**Article:**

Vishe, Pankaj Hemant and Gorodnik, Alexander (2018) Diophantine approximation for products of linear maps—Logarithmic improvements. Transactions of the AMS. ISSN: 1088-6850

https://doi.org/10.1090/tran/6953

# DIOPHANTINE APPROXIMATION FOR PRODUCTS OF LINEAR MAPS — LOGARITHMIC IMPROVEMENTS

ALEXANDER GORODNIK AND PANKAJ VISHE

Abstract. This paper is devoted to the study of a problem of Cassels in multiplicative Diophantine approximation which involves minimising values of a product of affine linear forms computed at integral points. It was previously known that values of this product become arbitrary close to zero, and we establish that, in fact, they approximate zero with an explicit rate. Our approach is based on investigating quantitative density of orbits of higher-rank abelian groups.

## 1. Introduction

Let $\langle u\rangle$ denote the distance of the real number $u$ to the nearest integer. The sequence $\langle qu\rangle$ with $q\in\mathbb{N}$ reflects how well $u$ is approximated by rational numbers. In particular, it is well-known that for every $Q\geq 1$ one can find $q\leq Q$ such that $\langle qu\rangle\leq 1/Q$, but there is a large set of numbers $u$ satisfying $\langle qu\rangle\geq c(u)/q$ for all $q$'s with some $c(u)>0$. The long-standing Littlewood conjecture concerns simultaneous approximation of a pair of real numbers $u,v\in\mathbb{R}$. It asserts that

$$\liminf_{q\to\infty} q\,\langle qu\rangle\,\langle qv\rangle=0 \tag{1}$$

holds for all $u,v\in\mathbb{R}$. This paper deals with the inhomogeneous version of this problem, namely, whether the following relation

$$\liminf_{|q|\to\infty} |q|\,\langle qu-\alpha\rangle\,\langle qv-\beta\rangle=0 \tag{2}$$

holds for $u,v,\alpha,\beta\in\mathbb{R}$. In this setting Cassels asked (see [4, p. 307]) whether there exists a pair $(u,v)$ for which the property (2) holds for all real numbers $\alpha,\beta$. This question was answered affirmatively by Shapira [16] who showed that this is true for almost all pairs $(u,v)$. He also gave an explicit example of a family of algebraic numbers $(u,v)$ satisfying this property, and showed that it fails if $u$ and $v$ are rationally dependent.

It is natural to ask for which pairs $(u,v)$ does the condition (1) hold and/or admit quantitative improvements. It follows from the results of Gallagher [8] that for almost every $(u,v)\in\mathbb{R}^2$,

$$\liminf_{q\to\infty}\,(\log q)^2 q\,\langle qu\rangle\,\langle qv\rangle=0. \tag{3}$$

Peck [14] showed if $1,u,v$ form a basis of a real cubic field, then

$$\liminf_{q\to\infty}\,(\log q) q\,\langle qu\rangle\,\langle qv\rangle<\infty. \tag{4}$$

Pollington and Velani [15] proved that (1) holds with an additional $\log q$ factor for a large set of pairs $(u,v)$, and Badziahin and Velani [2] conjectured that (4) holds for all real numbers $u$ and $v$.

Unlike in the *homogeneous* setting, literature on quantitative results in the *inhomogeneous* setting has been lacking. An old argument of Cassels readily implies that for almost all $(u,v,\alpha,\beta)\in\mathbb{R}^4$,

$$\liminf_{q\to\infty}(\log q)^2 q\,\langle qu-\alpha\rangle\,\langle qv-\beta\rangle=0$$

(see, for instance, [10, Theorem 3.3]). The case with $\alpha = 0$ was investigated by Haynes, Jensen and Kristensen in [11]. They proved that for all badly aproximable $u$, and $v$ contained in a set of badly approximable numbers of full Hausdorff dimension depending on $u$,

$$\liminf_{q\to\infty} (\log q)^{1/2-\epsilon} q \langle qu \rangle \langle qv - \beta \rangle = 0 \quad \text{with any } \epsilon > 0$$

holds for all $\beta$. Setting $\alpha = 0$ allowed in [11] to use tools developed in [15], but it seems unlikely that this approach could be applied when $\alpha$ is non-zero.

Apart from these results, no other quantitative improvements of the inhomogeneous property (2) are known to us. The aim of this paper is to establish the first quantitative improvement of (2) with arbitrary $\alpha, \beta$. In contrast with the existing analytical methods, dynamical ideas employed in this paper enable us to successfully deal with general $\alpha, \beta$ at a cost of a weaker logarithmic saving. The following theorem is a quantitative refinement of one of the main results from [16].

**Theorem 1.** *There exists $\delta > 0$ such that for almost all $(u, v) \in \mathbb{R}^2$,*

$$\liminf_{|q|\to\infty} (\log_{(5)} |q|)^{\delta} |q| \langle qu - \alpha \rangle \langle qv - \beta \rangle = 0$$

*holds for all $\alpha, \beta \in \mathbb{R}$. Here $\log_{(s)}$ denotes the $s$-th iterate of the function $x \mapsto \max(1, \log |x|)$.*

We note that our method, in principle, could also allow establishing this result for specific pairs $(u, v)$ provided that corresponding orbits satisfy a certain quantitative recurrence property.

In a subsequent paper [9], we also extend Theorem 1 to the $p$-adic setting motivated by the $p$-adic version of the Littlewood conjecture proposed by de Mathan and Teulié [6].

The setting of Theorem 1 can be considered as a particular case of a general problem of multiplicative Diophantine approximation for affine lattices (also called grids) in the Euclidean space $\mathbb{R}^d$. A grid in $\mathbb{R}^d$ is a subset of the form

$$\mathbb{Z}x_1 + \cdots + \mathbb{Z}x_d + w,$$

where $x_1, \ldots, x_d \in \mathbb{R}^d$ are linearly independent and $w \in \mathbb{R}^d$. To formulate this problem explicitly, we set $N(v) := v_1 v_2 \cdots v_d$ for a vector $v = {}^t(v_1, \ldots, v_d)$ in $\mathbb{R}^d$.

**Definition 1.1.** Let $\Lambda$ be a grid in $\mathbb{R}^d$ and $h : \mathbb{R}^+ \to [1, \infty)$ a function such that $h(x) \to \infty$ as $x \to \infty$.

(i) We say that $\Lambda$ is *multiplicatively approximable* if 0 is a non-trivial accumulation point of a sequence $N(v_n)$ with $v_n \in \Lambda$.

(ii) We say that $\Lambda$ is *$h$-multiplicatively approximable* if there exists a sequence $v_n \in \Lambda$ such that $v_n \to \infty$ and $0 < |N(v_n)| < h(\|v_n\|)^{-1}$, where $\| \cdot \|$ denotes the max norm.

We note that this notion is related to property (2) in the following way. For $u, v, \alpha, \beta \in \mathbb{R}$, we consider the grid

$$\Lambda(u, v, \alpha, \beta) := \{{}^t(x, xu - y - \alpha, xv - z - \beta) : \; x, y, z \in \mathbb{Z}\}. \tag{5}$$

It is easy to check that if the grid $\Lambda(u, v, \alpha, \beta)$ is multiplicatively approximable, then (2) holds. Moreover, assuming that the function $h$ is non-decreasing, if the grid $\Lambda(u, v, \alpha, \beta)$ is $h$-multiplicatively approximable, then

$$\liminf_{|q|\to\infty} h(|q|) |q| \langle qu - \alpha \rangle \langle qv - \beta \rangle \le 1.$$

It was also proved in [16] that for almost every lattice $\Lambda$ in $\mathbb{R}^d$, the grid $\Lambda + v$ is multiplicatively approximable for all $v \in \mathbb{R}^d$. Here we establish a quantitative refinement of this result.

**Theorem 2.** *There exists $\delta > 0$ such that for almost every lattice $\Delta$ in $\mathbb{R}^d$ with $d \ge 3$, every grid $\Delta + w$, $w \in \mathbb{R}^d$, is $h$-multiplicatively approximable with $h(x) = (\log_{(5)} x)^{\delta}$.*

We note that this theorem fails for $d = 2$ (see [5]).

The paper is organised as follows. In the following section we set up required notation and give a dynamical reformulation of the problem, which reduces our investigation to the study of a quantitative recurrence property for orbits of a higher-rank abelian group $A$ acting on the space of grids in the Euclidean space. However, it is not easy to establish this recurrence property directly, so in Section 3, we first investigate quantitative recurrence in a smaller space — the space of lattices. In particular, it would be crucial in the proof to establish recurrence to neighbourhoods of lattices with compact $A$-orbits. In Section 4, we discuss properties of compact orbits and relevant density results. Finally, in Section 5 we give a proof of the main theorems by performing local analysis in a neighbourhood of a grid whose corresponding lattice has compact $A$-orbit.

1.1. **Acknowledgements.** The authors would like to thank S. Velani for suggesting the problem and for his encouragement during the work on the project. The first author was supported by ERC grant 239606, and the second author was supported by EPSRC programme grant EP/J018260/1.

## 2. Preliminaries

In this section we introduce some basic notation regarding dynamics on the space of grids in $\mathbb{R}^d$ and give a dynamical reformulation of the above Diophantine approximation problem. We also introduce a collection of root subgroups that provides a convenient system of local coordinates.

2.1. **Space of grids.** Let $G$ denote the group of unimodular affine transformations of $\mathbb{R}^d$. Let us set $G_0 := \mathrm{SL}(d, \mathbb{R})$ and $V := \mathbb{R}^d$. Then $G \simeq V \rtimes G_0$. For $g \in G$, we write $g = (v, g_0)$ with $v \in V$ and $g_0 \in G_0$. We also set $\Gamma_0 := \mathrm{SL}(d, \mathbb{Z})$ and $\Gamma := \mathbb{Z}^d \rtimes \Gamma_0$. Then $\Gamma_0$ is lattice in $G_0$, and $\Gamma$ is lattice in $G$. The space $X := G_0/\Gamma_0$ can be identified with the space of unimodular lattices in $\mathbb{R}^d$, and the space $Y := G/\Gamma$ can be identified with the space of affine unimodular lattices, which are also called unimodular grids. For $x \in X$ we denote by $\Delta_x$ the corresponding lattice in $\mathbb{R}^d$, and for $y \in Y$, we denote by $\Lambda_y$ the corresponding grid. We denote by $\pi : Y \to X$ the natural factor map. We observe that $\Lambda_y = \Delta_{\pi(y)} + w$ for some $w \in V$. Moreover, $w$ can chosen to be uniformly bounded when $\pi(y)$ varies over bounded subsets of $X$.

