JOURNAL OF MODERN DYNAMICS
VOLUME 5, NO. 2, 2011, 203–235

doi: 10.3934/jmd.2011.5.203

LOCAL RIGIDITY OF HOMOGENEOUS PARABOLIC ACTIONS: I. A MODEL CASE

DANIJELA DAMJANOVIC AND ANATOLE KATOK
(Communicated by Raphaël Krikorian)

ABSTRACT. We show a weak form of local differentiable rigidity for the rank 2 abelian action of upper unipotents on $SL(2,\mathbb{R}) \times SL(2,\mathbb{R})/\Gamma$. Namely, for a 2-parameter family of sufficiently small perturbations of the action, satisfying certain transversality conditions, there exists a parameter for which the perturbation is smoothly conjugate to the action up to an automorphism of the acting group. This weak form of rigidity for the parabolic action in question is optimal since the action lives in a family of dynamically different actions. The method of proof is based on a KAM-type iteration and we discuss in the paper several other potential applications of our approach.

1. INTRODUCTION

1.1. **KAM/Harmonic analysis method in differentiable rigidity.** In this paper we develop an approach for proving local differentiable rigidity of higher-rank abelian groups, *i.e.*, $\mathbb{Z}^k \times \mathbb{R}^l$, $k + l \geq 2$, based on KAM-type iteration scheme that was first introduced in [5, 6]. In those papers local differentiable rigidity was proved for $\mathbb{Z}^k$, $k \geq 2$ actions by partially hyperbolic automorphisms of a torus. Here we deal with certain *nonhyperbolic* actions of $\mathbb{R}^2$ and $\mathbb{R}^3$, namely actions of certain unipotent subgroups of semisimple Lie groups on homogeneous spaces by translations. From the point of view of general classification of dynamical systems those actions are parabolic [10]. Before proceeding to specifics we will describe in general terms the version of the method adapted to the use for parabolic actions.

First, an important feature of the parabolic case, that sets it apart from all hyperbolic and many partially hyperbolic actions, comes from the fact that among *homogeneous* actions on semisimple groups unipotent ones are not stable under perturbation. Unipotent abelian subalgebras of a given dimension k of a semisimple Lie algebra $\mathfrak{g}$ have positive codimension d in the variety of k-dimensional abelian subalgebras of $\mathfrak{g}$. Thus among the nonhomogeneous perturbations of a unipotent $\mathbb{R}^k$ action α those conjugate to a (maybe different) unipotent action are expected to have positive codimension, greater or equal

Received March 30, 2010; revised April 25, 2011.
2000 *Mathematics Subject Classification:* 37C85.
Key words and phrases: Local rigidity, homogeneous actions, KAM method.
DD: Based on research supported by NSF grant DMS-0758555.
AK: Based on research supported by NSF grants DMS-0505539 and DMS-0803880.

than d. A natural way to find such perturbations is to consider d-parametric families $\tilde{\alpha}(\lambda)$ of perturbed $\mathbb{R}^k$ actions and look for an action conjugate to α or its homogeneous perturbation inside such a family. Thus, unlike the absolute rigidity established in [5, 6] (as well as earlier for higher-rank hyperbolic actions in [11]), we will look for conditional rigidity. Let us point out that in the most traditional KAM results dealing with perturbations of translations or linear flows on the torus (elliptic systems in the terminology of [10]) one also has to consider parametric families with the exception of the low-dimensional cases (dimension one for maps and two for flows) when rotation number serves as a modulus of conjugacy.

What is described below is a somewhat idealized scheme that highlights essential features of the method. While dealing with concrete problems, including those considered in this paper, certain steps may be combined and order of other steps may be reversed.

Step 1. Conjugacy equation is formally linearized at the target unperturbed system. Solution of the linearized equation is attempted by finding an inverse on a proper subspace of data. This step is carried out in Section 1.2 in great generality.

