Acta Math., 215 (2015), 127–159
DOI: 10.1007/s11511-015-0130-0

Mixing properties of commuting nilmanifold automorphisms

by

Alexander Gorodnik

University of Bristol
Bristol, U.K.

Ralf Spatzier

University of Michigan
Ann Arbor, MI, U.S.A.

1. Introduction

Given a measure-preserving action of a (discrete) group Γ on a probability space (X,μ), we say that this action is $(s+1)$-*mixing* if for every $f_0,...,f_s\in L^\infty(X)$ and $\gamma_0,...,\gamma_s\in\Gamma$,

$$\int_X\left(\prod_{i=0}^s f_i(\gamma_i x)\right)d\mu(x)\longrightarrow\prod_{i=0}^s\int_X f_i\,d\mu \tag{1.1}$$

as $\gamma_{i_1}\gamma_{i_2}^{-1}\to\infty$ for all $i_1\neq i_2$. In particular, 2-mixing corresponds to the usual notion of mixing. It was discovered by Ledrappier [13] that 2-mixing does not imply 3-mixing for $\mathbb{Z}^2$-actions. In this paper we will be interested in mixing of higher order for group actions. Mixing of all orders is a very widespread phenomenon for 1-parameter actions. In particular, it is known to hold for many transformations satisfying some hyperbolicity assumptions. However, this is a measurable property that might arise for a multitude of other reasons which are not well understood. For instance, the horocyclic flow provides an example of a parabolic dynamical system which is mixing of all orders. A well-known longstanding question of Rokhlin asks whether mixing of order 2 implies mixing of all orders for a general measure-preserving transformation.

Very little is known about higher-order mixing for actions of large groups. We are only aware of two general families of actions of large groups on manifolds where the multiple mixing has been established—$\mathbb{Z}^l$-actions by automorphisms on compact abelian groups and actions of simple Lie groups. K. Schmidt and Ward [22] proved that 2-mixing $\mathbb{Z}^l$-action by automorphisms on compact connected abelian groups are mixing of

A. G. was supported in part by EPSRC grant EP/H000091/1 and ERC grant 239606. R. S. was supported in part by NSF grant DMS-0906085.

all orders, and Mozes [16] established mixing of all orders for ergodic actions of connected semisimple Lie groups with finite centre.

In this paper we investigate mixing properties of $\mathbb{Z}^l$-actions by automorphisms on compact nilmanifolds. We prove that for such actions, 2-mixing implies mixing of all orders and establish quantitative estimates for 2-mixing and 3-mixing.

1.1. Main results

Let G be a simply connected nilpotent group and Λ be a discrete cocompact subgroup. We call the space $X=G/\Lambda$ a *compact nilmanifold*. We denote by $\mathrm{Aut}(X)$ the group of continuous automorphisms α of G such that $\alpha(\Lambda)=\Lambda$. Then $\mathrm{Aut}(X)$ naturally acts on X and preserves the Haar probability measure μ on X.

Our first main result concerns exponential 3-mixing. In order to obtain any quantitative estimate in (1.1), it is necessary to work in a class of sufficiently regular functions. We denote by $C^\theta(X)$ the space of Hölder functions with exponent θ, defined with respect to a Riemannian metric on X.

THEOREM 1.1. *Let $\alpha:\mathbb{Z}^l\to\mathrm{Aut}(X)$ be an action on a compact nilmanifold X such that every $\alpha(z)$, $z\neq 0$, is ergodic. Then there exists $\eta=\eta(\theta)>0$ such that for every $f_0, f_1, f_2\in C^\theta(X)$ and $z_0, z_1, z_2\in\mathbb{Z}^l$,*

$$\begin{aligned}\int_X & f_0(\alpha(z_0)x)f_1(\alpha(z_1)x)f_2(\alpha(z_2)x)\,d\mu(x)\\ &=\left(\int_X f_0\,d\mu\right)\left(\int_X f_1\,d\mu\right)\left(\int_X f_2\,d\mu\right)+O(N(z_0,z_1,z_2)^{-\eta}\|f_0\|_{C^\theta}\|f_1\|_{C^\theta}\|f_2\|_{C^\theta}),\end{aligned}$$

where $N(z_0,z_1,z_2)=\exp(\min_{i\neq j}\|z_i-z_j\|)$.

We note that this result is new even for the case of toral automorphisms. Previously, quantitative 2-mixing was established for toral automorphisms in [14] and for automorphisms of more general compact abelian groups in [15]. Mixing of all orders for ergodic commuting toral automorphisms was established in [22]. The argument in [22] relies on finiteness of the number of non-degenerate solutions of S-unit equations established in [19]. Although there are explicit estimates on the number of such solutions, these estimates are not sufficient to derive any quantitative estimate for 3-mixing because it is also essential to know how the sets of solutions depend on the coefficients. In order to prove Theorem 1.1, we use more delicate Diophantine estimates for linear forms in logarithms of algebraic numbers established in [24] (cf. Proposition 2.2 below).

We also prove mixing of all orders.

THEOREM 1.2. *Let* $\alpha:\mathbb{Z}^l\to\mathrm{Aut}(X)$ *be an action on a compact nilmanifold* X *such that every* $\alpha(z)$, $z\neq 0$, *is ergodic. Then, for every* $f_0,...,f_s\in L^\infty(X)$ *and* $z_0,...,z_s\in\mathbb{Z}^l$,

$$\int_X\left(\prod_{i=0}^s f_i(\alpha(z_i)x)\right)d\mu(x)=\prod_{i=0}^s\int_X f_i\,d\mu+o(1)$$

as $\min_{i\neq j}\|z_i-z_j\|\to\infty$. *Moreover, the convergence is uniform over families of Hölder functions* $f_0,...,f_s$ *such that* $\|f_0\|_{C^\theta},...,\|f_s\|_{C^\theta}\ll 1$.

This theorem extends the main result of [22] to general nilmanifolds. The proof in [22] utilises abelian Fourier analysis and properties of solutions of S-unit equations. Our approach is based on the study of distribution of images of polynomial maps in X. This reduces the proof to the investigation of certain Diophantine inequalities which are analysed using W. Schmidt's subspace theorem. In order to prove an effective version of Theorem 1.2, one would need to estimate the size of non-degenerate solutions of these Diophantine inequalities in terms of complexities of coefficients (cf. Proposition 3.1 below). However, this seems to be far out of reach of available techniques when $s>2$.

Finally, we discuss the problem of exponential mixing for shapes in $\mathrm{Aut}(X)$. This notion was introduced by K. Schmidt in [20] in order to better understand Ledrappier's examples [13] which are not mixing of higher order. A *shape* in $\mathrm{Aut}(X)$ is a collection of elements $\alpha_0,...,\alpha_s\in\mathrm{Aut}(X)$. We say that the shape is *mixing* if, for every $f_0,...,f_s\in L^\infty(X)$,

$$\int_X\left(\prod_{i=0}^s f_i(\alpha_i^n x)\right)d\mu(x)\longrightarrow\prod_{i=0}^s\int_X f_i\,d\mu$$

as $n\to\infty$. This property has been extensively studied in the context of commuting automorphisms of compact abelian groups (see, for instance, [6], [21, Chapter VIII], [26], and [27]).

We establish quantitative mixing for commuting Anosov shapes. We say that the shape $\alpha_0,...,\alpha_s$ is *Anosov* if $\alpha_{i_1}\alpha_{i_2}^{-1}$ is an Anosov map for all $i_1\neq i_2$.

THEOREM 1.3. *Let* X *be a compact nilmanifold and* $\alpha_0,...,\alpha_s\in\mathrm{Aut}(X)$ *be a commuting Anosov shape. Then there exists* $\varrho=\varrho(\theta)\in(0,1)$ *such that, for every* $f_0,...,f_s\in C^\theta(X)$ *and* $n\in\mathbb{N}$,

$$\int_X\left(\prod_{i=0}^s f_i(\alpha_i^n x)\right)d\mu(x)=\prod_{i=0}^s\int_X f_i\,d\mu+O\left(\varrho^n\prod_{i=0}^s\|f_i\|_{C^\theta}\right). \tag{1.2}$$

1.2. Applications to rigidity

The exponential mixing property played an important role in the program of classification of smooth Anosov higher-rank $\mathbb{Z}^l$-actions on compact manifolds. It is expected that all such actions can be built from actions by automorphisms on nilmanifolds. Fisher, Kalinin and Spatzier in [8] applied the exponential 2-mixing property and regularity results from [17] to extend their results for Anosov $\mathbb{Z}^l$-actions on tori to actions on nilmanifolds.

THEOREM 1.4. (Fisher, Kalinin, and Spatzier) *Let α be a C^∞-action of $\mathbb{Z}^l$, $l\geqslant 2$, on a compact nilmanifold X and let $\varrho:\mathbb{Z}^l\to\mathrm{Aut}(X)$ be the map induced by the action of $\alpha(\mathbb{Z}^l)$ on the fundamental group of X. Assume that there is a $\mathbb{Z}^2$ subgroup of $\mathbb{Z}^l$ such that $\varrho(z)$ is ergodic for every non-zero $z\in\mathbb{Z}^2$, and there is an Anosov element for α in each Weyl chamber of ϱ. Then α is C^∞-isomorphic to ϱ.*

In fact, this application to global rigidity was our original motivation to establish the exponential mixing property for nilmanifold automorphisms.

Recently, Rodriguez Hertz and Wang [18] generalised Theorem 1.4 and established a global rigidity result using only existence of a single Anosov element. Again, they crucially use the exponential mixing property, and reduce the problem to the prior result by showing existence of many Anosov elements.

We also use the exponential mixing property to establish cocycle rigidity for higher-rank $\mathbb{Z}^l$-actions by automorphisms of nilmanifolds, extending the results of Katok and Spatzier [11], [12]. A C^∞-*cocycle* is a C^∞-map $c:\mathbb{Z}^l\times X\to\mathbb{R}$ satisfying the identity

$$c(z_1+z_2,x)=c(z_1,z_2x)+c(z_2,x)\quad \text{for } z_1,z_2\in\mathbb{Z}^l \text{ and } x\in X.$$

Two cocycles c_1 and c_2 are called *smoothly cohomologous* if there exists $b\in C^\infty(X)$ such that

$$c_1(z,x)=c_2(z,x)+b(zx)-b(x)\quad \text{for } z\in\mathbb{Z}^l \text{ and } x\in X.$$

We call a cocycle *constant* if it does not depend on $x\in X$. We prove that cocycles over genuine higher-rank actions by automorphisms on nilmanifolds are smoothly cohomologous to constant cocycles. This phenomenon was first discovered by Katok and Spatzier in [11] for certain higher-rank Anosov actions. Using our methods, we generalise this cocycle rigidity theorem to actions by automorphisms on nilmanifolds. We emphasise that the action in the following theorem need not be Anosov.

THEOREM 1.5. *Let $\alpha:\mathbb{Z}^l\to\mathrm{Aut}(X)$ be an action on a compact nilmanifold X. Assume that there is a $\mathbb{Z}^2$ subgroup of $\mathbb{Z}^l$ such that $\alpha(z)$ is ergodic for every non-zero $z\in\mathbb{Z}^2$. Then every smooth $\mathbb{R}$-valued cocycle is smoothly cohomologous to a constant cocyle.*

For certain actions by partially hyperbolic left translations on homogeneous spaces G/Γ, where G is a semisimple Lie group and Γ is a lattice in G, a similar theorem was proved by Damjanovic and Katok [3]–[5] and Wang [25]. We note that these authors also prove Hölder versions of this result which are not amenable to our techniques. Furthermore, cocycle rigidity results are proven for small perturbations of these actions on G/Γ in [5] and [25]. Again we cannot obtain these results by our methods.

Acknowledgements

A. G. would like to thank the University of Michigan for hospitality during his visit when the work on this project had started. R. S. thanks the University of Bristol for hospitality and support during this work.

2. Exponential 3-mixing

In this section we prove Theorem 1.1. We start by setting up basic notation, which will be also used in subsequent sections. Then, in §§2.2–2.4, we collect some auxiliary estimates. The proof of Theorem 1.1 is divided into two parts. We first give a proof under an irreducibility condition in §2.5, and then in §2.6 prove Theorem 1.1 in general using an inductive argument.

We note that if the reader is only interested in exponential 2-mixing, then the results of §2.3 are not needed, and in §2.5, one only needs to consider case 1. This makes the proof much simpler.