2.2. **Dynamical approach to the multiplicative approximability property.** We show that the multiplicative approximability property can be reformulated in terms of dynamics of the group

$$A := \{a = \mathrm{diag}(a_1, \dots, a_d) : a_i > 0\}$$

acting on the space $Y$. More specifically, we show that the grid $\Lambda_y$ is $h$-multiplicatively approximable if the orbit $Ay$ visits certain shrinking subsets $\mathcal{W}(\vartheta, \varepsilon)$ of $Y$. Given $\varepsilon, \vartheta > 0$, we introduce the following non-empty open subset of $Y$

$$\mathcal{W}(\vartheta, \varepsilon) := \{y \in Y : \exists v \in \Lambda_y \text{ such that } \|v\| < \vartheta \text{ and } 0 < |N(v)| < \varepsilon\}.$$

We also denote by $\|\cdot\|$ the maximum norm on $\mathrm{Mat}(d, \mathbb{R})$, and for a subset $S$ of $\mathrm{Mat}(d, \mathbb{R})$, we set

$$S(T) := \{s \in S : \|s\| < T\}.$$

**Proposition 3.** *Let $h$ be a non-decreasing function such that $h(x) \to \infty$ as $x \to \infty$. Suppose that for $y \in Y$,*

$$\text{(WR)} \qquad \exists\, T_n \to \infty : \quad A(T_n)y \cap \mathcal{W}(T_n, h(T_n^d)^{-1}) \neq \emptyset.$$

*Then the grid $\Lambda_y$ is $h$-multiplicatively appoximable.*

*Proof.* It follows from our assumption that there exist sequences $a^{(n)} \in A(T_n)$ and $v^{(n)} \in \Lambda_y$ such that

$$|a_i^{(n)} v_i^{(n)}| < T_n \quad \text{for all } i,$$

and

$$|N(a^{(n)} v^{(n)})| = |N(v^{(n)})| \in (0, h(T_n^d)^{-1}).$$

This, in particular, implies that $0 \neq N(v^{(n)}) \to 0$, so that $v^{(n)} \to \infty$. We deduce from the first inequality that

$$|v_i^{(n)}| < \left(a_i^{(n)}\right)^{-1} T_n = \left(\prod_{j \neq i} a_j^{(n)}\right) T_n \leq T_n^d.$$

Hence, $\|v^{(n)}\| \leq T_n^d$, and since $h$ is non-decreasing, we conclude that

$$0 < |N(v^{(n)})| < h(\|v^{(n)}\|)^{-1}.$$

This proves that the grid $\Lambda_y$ is $h$-multiplicatively appoximable. $\square$

Proposition 3 suggests a dynamical approach to the problem of multiplicative approximation through analysing property (WR) — the quantitative recurrence property of $A$-orbits with respect to the sets $\mathcal{W}(\vartheta, \varepsilon)$ in $Y$.

2.3. **Root subgroups.** The crucial ingredient in understanding dynamics of the $A$-action on the spaces $X$ and $Y$ are the root subgroups, which we now introduce. The adjoint action of $A$ on the Lie algebra of $G$ is diagonalisable, and we denote by $\Phi(G)$ the set of roots of $A$ which is the set of non-trivial eigencharacters of $A$ appearing in this action. For each $\alpha \in \Phi(G)$, there is a one-parameter root subgroup $U_\alpha = \{u_\alpha(t)\}_{t \in \mathbb{R}} \subset G$ such that

$$a u_\alpha(t) a^{-1} = u_\alpha(\alpha(a) t) \quad \text{for } a \in A \text{ and } t \in \mathbb{R}.$$

More explicitly, the set of roots consists of

$$\alpha_{ij}(a) = a_i a_j^{-1} \text{ for } 1 \leq i \neq j \leq d \text{ and } \beta_i(a) = a_i \text{ for } 1 \leq i \leq d.$$

The corresponding root subgroups are the groups of affine transformations defined by

$$u_{ij}(t)u = u + t u_j e_i \text{ and } v_i(t)u = u + t e_i \qquad \text{for } u \in \mathbb{R}^d,$$

where $e_1, \ldots, e_d$ denotes the standard basis of $\mathbb{R}^d$. We denote the set of roots of the first type by $\Phi(G_0)$ and the set of roots of the second type by $\Phi(V)$. With a suitable ordering, the product maps

$$A \times \prod_{\alpha \in \Phi(G_0)} \mathbb{R} \to G_0 : \ (a, t_\alpha : \alpha \in \Phi(G_0)) \mapsto a \left( \prod_{\alpha \in \Phi(G_0)} u_\alpha(t_\alpha) \right),$$

$$\prod_{\alpha \in \Phi(V)} \mathbb{R} \to V : \ (t_\alpha : \alpha \in \Phi(G_0)) \mapsto \prod_{\alpha \in \Phi(V)} u_\alpha(t_\alpha),$$

and

$$A \times \prod_{\alpha \in \Phi(G)} \mathbb{R} \to G : \ (a, t_\alpha : \alpha \in \Phi(G)) \mapsto \left( \prod_{\alpha \in \Phi(V)} u_\alpha(t_\alpha) \right) a \left( \prod_{\alpha \in \Phi(G_0)} u_\alpha(t_\alpha) \right)$$

are diffeomorphisms in neighbourhoods of the origins. We set

$$\mathcal{U}_{G_0}(\varepsilon) := \{a \in A : \|a - e\| < \varepsilon\} \cdot \prod_{\alpha \in \Phi(G_0)} \{u_\alpha(t_\alpha) : |t_\alpha| < \varepsilon\},$$

$$\mathcal{U}_V(\varepsilon) := \prod_{\alpha \in \Phi(V)} \{u_\alpha(t_\alpha) : |t_\alpha| < \varepsilon\}, \tag{6}$$

$$\mathcal{U}_G(\varepsilon) := \mathcal{U}_V(\varepsilon) \mathcal{U}_{G_0}(\varepsilon).$$

Then $\mathcal{U}_{G_0}(\varepsilon)$, $\mathcal{U}_V(\varepsilon)$, and $\mathcal{U}_G(\varepsilon)$ define neighbourhoods of identity in the groups $G_0$, $V$, and $G$ respectively. We also consider the neighbourhoods of identity

$$\begin{aligned} \mathcal{O}_{G_0}(\varepsilon) &:= \{g \in G_0 : \ \|g - e\| < \varepsilon\}, \\ \mathcal{O}_V(\varepsilon) &:= \{v \in V : \ \|v\| < \varepsilon\}, \\ \mathcal{O}_G(\varepsilon) &:= \{(v, g) \in G : \ \|v\| < \varepsilon, \ \|g - e\| < \varepsilon\}. \end{aligned} \tag{7}$$

It is easy to check that there exists $c_0 > 0$ such that for every $\varepsilon \in (0, 1)$,

$$\mathcal{U}_{G_0}(\varepsilon) \subset \mathcal{O}_{G_0}(c_0\,\varepsilon), \quad \mathcal{U}_V(\varepsilon) \subset \mathcal{O}_V(c_0\,\varepsilon), \quad \mathcal{U}_G(\varepsilon) \subset \mathcal{O}_G(c_0\,\varepsilon). \tag{8}$$

While establishing quantitative recurrence of $A$-orbits to the sets $\mathcal{W}(\vartheta, \varepsilon)$ is the crux of the proof of our main results, it turns out that analogous recurrence property is easy to verify for orbits of the root subgroups. In fact, as an intermediate step in the proof, we will have to establish recurrence to smaller sets which are defined as

$$\mathcal{W}(\vartheta, \varepsilon_1, \varepsilon_2) := \{y \in Y : \exists v \in \Lambda_y \text{ such that } \|v\| < \vartheta \text{ and } \varepsilon_1 < |N(v)| < \varepsilon_2\}$$

for $\vartheta > 0$ and $0 < \varepsilon_1 < \varepsilon_2$.

**Lemma 4.** *Let $\alpha \in \Phi(G)$. For every $\varepsilon_1, \varepsilon_2 \in (0, 1)$, $\varepsilon_1 < \varepsilon_2$, and $y \in Y$, there exist positive $\vartheta = O_{\pi(y)}(1)$, positive $t_+ = O_{\pi(y)}(1)$, and negative $t_- = O_{\pi(y)}(1)$ such that*

$$u_\alpha(t_+)y \in \mathcal{W}(\vartheta, \varepsilon_1, \varepsilon_2) \quad \text{and} \quad u_\alpha(t_-)y \in \mathcal{W}(\vartheta, \varepsilon_1, \varepsilon_2).$$

*Proof.* We first note that the grid $\Lambda_y$ can be written as $\Lambda_y = \Delta_{\pi(y)} + w$, where $w$ belongs to a fixed fundamental domain for the lattice $\Delta_{\pi(y)}$. In particular, $\|w\| = O_{\pi(y)}(1)$.

Let us show that $v_i(t)y \in \mathcal{W}(\vartheta, \varepsilon_1, \varepsilon_2)$ for some positive $\vartheta = O_{\pi(y)}(1)$ and some positive $t = O_{\pi(y)}(1)$. Using that $\|w\| = O_{\pi(y)}(1)$ and adding a suitable vector from the lattice $\Delta_{\pi(x)}$, one can show that there exists a vector $z \in \Lambda_y = w + \Delta_{\pi(y)}$ such that

$$\|z\| = O_{\pi(y)}(1), \quad z_i < 0, \quad |z_k| \geq 1 \quad \text{for all } k.$$

Indeed, since $\Delta_{\pi(y)}$ is a lattice, there exists $s \in \Delta_{\pi(y)}$ satisfying $s_i < 0$ and $s_k \neq 0$ for all $k$. Then we can choose $z$ of the form $z = w + \ell s$ with a suitable $\ell \in \mathbb{N}$. We have to choose $t$ so that the inequalities

$$\varepsilon_1 < |N(v_i(t)z)| < \varepsilon_2$$

hold. Since $N(v_i(t)z) = N(z) + tN_i(z)$ where $N_i(z) := \prod_{j \neq i} z_i$, these inequalities are equivalent to

$$\varepsilon_1 |N_i(z)|^{-1} < |z_i + t| < \varepsilon_2 |N_i(z)|^{-1}.$$

Hence, we can take $t$ from the interval $(\varepsilon_1 |N_i(z)|^{-1} - z_i, \varepsilon_2 |N_i(z)|^{-1} - z_i)$. Due to our choice of $z$, we have $t > 0$ and $t = O_{\pi(y)}(1)$. Also, $\|v_i(t)z\| = O_{\pi(y)}(1)$. Hence, it follows that $v_i(t)y \in \mathcal{W}(\vartheta, \varepsilon_1, \varepsilon_2)$ with some $\vartheta = O_{\pi(y)}(1)$ as required. Similarly, one can also show that there exists negative $t$ satisfying $v_i(t)y \in \mathcal{W}(\vartheta, \varepsilon_1, \varepsilon_2)$.