Step 2. *Obstructions* for solving the linearized conjugacy equation for a particular element of the action are found.

There are *infinitely many* obstructions for solving the linearized (cohomological) equations for an individual generator of the action. This is a crucial difference with standard KAM for translations where obstruction is *unique*. In some cases (including those considered in this paper) those obstructions are *invariant distributions* for the action element. This has to do with parabolicity of our examples. If some hyperbolicity is present, as in [5, 6], the linearized equations are twisted in some way and obstructions have a somewhat different form.

While obstructions can be described by different means, for our purposes, dual description using harmonic analysis is preferred.

Step 3. All but finitely many obstructions vanish due to the commutation relations. This is sometimes called the *higher-rank trick*:

- At this step harmonic analysis is used crucially.
- Fourier analysis for actions by torus automorphisms.
- Unitary group representations for homogeneous actions.

See Section 1.4 for discussion of this step for our examples. The "philosophy" of the higher-rank trick is universal, its execution is highly case-sensitive.

Step 4. The remaining finitely many parameters are absorbed by allowing

(i) automorphism of the group, or
(ii) a standard perturbation, or
(iii) adjusting parameters.

Accordingly, rigidity comes in different flavors.

Step 5. (a) Linearized equation is solved. This involves glueing of solutions constructed in certain invariant subspaces of functional spaces such as
- Cyclic spaces of characters for the torus.
- Irreducible representation spaces for homogeneous actions.

(b) *Tame estimates* are obtained for the solution. This means finite loss of regularity in the chosen collection of norms in the Fréchet spaces, such as C^r or Sobolev norms.

Step 6. The perturbation can be split into two terms due to the commutation relations: one for which the linearized equations are satisfied, and the other "quadratically small" with tame estimates. This step also uses harmonic analysis.

Step 7. Conjugacy provided by the solution of the linearized equation transforms the (modified in the conditional case) perturbed action into an action quadratically close to the target with a fixed loss of regularity in the estimates.

Step 8. The process is iterated and the product of conjugacies converges to a conjugacy between the (modified) original and perturbed action (with adjusted parameters if necessary).

The last two steps are completely independent of the specific ways previous steps are performed and depend only on the conclusions reached at those steps.

Since we deal with higher-rank actions the basic iteration scheme is somewhat similar to that applied by Moser [14] to commuting rotations of the circle (which is the basis of the scheme used later in [6]).

1.2. **Conjugacy problem and linearization.** Now we describe Step 1 in complete generality. Let $G = G^1 \times G^2 \times \cdots \times G^m$, where G^j, $j = 1, \dots, m$, are simple Lie groups. Let $X = G/\Gamma$ where Γ is an irreducible cocompact lattice in G. Let $\mathfrak{g}$ denote the Lie algebra of G and $\mathfrak{g}^j$ Lie algebras of G^j. Let $U_1, \dots, U_d$ be some commuting elements of $\mathfrak{g}$. We consider an action $\alpha\colon \mathbb{R}^d \times X \to X$ by left translations on X whose generating vector fields are $U_1, \dots, U_d$, namely $U_k = \frac{\partial \alpha(t_1,t_2,x)}{\partial t_k}\big|_{(0,0)}$.

A smooth perturbation $\tilde{\alpha}$ of the action α is generated by commuting vector fields $\tilde{U}_k = U_k + W_k$, $k = 1, \dots, d$, on X which are small perturbations of U_k's.

