Vol. 103, No. 1 DUKE MATHEMATICAL JOURNAL

DISTRIBUTION OF ALMOST DIVISION POINTS

SHOU-WU ZHANG

1. Introduction. In [10], we proved an equidistribution theorem for small points on abelian varieties, based on the ideas in [7] and [8]. In this paper, we want to generalize this result to *almost division points*. In the following, we describe our main theorem and its application to the discreteness of almost division points on subvarieties.

Let A be an abelian variety defined over a number field K. Let x_n $(n \in \mathbb{N})$ be a sequence of distinct points in $A(\bar{K})$. We assume this is a sequence of almost division points, which means

$$\lim_{n\to\infty} \sup_{\sigma\in G} \left\| x_n^\sigma - x_n \right\| = 0.$$

Here, $G = \mathrm{Gal}(\bar{K}/K)$, and $\|\cdot\|$ is the square root of the Neron-Tate height function, with respect to some ample and symmetric line bundle on A. Obviously, the notion of almost division does not depend on the choice of the Neron-Tate height functions. If we drop the limit in the above equality, then all x_n are division points for $A(K)$.

We fix an embedding $\sigma : \bar{K} \to \mathbb{C}$; then $A(\bar{K})$ can be considered a subgroup of $A(\mathbb{C}) := A_\sigma(\mathbb{C})$. The Galois orbits x_n^G therefore define a sequence δx_n^G of probability measures on $A(\mathbb{C})$; if f is a continuous function on $A(\mathbb{C})$, then

$$\int_{A(\mathbb{C})} f\,\delta x_n^G = \frac{1}{\left|x_n^G\right|} \sum_{y\in x_n^G} f(y).$$

In this paper, we address the convergence of δx_n^G. More precisely, we want to know whether there is a measure $d\mu$ on $A(\mathbb{C})$ such that, for any continuous function f on $A(\mathbb{C})$,

$$\lim_{n\to\infty} \int_{A(\mathbb{C})} f\,dx_n^G = \int_{A(\mathbb{C})} f\,d\mu.$$

Obviously, such a measure $d\mu$ does not exist in general; but, since the space of the continuous functions on $A(\mathbb{C})$ can be topologically generated by countably many functions, $d\mu$ does exist if $(x_n, n \in \mathbb{N})$ is replaced with a subsequence. So, our purpose becomes to describe the following:

- the property of the sequence $(x_n, n \in \mathbb{N})$ which can be obtained by replacing it with a subsequence;

Received 5 November 1998. Revision received 9 June 1999.
2000 *Mathematics Subject Classification*. Primary 11G, 14G.
This research was supported by National Science Foundation grant number DMS-9796021 and by a Sloan Research Fellowship.

- the measure $d\mu$.

Let B be an abelian subvariety of A. We define the degree $d_B(x_n)$ of x_n modulo B as the degree $[K(\bar{x}_n) : K]$, where $\bar{x}_n$ is the image of x_n in A/B. Since A has only countably many abelian subvarieties, if we replace $(x_n, n \in \mathbb{N})$ with a subsequence, we may assume that, for any abelian subvariety B, either $d_B(x_n)$ remains bounded or $\lim_{n\to\infty} d_B(x_n) = \infty$. Obviously, there is a minimal abelian subvariety C such that $d_C(x_n)$ remains bounded. This C is unique.

Let y_n denote the image of x_n in A/C, via the projection

$$\pi : A \longrightarrow A/C.$$

Then the elements $y_n^\sigma - y_n$ with $n \in \mathbb{N}, \sigma \in G$ have the bounded degree and heights going to 0. By the Northcott theorem, these elements are in a finite list of torsion points for n sufficiently large. With $(y_n, n \in \mathbb{N})$ replaced by a subsequence, we may assume the following:

- there is a fixed subset T of torsion points such that, for any n,

$$\left\{ y_n^\sigma - y_n : \sigma \in G \right\} = T;$$

- the sequence $(y_n, n \in \mathbb{N})$ has a limit $b \in A(\mathbb{C})/C(\mathbb{C})$ in $\mathbb{C}$-topology.

We call x_n $(n \in \mathbb{N})$, obtained in the above manner, a sequence of almost division points with the *coherent limit* $(C, b+T)$. The following is the main result of this paper.

