MORDELL-LANG PLUS BOGOMOLOV

BJORN POONEN

1. Introduction

Let k be a number field. Let A be a semiabelian variety over k; i.e., an extension

$$0 \to T \to A \to A_0 \to 0,$$

where T is a torus, and A_0 is an abelian variety. Let $h : A(\overline{k}) \to \mathbf{R}$ be a canonical height[1]. For $\epsilon > 0$, let $B_\epsilon = \{ z \in A(\overline{k}) \mid h(z) < \epsilon \}$. Let Γ be a finitely generated subgroup of $A(\overline{k})$. The division group Γ' is defined as

$$\Gamma' := \{ x \in A(\overline{k}) \mid \text{ there exists } n \geq 1 \text{ such that } nx \in \Gamma \}.$$

Define

$$\Gamma'_\epsilon := \Gamma' + B_\epsilon = \{ \gamma + z \mid \gamma \in \Gamma', h(z) < \epsilon \}.$$

We may visualize Γ'_ϵ as a fattening of Γ', a "slab" in the height topology on $A(\overline{k})$.

Let X be a geometrically integral closed subvariety of A. Let $X_{\overline{k}}$ denote $X \times_k \overline{k}$, and so on. Define a *semiabelian subvariety* of a semiabelian variety A to be a subvariety $B \subseteq A$ such that the group structure on A restricts to give B the structure of a semiabelian variety. Then we make the following conjecture:

Conjecture 1 (Mordell-Lang + Bogomolov)**.** *If $X_{\overline{k}}$ is not a translate of a semiabelian subvariety of $A_{\overline{k}}$ by a point in Γ', then for some $\epsilon > 0$, $X(\overline{k}) \cap \Gamma'_\epsilon$ is not Zariski dense in X.*

Applying Conjecture 1 recursively to the components of the Zariski closure of $X(\overline{k}) \cap \Gamma'_\epsilon$ shows that the apparently stronger Conjecture 2 below is in fact equivalent.

Conjecture 2. *There exists $\epsilon > 0$ such that $X(\overline{k}) \cap \Gamma'_\epsilon$ is contained in a finite union $\bigcup Z_j$ where each Z_j is a translate of a semiabelian subvariety of $A_{\overline{k}}$ by a point in Γ' and $Z_j \subseteq X_{\overline{k}}$.*

The main achievement of this paper is a proof of these conjectures in the case that A is *almost split*, i.e., isogenous to the product of an abelian variety and a torus. Note that these conjectures contain the "Mordell-Lang conjecture" (which is completely proven) and the "generalized Bogomolov conjecture," (which has been proved when A is almost split).

Date: July 30, 1999.

1991 *Mathematics Subject Classification.* Primary 11G35; Secondary 11G10, 14G40.

Key words and phrases. Mordell-Lang conjecture, Bogomolov conjecture, small points, equidistribution, semiabelian variety, division group.

Much of this research was done while the author was visiting the Isaac Newton Institute, thanks to a fellowship from the Gabriella and Paul Rosenbaum Foundation. In Berkeley, the author is supported by NSF grant DMS-9801104 and by a Sloan Research Fellowship.

[1]We will give a precise definition in Section 3. For now, let us remark that if A is abelian, then h may be taken as a Néron-Tate canonical height associated to a symmetric ample line bundle, and if $A \cong \mathbf{G}_m^n$ is a split torus, then h may taken as the sum of the naive heights of the coordinates.

Our proof does not yield new proofs of either of these, however, because it uses both, as well as an equidistribution theorem. (We will discuss these major ingredients and their history in Sections 2 and 5.) On the other hand, we do not need the full strength of the Mordell-Lang conjecture: we use only the "Mordellic" (finitely generated) part. Hence our proof gives a new reduction of the full Mordell-Lang conjecture to the Mordellic part, at least for almost split semiabelian varieties.

We need the almost split hypothesis only because it is under this hypothesis that the Bogomolov conjecture and an equidistribution theorem have been proved so far [CL2]. In fact, if we assume the Bogomolov conjecture and an equidistribution theorem for general semiabelian varieties, then we can prove our conjecture entirely.

Remarks.

1. After seeing an earlier version of this paper in which a result only for $X \cap (\Gamma + B_\epsilon)$ was proved instead of for $X \cap (\Gamma' + B_\epsilon)$ (the latter being only conjectured), Shou-Wu Zhang has independently discovered a proof of the division group case [Zh4], again under the almost split hypothesis. He proves first a theorem about "equidistribution of almost division points," and then derives the result from this.
2. Moriwaki [Mo] has proved a generalization of the Bogomolov and equidistribution theorems in which the number field k is replaced by an arbitrary finitely generated field extension of $\mathbf{Q}$. His result is not the function field analogue of the number field result: instead he chooses a base variety having the finitely generated field as function field, and defines his height function as a combination of the geometric height (from the generic fiber) and contributions from the other fibers. It would be interesting to investigate whether the Mordell-Lang and Bogomolov conjectures can be merged over finitely generated fields.

2. The Mordell-Lang conjecture

By the "Mordell-Lang conjecture" we mean the theorem below, which in this form was conjectured by Lang [La] and proved by McQuillan [McQ], following work by Faltings, Vojta, Hindry, Raynaud, and many others.