Given a vector field Y and a diffeomorphism h, set

$$h_* Y(x) := (Dh)_{h^{-1}(x)} Y \circ h^{-1}(x).$$

Define operators $\mathfrak{L}$ and $\mathfrak{M}$ in the following way:

$$\text{Vect}^\infty(X) \xrightarrow{\mathfrak{L}} \text{Vect}^\infty(X)^d \xrightarrow{\mathfrak{M}} \text{Vect}^\infty(X)^{d\times d} \tag{1.1}$$

$$H \xrightarrow{\mathfrak{L}} \big((h_1)_* U_1, \dots, (h_1)_* U_d\big), \qquad (Y_1, \dots, Y_d) \xrightarrow{\mathfrak{M}} ([Y_i, Y_j])_{d\times d} \tag{1.2}$$

where $h_1 = \exp H$ and $\text{Vect}^\infty(X)$ denotes the space of C^∞ vector fields on X. Obviously, $\mathfrak{M} \circ \mathfrak{L} = 0$. Denote by $\mathfrak{L} \to \mathfrak{M}$ the nonlinear sequence of operators defined by (1.1) and (1.2).

Linearizing the sequence $\mathfrak{L} \to \mathfrak{M}$ "at the target", in other words at $H_0 = 0$ and at $(U_1, \dots, U_d) \in \mathrm{Vect}^\infty(X)^d$, we obtain the linearized sequence $L \to M$ as follows:

$$\mathrm{Vect}^\infty(X) \xrightarrow{L} \mathrm{Vect}^\infty(X)^d \xrightarrow{M} \mathrm{Vect}^\infty(X)^{d\times d} \tag{1.3}$$

$$H \xrightarrow{L} \big(\mathscr{L}_{U_1} H, \dots, \mathscr{L}_{U_d} H\big), \qquad (W_1, \dots, W_d) \xrightarrow{M} \big(\mathscr{L}_{U_j} W_i - \mathscr{L}_{U_i} W_j\big)_{d\times d}. \tag{1.4}$$

Again, our action is abelian, thus $M \circ L = 0$.

The following proposition is immediate from the fact that M is simply a linearization of the commutation condition and that we are interested in perturbations in the space of *commutative* actions. It relies on the simple estimate (2.1).

PROPOSITION 1.1. *If $\tilde{U} = U + W$ is a d-tuple of commuting vector fields, then*

$$\|M(W)\|_r \le C\|W\|_0 \|W\|_{r+1}$$

for $r \ge 0$.

This shows that the image of M is quadratically small with respect to W, so having a tame splitting (see Section 2 for definition of "tame") for the $L \to M$ sequence is the same as having an approximate solution to the linearized version of the conjugacy problem.

Finally, one may try to reduce the sequence $L \to M$ to a simpler sequence, for functions rather than vector fields

$$C^\infty(X) \xrightarrow{\mathscr{L}} C^\infty(X)^d \xrightarrow{\mathscr{M}} C^\infty(X)^{d\times d},$$

where by $\mathscr{L} \to \mathscr{M}$ we denote the following sequence of operators:

$$f \xrightarrow{\mathscr{L}} \big(\mathscr{L}_{U_1} f, \dots, \mathscr{L}_{U_d} f\big), \qquad (g_1, \dots, g_d) \xrightarrow{\mathscr{M}} \big(\mathscr{L}_{U_j} g_i - \mathscr{L}_{U_i} g_j\big)_{d\times d}.$$

Now the general strategy of approaching the local perturbation problem for a general $\mathbb{R}^k$ action can be roughly summarized as

$$\mathscr{L} \to \mathscr{M} \text{ splits} \implies L \to M \text{ splits} \implies \mathfrak{L} \to \mathfrak{M} \text{ is exact.} \tag{1.5}$$

Note that all the splittings above, as well as the exactness, should be *tame*, see Section 2 for exact definitions.

The paper is dedicated to proving the first statement that $\mathscr{L} \to \mathscr{M}$ splits and the two implications in (1.5) for certain actions under some additional conditions: certain small modifications and smaller splitting spaces. These small modifications will be simply coordinate changes in the acting group and considering smaller spaces is necessary because actions we deal with have nontrivial first cohomology, see Step 4 of the general scheme. The generality in which these statements are proved increases from left to right.

