}}| \leq \frac{(\deg P)!}{\lambda_1! \cdots \lambda_n!} M(P)$$

Exercise 7.2. *Let $\varphi_A : \mathbb{G}_m^n \to \mathbb{G}_m^n$ be an isogeny. Then $\hat{h}(\varphi_A^{-1}(V)) = \hat{h}(V)$.*
Hint. *Elementary divisors.*

Proposition 7.3. *Let V be a subvariety of $\mathbb{G}_m^n$. The following hold:*

1.
$$\deg([\ell]^{-1} V) = \ell^{\operatorname{codim}(V)} \deg(V),$$

2.
$$\hat{h}([\ell]^{-1}V) = \ell^{\mathrm{codim}(V)-1}\hat{h}(V),$$

3.
$$\deg([\ell]V) = \frac{\ell^{\dim V}\deg V}{|\ker[\ell]\cap \mathrm{Stab}(V)|},$$

4.
$$\hat{h}([\ell]V) = \frac{\ell^{\dim V+1}\hat{h}(V)}{|\ker[\ell]\cap \mathrm{Stab}(V)|}.$$

We sketch how to prove these properties in the special case $\mathrm{codim}(V)=1$; write V as the zero locus of some polynomial F.

Proof. For the first part, notice that $[\ell]^{-1}(V)$ is defined by $F(x_1^\ell,\cdots,x_n^\ell)=0$, so it clearly has degree equal to $\ell\deg(V)$.

For the rest, one combines the following:

- $[\ell]^{-1}[\ell](V) = \bigcup_{\zeta\in\ker[\ell]}\zeta V$
- $M(P(x_1^\ell,\dots,x_n^\ell)) = M(P)$
- $\deg(P(x_1^\ell,\dots,x_n^\ell)) = \ell\deg P$
- The class formula for the action of $\ker[\ell]$ on $\{\xi V : \xi\in\ker[\ell]\}$

□

Corollary 7.4. *The normalised height satisfies*
$$\hat{h}(V) = \frac{\hat{h}([\ell]V)\deg(V)}{\ell\deg([\ell]V)}$$

7.1 Essential minimum

Recall (definitions 1.11 and 6.4) that we defined the essential minimum of a subvariety V as
$$\mu_{\mathrm{ess}}(V) = \inf\{\vartheta>0 \mid \overline{V(\vartheta)} = V\},$$
where $V(\vartheta) = \{\underline{\alpha}\in V \mid \hat{h}(\underline{\alpha})\le\vartheta\}$.

> **Remark 7.5.** The following properties of the essential minimum are straightforward to check:

- $\mu_{\mathrm{ess}}([\ell]V) = \ell\mu_{\mathrm{ess}}(V)$
- $\mu_{\mathrm{ess}}(V^\sigma) = \mu_{\mathrm{ess}}(V)$ for all $\sigma\in\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$

We had also stated Zhang's inequality (Theorem 6.14):
$$\frac{\hat{h}(V)}{(1+\dim V)\deg(V)} \le \mu_{\mathrm{ess}}(V) \le \frac{\hat{h}(V)}{\deg V}.$$

7.1.1 Siegel's lemma

Let $V \subseteq \mathbb{G}_m^n$ be a $\mathbb{Q}$-irreducible variety, and let $\mathcal{P} \subseteq \mathbb{Q}[x_0, \ldots, x_n]$ be the defining ideal of V in $\mathbb{P}_n$. Define

$$r := H\left(\mathcal{P}^T; L\right) = \dim_{\mathbb{Q}} \left(\frac{\mathbb{Q}[x_0, \ldots, x_n]}{[\mathcal{P}^T]_L} \right).$$

Suppose $r < N := \binom{L+n}{n}$ and set $k = \frac{r}{N-r}$ (Dirichlet's exponent). Then there exists $F \in \mathbb{Q}[\underline{x}]$, of degree at most L, vanishing on V with multiplicity at least T and such that

$$h(F) \leq k\left((T+n)\log(L+1) + L\mu_{\mathrm{ess}}(V)\right).$$

Remark 7.6. There is also an *absolute* version of this lemma, where one only requires V to be defined over $\overline{\mathbb{Q}}$, but loses all control on the field of definition of F (even if V is defined over a number field of controlled degree, say).

Remark 7.7. In the special case when V is given by $\{F = 0\}$ with $\deg F = D$, one has

$$k \leq n\left(1 + \frac{DT}{L - DT}\right)^n \frac{DT}{L - DT};$$

if furthermore $L \geq (n+1)DT$, then $k \leq 3nDT/L$. Indeed, in this case one can compute k exactly:

$$k = \frac{\binom{L+n}{n} - \binom{L-DT+n}{n}}{\binom{L-DT+n}{n}}.$$

7.1.2 Proof of Zhang's inequality

Let V be the hypersurface given by $\{P = 0\}$. We may assume $P \notin \mathbb{Z}[\underline{x}]$ irreducible of degree D.

Lower bound on $\mu_{\mathrm{ess}}(V)$. By Siegel with $T = 1$, there exists $F_L \in \mathbb{Z}[\underline{x}]$ of degree at most L (with $P \mid F_L$, since F_L vanishes on V) such that

$$h(F_L) \leq kL\left(\frac{(n+1)\log(L+1)}{L} + \mu_{\mathrm{ess}}(V)\right),$$

where $kL \leq n\left(1 + \frac{D}{L-D}\right)\frac{DL}{L-D} \xrightarrow{L\to\infty} nD$, and where the term $\frac{(n+1)\log(L+1)}{L}$ goes to 0 as $L \to \infty$. Hence

$$\hat{h}(V) \leq \hat{h}(F_L) \leq h(F_L) \lesssim nD\mu_{\mathrm{ess}}(V)$$

Upper bound on $\mu_{\mathrm{ess}}(V)$. Let's assume for simplicity that $n = 2$.

Lemma 7.8. *Let $P \in \mathbb{C}[x]$ be a polynomial of degree D and let p be a prime number. Then*

$$\prod_{\substack{\omega^p = 1 \\ \omega \neq 1}} |P(\omega)| \leq p^D M(P)^{p-1}.$$

Proof. Left as an exercise for the reader. □

Let now $P \in \mathbb{C}[x, y]$. Given $\omega \in \mathbb{C}$, let $M(P(\omega, y))$ be the Mahler measure of the polynomial $y \mapsto P(\omega, y)$. By the previous lemma, for any prime p we have

$$\frac{1}{p-1}\sum_{\substack{\omega^p = 1 \\ \omega \neq 1}} \log|P(\omega, y)| \leq \log M(P(\omega, y)) + D_y \frac{\log p}{p-1},$$

where D_y is the degree of P in y.

It follows that, for p large and ω primitive p-th root of unity, we have

$$\min\{h(\alpha) \mid P(\omega,\alpha)=0\} \leq \frac{1}{p-1}\sum_{\alpha: P(\omega,\alpha)} h(\alpha) \leq \frac{\hat{h}(V)}{D_y}+\varepsilon.$$

Since the points (ω,α) thus obtained are an infinite set, this, implies that $\mu_{\mathrm{ess}}(V) \leq \frac{\hat{V}}{D_y}$. To obtain the upper bound

$$\mu_{\mathrm{ess}}(V) \leq \frac{\hat{h}(V)}{\deg V}$$

one may consider the hypersurface defined by $Q = P(xy, y)$: then $\deg_y Q = \deg P$, and $\hat{h}(V') = \hat{h}(V)$ by exercise 7.2.

8 30.05.2018 – Lower bounds for the essential minimum

Theorem 8.1. *Let $V \subset \mathbb{G}_m^2$ be a curve. If V is $\mathbb{Q}$-irreducible and is not a union of torsion cosets*[4]*, then*

$$\mu_{\mathrm{ess}}(V) \geq c(\deg V) > 0$$

Remark 8.2. One can replace the hypothesis '$\mathbb{Q}$-irreducible' by 'K-irreducible', at the cost of having the constant depend on the field.

If we want to avoid the dependence on the field of definition, we have the following result:

Theorem 8.3. *Suppose V is $\overline{\mathbb{Q}}$-irreducible and that V is not a torsion coset: then $\mu_{\mathrm{ess}}(V) \geq c(\deg V) > 0$.*

Proof of Theorem 8.1. The proof is analogous to that of Dobrowolski's theorem. In most questions of height, the crucial point is a metric inequality which is then inserted into the product formula. In this case, the inequality is that of Dobrowolski: given $\underline{\alpha} \in \mathbb{G}_m^n(\overline{\mathbb{Q}})$, let

$$F[x_1,\ldots,x_n], \deg(F) \leq L$$

be a polynomial that vanishes at $\underline{\alpha}$ with multiplicity at least T. Then for every prime p and for every place v dividing p we have

$$(\star) \quad |F(\underline{\alpha}^p)|_v \leq p^{-T}\max\{1, |\alpha_1|_v, \ldots, |\alpha_n|_v\}^{pL}$$

This is not obvious; it's easy for $T=1$ (Fermat's little theorem) but not for $T>1$. We won't give the details.

Now the assumption that V is not a union of torsion cosets implies that

$$\frac{1}{p}\mu_{\mathrm{ess}}(V) = \mu_{\mathrm{ess}}([p]^{-1}V) \neq \mu_{\mathrm{ess}}(V)$$

, for otherwise $\mu_{\mathrm{ess}}(V)$ would be zero. Set $D = \deg V$. In particular we can find $\underline{\alpha} \in V(\overline{\mathbb{Q}})$ such that

1. $h(\underline{\alpha}) \leq \mu_{\mathrm{ess}}(V) + \varepsilon$
2. $\underline{\alpha} \in V, \underline{\alpha}^p \notin V$

[4] translates of subgroups by torsion points

Applying ($\star$) by taking as F the minimal equation over $\mathbb{Z}$ of V and $T = 1$ and plugging the result into the product formula, we obtain

$$\mu_{\text{ess}}(V) + \varepsilon \geq h(\alpha) \geq \frac{\log p - h(F)}{pD}$$

Fixing F and choosing p to be so that the last fraction is positive we obtain some sort of lower bound.

Recall from yesterday that

$$h(F) \leq \hat{h}(F) + D \log 2 \leq D(\log 2 + \mu_{\text{ess}}(V))$$

where the second inequality is Zhang's theorem. All in all, one can conclude (notice that we can assume $\mu_{\text{ess}}(V)$ to be very small, for otherwise we are already done!) the proof, but the bound is quite weak: $C(D) \approx (De^D)^{-1}$. □

We describe a technique to obtain better bounds for the case of geometrically irreducible varieties:

Proof of Theorem 8.3. Let $F \in \overline{\mathbb{Q}}[x_1, \ldots, x_n]$, $\deg(L) \leq L$, vanishing at $\underline{\alpha}$ with multiplicity at least T. For every prime p, $\forall \zeta \in \ker[p]$, $\forall v \mid p$ one has

$$|F(\underline{\zeta\alpha})|_v \leq p^{-\frac{T}{p-1}} |F|_v \max\{1, |\alpha|_v\}^L \tag{1}$$

Proof of (1). Taylor's formula + $|\ell^{2\pi i/p} - 1|_v \leq p^{\frac{1}{p-1}}$. □

Let's try applying this inequality naïvely, without constructing any clever auxiliary functions. The fact that V is not a torsion coset implies that the stabiliser of V is a finite group. For simplicity, assume $\text{Stab}(V) = \{1\}$ (otherwise the proof is more technical). As before, there exists $\underline{\alpha} \in V(\overline{\mathbb{Q}})$ such that

1. $h(\underline{\alpha}) \leq \mu_{\text{ess}}(V) + \varepsilon$
2. $\underline{\alpha} \in V, \underline{\zeta\alpha}^p \notin V$ for some $\underline{\zeta} \in \ker[p]$ (in fact, even for *every* nontrivial $\zeta \in \ker[p] \setminus \{0\}$).

Reasoning as above, we obtain

$$\mu_{\text{ess}}(V) + \varepsilon \geq h(\alpha) \geq \frac{\frac{1}{p-1}\log p - h(F)}{D},$$

which gives **nothing**: for large p, this quantity is negative.

We now show that constructing a good auxiliary function leads (in both cases: both for varieties irreducible over $\mathbb{Q}$ and over $\overline{\mathbb{Q}}$) to an essentially optimal bound $C(D) = C_\varepsilon D^{-1-\varepsilon}$. Recall **Siegel's lemma**:

Lemma 8.4. *Let $V \subseteq \mathbb{G}_m^n$ be a curve of degree D, defined over $\mathbb{Q}$ and $\mathbb{Q}$-irreducible (respectively defined and irreducibile over $\overline{\mathbb{Q}}$). Let L, T be two parameters, with $L \geq (n+1)DT$. Then there exists a polynomial $F \in \mathbb{Q}[x_1, \ldots, x_n]$ (respectively in $\overline{\mathbb{Q}}[x_1, \ldots, x_n]$) such that F vanishes on V at order at least T, and*

$$h(F) \leq k((T+n)\log(L+1) + L\hat{\mu}_{ess}(V)),$$

where k is Dirichlet's exponent, for which we have

$$k \leq n\left(1 + \frac{DT}{L - DT}\right)\frac{DT}{L - DT} \leq \frac{3nDT}{L}.$$

Remark 8.5. The version in parentheses is called the **absolute Siegel's lemma**. If V is defined over the field K, one can ask that F be also defined over K: in this case, however, the upper bound on the height depends on the discriminant of K as well as on the other data. This is not good for our applications, since we do not want a dependence on K in the constant $C(D)$.

We now discuss both the case of fixed field of definition and the geometric case.

1. Case 1: $V/\mathbb{Q}$ is $\mathbb{Q}$-irreducible. Let N be a parameter and $p \in [N/2, N]$ be a prime.

   Extrapolation. We show that F (given by Siegel's lemma) vanishes on $[p]V$. Assume the contrary: then there is $\underline{\alpha}$ such that $h(\underline{\alpha}) \leq \mu_{\text{ess}}(V) + \varepsilon$, $F(\underline{\alpha}^p) = 0$. By the first metric inequality and the product formula we obtain

   $$0 \leq h(F) + n \log(L+1) + pLh(\underline{\alpha}) - T \log p.$$

   Combined with $h(\underline{\alpha}) \leq \mu_{\text{ess}}(V) + \varepsilon$ and the upper bound on $h(F)$ given by Siegel's lemma, this inequality yields

   $$T \log N \ll \left(1 + \frac{DT^2}{L}\right) \log L + (NL + DT)\hat{\mu}_{ess}(V),$$

   where the implicit constant depends only on n.

   Extrapolation – choice of parameters. One may choose $L \approx DT^2$ (even if DT^2/L is much smaller than 1, $1 + DT^2/L$ is still approximately 1). Suppose $\log L \ll \log D$ and $\log N \ll \log \log D$.