Theorem 3 (Mordell-Lang conjecture)**.** *Let A be a semiabelian variety over a number field k. Let Γ be a finitely generated subgroup of $A(\overline{k})$, and let Γ' be the division group. Let X be a geometrically integral closed subvariety of A that is not equal to the translate of a semiabelian subvariety (over $\overline{k}$). Then $X(\overline{k}) \cap \Gamma'$ is not Zariski dense in X.*

Remarks.

1. We obtain an equivalent statement if "translate of a semiabelian subvariety" is replaced by "translate of a semiabelian subvariety by a point in Γ'."
2. We obtain an equivalent statement if "not Zariski dense in X" is replaced by "contained in a finite union $\bigcup Z_j$ where each Z_j is a translate of semiabelian subvariety of $A_{\overline{k}}$, and $Z_j \subseteq X_{\overline{k}}$."
3. The theorem is true for *any* field k of characteristic 0: specialization arguments let one reduce to the number field case.
4. There are function field analogues: see [Hr] for example.

We will need only the weaker statement obtained by replacing Γ' by Γ in Theorem 3. This is sometimes called the "Mordellic part," because the special case where $X \subset A$ is a curve of genus ≥ 2 in its Jacobian is equivalent to Mordell's conjecture about the finiteness of $X(k)$.

3. Canonical heights on semiabelian varieties

For the rest of this paper, k denotes a number field. Let A be a semiabelian variety over k, fitting in an exact sequence

$$0 \to T \to A \xrightarrow{\rho} A_0 \to 0,$$

where T is a torus, and A_0 is an abelian variety. Enlarge k to assume that $T \cong \mathbf{G}_m^r$ over k.

To define canonical heights on A, we first compactify, as in [Vo]. (See also [CL1].) According to [Vo], there exist $\mathcal{M}_1, \ldots, \mathcal{M}_r$ in $\operatorname{Pic}^0(A_0)$ such that

$$A \cong \mathbf{P}'(\mathcal{O}_{A_0} \oplus \mathcal{M}_1) \times_{A_0} \cdots \times_{A_0} \mathbf{P}'(\mathcal{O}_{A_0} \oplus \mathcal{M}_r),$$

where $\mathbf{P}'(\mathcal{L} \oplus \mathcal{M})$ means the open subset of $\mathbf{P}(\mathcal{L} \oplus \mathcal{M})$ obtained by deleting the sections corresponding to the canonical projections to $\mathcal{L}$ and $\mathcal{M}$. Then

$$\overline{A} := \mathbf{P}(\mathcal{O}_{A_0} \oplus \mathcal{M}_1) \times_{A_0} \cdots \times_{A_0} \mathbf{P}(\mathcal{O}_{A_0} \oplus \mathcal{M}_r)$$

is a suitable compactification of A. Let $\mathcal{L}_0$ denote the line bundle on $\overline{A}$ associated to the divisor $\overline{A} \setminus A$. Let $\mathcal{L}_1 = \rho^*\mathcal{M}$ where $\mathcal{M}$ is a symmetric ample line bundle on A_0. For $n \in \mathbf{Z}$, the multiplication-by-n map on A extends to a morphism $[n] : \overline{A} \to \overline{A}$. We have $[n]^*\mathcal{L}_0 \cong \mathcal{L}_0^{\otimes |n|}$ (the $n \geq 1$ case is on p. 140 of [Vo]) and $[n]^*\mathcal{L}_1 \cong \mathcal{L}_1^{\otimes n^2}$. Therefore for $x \in A(\overline{k})$, we define nonnegative functions

$$(1) \qquad h_0(x) := \lim_{n\to\infty} \frac{h_{\mathcal{L}_0}([n]x)}{|n|}, \quad h_1(x) := \lim_{n\to\infty} \frac{h_{\mathcal{L}_1}([n]x)}{n^2}, \quad \text{and} \quad h(x) := h_0(x) + h_1(x).$$

We record some properties of h:

Lemma 4. *Suppose $n \in \mathbf{Z}$, $\sigma \in G_k$, and $x \in A(\overline{k})$.*

1. *Up to $O(1)$, $h(x)$ equals a Weil height associated to the line bundle $\mathcal{L} := \mathcal{L}_0 \otimes \mathcal{L}_1$ on $\overline{A}$, which is ample.*
2. *$h_0([n]x) = |n| h_0(x)$.*
3. *$h_1([n]x) = n^2 h_1(x)$.*
4. *$|n| h(x) \leq h([n]x) \leq n^2 h(x)$.*
5. *$h(\sigma x) = h(x)$.*
6. *If $a \in A(\overline{k})$ is torsion, then $h(x + a) = h(x)$.*

Proof. For the ampleness of $\mathcal{L}$, see Lemma 3.1 of [Vo]. Parts (1) through (5) are now trivial, and part (6) follows from (2) and (3). □

In the case where A is almost split, we can circumvent the need for a compactification by choosing a surjective homomorphism $\phi : A \to \mathbf{G}_m^r$ and defining $h_0(x)$ by the sum of the naive heights of the r coordinates of $\phi(x)$. Note that $h_1(x)$ can be defined as the Néron-Tate height with respect to $\mathcal{M}$ of the image of x in A_0; no compactification is needed for this. The definition of h depends on $\mathcal{M}$ (and also on ϕ in the almost split definition), but the content of the conjectures we are considering is independent of the choices made, as we point out in the next section.