- The splitting of $\mathscr{L} \to \mathscr{M}$ is proved for certain type of examples (Example 1 for the next section) via the splitting construction in Section 3.
- The first implication in (1.5) is discussed in the Section 4 and relies very much on the fact that the actions we consider here are upper unipotent, but applies to all the examples we describe in this paper.

- The second implication in (1.5) is proved through the KAM-type iteration scheme, which is very general, in Sections 6 and 7.

1.3. **Description of some unipotent actions.** The choice of examples below is not accidental. They represent exactly those unipotent homogeneous actions of $\mathbb{R}^k$, $k \geq 2$, for which Step 2 of the general scheme has been performed based on representation theory for $SL(2,\mathbb{R})$ [8][1] and $SL(2,\mathbb{C})$ [12]. Steps 3, 4, and 5(a) work for all three cases. Steps 5(b) and 6 for Example 1 below are carried out in Sections 3 and 4. Only the first part (tame estimates and splitting for functions) is specific for that example. Reduction from vector fields to functions (Section 4) is applicable to all three cases and is valid in even more general setting. The last two steps are carried out in Sections 6 and 7 and, as we already pointed out, are based only on the conclusions of the previous steps and are hence of general character.

We will discuss the remaining steps (Steps 5(b) and 6) for Examples 2 and 3 as well as other applications of our method, both in progress and potential, in Section 9.

EXAMPLE 1 (on $SL(2,\mathbb{R}) \times SL(2,\mathbb{R})$)**.** This is the model referred to in the title of the paper. Let $G = SL(2,\mathbb{R}) \times SL(2,\mathbb{R})$, Γ be an irreducible cocompact lattice in G, $X = G/\Gamma$, $\mathfrak{g}^j = sl(2,\mathbb{R})$ $(j = 1,2)$, and $\mathfrak{g} = \mathrm{Lie}\, G = sl(2,\mathbb{R}) \times sl(2,\mathbb{R})$.

The Lie algebra $\mathfrak{g}^j$ has a basis U_+^j, U_0^j, and U_-^j for $j = 1,2$, where

$$U_+^1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad U_+^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$U_0^1 = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad U_0^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & -\frac{1}{2} \end{pmatrix}$$

$$U_-^1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad U_-^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Our model parabolic action

$$\alpha_{\mathbb{R}\mathbb{R}} \colon \mathbb{R}^2 \times X \to X$$

is the action by left translations on X generated by commuting unipotent elements

$$U_1 = U_+^1, \quad U_2 = U_+^2.$$

[1]The quoted paper is a pioneering work where classification of irreducible unitary representations of a simple Lie group has been applied to a dynamical cohomology problem. In earlier work only more general information such as speed of decay of matrix coefficients was used. The latter is quite useful in hyperbolic and partially hyperbolic problems but is not sufficient in the parabolic setting since decay of correlations for smooth functions is too slow.

More precisely, if (t_1, t_2) denote coordinates in $\mathbb{R}^2$ then

$$\alpha_{\mathbb{R}\mathbb{R}}(t_1, t_2, x) = \begin{pmatrix} 1 & t_1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & t_2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot x. \tag{1.6}$$

We denote the pair of vector fields (U^1_+, U^2_+) generating the action $\alpha_{\mathbb{R}\mathbb{R}}$ by U and the pair (U^1_-, U^2_-) by V.

EXAMPLE 2 (on $SL(2,\mathbb{R}) \times SL(2,\mathbb{C})$)**.** Let $G = SL(2,\mathbb{R}) \times SL(2,\mathbb{C})$, Γ be an irreducible cocompact lattice in G, $X = G/\Gamma$, $\mathfrak{g}^1 = sl(2,\mathbb{R})$, and $\mathfrak{g}^2 = sl(2,\mathbb{C})$. The Lie algebra $\mathfrak{g}^1$ is generated by U^1_+, U^1_-, and U^1_0, as in the example above, while $\mathfrak{g}^2$ is generated by U^2_+, $(U^2_+)' = iU^2_+$, U^2_-, $(U^2_-)' = iU^2_-$, U^2_0, and $(U^2_0) = iU^2_0$, where

$$U^2_+ = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad U^2_- = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad U^2_0 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & -\frac{1}{2} \end{pmatrix}.$$