   Notation 8.6. *We write $x \lll y$ to mean that there exist constants A, B such that $x \leq Ay + B \log \log D$.*

   We wish to show, by contradiction, that

   $$(NL + DT)\hat{\mu}_{ess}(V) \gg \log D.$$

   Suppose the contrary.

   With our choice of parameters we have $T \lll \log D$. Now we choose $T \leq C \log D$ to obtain a contradiction. This allows us to conclude that the auxiliary function vanishes on $[p]V$, and finishes the extrapolation phase.

   Zeroes lemma. Given that F vanishes on $[p]V$ for all the primes in the interval $[N/2, N]$, one has that

   $$\deg \left(\bigcup_{N/2 \leq p \leq N} [p]V \right) \leq L.$$

   Outside of an exceptional set of primes[5], which we just ignore, one obtains

   $$\deg \left(\bigcup_{N/2 \leq p \leq N} [p]V \right) \geq \sum_{N/2 \leq p \leq N} \deg([p]V) \ggg ND.$$

   Here we have used

   $$\deg([p]V) = \frac{p^{\dim V} \deg V}{|\ker[p] \cap \operatorname{Stab}(V)|} \geq \deg(V),$$

   which follows from the fact that $\dim \operatorname{Stab}(V) \leq n - 1$, so $|\ker[p] \cap \operatorname{Stab}(V)| \leq p^{n-1}$. Now $L \approx DT^2 \approx D(\log D)^2$, so it suffices to choose $N \approx C(\log D)^2$ with C sufficiently large. This leads to a contradiction, hence the assumption $(NL + DT)\hat{\mu}_{ess}(V) \ll \log D$ must be false. Hence $\mu_{\text{ess}}(V) \gg \frac{\log D}{NL} \gg (D(\log D)^3)^{-1}$.

[5] the number of which is negligible with respect to $N/\log N$

2. Case 2: $V/\overline{\mathbb{Q}}$ is geometrically irreducible. We only consider the case $n = 2$, i.e. V is a curve in $\mathbb{G}_m^2$. The group $\mathrm{Stab}(V)$ is finite, and to simplify the presentation we assume that it is trivial.

   Extrapolation. We claim that the auxiliary function F vanishes on $\ker[p]V$, for $N/2 \leq p \leq N$ (at least if the parameters L, T, N satisfy certain inequalities). By contradiction assume the contrary, and choose $\underline{\alpha} \in V(\overline{\mathbb{Q}})$ such that

   (a) $h(\underline{\alpha}) \leq \mu_{\mathrm{ess}}(V) + \varepsilon$

   (b) $\exists \underline{\zeta} \in \ker[p]$ such that $F(\underline{\zeta}\underline{\alpha}) \neq 0$.

   The second metric inequality (1), combined with the product formula, yields

   $$0 \leq h(F) + n\log(L+1) + Lh(\alpha) - \frac{T}{p}\log p,$$

   which implies

   $$\frac{T}{N}\log N \ll \left(1 + \frac{DT^2}{L}\right)\log L + (L + DT)\,\mu_{\mathrm{ess}}(V).$$

   Extrapolation – choice of parameters. As above, we choose $L \approx DT^2$ and we ask that the inequalities $\log L \ll \log D$, $\log N \ll \log\log D$ hold. Again we forget all terms in $\log\log D$; assume that $(L + DT)\mu_{\mathrm{ess}}(V)\log D$. The conclusion is

   $$\frac{T}{N} \ll \log D.$$

   Choose $T \approx CN\log D$; this gives a contradiction, hence we may assume that the auxiliary function vanishes on all the translates of V by points in $\ker[p]$.

   Zeroes lemma.

   $$L \geq \deg\left(\bigcup_{\underline{\zeta}\in\ker[p], N/2\leq p\leq N} \underline{\zeta}V\right) \gg \sum_{p}\sum_{\underline{\zeta}\in\ker[p]} D \gg N^3 D,$$

   up to factors of $\log(N) \ll \log\log D$. Now one finishes the proof by choosing $L \approx DT^2 \approx N^2D(\log D)^2$ and $N \approx C(\log D)^2$; the conclusion is

   $$\mu_{\mathrm{ess}}(V) \gg \frac{1}{D(\log D)^5}.$$

   Remark 8.7. Notice that the proof **cannot** work for $n = 1$, because the result would amount to an absolute lower bound on the canonical height of algebraic numbers. Indeed, the crucial point where $n = 2$ is used is the fact that $|\ker[p]| = p^2$ in $\mathbb{G}_m^2$, but $|\ker[p]| = p$ in $\mathbb{G}_m$.

□

8.1 Plan for the remaining lectures

1. Zilber's conjecture and a panorama of related problems. Zilber's conjecture implies the toric case of Manin-Mumford (which is a theorem, and is even effective). Zilber's conjecture implies Lang's conjecture (which is also a theorem, but is ineffective).

2. Motivations for Zilber's conjecture: together with Schanuel, it leads to a uniform version of Schanuel (which admits a better formulation in model theory).

3. Known cases of Zilber's conjecture (Bombieri-Masser-Zannier): Schinzel's conjecture.

4. Applications to problems of greatest common divisors of lacunary polynomials. Sketch of proof of Schinzel's conjecture.

9 31.05.2018 – Computing Mahler measures

9.1 Graeffe's method: Mahler measure of $P \in \mathbb{C}[x]$

Given $p(x) = \prod(x - x_i)$, one wants to compute $p^{[\ell]}(x) = \prod(x - x_i^\ell)$. This can be done computing suitable resultants; to simplify the computational part of the problem, it is useful to choose $\ell = 2^m$, so that the problem boils down to computing $p^{[2]}(x) = \prod(x - x_i^2)$.

Write $p(x) = A(x^2) + B(x^2)x$ and set

$$(\tau p)(x) := A(x)^2 - B(x)^2 x.$$

Then one has:

Remark 9.1. 1. $\deg(\tau p) \le \deg p$

2. $M(\tau p) = M(p)^2$, because the roots of (τp) are the squares of the roots of $p(x)$. Indeed, if y is a root of $p(x)$, then

$$(\tau p)(y^2) = (A(y^2) + B(y^2)y)(A(y^2) - B(y^2)y) = p(y)p(-y).$$

The fact that the roots of τp are the squares of those of p already implies $M(\tau p) = M(p)^2$; another way of getting at the same conclusion is to remember that Mahler's measure does not change under isogenies, hence

$$M(p)^2 = M(p(y))M(p(-y)) = M(\tau p(y^2)) = M(\tau p(y))$$

As a consequence of an exercise we already discussed, we obtain

$$M(p)^{2^m} \le \|\tau^{(m)}p\| \le 2^{\deg p} M(p)^{2^m},$$

and taking the 2^m-root of this inequality we get

$$M(p) \le \|\tau^{(m)}p\|^{1/2^m} \le 2^{\deg p/2^m} M(p),$$

hence the sequence $\|\tau^{(m)}p\|^{1/2^m}$ converges very quickly to $M(p)$.

9.1.1 Polynomials in two variables

The idea is the same (and can be generalised to an arbitrary number of variables). Write

$$P = A_0(x^2, y) + B_0(x^2, y)x$$

and set

$$P_1 = A_0(x, y)^2 - B_0(x, y)^2 x;$$

write this polynomial as

$$P_1(x, y) = A_1(x, y^2) + B_1(x, y^2)y$$

and set

$$\tau P = P_2 = A_1(x, y)^2 - B_1(x, y)^2 y.$$

Our previous remarks generalise: we have

$$\deg_x(\tau P) \le 2\deg_x(P), \quad \deg_y(\tau P) \le 2\deg_y(P)$$

and

$$M(\tau P) = M(P)^4.$$

We don't show this second equality, but it follows from a generalisation of the identity $\tau P(y^2) = P(y)P(-y)$ that held for a univariate polynomial.

Recall the 2-dimensional inequality for Mahler measures:

$$M(P) \le \|P\| \le 2^{\deg_x P + \deg_y P} M(P).$$

Applying this to $\tau^{(m)}P$ we obtain

$$M(P)^{4^m} \le \|\tau^{(m)}P\| \le 2^{2^m(\deg_x P + \deg_y P)} M(P)^{4^m},$$

and taking the 4^m-th root we obtain again that $\|\tau^{(m)}P\|^{1/4^m}$ converges to $M(P)$.

9.1.2 Essential minimum of a curve

Once the Mahler measure of a polynomial F is known (and therefore also the height of the corresponding curve $\mathcal{C}$ is known) we can estimate the essential minimum of $\mathcal{C}$ using Zhang's inequality:

$$\frac{\hat{h}(\mathcal{C})}{2\deg\mathcal{C}} \leq \mu_{\text{ess}}(\mathcal{C}) \leq \frac{\hat{h}(\mathcal{C})}{\deg\mathcal{C}}$$

9.1.3 A remark (Dvornicich)

One can use similar tricks (avoiding resultants) to write down polynomials that have as roots (say) the cubes of the roots of a given one. In this generalisation, one finds a formula that involves $P(y), P(\zeta y), P(\zeta^2 y)$ where ζ is a cube root of unity. The geometric reason for this is that $[\ell]^{-1}\mathcal{C}$ is a union of translates of a certain subvariety (morally defined by $P(\sqrt[\ell]{x})$) by ℓ-th roots of unity.

9.2 Zilber's conjecture (2002)

Conjecture 9.2. *Let $V \subseteq \mathbb{G}_m^n$ (V defined over $\mathbb{C}$) an irreducible variety, and let $T \subset \mathbb{G}_m^n$ be a torsion coset. Suppose that there exists an irreducible component Y of $V \cap T$ whose dimension is unlikely*[6] *large, namely*

$$\dim Y > \dim(V) - \operatorname{codim}(T)$$

In Zilber's terminology, Y is called an atypical component. Then there exists a torsion coset T', contained in a set S_V (finite and depending only on V), such that $Y \subseteq T'$.

9.2.1 Some special cases

1. Suppose $\dim T = 0$, that is, $T = \{\underline{\omega}\}$ where $\underline{\omega}$ is a torsion point. In this case $\operatorname{codim} T = n$, so the conjecture is only interesting for $\dim(Y) = 0 > \dim(V) - n$, so that "$Y$ atypical" means "Y nonempty", or equivalently $\underline{\omega} \in V$. Zilber says that $\underline{\omega}$ belongs to T', where T' lies in a finite set of torsion cosets. In particular, when V is a curve, Zilber's conjecture implies that there are only finitely many torsion points lying on V (unless V contains a torsion coset), that is, the toric case of Manin-Mumford.

2. $\dim T = 1$. In this case Zilber's conjecture implies Schinzel's conjecture on lacunary polynomials:

   Conjecture 9.3 (Schinzel)**.** *Let $F, G \in \mathbb{Z}[x_1, \ldots, x_n]$ be two relatively prime polynomials. Let $\underline{a} \in \mathbb{Z}^n$. Suppose there exists an algebraic numer $\underline{\xi}$, not a root of unity, such that $F(\xi^{a_1}, \ldots, \xi^{a_n}) = G(\xi^{a_1}, \ldots, \xi^{a_n}) = 0$. Then there exists a vector $\underline{b} \in \mathbb{Z}^n \setminus \{0\}$ such that*

   (a) $\max|b_j| \leq B(F,G)$*, a constant depending only on F, G and not on a;*

   (b) $\underline{a} \cdot \underline{b} = 0$ *(the vectors are orthogonal)*

   Let's clarify the connection. Define $X = \{F = G = 0\} \subset \mathbb{G}_m^n$ and $\underline{\alpha} = (\xi^{a_1}, \ldots, \xi^{a_n})$. The assumption that F, G are relatively prime implies that X is a complete intersection, hence of codimension 2. It might not be irreducible: take an irreducible component V of X with $\underline{\alpha} \in V$. Now observe that

   $$\underline{\alpha} \in T = \{(t^{a_1}, \ldots, t^{a_n}) \bigm| t \in \mathbb{G}_m\},$$

   so that $\underline{\alpha} \in V \cap T$, $\dim\{\underline{\alpha}\} = 0$, and

   $$\dim V - \operatorname{codim} T = \dim T - \operatorname{codim} V = 1 - 2 = -1.$$

[6] in italiano: *patologicamente*

Hence in this setting $Y = \{\underline{\alpha}\}$ is atypical. Assuming Zilber's conjecture, we know that $\underline{\alpha}$ belongs to a finite set of torsion cosets that depends only on F, G; enlarging these torsion cosets if necessary, we can assume that they are of codimension 1, hence given by a single equation. Equivalently, there exists a finite list of vectors[7] $\underline{b} \in \mathbb{Z}^n \setminus \{0\}$ and a finite list of roots of unity ω such that $\underline{\alpha}$ belongs to one of the finitely many torsion cosets $T' = \{x_1^{b_1} \cdots x_n^{b_n} = \omega\}$. Clearly this implies

$$\xi^{a_1 b_1 + \cdots + a_n b_n} = \omega,$$

and since ξ is not a root of unity while ω is, this is only possible if $a_1 b_1 + \cdots + a_n b_n = 0$, as we wanted to show.

Remark 9.4. Conjecture 9.3 is now a theorem of Zannier.

3. $\dim V = 1$. Intersect V with $\bigcup_{\substack{H \text{ algebraic subgroup} \\ \operatorname{codim} H \geq 2}} H$: by dimension reasons, what one expects is that this should be a finite set, unless V is contained in a torsion coset.

   This fact is true, and is a theorem of Bombieri-Masser-Zannier (1999) and Maurin (2008). More precisely, Bombieri-Masser-Zannier proved the theorem assuming that V is not contained in any coset (not necessarily torsion), and the work of Maurin subsequently allows one to replace "not contained in any coset" with the (optimal) hypothesis "not contained in a torsion coset".

9.2.2 Lang's conjecture

Conjecture 9.5 (Lang)**.** *Let $V\mathbb{G}_m^n$ be irreducible, $\Gamma \subseteq \mathbb{G}_m^n$ a subgroup of finite rank (*$\dim(\Gamma \otimes \mathbb{Q}) < \infty$*). Suppose $\overline{\Gamma \cap V} = V$: then V is a translate of a torus by a point of Γ.*

Remark 9.6. 1. Lang's conjecture implies the toric case of Manin-Mumford by taking $\Gamma = (\mathbb{G}_m^n)_{\text{tors}}$.

2. This conjecture is now a theorem of Michel Laurent

3. The known proofs are all intrinsically ineffective

4. There is an equivalent (but apparently stronger) formulation:

$$\overline{\Gamma \cap V} = \bigcup_{j=1}^{r} T_j \gamma_j,$$

   where every T_j is a torus and every γ_j is a point of Γ. In particular, if V does not contain translates of tori of positive dimension, then $\Gamma \cap V$ is finite.