4. Sequences of small points

A sequence of points $x_i \in A(\overline{k})$ is said to be a *sequence of small points* if $h(x_i) \to 0$. Hrushovski pointed out to the author that in formulating the Bogomolov conjecture, equidistribution conjecture, or Conjecture 1, all that matters is the notion of sequence of small points, and that this notion can be defined without using canonical heights. Using his idea, which we develop in the rest of this section, we see that the notion is very robust, depending only on A, even though there were many choices in Section 3 for a canonical height function.

We temporarily consider a more general situation. Let U be a geometrically integral quasi-projective variety over a number field k, equipped with a morphism $f : U \to U$. For integers $r \geq 1$, let $f^r : U \to U$ denote the r-th iterate of f. We assume the following condition on (U, f):

(∗) There exist a Weil height $h : U(\overline{k}) \to \mathbf{R}$ associated to some embedding $U \hookrightarrow \mathbf{P}^n$, an integer $r \geq 1$, and real numbers $M > 0$ and $c > 1$ such that $h(z) > M$ implies $h(f^r(z)) > ch(z)$.

If $z \in U(\overline{k})$, let $N(z)$ be the smallest integer $N \geq 1$ such that $h(f^N(z)) > M$, or ∞ if no such N exists. Northcott's finiteness theorem about the number of points of bounded height and degree implies that $N(z) = \infty$ if and only if z is *preperiodic* for f (i.e., has finite orbit under the iterates of f). For $i \geq 1$, let $z_i \in U(\overline{k})$. We say that $\{z_i\}_{i \geq 1}$ is a *sequence of small points* if $N(z_i) \to \infty$ in $\mathbf{P}^1(\mathbf{R})$.

Proposition 5. *Suppose (U, f) satisfies (∗), with h, r, M, and c.*

1) *If h' is another Weil height, corresponding to another projective embedding, then there exist r', M', and c' as in (∗) for h'.*
2) *For any such choices, the notion of sequence of small points obtained is the same.*
3) *If (U, g) also satisfies (∗), and $fg = gf$, then the notion obtained using g is the same as that using f.*
4) *If (U', f') also satisfies (∗), and $\psi : U \to U'$ satisfies $\psi f = f' \psi$, then ψ maps sequences of small points to sequences of small points.*

Proof. This follows from elementary properties of heights, such as the fact that Weil heights corresponding to two different embeddings are bounded by linear functions in each other. □

Remarks.

1. Condition (∗) is satisfied for (U, f) if there exists an integral projective variety V containing U as an open dense subset, and an ample line bundle $\mathcal{L}$ on V such that f extends to a morphism $\overline{f} : V \to V$ and a height associated to $\mathcal{N} := \overline{f}^* \mathcal{L} \otimes \mathcal{L}^{\otimes -q}$ in $(\operatorname{Pic} V) \otimes \mathbf{Q}$ is bounded below for some $1 < q \in \mathbf{Q}$. The condition on $\mathcal{N}$ is satisfied, for instance, if $\mathcal{N}$ is the pullback of an ample sheaf under some morphism of varieties.
2. Our situation is only slightly more general than that considered in [CS] and the introduction of [Zh1], which consider projective varieties X equipped with $f : X \to X$ and $\mathcal{L}$ in $\operatorname{Pic} V$ or $(\operatorname{Pic} V) \otimes \mathbf{R}$ satisfying $f^* \mathcal{L} = \mathcal{L}^{\otimes d}$ for some $d > 1$. With our weaker assumptions, we have apparently lost the ability to define a "canonical height" using f, but the notion of "sequence of small points" is still definable.

When $U = A$ is a semiabelian variety as in Section 3 (with split torus), we may take $V = \overline{A}$, $\mathcal{L} = \mathcal{L}_0 \otimes \mathcal{L}_1$, and $\overline{f} = [m]$ for any integer $m \geq 2$. Remark 1 above with $q = m$ and $\mathcal{N} = \mathcal{L}_1^{\otimes (m^2 - m)}$, shows that $(A, [m])$ satisfies (∗) for $m \geq 2$. Proposition 5 shows that the resulting notion of sequence of small points depends only on A. Moreover it is easy to see that this notion agrees with the earlier ones defined using canonical heights.

5. The Bogomolov and equidistribution conjectures

We fix a canonical height h on each semiabelian variety A as in Section 3. For $\epsilon > 0$, define $B_\epsilon = \{ z \in A(\overline{k}) \mid h(z) < \epsilon \}$. The conjectures below are well known.

Conjecture 6 (Bogomolov conjecture for semiabelian varieties)**.** *Let A be a semiabelian variety over a number field k. Let X be a geometrically integral closed subvariety of A, such that $X_{\overline{k}}$ is not the translate of a semiabelian subvariety of $A_{\overline{k}}$ by a torsion point. Then there exists $\epsilon > 0$ such that $X(\overline{k}) \cap B_\epsilon$ is not Zariski dense in X.*

Bogomolov's original conjecture was for a curve of genus ≥ 2 in its Jacobian. This case was proved by Ullmo [Ul], using the equidistribution results of [SUZ]. For A an abelian variety, Conjecture 6 was proved by Zhang in [Zh3], and another proof (not entirely independent) was given shortly thereafter by David and Philippon [DP]. Zhang proved also the case where A is a torus [Zh2]. Recently, a proof for the case where A is almost split was announced by Chambert-Loir [CL2].

If X is a geometrically integral variety, then a sequence of points z_i is said to be *generic* in X if $z_i \in X(\overline{k})$ for all i, and the z_i converge to the generic point of $X_{\overline{k}}$. The latter condition means that each closed subvariety Y of $X_{\overline{k}}$ other than $X_{\overline{k}}$ itself contains at most finitely many points of the sequence.