The proof that $u_{ij}(t)y \in \mathcal{W}(\vartheta, \varepsilon_1, \varepsilon_2)$ for some positive $\vartheta = O_{\pi(y)}(1)$ and some positive $t = O_{\pi(y)}(1)$ follows similar lines. Since $\|w\| = O_{\pi(y)}(1)$, we can add to $w$ a vector from the lattice $\Delta_{\pi(x)}$ to show existence of $z \in \Lambda_y = w + \Delta_{\pi(x)}$ satisfying

$$\|z\| = O_{\pi(y)}(1), \quad z_i > 0, \quad z_j < 0, \quad |z_k| \geq 1 \quad \text{for all } k.$$

Since $N(u_{ij}(t)z) = N(z) + tz_jN_i(z)$, the inequalities

$$\varepsilon_1 < |N(u_{ij}(t)z)| < \varepsilon_2$$

are equivalent to

$$\varepsilon_1 |N_i(z)|^{-1} |z_j|^{-1} < |z_i z_j^{-1} + t| < \varepsilon_2 |N_i(z)|^{-1} |z_j|^{-1},$$

so that we can take $t$ from the interval $(\varepsilon_1 |N_i(z)|^{-1} |z_j|^{-1} - z_i z_j^{-1}, \varepsilon_2 |N_i(z)|^{-1} |z_j|^{-1} - z_i z_j^{-1})$. Then $t > 0$ and $t = O_{\pi(y)}(1)$. Also, it is clear that $\|u_{ij}(t)z\| = O_{\pi(y)}(1)$. Hence, it follows that $u_{ij}(t)y \in \mathcal{W}(\vartheta, \varepsilon_1, \varepsilon_2)$ with some $\vartheta = O_{\pi(y)}(1)$. The argument with negative $t$ is similar. □

## 3. Quantitative recurrence estimates

Quantitative recurrence plays an important role in the theory of Diophantine approximation. In particular, this connection was realised in Sullivan's work [17] and its subsequent generalization [12] by Kleinbock and Margulis. While these papers deal with recurrence to shrinking cuspidal neighbourhoods, we have to investigate visits of $A$-orbits to shrinking neighbourhoods of specific points inside the space $X$. The idea of our approach, which uses exponential mixing, is similar to [12], but it will be essential to establish recurrence to neighbourhoods of particular shape with respect to the root coordinate system introduced in Section 2.3. Namely, we consider neighbourhoods of $x \in X$ defined by $\mathcal{U}_\varepsilon(x) := \mathcal{U}_{G_0}(\varepsilon)x$, where $\mathcal{U}_{G_0}(\varepsilon)$ is defined in (6), and $\mathcal{U}^*_\varepsilon(x) := \mathcal{U}^*_{G_0}(\varepsilon)x$, where

$$\mathcal{U}^*_{G_0}(\varepsilon) := \{a \in A : \|a - e\| < 2\varepsilon\} \cdot \prod_{\alpha \in \Phi(G_0)} \{u_\alpha(t_\alpha) : |t_\alpha| < \varepsilon\}.$$

The main goal of this section is to prove the following proposition.

**Proposition 5.** *Let $x_0 \in X$, and $a_t$ be a non-trivial one-parameter subgroup of $A$. Then there exists a constant $\beta > 0$, such that for almost every $x \in X$ and every $T > T_0(x)$,*

$$a_t x \in \mathcal{U}_{T^{-\beta}}(x_0) \backslash \mathcal{U}^*_{T^{-\beta}/2}(x_0) \quad \text{for some } t \in [0, T].$$

We denote by $\mu$ the normalised invariant measure on the space $X$ and consider a family of averaging operators

$$A_T : L^2(X) \to L^2(X) : f \mapsto \frac{1}{T}\int_0^T f(a_t x)dt.$$

We begin by proving an $L^2$-estimate for the operators $A_T$.

**Lemma 6.** *For every $T \geq 1$ and $f \in C_c^\infty(X)$,*

$$\left\| A_T(f) - \int_X f\, d\mu \right\|_2 \ll T^{-1/2} S(f),$$

*where $S(f)$ denotes a suitable Sobolev norm.*

*Proof.* We recall the exponential mixing property (see, for instance, [12, 3.5]): there exists $\alpha > 0$ such that for every $f_1, f_2 \in C_c^\infty(X)$,

$$\int_X f_1(a_t x) f_2(x)\, d\mu(x) = \left(\int_X f_1\, d\mu\right)\left(\int_X f_2\, d\mu\right) + O\left(e^{-\alpha|t|} S(f_1) S(f_2)\right). \tag{9}$$

This property will be used to establish the required $L^2$-bound. Without loss of generality, we can assume that $\int_X f\, d\mu = 0$. Then using (9), we deduce that

$$\begin{aligned}
\|A_T(f)\|_2^2 &= T^{-2}\int_{(t,s)\in[0,T]^2}\int_X f(a_t x)\overline{f}(a_s x)\, d\mu(x)\, dt\, ds \\
&= T^{-2}\int_{(t,s)\in[0,T]^2}\int_X f(a_{t-s} x)\overline{f}(x) d\mu(x)\, dt\, ds \\
&\ll T^{-2}\left(\int_0^T\int_0^T e^{-\alpha|t-s|}\, dt\, ds\right) S(f)^2 \\
&\ll T^{-1} S(f)^2,
\end{aligned}$$

which completes the proof. □

We are now ready to apply a standard Borel-Cantelli type argument to prove Proposition 5.

*Proof of Proposition 5.* Let $\beta \in (0,1)$, to be specified later, and

$$\Omega_T := \{x \in X : a_t x \notin \mathcal{U}_{T^{-\beta}/2^\beta}(x_0) \backslash \mathcal{U}^*_{T^{-\beta}/2}(x_0) \text{ for all } 0 \le t \le T\}.$$

It is obvious from the definition that the neighbourhoods $\mathcal{U}_{G_0}(\varepsilon)$ and $\mathcal{U}^*_{G_0}(\varepsilon)$ are $\varepsilon$-boxes with respect to a suitable smooth coordinate system, so that we can choose a non-negative compactly supported function $f_T$ such that

$$\mathrm{supp}(f_T) \subset \mathcal{U}_{T^{-\beta}/2^\beta}(x_0) \backslash \mathcal{U}^*_{T^{-\beta}/2}(x_0), \qquad \int_X f_T \, d\mu = 1, \qquad S(f_T) \ll T^{c\beta}$$

with some fixed $c > 0$, determined by the Sobolev norm. We observe that for $x \in \Omega_T$, $A_T(f_T)(x) = 0$, so that

$$\int_{\Omega_T} \left| A_T(f_T) - \int_X f_T \, d\mu \right|^2 d\mu = |\Omega_T|.$$

On the other hand, by Lemma 6,

$$\int_{\Omega_T} \left| A_T(f_T) - \int_X f_T \, d\mu \right|^2 d\mu \le \left\| A_T(f_T) - \int_X f_T \, d\mu \right\|_2^2 \ll (T^{-1/2} S(f_T))^2 \ll T^{2c\beta - 1}.$$

We pick $\beta \in (0,1)$ sufficiently small to make the last exponent negative. Then the above estimates imply that

$$|\Omega_T| \ll T^{-\epsilon}$$

with some $\epsilon > 0$. Hence, it follows from the Borel-Cantelli lemma that the limsup of the sets $\Omega_{2^k}$ has measure zero. This means that for almost every $x \in X$, we have $x \notin \Omega_{2^k}$ for all $k \ge k_0(x)$, i.e., for all sufficiently large $k$, there exists $t \in [0, 2^k]$ such that

$$a_t x \in \mathcal{U}_{2^{-k\beta}/2^\beta}(x_0) \backslash \mathcal{U}^*_{2^{-k\beta}/2}(x_0).$$

Finally, given general $T \ge 1$, we choose $k$ so that $2^k \le T < 2^{k+1}$. Then $[0, 2^k] \subset [0, T]$. Hence, for all sufficiently large $T$, there exists $t \in [0, T]$ such that

$$a_t x \in \mathcal{U}_{2^{-k\beta}/2^\beta}(x_0) \backslash \mathcal{U}^*_{2^{-k\beta}/2}(x_0) \subset \mathcal{U}_{T^{-\beta}}(x_0) \backslash \mathcal{U}^*_{T^{-\beta}/2}(x_0).$$

This completes the proof. □

Proposition 5 is sufficient for the proof of Theorem 2, but for the proof of Theorem 1 we need a more refined recurrence property. We consider the one-parameter subgroup

$$a_t := \mathrm{diag}(e^{-(d-1)t}, e^t, \dots, e^t),$$

and denote by $U$ the expanding horospherical subgroup of $G_0$ for $a_t$ defined by

$$U := \{g \in G_0 : a_t^{-1} g a_t \to e \text{ as } t \to \infty\}. \tag{10}$$

We note that $U \simeq \mathbb{R}^{d-1}$ and the group $U$ is generated by the root subgroups $U_{21}, \dots, U_{d1}$. We prove a recurrence result for orbits starting from points in $Ux \subset X$.

**Proposition 7.** *Let $x_0, x \in X$. Then there exists a constant $\beta > 0$, such that for almost every $u \in U$ and every $T > T_0(u)$,*

$$a_t u x \in \mathcal{U}_{T^{-\beta}}(x_0) \backslash \mathcal{U}^*_{T^{-\beta}/2}(x_0) \quad \text{for some } t \in [0, T].$$

*Proof.* We note that it will be sufficient to prove Proposition 7 for almost all $u$ contained an open neighbourhood $U_0$ of identity in $U$. Our first goal is to prove an analogue of Lemma 6 for averages along $U_0 x$.