5. Zilber has shown that his conjecture implies Lang's (this is shown in Zilber's original paper about his conjecture).

6. There are versions of Manin-Mumford and of conjecture 9.5 for abelian varieties; they are now (ineffective) theorems, due respectively to Michel Raynaud and to Faltings.

9.3 Transcendence

9.3.1 Historical introduction

1873 Hermite shows that $e \notin \overline{\mathbb{Q}}$

1882 Lindemann shows that for every $\alpha \in \overline{\mathbb{Q}}^{\times}$ the number e^{α} is not algebraic. In particular, taking $\alpha = 2\pi i$ shows that π is transcendental (and squaring the circle is proven to be impossible).

[7] in particular, $\|b\|_\infty$ is bounded by a constant depending only on F, G

1885 Lindemann and Weierstrass prove that if $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}}$ are $\mathbb{Q}$-linearly independent, then $e^{\alpha_1}, \dots, e^{\alpha_n}$ are algebraically independent. This can be expressed in terms of transcendence degrees: $\operatorname{trdeg}_{\overline{\mathbb{Q}}} \mathbb{Q}(\alpha_1, \dots, \alpha_n, e^{\alpha_1}, \dots, e^{\alpha_n}) = n$.

1966 Schanuel conjectures the following: let $\alpha_1, \dots, \alpha_n$ be $\mathbb{Q}$-linearly independent. Then

$$\operatorname{trdeg}_{\mathbb{Q}} \mathbb{Q}(\alpha_1, \dots, \alpha_n, e^{\alpha_1}, \dots, e^{\alpha_n}) \geq n.$$

There is also an equivalent version, which can be stated as follows: let $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ be arbitrary. Then

$$\operatorname{trdeg}_{\mathbb{Q}} \mathbb{Q}(\underline{\alpha}, e^{\underline{\alpha}}) \geq \dim_{\mathbb{Q}} \langle \alpha_1, \dots, \alpha_n \rangle_{\mathbb{Q}}.$$

Finally, there is also a weaker version which is also interesting to consider, namely: let $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ be arbitrary and $e^{\alpha_1}, \dots, e^{\alpha_n}$ be multiplicatively independent. Then

$$\operatorname{trdeg}_{\mathbb{Q}} \mathbb{Q}(\alpha_1, \dots, \alpha_n, e^{\alpha_1}, \dots, e^{\alpha_n}) \geq n.$$

Remark 9.7.

- The 'weak' version of Schanuel's conjecture "can't determine" whether π and e are algebraically independent or not.
- The full Schanuel conjecture implies that π, e are algebraically independent: take $\alpha_1 = 1, \alpha_2 = 2\pi i$ to get

$$2 \leq \operatorname{trdeg}_{\mathbb{Q}} \mathbb{Q}(\alpha_1, \alpha_2, e^{\alpha_1}, e^{\alpha_2}) = \operatorname{trdeg}_{\mathbb{Q}} \mathbb{Q}(2\pi i, e).$$

- There are models of $(\mathbb{C}, \exp)$ in which the weak version of Schanuel holds, but the strong one does not.

9.3.2 Geometric version of Schanuel's conjecture

Let $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ be $\mathbb{Q}$-linearly independent. Let $V \subset \mathbb{G}_m^n$ be defined over $\mathbb{Q}$ with $\dim V < n$. Suppose that $(\underline{\alpha}, \exp(\underline{\alpha})) \in V$: then $\alpha_1, \dots, \alpha_n$ are $\mathbb{Q}$-linearly dependent.

In model theory, this last condition is not considered very nice, because one needs to write it as

$$\exists a_1, \dots, a_n \in \mathbb{Z}, a_1^2 + \cdots + a_n^2 \neq 0, \text{ such that } a_1\alpha_1 + \cdots + a_n\alpha_n = 0.$$

Quantifying over the integers is always dangerous, so we (as in 'the model theorists') don't like it. This is why one also considers uniform versions of Schanuel's conjecture, which are stated as follows:

Conjecture 9.8 (Uniform Schanuel)**.** *Let $V/\mathbb{Q}$, $V \subset \mathbb{C}^n \times \mathbb{G}_m^n$, $\dim V < n$. Suppose that $\underline{\alpha} \in \mathbb{C}^n$ is such that $(\underline{\alpha}, \exp(\underline{\alpha}))$ belongs to V. Then $\underline{\alpha}$ belongs to L, a linear subspace that belongs to a finite set depending only on V. In particular, $\exp(\underline{\alpha})$ belongs to T, a torsion coset taken from a finite set that depends only on V.*

Theorem 9.9 (Zilber)**.** *Assume conjecture 9.2. Then Schanuel's conjecture implies the uniform Schanuel conjecture.*

Proof. We show something weaker, namely that $\exp(\underline{\alpha})$ belongs to T, a torsion coset taken from a finite set that depends only on V.

Consider the projection $\pi : V \to \pi(V) \subseteq (\mathbb{C}^\times)^n$. Let d be the dimension of the generic fiber, which is $\dim V - \dim \pi(V) < n - \dim \pi(V)$. Let

$$V' = \{(\underline{x}, \underline{y}) \in V \mid \dim \pi^{-1}(\underline{y}) > d\}.$$

It is a (proper) closed subset of V. By induction (details omitted) we can assume that $(\underline{\alpha}, \exp(\underline{\alpha})$ is not in V'. Let

$$\Lambda = \{\underline{\lambda} \in \mathbb{Z}^n \mid \sum \lambda_i \alpha_i = 0\};$$

it is a saturated lattice. One has $\dim \Lambda = n - \dim_{\mathbb{Q}} \langle \underline{\alpha} \rangle_{\mathbb{Q}}$; if we assume Schanuel's conjecture, this number is strictly positive. Define

$$T = H_\Lambda = \{\underline{y} \in (\mathbb{C}^\times)^n \mid \prod y_i^{\lambda_i} = 1 \quad \forall \lambda \in \Lambda\},$$

so that T is a torus with $\operatorname{codim} T = \dim \Lambda$, hence

$$\dim T = \dim_{\mathbb{Q}} \langle \underline{\alpha} \rangle_{\mathbb{Q}}.$$

Observe that by definition $\exp(\underline{\alpha})$ belongs to T.

The equivalent version of Schanuel gives

$$\operatorname{trdeg}_{\mathbb{Q}}(\underline{\alpha}, \exp(\underline{\alpha})) \geq \dim_{\mathbb{Q}} \langle \underline{\alpha} \rangle_{\mathbb{Q}} = \dim T.$$

On the other hand,

$$\operatorname{trdeg}_{\mathbb{Q}} \mathbb{Q}(\underline{\alpha}, \exp(\underline{\alpha})) = \operatorname{trdeg}_{\mathbb{Q}} \mathbb{Q}(\exp(\underline{\alpha})) + \operatorname{trdeg}_{\mathbb{Q}(\exp(\underline{\alpha}))} \mathbb{Q}(\underline{\alpha}, \exp(\underline{\alpha})).$$

Let Y be an irreducible component (defined over $\mathbb{Q}$, and of maximal dimension) of $\pi(V) \cap T$ containing $\exp(\underline{\alpha})$. We have

$$\begin{aligned} \operatorname{trdeg}_{\mathbb{Q}} \mathbb{Q}(\underline{\alpha}, \exp(\underline{\alpha})) &= \operatorname{trdeg}_{\mathbb{Q}} \mathbb{Q}(\exp(\underline{\alpha})) + \operatorname{trdeg}_{\mathbb{Q}(\exp(\underline{\alpha}))} \mathbb{Q}(\underline{\alpha}, \exp(\underline{\alpha})) \\ &\leq \dim Y + \dim \pi^{-1}(\exp(\underline{\alpha})) \\ &= \dim Y + d, \end{aligned}$$

where we have used the fact that the fiber over $\pi(\underline{\alpha})$ contains $\underline{\alpha}$, so it gives an upper bound for its transcendence degree (over the field $\mathbb{Q}(\exp(\underline{\alpha}))$, which is the field over which $\exp(\underline{\alpha}) = \pi(\underline{\alpha}, \exp(\underline{\alpha}))$ is defined). Hence we get

$$\dim Y \geq \dim T - d > \dim T - (n - \dim \pi(V)) = \dim \pi(V) - \operatorname{codim} T,$$

so Y is atypical, and $Y \subseteq T'$, finite set depending only on V. Some more work would then be necessary to get rid of the $2\pi i$ which comes up when taking logarithms. □

10 05.06.2018 – GCD of lacunary polynomials

Recall the statement of the conjecture: $\forall V \subset \mathbb{C}^n \times (\mathbb{C}^\times)^n$, V defined over $\overline{\mathbb{Q}}$ and $\overline{\mathbb{Q}}$-irreducible, $\dim V < n$. Assume that $V \cap \Gamma_{\exp} \neq \emptyset$, where $\Gamma_{\exp}$ is the graph of the exponential function. Then there exists a $\mathbb{Q}$-linear subspace $H \subsetneq \mathbb{C}^n$ such that $\alpha \in H$

The problem for people working in Model theory is the existential quantifier; we have seen last time that assuming Zilber (and Schanuel) one might prove a uniform version of Schanuel, namely $\forall V \subseteq \mathbb{C}^n \times (\mathbb{C}^\times)^n$ there exists a finite set S of proper $\mathbb{Q}$-linear subspaces such that $V \cap \Gamma_{\exp} \neq \emptyset$ there exists $H \in S$ such that $\underline{\alpha} \in H$.

Remark 10.1. Combining the theorem of Lindemann-Weierstrass and the toric case of Manin-Mumford (due to Michel Laurent) one gets that for all $V \subseteq (\mathbb{C}^\times)^n$, $V/\overline{\mathbb{Q}}$ irreducible the following holds:

$$V \cap \exp(\overline{\mathbb{Q}}^n) = T_1 \cap \exp(\overline{\mathbb{Q}}^n) \cup \cdots \cup T_k \cap \exp(\overline{\mathbb{Q}}^n)$$

where $T_1, \ldots, T_k$ are the finitely many[8] subtori of V.

Proof. Exercise (same strategy used to prove that Zilber+Schanuel implies uniform Schanuel). □

[8] by Manin-Mumford

Theorem 10.2 (Bombieri-Zannier 1998, Schinzel's conjecture). *Let V be a subvariety of $\mathbb{G}_m^n$ defined over a number field*[9] *K of degree $\leq \delta$ by polynomials of naïve height at most h_0 and of degree $\leq d_0$. Let W be a $\overline{\mathbb{Q}}$-irreducible component of V of codimension at least 2. Let T be a torsion coset of dimension 1. Suppose there is a point $\underline{\alpha} \in W \cap T$; in the language of the last lecture, $\underline{\alpha}$ is an atypical component of the intersection. Then there exists a torsion coset T' with $\deg(T') \leq B(V)$ such that $\underline{\alpha} \in T'$.*

Remark 10.3. Bounding the degree is equivalent to imposing finiteness of the possible T'.

Remark 10.4 (Schinzel). B **must** depend on h_0: take $W = V_a$ defined by the equations $x - 2 = y - 2^a = 0$, where $a \in \mathbb{N}$ is a parameter. Consider the torsion coset (in fact, the torus) T_a parametrised by $\{(t, t^a) : t \in \mathbb{G}_m\}$. Then $\underline{\alpha} = (2, 2^a) \in W \cap T_a$. If $\underline{\alpha}$ belongs to T', then necessarily $\det(T') \gg a$.

10.1 Application to GCDs of lacunary polynomials

Theorem 10.5. *Let $F, G \in \mathbb{Z}[x_1, \ldots, x_n]$ and $\underline{a} \in \mathbb{Z}^n \setminus \{0\}$. Set $f_{\underline{a}} = F(t^{\underline{a}}), g_a(\underline{a}) = G(t^{\underline{a}})$. Then there exist $k \in \{0, \ldots, n-1\}$ and morphism ρ, ψ such that the following diagram is commutative*

$$\begin{tikzcd}
\mathbb{G}_m \arrow[r, "\varphi_{\underline{a}}"] \arrow[d, "\varphi'_{\underline{a}}:=\rho\circ\varphi_{\underline{a}}"'] & \mathbb{G}_m^n \arrow[dl, bend left, "\rho"] \\
\mathbb{G}_m^{n-k} \arrow[ur, "\psi"'] &
\end{tikzcd}$$

and such that

1. *ρ, ψ have size at most $B(F, G)$, where the size of a morphism is the maximum of the absolute values of the coefficients in the matrices that define them. The crucial part here is that $B(F, G)$ is independent of $\underline{a}$.*

2. *set $P := \gcd(F \circ \psi, G \circ \psi)$, where $F \circ \psi$ actually means the pullback on the rings of regular functions*[10]*. By commutativity of the diagram, $h := P \circ \rho \circ \varphi_{\underline{a}}$ divides $\gcd(f_{\underline{a}}, g_{\underline{a}})$. Here comes the punchline: one has that $\frac{\gcd(f,g)}{h}$ is a product of cyclotomic factors.*

10.2 Proof: theorem 10.2 implies theorem 10.5

Let $\xi \in \overline{\mathbb{Q}}^\times$ be a common root of $f_{\underline{a}}$ and $g_{\underline{a}}$ which is not a root of unity. Define $V = \{F = G = 0\} \subseteq \mathbb{G}_m^n$.

One has $\varphi_{\underline{a}} \in V \cap T_a$, where T_a is the torus $\operatorname{Im} \varphi_{\underline{a}}$.

Define W to be an irreducible component of V of dimension 2 containing ξ. Notice that this needs component needs not exist, since V might even be of codimension 1 (e.g. if $F = G$).

By theorem 10.2 there exists a torsion coset T' (wlog of codimension 1), with degree $\deg T' \leq B(V) = B(F, G)$, such that $\varphi_{\underline{a}}(\xi) \in T'$.

Remark 10.6. Since ξ is not a root of unity, T' is a torus and $\operatorname{Im} \varphi_{\underline{a}} \subseteq T'$. To see this, notice that $\varphi_{\underline{a}}(\xi) \in T'$ means $\xi^{a_1\lambda_1} \cdots \xi^{a_n\lambda_n} = \omega$ with ω a root of unity; since ξ is not a root of unity, this implies $\omega = 1$ and $\sum a_i\lambda_i = 0$.

Let τ be an automorphism of $\mathbb{G}_m^n$ of size $\ll_{F,G} 1$ that sends T' to $\{x_n = 1\}$. Let $\iota : \mathbb{G}_m^{n-1} \hookrightarrow \mathbb{G}_m^n$ be the injectiion $\underline{x} \mapsto (\underline{x}, 1)$ and $\pi : \mathbb{G}_m^n \to \mathbb{G}_m^{n-1}$ be the projection on the first $n-1$ coordinates.