Recall that if μ_i for $i \geq 1$ and μ are probability measures on a metric space X, then one says that the μ_i *converge weakly* to μ if for every bounded continuous function f on X, $\lim_{i\to\infty} \int f \mu_i = \int f \mu$. For the following, we fix an embedding $\sigma : \overline{k} \hookrightarrow \mathbf{C}$, and let A_σ denote the semiabelian variety over $\mathbf{C}$ obtained by base extension by σ. Let $G_k = \mathrm{Gal}(\overline{k}/k)$. For $z \in A(\overline{k})$, define $G_k(z) := \{ gz : g \in G_k \}$.

Conjecture 7 (Equidistribution conjecture)**.** *Let A be a semiabelian variety over a number field k. Let $\{z_i\}$ be a sequence in $A(\overline{k})$, generic in A, with $h(z_i) \to 0$. Let μ_i be the uniform probability measure on the finite set $\sigma(G_k(z_i)) \subset A_\sigma(\mathbf{C})$. Then the μ_i converge weakly to the normalized Haar measure μ on the maximal compact subgroup A^0 of $A_\sigma(\mathbf{C})$.*

Remarks.

1. The subgroup A^0 also equals the closure of $A_\sigma(\mathbf{C})_{\mathrm{tors}}$ in the complex topology. It is all of $A(\mathbf{C})$ if A is abelian, and it is a "polydisc" if A is a torus.
2. Some authors, when they speak of the equidistribution conjecture, replace the word "generic" by the weaker hypothesis "strict," which means that each translate of a semiabelian subvariety by a torsion point contains at most finitely many terms of the sequence. The resulting statement hence combines the Bogomolov conjecture with what we are calling the equidistribution conjecture.

The cases where A is an abelian variety or a torus are proved in [SUZ] and [Bi], respectively. A proof for the case where A is an almost split semiabelian variety has been announced by Chambert-Loir [CL2]. Bilu suggested in [Bi] that for any semiabelian variety A, Haar measure on A^0 should be the limit measure, but the result is not yet proved.

We recommend [Ul] and the survey by Abbès [Ab] for the history of the Bogomolov and equidistribution conjectures.

6. Statement of the result

Let $\mathcal{S}(A)$ be the set of semiabelian varieties over number fields that can be obtained from a semiabelian variety A by taking algebraic subgroups, quotients, and products, and changing the field of definition.

Theorem 8. *Let A be a semiabelian variety over a number field k. Assume that Conjectures 6 and 7 hold for all $B \in \mathcal{S}(A)$. Then Conjectures 1 and 2 hold for subvarieties X in A.*

Corollary 9. *Conjectures 1 and 2 hold when A is almost split.*

Proof. One checks that all $B \in \mathcal{S}(A)$ are almost split. Chambert-Loir [CL2] proved Conjectures 6 and 7 for almost split semiabelian varieties. □

7. Properties of small points

In preparation for the proof of Theorem 8, we derive a few more properties of sequences of small points. Throughout this section, A denotes a semiabelian variety over a number field k. Let $\beta_n : A^n \to A^{n-1}$ be the map sending $(x_1, \ldots, x_n)$ to $(x_2 - x_1, x_3 - x_1, \ldots, x_n - x_1)$. Up to an automorphism of A^{n-1}, this is the same as the map α_n used in [Zh3].

Let $\operatorname{Diff}_n(x)$ be the (arbitrarily ordered) list of $[k(x) : k]^n$ elements of A^{n-1} obtained by applying β_n to the elements of $G_k(x)^n$. Given any sequence $x_1, x_2, \ldots$ of points in $A(\overline{k})$, let $\mathcal{D}_n = \mathcal{D}_n(\{x_i\})$ denote the sequence obtained by concatenating $\operatorname{Diff}_n(x_1)$, $\operatorname{Diff}_n(x_2)$, $\ldots$.

Lemma 10. *Let A and B be semiabelian varieties over a number field k. Let $\{x_i\}$ be a sequence of small points in A, and let $\{y_i\}$ be a sequence of small points in B.*

1. *The sequence $\{(x_i, y_i)\}$ is a sequence of small points in $A \times B$.*
2. *If $f : A \to B$ is a homomorphism, then $\{f(x_i)\}$ is a sequence of small points in B.*
3. *If $A = B$, then $\{x_i + y_i\}$ and $\{x_i - y_i\}$ are sequences of small points in A.*
4. *For any $n \geq 2$, $\mathcal{D}_n = \mathcal{D}_n(\{x_i\})$ is a sequence of small points in A^{n-1}.*
5. *If $[k(x_i) : k]$ is bounded, then there is a finite subset $T \subset A(\overline{k})_{\text{tors}}$ containing all but finitely many of the x_i.*

Proof. Property 1 is immediate from the definition in Section 4, and (2) follows from part 4) of Proposition 5. Property (3) follows from (1) and (2), and (4) follows from (1) and (3) and part (5) of Lemma 4.

Finally we prove (5). By Northcott's theorem, the sequence involves only finitely many points. Since $h(x_i) \to 0$, there is a finite set T of points of zero height such that $x_i \in T$ for $i \gg 0$. Northcott implies that points of zero height are torsion. □

Next we have a sequence of lemmas leading up to Lemma 14, which is the only other result from this section that will be used later.