We introduce a complementary to $U$ subgroup

$$W := \{g \in G_0 : a_t g a_t^{-1} \text{ is bounded as } t \to \infty\}.$$

The product map $W \times U \to G_0$ is a diffeomorphism in a neighbourhood of identity. We fix a right-invariant Riemannian metric on $G_0$ which also defines a metric on $X = G_0/\Gamma_0$. Let $W_\sigma$ denote the open $\sigma$-neighbourhood of identity in $W$. We assume that $\sigma$ and $U_0$ are sufficiently small, so that the product map $W_\sigma \times U_0 \to G_0$ is a diffeomorphism onto its image, and the projection map $g \mapsto gx$, $g \in W_\sigma U_0$, is one-to-one. Let $X_\sigma := W_\sigma U_0 x \subset X$. We note that the invariant measure on $W_\sigma U_0 \subset G_0$ is the image under the product map of a left invariant measure on $W_\sigma$ and a right invariant measure on $U_0$. After suitable normalisation, this measure projects to the measure $\mu$ on $X_\sigma$. This implies that for every $f \in C_c^\infty(X)$,

$$\begin{aligned}(11)\qquad \left\| A_T(f)(wux) - \int_X f\,d\mu \right\|_{L^2(W_\sigma \times U_0)} &= \left\| A_T(f) - \int_X f\,d\mu \right\|_{L^2(X_\sigma)} \\ &\le \left\| A_T(f) - \int_X f\,d\mu \right\|_{L^2(X)} \ll T^{-1/2} S(f),\end{aligned}$$

where in the last estimate we used Lemma 6.

We observe that for every $wux \in W_\sigma U_0 x$ and every $t > 0$,

$$d(a_t wux, a_t ux) \le d(a_t w a_t^{-1}, e) \ll \sigma.$$

Hence, it follows from the Sobolev embedding theorem that for a suitable Sobolev norm $S$,

$$|f(a_t wux) - f(a_t ux)| \ll \sigma S(f), \quad f \in C_c^\infty(X).$$

This also implies that $|A_T(f)(wux) - A_T(f)(ux)| \ll \sigma S(f)$, and

$$(12)\qquad \|A_T(f)(wux) - A_T(f)(ux)\|_{L^2(W_\sigma \times U_0)} \ll \sigma |W_\sigma|^{1/2} S(f).$$

Combining (11) and (12), we deduce that

$$\left\| A_T(f)(ux) - \int_X f\,d\mu \right\|_{L^2(W_\sigma \times U_0)} \ll (T^{-1/2} + \sigma |W_\sigma|^{1/2}) S(f).$$

Hence,

$$\begin{aligned}\left\| A_T(f)(ux) - \int_X f\,d\mu \right\|_{L^2(U_0)} &= |W_\sigma|^{-1/2} \left\| A_T(f)(ux) - \int_X f\,d\mu \right\|_{L^2(W_\sigma \times U_0)} \\ &\ll (|W_\sigma|^{-1/2} T^{-1/2} + \sigma) S(f) \\ &\ll (\sigma^{-\dim(W)/2} T^{-1/2} + \sigma) S(f),\end{aligned}$$

and taking $\sigma = T^{-\epsilon}$ for sufficiently small $\epsilon > 0$, we conclude that for all $T \ge 1$ and $f \in C_c^\infty(X)$,

$$\left\| A_T(f)(ux) - \int_X f\,d\mu \right\|_{L^2(U_0)} \ll T^{-\alpha} S(f)$$

with some $\alpha > 0$.

Finally, we note that the last estimate is a complete analogue of Lemma 6 for averages along $U_0 x$. Now we can apply exactly the same argument as in the proof of Proposition 5 to conclude that almost every $u \in U_0$ satisfies the claim of the proposition. □

## 4. Compact $A$-orbits

Our argument is based on studying the distribution of $A$-orbits of points in a neighbourhood of a point with compact $A$-orbit. This idea goes back to the papers of Furstenberg [7] and Berend [3], and in the context of Cartan actions it was developed by Lindenstrauss, Weiss [13] and Shapira [16]. It would be sufficient for our purposes to know that there exists $x_0 \in X$ with a compact $A$-orbit. In fact, it is known that every order in a totally real number field gives rise to a compact $A$-orbit (see, for instance, [13, Sec.6] for details).

From now on we fix $x_0 \in X$ such that $Ax_0$ is compact. Let $B := \mathrm{Stab}_A(x_0)$. It is a discrete cocompact subgroup of $A$. The group $B$ acts on the fiber $\pi^{-1}(x_0)$ which can be naturally identified with the torus $\mathbb{R}^d/\Delta_{x_0}$, where $\Delta_{x_0}$ denotes the lattice corresponding to $x_0$. Every $y \in \pi^{-1}(x_0)$ corresponds to a grid $\Lambda_y = \Delta_{x_0} + v$ with $v \in V$. We say that $y \in \pi^{-1}(x_0)$ is *$q$-rational* if $qv \in \Delta_{x_0}$. We note that $B$ preserves the set of $q$-rational elements which has cardinality $q^d$. Hence, if $y \in \pi^{-1}(x_0)$ is $q$-rational, then the subgroup $B_1 := \mathrm{Stab}_A(y)$ has finite index in $B$, namely $|B : B_1| \le q^d$. In particular, this implies the following approximation property.

**Lemma 8.** *There exists $c > 0$ such that for every $a \in A$, one can choose $b \in B_1$ satisfying*

$$\|ab^{-1}\| \le \exp(c\, q^d).$$

*Proof.* We fix a generating set $b_1, \dots, b_{d-1}$ of $B$. Then $b_i^{|B:B_1|} \in B_1$. Hence, it follows that every $a \in A$ can be written as $a = b_0 \left(\prod_{i=1}^{d-1} b_i^{\ell_i}\right) b$ where $b_0$ is in a fixed compact fundamental domain of $B$ in $A$, $0 \le \ell_i < |B : B_1|$, and $b \in B_1$. This implies the lemma. □

Our argument involves study of dynamics of the action for the groups $B$ and $B_1$ in a neighbourhood of the fiber $\pi^{-1}(x_0)$. The crucial part will be played by two quantitative density results that we now state. The first result (Theorem 9), which was proved by Z. Wang [18], establishes quantitative density in the fibers, and the second result (Proposition 10), which is deduced from Baker's Theorem, will be used to prove density along orbits of root subgroups.

We say that $y \in \pi^{-1}(x_0)$ is Diophantine of exponent $k$ if $\Lambda_y = \Delta_{x_0} + v$ and for some $c > 0$,

$$|qv - z| \ge c\, q^{-k+1} \quad \text{for every } q \ge 2 \text{ and } z \in \Delta_{x_0}. \tag{13}$$

The following theorem allows us to establish quantitative density in fibers $\pi^{-1}(x_0)$ of the space $Y$ under a Diophantine condition.

**Theorem 9** (Z. Wang [18])**.** *There exist $Q_0, \sigma > 0$ and $c = c(x_0) > 0$ such that for every $y \in \pi^{-1}(x_0)$ and $k > 1$ satisfying* (13) *and $Q \ge Q_0$, the set $B(Q^{k+2})y$ is $(\log_{(3)} Q)^{-\sigma}$-dense in the torus $\pi^{-1}(x_0)$.*

We note that this result is stated in [18] (see [18, Theorem 1.10]) for the standard torus $\mathbb{R}^d/\mathbb{Z}^d$ and balls defined by the Mahler measure, but it is straightforward to extend it to our setting.

On the other hand, if the point $y$ in the fiber is close to a $q$-rational point, we will analyse action of the group $B_1$ on orbits of the root subgroups and use the following proposition.

**Proposition 10.** *Assuming $d \ge 3$, there exists $\eta > 1$ such that given $\alpha \in \Phi(G)$ and a subgroup $B_1$ of $B$ of index $q$, for every $M \ge 1$ and $t > 0$, there exists $a \in B_1$ such that*

$$|\alpha(a) - t| \ll qtM^{-1} \quad \textit{and} \quad \log\|a\| \ll |\log t| M^{\eta+1}.$$

We note that $\eta$ is precisely the exponent appearing in Baker's estimate (16).

In the proof of Proposition 10 we use the following lemma.

**Lemma 11.** *Let $S$ be a multiplicative subgroup of $\mathbb{R}^+$ generated by multiplicatively independent algebraic numbers $\lambda_1$ and $\lambda_2$. Then there exists $\eta > 1$ such that for every $M \ge 1$ and $t > 0$, there exists $s = \lambda_1^{\ell_1}\lambda_2^{\ell_2} \in S$ satisfying*

$$|s - t| \ll tM^{-1} \quad \textit{and} \quad |\ell_1|, |\ell_2| \ll |\log t| M^{\eta+1}.$$

*Moreover, if $S_1$ is a subgroup of $S$ of index $q \le M$, then there exists $s = \lambda_1^{\ell_1}\lambda_2^{\ell_2} \in S_1$ satisfying*

$$|s - t| \ll qtM^{-1} \quad \textit{and} \quad |\ell_1|, |\ell_2| \ll |\log t| M^{\eta+1}.$$

*Proof.* We set $a_1 = \log\lambda_1$ and $a_2 = \log\lambda_2$. By Minkowski's theorem, for every $M \ge 1$, there exists $(n_1, n_2) \in \mathbb{Z}^2 \backslash \{(0,0)\}$ such that

$$|n_1a_1 + n_2a_2| \le M^{-1} \quad \text{and} \quad |n_1|, |n_2| \ll M. \tag{14}$$

We set $a := n_1a_1 + n_2a_2$. We note that $a \ne 0$ because $\lambda_1$ and $\lambda_2$ are assumed to be multiplicatively independent. It is clear that we can arrange $a > 0$. Let $b := \lceil \frac{M^{-1}}{a} \rceil a$. Then

$$M^{-1} \le b < \left(\frac{M^{-1}}{a} + 1\right) a \le 2M^{-1}. \tag{15}$$

It follows from Baker's Theorem (see, for instance, [1, Ch. 3]) that there exists $\eta > 1$ such that for all $(m_1, m_2) \in \mathbb{Z}^2 \backslash \{(0,0)\}$,

$$|m_1 a_1 + m_2 a_2| \geq \max(|m_1|, |m_2|)^{-\eta}. \tag{16}$$

Hence, we deduce from (14) and (16) that $\lceil \frac{M^{-1}}{a} \rceil \ll M^{\eta-1}$, so that $b = \ell_1 a_1 + \ell_2 a_2$ with $|\ell_1|, |\ell_2| \ll M^\eta$. It follows from (15) that the set $\{ib : |i| \leq LM\}$ forms a $2M^{-1}$-net of the interval $[-L, L]$. Hence, for every $t > 0$, there exists $d = i\ell_1 a_1 + i\ell_2 a_2$ such that

$$|d - \log t| \ll M^{-1} \quad \text{and} \quad |i\ell_1|, |i\ell_2| \ll |\log t| M^{\eta+1}.$$

This implies that $|e^d - t| \ll tM^{-1}$, as required.