2.1. Notation

Let G be a connected simply connected nilpotent Lie group, Λ be a discrete cocompact subgroup, and $X=G/\Lambda$ be the corresponding nilmanifold equipped with the invariant probability measure μ. We fix a right-invariant Riemannian metric d on G which also defines a Riemannian metric on X. Let $\mathcal{L}(G)$ be the Lie algebra of G and $\exp\colon\mathcal{L}(G)\to G$ be the exponential map. The lattice subgroup Λ defines a rational structure on $\mathcal{L}(G)$. For a field $K\supset\mathbb{Q}$, we denote by $\mathcal{L}(G)_K$ the corresponding Lie algebra over K. Denoting the commutator subgroup by G', let $\pi\colon G\to G/G'$ be the factor map. We also have the corresponding map $D\pi\colon\mathcal{L}(G)\to\mathcal{L}(G/G')$. We fix an identification $G/G'\simeq\mathcal{L}(G/G')\simeq\mathbb{R}^d$ that respects the rational structures.

Every automorphism β of G defines a Lie-algebra automorphism $D\beta\colon\mathcal{L}(G)\to\mathcal{L}(G)$ such that $\beta\circ\exp=\exp\circ D\beta$. If $\beta(\Lambda)=\Lambda$, then $D\beta$ preserves the rational structure of $\mathcal{L}(G)$ defined by Λ. In particular, given an action $\alpha\colon\mathbb{Z}^l\to\mathrm{Aut}(X)$ on the nilmanifold $X=G/\Lambda$, we obtain a homomorphism $D\alpha\colon\mathbb{Z}^l\to\mathrm{GL}(\mathcal{L}(G)_{\mathbb{Q}})$.

For a multiplicative character $\chi:\mathbb{Z}^l\to\mathbb{C}^\times$, we set

$$\mathcal{L}_\chi := \{u\in\mathcal{L}(G)\otimes\mathbb{C} : D\alpha(z)u=\chi(z)u \text{ for } z\in\mathbb{Z}^l\}.$$

Let $\mathcal{X}(\alpha)$ denote the set of characters χ appearing in the action $D\alpha$ on $\mathcal{L}(G)$, and $\mathcal{X}'(\alpha)\subset\mathcal{X}(\alpha)$ be the set of characters appearing in the action on $\mathcal{L}(G)/\mathcal{L}(G)'$.

2.2. Estimates on Lyapunov exponents

Since $\alpha(\mathbb{Z}^l)$ preserves the rational structure on $\mathcal{L}(G)$ defined by the lattice Λ, it follows that each character χ in $\mathcal{X}(\alpha)$ is of the form $\chi(z)=u_1^{z_1} \dots u_l^{z_l}$ where the u_i's are algebraic numbers. The Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ naturally acts on $\mathcal{X}(\alpha)$ and $\mathcal{X}'(\alpha)$. Let $\mathcal{X}_0\subset\mathcal{X}'(\alpha)$ be one of the Galois orbits.

Lemma 2.1. *Suppose that every $\alpha(z)$, $z\neq0$, acts ergodically on X. Then there exists $c>0$ such that*

$$\max_{\chi\in\mathcal{X}_0}|\chi(z)|\geqslant\exp(c\|z\|) \quad \textit{for all } z\in\mathbb{Z}^l.$$

Proof. By [2, Theorem 5.4.13], $\Lambda G'/G'$ is a lattice in G/G', and the action α defines the action on the torus $T:=G/\Lambda G'$ by linear automorphisms. Let V be the subspace of $\mathcal{L}(G)/\mathcal{L}(G)'$ spanned by the χ-eigenspaces with $\chi\in\mathcal{X}_0$. Clearly, this subspace is invariant under $\alpha(\mathbb{Z}^l)$ and is defined over $\mathbb{Q}$. Hence, it defines an α-invariant subtorus $T_{\mathcal{X}_0}$ of T. Since $\alpha(z)|_T$ is ergodic when $z\neq0$, it follows that the corresponding linear map has no roots of unity as eigenvalues. This implies that $\alpha(z)|_{T_{\mathcal{X}_0}}$ is also ergodic.

Consider a linear map $\ell:\mathbb{R}^l\to\mathbb{R}^{|\mathcal{X}_0|}$ which is defined for $z\in\mathbb{Z}^l$ by

$$\ell(z):=(\log|\chi(z)|:\chi\in\mathcal{X}_0)$$

and extended to $\mathbb{R}^l$ by linearity. Since for every $z\in\mathbb{Z}^l\setminus\{0\}$, the automorphism $\alpha(z)$ acting on $T_{\mathcal{X}_0}$ is ergodic, we have $\ell(z)\neq0$ by [9, Lemma 3.2]. Hence, $\ell|_{\mathbb{Z}^l}$ is injective.

We also claim that $\ell(\mathbb{Z}^l)$ is discrete. We consider the embedding $\mathbb{Z}^l\to\mathrm{GL}(V)$ defined by α. Since $\alpha(\mathbb{Z}^l)$ preserves the integral lattice in V corresponding to the torus $T_{\mathcal{X}_0}$, it follows that the image of this embedding is discrete. In other words, the subset

$$\{(\chi(z):\chi\in\mathcal{X}_0):z\in\mathbb{Z}^l\}$$

of $(\mathbb{C}^\times)^{|\mathcal{X}_0|}$ is discrete. Since the kernel of the natural homomorphism $(\mathbb{C}^\times)^{|\mathcal{X}_0|}\to\mathbb{R}^{|\mathcal{X}_0|}$ defined by $s\mapsto\log|s|$ is compact, this implies that $\ell(\mathbb{Z}^l)$ is discrete, as claimed. Since $\ell(\mathbb{Z}^l)$ is discrete and has rank l, it follows that the space $\ell(\mathbb{R}^l)$ has dimension l, and, in particular, the map ℓ is injective. Therefore, by compactness, there exists $c>0$ such that, for every $z\in\mathbb{R}^l$, we have

$$\max_{\chi\in\mathcal{X}_0}\log|\chi(z)|\geqslant c\|z\|.$$

This implies the lemma. □

Lemma 2.1 shows that in Theorem 1.1 we may replace $N(z_0,z_1,z_2)$ by

$$\exp\Big(\min_{i\neq j}\max_{\chi\in\mathcal{X}_0}|\chi(z_i-z_j)|\Big).$$

2.3. Diophantine estimates

Recall that the (*absolute*) *height* of an algebraic number u is defined by

$$H(u)=\bigg(\prod_v\max\{1,|u|_v\}\bigg)^{1/[\mathbb{Q}(u):\mathbb{Q}]},$$

where $|\cdot|_v$ denote suitably normalised absolute values of the field $\mathbb{Q}(u)$. When u is an algebraic integer, the height can be computed as

$$H(u)=\bigg(\prod_i\max\{1,|u_i|\}\bigg)^{1/[\mathbb{Q}(u):\mathbb{Q}]},$$

where the u_i's denote all the Galois conjugates of u.

The following result is deduced from the work of Waldschmidt [24, Corollary 10.1].

PROPOSITION 2.2. *Let $u_1,\dots u_l,u\in\mathbb{C}$ be algebraic numbers and $z=(z_1,\dots,z_l)\in\mathbb{Z}^l$. Then there exist $c_1,c_2,c_3>1$, depending on $u_1,\dots,u_l$ and $[\mathbb{Q}(u):\mathbb{Q}]$, such that, assuming that*

$$\|z\|\geqslant\log(c_2H(u)) \tag{2.1}$$

and

$$u_1^{z_1}\dots u_l^{z_l}u\neq 1,$$

we have the estimate

$$|u_1^{z_1}\dots u_l^{z_l}u-1|\geqslant\exp\bigg(-c_1\log(c_2H(u))\log\bigg(\frac{c_3\|z\|}{\log(c_2H(u))}\bigg)\bigg). \tag{2.2}$$

Surprisingly, it turns out that the term $\log(c_2H(u))$ in the denominator is essential to establish exponential 3-mixing (cf. (2.28)–(2.30) below).

Proof. We note that, since $H(u)\geqslant 1$ and (2.1) holds, the right-hand side of (2.2) is bounded from above by

$$\exp(-c_1 \log c_2 \log c_3).$$

Taking the constants sufficiently large, we may arrange that this quantity is bounded by $\frac{1}{2}$. Then (2.2) trivially holds when $|u_1^{z_1} \dots u_l^{z_l} u-1|\geqslant \frac{1}{2}$, and without loss of generality we assume that $|u_1^{z_1} \dots u_l^{z_l} u-1|\leqslant \frac{1}{2}$.

Let log denote the principle value of the (complex) logarithm. There exists $z_0\in\mathbb{Z}$ such that $|z_0|\ll\|z\|$ and

$$T := \log(u_1^{z_1} \dots u_n^{z_n} u) = \pi i z_0 + z_1 \log u_1 + \dots + z_l \log u_l + \log u.$$

It is convenient to set $u_0=-1$, so that $\log u_0=\pi i$. (Here, but not elsewhere, i denotes the imaginary unit.) Let $S:=u_1^{z_1} \dots u_l^{z_l} u$. Since $|S-1|\leqslant \frac{1}{2}$,

$$|T| = |\log S| \leqslant 2|S-1|.$$

Hence, it is sufficient to establish a lower bound for $|T|$. Note that, since $S\neq 1$, we have $T\neq 0$. For this purpose, we use [24, Corollary 10.1], which we now recall. We note that the result in [24] is stated using the logarithmic height while here we use the exponential height. For simplicity, we take $E=e$ and $f=1$.

Let $D=[\mathbb{Q}(u_0, \dots, u_l, u):\mathbb{Q}]$, $A_0, \dots, A_l$ and B be numbers, greater than e, such that

$$H(u_i) \leqslant A_i, \;\; i=0, \dots, l, \quad H(u)\leqslant B \quad \text{and} \quad \sum_{i=0}^{l} \frac{|\log u_i|}{\log A_i} + \frac{|\log u|}{\log B} \leqslant e^{-1}(l+2)D. \tag{2.3}$$

We set

$$\begin{aligned}
A &= \max\{A_0, \dots, A_l, B\}, \\
M &= \max_{i=0,\dots,l} \left(\frac{1}{\log A_i} + \frac{|z_i|}{\log B} \right), \\
Z_0 &= \max\{7+3\log(l+2), \log D\}, \\
G_0 &= \max\{4(l+2)Z_0, \log M, \log D\}, \\
U_0 &= \max\{D^2 \log A, D^{l+4} G_0 Z_0 \log A_0 \dots \log A_l \log B\}.
\end{aligned}$$

Then, according to [24, Corollary 10.1],

$$|T| \geqslant \exp(-cU_0), \tag{2.4}$$

where c is an explicit positive constant depending only on n. We set $B:=c_2H(u)$ with $c_2>1$. We note that

$$|\log u|^2 \leqslant \pi^2+(\log|u|)^2 \leqslant \pi^2+[\mathbb{Q}(u):\mathbb{Q}]^2(\log H(u))^2.$$

Therefore, taking A_i sufficiently large, depending on u_i, and sufficiently large c_2, we may arrange that (2.3) holds. If c_2 is sufficiently large, $A=B$. Under the assumption (2.1), we have $M\leqslant c_3\|z\|/\log B$ with sufficiently large c_3 and also

$$G_0=\log M \leqslant \log\left(\frac{c_3\|z\|}{\log B}\right).$$

Moreover, if c_3 is sufficiently large, then

$$U_0 \ll \log\left(\frac{c_3\|z\|}{\log B}\right)\log B.$$

Therefore, estimate (2.4) implies that

$$|T| \geqslant \exp\left(-c_1\log\left(\frac{c_3\|z\|}{\log B}\right)\log B\right),$$

where c_1 is an explicit positive constant. This completes the proof of the proposition. □

2.4. Equidistribution of box maps

A *box map* is an affine map

$$\iota: B:=[0,T_1]\times\ldots\times[0,T_k]\longrightarrow \mathcal{L}(G)$$

of the form

$$\iota:(t_1,\ldots,t_k)\longmapsto v+t_1w_1+\ldots+t_kw_k, \tag{2.5}$$

with $v,w_1,\ldots,w_k\in\mathcal{L}(G)$. We shall use the following result, which is a variation of our theorem [9, Theorem 2.1], that implies equidistribution of box maps under suitable Diophantine conditions. This result is based on the work of Green and Tao [10].

We denote by $|B|$ the k-dimensional volume of the box B.