Lemma 11. *Given $a \in A(\overline{k})$, there exists $M_a > 0$ such that if $x \in A(\overline{k})$ and $h(x) > M_a$, then $h([2]x + a) > (3/2)h(x)$.*

Proof. Translation-by-a extends to a morphism $\tau_a : \overline{A} \to \overline{A}$, and $\tau_a^* \mathcal{L}_0 = \mathcal{L}_0$. It follows that $h_0(x + a) = h_0(x) + O(1)$, where the $O(1)$ depends on a. On the other hand, since h_1 is a quadratic function, $h_1(x+a) = h_1(x) + O(h_1(x)^{1/2}) + O(1)$. Using Lemma 4 we obtain

$$\begin{aligned} h([2]x+a) &= h_0([2]x+a) + h_1([2]x+a) \\ &= [2h_0(x) + O(1)] + \left[4h_1(x) + O(h_1(x)^{1/2}) + O(1)\right] \\ &= 2h(x) + \left[2h_1(x) + O(h_1(x)^{1/2}) + O(1)\right] \\ &\geq 2h(x) + O(1). \end{aligned}$$

□

Lemma 12. *Let Γ be a finitely generated subgroup of $A(\overline{k})$, and let $\{x_i\}$ be a sequence in Γ'. If the image of $\{x_i\}$ in $\Gamma' \otimes \mathbf{R} = \Gamma \otimes \mathbf{R}$ converges to zero in the usual real vector space topology, then $\{x_i\}$ is a sequence of small points.*

Proof. Let $S := \{\gamma_1, \gamma_2, \ldots, \gamma_n\} \subset \Gamma$ be a $\mathbf{Z}$-basis for $\Gamma/\Gamma_{\text{tors}}$. Let $U = \{\sum \epsilon_i \gamma_i : \epsilon_i \in \{-1, 0, 1\}\}$. Let $f_1, f_2, \ldots, f_u$ be the maps $A \to A$ of the form $x \mapsto [2]x + a$ for $a \in U$. Applying Lemma 11 to all $a \in U$ yields $M > 0$ such that $h(x) > M$ implies $h(f_i(x)) > (3/2)h(x)$ for all i.

Let B be the subset of elements of Γ' whose image in $\Gamma \otimes \mathbf{R}$ have coordinates (with respect to the basis S) bounded by 1 in absolute value. For every $b_0 \in B$, there exists i such that $b_1 := f_i(b_0) \in B$, and then there exists j such that $b_2 := f_j(b_1) \in B$, and so on. The intersection I of B with the finitely generated subgroup generated by b_0 and S is finite, and $b_i \in I$ for all i. But if $h(b_0) > M$, then $h(b_{m+1}) > (3/2)h(b_m)$ for all m, and in particular, the b_i would be all distinct. This contradicts the finiteness of I, so $h(b_0) \leq M$ for all $b_0 \in B$. The lemma now follows from part (4) of Lemma 4. □

Remark. The converse to Lemma 12 is true, but we do not need it.

Lemma 13. *Let Γ be a finitely generated subgroup of $A(\overline{k})$. Then $A(k) \cap \Gamma'$ is a finitely generated group.*

Proof. It suffices to show that there is a neighborhood U of 0 in $\Gamma \otimes \mathbf{R}$ in the real topology such that the set of elements of $A(k) \cap \Gamma'$ that map into U is finite. If no such U exists, then we have an infinite sequence of distinct points $x_i \in A(k) \cap \Gamma'$ whose images in $\Gamma \otimes \mathbf{R}$ tend to 0 in the real topology. By Lemma 12, $h(x_i) \to 0$. This contradicts Northcott's theorem. □

Lemma 14. *Let Γ be a finitely generated subgroup of $A(\overline{k})$. Suppose $\{x_i\}$ is a sequence in $A(k)$, and $x_i = \gamma_i + z_i$ where $\gamma_i \in \Gamma'$, and $z_i \in A(\overline{k})$ with $h(z_i) \to 0$. Then there is a finitely generated subgroup of Γ' containing all but finitely many of the x_i.*

Proof. We may enlarge k to assume $\Gamma \subset A(k)$. By part (4) of Lemma 4 we can choose integers $n_i \geq 1$ tending to infinity slowly enough that if $1 \leq m_i \leq n_i$, then $h(m_i z_i) \to 0$. By elementary diophantine approximation (the pigeonhole principle), there exist integers m_i with $1 \leq m_i \leq n_i$, and $\nu_i \in \Gamma$ such that the images of $m_i\gamma_i - \nu_i$ in $\Gamma \otimes \mathbf{R}$ approach zero as $i \to \infty$. Then $h(m_i\gamma_i - \nu_i) \to 0$ by Lemma 12, but $h(m_i z_i) \to 0$ also, so by part (3) of Lemma 10, $h(m_i x_i - \nu_i) \to 0$. On the other hand, $m_i x_i - \nu_i \in A(k)$, so by part (5) of Lemma 10, $m_i x_i - \nu_i$ is torsion and $x_i \in \Gamma'$ for all but finitely many i. Finally, Lemma 13 implies that there is a finitely generated subgroup of Γ' containing all but finitely many x_i. □