To prove the second part of the lemma, we apply the above argument to the elements $\lambda_1^q$ and $\lambda_2^q$ that belong to the subgroup $S_1$. It follows from (14) that there exists $(n_1, n_2) \in \mathbb{Z}^2 \backslash \{(0,0)\}$ such that

$$|n_1(qa_1) + n_2(qa_2)| \leq qM^{-1} \quad \text{and} \quad |n_1|, |n_2| \ll M.$$

We set $a := n_1(qa_1) + n_2(qa_2)$ and $b := \lceil \frac{qM^{-1}}{a} \rceil a$. Then it follows from (16) that $a \geq qM^{-\eta}$. We proceed exactly as in the previous paragraph to prove the second part of the lemma. $\square$

*Proof of Proposition 10.* We write $x_0 = g_0 \Gamma_0$ for some $g_0 \in G_0$. Then $B = A \cap g_0 \Gamma_0 g_0^{-1}$. It follows that entries of elements in $B$ are eigenvalues of matrices from $\Gamma_0 = \mathrm{SL}(d, \mathbb{Z})$. Hence, entries of elements in $B$ are algebraic numbers. In particular, the group $\alpha(B)$ consists of algebraic numbers. We apply Lemma 11 to a subgroup of this group. It was proven in [16, Cor. 3.3] that $\alpha(B)$ is dense in $\mathbb{R}^+$ for every $\alpha \in \Phi(G)$. Since $B$ is finitely generated, this implies that $\alpha(B)$ must contain two multiplicatively independent elements. Now Proposition 10 follows directly from Lemma 11. $\square$

## 5. Proof of the main theorems

The proof of the main theorems will use the dynamical approach to the 'multiplicatively approximable' property provided by Proposition 3. More explicitly, we will establish that for points $y$ in the space of grids $Y$, their orbits $A(R)y$ visit the shrinking sets $\mathcal{W}(\vartheta, (\log_{(5)} R)^{-\zeta})$, provided that $R$ is sufficiently large. As the first step, we use the results from Section 3 to deduce that the projected orbits $A(R)x$, with $x = \pi(y)$, in the space of lattices $X$ visit shrinking neighbourhoods of any given point $x_0$ in $X$. We apply this observation when the point $x_0$ has compact $A$-orbit. This will allow us to analyse behaviour of $A$-orbits locally in a neighbourhood of the fiber $\pi^{-1}(x_0)$. The crucial step of the proof is the following proposition:

**Proposition 12.** *Assume that $d \geq 3$. Let $x_0 \in X$ such that $Ax_0$ is compact and $x \in X$. We assume that for fixed $\nu, \beta > 0$ and all sufficiently large $T$,*

$$\exists a_0 \in A : \quad \|a_0\| \leq e^{\nu T} \quad \text{and} \quad a_0 x \in \mathcal{U}_{T^{-\beta}}(x_0) \backslash \mathcal{U}^*_{T^{-\beta}/2}(x_0). \tag{17}$$

*Then there exist $\vartheta, \delta > 0$, independent of $x$, such that for any $y \in \pi^{-1}(x)$ and all sufficiently large $R$,*

$$A(R)y \cap \mathcal{W}(\vartheta, (\log_{(5)} R)^{-\delta}) \neq \emptyset. \tag{18}$$

We begin by investigating how the recurrence property in Proposition 12 changes under small perturbations of the base point $y$. It will be convenient to consider the family of neighbourhoods of $y$ in $Y$ defined by $\mathcal{O}_\varepsilon(y) := \mathcal{O}_G(\varepsilon)y$, where $\mathcal{O}_G(\varepsilon)$ is defined in (7).

**Lemma 13.** *Let $\vartheta \geq 1$, $0 < \varepsilon < \varepsilon_0(\vartheta)$, $\varepsilon^{1/4} \leq \varepsilon_1 < \varepsilon_2$, and $a \in A(\varepsilon^{-1/(2d)})$. Then for every $y \in Y$, if*

$$ay \in \mathcal{W}(\vartheta, \varepsilon_1, \varepsilon_2), \tag{19}$$

*then for all $y' \in \mathcal{O}_\varepsilon(y)$,*

$$ay' \in \mathcal{W}(3\vartheta, 2^{-1}\varepsilon_1, 2\varepsilon_2).$$

*Proof.* Since $y' \in \mathcal{O}_\varepsilon(y)$, we can write $y' = hy$ with $h \in \mathcal{O}_G(\varepsilon)$. The element $h$ can be written as $h = (v, g)$ with $v \in V$ satisfying $\|v\| < \varepsilon$ and $g \in G_0$ satisfying $\|g - e\| < \varepsilon$. Then $ay' = (aha^{-1})ay$ where $aha^{-1} = (av, aga^{-1})$. We observe that for $x \in \mathrm{Mat}(d, \mathbb{R})$, we have

$$\|axa^{-1}\| \le \|a\| \cdot \|a^{-1}\| \cdot \|x\| \le \|a\|^d \cdot \|x\|,$$

so that since $a \in A(\varepsilon^{-1/(2d)})$, we deduce that

$$\|aga^{-1} - e\| \le \varepsilon^{-1/2}\|g - e\| < \varepsilon^{1/2}.$$

Also $\|av\| \le \|a\|\,\|v\| < \varepsilon^{1/2}$.

According to our assumption (19), there exists $z \in \Lambda_{ay}$ such that $\|z\| < \vartheta$ and $\varepsilon_1 < N(z) < \varepsilon_2$. Then the vector $w := (aha^{-1})z$ belongs to $\Lambda_{ay'}$, and

$$w = (aga^{-1})z + av = z + ((aga^{-1}) - e)z + av.$$

This implies that $\|w\| < 3\vartheta$, and $w = z + O_\vartheta(\varepsilon^{1/2})$, so that $N(w) = N(z) + O_\vartheta(\varepsilon^{1/2})$. Hence, $ay' \in \mathcal{W}(3\vartheta, 2^{-1}\varepsilon_1, 2\varepsilon_2)$ for $\varepsilon \in (0, \varepsilon_0(\vartheta))$. □

*Proof of Proposition 12.* We write $a_0 x = g_0 x_0$ with $g_0 \in \mathcal{U}_{G_0}(T^{-\beta}) \backslash \mathcal{U}^*_{G_0}(T^{-\beta}/2)$. The element $g_0$ has a decomposition

$$g_0 = c \prod_{\alpha \in \Phi(G_0)} u_\alpha(t_\alpha),$$

where $c \in A$ with $\|c - e\| < T^{-\beta}$, $|t_\alpha| < T^{-\beta}$ for all $\alpha$, and $|t_{\alpha_0}| \ge T^{-\beta}/2$ for some $\alpha_0 \in \Phi(G_0)$. In particular, $g_0 = e + O(T^{-\beta})$. For every $y \in \pi^{-1}(x)$,

$$a_0 y = g_0 y_0 \tag{20}$$

with some $y_0 \in \pi^{-1}(x_0)$. The point $y_0$ corresponds to the grid $\Delta_{x_0} + g_0^{-1} w$ with some $w \in V$. Since $g_0$ is bounded, $w$ can be chosen to lie in a fixed bounded subset of $V$, depending only on $\Delta_{x_0}$. Although we don't have any control over $w$, we may assume (after changing (20)) that either

$$\|w\| \ge T^{-\beta} \quad \text{or} \quad w = 0. \tag{21}$$

More precisely, we modify (20) as follows. If $\|w\| < T^{-\beta}$, then $w = \prod_{\alpha \in \Phi(V)} u_\alpha(t_\alpha)$ with $|t_\alpha| \ll T^{-\beta}$. The element $g := (w, g_0) \in G$ can be written as

$$g = \left( \prod_{\alpha \in \Phi(V)} u_\alpha(t_\alpha) \right) c \left( \prod_{\alpha \in \Phi(G_0)} u_\alpha(t_\alpha) \right), \tag{22}$$

and we obtain from (20) that

$$a_0 y = g y_0', \tag{23}$$

where the point $y_0' := (-g_0^{-1} w, e) y_0$ corresponds to the grid $\Delta_{x_0}$. On the other hand, if $\|w\| \ge T^{-\beta}$, we set $g := g_0$ and $y_0' := y_0$. We conclude that in both cases we obtain (23) with $g$ as in (22), where $y_0'$ corresponds to a grid $\Delta_{x_0} + g^{-1} w$ with $w$ satisfying (21). We note that $g = e + O(T^{-\beta})$, and $g$ has its $G_0$-component equal to $g_0$.

Since behaviour of the orbit $Ay$ depends crucially on the Diophantine properties of the vector $w$ with respect to the lattice $\Delta_{x_0}$, the proof naturally splits into the following three subcases:

1. *$w$ is Diophantine:* for every $q \ge 2$ and $w_0 \in \Delta_{x_0}$, $|qw - w_0| \ge c(x_0)\, q^{-k+1}$, where $c(x_0)$ is as in Theorem 9.
2. *$w$ is close to a torsion point with small period:* there exist $q \ge 2$ with $q \le L$ and $w_0 \in \Delta_{x_0}$ such that $|qw - w_0| < c(x_0)\, q^{-k+1}$.
3. *$w$ is close to a torsion point with large period:* for every $q \ge 2$ with $q \le L$ and $w_0 \in \Delta_{x_0}$, $|qw - w_0| \ge c(x_0)\, q^{-k+1}$, but there exist $q \in \mathbb{N}$ and $w_0 \in \Delta_{x_0}$ such that $|qw - w_0| < c(x_0)\, q^{-k+1}$.

The parameters $k$ and $L$ appearing above will be chosen of the form $k = k(T) \to \infty$ and $L = L(T) \to \infty$ as $T \to \infty$, and they will be specified in the course of the proof.