THEOREM 2.3. *Fix $0<\theta\leqslant 1$. There exist $L_1,L_2>0$ such that for every $\delta\in(0,\delta_0)$ and every box map $\iota: B\to\mathcal{L}(G)$, one of the following conditions holds*:

(i) *For every θ-Hölder function $f: X\to\mathbb{R}$, $u\in\mathcal{L}(G)$, and $g\in G$,*

$$\left|\frac{1}{|B|}\int_B f(\exp(u)\exp(\iota(t))g\Lambda)\,dt-\int_X f\,d\mu\right| \leqslant \delta\|f\|_{C^\theta}.$$

(ii) *There exists $z\in\mathbb{Z}^d\setminus\{0\}$ such that*

$$\|z\|\ll\delta^{-L_1}\quad \textit{and}\quad |\langle z, D\pi(w_i)\rangle| \ll \frac{\delta^{-L_2}}{T_i}\ \textit{for all}\ i=1,\ldots,k.$$

Proof. In the case of Lipschitz functions f, this is [9, Theorem 2.1], and the analogous result for Hölder functions can be deduced by a standard approximation argument. Indeed, suppose that for some $f\in C^\theta(X)$, $u\in\mathcal{L}(G)$, and $g\in G$,

$$\left|\frac{1}{|B|}\int_B f(\exp(u)\exp(\iota(t))g\Lambda)\,dt-\int_X f\,d\mu\right|>\delta\|f\|_{C^\theta}. \tag{2.6}$$

Then one can find a Lipschitz function f_ε such that

$$\|f_\varepsilon-f\|_{C^0}\leqslant\varepsilon^\theta\|f\|_{C^\theta}\quad\text{and}\quad\|f_\varepsilon\|_{\mathrm{Lip}}\ll\varepsilon^{-\dim(X)-1}\|f\|_{C^0}$$

(see, for instance, [9, Lemma 2.4]). Then taking $\varepsilon=\left(\frac{1}{3}\delta\right)^{1/\theta}$, we deduce from (2.6) that

$$\begin{aligned}\left|\frac{1}{|B|}\int_B f_\varepsilon(\exp(u)\exp(\iota(t))g\Lambda)\,dt-\int_X f_\varepsilon\,d\mu\right|&>(\delta-2\varepsilon^\theta)\|f\|_{C^\theta}\\&\gg\varepsilon^{\dim(X)+1}(\delta-2\varepsilon^\theta)\|f_\varepsilon\|_{\mathrm{Lip}}\\&\gg\delta^{(\dim(X)+1)/\theta+1}\|f_\varepsilon\|_{\mathrm{Lip}}.\end{aligned}$$

Now the theorem for Lipschitz functions implies that (ii) holds with some $L_1,L_2>0$ depending on θ. □

2.5. 3-mixing under an irreducibility condition

The action of $D\alpha(\mathbb{Z}^l)$ preserves the rational structure on $\mathcal{L}(G)$ defined by the lattice Λ. In particular, it follows that each character χ in $\mathcal{X}(\alpha)$ is of the form $\chi(z)=u_1^{z_1}\dots u_l^{z_l}$ where the u_i's are algebraic numbers. The Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ naturally acts on $\mathcal{X}(\alpha)$ and on $\mathcal{L}(G)_{\overline{\mathbb{Q}}}$. We fix an orbit $\mathcal{X}_0\subset\mathcal{X}'(\alpha)$ of the Galois group and for each $\chi\in\mathcal{X}_0$, we fix a vector $w_\chi\in\mathcal{L}_\chi$ whose coordinates are algebraic integers, so that the vectors w_χ are also conjugate under the action of the Galois group. Let $W_\mathbb{C}$ be the Lie subalgebra of $\mathcal{L}(G)\otimes\mathbb{C}$ generated by the vectors w_χ, $\chi\in\mathcal{X}_0$, and $W=W_\mathbb{C}\cap\mathcal{L}(G)$. We also fix a basis $\{w_i\}_i$ of W.

In this section, we prove Theorem 1.1 under the irreducibility assumption that $D\pi(W)$ is not contained in a proper rational subspace. Let $\bar{w}_\chi=D\pi(w_\chi)$, $\chi\in\mathcal{X}_0$. We observe that under this assumption the coordinates of each of the vectors $\bar{w}_\chi$ are linearly independent over $\mathbb{Q}$. Indeed, if we suppose that $\langle a,\bar{w}_\chi\rangle=0$ for some $a\in\mathbb{Q}^d\setminus\{0\}$, then applying the action of the Galois group, we deduce that $\langle a,\bar{w}_\chi\rangle=0$ for all $\chi\in\mathcal{X}_0$. Since $D\pi(W)$ is spanned over $\mathbb{C}$ by the vectors w_χ, $\chi\in\mathcal{X}_0$, this would imply that $D\pi(W)$ is contained in a proper rational subspace, which contradicts our assumption.

Let

$$N=N(z_1,z_2,z_3):=\exp\Bigl(\min_{i\neq j}\|z_i-z_j\|\Bigr). \tag{2.7}$$

Without loss of generality, we may assume that $z_0=0$ and $N=\exp(\|z_1-z_2\|)$. We set $\varepsilon=N^{-\varkappa}$, where $\varkappa>0$ is a fixed parameter which is sufficiently small and will be specified later (see (2.21), (2.23), (2.26) and (2.30) below).

We fix a fundamental domain $F\subset G$ for $X=G/\Lambda$ and set $E=\exp^{-1}(F)$. As in [9, §3], we may arrange that E is bounded and has piecewise smooth boundary. Since the Haar measure on G is the image under exp of a suitably normalised Lebesgue measure on $\mathcal{L}(G)$ [2, Theorem 1.2.10], we obtain

$$\begin{aligned}\int_X f_0(x)f_1(\alpha(z_1)x)f_2(\alpha(z_2)x)\,d\mu(x)\\ =\int_E f_0(\exp(u)\Lambda)f_1(\exp(D\alpha(z_1)u)\Lambda)f_2(\exp(D\alpha(z_2)u)\Lambda)\,du.\end{aligned}\tag{2.8}$$

We choose a basis of $\mathcal{L}(G)$ that contains the basis $\{w_i\}_i$ of W and tessellate $\mathcal{L}(G)$ by cubes C of size ε with respect to this basis. Since E has piecewise smooth boundary, we obtain

$$\left|E\setminus\bigcup_{C\subset E}C\right|\ll\varepsilon,\tag{2.9}$$

and

$$\begin{aligned}&\int_E f_0(\exp(u)\Lambda)f_1(\exp(D\alpha(z_1)u)\Lambda)f_2(\exp(D\alpha(z_2)u)\Lambda)\,du\\ &\quad=\sum_{C\subset E}\int_C f_0(\exp(u)\Lambda)f_1(\exp(D\alpha(z_1)u)\Lambda)f_2(\exp(D\alpha(z_2)u)\Lambda)\,du\\ &\qquad+O(\varepsilon\|f_0\|_{C^\theta}\|f_1\|_{C^\theta}\|f_2\|_{C^\theta}).\end{aligned}\tag{2.10}$$

For every cube C, we pick a point $u_C\in C$. Then, since f_0 is θ-Hölder,

$$\begin{aligned}&\int_C f_0(\exp(u)\Lambda)f_1(\exp(D\alpha(z_1)u)\Lambda)f_2(\exp(D\alpha(z_2)u)\Lambda)\,du\\ &\quad=f_0(\exp(u_C)\Lambda)\int_C f_1(\exp(D\alpha(z_1)u)\Lambda)f_2(\exp(D\alpha(z_2)u)\Lambda)\,du\\ &\qquad+O(\varepsilon^\theta\|f_0\|_{C^\theta}\|f_1\|_{C^\theta}\|f_2\|_{C^\theta}).\end{aligned}\tag{2.11}$$

We decompose each cube C as $C=B'+B$, where B is a cube in W and B' is a cube in the complementary subspace.

We claim that, for sufficiently small $\varkappa>0$ and all sufficiently large N defined in (2.7),

$$\begin{aligned}\frac{1}{|B|}\int_B f_1(\exp(v_1+D\alpha(z_1)b)\Lambda)f_2(\exp(v_2+D\alpha(z_2)b)\Lambda)\,db\\ =\left(\int_X f_1\,d\mu\right)\left(\int_X f_2\,d\mu\right)+O(N^{-\varkappa}\|f_1\|_{C^\theta}\|f_2\|_{C^\theta}),\end{aligned}\tag{2.12}$$

uniformly over the cubes B and $v_1, v_2 \in \mathcal{L}(G)$.

Suppose first that (2.12) holds. Then using uniformity over v_1 and v_2, we deduce that

$$\begin{aligned}
&\frac{1}{|C|}\int_C f_1(\exp(D\alpha(z_1)u)\Lambda)f_2(\exp(D\alpha(z_2)u)\Lambda)\,du \\
&\quad= \frac{1}{|B'|\,|B|}\int_{B'}\int_B f_1(\exp(D\alpha(z_1)b'+D\alpha(z_1)b)\Lambda)f_2(\exp(D\alpha(z_2)b'+D\alpha(z_2)b)\Lambda)\,db\,db' \\
&\quad= \left(\int_X f_1\,d\mu\right)\left(\int_X f_2\,d\mu\right)+O(N^{-\varkappa}\|f_1\|_{C^\theta}\|f_2\|_{C^\theta}).
\end{aligned}$$

Combining this estimate with (2.10) and (2.11), we obtain

$$\begin{aligned}
&\int_E f_0(\exp(u)\Lambda)f_1(\exp(D\alpha(z_1)u)\Lambda)f_2(\exp(D\alpha(z_2)u)\Lambda)\,du \\
&\qquad= \left(\sum_{C\subset E} f_0(\exp(u_C)\Lambda)|C|\right)\left(\int_X f_1\,d\mu\right)\left(\int_X f_2\,d\mu\right) \\
&\qquad\quad +O((N^{-\varkappa}+\varepsilon^\theta)\|f_0\|_{C^\theta}\|f_1\|_{C^\theta}\|f_2\|_{C^\theta}).
\end{aligned}$$

Since f is θ-Hölder and (2.9) holds,

$$\begin{aligned}
\sum_{C\subset E} f_0(\exp(u_C)\Lambda)|C| &= \sum_{C\subset E}\int_C f_0(\exp(u)\Lambda)\,du+O(\varepsilon^\theta\|f_0\|_{C^\theta}) \\
&= \int_E f_0(\exp(u)\Lambda)\,du+O((\varepsilon+\varepsilon^\theta)\|f_0\|_{C^\theta}) \\
&= \int_X f_0\,d\mu+O(\varepsilon^\theta\|f_0\|_{C^\theta}).
\end{aligned}$$

Hence,

$$\begin{aligned}
&\int_E f_0(\exp(u)\Lambda)f_1(\exp(D\alpha(z_1)u)\Lambda)f_2(\exp(D\alpha(z_2)u)\Lambda)\,du \\
&\quad= \left(\int_X f_0\,d\mu\right)\left(\int_X f_1\,d\mu\right)\left(\int_X f_2\,d\mu\right)+O(N^{-\varkappa\theta}\|f_0\|_{C^\theta}\|f_1\|_{C^\theta}\|f_2\|_{C^\theta}).
\end{aligned} \tag{2.13}$$

This proves the required estimate when N is sufficiently large, and it is also clear that this estimate holds for N in bounded intervals. Hence, Theorem 1.1 follows. Now it remains to prove the claim (2.12).

To prove (2.12), we apply Theorem 2.3 to the nilmanifold $X\times X=(G\times G)/(\Lambda\times\Lambda)$ with $\delta=N^{-\varkappa}$. We assume that N is sufficiently large, so that Theorem 2.3 applies. Let $f=f_1\otimes f_2$. Clearly,

$$\int_{X\times X} f\,d(\mu\otimes\mu)=\left(\int_X f_1\,d\mu\right)\left(\int_X f_1\,d\mu\right) \quad\text{and}\quad \|f\|_{C^\theta}\ll\|f_1\|_{C^\theta}\|f_2\|_{C^\theta}.$$

We consider the map $\iota: [0,\varepsilon]^k \to \mathcal{L}(G)$ defined by

$$\iota(t) = \left(v_1' + \sum_{i=1}^{k} t_i D\alpha(z_1) w_i, v_2' + \sum_{i=1}^{k} t_i D\alpha(z_2) w_i \right),$$

with suitably chosen $v_1', v_2' \in \mathcal{L}(G)$, so that

$$\int_B f_1(\exp(v_1 + D\alpha(z_1)b)\Lambda) f_2(\exp(v_2 + D\alpha(z_2)b)\Lambda)\, db = \int_{[0,\varepsilon]^k} f(\iota(t)\Lambda)\, dt.$$

It is sufficient to prove that

$$\varepsilon^{-k} \int_{[0,\varepsilon]^k} f(\iota(t)\Lambda)\, dt = \int_{X\times X} f\, d(\mu\otimes\mu) + O(\delta \|f\|_{C^\theta}).$$

Applying Theorem 2.3, we deduce that either

$$\left| \varepsilon^{-k} \int_{[0,\varepsilon]^k} f(\iota(t)\Lambda)\, dt - \int_{X\times X} f\, d(\mu\otimes\mu) \right| \leqslant \delta \|f\|_{C^\theta}, \tag{2.14}$$

or there exists $(a_1, a_2) \in (\mathbb{Z}^d)^2 \setminus \{(0,0)\}$ such that

$$\max\{\|a_1\|, \|a_2\|\} \ll \delta^{-L_1} = N^{\varkappa L_1} \tag{2.15}$$

and

$$|\langle a_1, (D\pi) D\alpha(z_1) w_i\rangle + \langle a_2, (D\pi) D\alpha(z_2) w_i\rangle| \ll \frac{\delta^{-L_2}}{\varepsilon} = N^{\varkappa(L_2+1)} \tag{2.16}$$

for all $i = 1, ..., k$.