Our example is the action $\alpha_{\mathbb{R}\mathbb{C}}$ of $\mathbb{R}^3$ on X, by left translations, generated by commuting unipotent elements

$$U_1 = U^1_+, \quad U_2 = U^2_+, \quad U_3 = (U^2_+)'.$$

We denote the triple of vector fields $(U^1_+, U^2_+, (U^2_+)')$ generating the action $\alpha_{\mathbb{R}\mathbb{C}}$ by U, and the triple $(U^1_-, U^2_-, (U^2_-)')$ by V.

EXAMPLE 3 (on $SL(2,\mathbb{C})$)**.** Let $G = SL(2,\mathbb{C})$, Γ be an irreducible cocompact lattice in G, $X = G/\Gamma$. In this case $m = 1$, $\mathfrak{g} = \mathfrak{g}^1 = sl(2,\mathbb{C})$, and $\mathfrak{g}^1$ is generated by U^1_+, $(U^1_+)' = iU^1_+$, U^1_-, $(U^1_-)' = iU^1_-$, U^1_0, and $(U^1_0)' = i(U^1_0)'$, where

$$U^1_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad U^1_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad U^1_0 = \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix}.$$

Our third example is the action $\alpha_{\mathbb{C}}$ of $\mathbb{R}^2$ on X by left translations, generated by commuting unipotent elements

$$U_1 = U^1_+, \quad U_2 = (U^1_+)'.$$

This example possesses a peculiar feature that will be explained and discussed in Section 9.1. For now we just mention that If the first Betti number of Γ is not zero there are additional obstructions in the cohomology problem and they correspond to certain specific classes of perturbations of action $\alpha_{\mathbb{C}}$.

1.4. **Cocycle rigidity of parabolic (unipotent) actions.** David Mieczkowski [12] studied real-valued cocycles over the actions of unipotent subgroups described above. He proved vanishing of the obstructions and constructed smooth solutions in all three cases, obtained tame estimates for both product cases (actions $\alpha_{\mathbb{R}\mathbb{R}}$ in Example 1 and $\alpha_{\mathbb{R}\mathbb{C}}$ in Example 2) and for the former also gave a partial proof for a splitting of an almost cocycle into a cocycle and a small residue with tame estimates.

The starting point in [12] for Examples 1 and 2 is the description of obstructions and solution of cohomological equations for horocycle flows (upper-triangular unipotent action on $SL(2,\mathbb{R})/\Gamma$) by L. Flaminio and G. Forni [8]. In those cases presence of one $SL(2,\mathbb{R})$ factor is essential. Solutions are constructed for the unipotent element from that factor and estimates are carried out. Additional generator(s) serve an auxiliary purpose to guarantee the vanishing of obstructions for the solution of the cohomological equations for the first factor. In fact, specific choices of the second factor in these two cases has to do with the fact that those are the only cases where irreducible lattices exist in the product. Example 3 on $SL(2,\mathbb{C})/\Gamma$ is completely different as we will remark later in Section 9.1.

The space $L^2(SL(2,\mathbb{R})/\Gamma)$ is decomposed into the direct sum of irreducible representation spaces for $SL(2,\mathbb{R})$; representations of principal, complementary and discrete series are treated separately; each carries obstructions for solutions of the cohomological equation

$$\mathscr{L}_U h = g$$

for the nilpotent $U = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) \in \mathfrak{sl}(2,\mathbb{R})$.

Thus, not only general perturbations but already *time changes* of the horocycle flow are highly unstable (infinitely many moduli of smooth conjugacy) in contrast with the Diophantine elliptic case.