Now we investigate each of these cases separately:

Case 1: *$w$ is Diophantine.* It follows from Theorem 9 that for sufficiently large $Q$, the set $B(Q^{k+2})w$ is $(\log_{(3)} Q)^{-\sigma}$-dense in $V/\Delta_{x_0}$. We observe that $g^{-1}w = w + O(T^{-\beta})$. Let us assume that $Q^{k+2} \le T^{\beta/2}$ (in fact, later in the proof we will have to impose a much stronger restriction). Then the set $B(Q^{k+2})g^{-1}w$ is $2(\log_{(3)} Q)^{-\sigma}$-dense in $V/\Delta_{x_0}$, when $Q$ is sufficiently large. We observe that for $a \in B = \mathrm{Stab}_A(x_0)$,

$$a\Lambda_{y_0} = \Lambda_{ay_0} = a(\Delta_{x_0} + g^{-1}w) = \Delta_{x_0} + ag^{-1}w.$$

Hence, the set $B(Q^{k+2})\Lambda_{y_0}$ is also $2(\log_{(3)} Q)^{-\sigma}$-dense in $V$. In particular, this implies that there exist $a \in B(Q^{k+2})$ and $z \in \Lambda_{ay_0}$ such that

$$(\log_{(3)} Q)^{-\sigma} < z_i < 5(\log_{(3)} Q)^{-\sigma} \quad \text{for every } i.$$

These inequalities imply that

$$\|z\| < 1 \qquad \text{and} \qquad (\log_{(3)} Q)^{-d\sigma} < N(z) < 5^d(\log_{(3)} Q)^{-d\sigma},$$

so that

$$ay_0 \in \mathcal{W}(1, (\log_{(3)} Q)^{-d\sigma}, 5^d(\log_{(3)} Q)^{-d\sigma}) \tag{24}$$

assuming $Q$ is sufficiently large.

Next we claim that an analogous inclusion holds for the point $agy_0$ as well. Since $gy_0 \in \mathcal{O}_{c_0 T^{-\beta}}(y_0)$ for some fixed $c_0 > 0$ (see (8)) and $a \in B(Q^{k+2})$, we will be able to deduce this from Lemma 13 provided that $Q$ is not too large. Taking this into account, we choose $Q$ so that $Q^{k+2} = (c_0 T^{\beta})^{1/(2d)}$. We note that if $k = o(\log T)$ as $T \to \infty$, then $Q \to \infty$ as $T \to \infty$ and (24) holds assuming $T$ is sufficiently large. Then Lemma 13 implies that if $T$ is sufficiently large,

$$aa_0y = agy_0 \in \mathcal{W}(3, 2^{-1}(\log_{(3)} Q)^{-d\sigma}, 2 \cdot 5^d(\log_{(3)} Q)^{-d\sigma}).$$

Since $Q^{k+2} = (c_0 T^{\beta})^{1/(2d)}$,

$$aa_0 \in B(Q^{k+2})A(e^{\nu T}) \subset A(e^{(\nu+1)T}).$$

Hence, taking $R = e^{(\nu+1)T}$, we deduce that for all sufficiently large $R$,

$$A(R)y \cap \mathcal{W}(3, f(R)^{-d\sigma}) \neq \emptyset, \tag{25}$$

where

$$f(R) := (2 \cdot 5^d)^{-1/(d\sigma)} \log_{(3)} \left( c_0^{\frac{1}{2d(k+2)}} \left( \frac{\log R}{\nu+1} \right)^{\frac{\beta}{2d(k+2)}} \right) \gg \log \left( \log_{(3)} R - \log k \right).$$

In order for (25) to give a non-trivial estimate, we have to choose $k$ such that

$$\log k = \log_{(3)} R - s(R) \quad \text{with } s(R) \to \infty. \tag{26}$$

We note that this choice of $k$, in particular, implies that $k = o(\log T)$, which we have used above. Finally, we deduce from (25) that for all sufficiently large $R$,

$$A(R)y \cap \mathcal{W}(3, O((\log s(R))^{-d\sigma})) \neq \emptyset. \tag{27}$$

Case 2: *$w$ is close to a torsion point with small period.* We start by modifying equation (23). We observe that

$$g = c \left( \prod_{\alpha \in \Phi(V)} u_\alpha(t'_\alpha) \right) \left( \prod_{\alpha \in \Phi(G_0)} u_\alpha(t_\alpha) \right),$$

where $t'_\alpha = \alpha(c)^{-1}t_\alpha = O(T^{-\beta})$. Replacing $a_0$ by $a_0c^{-1}$, we may assume without loss of generality that (23) holds with $c = e$. We denote by $y''_0$ the element of $Y$ that corresponds to the grid $\Delta_{x_0} + q^{-1}w_0$. Then

$$a_0y = gy'_0 = hy''_0 \quad \text{where } h := (w - q^{-1}gw_0, e)g \in G. \tag{28}$$

We observe that

$$\|w - q^{-1}gw_0\| \le \|w - q^{-1}w_0\| + \|q^{-1}w_0 - q^{-1}gw_0\| \ll \max(T^{-\beta}, q^{-k}).$$

We decompose $h$ into a product with respect to root subgroups:

$$h = \prod_{\alpha\in\Phi(G)} u_\alpha(t_\alpha), \tag{29}$$

where $|t_\alpha| \ll \max(T^{-\beta}, q^{-k})$ for all roots $\alpha$. We recall that $k$ is chosen so that $k = o(\log T)$ (see Case 1), Hence, it follows that $|t_\alpha| \ll 2^{-k}$ for all roots $\alpha$, when $T$ is sufficiently large. Let $B_1 := \mathrm{Stab}_A(y''_0)$. Since $B_1$ is precisely the stabiliser in $B$ of the $q$-rational point $q^{-1}w_0$ in $V/\Delta_{x_0}$, it follows that $|B : B_1| \le q^d$. We observe that for $b \in B_1$, we have

$$bhy''_0 = (bhb^{-1})y''_0 = \left(\prod_{\alpha\in\Phi(G)} u_\alpha(\alpha(b)t_\alpha)\right) y''_0. \tag{30}$$

Our argument is based on picking suitable elements $b \in B_1$ which contract some of the coordinates $t_\alpha$. This will allow to reduce the complexity of the product in (30). A useful tool for achieving this is the following elementary lemma.

**Lemma 14.** *Let $v_1, \dots, v_s$ be a collection of vectors in a vector space $V$ such that for all $i \ne j$, $v_i$ and $v_j$ are linearly independent. Assume that there exists $L \in V^*$ such that $L(v_i) > 0$ for all $i$. Then there exists $v_j$ such that*

- *for some $S_1 \in V^*$, $S_1(v_j) > 0$ and $S_1(v_i) < 0$ with $i \ne j$,*
- *for some $S_2 \in V^*$, $S_2(v_j) = 0$ and $S_2(v_i) < 0$ with $i \ne j$.*

*Proof.* Let $\bar v_i$ denote the positive multiple of $v_i$ such that $L(\bar v_i) = 1$. We denote by $\mathcal{C}$ the closed convex hull of the points $\bar v_i$. Let $\bar v_j$ be an extreme point of $\mathcal{C}$. Then there exists a hyperplane $\mathcal{H}$ in $L = 1$ which separates $\bar v_j$ and $\bar v_i$, $i \ne j$. It is sufficient to pick $S_1 \in V^*$ such that $\{S_1 = 0\} \cap \{L = 1\} = \mathcal{H}$ with a suitable sign. The proof of the second part is similar. □

We note that conjugating by elements from $B_1$, one can only achieve precision $e^{O(q^d)}$ (cf. Lemma 8), but the coordinates $t_\alpha$ are of order $O(2^{-k})$, so that our argument, which is presented below, could only work provided that $k \ge O(q^d)$. Since $q \le L$, it is sufficient to assume that the parameter $L = L(T)$ satisfies

$$L^d = o(k) \quad \text{as } T \to \infty. \tag{31}$$

In view of (30), we have to analyse the behaviour of $\alpha(b)t_\alpha$, $\alpha \in \Phi(G)$, with $b \in B_1$. Now we construct an explicit $b \in B_1$ which contracts some of the factors in (30). Since $B$ is a lattice in $A$, there exists an element $a \in B$ such that $\alpha(a) \ne 1$ for all $\alpha \in \Phi(G)$. In particular, we have a decomposition

$$\Phi(G) = \Phi^+ \sqcup \Phi^- \quad \text{where } \Phi^+ := \{\alpha : \alpha(a) > 1\} \text{ and } \Phi^- := \{\alpha : \alpha(a) < 1\}.$$

This decomposition is non-trivial because $\prod_\alpha \alpha(a) = 1$. We recall that there exists $\alpha_0 \in \Phi(G_0)$ such that $|t_{\alpha_0}| \ge T^{-\beta}/2$. Replacing $a$ by $a^{-1}$ if necessary we may assume that $\alpha_0 \in \Phi^+$. Let us pick the maximal exponent $i$ such that $\alpha(a)^{i|B:B_1|}|t_\alpha| \le 1$ for all $\alpha \in \Phi(G)$ and set $b := a^{i|B:B_1|}$. Clearly, $b \in B_1$. Also since $\alpha_0(a)^{i|B:B_1|}T^{-\beta}/2 \le 1$, it follows that $i|B : B_1| \ll \log T$, so that

$$\|b\| \le T^{O(1)}.$$

It follows from our choice of the exponent $i$ that there exists $\alpha_1 \in \Phi(G)$ such that

$$\alpha_1(a)^{(i+1)|B:B_1|}|t_{\alpha_1}| = \alpha_1(b)\alpha_1(a)^{|B:B_1|}|t_{\alpha_1}| > 1.$$

Hence,

$$\alpha_1(b)|t_{\alpha_1}| \geq e^{-O(q^d)}. \tag{32}$$

On the other hand, for all $\alpha \in \Phi^-$, we have $\alpha(b)|t_\alpha| < |t_\alpha| = O(2^{-k})$. We conclude that

$$h_1 := bhb^{-1} = \prod_{\alpha\in\Phi(G)} u_\alpha(s_\alpha), \tag{33}$$

where $|s_\alpha| \leq 1$ for all $\alpha$, $|s_{\alpha_1}| \geq e^{-O(q^d)}$, and $|s_\alpha| \leq \omega\, 2^{-k}$ for fixed $\omega > 0$ and all $\alpha \in \Phi^-$.