We shall show that if $\varkappa > 0$ is sufficiently small and N is sufficiently large, then either (2.15) or (2.16) fails. Suppose that both (2.15) and (2.16) holds. Since each of the vectors w_χ, $\chi \in \mathcal{X}_0$, is a linear combination of vectors w_i, we deduce from (2.16) that

$$|\langle a_1, (D\pi) D\alpha(z_1) w_\chi\rangle + \langle a_2, (D\pi) D\alpha(z_2) w_\chi\rangle| \ll N^{\varkappa(L_2+1)} \quad \text{for all } \chi \in \mathcal{X}_0. \tag{2.17}$$

As $D\alpha(z) w_\chi = \chi(z) w_\chi$ and $\bar{w}_\chi = D\pi(w_\chi)$, (2.17) becomes

$$|\chi(z_1)\langle a_1, \bar{w}_\chi\rangle + \chi(z_2)\langle a_2, \bar{w}_\chi\rangle| \ll N^{\varkappa(L_2+1)} \quad \text{for all } \chi \in \mathcal{X}_0. \tag{2.18}$$

We divide the argument into three cases.

Case 1. $a_1 = 0$. Then $a_2 \neq 0$ and $\langle a_2, \bar{w}_\chi\rangle \neq 0$. Moreover, by [1, Theorem 7.3.2],

$$|\langle a_2, \bar{w}_\chi\rangle| \gg \|a_2\|^{-d-1} \gg N^{-\varkappa L_1(d+1)}. \tag{2.19}$$

By Lemma 2.1, there exists $\chi\in\mathcal{X}_0$ such that $|\chi(z_2)|\geqslant N^c$ with fixed $c>0$. Hence, it follows from (2.18) that

$$|\langle a_2,\bar{w}_\chi\rangle|\ll N^{\varkappa(L_2+1)-c}. \tag{2.20}$$

We assume that the parameter $\varkappa>0$ satisfies

$$-\varkappa L_1(d+1)>\varkappa(L_2+1)-c. \tag{2.21}$$

Comparing (2.19) and (2.20), we get a contradiction if N is sufficiently large. Hence, we may assume that $a_1\neq 0$.

Case 2. *$a_1\neq 0$ and $\chi(z_1)\langle a_1,\bar{w}_\chi\rangle+\chi(z_2)\langle a_2,\bar{w}_\chi\rangle=0$ for some $\chi\in\mathcal{X}_0$.* As the Galois group acts transitively on the set $\mathcal{X}_0$, it follows that this equality holds for all $\chi\in\mathcal{X}_0$. By Lemma 2.1, there exists $\chi\in\mathcal{X}_0$ such that $|\chi(z_2-z_1)|\geqslant N^c$ with fixed $c>0$. Then

$$|\langle a_2,\bar{w}_\chi\rangle|=|\chi(z_1-z_2)|\,|\langle a_1,\bar{w}_\chi\rangle|\geqslant N^c|\langle a_1,\bar{w}_\chi\rangle|. \tag{2.22}$$

Since $a_1\neq 0$, we have $\langle a_1,\bar{w}_\chi\rangle\neq 0$, and by [1, Theorem 7.3.2],

$$|\langle a_1,\bar{w}_\chi\rangle|\gg\|a_1\|^{-d-1}\gg N^{-\varkappa L_1(d+1)}.$$

On the other hand,

$$|\langle a_2,\bar{w}_\chi\rangle|\ll\|a_2\|\ll N^{\varkappa L_1}.$$

Hence, we deduce that

$$N^{-\varkappa L_1(d+1)+c}\ll N^{\varkappa L_1}.$$

We choose the parameter $\varkappa>0$ so that

$$-\varkappa L_1(d+1)+c>\varkappa L_1. \tag{2.23}$$

Then when N is sufficiently large, we get a contradiction.

Case 3. *$a_1\neq 0$ and $\chi(z_1)\langle a_1,\bar{w}_\chi\rangle+\chi(z_2)\langle a_2,\bar{w}_\chi\rangle\neq 0$ for all $\chi\in\mathcal{X}_0$.* This is the most difficult part of the proof.

Since $a_1\neq 0$, we have $\langle a_1,\bar{w}_\chi\rangle\neq 0$, and, by [1, Theorem 7.3.2],

$$|\langle a_1,\bar{w}_\chi\rangle|\gg\|a_1\|^{-d-1}\gg N^{-\varkappa L_1(d+1)}.$$

We set $u=-\langle a_2,\bar{w}_\chi\rangle/\langle a_1,\bar{w}_\chi\rangle$. By Lemma 2.1, there exists $\chi\in\mathcal{X}_0$ such that $|\chi(z_1)|\gg N^c$ with fixed $c>0$. It follows from (2.18) that for this χ, we have the estimate

$$|\chi(z_2-z_1)u-1|\ll\frac{N^{\varkappa(L_2+1)}}{|\chi(z_1)|\,|\langle a_1,\bar{w}_\chi\rangle|}\ll N^{\varkappa(L_2+1+L_1(d+1))-c}. \tag{2.24}$$

Let $K_1 := \varkappa(L_2+1+L_1(d+1))-c$.

Next, we compare this estimate with the lower estimate provided by Proposition 2.2. We note that

$$H(u) = \prod_v \max\{|\langle a_1, \overline{w}_\chi\rangle|_v, |\langle a_2, \overline{w}_\chi\rangle|_v\}^{1/[\mathbb{Q}(u):\mathbb{Q}]}.$$

For all non-Archimedian places v,

$$|\langle a_i, \overline{w}_\chi\rangle|_v \leqslant 1,$$

and for all Archimedian v,

$$|\langle a_i, \overline{w}_\chi\rangle|_v \ll \|a_i\| \ll N^{\varkappa L_1}.$$

Therefore,

$$H(u) \ll N^{K_2}, \tag{2.25}$$

where $K_2 := \varkappa L_1'$ with fixed $L_1' > 0$. We take the parameter $\varkappa > 0$ so that

$$K_2 = \varkappa L_1' < 1. \tag{2.26}$$

Then assuming that N is sufficiently large, we obtain

$$\log(c_2 H(u)) \leqslant \log(c_2' N^{K_2}) \leqslant \log N, \tag{2.27}$$

where $c_2' > 1$ depends on the implicit constant in the estimate (2.25). We recall that we have chosen the indices so that

$$\log N = \|z_2 - z_1\|.$$

Since (2.27) holds, Proposition 2.2 applies, and we deduce that

$$|\chi(z_2-z_1)u-1| \geqslant \exp\left(-c_1 \log(c_2 H(u)) \log\left(\frac{c_3\|z_2-z_1\|}{\log(c_2 H(u))}\right)\right).$$

Without loss of generality, we may assume that $c_3 > e$. Since the function $x \mapsto x\log(C/x)$ is increasing for $x \leqslant C/e$, we deduce that

$$\begin{aligned} |\chi(z_2-z_1)u-1| &\geqslant \exp\left(-c_1 \log(c_2' N^{K_2}) \log\left(\frac{c_3 \log N}{\log(c_2' N^{K_2})}\right)\right) \\ &\geqslant \exp(-c_1 \log(c_2' N^{K_2}) \log(c_3 K_2^{-1})). \end{aligned} \tag{2.28}$$

Comparing (2.24) and (2.28), we conclude that

$$K_2' \log N + M_2 \leqslant K_1 \log N + M_1, \tag{2.29}$$

where $K_2' := -c_1 K_2 \log(c_3 K_2^{-1})$, $M_2 := -c_1 \log c_2' \log(c_3 K_2^{-1})$, and M_1 is determined by the implicit constant in (2.24). We observe that, as $\varkappa \to 0^+$, $K_2' \to 0^-$ and $K_1 \to -c < 0$. Therefore, taking the parameter $\varkappa > 0$ sufficiently small, we may arrange that

$$K_2' > K_1. \tag{2.30}$$

Then when N is sufficiently large, (2.29) fails. This shows that either (2.15) or (2.16) fails, and (2.14) holds when N is sufficiently large. Now we have verified the claim (2.12) and completed the proof of Theorem 1.1 under the irreducibility condition.

In order to prove Theorem 1.1 in general, we observe that using the same argument, one can deduce the following more general version of the estimate (2.12): for all sufficiently large N defined in (2.7),

$$\begin{aligned} \frac{1}{|B|} \int_B & f_1(h_1 \beta_1(\exp(v_1 + D\alpha(z_1)b))\Lambda) f_2(h_2 \beta_2(\exp(v_2 + D\alpha(z_2)b))\Lambda)\, db \\ &= \left(\int_X f_1\, d\mu\right)\left(\int_X f_2\, d\mu\right) + O(N^{-\varkappa} \|f_1\|_{C^\theta} \|f_2\|_{C^\theta}) \end{aligned} \tag{2.31}$$

uniformly over the cubes B, $h_1, h_2 \in G$, $v_1, v_2 \in \mathcal{L}(G)$, and automorphisms β_1 and β_2 of G which act trivially on G/G'. Indeed,

$$\begin{aligned} \int_B & f_1(h_1 \beta_1(\exp(v_1 + D\alpha(z_1)b))\Lambda) f_2(h_2 \beta_2(\exp(v_2 + D\alpha(z_2)b))\Lambda)\, db \\ &= \int_B f_1(h_1 \exp(D\beta_1(v_1) + D\beta_1 D\alpha(z_1)b)\Lambda) f_2(h_2 \exp(D\beta_2(v_2) + D\beta_2 D\alpha(z_2)b)\Lambda)\, db, \end{aligned}$$

and to prove (2.31), we can apply Theorem 2.3 to the map

$$\iota : t \longmapsto \left(v_1' + \sum_{i=1}^k t_i D\beta_1 D\alpha(z_1) w_i,\, v_2' + \sum_{i=1}^k t_i D\beta_2 D\alpha(z_2) w_i\right).$$

As in the above proof, either (2.31) holds, or an analogue of (2.16) holds, but since $D\pi D\beta_i = D\pi$, this reduces to exactly the same estimate as (2.16). Therefore, (2.31) follows. Now we combine (2.31) with the argument (2.8)–(2.13) to deduce that

$$\begin{aligned} \int_X & f_0(x) f_1(h_1 \beta_1(\alpha(z_1)(x))) f_2(h_2 \beta_2(\alpha(z_2)(x)))\, d\mu(x) \\ &= \left(\int_X f_0\, d\mu\right)\left(\int_X f_1\, d\mu\right)\left(\int_X f_2\, d\mu\right) + O(N^{-\varkappa\theta} \|f_0\|_{C^\theta} \|f_1\|_{C^\theta} \|f_2\|_{C^\theta}) \end{aligned} \tag{2.32}$$

uniformly over $h_1, h_2 \in G$ and automorphisms β_1 and β_2 of G that preserve Λ and act trivially on G/G'. We will use this estimate to establish Theorem 1.1 in general.

2.6. 3-mixing in general

Let W be the Lie subalgebra of $\mathcal{L}(G)$ introduced in §2.5. By [23, Chapter 5, §5], there exists a closed connected normal subgroup M of G such that $M/(M\cap\Lambda)$ is compact, and

$$\overline{\exp(W)g\Lambda}=Mg\Lambda \quad \text{for almost every } g\in G.$$

Since we may replace the lattice Λ by its conjugate, we assume that

$$\overline{\exp(W)\Lambda}=M\Lambda.$$

We note that the group M satisfies the following properties:

(i) M is $\alpha(\mathbb{Z}^l)$-invariant;

(ii) $D\pi(W)$ is not contained in a proper rational subspace of $\mathcal{L}(M/M')$;

(iii) $[G,M]\subset M'$.

Properties (i)–(iii) can be verified exactly as in the proof of [9, Lemma 3.4].