8. Measure-theoretic lemmas

Lemma 15. *Let V be a projective variety over $\mathbf{C}$. Let S be a connected quasi-projective variety over $\mathbf{C}$. Let $\mathcal{Y} \to V \times S$ be a closed immersion of S-varieties, where $\mathcal{Y} \to S$ is flat with d-dimensional fibers. For $i \geq 1$, let $s_i \in S(\mathbf{C})$ and let $Y_i \subset V$ be the fiber of $\mathcal{Y} \to S$ above s_i. Let μ_i be a probability measure supported on $Y_i(\mathbf{C})$. If the μ_i converge weakly to a probability measure μ on $V(\mathbf{C})$, then the support of μ is contained in a d-dimensional Zariski closed subvariety of V.*

Proof. We may assume that $V = \mathbf{P}^n$ and that $\mathcal{Y} \to \mathbf{P}^n_S$ is the universal family over a Hilbert scheme S. Since S is projective over $\mathbf{C}$, we may pass to a subsequence to assume that the s_i converge in the complex topology to $s \in S(\mathbf{C})$. By compactness of $\mathcal{Y}(\mathbf{C})$, μ must be supported on the fiber $\mathcal{Y}_s \subset V$. □

Remark. The hypotheses can be weakened. It is enough to assume that the Y_i form a "limited family" of closed subvarieties of V of dimension $\leq d$, because then there are finitely many possibilities for their Hilbert polynomials [Gr, Théorème 2.1]. The limited family condition holds, for instance, if the Y_i are reduced and equidimensional of dimension d, and $\deg Y_i$ (with respect to some fixed embedding $V \hookrightarrow \mathbf{P}^n$) is bounded [Gr, Lemme 2.4].

Lemma 16. *Let V and S be quasi-projective varieties over $\mathbf{C}$, with S integral. Let $\mathcal{Y}$ be a subvariety of $V \times S$. Let $s_1, s_2, \ldots$ be a sequence in $S(\mathbf{C})$, Zariski dense in S. Let μ_i be a probability measure with support contained in the fiber of $\pi : \mathcal{Y} \to S$ above s_i, considered as subvariety of V. Suppose the μ_i converge weakly to a probability measure μ on $V(\mathbf{C})$. Then the support of μ is contained in a subvariety of V of dimension* $\dim \mathcal{Y} - \dim S$.

Proof. Choose an embedding $V \hookrightarrow \mathbf{P}^m$. Without loss of generality, we may replace V and $\mathcal{Y}$ by their closures in $\mathbf{P}^m$ and $\mathbf{P}^m \times S$, respectively. Replacing S by a dense open subset U and $\mathcal{Y}$ by $\pi^{-1}(U)$, and passing to a subsequence, we may assume that $\mathcal{Y} \to S$ is flat. The result now follows from Lemma 15. □

Lemma 17. *Retain the assumptions of the previous lemma, but assume in addition that V is a semiabelian variety over $\mathbf{C}$, and that $\mathcal{Y}$ is not Zariski dense in $V \times S$. Then μ does not equal the normalized Haar measure on the maximal compact subgroup V^0 of $V(\mathbf{C})$.*

Proof. We have $\dim \mathcal{Y} - \dim S < \dim V$, so by the previous lemma, μ is supported on a subvariety of V of positive codimension. But V^0 is Zariski dense in V, since its Zariski closure is an algebraic subgroup of V containing all its torsion. □

9. Proof of Theorem 8

The reduction of Conjecture 2 to Conjecture 1 mentioned in the introduction applies, since we are assuming Conjectures 6 and 7 for all $B \in \mathcal{S}(A)$; therefore we need prove only Conjecture 1. The proof will proceed through various reductions; to aid the reader, we $\boxed{\text{box}}$ cumulative assumptions and other partial results to be used later in proof.

Let G be the group of translations preserving X; i.e., the largest algebraic subgroup of A such that $X+G = X$. We may assume $\boxed{\dim G = 0}$, since otherwise we consider $X/G \hookrightarrow A/G$ and use part (2) of Lemma 10. We may also enlarge k to assume that $\boxed{\Gamma \subset A(k)}$.

If Conjecture 1 is false, then there exists a sequence $x_i = \gamma_i + z_i \in X(\overline{k})$, generic in X, with $\gamma_i \in \Gamma'$, $z_i \in A(\overline{k})$, and $h(z_i) \to 0$. For $\sigma, \tau \in G_k$,

$$\sigma x_i - \tau x_i = (\sigma\gamma_i - \tau\gamma_i) + (\sigma z_i - \tau z_i).$$

Some multiple of γ_i is in $\Gamma \subset A(k)$, so $\sigma\gamma_i - \tau\gamma_i$ is torsion. Part (4) of Lemma 10 implies that $\mathcal{D}_2(\{z_i\})$ is a sequence of small points, so by part (6) of Lemma 4, $\mathcal{D}_2 := \mathcal{D}_2(\{x_i\})$ also is a sequence of small points. Then by part (1) of Lemma 10, $\mathcal{D}_n$ is a sequence of small points for each $n \geq 2$. Fix $\boxed{n > \dim A}$.

By repeated application of Conjecture 6, we may discard finitely many of the x_i in order to assume that the Zariski closure $\overline{\mathcal{D}}_n$ of $\mathcal{D}_n$ in $A^{n-1}_{\overline{k}}$ is a finite union $\bigcup_{j=1}^s (B_j + t_j)$ where B_j is a semiabelian subvariety of $A^{n-1}_{\overline{k}}$, and $t_j \in A^{n-1}(\overline{k})$ is a torsion point. If we replace X by the image of X under multiplication by a positive integer N, and replace each x_i by Nx_i, then $\bigcup_{j=1}^s B_j$ is unchanged. If we pass to a subsequence of the x_i or enlarge k, then the new $\bigcup_{j=1}^s B_j$ can only be smaller. Since $A^{n-1}_{\overline{k}}$ is noetherian, we may assume without loss of generality that these operations are done so as to make $\bigcup_{j=1}^s B_j$ minimal. Moreover, by multiplying by a further integer N we may assume that $t_j = 0$ for all j. Now, any further operations of the types above will leave $\overline{\mathcal{D}}_n$ unchanged, equal to $\bigcup_{j=1}^s B_j$. Enlarging k, we may assume that each B_j is defined over k.