[9] in fact, we want all geometrically irreducible components to also be defined over that same number field

[10] concretely, if $y_1, \ldots, y_{n-k}$ are variables on gm^{n-k}, $F \circ \psi$ is a polynomial in the y_i, given by $F(m_1, \ldots, m_n)$, where the m_j are monomials in the y_i

It follows that $\mathrm{Imm}(\tau \circ \varphi_a) \subseteq \{x_n = 1\}$, and we have

$$\begin{tikzcd}
\mathbb{G}_m \arrow[r, "\varphi_{\underline{a}}"] \arrow[d, "\pi\circ\tau\circ\varphi_{\underline{a}}"'] & \mathbb{G}_m^n \\
\mathbb{G}_m^{n-1} \arrow[ur, "\tau^{-1}\circ\iota"'] &
\end{tikzcd}$$

Remark 10.7. The construction depends only on T' (which is a torus and satisfies $\mathrm{Im}(\varphi_{\underline{a}}) \subseteq T'$.

We now continue inductively with n replaced by $n-1$ and F, G replaced by $F\circ\tau^{-1}\circ\iota$, $G\circ\tau^{-1}\circ\iota$ (and $\varphi_{\underline{a}}$ replaced by $\pi \circ \tau \circ \varphi_{\underline{a}}$. Repeating the same argument, either there is a subvariety W as above (and then I do exactly the same as above), or there is no such W, and then I stop. When I stop, it means that $\varphi_{\underline{a}}$ belongs to a component of codimension 1, that is, it lies on some component defined by a nontrivial factor of the $\gcd(F, G)$ (or rather, of the polynomials found by changing variables).

10.2.1 Ok, but how do we do it concretely?

In practice one does not have ξ. Choose a torus T' of degree at most $B(F, G)$ and such that $\mathrm{Im}\,\varphi_{\underline{a}} \subseteq T'$, and I apply the construction described above. Repeat until possible.

10.3 Sketch of proof of theorem 10.2

Write

$$T = \{(\zeta_1 t^{a_1}, \dots, \zeta_n t^{a_n} \mid t \in \mathbb{G}_m\}, \quad \underline{\zeta} \in (\mathbb{G}_m^n)_{\mathrm{tors}}.$$

10.3.1 Digression on the torsion coset locus of a variety

Let V be a subvariety of $\mathbb{G}_m^n$ and define V^u as the union of all the torsion cosets contained in V. Further define:

1. V^a to be the union of all the cosets (of dimension > 0) contained in V
2. $V^* = V \setminus V^u$
3. $V^0 = V \setminus V^a$

A possible formulation of Manin-Mumford is that V^u is the union of the maximal torsion cosets of V. Recall the following theorem:

Theorem 10.8 (Laurent). *Let $V \subseteq \mathbb{G}_m^n$ be defined by*

$$f_\ell(\underline{x}) = \sum_{\underline{\lambda}\in I} a_{\ell,\underline{\lambda}} x^{\underline{\lambda}} = 0$$

for $\ell = 1, \dots, t$. If $H \subseteq V$ is a maximal subgroup, there exists a subgroup Λ of $\mathbb{Z}^n$ generated by the vectors of

$$D(I) = \{\underline{\lambda} - \underline{\mu} \mid \underline{\lambda}, \underline{\mu} \in I\}$$

and such that $H = H_\Lambda$.

Remark 10.9. n particular, the cosets $T = aH \subseteq V$ are all obtained from subgroups contained in $a^{-1}V$ (and $D(I)$ does not change).

Example 10.10. Let $V : \{\sum_{i=0}^n a_{\ell,i} x_i = 0\}$ for $\ell = 1, \dots, t$, with $x_0 = 1$. Given a partition $\mathcal{P} = \Lambda_1 \cup \cdots \cup \Lambda_k$ of $\{0, \dots, n\}$ let $\sim$ be the associated equivalence relation ($i \sim j$ iff i, j belong to the same Λ_r). Then the set

$$H_{\mathcal{P}} = \{\underline{x} : x_i = x_j \text{ if } i \sim j\}$$

is a subgroup of $\mathbb{G}_m^n$, and its cosets $H\underline{g}$ are defined by

$$x_i g_i^{-1} = x_j g_j^{-1} \quad \text{if } i \sim j.$$

Such a coset is contained in V iff

$$\forall \underline{x} \in H_{\mathcal{P}} g, \quad \sum_{i \in \Lambda_1} a_{\ell,i} x_i = \cdots = \sum_{i \in \Lambda_k} a_{\ell,i} x_i = 0$$

In other words, points in V^a are *degenerate* solutions – those for which there is a proper subsum equal to 0.

Remark 10.11. In the previous example, if $i \sim 0$ one needs to have $x_i = x_0 = 1$ (so these 'extra equations' that we didn't write down explicitly, but are there)

10.4 Back to the proof of theorem 10.2

There is an easy case: $\underline{\alpha} \in W^a$. Write

$$W^a = \bigcup_{i=1}^{k} \bigcup_{a \in S_i} H_i a,$$

with k finite, H_i tori, and S_i potentially infinite.

Example 10.12. In $\mathbb{G}_m^3$, taking $V = \{x_1 + x_2 + x_3 = 0\}$ and $H = \{(t,t,t) : t \in \mathbb{G}_m\}$, then $V = \bigcup_{\underline{\alpha} \in V} H\alpha$ and $V^0 = \emptyset$ (even more directly, take $V = W \times \mathbb{G}_m \subseteq \mathbb{G}_m^N$: then $V^0 = \emptyset$).

It follows that there exists a torus H (of positive dimension $n - h > 0$), lying in a **finite** set depending only on V, and a point $g \in \mathbb{G}_m^n$, such that $\underline{\alpha} \in Hg \subseteq W$. Of course one could take $g = \underline{\alpha}$. After applying an automorphism of $\mathbb{G}_m^n$ which depends only on V we may assume that $H = \{1\}^h \times \mathbb{G}_m^{n-h}$. Let $W' = \{\underline{x} \in \mathbb{G}_m^h : (x_1, \dots, x_n) \times \mathbb{G}_m^{n-h} \subseteq W\}$. We now know $W' \times \mathbb{G}_m^{n-h} \subseteq W$ and $(\alpha_1, \dots, \alpha_n) \in V'$. We'd like to proceed by induction: we notice that

$$\dim(W') + (n-h) \leq \dim W \leq n-2,$$

hence $\dim(W') \leq h-2$ and W' has codimension at least 2. This allows us to apply the inductive hypothesis (notice that the case $n=2$ is trivial since $\dim W = 0$).

Hence we may assume $\underline{\alpha} \in W^0$. We may further assume that $\underline{\alpha}$ is not a torsion point. Indeed, if it is, using the fact that W^0 does not contain positive-dimension torsion subvarieties we obtain that α is contained in a 0-dimensional torsion subvariety of W^0, but there are only finitely many of these.

11 06.06.2018 – GCD of lacunary polynomials

11.1 An example: GCD of two lacunary polynomials

Consider the two polynomials

$$f(t) = t^{u+v} - 5t^u + 2t^{v+1} + 6t^v - 5t - 15$$

$$g(t) = 3t^{u+v} - 2t^u + 3t^{v+1} + 9t^v - 2t - 6$$

We linearise the problem by considering linear polynomials

$$F(x_1, \dots, x_5) = 2x_1 - 5x_2 + 2x_3 + 6x_4 - 5x_5 - 15G(x_1, \dots, x_6) = 3x_1 - 2x_2 + 3x_3 + 9x_4 - 2x_5 - 6$$

The vector $\underline{a}$ is $(u+v, u, v+1, v, 1)$.

1. We look for a 'small' torus that contains the image of $\varphi_{\underline{a}}$, namely $\{(t^{u+v}, t^u, t^{v+1}, t^v, t\}$. We may take
$$T' = \{\underline{x} \in \mathbb{G}_m^5 : x_1 x_2^{-1} x_4^{-1} = 1\}.$$
We then have
$$\begin{array}{cccc} \tau: & \mathbb{G}_m^5 & \to & \mathbb{G}_m^5 \\ & \underline{x} & \mapsto & (x_3, x_4, x_3, x_5, x_1 x_2^{-1} x_4^{-1}) \end{array}$$
and
$$\begin{array}{cccc} \tau^{-1}: & \mathbb{G}_m^5 & \to & \mathbb{G}_m^5 \\ & \underline{y} & \mapsto & (y_1 y_2 y_5, y_1, y_3, y_2, y_4) \end{array}$$
The map ψ (or rather, the 'piece' of ψ from $\mathbb{G}_m^4$ to $\mathbb{G}_m^5$) is obtained as $\tau^{-1} \circ \iota$, where $\iota(y_1, \dots, y_4) = (y_1, \dots, y_4, 1)$, and is therefore given by $\psi(y_1, \dots, y_4) = (y_1 y_2, y_1, y_3, y_2, y_4)$. Finally, $\varphi'_{\underline{a}}$ is obtained as $\rho \circ \varphi_{\underline{a}} = \pi \circ \tau \circ \varphi_{\underline{\alpha}}$, where $\pi : \mathbb{G}_m^5 \to \mathbb{G}_m^4$ is the projection on the first 4 components. We have therefore described the following commutative diagram:
$$\begin{tikzcd}
\mathbb{G}_m^5 & & \mathbb{G}_m \arrow[ll, "{(t^{u+v},t^u,t^{v+1},t)}"'] \arrow[dll, "{(t^u,t^v,t^{v+1},t)}"] \\
\mathbb{G}_m^4 \arrow[u, "{\underline{y} \mapsto (y_1 y_2, y_1, y_3, y_2, y_4)}"] & &
\end{tikzcd}$$

2. Again we look for a small torus containing the image of $\varphi'_{\underline{a}} = \{(t^u, t^v, t^{v+1}, t) \bigm| t \in \mathbb{G}_m\}$. We may take
$$T' = \{y \in \mathbb{G}_m^4 : y_2 y_3^{-1} y_4 = 1\}.$$
We now want to change variables so as to bring this torus to $\{y_4 = 1\}$; we take
$$\begin{array}{cccc} \tau: & \mathbb{G}_m^4 & \to & \mathbb{G}_m^4 \\ & \underline{y} & \to & (y_1, y_2, y_4, y_2 y_3^{-1} y_4) \end{array}$$
with inverse
$$\begin{array}{cccc} \tau^{-1}: & \mathbb{G}_m^4 & \to & \mathbb{G}_m^4 \\ & \underline{z} & \to & (z_1, z_2, z_2 z_3 z_4^{-1}, z_3) \end{array}$$
Now $\iota(z_1, z_2, z_3) = (z_1, z_2, z_3, 1)$ and the new 'piece' of τ is $\psi^{-1} \circ \iota$. We obtain the new diagram
$$\begin{tikzcd}
\mathbb{G}_m^5 & & \mathbb{G}_m \arrow[ll, "{(t^{u+v},t^u,t^{v+1},t)}"'] \arrow[dll, "{(t^u,t^v,t^{v+1},t)}"] \arrow[ddll, bend left, "{(t^u,t^v,1)}"] \\
\mathbb{G}_m^4 \arrow[u, "{\underline{y} \mapsto (y_1 y_2, y_1, y_3, y_2, y_4)}"] & & \\
\mathbb{G}_m^3 \arrow[u, "{\underline{z} \mapsto (z_1, z_2, z_2 z_3, z_3)}"] & &
\end{tikzcd}$$

3. Now we should look for a 'small' torus containing the image of $\varphi''_{\underline{a}}$, which is $\{(t^u, t^v, t) : t \in \mathbb{G}_m\}$. If u, v are large enough[11], there is no such 'small' torus, so this phase of the algorithm is finished.

4. We then write $F(x), G(x)$ in the new coordinates:
$$F(x) = 2z_1 z_2 - 5z_1 + 2z_2 z_3 + 6z_2 - 5z_3 - 15 = \tilde{F}(z)$$
$$G(x) = 3z_1 z_2 - z_1 + 3z_2 z_3 + 9z_2 - 2z_3 - 6 = \tilde{G}(z)$$

5. We compute $P = \gcd(\tilde{F}, \tilde{G}) = z_1 + z_3 + 3$

6. The conclusion is that, up to cyclotomic factors (and for $u, v \gg 1$), the gcd of f, g is $P(t^u, t^v, t) = t^u + t + 3$.

[11] and 'sufficiently independent' – for example, $u = v$ leads to a small torus

11.2 Continuation of the proof of theorem 10.2

11.2.1 Geometry of numbers

Let Λ be a lattice of rank r in $\mathbb{Z}^n$. Let $W = \Lambda \otimes_{\mathbb{Z}} \mathbb{R} \subseteq \mathbb{R}^n$, $\dim_{\mathbb{R}}(W) = r$. Consider the associated quadratic form

$$Q(x_1, \dots, x_r) = \| \sum_{j=1}^{r} x_j \underline{\lambda}^{(j)} \|_2^2 = \sum_{i,j} q_{i,j}(\underline{\lambda}) x_i x_j$$

It is possible to show that the following three numbers agree:

1. $\operatorname{vol}\left(\left\{\sum_{i=1}^{r} t_i \lambda^{(i)} \;\middle|\; 0 \leq t_i < 1\right\}\right)$
2. $|\det q_{ij}|$
3. $\left\|\text{determinants } r \times r \text{ minors of } \left(\lambda_j^{(i)}\right)_{i,j}\right\|_2^2$ (squared L^2-norm of the vector whose entries are the determinants of $r \times r$ minors of $\left(\lambda_j^{(i)}\right)_{i,j}$

Furthermore, these three numbers are independent of the choice of the basis $\lambda^{(j)}$; their common value is denoted by $\operatorname{vol}(W/\Lambda)$, or simply by $\operatorname{vol}(\Lambda)$. The following hold:

1. $\operatorname{vol}(\Lambda) \leq \prod_{j=1}^{r} \|\lambda^{(j)}\|_2$ (Hadamard's inequality)
2. $\operatorname{vol}(\Lambda^\vee) = \operatorname{vol}(\Lambda)$
3. there exist $u^{(1)}, \dots, u^{(r)} \in \Lambda$, linearly independent (but not necessarily a basis), such that $\prod_{j=1}^{r} \|u^{(j)}\| \ll \operatorname{vol}(\Lambda)$ (Minkowski's theorem)

11.2.2 Application to the construction of 'small' tori

We've seen yesterday that we may assume $\underline{\alpha} \in V^0$ and $\underline{\alpha} \notin (\mathbb{G}_m^n)_{\text{tors}}$.