Let us introduce a parameter $\zeta \in (0, 1)$ which will be specified later (see (47) below). Since $G$ has only $d^2$ roots, there exists $1 \leq \ell \leq d^2 + 1$ such that no coordinates $|s_\alpha|$ are contained in the interval $((\omega\, 2^{-k})^{\zeta^{\ell-1}}, (\omega\, 2^{-k})^{\zeta^\ell}]$. We decompose the set of roots as $\Phi(G) = \Phi_1 \sqcup \Phi_2$ where $\Phi_1$ consists of $\alpha$ such that $|s_\alpha| \geq (\omega\, 2^{-k})^{\zeta^\ell}$, and $\Phi_2$ consists of $\alpha$ such that $|s_\alpha| \leq (\omega\, 2^{-k})^{\zeta^{\ell-1}}$. We note that $\Phi^- \subset \Phi_2$. Also, it follows from (32) and (31) that $\alpha_1 \in \Phi_1$. In particular, $\Phi_1$ is not empty. We observe that for $\alpha \in \Phi_2$ and bounded $g \in G$, we have $gu_\alpha(s_\alpha) = vg$ where $v = e + O(2^{-\zeta^{\ell-1}k})$. Therefore, we can rearrange the terms in the product (29), so that

$$h_1 = v_1 h_2 \qquad \text{where} \quad v_1 = e + O(2^{-\zeta^{\ell-1}k}) \;\text{ and }\; h_2 := \prod_{\alpha\in\Phi_1} u_\alpha(s_\alpha). \tag{34}$$

Now we apply Lemma 14 to the set $\Phi_1$ considered as a subset of the dual $A^*$ of $A$. The condition of the lemma holds because $\Phi_1 \subset \Phi^+$. Hence, we deduce that there exist $a_1, a_2 \in A = (A^*)^*$ and $\alpha_2 \in \Phi_1$ such that $\alpha_2(a_1) > 1$ and $\alpha(a_1) < 1$ for all $\alpha \in \Phi_1\backslash\{\alpha_2\}$, and $\alpha_2(a_2) = 1$ and $\alpha(a_2) < 1$ for all $\alpha \in \Phi_1\backslash\{\alpha_2\}$. Rescaling $a_1$, we arrange that $\alpha_2(a_1) = |s_{\alpha_2}|^{-1}$. Since $|s_{\alpha_2}|^{-1} \ll 2^{\zeta^\ell k}$, there exists a constant $c_1 > 0$ such that $\|a_1\| \leq 2^{c_1\zeta^\ell k}$. Moreover, $c_1$ depends only on the initial choice of $a_1$, so that it is uniform. Furthermore, we also rescale $a_2$ so that $\|a_2\| \leq 2^{c_1\zeta^\ell k}$ and $\alpha(a_2) < 2^{-\delta k}$ with some fixed $\delta > 0$ (depending on $\zeta$) for all $\alpha \in \Phi_1\backslash\{\alpha_2\}$. Then

$$\|a_1 a_2\| \leq \|a_1\|\|a_2\| \leq 2^{2c_1\zeta^\ell k}, \tag{35}$$

and

$$\alpha_2(a_1a_2)|s_{\alpha_2}| = 1, \quad \alpha(a_1a_2)|s_\alpha| < 2^{-\delta k} \text{ for all } \alpha \in \Phi_1\backslash\{\alpha_2\}. \tag{36}$$

By Lemma 8, there exists $b_2 \in B_1$ such that

$$\|(a_1a_2)b_2^{-1}\| \leq e^{O(q^d)}. \tag{37}$$

It follows from (35) and (31) that

$$\|b_2\| \leq 2^{2c_1\zeta^\ell k} e^{O(q^d)} = 2^{(2c_1\zeta^\ell + o(1))k}. \tag{38}$$

We have

$$h_3 := b_2 h_2 b_2^{-1} = \prod_{\alpha\in\Phi_1} u_\alpha(r_\alpha),$$

where $r_\alpha = \alpha(b_2)s_\alpha$. By (36), (37) and (31),

$$|r_{\alpha_2}| \geq e^{-O(q^d)}, \quad |r_\alpha| \leq 2^{-\delta k} e^{O(q^d)} = 2^{-(\delta - o(1))k} \text{ for all } \alpha \in \Phi_1\backslash\{\alpha_2\}.$$

Arguing as in (34), we deduce that

$$h_3 = v_2 u_{\alpha_2}(r_{\alpha_2}) \qquad \text{where} \quad v_2 = e + O(2^{-(\delta - o(1))k}). \tag{39}$$

Let us assume that $r_{\alpha_2} > 0$ since the other case can be treated similarly. By Lemma 4, for every $0 < \varepsilon_1 < \varepsilon_2 < 1$, there exists positive $t_+ = O_{\pi(y_0'')}(1) = O_{x_0}(1)$ such that

$$u_{\alpha_2}(t_+)y_0'' \in \mathcal{W}(\vartheta, \varepsilon_1, \varepsilon_2), \tag{40}$$

where $\vartheta = O_{\pi(y_0'')}(1) = O_{x_0}(1)$. By Proposition 10, for every $M \geq q^d$, there exists $b_3 \in B_1$ such that

$$|\alpha_2(b_3) - t_+/r_{\alpha_2}| \ll q^d (t_+/r_{\alpha_2}) M^{-1} \ll q^d r_{\alpha_2}^{-1} M^{-1}, \tag{41}$$

and

$$\|b_3\| \le e^{O(|\log(t_+/r_{\alpha_2})|M^{\eta+1})} = e^{O(q^d M^{\eta+1})}. \tag{42}$$

Since in the next step we will apply Lemma 13 with $a = b_3$, $y = u_{\alpha_2}(r_{\alpha_2})y_0''$ and $y' = v_2 y$, we have to take $b_3$ so that

$$\|b_3\| < \|e - v_2\|^{-1/(2d)}. \tag{43}$$

To arrange this (see (39)), we can take $b_3$ of size

$$\|b_3\| \le \omega_1 2^{(\delta - o(1))k/(2d)} \tag{44}$$

with sufficiently small $\omega_1 > 0$. Moreover, in the next step, we will apply Lemma 13 with $a = b_3 b_2$, $y = h_2 y_0''$ and $y' = v_1 y$. Hence, we also have to arrange that

$$\|b_3 b_2\| < \|e - v_1\|^{-1/(2d)}. \tag{45}$$

For this purpose, we can choose $b_3$ such that

$$\|b_3 b_2\| \le \omega_2 2^{\zeta^{\ell-1}k/(2d)} \tag{46}$$

with sufficiently small $\omega_2 > 0$ (see (34)). We pick the parameter $\zeta$ so that

$$\zeta < 1/(8c_1 d). \tag{47}$$

Then in view of (38), $\|b_2\| \le 2^{\zeta^{\ell-1}k/(4d)}$ for sufficiently large $T$. Hence, if take $b_3$ of size

$$\|b_3\| \le \omega_2 2^{\zeta^{\ell-1}k/(4d)}, \tag{48}$$

then (46) holds. Now let us take $M$ such that

$$q^d M^{\eta+1} \le \delta' k \quad \text{with } \delta' > 0. \tag{49}$$

If $\delta'$ is sufficiently small, then (42) implies that (44) and (48) hold, and consequently (43) and (45) hold. The parameter $M$ in (49) can be chosen to satisfy $M \gg (k/q^d)^{1/(\eta+1)}$. Then it follows from (41) that

$$|\alpha_2(b_3)r_{\alpha_2} - t_+| \ll \frac{q^{d+d/(\eta+1)}}{k^{1/(\eta+1)}} \le \theta, \tag{50}$$

where $\theta := \frac{L^{d(\eta+2)/(\eta+1)}}{k^{1/(\eta+1)}}$. In order for this to give a non-trivial estimate, we have to require that the parameter $L$ is chosen so that

$$L^{d(\eta+2)} = o(k) \quad \text{as } T \to \infty, \tag{51}$$

which is a strengthening of our previous assumption (31). We assume that (51) holds. Then, in particular, $\theta \to 0$ as $T \to \infty$, and $M \ge q^d$ for sufficiently large $T$, which was required for (41) to hold.

Now in (40) we choose $\varepsilon_1 = c_1\theta$ and $\varepsilon_2 = c_2\theta$ with some $0 < c_1 < c_2$. Then since $u_{\alpha_2}(t_+)y_0'' \in \mathcal{W}(\vartheta, \varepsilon_1, \varepsilon_2)$, there exists $z \in \Lambda_{y_0''}$ such that

$$\|u_{\alpha_2}(t_+)z\| < \vartheta \quad \text{and} \quad c_1\theta < |N(u_{\alpha_2}(t_+)z)| < c_2\theta. \tag{52}$$

It follows from (50) that

$$u_{\alpha_2}(\alpha_2(b_3)r_{\alpha_2})z = u_{\alpha_2}(t_+)z + O(\theta\vartheta).$$

Hence, we conclude that the vector $u_{\alpha_2}(\alpha_2(b_3)r_{\alpha_2})z$ also satisfies bounds of the form (52) (provided that the constants $c_1$ and $c_2$ are sufficiently large), Namely, we deduce that

$$\|u_{\alpha_2}(\alpha_2(b_3)r_{\alpha_2})z\| < 2\vartheta$$

if $T$ is sufficiently large, and

$$c_1'\theta < |N(u_{\alpha_2}(\alpha_2(b_3)r_{\alpha_2})z)| < c_2'\theta$$

for some $c_1', c_2' > 0$. Thus it follows that

$$b_3 u_{\alpha_2}(r_{\alpha_2})y_0'' = u_{\alpha_2}(\alpha_2(b_3)r_{\alpha_2})y_0'' \in \mathcal{W}(2\vartheta, c_1'\theta, c_2'\theta).$$

Since (43) holds, we can apply Lemma 13 with $a = b_3$ and $y = u_{\alpha_2}(r_{\alpha_2})y_0''$ to deduce that

$$b_3b_2h_2y_0'' = b_3h_3y_0'' = b_3v_2u_{\alpha_2}(r_{\alpha_2})y_0'' \in \mathcal{W}(6\vartheta, 2^{-1}c_1'\,\theta, 2c_2'\,\theta)$$

for sufficiently large $T$. Since (45) holds, we can apply again Lemma 13 with $a = b_3b_2$ and $y = h_2y_0''$ to conclude that

$$(b_3b_2)bhy_0'' = (b_3b_2)h_1y_0'' = (b_3b_2)v_1h_2y_0'' \in \mathcal{W}(18\vartheta, 4^{-1}c_1'\,\theta, 4c_2'\,\theta)$$

for sufficiently large $T$. Finally, combining this with (28), we deduce that for $a := a_0b_3b_2b$, we have

$$ay = (b_3b_2)bhy_0'' \in \mathcal{W}(18\vartheta, 4^{-1}c_1'\,\theta, 4c_2'\,\theta).$$

We note that

$$\|a\| \le \|a_0\|\,\|b_3b_2\|\,\|b\| \ll e^{\nu T}\, 2^{\zeta^{\ell-1}k/(2d)}\, T^{O(1)} \le e^{(\nu+1)T}$$

for sufficiently large $T$. This proves that

$$A(e^{(\nu+1)T})y \cap \mathcal{W}\left(18\vartheta, 4c_2'\frac{L^{d(\eta+2)/(\eta+1)}}{k^{1/(\eta+1)}}\right) \ne \emptyset \tag{53}$$

for sufficiently large $T$.