For an $\mathbb{R}^2$ action generated by unipotents $\exp U_1$ and $\exp U_2$, the basic *cocycle condition* has the form

$$\mathscr{L}_{U_2} f = \mathscr{L}_{U_1} g$$

(and similarly for more generators). Integrals with respect to the Haar measure are obvious obstructions and solution, if it exists, is unique up to an arbitrary constant.

The higher-rank trick is applied in each irreducible representation space (except for the trivial one and a single other exception for $G = SL(2,\mathbb{C})$) and allows to find h such that

$$\mathscr{L}_{U_1} h = f, \qquad \mathscr{L}_{U_2} h = g$$

with tame estimates for an appropriate family of Sobolev norms.

For $G = SL(2,\mathbb{R}) \times SL(2,\mathbb{R})$ or $SL(2,\mathbb{R}) \times SL(2,\mathbb{C})$, one uses *algebraic description* for the irreducible representations, similarly to Flaminio and Forni [8], *i.e.,* constructs convenient orthogonal bases in each irreducible space and describes the generators of the action in those bases. But for $G = SL(2,\mathbb{C})$ one uses *analytic description* of the irreducible representations as actions in certain Hilbert spaces of functions on $\mathbb{R}^2$ and the Riemann sphere.

Mieczkowski gave tame estimates with respect to standard Sobolev norms for both product cases. But for the $SL(2,\mathbb{C})$ case Mieczkowski only obtained tame estimates with respect to *incomplete* Sobolev norms based on two embeddings of $SL(2,\mathbb{R})$ into $SL(2,\mathbb{C})$.

Spectral gap is essential for obtaining tame estimates in any of the cases. It plays the role of Diophantine conditions.

1.5. **The main result.** As we already mentioned, in this paper we carry out the program outlined in Section 1.1 for the action $\alpha_{\mathbb{R}\mathbb{R}}$ in Example 1.

Let $U = (U^1_+, U^2_+)$ be generators of the action $\alpha_{\mathbb{R}\mathbb{R}}$ in Example 1. Recall that $V = (U^1_-, U^2_-)$ denotes the pair of "opposite" unipotent elements. Given a family of actions generated by vector fields $\tilde{U}(\lambda) = (\tilde{U}_1(\lambda), \tilde{U}_2(\lambda))$, we define a map $\Phi(\tilde{U})\colon \mathbb{R}^2 \to \mathbb{R}^2$ such that $\Phi(\tilde{U})$ takes each $\lambda \in \mathbb{R}^2$ to the pair of real numbers $(\mathrm{Ave}^1_-(\tilde{U}_1(\lambda)), \mathrm{Ave}^2_-(\tilde{U}_2(\lambda)))$ which are averages of $\tilde{U}_1(\lambda)$ and $\tilde{U}_2(\lambda)$, respectively, in the directions U^1_- and U^2_- with respect to the Haar measure.

We use the following norms for a family $\tilde{U}(\lambda) = (\tilde{U}_1(\lambda), \tilde{U}_2(\lambda))$ of pairs of vector fields on X:

- For a fixed λ, $\|\tilde{U}(\lambda)\|_r$ denotes the maximum of the C^r norms of $\tilde{U}_1(\lambda)$ and $\tilde{U}_2(\lambda)$.
- $\|\tilde{U}(\lambda)\|_{r,(m)}$ is the maximum of the C^r norms of all the partial derivatives of $\tilde{U}$ in the λ variable of order less or equal to m.
- For the "averages" map $\Phi(\tilde{U})\colon \mathbb{R}^2 \to \mathbb{R}^2$, the notation $\|\Phi(\tilde{U})\|_{(r)}$ stands for the usual C^r norm.

THEOREM 1.2 (Main Theorem)**.** *Let $U = (U^1_+, U^2_+)$ be as in the Example 1, and let $\alpha_{\mathbb{R}\mathbb{R}}$ be the $\mathbb{R}^2$ action generated by U, see* (1.6).