Let $T = \{(\zeta_1 t^{a_1}, \dots, \zeta_n t^{a_n}) \mid t \in \mathbb{G}_m\}$, where $\zeta \in (\mathbb{G}_m)^n_{\text{tors}}$ and $(a_1, \dots, a_n)$ is primitive. Recall that the hypothesis is $\underline{\alpha} \in W \cap T$, $W \subseteq V$, $\operatorname{codim} W \geq 2$.

Write $\Lambda = \langle \underline{a} \rangle \subseteq \mathbb{Z}^n$. We have $\operatorname{vol}(\Lambda^\vee) = \operatorname{vol}(\Lambda) \ll \|a\|$, and by Minkowski there are vectors $u^{(1)}, \dots, u^{(n-1)} \in \Lambda^\vee$, linearly independent, such that

$$\|u^{(1)}\| \cdots \|u^{(n-1)}\| \ll \|a\|.$$

We may assume $\|u^{(1)}\| \leq \cdots \leq \|u^{(n-1)}\|$.

Let T_2 be the torus defined by

$$\underline{x}^{u^{(1)}} = \cdots = \underline{x}^{u^{(n-2)}} = \underline{\omega},$$

where $\underline{\omega}$ can be chosen to be defined over $\mathbb{Q}(\underline{\zeta})$.

Since $u^{(1)}, \dots, u^{(n-2)}$ are in $\Lambda^\vee$, $T \subseteq T_2$, and moreover

$$\deg T_2 \ll \|u^{(1)}\| \cdots \|u^{(n-2)}\| \ll \|a\|^{(n-2)/(n-1)},$$

where the last inequality is obtained by combining

$$\|u^{(1)}\| \cdots \|u^{(n-2)}\| \leq \|u^{(n-1)}\|^{n-2}$$

and

$$\|a\| \gg \|u^{(1)}\| \cdots \|u^{(n-1)}\| \gg (\|u^{(1)}\| \cdots \|u^{(n-2)}\|)^{1+\frac{1}{n-2}}.$$

11.2.3 First case: $\underline{\alpha}$ is an isolated component of $W \cap T_2$

Let K be the field of definition of W, $[K:\mathbb{Q}] = \delta$. Let $E = K(\zeta_1, \ldots, \zeta_n)$. We have that W and T_2 are both defined over E, and for every $\sigma : E(\alpha) \hookrightarrow \mathbb{C}$, $\sigma|_E = \mathrm{id}$, one has $\sigma(\alpha) \in W \cap T_2$. It follows that, setting $D := [E(\underline{\alpha}) : E]$ we have $D \leq \deg(W \cap T_2) \ll \|a\|^{1-\frac{1}{n-1}} \deg(W)$.

By a special case of the Bounded Height Conjecture[12], the fact that

$$\underline{\alpha} \in V^0 \cap T \subseteq V^0 \cap \bigcup \{\text{torsion cosets of dimension } 1\}$$

implies $h(\underline{\alpha}) \ll 1$.

We also have a lower bound on the height of $\underline{\alpha}$. Write $\underline{\alpha} = (\zeta_1 \xi^{a_1}, \ldots, \zeta_n \xi^{a_n})$, where ξ is not a torsion point (=root of unity). Then we have

$$h(\underline{\alpha}) \gg \sum_{i=1}^{n} h(\underline{\alpha}_i) = \sum_i |a_i| h(\xi) \gg \|a\| h(\xi).$$

On the other hand, the fact that a is primitive implies $K(\xi_1, \ldots, \xi_n, \underline{\alpha}) = K(\zeta_1, \ldots, \zeta_n, \xi)$, from which we obtain

$$[E(\underline{\alpha}) : E] = [E(\xi) : E].$$

Thus we have an inequality of the form $[E(\xi) : \mathbb{Q}(\xi_1, \ldots, \xi_n)] \leq \delta D$. We would like to apply Dobrowolski, but unfortunately this would involve the degree of ξ over $\mathbb{Q}$, which depends on the roots of unity $\zeta_1, \ldots, \zeta_n$. A finer version of Dobrowolski due to Amoroso-Zannier yields $h(\xi) \gg_{\varepsilon,\delta} D^{-1-\varepsilon}$ (the point here is that one controls the degree of ξ over some cyclotomic extension).

Putting everything together we obtain

$$D^{-1-\varepsilon} \|a\| \ll h(\underline{\alpha}) \ll 1,$$

hence

$$\|a\|^{1/(1+\varepsilon)} \ll D \ll \|a\|^{1-\frac{1}{n-1}} \deg(W),$$

which in particular gives $\|a\|^{\frac{1-\varepsilon'}{n-1}} \ll \deg(W)$, where

$$\varepsilon' = (n-1)\left(1 - \frac{1}{1+\varepsilon}\right) < 1 \quad \text{if } \varepsilon < \frac{1}{n-2}.$$

The conclusion is that the degree of T is limited in terms of $\deg(W)$, hence we may simply choose $T' = T$.

12 07.06.2018 – Conclusion of the proof of theorem 10.2

Recall that we are trying to prove:

Theorem 12.1 (Bombieri-Zannier 1998, Schinzel's conjecture; see theorem 10.2)**.** *Let V be a subvariety of $\mathbb{G}_m^n$ defined over a number field[13] K of degree $\leq \delta$ by polynomials of naïve height at most h_0 and of degree $\leq d_0$. Let W be a $\overline{\mathbb{Q}}$-irreducible component of V of codimension at least 2. Let T be a torsion coset of dimension 1. Suppose there is a point $\underline{\alpha} \in W \cap T$: then there exists a torsion coset T' with $\deg(T') \leq B(V)$ such that $\underline{\alpha} \in T'$.*

So far we have seen the following:

1. we may assume $W = V$: the field of definition of W and its degree can be estimated in terms of those of V;

[12] which is now a theorem of Habegger, proven by Zannier in the special case needed here

[13] in fact, we want all geometrically irreducible components to also be defined over that same number field

2. we may assume that $\underline{\alpha}$ is in V^0 and $\underline{\alpha}$ is not a torsion point;

3. Let $\underline{\alpha} \in V^0$ and let $T = \{(\zeta_1 t^{a_1}, \dots, \zeta_n t^{a_n} \mid t \in \mathbb{G}_m\}$ with $\underline{\zeta} \in (\mathbb{G}_m^n)_{\text{tors}}$ and $\underline{a} \in \mathbb{Z}^n \setminus \{0\}$ primitive. By methods of geometry of numbers we have shown: there exists a torsion coset T_2 of dimension 2, $T_2 \supseteq T$, such that $\deg(T_2) \ll |a|^{1-1/(n-1)}$ (whereas $\deg T \approx |a|$)

4. We have seen how to handle the first case: $\{\underline{\alpha}\}$ is an isolated component of $V \cap T_2$. In this case,
$$\deg(T)^{\frac{1-\varepsilon}{n-1}} \ll_{\varepsilon,\delta,d_0} 1 + h_0.$$
In proving this, we have assumed a case of the Bounded Height Conjecture: $\forall \underline{\alpha} \in V^0 \cap \bigcup\{\text{torsion cosets of dimension } 1\}$, then $h(\underline{\alpha}) \ll_{\delta,d_0} 1 + h_0$. We have also used a relative version of Dobrowolski.

We now consider the second case.

12.1 Anomalous intersections and the structure theorem

Definition 12.2. *Let V be irreducible and let T be a translate*[14]*. We say that an irreducible subvariety Y of $V \cap T$ is* ***anomalous*** *(or more precisely V-anomalous) if $\dim Y > 0$ and $\dim Y > \dim V - \operatorname{codim}(T)$.*

Remark 12.3. This is not the same as atypical, because we are considering all translates and not just torsion cosets. Notice that Y atypical of positive dimension implies Y anomalous. Furthermore, every translate Y of positive dimension which is contained in V is anomalous.

Definition 12.4. *We set $V^{0,a} = V \setminus \bigcup\{\text{anomalous subvarieties}\} \subset V^0$.*

Theorem 12.5 (Corollary of the structure theorem)**.** *Any V-anomalous varieties are contained in Hg, where H is a proper subtorus of $\mathbb{G}_m^n$ of degree controlled by $\deg(V)$ and g is a point of $\mathbb{G}_m^n$ (not a torsion point, and on whose height we say nothing).*

Coming back to the proof of theorem 10.2, suppose that $\underline{\alpha}$ belongs to a component Y of dimension ≥ 1 of $V \cap T_2$. Then $\dim Y \geq 1 > 2 - 2 \geq \dim(T_2) - \operatorname{codim}(V)$, so Y is anomalous. By theorem 12.5, there exists a translate $T' = Hg$ with $Y \subseteq T'$ and $\deg T' \ll_{n,d_0} 1$. Everything is going (almost) smoothly: we would be done if we knew that g is a torsion point.

Notice that $Y \subseteq T_2 \cap T'$, which is a union of translates. Let K be an irreducible component of $T_2 \cap T'$ that contains Y; the dimension of K is 1 or 2. If $\dim K = 1$, then $Y = K$ (because $\dim Y = 1$ and K is irreducible), hence $\underline{\alpha} \in K = Y \subseteq V$, and since K is a translate this contradicts the fact that $\underline{\alpha}$ was taken in V^0. Hence K is of dimension 2, and therefore is a component of T_2, and in particular a torsion coset. Thus $K \subseteq T'$ contains a torsion point, and therefore T' is itself a torsion coset. □

12.2 Remarks on the dependence of the result on the height

In case 2 we have found a torsion coset of bounded degree with the additional property that the bound is independent of the height of V. With some care, one obtains the following: suppose $\exists \underline{\alpha} \in W \cap T$, where $\underline{\alpha} \in V^0$ is not a torsion point. Then:

1. either $\deg(T)^{\frac{1-\varepsilon'}{n-1}} \ll_{n,\delta,d_0} 1 + h_0$ (so T itself has small height,

2. or there exists a torsion coset T' with $\deg(T') \ll_{n,d_0} 1$ (that is, the degree of T' is bounded also uniformly in the height)

Notice however that we have the assumption $\underline{\alpha} \in V^0$ (and not torsion), and that this condition is necessary, as shown by the following example.

[14] not necessarily by a torsion point

Example 12.6 ($\alpha \in V^0$ is necessary for uniformity)**.** Let $V = \{(2, 2^a)\} \times \mathbb{G}_m \subset \mathbb{G}_m^3$ and notice that $V^0 = \emptyset$. The point $\underline{\alpha} = \left(2, 2^a, 2^b\right)$ belongs to $V \cap T$, where $T = \{(t, t^a, t^b) : t \in \mathbb{G}_m\}$. The degree of T is essentially $\max\{|a|, |b|\}$, so the first conclusion cannot hold ($h_0 \approx \log a$, so $\deg(T)$ is certainly not bounded uniformly in terms of h_0 when $b \to \infty$). On the other hand, $\underline{\alpha}$ does not belong to a torus of height bounded uniformly in both a and b (when $a, b \to \infty$), so the second conclusion also does not hold.

Remark 12.7. The problem in the previous example is that when one make the reduction to the case $\underline{\alpha} \in V^0$ the coordinate 2^b of the torus simply disappears. This, however, is a purely technical problem, and we'll say later how to go around it.

12.3 Application to the study of multiple roots of lacunary polynomials

We have already discussed the following theorem:

Theorem 12.8. *Let $F, G \in \mathbb{Z}[x_1, \ldots, x_n]$ be linear polynomials, and let $\underline{a} \in \mathbb{Z}^n \setminus \{0\}$ a primitive vector. Assume that $|a|^{1/2(n-1)} > c(n)(1 + \max\{\log|f_i|, \log|g_i|\}$. Set*

$$f_{\underline{a}}(t) = F(t^{\underline{a}}) = f_0 + f_1 t^{a_1} + \cdots + f_n t^{a_n}$$

and

$$g_{\underline{a}}(t) = G(t^{\underline{a}}) = g_0 + g_1 t^{a_1} + \cdots + g_n t^{a_n}.$$

There exist morphisms $\rho : \mathbb{G}_m^n \to \mathbb{G}_m^{n-k}$ and $\psi : \mathbb{G}_m^{n-k} \to \mathbb{G}_m$, of size bounded by a function of n alone, such that setting $\varphi_{\underline{a}}(t) = t^{\underline{a}}$ and $\varphi'_{\underline{a}} = \rho \circ \varphi_{\underline{a}}$ one has

1. *$\varphi_{\underline{a}} = \psi \circ \varphi'_{\underline{a}}$*
2. *setting $P = \gcd(F \circ \psi, G \circ \psi)$ and $h = P \circ \varphi'_{\underline{a}}$, we have that if*

$$\frac{\gcd(f_{\underline{a}}, g_{\underline{a}})}{h}$$

vanishes on ξ, then:

  (a) either ξ is a root of unity

  (b) or there exists a proper subset Λ of $\{0, \ldots, n\}$ such that

$$\sum_{i \in \Lambda} f_i \xi^{a_i} = \sum_{i \in \Lambda} g_i \xi^{a_i} = 0. \quad (\star)$$

Remark 12.9. $(\star)$ might happen because it is possible that *some* coefficients of $\underline{a}$ stay small even though $|a| \to \infty$.

Corollary 12.10 (Multiple roots of lacunary polynomials)**.** *With the same setting, there exist a finite number of ρ, ψ as above, each of size at most $c(n, F)$, such that if ξ is a multiple root of $f_{\underline{a}}(t)$ and is not a root of unity, then there exists a polynomial P such that $P^2 \mid F \circ \psi$ and ξ is a root of $\underline{P} \circ \underline{P} \circ \varphi'_{\underline{a}}$.*

Proof. The idea is of course to use the derivative criterion for multiple roots (more precisely, if ξ is a double root of $f_{\underline{a}}$, then ξ is a root of both $f_{\underline{a}}$ and $tf'_{\underline{a}}$).

Given $F(\underline{x}) = f_0 + f_1 x_1 + \ldots + f_n x_n$, we take $G(\underline{x}) = f_1 a_1 x_1 + \ldots + f_n a_n x_n$. The only problem is that the subgroup (that is, $\underline{a}$) shows up in the coefficients.