We give the proof of Theorem 1.1 using induction on the dimension of X. For this, we use that $X=G/\Lambda$ fibers over $Y=G/M\Lambda$ with fibers isomorphic to

$$R=M\Lambda/\Lambda\simeq M/(M\cap\Lambda).$$

The invariant measure on X can be decomposed as

$$\int_X f\,d\mu=\int_Y\int_R f(yr)\,d\mu_R(r)\,d\mu_Y(y), \quad f\in C(X),$$

where μ_Y and μ_R are normalised invariant measure on Y and R, respectively. Since the fibration is $\alpha(\mathbb{Z}^l)$-equivariant (by (i)),

$$\begin{aligned}\int_X & f_0(x)f_1(\alpha(z_1)x)f_2(\alpha(z_2)x)\,d\mu(x)\\ &=\int_Y\left(\int_R f_0(yr)f_1(\alpha(z_1)(y)\alpha(z_1)(r))f_2(\alpha(z_2)(y)\alpha(z_2)(r))\,d\mu_R(r)\right)d\mu_Y(y) \qquad (2.33)\\ &=\int_F\left(\int_R f_0(gr)f_1(\alpha(z_1)(g)\alpha(z_1)(r))f_2(\alpha(z_2)(g)\alpha(z_2)(r))\,d\mu_R(r)\right)dm_F(g),\end{aligned}$$

where $F\subset G$ is a bounded fundamental domain for $G/M\Lambda$, and m_F is the measure on F induced by μ_Y. We shall show that for N defined in (2.7) and some $\eta>0$,

$$\begin{aligned}\int_R & f_0(gr)f_1(\alpha(z_1)(g)\alpha(z_1)(r))f_2(\alpha(z_2)(g)\alpha(z_2)(r))\,d\mu_R(r)\\ &=\left(\int_R f_0(gr)\,d\mu_R(r)\right)\left(\int_R f_1(\alpha(z_1)(g)r)\,d\mu_R(r)\right)\left(\int_R f_2(\alpha(z_2)(g)r)\,d\mu_R(r)\right) \qquad (2.34)\\ &\quad+O(N^{-\eta}\|f_0\|_{C^\theta}\|f_1\|_{C^\theta}\|f_2\|_{C^\theta})\end{aligned}$$

uniformly over $g\in F$.

Suppose that (2.34) holds. Then, combining (2.33) and (2.34), we obtain

$$\begin{aligned}\int_X & f_0(x)f_1(\alpha(z_1)x)f_2(\alpha(z_2)x)\,d\mu(x)\\ &=\int_Y \bar f_0(y)\bar f_1(\alpha(z_1)y)\bar f_2(\alpha(z_2)y)\,d\mu_Y(y)+O(N^{-\eta}\|f_0\|_{C^\theta}\|f_1\|_{C^\theta}\|f_2\|_{C^\theta}),\end{aligned}$$

where the functions $\bar f_i$ on Y are defined by

$$y\longmapsto \int_R f_i(yr)\,d\mu_R(r).$$

Since $\dim(Y)<\dim(X)$, it follows from the inductive assumption that, for some $\eta>0$,

$$\begin{aligned}\int_Y & \bar f_0(y)\bar f_1(\alpha(z_1)y)\bar f_2(\alpha(z_2)y)\,d\mu_Y(y)\\ &=\left(\int_Y \bar f_0\,d\mu_Y\right)\left(\int_Y \bar f_1\,d\mu_Y\right)\left(\int_Y \bar f_2\,d\mu_Y\right)+O(N^{-\eta}\|\bar f_0\|_{C^\theta}\|\bar f_1\|_{C^\theta}\|\bar f_2\|_{C^\theta})\\ &=\left(\int_X f_0\,d\mu\right)\left(\int_X f_1\,d\mu\right)\left(\int_X f_2\,d\mu\right)+O(N^{-\eta}\|f_0\|_{C^\theta}\|f_1\|_{C^\theta}\|f_2\|_{C^\theta}),\end{aligned}$$

and this completes the proof of Theorem 1.1. Hence, it remains to prove (2.34).

To prove (2.34), we write

$$\alpha(z_i)(g)=a_im_i\lambda_i \quad \text{with } a_i\in F,\ m_i\in M \text{ and } \lambda_i\in\Lambda,\ i=1,2.$$

Then

$$\begin{aligned}\int_R & f_0(gr)f_1(\alpha(z_1)(g)\alpha(z_1)(r))f_2(\alpha(z_2)(g)\alpha(z_2)(r))\,d\mu_R(r)\\ &=\int_R f_0(gr)f_1(a_1m_1\beta_1(\alpha(z_1)(r)))f_2(a_2m_2\beta_2(\alpha(z_2)(r)))\,d\mu_R(r),\end{aligned}$$

where the β_i's are the maps induced by the automorphisms $m\mapsto\lambda_im\lambda_i^{-1}$. We observe that because of (ii), $W\subset\mathcal{L}(M)$ satisfies the irreducibility assumption of §2.5, and by (iii), the automorphisms β_i act trivially on M/M'. Hence, (2.32) holds. We apply (2.32) to the functions on R defined by

$$\phi_0(r):=f_0(gr) \quad\text{and}\quad \phi_i(r):=f_i(a_ir),\quad i=1,2.$$

This gives

$$\begin{aligned}\int_R & \phi_0(r)\phi_1(m_1\beta_1(\alpha(z_1)r))\phi_2(m_2\beta_2(\alpha(z_2)r))\,d\mu_R(r)\\ &=\left(\int_R \phi_0\,d\mu_R\right)\left(\int_R \phi_1\,d\mu_R\right)\left(\int_R \phi_2\,d\mu_R\right)+O(N^{-\eta}\|\phi_0\|_{C^\theta}\|\phi_1\|_{C^\theta}\|\phi_2\|_{C^\theta})\\ &=\left(\int_R f_0(gr)\,d\mu_R(r)\right)\left(\int_R f_1(\alpha(z_1)(g)r)\,d\mu_R(r)\right)\left(\int_R f_2(\alpha(z_2)(g)r)\,d\mu_R(r)\right)\\ &\quad +O(N^{-\eta}\|f_0\|_{C^\theta}\|f_1\|_{C^\theta}\|f_2\|_{C^\theta}).\end{aligned}$$

This implies (2.34) and completes the proof of Theorem 1.1.

3. Higher-order mixing

The aim of this section is to prove Theorem 1.2. We shall use the notation introduced in §2.1. In §3.1 we prepare Diophantine estimates. Then in §3.2 we give a proof of Theorem 1.2 under an irreducibility condition, and in §3.3 we give a proof in general using an inductive argument.

We note that it is sufficient to prove Theorem 1.2 for a collection of functions $f_i\in L^\infty(X)$ which is dense in $L^1(X)$. Hence, we may assume that $f_0, ..., f_s\in C^\theta(X)$. Furthermore, we may assume that $z_0=0$.

3.1. Diophantine estimates

Let K be a number field and S be a finite set of places of K containing all the archimedean places. We denote by U_S the ring of S-units, namely, the group of elements x in K such that $|x|_v=1$ for $v\notin S$. For a vector $\bar{x}\in K^s$, we define its (relative) height by

$$H(\bar{x})=\prod_v \max\{1, \|\bar{x}\|_v\},$$

where v runs the set of all places of K, and $\|\bar{x}\|_v=\max_i |x_i|_v$.

Proposition 3.1. *Let $v\in S$ and $b_1, ... b_s\in K\setminus\{0\}$. Then for every $\varepsilon>0$, the inequality*

$$\left|b_1+\sum_{j=2}^s b_j x_j\right|_v < H(\bar{x})^{-\varepsilon} \tag{3.1}$$

has finitely many solutions $\bar{x}\in U_S$ such that no proper subsum of $b_1+\sum_{j=2}^s b_j x_j$ vanishes.

We call such solutions of (3.1) *non-degenerate.*

We give a simple proof of the proposition which is based on the classical W. Schmidt subspace theorem. We note that this proposition is closely related to results about finiteness of the number of solutions of S-unit equations. For S-unit equations the number of non-degenerate solutions can be estimated explicitly. For instance, we refer to a remarkably uniform bound in [7]. Since explicit bounds on the number of solutions do not play any role in our arguments, we do not pursue this direction here.

Proof. We prove the proposition by induction on s. Note that, when $s=1$, the statement holds trivially because there are only finitely many solutions of $H(\bar{x})<c$.

Given a solution $\bar{x}$ of (3.1), we set $\bar{y}=(1, \bar{x})$, and we denote by $j_v=j_v(\bar{x})$ the first index j_v such that

$$|y_{j_v}|_v=\|\bar{y}\|_v\geqslant 1.$$

Partitioning the set of solutions according to the index j_v, we may assume that this index is fixed.

We introduce a family of linear forms $L_{wj}(\bar{y})$, with $w\in S$ and $j=1,...,s$, defined by

$$L_{wj}(\bar{y}) = \begin{cases} y_j, & \text{if } (w,j)\neq(v,j_v), \\ \sum_{j=1}^{s} b_j y_j, & \text{if } (w,j)=(v,j_v). \end{cases}$$

Then, if $\bar{y}=(1,\bar{x})$ corresponds to a solution of (3.1),

$$\prod_{j=1}^{s} |L_{wj}(\bar{y})|_w = \prod_{j=1}^{s} |y_j|_w, \quad w\neq v,$$
$$\prod_{j=1}^{s} |L_{vj}(\bar{y})|_v = |L_{vj_v}(\bar{y})|_v \prod_{j\neq j_v} |y_j|_v < H(\bar{y})^{-\varepsilon} \prod_{j=1}^{s} |y_j|_v,$$

and, by the product formula,

$$\prod_{w\in S} \prod_{j=1}^{s} |L_{wj}(\bar{y})|_w < H(\bar{y})^{-\varepsilon}. \tag{3.2}$$

By the W. Schmidt subspace theorem [1, Corollary 7.2.5], all the solutions of (3.2) are contained in a finite union of proper linear subspaces of K^s. Partitioning solutions of (3.1) according to these subspaces, we may assume that these solutions additionally satisfy a non-trivial linear relation

$$c_1 + \sum_{j=2}^{s} c_j x_j = 0 \tag{3.3}$$

with $c_1, ..., c_s \in K$.

Suppose that $c_1\neq 0$. Given a solution $\bar{x}$ of (3.3), we pick a minimal $J\subset\{2,...,s\}$ such that

$$c_1 + \sum_{j\in J} c_j x_j = 0. \tag{3.4}$$

Then no proper subsum in (3.4) vanishes. It follows from the finiteness of the number of non-degenerate solutions of unit equations [1, Theorem 7.4.2] that x_j, $j\in J$, varies over a finite set. This shows that for every solution $\bar{x}$ of (3.3) there exists $j_0=2,...,s$ such that x_{j_0} belongs to a fixed finite set. Hence, in order to prove finiteness of non-degenerate solutions (3.1), we may assume that x_{j_0} is fixed. Then (3.1) becomes

$$\left| b_1 + b_{j_0} x_{j_0} + \sum_{j\neq j_0} b_j x_j \right|_v < H(\bar{x})^{-\varepsilon}. \tag{3.5}$$

Since we are assuming that no proper subsum in (3.1) vanishes, $b_1+b_{j_0}x_{j_0}\neq 0$ and no proper subsum in (3.5) vanishes either. Let $\bar{x}'=(x_j:j\neq j_0)$. Then $H(\bar{x}')\leqslant H(\bar{x})$. Hence, by the inductive assumption, the number of non-degenerate solutions $\bar{x}'$ of (3.5) is finite, and this implies the proposition in this case.

Now suppose that $c_1=0$ in (3.3). One of $c_2, ..., c_s$ is non-zero, and for simplicity, we assume that $c_s\neq 0$. Then combining (3.1) with (3.3), we obtain that

$$\left| b_1+\sum_{j=2}^{s-1}(b_j-c_jb_sc_s^{-1})x_j \right|_v < H(\bar{x})^{-\varepsilon}. \tag{3.6}$$

Given a solution $\bar{x}$ of (3.6), we pick a minimal $J\subset\{2, ..., s-1\}$ such that

$$\left| b_1+\sum_{j\in J}(b_j-c_jb_sc_s^{-1})x_j \right|_v < H(\bar{x})^{-\varepsilon}, \tag{3.7}$$

and no proper subsum of $b_1+\sum_{j\in J}(b_j-c_jb_sc_s^{-1})x_j$ vanishes. Let $\bar{x}'=(x_j:j\in J)$. Since $H(\bar{x}')\leqslant H(\bar{x})$, it follows from the inductive hypothesis that $\bar{x}'$ belongs to a fixed finite set. This proves that for every solution $\bar{x}$ of (3.1) there exists $j_0=2, ..., s$ such that x_{j_0} belongs to a fixed finite set. Now we can finish the argument as in the previous paragraph, and this completes the proof of the proposition. □

3.2. Higher-order mixing under irreducibility condition

We define the subspace W in $\mathcal{L}(G)$, the set of characters $\mathcal{X}_0$ and the eigenvectors w_χ with $\chi\in\mathcal{X}_0$ as in §2.5.

