TERM 1 READING GROUP, 2024-25
Equidistribution
Week 4 - Bogomolov's conjecture for the torus

Edison Au-Yeung

(This document follows Chapter 3 and Section 4.3 of *Heights in Diophantine Geometry* by Bombieri and Gubler, mostly consists of annotations or elaborations of the original text.)

1. Introduction and notation

Bogomolov's conjecture is a statement that about the height of algebraic points on curves, claiming that the number of points of small height on a curve should be finite. In 1992, Zhang proved the case for curves in $\mathbb{G}_m^n$:

Theorem 1.1. *Let X be a closed subvariety of $\mathbb{G}_m^n$ defined over a number field K and let X^* be the complement in X of the union of all torsion cosets $\varepsilon H \subseteq X$. Let $f_i \in K[\mathbf{x}]$ be a set of polynomials of degree at most d defining X. Then the height of points $P \in X^*$ has a positive lower bound, depending on $n, d, [K : \mathbb{Q}]$, and $\max h(f_i)$.*

The original proof of the above statement is effective, meaning the finite set of points can be effectively determined for every $P \in \mathbb{G}_m^n$. On the other hand, there is an alternative proof, which relies on the idea that points of small height under the action of Galois conjugation tend to become equidistributed with respect to a suitable measure. Although this proof is ineffective, we shall follow this approach and use Bilu's Theorem to establish the result.

Throughout the study group, we will use the following notations:

1. The standard height of $\boldsymbol{x} = (x_1, x_2, \cdots, x_n) \in \mathbb{G}_m^n$: $\hat{h}(\boldsymbol{x}) = \sum_{i=1}^n h(x_i)$.
2. Space of complex valued continuous functions on X with compact support: $C_C(X)$.
3. The unit circle in $\mathbb{C}$: $\mathbb{T} = \{e^{i\theta} | 0 \leq \theta \leq 2\pi\}$.
4. $\log^+(\cdot) = \max\{0, \log(\cdot)\}$.

2. Notations and facts about $\mathbb{G}_m^n$

In this section, we will cover some basic facts about $\mathbb{G}_m^n$, mainly for unravelling the terms in the theorem which might look unfamiliar to certain readers. Let K be a field of characteristic 0. We identify $G := \mathbb{G}_m^n$ over a field K as an affine variety with the Zariski open subset

$$\{(x_1, ..., x_n) \in K^n : x_1...x_n \neq 0\}$$

of affine space $\mathbb{A}_K^n$, with multiplication

$$(x_1, x_2, ..., x_n) \cdot (y_1, y_2, ..., y_n) = (x_1y_1, x_2y_2, ..., x_ny_n).$$

The identity element of this group structure is of course, $\mathbf{1}_n = (1, 1, ..., 1)$.

Below are some definitions about the subgroups of G:

- Algebraic subgroup: an algebraic subgroup of G is a Zariski closed subgroup.
- Linear torus: a linear torus H is an algebraic subgroup which is geometrically irreducible.
- Torsion coset: it is a coset gH of a linear torus H of positive dimension. In addition, we call gH a torsion coset if g is a torsion point in G (i.e. g is a point of finite order in G).

There is a particular algebraic subgroup of G that we will use in the proof of Theorem 1.1: let $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{Z}^n$, and for $\boldsymbol{x} = (x_1, x_2, \dots, x_n) \in G$, we write $\boldsymbol{x}^{\boldsymbol{\lambda}} = x_1^{\lambda_1} x_2^{\lambda_2} \dots x_n^{\lambda_n}$. Recall that a subgroup Λ of $\mathbb{Z}^n$ is called a lattice if it is a subgroup of rank n – this determines an algebraic subgroup.

$$H_\Lambda \coloneqq \{\boldsymbol{x} \in G \colon \boldsymbol{x}^{\boldsymbol{\lambda}} = 1 \; \forall \lambda \in \Lambda\}.$$

Proposition 2.1. *For a subgroup Λ of $\mathbb{Z}^n$ of rank $n - r$, the following properties are equivalent:*

(a) H_Λ is a linear torus;

(b) H_Λ is isomorphic to $\mathbb{G}_m^r$;

(c) H_Λ is irreducible.