THEOREM 1.1. *Let $x_n, n \in \mathbb{N}$ be a sequence of almost division points with the coherent limit $(C, b+T)$ as above. Then δx_n^G converges to the measure*

$$d\mu = \frac{1}{|T|} \sum_{t\in T} \delta_{\pi^{-1}(b+t)},$$

where π is the projection $A \to A/C$ and $\delta_{\pi^{-1}(b+t)}$ is the $C(\mathbb{C})$-invariant probability measure supported in $\pi^{-1}(b+t)$.

As an application, we show the following theorem about subvarieties.

THEOREM 1.2. *Let X be a subvariety of $A_{\bar{K}}$ that is not a translation of an abelian subvariety. Then there is an $\epsilon > 0$ such that the subset*

$$\left\{ x \in X(\bar{K}) : d\big(x, A(K) \otimes \mathbb{R}\big) \le \epsilon \right\}$$

is not Zariski-dense. Here the distance function in $A(\bar{K}) \otimes \mathbb{R}$ is given by a fixed Neron-Tate height pairing.

Remarks. (1) Theorem 1.2 has previously been conjectured by B. Poonen. Recently, he proved it independently in [5], using a slightly different argument. When $\dim X = 1$, the above theorem is a special case of a conjecture in [10].

(2) Since the above subset contains the division points of the Mordell-Weil group, the above theorem therefore implies the Mordell-Lang conjecture (see [6], [4]). As in

M. Raynaud's proof, we assume Faltings's theorem on Lang's conjecture. However, our arguments are purely in "height theory," rather than Galois theory, on torsion or division points.

(3) Both Theorem 1.1 and Theorem 1.2 can be generalized to semiabelian varieties if we assume the equidistribution theorem in this case. Here, $A(K)$ is replaced by any finitely generated subgroup Γ of $A(K)$, and almost division points (with respect to Γ) mean that

$$\lim_{n\to\infty} d(x_n; \Gamma\otimes\mathbb{Q}) = 0.$$

This is true, for example, for multiplicative group by a result of Y. Bilu [1], and for the split case communicated to me by A. Chambert-Loir.

What can we say about the points with large distance to $A(K)\otimes\mathbb{Q}$? Using Faltings's proof in [2] and [3], we can strengthen Theorem 1.2 to the following theorem.

THEOREM 1.3. *There are positive numbers α and β such that the subset*

$$\left\{x \in X(\bar{K}) : d\big(x, A(K)\otimes\mathbb{R}\big) \le \alpha\|x\| + \beta\right\}$$

is not Zariski-dense in X.

By Theorem 1.2, it suffices to prove that the subset

$$\left\{x \in X(\bar{K}) : \|x\| \ge H,\ d\big(x, A(K)\otimes\mathbb{R}\big) \le \epsilon\|x\|\right\}$$

is not Zariski-dense for some positive numbers H and ϵ. Since

$$d\big(x, A(K)\otimes\mathbb{R}\big) = \|x\| \inf{}^{v\in A(K)\otimes\mathbb{R}}_{\|v\|=1} \sin\angle(x, v),$$

where $\angle(x, v) \in [0, \pi]$ denotes the angle between x and v, it suffices to show that, for any unit vector v of $A(K)\otimes\mathbb{R}$, the set

$$\left\{x \in X(\bar{K}) : \|x\| \ge H,\ \angle(x, v) \le \epsilon\right\}$$

is not Zariski-dense for some $\epsilon > 0$. If this is not true, then we have a unit vector $v \in A(K)\otimes\mathbb{R}$, and a sequence of points $(y_n, n \in \mathbb{N})$ such that we have the following:

- $\lim_{n\to\infty} \|y_n\| = \infty$;
- $\lim_{n\to\infty} \angle(y_n, v) = 0$;
- the set $\{y_n, n \in N\}$ has finite intersection with any proper subvariety of X.

Now we can copy Faltings's proof [3, Theorem 4.1] for points $x_1, \ldots, x_m$ chosen in $\{y_n, n \in N\}$. The only difference is that x_m's are no longer defined over K.

Acknowledgment. I want to thank Bjorn Poonen for interesting discussions and for his lemma on the limit measures, which is used in my proof of Theorem 1.2.

2. Some reductions. In this section we want to reduce Theorem 1.1 into a statement about the equidistribution of small points. First, we notice that the condition *almost division with the coherent limit* $(C, b+T)$ on x_n and the convergence of δx_n^G depend only on the sets $\{x_n, n \in \mathbb{N}\}$ and $\{\delta x_n^G, n \in \mathbb{N}\}$, respectively, rather than particular orders put on them. So, we may simply talk about notions of a *set of almost division points with the coherent limit* $(C, b+T)$ and *convergence of set of measures*.