<u>Case 3:</u> *$w$ is close to a torsion point with large period.* We consider the set

$$\mathcal{D}_{x_0}(k,L) := \{z \in V/\Delta_{x_0} : \|z - w_0\| < q^{-k} \text{ for some } w_0 \in q^{-1}\Delta_{x_0} \text{ and } q \ge L\}.$$

Let $\mathrm{diam}^*(S)$ denote the supremum of diameters of the connected components of the set $S$. We observe that

$$\mathrm{diam}^*(\mathcal{D}_{x_0}(k,L)) \le \sum_{q\ge L}\sum_{w_0\in q^{-1}\Delta_0} 2q^{-k} \ll \sum_{q\ge L} q^{d-k} \ll L^{d+1-k}. \tag{54}$$

We recall that by (21) either $w = 0$ or $\|w\| \ge T^{-\beta}$. Hence, according to our assumption in Case 3, we must have $\|w\| \ge T^{-\beta}$. Without loss of generality, let us assume that $|w_1| \ge T^{-\beta}$. We consider the one-parameter subgroup $a(t) := \mathrm{diag}(e^t, e^{-t}, 1, \ldots, 1)$ of $A$. We observe that

$$\|a(t)w - a(0)w\| \ge (e^t - 1)|w_1| \ge (e^t - 1)T^{-\beta}.$$

Hence,

$$\mathrm{diam}(a([0, \log(1 + T^{-\beta})])w) \ge T^{-2\beta}. \tag{55}$$

We choose the parameter $L$ so that

$$L^{d+1-k} < \omega\, T^{-2\beta} \tag{56}$$

with sufficiently small $\omega > 0$. Then comparing (54) and (55), we deduce that there exists $a(t)$ with $\|a(t) - e\| \ll T^{-\beta}$ such that $a(t)w \notin \mathcal{D}_{x_0}(k,L)$. We replace (23) by

$$a(t)a_0y = a(t)gy_0'$$

where the point $y_0'$ corresponds to the grid $\Delta_{x_0} + (a(t)g)^{-1}a(t)w$. Hence, if we replace $a_0$ by $a(t)a_0$ and $g$ by $a(t)g$, we obtain (23) with $w$ satisfying either the condition of Case 1 or the condition of Case 2. Hence, we can reduce the proof to the situations considered in Cases 1 or 2. This reduction is possible provided that (56) holds, so that we can choose

$$L \ll T^{2\beta/(k-d-1)}. \tag{57}$$

Then (53) becomes

$$A(e^{(\nu+1)T})y \cap \mathcal{W}\left(18\vartheta, O\left(\frac{T^{\frac{2\beta d(\eta+2)}{(\eta+1)(k-d-1)}}}{k^{1/(\eta+1)}}\right)\right) \ne \emptyset. \tag{58}$$

Finally, we complete the proof of the theorem by combining the estimates obtained in Cases 1 and 2, and optimising the parameter $k = k(T)$. We recall that $k$ is required to satisfy (26) with $R = e^{(\nu+1)T}$, so that $k = \frac{\log T}{\rho(T)}$ for some $\rho(T) \to \infty$. Moreover, $k$ has to satisfy (51). Since we are assuming that $k \to \infty$, it follows from (57) that (51) holds provided that $T^{\delta''/k} = o(k)$ with some $\delta'' > 2\beta d(\eta + 2)$. Therefore, the parameter $k$ has to be chosen so that

$$\frac{\delta''}{k} \log T - \log k = \delta'' \rho(T) - \log\log T + \log \rho(T) \to -\infty.$$

Hence, we can take $\rho(T) = \kappa \log\log T$ for sufficiently small $\kappa > 0$. Then (26) holds with $s(R) \gg \log_{(4)} R$, where $R = e^{(\nu+1)T}$. Hence, (27) implies that for sufficiently large $R$,

$$A(R)y \cap \mathcal{W}(3, O((\log_{(5)} R)^{-d\sigma})) \neq \emptyset.$$

This proves the proposition in Case 1.

Also with this choice of $\rho(T) = \kappa \log\log T$, provided that $\kappa$ is chosen sufficiently small, we obtain that

$$\begin{aligned} \log\left(\frac{T^{\frac{2\beta d(\eta+2)}{(\eta+1)(k-d-1)}}}{k^{1/(\eta+1)}}\right) &\leq \frac{2\beta d(\eta+2)}{(\eta+1)k}\log T - \frac{1}{\eta+1}\log k \\ &= (\eta+1)^{-1}(2\beta d(\eta+2)\rho(T) - \log\log T + \log\rho(T)) \\ &\leq -\varsigma \log\log T \end{aligned}$$

with $\varsigma < (\eta+1)^{-1}$. Hence, (58) implies that

$$A(e^{(\nu+1)T})y \cap \mathcal{W}(18\vartheta, O((\log T)^{-\varsigma})) \neq \emptyset,$$

which proves the proposition in Case 2 and completes the proof of the theorem. □

*Proof of Theorem 1* . We consider the family of grids

$$\Lambda(u, v, \alpha, \beta) := \{{}^t(x, xu - y - \alpha, xv - z - \beta) : x, y, z \in \mathbb{Z}\}.$$

We note that the lattices $\Lambda(u, v, 0, 0)$ with $(u, v) \in \mathbb{R}^2$ are precisely the lattices in the orbit $U\mathbb{Z}^3$, where $U$ is the expanding horospherical subgroup defined in (10). Hence, it follows from Proposition 7 that for almost every $(u, v) \in \mathbb{R}^2$, the lattice $\Lambda(u, v, 0, 0)$ satisfies the assumption of Proposition 12. Therefore, by this proposition, the grid $\Lambda(u, v, \alpha, \beta)$ with arbitrary $\alpha, \beta \in \mathbb{R}$ has property (WR) with $h(T) = (\log_{(5)} T)^\delta$. Thus, by Proposition 3, the grid $\Lambda(u, v, \alpha, \beta)$ is $h$-multiplicatively approximable. This implies that there exists a sequence $v_n = {}^t(q_n, q_n u - r_n - \alpha, q_n v - s_n - \beta)$ with $q_n, r_n, s_n \in \mathbb{Z}$ such that $v_n \to \infty$, and

$$0 < (\log_{(5)} \|v_n\|)^\delta |q_n||q_n u - r_n - \alpha||q_n v - s_n - \beta| < 1.$$

In particular, it follows that $0 \neq |q_n||q_n u - r_n - \alpha||q_n v - s_n - \beta| \to 0$. Since $r_n, s_n \in \mathbb{Z}$, this can only happen if $|q_n| \to \infty$. We observe that $\|v_n\| \geq |q_n|$, so that

$$(\log_{(5)} |q_n|)^\delta |q_n| \langle q_n u - \alpha\rangle \langle q_n v - \beta\rangle \leq (\log_{(5)} \|v_n\|)^\delta |q_n||q_n u - r_n - \alpha||q_n v - s_n - \beta|.$$

Hence, we deduce that

$$\liminf_{|q|\to\infty} (\log_{(5)} |q|)^\delta |q| \langle qu - \alpha\rangle \langle qv - \beta\rangle \leq 1,$$

and Theorem 1 follows. □

*Proof of Theorem 2.* The proof of Theorem 2 is similar. We note that it is sufficient to prove the theorem for almost every unimodular lattice $\Delta$ because general lattices can be obtained by rescaling. By Proposition 5, almost every $x \in X$ satisfies the assumptions of Proposition 12. Let $\Delta$ denote the lattice corresponding to some $x$ for which Proposition 12 applies. Then combining Proposition 12 and Proposition 3, we deduce that for every $w \in \mathbb{R}^d$, the grid $\Delta + w$ is $h$-multiplicatively approximable with $h(T) = (\log_{(5)} T)^\delta$. This completes the proof of the theorem. □

## References


[1] Baker, A. *Transcendental number theory.* Second edition. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1990.

[2] Badziahin D., Velani S., *Multiplicatively badly approximable numbers and generalised Cantor sets*. Adv. Math. 228 (2011), no. 5, 2766–2796.

[3] Berend B., *Multi-invariant sets on tori.* Trans. Amer Math. Soc. 280 (1983), 509–532.

[4] Cassels J. W. S., *An introduction to diophantine approximation*, Hafner Publishing Co., New York, 1972.

[5] Davenport, H. *Indefinite binary quadratic forms, and Euclid's algorithm in real quadratic fields.* Proc. London Math. Soc. 53 (1951), 65–82.

[6] de Mathan B., Teulié O., Problèmes diophantiens simultanès, Monatsh. Math. 143 (2004), no. 3, 229–245.

[7] Furstenberg, H. *Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation.* Theory of Computing Systems Vol.

[8] Gallagher P., *Metric simultaneous Diophantine aproximations*, J. London Math. Soc. 37 (1962), 387–390.

[9] Gorodnik A., Vishe P., *Mixed inhomogeneous Littlewood conjecture and quantitative improvements*, in preparation.

[10] Harrap S., *Twisted inhomogeneous Diophantine approximation and badly approximable sets*, Acta Arith. 151 (2012), 55–82.

[11] Haynes A., Jensen J. L., Kristensen S., *Metrical musings on Littlewood and friends*, Proc. Amer. Math. Soc. 142 (2014), 457–466.

[12] Kleinbock D., Margulis G., *Logarithm laws for flows on homogeneous spaces*, Inv. Math. 138 (1999), 451–494.

[13] Lindenstrauss E., Weiss B., *On sets invariant under the action of the diagonal group.* Ergodic Theory Dynam. Systems 21 (2001), 1481–1500.

[14] Peck L. G., *Simultaneous rational approximations to algebraic numbers*, Bull. Amer. Math. Soc. 67 (1961), 197–201.

[15] Pollington A., Velani S., *On a problem in simultaneous Diophantine approximation: Littlewood's conjecture*, Acta Math. 185 (2000), 287–306.

[16] Shapira U., *A solution to a problem of Cassels and Diophantine properties of cubic numbers*, Ann. Math. 173 (2011), 1–14.

[17] Sullivan, D., *Disjoint spheres, approximation by imaginary quadratic numbers, and the logarithm law for geodesics.* Acta Math. 149 (1982), 215–237.

[18] Wang, Z., *Quantitative density under higher rank abelian algebraic toral actions*, Int. Math. Res. Not. 2011, no. 16, 3744–3821.