3. Proof of Bogomolov conjecture – setup

We would like to use Bilu's theorem on equidistribution:

Theorem 3.1. *Let $(\xi_i)_{i \in \mathbb{N}}$ be an infinite sequence of distinct non-zero algebraic numbers such that $h(\xi_i) \to 0$ as $i \to \infty$. Then the sequence $(\delta_{\xi_i})_{i \in \mathbb{N}}$ converges in the weak-* topology to the uniform probability measure $\mu_{\mathbb{T}} := \frac{d\theta}{2\pi}$ on the unit circle $\mathbb{T} := \{e^{i\theta} \mid 0 \le \theta < 2\pi\}$ in $\mathbb{C}$.*

Before we proceed, we should note that to apply Bilu's theorem, we require all the algebraic numbers in the sequence $(\xi_i)_{i \in \mathbb{N}}$ to be *distinct*. However, this is actually unecessary and can be relaxed to two conditions:

1. $h(\xi) \to 0$.
2. **No root of unity** in the sequence $(\xi_i)_{i \in \mathbb{N}}$ is repeated **infinitely often**.

This is because the most important fact in the proof of Bilu's theorem is the degree of the minimal polynomial of ξ_i, d_i, tends to infinity as $i \to \infty$. By Kronecker's theorem, we know that the second condition has to hold.

Theorem 3.2 (Kronecker's theorem)**.** *The height of $\zeta \in \overline{\mathbb{Q}}^\times$ is 0 if and only if ζ is a root of unity.*

We also note that we need a sequence of algebraic numbers in order to apply Bilu's theorem, instead of a sequence of points in variety. To do so, we can consider the associated sequence

$(\chi(\boldsymbol{\xi}_i))_{i\in\mathbb{N}}$, where $\chi(\boldsymbol{x}) = x_1^{m_1} x_2^{m_2} ... x_n^{m_n}$ is a non-trivial character $\chi \colon \mathbb{G}_m^n \to \mathbb{C}$. This is because when a character is defined over $\mathbb{C}$, every such character value is an algebraic integer.

4. Proof of Bogomolov conjecture

This approach proves the theorem by contradiction: suppose we have an infinite sequence of distinct points $\boldsymbol{\xi}_i \in X^*$ with $\hat{h}(\boldsymbol{\xi}_i) \to 0$.

Strategy: Mimick the construction in the proof of Bilu's theorem by defining the probability measure

$$\boldsymbol{\delta}_{\boldsymbol{\xi}} = \frac{1}{[\mathbb{Q}(\boldsymbol{\xi}) : \mathbb{Q}]} \sum_{\sigma : \mathbb{Q}(\boldsymbol{\xi}) \to \mathbb{C}} \delta_{\sigma \boldsymbol{\xi}}$$

associated to the Galois orbit of $\boldsymbol{\xi}$. As in the proof of Bilu's theorem, consider a weak-* limit measure μ of the sequence $(\boldsymbol{\delta}_{\boldsymbol{\xi}_i})_{i\in\mathbb{N}}$ and the same argument should show that μ is supported in $\mathbb{T}^n$. This will then lead to a contradiction related to the density of torsion points in $\mathbb{G}_m^n$.

Now for any non-trivial character $\chi(\boldsymbol{x}) = x_1^{m_1} \cdots x_n^{m_n}$ of $(\mathbb{C}^*)^n$, consider the associated sequence $(\chi(\boldsymbol{\xi}_i))_{i\in\mathbb{N}}$. As a sanity check, we verify that Bilu's theorem is really applicable to the sequence $\{\chi(\boldsymbol{\xi}_i)\}_{i\in\mathbb{N}}$:

$$\begin{aligned} h(\chi(\boldsymbol{\xi}_i)) &= h(x_1^{m_1} \cdots x_n^{m_n}) \\ &\leq (\max_j |m_j|) \hat{h}(\boldsymbol{\xi}_i) \to 0. \end{aligned}$$

As discussed in Section 2, we have two cases due to the hypothesis of Bilu's theorem.

Case I: For every non-trivial character χ, the sequence $(\chi(\boldsymbol{\xi}_i))_{i\in\mathbb{N}}$ ultimately consists of distinct elements.

In this case, we want to show that

$$\int_{\mathbb{T}^n} \chi(\boldsymbol{x}) d\mu = 0$$

since this tells us that μ is a probability measure (again, the limit of a sequence of probability measure need not to be a probability measure).