In this section we assume that $D\pi(W)$ is not contained in any proper rational subspace. Let $\{w_1, ..., w_k\}$ be a fixed basis of W. Consider a box map

$$\begin{aligned} \iota: B &\longrightarrow \mathcal{L}(G), \\ t &\longmapsto \sum_{i=1}^{k} t_iw_i, \end{aligned}$$

where $B=[0,T_1]\times...\times[0,T_k]$.

Lemma 3.2. *Let $f_1, ..., f_s\in C(X)$, $u_1, ..., u_s\in\mathcal{L}(G)$, $\beta_1, ..., \beta_s$ be automorphisms of G such that $\beta_i=\mathrm{id}$ of G/G', $z_1, ... z_s\in\mathbb{Z}^l$, $v_1, ..., v_s\in\mathcal{L}(G)$, and $x_1, ..., x_s\in X$. Then*

$$\frac{1}{|B|}\int_B\left(\prod_{i=1}^{s} f_i(\exp(u_i)\beta_i(\alpha(z_i)(\exp(v_i+\iota(t))))x_i)\right)dt=\prod_{i=1}^{s}\int_X f_i\,d\mu+o(1)$$

as $\min\{\|z_i\|, \|z_i-z_j\|:i\neq j\}\to\infty$. Moreover, the convergence is uniform over $u_1, ..., u_s$, $\beta_1, ..., \beta_s$, $v_1, ..., v_s$, $x_1, ... x_s$, and functions $f_1, ..., f_s$ with $\|f_i\|_{C^\theta}\ll 1$.

Proof. It is sufficient to prove the claim when $f_1, ..., f_s$ belong to a dense family of functions in $C(X)$. Hence, without loss of generality, we may assume that the functions are Hölder with exponent θ.

To prove the lemma, we apply Theorem 2.3 to the product nilmanifold $X^s = G^s/\Lambda^s$. Suppose that the claim of the lemma fails. Then there exist $\delta \in (0, \delta_0)$ and sequences $f_1^{(n)}, ..., f_s^{(n)} \in C^\theta(X)$, $u_1^{(n)}, ..., u_s^{(n)} \in \mathcal{L}(G)$, $\beta_1^{(n)}, ..., \beta_s^{(n)}$ which satisfy $\beta_i^{(n)} = \mathrm{id}$ on G/G', $z_1^{(n)}, ..., z_s^{(n)} \in \mathbb{Z}^l$ satisfying

$$\min\{\|z_i^{(n)}\|, \|z_i^{(n)} - z_j^{(n)}\| : i \neq j\} \to \infty \quad \text{as } n \to \infty,$$

$v_1^{(n)}, ..., v_s^{(n)} \in \mathcal{L}(G)$, and $x_1^{(n)}, ..., x_s^{(n)} \in X$ such that

$$\begin{aligned} &\left| \frac{1}{|B|} \int_B \left(\prod_{i=1}^s f_i^{(n)}(\exp(u_i^{(n)}) \beta_i^{(n)}(\alpha(z_i^{(n)})(\exp(v_i^{(n)} + \iota(t)))) x_i^{(n)}) \right) dt - \prod_{i=1}^s \int_X f_i^{(n)}\, d\mu \right| \\ &\qquad > \delta \prod_{i=1}^s \|f_i^{(n)}\|_{C^\theta}. \end{aligned}$$

We set $f^{(n)} = f_1^{(n)} \otimes ... \otimes f_s^{(n)} : X^s \to \mathbb{R}$, $u^{(n)} = (u_1^{(n)}, ..., u_s^{(n)}) \in \mathcal{L}(G)^s$,

$$\begin{aligned} \iota^{(n)} &: B \longrightarrow \mathcal{L}(G)^s, \\ & \quad t \longmapsto (D\beta_1^{(n)} D\alpha(z_1^{(n)})(v_1^{(n)} + \iota(t)), ..., D\beta_s^{(n)} D\alpha(z_s^{(n)})(v_s^{(n)} + \iota(t))), \end{aligned}$$

and $x^{(n)} = (x_1^{(n)}, ..., x_s^{(n)}) \in X^s$. Then

$$\|f^{(n)}\|_{C^\theta} \ll \prod_{i=1}^s \|f_i^{(n)}\|_{C^\theta}$$

and

$$\left| \frac{1}{|B|} \int_B f^{(n)}(\exp(u^{(n)}) \exp(\iota^{(n)}(t)) x^{(n)})\, dt - \int_X f^{(n)}\, d\mu^{\otimes s} \right| \gg \delta \|f^{(n)}\|_{C^\theta}.$$

It follows from Theorem 2.3 that there exists $(a_1^{(n)}, ..., a_s^{(n)}) \in (\mathbb{Z}^d)^s \setminus \{0\}$ such that

$$\|a_1^{(n)}\|, ..., \|a_s^{(n)}\| \ll \delta^{-L_1} \ll 1 \tag{3.8}$$

and

$$\left| \sum_{j=1}^s \langle a_j^{(n)}, D\pi D\beta_j^{(n)} D\alpha(z_j^{(n)})(w_i) \rangle \right| \ll \frac{\delta^{-L_2}}{T_i} \ll 1 \quad \text{for all } i = 1, ..., k. \tag{3.9}$$

Since $\beta_j^{(n)}=\mathrm{id}$ on G/G', we have $D\pi D\beta_j^{(n)}=D\pi$. We rewrite (3.9) in terms of vectors w_χ, $\chi\in\mathcal{X}_0$, that satisfy $D\alpha(z)w_\chi=\chi(z)w_\chi$ for $z\in\mathbb{Z}^l$. Since each w_χ can be written as a linear combination of the w_i's, it follows from (3.9) that

$$\left|\sum_{j=1}^{s}\chi(z_j^{(n)})\langle a_j^{(n)}, D\pi(w_\chi)\rangle\right| \ll 1 \quad \text{for all } \chi\in\mathcal{X}_0. \tag{3.10}$$

We observe that, because of (3.8), the tuple $(a_1^{(n)},...,a_s^{(n)})$ varies over a finite set. Hence, passing to a subsequence, we may assume that (3.10) holds for a fixed tuple $(a_1,...,a_s)\in(\mathbb{Z}^d)^s\setminus\{0\}$. After changing indices, we may assume that $a_j\neq 0$ for $j=1,...,s'$, and $a_j=0$ for $j>s'$. We note that this implies that

$$b_j := \langle a_j, D\pi(w_\chi)\rangle \neq 0 \quad \text{for all } j=1,...,s' \text{ and } \chi\in\mathcal{X}_0.$$

Indeed, if $\langle a_j, D\pi(w_\chi)\rangle=0$ for some j and χ, then taking Galois conjugates we obtain that $\langle a_j, D\pi(w_\chi)\rangle=0$ for all $\chi\in\mathcal{X}_0$. This implies that $D\pi(W)$ is contained in a proper rational subspace and contradicts the irreducibility assumption.

We may cancel vanishing subsums from (3.10), and passing to a subsequence, we may assume that no proper subsum in (3.10) vanishes.

Passing to a subsequence and changing indices, we may also assume that

$$\max_{j=1,\dots,s'}\|z_j^{(n)}\| = \|z_1^{(n)}\|. \tag{3.11}$$

By Lemma 2.1, there exists fixed $c>0$ such that

$$\max_{\chi\in\mathcal{X}_0}|\chi(z)| \geqslant \exp(c\|z\|), \quad z\in\mathbb{Z}^l. \tag{3.12}$$

Hence, passing to a subsequence, we may assume that

$$|\chi_0(z_1^{(n)})| \geqslant \exp(c\|z_1^{(n)}\|)$$

holds with a fixed $\chi_0\in\mathcal{X}_0$. For this χ_0, (3.10) gives

$$\left|b_0+\sum_{j=1}^{s'} b_j x_j^{(n)}\right| \ll \exp(-c\|z_1^{(n)}\|), \tag{3.13}$$

where $x_j^{(n)}:=\chi_0(z_j^{(n)}-z_1^{(n)})$. It is clear that the b_j's and $x_j^{(n)}$'s are S-units in a fixed number field, and to derive a contradiction, we apply the estimate of Proposition 3.1. We observe that there exists $c_v>1$, $v\in S$, such that

$$|\chi_0(z)|_v \leqslant \exp(c_v\|z\|), \quad z\in\mathbb{Z}^l.$$

Hence,

$$|x_j^{(n)}|_v=|\chi_0(z_j^{(n)}-z_1^{(n)})|_v\leqslant \exp(c_v\|z_j^{(n)}-z_1^{(n)}\|),$$

and, by (3.11),

$$\begin{aligned}H(\bar{x}^{(n)})&=\prod_{v\in S}\max\{1,|x_1^{(n)}|_v,...,|x_{s'}^{(n)}|_v\}\\&\leqslant \exp\biggl(\biggl(\sum_{v\in S}c_v\biggr)\max_{1\leqslant j\leqslant s'}\|z_j^{(n)}-z_1^{(n)}\|\biggr)\leqslant \exp\biggl(2\biggl(\sum_{v\in S}c_v\biggr)\|z_1^{(n)}\|\biggr).\end{aligned}$$

It follows from (3.13) that

$$\left|b_0+\sum_{j=1}^{s'}b_jx_j^{(n)}\right|\ll H(\bar{x}^{(n)})^{-\varepsilon}, \tag{3.14}$$

with fixed $\varepsilon>0$. According to our construction, no proper subsum in (3.14) vanishes. Hence, it follows from Proposition 3.1 that $x^{(n)}=\chi_0(z_j^{(n)}-z_1^{(n)})$ runs over a finite set. As all elements in $\mathcal{X}_0$ are conjugate under the Galois action, it follows that $\chi(z_j^{(n)}-z_1^{(n)})$, with $\chi\in\mathcal{X}_0$, also runs over a finite set. In particular,

$$\max_{\chi\in\mathcal{X}_0}|\chi(z_j^{(n)}-z_1^{(n)})|\ll 1.$$

On the other hand, by (3.12),

$$\max_{\chi\in\mathcal{X}_0}|\chi(z_j^{(n)}-z_1^{(n)})|\to\infty.$$

This contradiction proves the lemma. □

Now we prove Theorem 1.2 under the irreducibility condition. Without loss of generality, we may assume that $z_0=0$. We fix a fundamental domain $F\subset G$ for G/Λ and set $E=\exp^{-1}(F)$. We may arrange that E is bounded and has piecewise smooth boundary. Then

$$\int_X f_0(x)\biggl(\prod_{i=1}^s f_i(\alpha(z_i)(x))\biggr)\,d\mu(x)=\int_E f_0(\exp(u)\Lambda)\biggl(\prod_{i=1}^s f_i(\exp(D\alpha(z_i)u)\Lambda)\biggr)\,du,$$

where du denotes a suitably normalised Lebesgue measure on $\mathcal{L}(G)$. We choose a basis of $\mathcal{L}(G)$ that contains the fixed basis $\{w_i\}_i$ of W and tessellate $\mathcal{L}(G)$ by cubes C of size ε with respect to this basis. Then

$$\left|E\setminus\bigcup_{C\subset E}C\right|\ll\varepsilon. \tag{3.15}$$

For all $u_1, u_2 \in C$,

$$|f_0(\exp(u_1)\Lambda)-f_0(\exp(u_2)\Lambda)| \ll \varepsilon^\theta. \tag{3.16}$$

Here and later in the argument the implied constants may depend on the Hölder norms of $f_0, ..., f_s$. For every cube C, we pick a point $u_C \in C$. Then it follows from (3.15) and (3.16) that

$$\begin{aligned}
&\int_E f_0(\exp(u)\Lambda)\left(\prod_{i=1}^s f_i(\exp(D\alpha(z_i)u)\Lambda)\right) du \\
&\quad = \sum_{C\subset E} \int_C f_0(\exp(u)\Lambda)\left(\prod_{i=1}^s f_i(\exp(D\alpha(z_i)u)\Lambda)\right) du + O(\varepsilon) \\
&\quad = \sum_{C\subset E} f_0(\exp(u_C)\Lambda) \int_C \left(\prod_{i=1}^s f_i(\exp(D\alpha(z_i)u)\Lambda)\right) du + O(\varepsilon^\theta).
\end{aligned} \tag{3.17}$$

Each cube C in the above sum can be written as $C=B'+B$, where B is a cube in W and B' is a cube in the complementary subspace.