Repeating the same procedure with 2 instead of n, we may minimize $\overline{\mathcal{D}}_2 = \bigcup_{j=1}^u C_j$, and assume that each C_j is a semiabelian subvariety of A. We now show that there is only one C_j. By the pigeonhole principle, there is some C_j, say $C := C_1$, such that for infinitely many i, at least a fraction $1/u$ of the elements of $G_k(x_i)^2$ are mapped by β_2 into C. Passing to a subsequence of the x_i, we may suppose that this holds for *all* i. Let $\pi : A \to A/C$ be the projection, and let $y_i = \pi(x_i)$. If ℓ is a finite Galois extension of k containing $k(x_i)$, then it follows that $\sigma y_i - \tau y_i = 0$ for at least a fraction $1/u$ of the pairs $(\sigma, \tau) \in \mathrm{Gal}(\ell/k)^2$. Thus the subgroup of $\mathrm{Gal}(\ell/k)$ stabilizing y_i must have index at most u. Hence $\deg_k(y_i) = \#G_k(y_i) \leq u$. Then $\mathcal{D}_2(\{y_i\})$ consists of points of degree bounded by u^2. On the other hand, $\mathcal{D}_2(\{y_i\})$ is a sequence of small points in A/C, by parts (2) and (4) of Lemma 10. By part (5) of Lemma 10, there is a finite subset T of torsion points of A/C such that $\mathrm{Diff}_2(y_i) \subseteq T$ for all sufficiently large i. Passing to a subsequence of the x_i, and multiplying everything by an integer N to kill T, we may assume that $\mathrm{Diff}_2(y_i) = \{0\}$ for all i. Then $\mathrm{Diff}_2(x_i) \subseteq C$, so $\overline{\mathcal{D}}_2 \subseteq C$, and hence $\boxed{\overline{\mathcal{D}}_2 = C}$ by definition of C.

We next show that there is only one B_j, and that it equals C^{n-1}. By definition of $\mathcal{D}_2$ and $\mathcal{D}_n$, we have $B_j \subseteq C^{n-1}$ for each j. By the pigeonhole principle, there is some B_j, say $B := B_1$, such that for infinitely many i, at least a fraction $1/s$ of the elements of $G_k(x_i)^n$ are mapped by β_n into B. Passing to a subsequence, we may suppose that this holds for *all* i. For $1 \leq q \leq n-1$, define the "coordinate axis" $C_{(q)} = 0 \times \cdots \times 0 \times C \times 0 \times \cdots \times 0$, with C in the q-th place. Let $B_{(q)} = B \cap C_{(q)}$. By the pigeonhole principle again, given i, there exist $w_1, w_2, \ldots, w_q, w_{q+2}, \ldots, w_n \in G_k(x_i)$ such that $\beta_n(w_1, w_2, \ldots, w_q, \zeta, w_{q+2}, \ldots, w_n) \in B$ for at least a fraction $1/s$ of the elements ζ of $G_k(x_i)$. Subtracting, we find that $\beta_n(0, 0, \ldots, 0, \zeta - \zeta', 0, \ldots, 0) \in B_{(q)}$ for at least a fraction $1/s^2$ of the pairs of elements ζ, ζ' of $G_k(x_i)$. As before, this implies (after passing to a subsequence and multiplying by a positive integer again) that the image y_i of x_i in $C/B_{(q)}$ satisfies $\mathrm{Diff}_2(y_i) = \{0\}$. Then $\overline{\mathcal{D}}_2 \subseteq B_{(q)} \subseteq C$, so $B_{(q)} = C$. This holds for all q, so $\boxed{\overline{\mathcal{D}}_n = B = C^{n-1}}$.

If $C = \{0\}$, then $\mathcal{D}_2 = \{0\}$, and then by definition of $\mathcal{D}_2$, $x_i \in X(k)$ for all i. Lemma 14 implies that all but finitely many x_i are contained in a finitely generated subgroup $\tilde{\Gamma}$ of Γ'. The Mordellic part of Theorem 3 applied to $\tilde{\Gamma}$ implies that $X_{\overline{k}}$ is a translate of a semiabelian subvariety. Moreover, it is a translate by a point in Γ', since $X(\overline{k}) \cap \Gamma'$ contains points, namely most of the x_i. This contradicts the assumption on X. Therefore we may assume $\boxed{\dim C \geq 1}$.

Let $S = \pi(X) \subseteq A/C$. Note that S is integral. Consider the fibered power $X_S^n := X \times_S X \times_S \cdots \times_S X$ as a subvariety of X^n.