We prove this as follows. Fix $c > 0$ and let $f \in C_c(\mathbb{C}^*)$ be the identity $f(x) = x$ in the neighborhood $|\log|x|| < c$ of $\mathbb{T}$. Then we construct a function

$$f_\chi(\boldsymbol{x}) = f(x_1^{m_1}) \cdots f(x_n^{m_n})$$

which has compact support in $(\mathbb{C}^*)^n$ and coincides with $\chi(\boldsymbol{x})$ in a neighborhood of $\mathbb{T}^n$. Therefore, we have

$$\int_{\mathbb{T}^n} \chi(\boldsymbol{x}) d\mu = \int_{(\mathbb{C}^*)^n} f_\chi(\boldsymbol{x}) d\mu$$
$$= \lim_{i\to\infty} \frac{1}{[\mathbb{Q}(\boldsymbol{\xi}_i):\mathbb{Q}]} \sum_{\sigma:\mathbb{Q}(\boldsymbol{\xi}_i)\to\mathbb{C}} f_\chi(\sigma\boldsymbol{\xi}_i). \qquad \text{(by weak-*convergence)}$$

We would like to replace $f_\chi(\boldsymbol{x})$ by $\chi(\boldsymbol{x})$ in the last sum. This is because by Bilu's theorem, the measure μ_χ determined by the sequence $(\chi(\boldsymbol{\xi}_i))_{i\in\mathbb{N}}$ (i.e. we are considering the sequence of measures $\{\boldsymbol{\delta}_{\chi(\boldsymbol{\xi})}\}_{i\in\mathbb{N}}$) is the uniform measure on $\mathbb{T} = \chi(\mathbb{T}^n)$. but this step requires justification, because $f(\chi(\boldsymbol{x}))$ is not compactly supported in $(\mathbb{C}^*)^n$ ($f(\chi(\boldsymbol{x}))$ is a monomial). To deal with this problem, we will break up the sum by

$$\frac{1}{[\mathbb{Q}(\boldsymbol{\xi}_i):\mathbb{Q}]} \sum_{\sigma:\mathbb{Q}(\boldsymbol{\xi}_i)\to\mathbb{C}} f_\chi(\sigma\boldsymbol{\xi}_i) = \frac{1}{[\mathbb{Q}(\chi(\boldsymbol{\xi}_i)):\mathbb{Q}]} \sum_{\tau:\mathbb{Q}(\chi(\boldsymbol{\xi}_i))\to\mathbb{C}} f(\tau\chi(\boldsymbol{\xi}_i))$$
$$+ \frac{1}{[\mathbb{Q}(\boldsymbol{\xi}_i):\mathbb{Q}]} \sum_{\sigma:\mathbb{Q}(\boldsymbol{\xi}_i)\to\mathbb{C}} \left(f_\chi(\sigma\boldsymbol{\xi}_i) - f(\chi(\sigma\boldsymbol{\xi}_i))\right).$$

Remark 4.1. *A trivial remark: if we expand the terms a little bit, we would see why the above is an equality.*

- $f(\chi(\sigma\boldsymbol{\xi}_i)) = f((\sigma\xi_{i1})^{m_1}\cdots(\sigma\xi_{in})^{m_n}) = f(\sigma(\xi_{i1}^{m_1})\cdots\sigma(\xi_{in}^{m_n}))$.
- $f(\tau\chi(\boldsymbol{\xi}_i)) = f(\tau(\xi_{i1}^{m_1}\cdots\xi_{in}^{m_n})) = f(\tau(\xi_{i1}^{m_1})\cdots\tau(\xi_{in}^{m_n}))$.

Let $m = \max|m_j|$ and $M = \max|f|$. For the second sum, note that the summand $f_\chi(\sigma\boldsymbol{\xi}_i) - f(\chi(\sigma\boldsymbol{\xi}_i))$:

- can be bounded above by $M + M^n$ since

$$|f_\chi(\sigma\boldsymbol{\xi}_i) - f(\chi(\sigma\boldsymbol{\xi}_i))| \le |f_\chi(\sigma\boldsymbol{\xi}_i)| + |f(\chi(\sigma\boldsymbol{\xi}_i))|$$
$$\le \left|f(\sigma x_1^{m1})f(\sigma x_2^{m2})\cdots f(\sigma x_n^{mn})\right| + M$$
$$\le M^n + M$$

- $f_\chi(\sigma\boldsymbol{\xi}_i) = f(\chi(\sigma\boldsymbol{\xi}_i))$ unless $|\log|x_j|^m| > c \implies |\log|x_j|| > c/m$ for some j, since $f(x) = x$ in the neighbourhood if $|\log|x_j|| \le c/m$ for every j by definition.