First reduction. We may reduce the theorem to the case $T = \{0\}$. Indeed, since y_n is included in a finite generated subgroup of $A(\bar{K})$, we may find a finite extension L of K, such that all $y_n (n \in \mathbb{N})$ and all $t \in T$ are rational over L. Now we can apply the theorem to $A_L = A \otimes L$ and the sets

$$\bigcup_n \{x \in x_n^G,\ \pi(x) = y_n + t\}, \quad t \in T.$$

These are sets of almost division points with the coherent limits $(C, b+t)$ $(t \in T)$.

Second reduction. Now we assume $T = \{0\}$ and reduce the theorem to the equidistribution of D_n in $C(\mathbb{C})$, where

$$D_n := \frac{1}{|Gx_n|^2} \sum_{x, y \in Gx_n} \delta_{x-y}, \quad n \in \mathbb{N}.$$

Indeed, if δx_n^G does not converge to $d\mu$, then, after replacing x_n by a subsequence, we may assume that δx_n^G converges to a measure $d\mu^*$ not equal to $d\mu$. It is easy to show that $d\mu^*$ is supported in $\pi^{-1}(s)$ and that D_n has the limit measure defined by

$$f \longrightarrow \int_{A(\mathbb{C})} \int_{A(\mathbb{C})} f(x-y)\, d\mu^*(x)\, d\mu^*(y),$$

for any continuous function f on $A(\mathbb{C})$. Assume D_n is equidistributed; then we must have

$$\int_{A(\mathbb{C})} \int_{A(\mathbb{C})} f(x-y)\, d\mu^*(x)\, d\mu^*(y) = \int_{C(\mathbb{C})} f\, dx,$$

where dx is the Haar measure on $C(\mathbb{C})$. Let a be a fixed point in $\pi^{-1}(b)$; then $\pi^{-1}(b) = a + C$. So $d\mu^*$ is induced by a measure $d\mu'$ such that

$$\int_{A(\mathbb{C})} f\, d\mu^* = \int_{C(\mathbb{C})} f(a+x)\, d\mu'(x).$$

Since every continuous function on $a+C$ can be extended to a continuous function on $A(\mathbb{C})$, the above formula implies that

$$\int_{C(\mathbb{C})} \int_{C(\mathbb{C})} f(x-y)\, d\mu'(x)\, d\mu'(y) = \int_{C(\mathbb{C})} f\, dx$$

for any continuous function f on $C(\mathbb{C})$. In particular, if χ is any nontrivial character on $C(\mathbb{C})$, we obtain

$$\left| \int_{C(\mathbb{C})} \chi\, d\mu' \right|^2 = 0.$$

In other words,

$$\int_{C(\mathbb{C})} \chi \, d\mu' = \int_{C(\mathbb{C})} \chi \, dx$$

for all characters. Since the space of continuous function on $C(\mathbb{C})$ is generated topologically by all characters of $C(\mathbb{C})$, we must have $d\mu' = dx$. In other words, the measure $d\mu^*$ is $C(\mathbb{C})$-invariant. We therefore obtain a contradiction.

3. Equidistribution of D_n. For Theorem 1.1, it remains to prove the following proposition.

PROPOSITION 3.1. *The sequence of measures D_n converges to the Haar measure dx of $C(\mathbb{C})$.*

Proof. First, we note that D_n is a linear combination of the uniform probability measures of some Galois orbits of small points. Indeed, for each n, let H_n be a finite quotient that corresponds to a finite Galois extension L_n of K such that x_n is rational over L_n. Then we can rewrite Dx_n as

$$Dx_n = \frac{1}{|H_n|^2} \sum_{\sigma \in H_n} \sum_{\tau \in H_n} \delta_{(x_n^\sigma - x_n)^\tau} = \frac{1}{\left|x_n^G\right|} \sum_{x \in x_n^G} \delta(x - x_n)^G.$$

Here, as before, $\delta(x - x_n)^G$ denotes the uniform probability measure of the Galois orbits of $x - x_n$.