Let

$$N = N(z_1, ..., z_s) := \min\{\|z_i\|, \|z_i - z_j\| : i \neq j\}.$$

It follows from Lemma 3.2 that

$$\frac{1}{|B|}\int_B \left(\prod_{i=1}^s f_i(\exp(v_i + D\alpha(z_i)(b))\Lambda)\right) db \to \prod_{i=1}^s \int_X f_i \, d\mu \tag{3.18}$$

as $N\to\infty$, uniformly over $v_1, ... v_s \in \mathcal{L}(G)$ and the cubes B (note that all the cubes are translates of a fixed cube). Hence, it follows that

$$\begin{aligned}
&\frac{1}{|C|}\int_C \left(\prod_{i=1}^s f_i(\exp(D\alpha(z_i)(u))\Lambda)\right) du \\
&\quad = \frac{1}{|B'|\,|B|}\int_{B'}\int_B \left(\prod_{i=1}^s f_i(\exp(D\alpha(z_i)(b') + D\alpha(z_i)(b))\Lambda)\right) db\, db' \to \prod_{i=1}^s \int_X f_i\, d\mu
\end{aligned}$$

as $N\to\infty$. Combining this with (3.17), we deduce that

$$\begin{aligned}
&\int_E f_0(\exp(u)\Lambda)\left(\prod_{i=1}^s f_i(\exp(D\alpha(z_i)u)\Lambda)\right) du \\
&\quad = \left(\sum_{C\subset E} f_0(\exp(u_C)\Lambda)|C|\right)\prod_{i=1}^s \int_X f_i\, d\mu + \left(\sum_{C\subset E} |C|\right) o(1) + O(\varepsilon^\theta),
\end{aligned}$$

where clearly $\sum_{C\subset E}|C|=O(1)$. Moreover, using (3.16) and (3.15), we deduce that

$$\begin{aligned}\sum_{C\subset E} f_0(\exp(u_C)\Lambda)|C| &= \sum_{C\subset E}\left(\int_C f_0(\exp(u)\Lambda)\,du+O(|C|\varepsilon^\theta)\right)\\ &= \int_E f_0(\exp(u)\Lambda)\,du+O(\varepsilon^\theta)=\int_X f_0\,d\mu+O(\varepsilon^\theta).\end{aligned}$$

This implies that

$$\begin{aligned}\int_X f_0(x)\left(\prod_{i=1}^s f_i(\alpha(z_i)(x))\right)d\mu(x) &= \int_E f_0(\exp(u)\Lambda)\left(\prod_{i=1}^s f_i(\exp(D\alpha(z_i)u)\Lambda)\right)du\\ &= \prod_{i=0}^s\int_X f_i\,d\mu+o(1)+O(\varepsilon^\theta)\end{aligned}$$

as $N\to\infty$. This proves Theorem 1.2 under the irreducibility condition. It is clear from the proof that convergence is uniform provided that $\|f_0\|_{C^\theta},...,\|f_s\|_{C^\theta}\ll 1$.

3.3. Higher-order mixing in general

We will apply an inductive argument which uses the result of §3.2 as a base case. In fact, we note that the argument in §3.2 implies that

$$\int_X f_0(x)\left(\prod_{i=1}^s f_i(h_i\beta_i(\alpha(z_i)(x)))\right)d\mu(x)=\prod_{i=0}^s\int_X f_i\,d\mu+o(1) \tag{3.19}$$

as $N=N(z_1,...,z_s)\to\infty$, uniformly over functions $f_0,...,f_s$ with $\|f_i\|_{C^\theta}\ll 1$, $h_1,...,h_s\in G$ and automorphisms $\beta_1,...,\beta_s$ of G that preserve Λ and act trivially on G/G'. Indeed, Lemma 3.2 implies that in (3.18) we more generally have

$$\frac{1}{|B|}\int_B\left(\prod_{i=1}^s f_i(h_i\beta_i(\exp(v_i+D\alpha(z_i)(b)))\Lambda)\right)db=\prod_{i=1}^s\int_X f_i\,d\mu+o(1)$$

as $N\to\infty$, uniformly over $f_0,...,f_s$ with $\|f_i\|_{C^\theta}\ll 1$, $h_1,...,h_s\in G$, automorphisms $\beta_1,...,\beta_s$, and $v_1,...\,v_s\in\mathcal{L}(G)$. Then the rest of the argument carries over and implies (3.19).

Let W be the Lie subalgebra of $\mathcal{L}(G)$ introduced in §2.5. By [23, §5.5], there exists a closed connected normal subgroup M of G such that $M/(M\cap\Lambda)$ is compact, and

$$\overline{\exp(W)g\Lambda}=Mg\Lambda\quad\text{for almost every } g\in G.$$

Since we may replace the lattice Λ by its conjugate, we assume that

$$\overline{\exp(W)\Lambda}=M\Lambda.$$

We note that the group M satisfies the following properties:

(i) M is $\alpha(\mathbb{Z}^l)$-invariant;

(ii) $D\pi(W)$ is not contained in a proper rational subspace of $\mathcal{L}(M/M')$;

(iii) $[G,M]\subset M'$.

Properties (i)–(iii) can be verified exactly as in the proof of [9, Lemma 3.4].

To apply induction, we observe that the nilmanifold $X=G/\Lambda$ fibers over the nilmanifold of $Y=G/M\Lambda$ with fibers isomorphic to $R=M\Lambda/\Lambda\simeq M/(M\cap\Lambda)$. The invariant measure μ on X decomposes as

$$\int_X f\,d\mu=\int_Y\int_R f(yr)\,d\mu_R(r)\,d\mu_Y(y),\quad f\in C(X),$$

where μ_Y and μ_R denote the normalised invariant measures on Y and R, respectively. It follows from (i) that the fibration $X\to Y$ is $\alpha(\mathbb{Z}^l)$-equivariant. Hence, we obtain

$$\begin{aligned}&\int_X f_0(x)\biggl(\prod_{i=1}^s f_i(\alpha(z_i)(x))\biggr)\,d\mu(x)\\&\quad=\int_Y\biggl(\int_R f_0(yr)\biggl(\prod_{i=1}^s f_i(\alpha(z_i)(y)\alpha(z_i)(r))\biggr)\,d\mu_R(r)\biggr)\,d\mu_Y(y)\qquad(3.20)\\&\quad=\int_F\biggl(\int_R f_0(gr)\biggl(\prod_{i=1}^s f_i(\alpha(z_i)(g)\alpha(z_i)(r))\biggr)\,d\mu_R(r)\biggr)dm_F(g),\end{aligned}$$

where $F\subset G$ is a bounded fundamental set for $G/M\Lambda$, and m_F is the measure on F induced by μ_Y. We write

$$\alpha(z_i)(g)=a_im_i\lambda_i\quad\text{with } a_i\in F,\ m_i\in M\text{ and }\lambda_i\in\Lambda,\ i=1,...,s.$$

Then

$$\begin{aligned}&\int_R f_0(gr)\biggl(\prod_{i=1}^s f_i(\alpha(z_i)(g)\alpha(z_i)(r))\biggr)\,d\mu_R(r)\\&\qquad=\int_R f_0(gr)\biggl(\prod_{i=1}^s f_i(a_im_i\beta_i(\alpha(z_i)(r)))\biggr)\,d\mu_R(r),\end{aligned}$$

where the β_i's are the transformations of S induced by the automorphisms $m\mapsto\lambda_im\lambda_i^{-1}$ of M. By (ii), the subspace $W\subset\mathcal{L}(M)$ satisfies the irreducibility assumption of §3.2, and by (iii), the automorphisms β_i act trivially on M/M'. Let

$$\phi_0(r):=f_0(gr)\quad\text{and}\quad\phi_i(r):=f_i(a_ir),\ \ i=1,...,s.$$

Since $F \subset G$ is bounded, we have $\|\phi_i\|_{C^\theta} \ll \|f_i\|_{C^\theta} \ll 1$. Hence, it follows from (3.19) that

$$\int_R \phi_0(r) \left(\prod_{i=1}^s \phi_i(m_i \beta_i(\alpha(z_i)(r))) \right) d\mu_R(r) = \prod_{i=0}^s \int_R \phi_i \, d\mu_R + o(1)$$

as $N \to \infty$, uniformly over $g \in F$, $m_1, ..., m_s \in M$, and the automorphisms $\beta_1, ..., \beta_s$. Since $a_i M\Lambda = \alpha(z_i)(g) M\Lambda$, this implies that

$$\int_R f_0(gr) \left(\prod_{i=1}^s f_i(\alpha(z_i)(g)\alpha(z_i)(r)) \right) d\mu_R(r) = \prod_{i=0}^s \int_R f_i(\alpha(z_i)(g)r) \, d\mu_R(r) + o(1) \quad (3.21)$$

as $N \to \infty$, uniformly over $g \in F$. Let $\bar{f}_i$ be the function on Y defined by

$$y \longmapsto \int_R f_i(yr) \, d\mu_R(r).$$

Combining (3.20) and (3.21), we deduce that

$$\int_X f_0(x) \left(\prod_{i=1}^s f_i(\alpha(z_i)(x)) \right) d\mu(x) = \int_Y \bar{f}_0(y) \left(\prod_{i=1}^s \bar{f}_i(\alpha(z_i)(y)) \right) d\mu_Y(y) + o(1)$$

as $N \to \infty$. Finally, it follows by induction on $\dim(X)$ that

$$\int_Y \bar{f}_0(y) \left(\prod_{i=1}^s \bar{f}_i(\alpha(z_i)(y)) \right) d\mu_Y(y) = \prod_{i=0}^s \int_Y \bar{f}_i \, d\mu_Y + o(1) = \prod_{i=0}^s \int_X f_i \, d\mu + o(1).$$

The above argument implies uniform convergence provided that $\|f_0\|_{C^\theta}, ..., \|f_s\|_{C^\theta} \ll 1$. This completes the proof of Theorem 1.2.

4. Exponential mixing of shapes

While we have proved exponential 2-mixing and 3-mixing for $\mathbb{Z}^l$-actions by automorphisms on nilmanifolds, we do not know if exponential mixing of higher orders holds for them in general. This would require a quantitative version of Proposition 3.1 which seems to be out of reach of available number-theoretic methods. Nonetheless, we prove a weak form of exponential mixing where the error term is controlled by

$$N_*(z_0, ..., z_s) := \min_{\chi \in \mathcal{X}(\alpha)} \min_{\substack{|\chi(z_i - z_j)| \geqslant 1 \\ i \neq j}} |\chi(z_i - z_j)|$$

with notation as in §2.1. This, in particular, implies exponential mixing for Anosov shapes—Theorem 1.3.

THEOREM 4.1. *Let* $\alpha\colon\mathbb{Z}^l\to\mathrm{Aut}(X)$ *be an action on a compact nilmanifold* X *such that every* $\alpha(z)$, $z\neq 0$, *is ergodic. Then there exists* $\eta=\eta(\theta)>0$ *such that, for every* $f_0, ..., f_s\in C^\theta(X)$ *and* $z_0, ..., z_s\in\mathbb{Z}^l$,

$$\int_X\left(\prod_{i=0}^s f_i(\alpha(z_i)x)\right)d\mu(x)=\prod_{i=0}^s\int_X f_i\,d\mu+O\left(N_*(z_0, ..., z_s)^{-\eta}\prod_{i=0}^s\|f_i\|_{C^\theta}\right). \tag{4.1}$$

Proof. We adapt the method of the proof of [9, Theorem 1.2] from our previous paper. Since the proof is quite similar to the argument in this paper, we will only give an outline.

We take a character $\chi\in\mathcal{X}(\alpha)$ and the corresponding eigenvector $w\in\mathcal{L}(G)$. If χ is real, we denote by W the corresponding 1-dimensional eigenspace. Otherwise, we denote by W the 2-dimensional subspace $\langle w,\bar{w}\rangle\cap\mathcal{L}(G)$. Then $D\alpha(z)|_W=r(z)\omega(z)$, where $r(z)=|\chi(z)|$ and $\omega(z)$ is a rotation. We assume in addition that $\chi^2\notin\mathcal{X}(\alpha)$. Then W is closed under the Lie bracket.