Suppose, for ease of exposition, that $(\star)$ is not satisfied. We may then apply the previous theorem (unless $\underline{a}$ is very small, the inequality $|a|^{1/2(n-1)} > c(n)(1 + \max\{\log|f_i|, \log|g_i|\}$ will be satisfied): after a bounded change of variables, I have new polynomials (still denoted by F, G)

$$F(\underline{x}) = f_0 + f_1 y^{\underline{b}_1} + \cdots + f_n y^{\underline{b_n}}$$

$$G(\underline{x}) = a_1 f_1 y^{\underline{b}_1} + \cdots + f_n a_n y^{\underline{b_n}}$$

where $\underline{b}_1, \dots, \underline{b}_n \in \mathbb{Z}^{n-k}$ and $\underline{y} = (y_1, \dots, y_{n-k})$. Equivalently,

$$f_{\underline{a}} = F(t^{\underline{a}'}), \quad tg_{\underline{a}} = G(t^{\underline{a}'})$$

where $\underline{a} = B\underline{a}'$ for some integral matrix $B \in \mathrm{Mat}_{n,n-k}(\mathbb{Z})$. Now ξ is a root of $h(t) = P(t^{\underline{a}'})$, where $P = \gcd(F, G)$. We would like to be more precise and show that P is a multiple factor of F.

Let $P = \sum P_{\underline{b}} y^{\underline{b}}$, so that $h(t) = \sum P_{\underline{b}} t^{\langle \underline{b}, \underline{a}' \rangle}$, and consider the differential operator

$$\Delta := a_1' y_1 \frac{\partial}{\partial y_1} + \cdots + a_{n-k}' y_{n-k} \frac{\partial}{\partial y_{n-k}}.$$

Remark 12.11. The reason to introduce Δ is that for $\underline{b} \in \mathbb{Z}^{n-k}$ the polynomial $y^{\underline{b}}$ is an eigenvector of Δ corresponding to the eigenvalue $\langle \underline{b}, \underline{a}' \rangle$.

We obtain

$$\Delta F = \sum f_i \langle \underline{b}_i, \underline{a}' \rangle y^{\underline{b}_i} = \sum f_i a_i y^{\underline{b}_i} = G.$$

Hence (since we have $P \mid F$ and $P \mid \Delta F$) we obtain that either $P^2 \mid F$ or P is an eigenvector for Δ. But in the latter case $\lambda P = \Delta P = \sum P_{\underline{b}} \langle \underline{b}, \underline{a}' \rangle y^{\underline{b}}$, which implies that $\langle \underline{b}, \underline{a}' \rangle$ is constant, hence $h(t)$ is a monomial, contradiction (a monomial does not have roots in $\mathbb{G}_m$). We deduce as desired that $P^2 \mid F$. □

13 12.06.2018 – Tools used in the proof of theorem 10.2

Let $V \subseteq \mathbb{G}_m^n$ be a $\overline{\mathbb{Q}}$-irreducible subvariety of dimension r. In proving theorem 10.2 we assumed the following two results:

1. Bounded height: let $\underline{\alpha} \in V^0 \cap T$ with T of codimension 1: then $h(\underline{\alpha}) \ll 1$
2. Given Y a V-anomalous component, there exists a translate T which contains Y and which satisfies $\deg T \ll_n (\deg V)^{\mu(n,r)}$

Today we start working towards the proof of these results. Some notation:

Notation 13.1. *1. $V \subseteq \mathbb{G}_m^n$ is a $\overline{\mathbb{Q}}$-irreducible subvariety of dimension r*

*2. $Y \subseteq V$ is a **maximal** V-anomalous subvariety: Y is contained in a translate $H\underline{g}$, where H is a torus of codimension h, $\underline{g}$ is a point in $\mathbb{G}_m^n$, and $s := \dim Y > \dim V - \mathrm{codim}(H)$, that is, $s > r - h$.*

Remark 13.2. 1. V is contained in a (proper) translate if and only if V is V-anomalous

2. if Y is V-anomalous, we may assume $s = r - h + 1$ (simply eliminate equations from the torus until its dimension is the correct one).

Theorem 13.3 (Special case of the structure theorem)**.** *Let Y be V-anomalous. There exists $H\underline{g}$ containing Y and such that $\deg H \ll (\deg V)^{\mu(n,r)}$.*

Lemma 13.4 (Lemma 1)**.** *The theorem is true if $Y = V$.*

Sketch of proof. We know that V is contained in a translate. Let $x_1, \dots, x_n \in \mathbb{C}(V)$ be multiplicately independent modulo constants. Let $y_1, \dots, y_r$ be a transcendence basis given by sufficiently generic linear combinations of $x_1, \dots, x_n$.

$$\begin{tikzcd}
\mathbb{C}(V) \arrow[d, dash, "\Delta:=\deg(V)"'] \\
\mathbb{C}(y_1, \dots, y_r) \arrow[d, dash, "\text{purely transcendental}"'] \\
\mathbb{C}
\end{tikzcd}$$

On the field $\mathbb{C}(y_1, \dots, y_r)$ we have a set of nonarchimedean valuations that satisfy the product formula. Writing $R = P_1^{\ell_1} \cdots P_k^{\ell_k} \in \mathbb{C}[\underline{y}]$ for the factorisation in irreducibles, the product formula is simply the statement

$$\sum_{i=1}^{k} \ell_i \deg(P_i) + (-\deg(R)) = 0.$$

We define $h(R)$ to be its degree; if R is a rational function, $R = \frac{\prod P_i^{e_i}}{\prod Q_j^{f_j}}$, then $h(R) = \sum f_i \deg(Q_i) + \max\{0, \deg R\}$. The following (trivially) hold:

1. $h(R) = 0 \Leftrightarrow R \in \mathbb{C}$;
2. if $R \notin \mathbb{C}$, then $h(R) \geq 1$.

All these absolute values on $\mathbb{C}(\underline{y})$ can be extended to absolute values on $\mathbb{C}(V)$, thus giving a *geometric* height on $\mathbb{C}(V)$.

Remark 13.5 (Lehmer's conjecture in the geometric setting)**.** For all $\gamma \in \mathbb{C}(V) \setminus \mathbb{C}$ we have $h(\gamma) \geq \Delta^{-1}$

Proof. Let γ be as in the statement, with conjugates $\gamma_1, \dots, \gamma_\Delta$. let σ be an elementary symmetric function in $\gamma_1, \dots, \gamma_\Delta$ such that $\vartheta := \sigma(\gamma_1, \dots, \gamma_\Delta) \notin \mathbb{C}$ (this is possible because γ is not in $\mathbb{C}$: if all the). We obtain

$$1 \leq h(\vartheta) \leq h(\gamma_1) + \cdots + h(\gamma_\Delta),$$

using the fact that $\sigma(\cdots)$ is a sum of monomials of degree at most Δ. The coefficients are irrelevant (they are in $\mathbb{C}$, hence they have height 0), so

$$|\sum c_{\underline{\lambda}} \underline{\gamma}^{\underline{\lambda}}| \leq \max \underline{\gamma}^{\underline{\lambda}} \leq \prod_{j=1}^{\Delta} \max\{1, |\gamma_j|\}^{\lambda_j}.$$

Using $\sum_{j=1}^{\Delta} \lambda_j \leq \Delta$ we obtain the desired estimate. Finally, since $\gamma_1, \dots, \gamma_\Delta$ are conjugate, we obtain

$$1 \leq h(\vartheta) \leq \sum h(\gamma_i) = \Delta h(\gamma).$$

□

We will admit the following lemmas:

Lemma 13.6. $h(x_1), \dots, h(x_n) \leq 1$

In our case, the x_i are multiplicatively dependent on V (modulo constants); using this fact and elementary geometry of numbers, one obtains:

Lemma 13.7. *For every $L \in \mathbb{N}$, there exists $\underline{a} \in \mathbb{Z}^n \setminus \{0\}$, $|a| \leq L$, such that*

$$h(\underline{x}^{\underline{a}}) \leq cL^{-1/(n-1)} \max\{h(x_i)\} \underbrace{\leq}_{\textit{previous lemma}} cL^{-1/(n-1)}.$$

Choosing $L = 1 + [c\Delta^{n-1}]$ in the previous lemma we find an $\underline{a}$ of bounded degree and such that $h(\underline{x}^{\underline{a}}) < \Delta^{-1}$. Since any non-constant function has height at least Δ^{-1}, this implies that $\underline{x}^{\underline{a}}$ is constant on V, hence that V is contained in the translate defined by $\underline{x}^{\underline{a}} = C_V$ for a suitable constant C_V.

□

13.1 Chow forms

Let $\underline{u}^j = (u_0^j, \ldots, u_n^j)$, for $j = 0, \ldots, r$, be $r+1$ 'packets' of $n+1$ variables. Consider the 'generic' linear forms

$$L_j = \sum_{i=0}^{n} ui^j x_i.$$

Let $\mathcal{P} = \mathcal{P}(V) \subseteq \mathbb{C}[x_0, \ldots, x_n]$ be the proper, prime, homogeneous ideal of V. We construct an ideal

$$\tilde{\mathcal{P}} = \left\{ F \in \mathcal{P}[u^0, \ldots, u^r] \;\middle|\; \exists M \geq 1, x_0^M F \in (\mathcal{P}, L_0, \ldots, L_r) \subseteq \mathbb{C}[\underline{x}, \underline{u}] \right\}$$

Theorem 13.8. *$\tilde{\mathcal{P}}$ is a multihomogeneous principal ideal; any generator F is called a* ***Chow form****.*

Remark 13.9. $F(\underline{u}^0, \ldots, \underline{u}^r) = 0$ if and only if $V \cap \{L_0 = \ldots = L_r = 0\} \neq \emptyset$ in $\mathbb{P}^n$.

Let now $S^0, \ldots, S^r$ be variable[15] antisymmetric matrices of size $(n+1) \times (n+1)$.

Remark 13.10. Let $\underline{\alpha} \in \mathbb{P}_n$. Suppose that $\underline{\alpha} S^j \underline{\alpha} = 0$: then the linear forms $S^j \underline{\alpha}$ vanish on $\underline{\alpha}$, and therefore, assuming that $\underline{\alpha}$ lies on V, we have $F(S^0\underline{\alpha}, \ldots, S^r\underline{\alpha}) = 0$.

Consider the generic polynomial[16] $F(S^0\underline{x}, \ldots, S^r\underline{x}) \in \mathbb{C}[\underline{x}][S^0, \ldots, S^r]$ and denote by P_1, ..., P_N all the coefficients of $F(S^0\underline{x}, \ldots, S^r\underline{x})$. Set $\mathcal{P}_1 = (P_1, \ldots, P_N)$.

Proposition 13.11. *$\mathcal{P}$ is the only isolated component of $\mathcal{P}_1$*

Furthermore, $\mathcal{P}$ is naturally endowed with a quasi-basis $P_1, \ldots, P_N$; one has $\deg P_j \leq \Delta(r+1)$. From now on, we dehomogeneise the P_j by taking $x_0 = 1$.

13.2 Jacobian matrices

Define

$$J(V) = \left(\frac{\partial P_i}{\partial x_j} \right)_{\substack{i=1,\ldots,N \\ j=1,\ldots,n}};$$

notice that we have de-homogeneised (there is no x_0). Given h n-tuples of complex numbers $\underline{z}^1, \ldots, \underline{z}^h \in \mathbb{C}^n$, we further define

$$J(\underline{z}^1, \ldots, \underline{z}^h) = \begin{pmatrix} & J(V) & \\ z_1^1/x_1 & \ldots & z_n^1/x_n \\ z_1^2/x_1 & \ldots & z_n^2/x_n \\ & \vdots & \\ z_1^h/x_1 & \ldots & z_n^h/x_n \end{pmatrix}$$

Remark 13.12. These extra rows correspond to the partial derivatives $\frac{\partial x^{\underline{z}}}{\partial x_j} = x^{\underline{z}} \frac{z_j}{x_j}$. Therefore $J(\underline{z}^1, \ldots, \underline{z}^h)$ corresponds to the Jacobian matrix of the intersection $V \cap$ translate.

Notation 13.13. *Given a matrix M with coefficients in $\mathbb{C}(x_1, \ldots, x_n)$ and $Y \subseteq V$, we write* $\operatorname{rank}_Y(M)$ *for the rank of M considered as a matrix in $\mathbb{C}(Y)$. In other words, we replace the x_i (considered as variables) with the x_i (considered as functions on Y).*

Remark 13.14. Notice that the Jacobian matrix of a translate depends only on the underlying torus and not on the point we translate by.

[15] that is: all the coefficients under the diagonal are free variables, all the coefficients *on* the diagonal are 0, and all those above the diagonal are the opposite of the corresponding free variables

[16] this is a polynomial in the coefficients of $\underline{x}$ and in all the variables that are the coefficients of the S^j

13.3 More on the proof of theorem 13.3

Recall our setting: Y is an s-dimensional subvariety of $\mathbb{G}_m^n$ contained in a translate $H\underline{g}$. Suppose H is given by $\underline{x}^{\underline{a}^1} = \cdots = \underline{x}^{\underline{a}^h} = 1$, and let $\varphi_i = \underline{x}^{\underline{a}^i}$ (considered as a function on Y). Let

$$\begin{array}{cccc} \varphi: & \mathbb{G}_m^n & \to & \mathbb{G}_m^h \\ & \underline{x} & \mapsto & (\varphi_1, \ldots, \varphi_h) \end{array}$$

and write w for $\varphi(\underline{g}) = \varphi(H\underline{g})$.

Lemma 13.15 (Lemma 2)**.**

$$\operatorname{rank}_Y J(\underline{a}^1, \ldots, \underline{a}^h, V) \leq n - r + h - 1 = n - s.$$

Proof. $H\underline{g}$ is given by the equations $\varphi_i = w_i$. It follows that

$$P_1, \ldots, P_N, P_{N+1} := \varphi_1 - w_1, \ldots, P_{N+h} := \varphi_h - w_h \in I(Y)$$

since $Y \subseteq V \cap H\underline{g}$. By the Jacobian criterion,

$$\operatorname{rank}_Y \left(\frac{\partial P_i}{\partial x_j} \right)_{\substack{i=1,\ldots,N+h \\ j=1,\ldots,n}} \leq n - \dim Y = n - s$$

Now using remark 13.12 we obtain

$$\left(\frac{\partial P_i}{\partial x_j} \right)_{\substack{i=1,\ldots,N+h \\ j=1,\ldots,n}} = J(\underline{a}^1, \ldots, \underline{a}^h; V),$$

and the lemma follows: notice that the rows we add differ from $\underline{x}^{\underline{z}} \frac{z_j}{x_j}$ by $\underline{x}^{\underline{z}}$, which are invertible functions on all of $\mathbb{G}_m^n$. □

Lemma 13.16 (Lemma 3)**.** *Suppose* $\operatorname{rank}_V J(\underline{a}^1, \cdots, \underline{a}^h; V) \leq n - r + h - 1$. *Then* $\varphi_1, \ldots, \varphi_h$ *are* ***algebraically*** *dependent on* V*.*

Proof. There are at least $N + h - (n - r + h - 1) = N - n + r + 1$ linear relations among the rows of $J(\underline{a}^1, \ldots, \underline{a}^h; V)$ with coefficients in $\mathbb{C}(V)$. Each such relation is of the form

$$\gamma_1 \frac{\partial P_1}{\partial x_j} + \cdots + \gamma_1 N \frac{\partial P_N}{\partial x_j} + \gamma_{N+1} \frac{a_j^{(1)}}{x_j} + \cdots + \gamma_{N+h} \frac{a_j^h}{x_j} = 0, \quad j = 1, \ldots, n. \tag{2}$$

Suppose by contradiction that $\varphi_1, \ldots, \varphi_h$ are algebraically independent over $\mathbb{C}$. This implies that the derivations $\frac{\partial}{\partial \varphi_k}$ (which are naturally defined on the transcendental extension $\mathbb{C}(\varphi_1, \ldots, \varphi_h)$ contained in $\mathbb{C}(V)$) extend to $\mathbb{C}(V)$: we now use this fact to prove the following lemma.