There are only countably many closed reduced subvarieties of C which are unions of varieties of the form $B + C[N]$, with B a proper abelian subvariety of C, and N a positive integer. We may find a sequence X_i of reduced closed subvarieties of C over K such that we have the following:

- each X_i is a union of subvarieties of the form $B + C[N]$;
- $X_i \subset X_{i+1}$;
- for any $B + C[N]$, there is an X_i including $B + C[N]$.

By the equidistribution theorem for small points [9], the set

$$\left\{\delta(y - x_n) : y \in x_n^G,\ y - x_n \notin X_n\right\}$$

of measures converges to the Haar measure of $C(\mathbb{C})$. We apply this fact to some subsequence of x_n.

For any proper closed subvariety X defined over K, we define a positive rational number $\alpha_{X,n}$ by the formula

$$\alpha_X = \frac{\left|x_n^G \cap X\right|}{\left|x_n^G\right|}.$$

We want to reduce the equidistribution of D_n to the fact

$$\lim_{n \to \infty} \alpha_{X,n} = 0,$$

for each X of the form $B+C[N]$. As for any two closed subvarieties X and Y of C,

$$\alpha_{X\cup Y,n} \le \alpha_{X,n} + \alpha_{Y,n},$$

we see that

$$\lim_{n\to\infty} \alpha_{X_i,n} = 0,$$

for any X_i. Now by the *diagonal process,* we may find a subsequence n_i of $\mathbb{N}$, such that

$$\lim_{i\to\infty} \alpha_{X_i,n_i} = 0.$$

If D_n does not have the limit $d\mu$, then there is a continuous function f, such that $\int f D_n$ does not converges to $\int f d\mu$. Since $\int f D_n$ has its absolute value bounded by $\|f\|_{\sup}$, if we replace x_n with a subsequence, we may assume that $\int f D_n$ converges to a number not equal to $\int f d\mu$. Now consider the expression

$$\int_{C(\mathbb{C})} f D_{n_i} = \frac{1}{\left|x_{n_i}^G\right|} \sum_{\substack{x\in x_{n_i}^G \\ x\in X_i}} \int_{C(\mathbb{C})} f\delta(x-x_{n_i})^G + \frac{1}{\left|x_{n_i}^G\right|} \sum_{\substack{x\in x_{n_i}^G \\ x\notin X_i}} \int_{C(\mathbb{C})} f\delta(x-x_{n_i})^G.$$

The first summand in the right-hand side has the absolute value bounded by α_{X_i,n_i} $\|f\|_{\sup}$, while the second summand approaches

$$\left(1-\alpha_{X_i,n_i}\right) \int_{C(\mathbb{C})} f\,dx,$$

by the equidistribution theorem of small points. The right-hand side therefore has the limit $\int_{C(\mathbb{C})} f\,dx$. This shows that the subsequence D_{n_i} $(i \in \mathbb{N})$ has the limit dx. This gives a contradiction.

Now we fix an X of the form $B+C[N]$ and prove that

$$\lim_{n\to\infty} \alpha_{X,n} = 0.$$

Replacing A by A/B, we may assume that $B=0$. Our assumption on C implies that $\deg x_n \to \infty$. Consider the multiplication by N:

$$u : x_n^G \longrightarrow (Nx_n)^G.$$

Then we have

$$\left|u^{-1}(Nx_n)\right| = \frac{\deg x_n}{\deg(Nx_n)},$$

since the Galois group G acts transitively on the set of fibers of u. It follows that, for $n \gg 0$,

$$\alpha_{X,n} = \frac{\left|u^{-1}(Nx_n)\right|}{\left|x_n^G\right|} = \frac{1}{\deg(Nx_n)} \le \frac{N^{2\dim C}}{\deg x_n} \longrightarrow 0. \qquad \square$$

4. Proof of Theorem 1.2. Replacing K by a large field, we may assume that X is defined over K. Assume that the theorem is not true. Then we can find a sequence (x_n, z_n) of pairs of points, with $x_n \in X(\bar{K})$ and $z_n \in A(K)\otimes\mathbb{Q}$ such that we have the following:

- $\lim_{n\to\infty}\|x_n - z_n\| = 0$;
- the sequence x_n converges to the generic point of X with respect to the Zariski topology on X.