Thus the second sum can be bounded above by

$$\sum_{j=1}^n \frac{1}{[\mathbb{Q}(\boldsymbol{\xi}_i):\mathbb{Q}]} \sum_{\substack{\sigma:\mathbb{Q}(\boldsymbol{\xi}_i)\to\mathbb{C} \\ |\log|\sigma\xi_{ij}|>c/m|}} (M + M^n) \le \sum_{j=1}^n \frac{1}{[\mathbb{Q}(\xi_{ij}):\mathbb{Q}]} \sum_{\substack{\sigma:\mathbb{Q}(\xi_{ij})\to\mathbb{C} \\ |\log|\sigma\xi_{ij}|>c/m|}} (M + M^n).$$

We have assumed $\hat{h}(\boldsymbol{\xi}_i) \to 0$. As in the proof of Bilu's theorem, this assumption implies that

$$\sum_{\substack{\sigma:\mathbb{Q}(\xi_{ij})\to\mathbb{C} \\ |\log|\sigma\xi_{ij}|>c/m|}} (M + M^n) = o(d_{ij}),$$

where $d_{ij} = [\mathbb{Q}(\xi_{ij}) : \mathbb{Q}]$. Therefore, the term above tends to 0 as $i \to \infty$, proving

$$\begin{aligned}
\int_{\mathbb{T}^n} \chi(x) d\mu &= \lim_{i\to\infty} \frac{1}{[\mathbb{Q}(\boldsymbol{\xi}_i) \colon \mathbb{Q}]} \sum_{\sigma:\mathbb{Q}(\boldsymbol{\xi}_i)\to\mathbb{C}} f_\chi(\sigma\boldsymbol{\xi}_i) \\
&= \lim_{i\to\infty} \frac{1}{[\mathbb{Q}(\chi(\boldsymbol{\xi}_i)) : \mathbb{Q}]} \sum_{\tau:\mathbb{Q}(\chi(\boldsymbol{\xi}_i))\to\mathbb{C}} f(\tau\chi(\boldsymbol{\xi}_i)) \\
&= \int_{\mathbb{T}} x d\mu_\chi(x) && \text{(by weak-* convergence)} \\
&= 0. && \text{(by Bilu's theorem)}
\end{aligned}$$

As in the proof of Bilu's theorem, it is clear that μ is a probability measure. It is a standard fact from functional analysis that the characters $\chi(x)$ restricted to $\mathbb{T}^n$ form an orthonormal basis of $L^2(\mathbb{T}^n)$, whence $\int_{\mathbb{T}^n} \chi(x)d\mu = 0$ shows that the restriction of μ to $\mathbb{T}^n$ is the uniform measure on $\mathbb{T}^n$.

Recall that μ is supported on X, but since it has to be the uniform measure on $\mathbb{T}^n$, it turns out that $\mathbb{T}^n$ is contained in the union of the conjugates of X over $\mathbb{Q}$. Since torsion points are Zariski dense in $\mathbb{G}_m^n$, this contradicts the assumption that X is a proper algebraic subvariety of $\mathbb{G}_m^n$.

Case II: There is a non-trivial character χ such that the sequence $(\chi(\xi_i))_{i\in\mathbb{N}}$ has an element ε_0 occurring infinitely many times.

In this case, we proceed by induction on n, the claim being trivial for $n = 0$. Since $h(\chi(\boldsymbol{\xi_i})) \to 0$, we have $h(\varepsilon_0) = 0$ and ε_0 is a root of unity by Kronecker's theorem. Let $\boldsymbol{\varepsilon}$ be a torsion point such that $\chi(\boldsymbol{\varepsilon}) = \varepsilon_0$ and replace X by $\varepsilon^{-1}X$ and $\{\boldsymbol{\xi_i}\}$ by $\{\boldsymbol{\varepsilon}^{-1}\boldsymbol{\xi}_i\}$. Now $(\boldsymbol{\varepsilon}^{-1}X)^* = \boldsymbol{\varepsilon}^{-1}(X^*)$ and multiplication by a torsion point does not change the height; therefore, there is no loss of generality in assuming that $\varepsilon_0 = 1$. Furthermore, going to an infinite subsequence of the sequence $(\boldsymbol{\xi}_i)_{i\in\mathbb{N}}$ if needed, we may also assume that there is a torsion point $\boldsymbol{\varepsilon'}$ such that $\{\boldsymbol{\varepsilon'}\boldsymbol{\xi}_i\}$ is contained in the connected component of the identity of the kernel of χ, say H. Now H is a proper subtorus of $\mathbb{G}_m^n$ and we may replace X, $\mathbb{G}_m^n$ by $\varepsilon' X \cap H$ and H, and then use induction since H is isomorphic to G_m^r for $r < n$. $\square$