Lemma 13.17. *We have*

$$\gamma_1 \frac{\partial P_1}{\partial x_j} + \cdots + \gamma_N \frac{\partial P_N}{\partial x_j} = 0 \text{ on } V,$$

and this for every $j = 1, \ldots, n$*.*

Proof. We have $P_i(x_1, \ldots, x_n) \equiv 0$ on V since the P_i belong to the ideal of V, so on V we have the equation

$$0 = \frac{\partial P_i(x)}{\partial \varphi_\ell} = \sum_{j=1}^n \frac{\partial P_i}{\partial x_j} \frac{\partial x_j}{\partial \varphi_\ell}.$$

From equation (2), multiplying by $\frac{\partial x_j}{\partial \varphi_\ell}$ and summing over j,

$$0 = \gamma_{N+1} \sum_{j=1}^n \frac{a_j^{(i)}}{x_j} \frac{\partial x_j}{\partial \varphi_\ell} + \cdots + \gamma_{N+h} \sum_{j=1}^n \frac{a_j^{(h)}}{x_j} \frac{\partial x_j}{\partial \varphi_\ell} \quad \ell = 1, \ldots, h$$

On the other hand,

$$\delta_{i,\ell} = \frac{\partial \varphi_i}{\partial \varphi_\ell} = \underline{x}^{\underline{a}_i} \sum_{j=1}^{n} \frac{a_j^{(i)}}{x_j} \frac{\partial x_j}{\partial \varphi_\ell} = \varphi_i \sum_{j=1}^{n} \frac{a_j^{(i)}}{x_j} \frac{\partial x_j}{\partial \varphi_\ell};$$

replacing $\sum_{j=1}^{n} \frac{a_j^{(i)}}{x_j} \frac{\partial x_j}{\partial \varphi_\ell}$ with $\delta_{i,\ell}\varphi_i^{-1}$ in the previous equation we obtain $0 = \gamma_{N+\ell}\varphi_\ell^{-1}$ for $\ell = 1, \ldots, n$, which is what we wanted to show.

□

These $N - n + r + 1$ independent relations remain independent, hence $\operatorname{rank} J(V) \leq N - (N - n + r + 1) = n - r - 1$. Given that $P_1, \ldots, P_n$ is a quasi-basis of $\mathcal{P}$, by the Jacobian criterion we get a contradiction.

More precisely: if $\mathcal{P}_1 = \mathcal{P}$, this is really a contradiction with the Jacobian criterion. Otherwise, $\mathcal{P}$ is the only isolated component of $\mathcal{P}_1$, so there exists a $Q \notin \mathcal{P}$ such that $\mathcal{P} \subseteq Q^{-1}\mathcal{P}_1$, and the same argument applies. □

Lemma 13.18 (Lemma 4). *Let, as before, $Y \subseteq V$ be anomalous of dimension s. Let $Y \subseteq H\underline{g}$, where $H = \ker \varphi$ is of codimension h and*

$$\begin{array}{cccc} \varphi: & \mathbb{G}_m^n & \to & \mathbb{G}_m^h \\ & \underline{x} & \mapsto & (\underline{x}^{\underline{a}_1}, \ldots, \underline{x}^{\underline{a}_h}). \end{array}$$

As before, denote by φ_i the function $\underline{x}^{\underline{a}_i}$. Suppose that the following hold:

1. $Y \nsubseteq V_{\text{sing}}$
2. $\operatorname{trdeg} \mathbb{C}(\varphi_1, \ldots, \varphi_h) = h - 1$
3. *w is not a singular point of $\varphi(V)$.*

Then there exists $\underline{z} \in \mathbb{C}^n$, $\underline{z} \neq \underline{0}$, such that

$$\operatorname{rank}_Y J(\underline{z}; V) \leq n - r.$$

Proof. We start by remarking that $\varphi(V)$ is of dimension $h-1$ (by the assumption on the transcendence degree), hence $\varphi(V)$ is a hypersurface in $\mathbb{G}_m^h$. Classical results imply that there are derivations $\delta_1, \ldots, \delta_r$ on $\mathbb{C}(V)$ such that $\delta_\ell(x_j)$ is regular outside of V_{sing} and the vectors $(\delta_\ell(x_1), \ldots, \delta_\ell(x_n))$ (for $\ell = 1, \ldots, r$) are linearly independent over $\mathbb{C}(V)$.

Let $F(y_1, \ldots, y_h) = 0$ be the equation of $\varphi(V)$ in $\mathbb{G}_m^h$. We have $F(\varphi_1, \ldots, \varphi_h) \equiv 0$ on V, hence

$$0 = \delta_\ell \left(F(\varphi_1, \ldots, \varphi_h) \right) = \sum_{i=1}^{h} F_i(\underline{\varphi}) \delta_\ell(\varphi_i),$$

where $F_i = \dfrac{\partial F}{\partial x_i}$. On the other hand,

$$\delta_\ell(\varphi_i) = \varphi_i \sum_{j=1}^{n} \frac{a_j^{(i)}}{x_j} \delta_\ell(x_j).$$

It follows, as in the proof of the previous lemma,

$$\sum_{j=1}^{n} \sum_{i=1}^{h} F_i(\underline{\varphi}) \varphi_i \frac{a_j^{(i)}}{x_j} \delta_\ell(x_j) = 0 \text{ on } V.$$

Since Y is not contained in the singular locus, we can specialise[17] on Y and, setting $z_j = \sum_{i=1}^h F_i(\underline{w}) w_i a_j^{(i)}$, we obtain the relation

$$\sum_{j=1}^n \frac{z_j}{x_j} \delta_\ell(x_j) = 0 \quad \text{on } Y.$$

As for the Chow equations of V we have

$$0 = \delta_\ell(P(x_1, \dots, x_n)) = \sum_{j=1}^n \frac{\partial P_i}{\partial x_j} \delta_\ell(x_i) \quad \text{on } V, \text{ hence on } Y.$$

Combining the previous relations we obtain

$$J(\underline{z}; V)\ ^t\, (\delta_\ell(x_1), \dots, \delta_\ell(x_n)) = 0, \quad \ell = 1, \dots, r \text{ on } Y$$

Hence we have found r independent vectors in the kernel of $J(\underline{z}; V)$, which is therefore of rank at most $n - r$. It remains only to see that $\underline{z}$ is not the zero vector.

This follows combining three remarks: one is that $\underline{a}^{(1)}, \dots, \underline{a}^{(h)} \in \mathbb{Z}^n$ are linearly independent over $\mathbb{Q}$, hence also over $\mathbb{C}$. The second is that $w_i \neq 0$ for $i = 1, \dots, h$ (since $w \in \mathbb{G}_m^h$). Finally, there is an index i such that $F_i(\underline{w}) \neq 0$, because $\underline{w}$ is nonsingular in $\varphi(V)$. □

14 13.06.2018 – Proof of (the special case we need of) the structure theorem

We are in the process of proving the structure theorem; so far, we've established the following lemmas:

Lemma 14.1 (Lemma 1, cf. lemma 13.4)**.** *The theorem is true if* $Y = V$.

Lemma 14.2 (Lemma 2, cf. lemma 13.15)**.**

$$\operatorname{rank}_V J(\underline{a}^1, \dots, \underline{a}^h; V) \leq n - r + h - 1 = n - s.$$

Lemma 14.3 (Lemma 3, cf. lemma 13.16)**.** *Suppose* $\operatorname{rank}_V J(\underline{a}^1, \dots, \underline{a}^h; V) \leq n - r + h - 1$. *Then* $\varphi_1, \dots, \varphi_h$ *are algebraically dependent on* V.

Lemma 14.4 (Lemma 4, cf. lemma 13.18)**.** *Suppose* $Y \not\subseteq V_{\text{sing}}$ *and that* $\underline{w}$ *is not singular in* $\varphi(V)$. *If*

$$\operatorname{trdeg}_V \mathbb{C}(\varphi_1, \dots, \varphi_h) = h - 1,$$

then there exists $\underline{z} \in \mathbb{C}^n \setminus \{0\}$ *such that* $\operatorname{rank} J_Y(\underline{z}; V) \leq n - r$.

Today we complete the proof, using yet another lemma:

Lemma 14.5 (Lemma 5)**.** *If* $\exists \underline{z} \in \mathbb{C}^n \setminus \{0\}$ *such that* $\operatorname{rank}_V J(\underline{z}; V) \leq n - r$, *then* V *is* V*-anomalous.*

For the proof, see section 15.1.

Proof of the structure theorem. By induction on $r = \dim V$. Let $\Delta := \deg(V)$.

By lemma 14.1 we can assume $Y \subsetneq V$, hence $s \leq r - 1$ and in particular $r \geq 2$. Furthermore, by lemma 14.2 we have $\operatorname{rank}_Y J(\underline{a}^1, \dots, \underline{a}^h; V) \leq n - r + h - 1 = n - s$.

Suppose $\operatorname{rank}_V J(\underline{a}^1, \dots, \underline{a}^h; V) > n - r + h - 1$. Then there exists a minor F of $J(\underline{a}^1, \dots, \underline{a}^h; V)$, of size at least $\geq n - r + h$, which is not identically 0 on V but is identically 0 on Y. Let V' be an

[17] recall that w_i is the i-th coordinate of the image of V

irreducible component of $V \cap \{F = 0\}$ containing Y. Now $\dim V' = r-1$ and Y is (even more) anomalous in V'. Moreover, by Bézout we have

$$\deg(V') \leq \deg(F)\deg(V) \ll \Delta^2.$$

By induction, we can choose $\mu(n,r) = 2\mu(n, r-1)$.

Suppose instead $\operatorname{rank}_V J(\underline{a}^1, \ldots, \underline{a}^h; V) \leq n - r + h - 1$. By lemma 14.3, $\varphi_1, \ldots, \varphi_h$ are algebraically dependent on V, hence $\dim \varphi(V) \leq h-1$ (and in fact equality holds[18]). Let $W := \varphi(V)$; it is a hypersurface of $\mathbb{G}_m^h$. By the usual theorem on the dimension of the fibers of morphisms, there is a (Zariski) open dense subset U of W such that $\forall \underline{u} \in U$ and for every component Z of $V \cap \varphi^{-1}(\underline{u})$ we have

$$\dim Z = \dim V - \dim W = r - (h-1) = s.$$

By shrinking U if necessary, we may further assume $W_{\text{sing}} \cap U = \emptyset$. We want to apply Lemma 14.4 with $Y \leftarrow Z$.

Suppose for now that $Z \not\subseteq V_{\text{sing}}$. Then by Lemma 14.4 there exists $\underline{z} \in \mathbb{C}^n, \underline{z} \neq 0$, such that $\operatorname{rank}_{Z_1} J(\underline{z}; V) \leq n - r$. On the other hand, Lemma 14.5, combined with the fact that V is not V-anomalous[19], yields that $\operatorname{rank}_V J(z; V) > n - r$. Proceeding as above, we find that there exists a minor F of $J(\underline{z}; V)$, of rank $n-r$ and not identically 0 on V, which is $\equiv 0$ on Z.

If instead $Z \subseteq V_{\text{sing}}$, then I just fix an equation $F \equiv 0$ on V_{sing} but not identically zero on V (such an equation can be chosen of degree $\ll \Delta$ for the argument using Chow forms).

Hence, in any case, we have an equation $F \equiv 0$ on Z which is not identically zero on V and whose degree is $\ll \Delta$. We apply Bézout again: choose V', a component of $V \cap \{F = 0\}$ containing Z. The dimension of V' is $r-1$, and Z is anomalous in V'. By induction, $\exists \underline{a}_Z \in \mathbb{Z}^n \setminus \{0\}$ such that $|\underline{a}_Z| \ll \Delta^{2\mu(n,r-1)}$ and $\underline{x}^{\underline{a}_Z}$ is constant on Z.

Now the problem is that we want to choose an $\underline{a}$ which works for 'almost all' the Z. The union of the Z that we are interested in is $V \cap \varphi^{-1}(U)$, which is dense in V. Now $\underline{a}_Z$ depends on Z, and there is no changing this fact; however, $|a_Z|$ is bounded *uniformly in* Z, hence – by pigeonhole – there exists $\underline{a} \in \mathbb{Z} \setminus \{0\}$ (with the same bound $|\underline{a}| \ll \Delta^{2\mu(n,r-1)}$) such that $\varphi_{h+1} := x^{\underline{a}}$ is constant for any $Z \subseteq \Omega$, where Ω is dense in $V \cap \varphi^{-1}(U)$, and therefore in V.

Remark 14.6. Notice that this is not the same as saying that $x^{\underline{a}}$ is constant on Ω! Otherwise (by density) it would be constant on V, which is in general not possible because V is not (in general) contained in a translate.

Fact. Recall that $\varphi_1, \ldots, \varphi_h$ are algebraically dependent on V. We now claim that φ_{h+1} is also algebraically dependent from $\varphi_1, \ldots, \varphi_h$.