As the Galois group acts trivially on $A(K)\otimes\mathbb{Q}$ and preserves the Neron-Tate height pairing on $A(\bar{K})$, we see that $(x_n, n\in\mathbb{N})$ is a sequence of almost division points in $A(\bar{K})$. With $(x_n, n\in\mathbb{N})$ replaced by a subsequence, we may assume that this sequence has the coherent limit $(C, b+T)$. Again, we may enlarge K and replace $(x_n, n\in N)$ by a subsequence in the union of x_n^G, as in §2, we may assume that $T=\{0\}$. Now Theorem 1.1 implies that the measures δx_n^G converges to the Haar measure $d\mu$ in the fiber over b of the map $\pi : A\to B := A/C$.

Now consider the following diagram with $Y=\pi(X)$:

$$\begin{array}{ccc} X & \hookrightarrow & A \\ \downarrow & & \downarrow \\ Y & \hookrightarrow & B. \end{array}$$

Our assumption implies that the sequence $y_n=\pi(x_n)$ is rational over K and Zariski-dense in Y. By Faltings's theorem, Y must be the translate of an abelian subvariety. As X itself is not the translate of any abelian subvariety, $X\neq\pi^{-1}(Y)$. This implies that

$$\dim X < \dim Y + \dim C.$$

As the measures δx_n^G are supported in the fiber $X(\mathbb{C})$ over y_n, we obtain a contradiction by the following lemma.

Lemma 4.1 (B. Poonen). *Let $\pi : X\to Y$ be a surjective morphism of projective and integral complex varieties. Let $(y_n, n\in\mathbb{N})$ be a sequence of points in Y that converges to the generic point of Y, with respect to the Zariski topology. Let $d\mu_n$ be a sequence of probability measures of X that converges (weakly) to a measure $d\mu$ on X. Assume each $d\mu_n$ is supported in the fiber of π over y_n. Then $d\mu$ is supported in a closed subvariety of X of the dimension* $=\dim X - \dim Y$.

Proof. We take an embedding $X\to\mathbb{P}^N$ and consider $d\mu_n$ and $d\mu$ as measures in $\mathbb{P}^N$. Let P be the Hilbert polynomial of the generic fiber of π and let $\mathcal{X}\to\mathbb{P}^N\times\mathcal{Y}$ be the universal family of the subvarieties of $\mathbb{P}^N$ with the Hilbert polynomial P. Then, for n sufficiently large, π is flat at points over y_n, and the fiber X_{y_n} is therefore given by a point p_n in $\mathcal{Y}$. Since the Hilbert scheme $\mathcal{Y}$ is projective, p_n converges to a point p in $\mathcal{Y}$, with respect to the $\mathbb{C}$-topology. So the measure $d\mu$, as the limit of some measures on the fibers of $\mathcal{X}\to\mathcal{Y}$ over p_n, is supported in the fiber over p. This shows that the Zariski closure (as a subvariety of $\mathbb{P}^N$) of the support of $d\mu$ has the dimension less than or equal to $\dim X - \dim Y$. $\square$

References

[1] Y. Bilu, *Limit distribution of small points on algebraic tori,* Duke Math. J. **89** (1997), 467–476.

[2] G. Faltings, *Diophantine approximation on abelian varieties,* Ann. of Math. (2) **133** (1991), 549–576.

[3] ———, "The general case of S. Lang's conjecture" in *Barsatti Symposium in Algebraic Geometry (Abano Term, 1991),* Perspect. Math. **15**, Academic Press, San Diego, 1994, 175–182.

[4] M. McQuillan, *Division points on semiabelian varieties,* Invent. Math. **120** (1995), 143–159.

[5] B. Poonen, *Mordell-Lang plus Bogomolov,* Invent. Math. **137** (1999), 413–425.

[6] M. Raynaud, "Sous-variétés d'une variétés abélienne et points de torsion" in *Arithmetic & Geometry, Vol. I,* Progr. Math. **35**, Birkhäuser, Boston, 1983, 327–352.

[7] L. Szpiro, E. Ullmo, and S. Zhang, *Equirépartition des petits points des courbes,* Invent. Math. **127** (1997), 337–347.

[8] E. Ullmo, *Positivité et discrétion des points algébriques,* Ann. of Math. (2) **147** (1998), 167–179.

[9] S. Zhang, *Equidistribution of small points on abelian varietes,* Ann. of Math. (2) **147** (1998), 159–165.

[10] ———, "Small points and Arakelov theory" in *Proceedings of the International Congress of Mathematicians '98, Vol. II (Berlin, 1998),* Doc. Math. (1998), 217–225.

Department of Mathematics, Columbia University, New York, New York 10027 USA; szhang@math.columbia.edu