Let $\boxed{m = \dim C}$. Note that $1 \leq m \leq \dim A < n$. Let $\dim(X/S)$ denote the relative dimension; i.e., the dimension of the generic fiber of $X \to S$. Then $\boxed{\dim(X/S) < m}$, since otherwise X (being closed) would equal the entire inverse image of S under $A \to A/C$, and then $C \subseteq G$, contradicting $\dim G = 0$. Hence

$$\dim(X_S^n/S) = n \dim(X/S) \leq n(m-1) < m(n-1) = \dim(C^{n-1} \times S/S). \tag{2}$$

The homomorphism β_n restricts to a morphism $X_S^n \to C^{n-1}$. We also have the obvious morphism $X_S^n \to S$. Let $\mathcal{Y}$ denote the image of the product morphism $X_S^n \to C^{n-1} \times S$. Then $\dim \mathcal{Y} < \dim C^{n-1} \times S$, by (2).

Since $\sigma x_i - \tau x_i \in C(\overline{k})$ for all $\sigma, \tau \in G_k$ and all $i \geq 1$, concatenating the finite subsets $\beta_n(G_k(x_i)^n) \times \{\pi(x_i)\}$ of $C^{n-1} \times S$ yields a sequence of points $y_j = (c_j, s_j)$ in $\mathcal{Y}$. The c-sequence is simply $\mathcal{D}_n$, and each s_j equals $\pi(x_i)$ for some i. Since $\mathcal{D}_n$ is dense in C^{n-1}, we may pass to a subsequence of the y_j to assume that the c_j are generic in C^{n-1}. Now fix an embedding $\sigma : \overline{k} \hookrightarrow \mathbf{C}$, and let μ_j be the uniform probability measure on the finite subset $\sigma(G_k(c_j)) \subset C_\sigma^{n-1}(\mathbf{C})$. Conjecture 7 implies that the μ_j converge to the normalized Haar measure μ on the maximal compact subgroup of $C_\sigma^{n-1}(\mathbf{C})$.

On the other hand, μ_j is supported on $\sigma\beta_n(G_k(x_i)^n)$, which is contained in the fiber of $\mathcal{Y} \to S$ above s_j, when we consider the fiber as a subvariety of $V := C_\sigma^{n-1}$. Lemma 17 implies that the μ_j cannot converge to μ. This contradiction completes the proof of Theorem 8.

Acknowledgements

I thank Ahmed Abbès and Jean-Benoît Bost for many suggestions and corrections to an early draft, and the former for explaining to me the proof of the Bogomolov conjecture for abelian varieties. I also thank Ehud Hrushovski for the idea that one could define the notion of a sequence of small points using only Weil heights. Finally I thank Shou-Wu Zhang for some interesting conversations on related questions, and Antoine Chambert-Loir and Paul Vojta for discussions about compactifications and canonical heights on semiabelian varieties.

References

[Ab] Abbès, A., Hauteurs et discrétude, *Séminaire Bourbaki*, 49ème année, 1996-97 (Mars 1997), no. 825.

[Bi] Bilu, Y., Limit distribution of small points on algebraic tori, *Duke Math. J.* **89** (1997), no. 3, 465–476.

[CS] Call, G. and Silverman, J., Canonical heights on varieties with morphisms, *Compositio Math.* **89** (1993), no. 2, 163–205.

[CL1] Chambert-Loir, A., Géométrie d'Arakelov et hauteurs canoniques sur des variétés semi-abéliennes, to appear in *Math. Ann.*

[CL2] Chambert-Loir, A., Points de petite hauteur sur les variétés semi-abéliennes, preprint, 1998.

[DP] David, S. and Philippon, P., Minorations des hauteurs normalisées des sous-variétés de variétés abéliennes, in *Number theory (Tiruchirapalli, 1996)*, 333–364, Contemp. Math., 210, Amer. Math. Soc., Providence, RI.

[Gr] Grothendieck, A., Techniques de construction et théorèmes d'existence en géométrie algébrique IV: les schémas de Hilbert, no. 221 in *Fondements de la géométrie algébrique. [Extraits du Séminaire Bourbaki, 1957–1962.]*, Secrétariat Math., Paris, 1962.

[Hr] Hrushovski, E., The Mordell-Lang conjecture for function fields, *J. Amer. Math. Soc.* **9** (1996), no. 3, 667–690.

[La] Lang, S., Division points on curves, *Ann. Mat. Pura Appl. (4)* **70** (1965), 229–234.

[McQ] McQuillan, M., Division points on semi-abelian varieties, *Invent. Math.* **120** (1995), no. 1, 143–159.

[Mo] Moriwaki, A., Arithmetic height functions over function fields, preprint, July 1998.

[SUZ] Szpiro, L., Ullmo, E., and Zhang, S., Équirépartition des petits points, *Invent. Math.* **127** (1997), no. 2, 337–347.

[Ul] Ullmo, E., Positivité et discrétion des points algébriques des courbes, *Ann. of Math. (2)* **147** (1998), no. 1, 167–179.

[Vo] Vojta, P., Integral points on subvarieties of semiabelian varieties, I, *Invent. Math.* **126** (1996), no. 1, 133–181.

[Zh1] Zhang, S., Small points and adelic metrics, *J. Algebraic Geom.* **4** (1995), no. 2, 281–300.

[Zh2] Zhang, S., Positive line bundles on arithmetic varieties, *J. Amer. Math. Soc.* **8** (1995), no. 1, 187–221.

[Zh3] Zhang, S., Equidistribution of small points on abelian varieties, *Ann. of Math. (2)* **147** (1998), no. 1, 159–165.

[Zh4] Zhang, S., Distribution of almost division points, preprint, September 1998.

Department of Mathematics, University of California, Berkeley, CA 94720-3840, USA
E-mail address: `poonen@math.berkeley.edu`