Proof of the fact. We may assume that $\varphi_2, \ldots, \varphi_h$ are algebraically independent (the transcendence degree is $h-1$). Suppose by contradiction that $\varphi_2, \ldots, \varphi_h, \varphi_{h+1}$ are still algebraically independent. Using the converse of lemma 14.3, we obtain

$$\operatorname{rank}_V J(\underline{a}^2, \ldots, \underline{a}^{h+1}; V) > n - r + h - 1.$$

Therefore there exists a minor F of this matrix, of size a least $n-r+h$, such that F is not identically zero on V. It follows that there exists $\underline{t} \in \Omega$ such that $F(\underline{t}) \neq 0$. But $\underline{t} \in Z$ for one of the 'good' Z, hence φ_{h+1} is constant on Z. It follows that $\underline{t} \in Z$ belongs to a translate of $\varphi_2 = \ldots = \varphi_{n+1} = 1$. By lemma 14.2 we obtain

$$\operatorname{rank}_Z J(\underline{a}^2, \ldots, \underline{a}^{h+1}; V) \leq n - r + h - 1,$$

but F is a minor (of size $n-r+h-1$) of this same matrix, which implies $F \equiv 0$ on Z, contradiction (notice that $\underline{t}$ belongs to Z and $F(\underline{t}) \neq 0$). □

[18] this follows from the maximality of Y in the following way. If $\dim \varphi(V) < h-1$, then every component of $V \cap \varphi^{-1}(\underline{w})$ would have dimension at least $r - \dim(V) > r - (h-1) = s$. Now each of these components is V-anomalous, and one contains Y, contradiction

[19] otherwise we would have $Y = V$ by maximality

We are now in a position to finish the proof of the theorem.

$$\begin{tikzcd} \mathbb{G}_m^n \arrow[d, "\varphi"'] & Y \subseteq H\underline{g} \arrow[d] \\ \mathbb{G}_m^h & \underline{w} \end{tikzcd}$$

Remark 14.7. Y is a component of $V \cap \varphi^{-1}(\underline{w})$, for otherwise Y would not be maximal.

Write $V \cap \varphi^{-1}(\underline{w}) = Y \cup V_0$, where V_0 is the union of the other irreducible components. Take $\underline{y} \in Y \setminus V_0$ and set $w_{h+1} = \varphi_{h+1}(\underline{y})$. Also denote by $\tilde{\varphi}$ the morphism $(\varphi_1, \dots, \varphi_{h+1}) : \mathbb{G}_m^n \to \mathbb{G}_m^{h+1}$.

With this notation, $\underline{y} \in V \cap \tilde{\varphi}^{-1}(\underline{\tilde{w}}) \subseteq V \cap \varphi^{-1}(\underline{w})$. Hence $\underline{y}$ belongs to $\tilde{Y}$, irreducible component of $V \cap \tilde{\varphi}^{-1}(\tilde{w})$. Since $\tilde{Y}$ is contained in one of the irreducible components of $V \cap \varphi^{-1}(\underline{w})$ and $\underline{y}$ does not belong to V_0 (which is the union of all the components different from Y), we obtain $\tilde{Y} \subseteq Y$. We want to show that they are equal, which will imply the theorem.

Our Fact, combined with the algebraic dependence of $\varphi_1, \dots, \varphi_h$ on V, implies $\dim \tilde{\varphi}(V) \leq h-1$. By the fiber dimension theorem,

$$\dim \tilde{Y} \geq r - (h-1) = s = \dim Y,$$

which implies $Y = \tilde{Y}$ as desired. □

15 14.06.2018 – Bounded height conjecture and conclusion

15.1 Schanuel, Ax, and the proof of lemma 14.5

Schanuel's classical conjecture says the following:

$$(S_1)\ \alpha_1, \dots, \alpha_n \in \mathbb{C} \quad \mathbb{Q}-\text{linearly independent} \Rightarrow \operatorname{trdeg}_{\mathbb{Q}} \mathbb{Q}(\underline{\alpha}, \exp(\underline{\alpha})) \geq n.$$

Schanuel himself has formulated a geometric version of his conjecture:

$$(S_2)\ y_1, \dots, y_n \in t\mathbb{C}[[t]] \quad \mathbb{Q}-\text{linearly independent} \Rightarrow \operatorname{trdeg}_{\mathbb{C}(t)} \mathbb{C}(t, \underline{y}, \exp(\underline{y})) \geq n.$$

There is also a multi-dimensional version due to Ax:

$$\text{(Ax)} \qquad \begin{array}{c} y_1, \dots, y_n \in \mathbb{C}[[t_1, \dots, t_m]] \quad \mathbb{Q}\text{-linearly independent} \\ \text{implies} \\ \operatorname{trdeg}_{\mathbb{Q}} \mathbb{Q}(\underline{y}, \exp(\underline{y})) \geq n + \operatorname{rank}\left(\frac{\partial y_i}{\partial t_j}\right) \end{array}$$

Remark 15.1. (Ax) with $m = 0$ implies (S_1), and (S_2) implies (Ax) with $m = 1$ if $y_i(0) = 0$.

Less obviously, we have the following theorems of Ax:

Theorem 15.2 (Ax). *(Ax) is true if we have $y_i(0) = 0$ for $i = 1, \dots, n$*

Theorem 15.3 (Ax). *(Ax) is equivalent to (S_1)*

Lemma 15.4 (Cf. Lemma 14.5). *Suppose that there exists $\underline{z} \in \mathbb{C}^n \setminus \{\underline{0}\}$ such that* $\operatorname{rank}_V J(\underline{z}; V) \leq n-r$. *Then V is V-anomalous, that is, there exists $\underline{a} \in \mathbb{Z}^n \setminus \{\underline{0}\}$ such that $\varphi(\underline{x}) = \underline{x}^{\underline{a}}$ is constant on V.*

Remark 15.5. Notice that in fact $\operatorname{rank}_V J(\underline{z}; V) = n-r$, because $\operatorname{rank}_V J(\underline{z}; V) \geq \operatorname{rank}_V J(V) = n-r$.

Proof of lemma 14.5. $\operatorname{rank}_V J(V) = n - r$, so we have a linear relation (on V)

$$\frac{z_j}{x_j} = \gamma_1 \frac{\partial P_1}{\partial x_j} + \cdots + \gamma_N \frac{\partial P_N}{\partial x_j}$$

Here $\gamma_i \in \mathbb{C}(V)$, and the equality holds for $j = 1, \ldots, n$.

Take $\underline{\alpha} \in V \setminus V_{\text{sing}}$. Let $y_1, \ldots, y_r$ be generic linear polynomials in $x_1, \ldots, x_n$, and assume that the y_i vanish at α (in fact, I think we require that $\underline{\alpha}$ is the only common zero of the y_i, at least in a neighbourhood of $\underline{\alpha}$).

On the one hand, standard manipulations of derivatives[20] yield

$$\sum z_j \frac{1}{x_j} \frac{\partial x_j}{\partial y_\ell} = 0 \tag{3}$$

On the other, let $x_j = \alpha_j \exp(X_j)$, where $X_j \in \mathbb{C}[[y_1, \ldots, y_r]]$ are power series without constant term (we will apply Ax's theorem to the X_j). Rewriting (3) in terms of the X_j, we find

$$\frac{\partial}{\partial y_\ell} \left(z_1 X_1 + \cdots + z_n X_n \right) = 0 \quad \ell = 1, \ldots, r.$$

In particular, $z_1 X_1 + \cdots + z_n X_n$ is locally constant.

Hence $t := \operatorname{trdeg}_{\mathbb{C}} \mathbb{C}\left(X_1, \ldots, X_n, \exp(X_1), \ldots, \exp(X_n)\right) \leq (n-1) + \dim V$: this follows because there is a linear relation among the X_i and $\exp(X_j)$ parametrise V.

Theorem 15.2 implies $t \geq n + r'$, where $r' = \operatorname{rank}\left(\frac{\partial X_j}{\partial y_\ell}\right) = r$ (the last equality should be checked by an explicit computation, which we skip). This is a contradiction, which means that the hypothesis in Ax's theorem must fail. Hence $X_1, \ldots, X_n$ are $\mathbb{Q}$-linearly dependent, so $x_1, \ldots, x_n$ are multiplicatively dependent on V, which is what we want. □

15.2 Bounded height

Proposition 15.6. *$V \subseteq \mathbb{G}_m^n$. Then the points in $V^0 \cap \bigcup\{$torsion cosets of dimension $1\}$ have bounded height.*

Sketch of proof. Write V as $\{f_\ell = 0 : \ell = 1, \ldots, L\}$, where $f_\ell := \sum_{\underline{\lambda} \in I} a_{\ell,\underline{\lambda}} \underline{x}^{\underline{\lambda}} \in \mathbb{Z}[\underline{x}]$ (or in fact even $\overline{\mathbb{Q}}[\underline{x}]$).

Write $\underline{\alpha} = (\zeta_1 \xi_1^{a_1}, \ldots, \zeta_n \xi_n^{a_n}) \in V^0$, where the ζ_i are roots of unity[21] and $a_j \in \mathbb{Z}$. We want to show that $h(\underline{\alpha}) \ll 1$. The fact that $\underline{\alpha}$ is in V^0 implies that there is no sub-sum of f_ℓ that vanishes on $\underline{\alpha}$. This gives the following: if $c_1, \ldots, c_L$ are generic, and if we set $f = \sum_{i=1}^L c_i f_i = \sum_\lambda f_{\underline{\lambda}} \underline{x}^{\underline{\lambda}}$, then

$$\sum f_{\underline{\lambda}} \xi^{\langle \underline{a}, \underline{\lambda} \rangle} = 0$$

and there are no nonzero subsums. Denote by $m_1 > \cdots > m_k$ the integers $\langle \underline{a}, \underline{\lambda} \rangle$ and assume $m_k = 0$ (otherwise we just shift). We now apply the following lemma:

Lemma 15.7. $h(\xi) \ll m_1^{-1}$

The lemma is enough to conclude the proof: we have

$$h(\underline{\alpha}) \ll \max |a_j| h(\xi) \ll \frac{\max |a_j|}{m_1} \ll \frac{\max |a_j|}{\max_{\underline{\lambda}} |\langle \underline{a}, \underline{\lambda} \rangle|};$$

now if the vectors $\underline{\lambda}$ generate $\mathbb{Z}^n$ (or at least a full-rank sublattice) we have that $\max_{\underline{\lambda}} |\langle \underline{a}, \underline{\lambda} \rangle|$ is comparable with $\max |a_j|$, and we are done. If not, one makes a change of coordinaes to reduce to the case of a smaller ambient space. □

[20] using $\frac{\partial X_j}{\partial y_\ell} = \frac{1}{x_j} \frac{\partial x_j}{\partial y_\ell}$

[21] which we take equal to 1 for simplicity

Proof of lemma 15.7. For $r = 1, \dots, k-1$ let

$$\gamma_r = f_1 \xi^{m_1 - m_r} + \cdots + f_{r-1} \xi^{m_{r-1} - m_r} + f_r;$$

each of these expressions is nonzero, because there are no trivial subsums. On the other hand, $\sum_{j=1}^{k} f_j \xi^{m_j - m_r} = 0$, so

$$-\gamma_r = \frac{f_{r+1}}{\xi^{m_r - m_{r+1}}} + \cdots + \frac{f_k}{\xi^{m_r - m_k}}.$$

This imply that for every finite place v we have

$$|\gamma_r|_v \leq \max\{1, |\xi|_v\}^{-(m_r - m_{r+1})} \max_i |f_i|_v$$

while for v infinite

$$|\gamma_r|_v \leq k \max\{1, |\xi|_v\}^{-(m_r - m_{r+1})} \max_i |f_i|_v$$

Applying the product formula over any field that contains all the numbers of interest we obtain

$$0 \leq \log k - (m_r - m_{r+1}) h(\xi) + h(f),$$

or equivalently $(m_r - m_{r+1}) h(\xi) \leq \log k + h(f)$. Summing over $r = 1, \dots, k-1$ we obtain

$$m_1 h(\xi) = (m_1 - m_k) h(\xi) \leq k \log k + k h(f) \ll 1.$$

□

15.3 Complements

15.3.1 Bounded height conjecture, full version

$V \subseteq \mathbb{G}_m^n$ irreducible of dimension r. Let $\rho \geq r$. We say that $Y \subseteq V$ is ρ-anomalous if

1. $\dim Y > 0$
2. Y is contained in a translate $H\underline{g}$, where $\operatorname{codim}(H) = h$ and $s > \rho - h$.

Remark 15.8. In particular, Y is anomalous $\Leftrightarrow$ Y is r-anomalous, and Y is $(n-1)$-anomalous if and only if Y is a translate of positive dimension contained in V. Indeed,

$$s > n - 1 - h \Leftrightarrow s > \dim(H\underline{g}) - 1 \Rightarrow \dim Y \geq \dim(H\underline{g}),$$

and since $Y \subseteq H\underline{g}$ this means $Y = H\underline{g}$.

Set

$$V^{oa,\rho} := V \setminus \bigcup_{Y\,\rho-\text{anomalous}} Y.$$

Remark 15.9. One has $V^{oa,n-1} = V^0$ and $V^{oa} := V^{oa,r} = V \setminus \bigcup\{V - \text{anomalous curves}\}$

Theorem 15.10 (Habegger). *$V^{oa,\rho} \cap \bigcup_{\substack{T \text{ torsion coset} \\ \operatorname{codim} T \geq \rho}} T$ has bounded height.*

Remark 15.11. Theorem 15.10 implies proposition 15.6 by taking $\rho = n-1$.

In particular,

$$V^{oa} \cup \bigcup \{T \mid T \text{ torsion coset of } \operatorname{codim}(T) \geq \dim V\}$$

has bounded height.

15.3.2 Structure theorem, full version

$V \subseteq \mathbb{G}_m^n$ irreducible of dimension $r \geq 1$.

1. $\forall H$ torus of codimension h with $1 \leq h \leq r$, the union Z_H of the $Y \subseteq V \cap H\underline{g}$ (for $\underline{g}$ varying in $\mathbb{G}_m^n$) is a closed subset of $\mathbb{G}_m^n$. Moreover, HZ_H is not dense[22] in $\mathbb{G}_m^n$.
2. There exists a finite set Φ of tori of codimension h ($1 \leq h \leq r$), each of degree $\ll (\deg V)^{2^{r-1}(n-1)}$, such that every V-anomalous Y is a component of $V \cap H\underline{g}$ for some $H \in \Phi$ and $\underline{g} \in Z_H$.

We won't prove this, but we make some remarks:

1. the theorem implies that V^{oa} is open, because $V^{oa} = V \setminus \bigcup_{H \in \Phi} Z_H$
2. $\overline{HZ_H} \neq \mathbb{G}_m^n$. Let $H \in \Phi$, and let

$$\begin{array}{cccc} \phi: & gm^n & \to & \mathbb{G}_m^h \\ & \underline{x} & \mapsto & (\varphi_1, \ldots, \varphi_h) \end{array}$$

We have $Z_H = \{Y \mid \exists \underline{g} \in \mathbb{G}_m^n : Y \subseteq V \cap H\underline{g}, \dim Y = r - h + 1\}$. Given that φ contracts H to a point, $\varphi(Z_H)$ is the set of the images via φ of the points $g \in \mathbb{G}_m^n$ that we are considering. Now $\overline{HZ_H} \neq \mathbb{G}_m^n$ is equivalent to $\overline{\varphi(Z_H)} \neq \mathbb{G}_m^h$.

[22] this seems not to be true for $V = \mathbb{G}_m^n$, but I guess this is the only exception