We first treat the irreducible case: namely, when $D\pi(W)$ is not contained in a proper rational subspace. Without loss of generality, $N_*>1$. Then for all $i\neq j$, $|\chi(z_i-z_j)|\neq 1$ and after changing indexes, $|\chi(z_0)|<|\chi(z_1)|<...<|\chi(z_s)|$. We may also assume that $z_0=0$. We fix a basis of $\mathcal{L}(G)$ which contains a basis of W and tessellate $\mathcal{L}(G)$ by cubes of size ε with respect to this basis. Then the integral

$$\int_X f_0(x)\left(\prod_{i=1}^s f_i(\alpha(z_i)x)\right)d\mu(x)$$

can be approximated by a sum of the integrals

$$\int_C\left(\prod_{i=1}^s f_i(\exp(D\alpha(z_i)u)\Lambda)\right)du$$

with the error of size $O(\varepsilon^\theta\prod_{i=1}^s\|f\|_{C^\theta})$. Since $C=B'+B$, where B is a cube in W and B' is a cube in the complementary subspace, the above integral can be written as

$$\int_{B'}\int_B\left(\prod_{i=1}^s f_i(\exp(D\alpha(z_i)b'+D\alpha(z_i)b)\Lambda)\right)db\,db'.$$

For every cube B, we take a box map $\iota_B\colon[0,\varepsilon]^{\dim(W)}\to B$ that parameterises B. Then, by [9, Proposition 4.2], there exists $\varkappa>0$ such that

$$\begin{aligned}
&\frac{1}{|B|}\int_B\left(\prod_{i=1}^s f_i(\exp(v_i+D\alpha(z_i)b)\Lambda)\right)db\\
&\quad=\varepsilon^{-\dim(W)}\int_{[0,\varepsilon]^{\dim(W)}}\left(\prod_{i=1}^s f_i(\exp(v_i+r(z_i)\omega(z_i)\iota_B(t))\Lambda)\right)dt\\
&\quad=\prod_{i=1}^s\int_X f_i\,d\mu+O\left(\sigma^{-\varkappa}\prod_{i=1}^s\|f\|_{C^\theta}\right)
\end{aligned}$$

uniformly over $v_1, ..., v_s \in \mathcal{L}(G)$, where $\sigma = \min\{\varepsilon r(z_1), r(z_s)/r(z_{s-1}), ..., r(z_2)/r(z_1)\}$. We note that the Diophantine condition for the box map $t \mapsto \omega(z_i)\iota_B(t)$, which is required in [9, Proposition 4.2], is satisfied because W is spanned by vectors with algebraic coordinates and [1, Theorem 7.3.2] applies. Since this estimate is uniform over the v_i's, it follows that

$$\frac{1}{|C|}\int_C \left(\prod_{i=1}^s f_i(\exp(D\alpha(z_i)u)\Lambda)\right) du = \prod_{i=1}^s \int_X f_i\, d\mu + O\left(\sigma^{-\varkappa} \prod_{i=1}^s \|f\|_{C^\theta}\right),$$

and we deduce that

$$\int_X f_0(x)\left(\prod_{i=1}^s f_i(\alpha(z_i)x)\right) d\mu(x) = \prod_{i=1}^s \int_X f_i\, d\mu + O\left((\varepsilon^\theta + \sigma^{-\varkappa}) \prod_{i=1}^s \|f\|_{C^\theta}\right).$$

We refer to the proof of [9, Theorem 1.2] for details. Choosing $\varepsilon = r(z_1)^{-1/2}$ implies the claim of the theorem in the irreducible case.

To give a proof in general, we use induction on $\dim(X)$. This argument is very similar to §2.6. If $\overline{\exp(W)\Lambda} \neq X$, we consider the α-equivariant fibration $X \to Y$ defined by the closure. The above argument implies that

$$\int_X f_0(x)\left(\prod_{i=1}^s f_i(h_i\beta_i(\alpha(z_i)x))\right) d\mu(x) = \prod_{i=1}^s \int_X f_i\, d\mu + O\left(N_*^{-\eta} \prod_{i=1}^s \|f\|_{C^\theta}\right)$$

uniformly over $h_1, ..., h_s \in G$ and automorphisms $\beta_1, ..., \beta_s \in \operatorname{Aut}(X)$ that act trivially on G/G'. (In fact, this uniformity was part of [9, Proposition 4.2].) Then as in §2.6,

$$\begin{aligned}\int_X f_0(x)\left(\prod_{i=1}^s f_i(\alpha(z_i)x)\right) d\mu(x) \\ = \int_Y \bar{f}_0(x)\left(\prod_{i=1}^s \bar{f}_i(\alpha(z_i)x)\right) d\mu_Y(x) + O\left(N_*^{-\eta} \prod_{i=1}^s \|f\|_{C^\theta}\right),\end{aligned}$$

where $\bar{f}_0, ..., \bar{f}_s \in C^\theta(Y)$. Now the claim follows by induction on dimension. □

Remark 4.2. In the irreducible case of the above proof, we can replace $N_*(z_0, ..., z_s)$ by

$$N'_*(z_0, ..., z_s) := \max_{\chi \in \mathcal{X}'(\alpha)} \min_{\substack{|\chi(z_i - z_j)| \geqslant 1 \\ i \neq j}} |\chi(z_i - z_j)|,$$

which provides a better estimate.

Proof of Theorem 1.3. We note that if in Theorem 4.1 we assume that $\alpha(z_i - z_j)$ are Anosov for all $i \neq j$, then $N_*(z_0, ..., z_s) > 1$ and $N_*(nz_0, ..., nz_s) = N_*(z_0, ..., z_s)^n$. Hence, Theorem 1.3 follows directly from Theorem 4.1. □

5. Cocycle rigidity

We now apply exponential 2-mixing to prove smooth cocycle rigidity for genuinely higher-rank abelian actions by automorphisms of nilmanifolds—Theorem 1.5. The proof is based on the "higher-rank trick" from [11].

Let $\mathbb{Z}^l \to \mathrm{Aut}(X)$ be an action on a compact nilmanifold X and $c: \mathbb{Z}^l \times X \to \mathbb{R}$ be a cocycle. Assume that there is a rank-2 subgroup $\langle a, b\rangle$ of $\mathbb{Z}^l$ such that all its non-zero elements act ergodically on X.

First, we note that the map $c_0: \mathbb{Z}^l \mapsto \mathbb{R}$ defined by $z \mapsto \int_X c(z,x)\, d\mu(x)$ is a homomorphism by the cocycle property. Then $c - c_0$ is also a cocycle, and it will suffice to prove the theorem for $c - c_0$. Thus, we will assume that all functions $c(z, \cdot)$ for $z \in \mathbb{Z}^l$ satisfy $\int_X c(z,x)\, d\mu(x) = 0$.

Let $f(x) = c(a, x)$. We shall show that there exists $\phi \in L^2(X)$ such that

$$f = \phi \circ a - \phi. \tag{5.1}$$

We will apply our previous results [9, §6]. By [9, Theorem 6.1], it suffices to show that

$$\sigma^2 := \int_X f^2\, d\mu + 2\sum_{i=1}^{\infty} \int_X (f \circ a^i) f\, d\mu = \sum_{i=-\infty}^{\infty} \langle f \circ a^i, f\rangle = 0.$$

We note that the assumption of this theorem is verified in [9, §6] using exponential mixing of a. Let $h(x) = c(b^j, x)$. It follows from the cocycle property that

$$f \circ b^j - f = h \circ a - h.$$

Hence,

$$\sum_{i=-n}^{n} (f \circ a^i b^j - f \circ a^i) = \sum_{i=-n}^{n} (h \circ a^{i+1} - h \circ a^i) = h \circ a^{n+1} - h \circ a^{-n},$$

and it follows from exponential mixing that

$$\sigma^2 = \sum_{i=-\infty}^{\infty} \langle f \circ a^i b^j, f\rangle \quad \text{for every } j \in \mathbb{Z}.$$

On the other hand, by the exponential mixing for the group $\langle a, b\rangle$ established in Theorem 1.1,

$$\sum_{i,j \in \mathbb{Z}} \langle f \circ a^i b^j, f\rangle < \infty.$$

This implies that $\sigma^2 = 0$ and proves (5.1).

Now using the cocycle regularity result [9, Theorem 7.1] established in our previous paper, we deduce that (5.1) also has a C^∞ solution. Finally, since a acts ergodically, it follows from [11, Lemma 4.1] that

$$c(z, \cdot) = \phi \circ z - \phi \quad \text{for all } z \in \mathbb{Z}^l.$$

This completes the proof of Theorem 1.5.

References

[1] Bombieri, E. & Gubler, W., *Heights in Diophantine Geometry*. New Mathematical Monographs, 4. Cambridge Univ. Press, Cambridge, 2006.

[2] Corwin, L. J. & Greenleaf, F. P., *Representations of Nilpotent Lie Groups and their Applications*. Part I. *Basic Theory and Examples*. Cambridge Studies in Advanced Mathematics, 18. Cambridge Univ. Press, Cambridge, 1990.

[3] Damjanović, D., Central extensions of simple Lie groups and rigidity of some abelian partially hyperbolic algebraic actions. *J. Mod. Dyn.*, 1 (2007), 665–688.

[4] Damjanović, D. & Katok, A., Periodic cycle functionals and cocycle rigidity for certain partially hyperbolic $\mathbb{R}^k$ actions. *Discrete Contin. Dyn. Syst.*, 13 (2005), 985–1005.

[5] — Local rigidity of partially hyperbolic actions. II: The geometric method and restrictions of Weyl chamber flows on $\mathrm{SL}(n,\mathbb{R})/\Gamma$. *Int. Math. Res. Not. IMRN*, 2011 (2011), 4405–4430.

[6] Einsiedler, M. & Ward, T., Asymptotic geometry of non-mixing sequences. *Ergodic Theory Dynam. Systems*, 23 (2003), 75–85.

[7] Evertse, J.-H., Schlickewei, H. P. & Schmidt, W. M., Linear equations in variables which lie in a multiplicative group. *Ann. of Math.*, 155 (2002), 807–836.

[8] Fisher, D., Kalinin, B. & Spatzier, R., Global rigidity of higher rank Anosov actions on tori and nilmanifolds. *J. Amer. Math. Soc.*, 26 (2013), 167–198.

[9] Gorodnik, A. & Spatzier, R., Exponential mixing of nilmanifold automorphisms. *J. Anal. Math.*, 123 (2014), 355–396.

[10] Green, B. & Tao, T., The quantitative behaviour of polynomial orbits on nilmanifolds. *Ann. of Math.*, 175 (2012), 465–540; erratum in *Ann. of Math.*, 179 (2014), 1175–1183.

[11] Katok, A. & Spatzier, R. J., First cohomology of Anosov actions of higher rank abelian groups and applications to rigidity. *Inst. Hautes Études Sci. Publ. Math.*, 79 (1994), 131–156.

[12] — Subelliptic estimates of polynomial differential operators and applications to rigidity of abelian actions. *Math. Res. Lett.*, 1 (1994), 193–202.

[13] Ledrappier, F., Un champ markovien peut être d'entropie nulle et mélangeant. *C. R. Acad. Sci. Paris Sér. A-B*, 287 (1978), A561–A563.

[14] Lind, D. A., Dynamical properties of quasihyperbolic toral automorphisms. *Ergodic Theory Dynam. Systems*, 2 (1982), 49–68.

[15] Miles, R. & Ward, T., A directional uniformity of periodic point distribution and mixing. *Discrete Contin. Dyn. Syst.*, 30 (2011), 1181–1189.

[16] Mozes, S., Mixing of all orders of Lie groups actions. *Invent. Math.*, 107 (1992), 235–241; erratum in *Invent. Math.*, 119 (1995), 399.

[17] Rauch, J. & Taylor, M., Regularity of functions smooth along foliations, and elliptic regularity. *J. Funct. Anal.*, 225 (2005), 74–93.

[18] Rodriguez Hertz, F. & Wang, Z., Global rigidity of higher rank abelian Anosov algebraic actions. *Invent. Math.*, 198 (2014), 165–209.

[19] Schlickewei, H. P., S-unit equations over number fields. *Invent. Math.*, 102 (1990), 95–107.

[20] Schmidt, K., Mixing automorphisms of compact groups and a theorem by Kurt Mahler. *Pacific J. Math.*, 137 (1989), 371–385.

[21] — *Dynamical Systems of Algebraic Origin*. Progress in Mathematics, 128. Birkhäuser, Basel, 1995.

[22] Schmidt, K. & Ward, T., Mixing automorphisms of compact groups and a theorem of Schlickewei. *Invent. Math.*, 111 (1993), 69–76.

[23] Starkov, A. N., *Dynamical Systems on Homogeneous Spaces*. Translations of Mathematical Monographs, 190. Amer. Math. Soc., Providence, RI, 2000.

[24] Waldschmidt, M., Minorations de combinaisons linéaires de logarithmes de nombres algébriques. *Canad. J. Math.*, 45 (1993), 176–224.

[25] Wang, Z. J., Local rigidity of partially hyperbolic actions. *J. Mod. Dyn.*, 4 (2010), 271–327.

[26] Ward, T., Three results on mixing shapes. *New York J. Math.*, 3A (1997/98), 1–10.

[27] — Mixing and tight polyhedra, in *Dynamics & Stochastics*, IMS Lecture Notes Monogr. Ser., 48, pp. 169–175. Inst. Math. Statist., Beachwood, OH, 2006.

Alexander Gorodnik
School of Mathematics
University of Bristol
Howard House, Queens Ave.
Clifton, Bristol, BS8 1SD
U.K.
a.gorodnik@bristol.ac.uk

Ralf Spatzier
Department of Mathematics
University of Michigan
Ann Arbor, MI 48109
U.S.A.
spatzier@umich.edu

Received May 9, 2013
Received in revised form July 11, 2014