There exist $l > 0$, $\varepsilon > 0$, $K > 0$, and $C > 0$ such that for every family $\tilde{\alpha}(\lambda)$ of perturbations of $\alpha_{\mathbb{R}\mathbb{R}}$ where $\lambda \in B = B(0,\varepsilon) \subset \mathbb{R}^2$, generated by a pair of vector fields $\tilde{U}(\lambda) = (\tilde{U}_1(\lambda), \tilde{U}_2(\lambda))$ such that for $\lambda \in B$:

$$\|\tilde{U}\|_{0,(2)} \le K, \qquad \|\Phi(\tilde{U}) - \mathrm{Id}\|_{(1)} \le \frac{\varepsilon}{2C}, \quad \text{and} \qquad \|\tilde{U}(\lambda) - U\|_l \le \varepsilon$$

there exist $\bar{\lambda} \in B$, an automorphism π of the acting group $\mathbb{R}^2$ and $h \in \mathrm{Diff}^\infty X$ such that for all $x \in X$, $a \in \mathbb{R}^2$,

$$h(\tilde{\alpha}(\bar{\lambda})(a,x)) = \alpha_{\mathbb{R}\mathbb{R}}(\pi(a), h(x)).$$

The work on Examples 2 and 3 is in progress and appropriate local rigidity results will appear in a subsequent paper joint with Livio Flaminio. We refer to Section 9.1 for more detailed discussion of those examples.

A substantial part of the proof of the above result is presented in the form which can be used in even greater generality than the remaining two examples.

2. SOME ANALYTIC TOOLS

2.1. **Graded Fréchet spaces and tame operators.**

DEFINITION 2.1. By a graded Fréchet space, we mean a Fréchet space X with a collection of norms $\|\cdot\|_r$, $r \in \mathbb{N}_0$, such that $\|x\|_r \le \|x\|_{r+k}$ for every $r, k \ge 0$, for every $x \in X$.

DEFINITION 2.2. A map $L\colon X \to Y$ between two graded Fréchet spaces is r_0-*tame* if for every $x \in X$ and for every $r \ge 0$ we have that $\|Lx\|_r \le C_r \|x\|_{r+r_0}$

REMARK 2.3. The above notion of "tame" is used in literature for linear maps, and in this paper we will exclusively apply it to linear maps.

DEFINITION 2.4. Tame maps $L\colon X \to Y$ and $M\colon Y \to Z$ between graded Fréchet spaces are said to form a sequence of maps if $ML = 0$. We denote such a sequence of maps by $L \to M$. If $\operatorname{Im} L = \ker M$ we call the sequence exact. We call the sequence $L \to M$ r_0-*tamely exact* if there exists an r_0-tame map L' from $\ker M$ to X such that $LL' = \mathrm{Id}$ on $\ker M$.

DEFINITION 2.5. It is said that a sequence of tame maps $X \xrightarrow{L} Y \xrightarrow{M} Z$ has an r_0-*tame splitting on a subspace* $\bar{Y}$ of Y if there exists r_0-tame maps $L'\colon \bar{Y} \to X$ and $M'\colon Z \to \bar{Y}$ such that $LL' + M'M = \mathrm{Id}_{\bar{Y}}$.

2.2. **Fréchet spaces and gradings to be used in the proof of Theorem 1.2.** The space X is compact, thus the spaces of smooth functions and smooth vector fields are graded Fréchet spaces. We will use the following notation:

1. $C^\infty(X)$ denotes the graded Fréchet space of C^∞ functions with grading given either by the usual C_r norms or by the Sobolev norms. We will always specify which grading we use, and the chosen grading will be denoted by $\|\cdot\|_r$. Similarly, $C^\infty(X)^d$ and $C^\infty(X)^{d\times d}$ denote d-tuples and $d\times d$ matrices, respectively, of C^∞-functions on X.
2. $\mathrm{Vect}^\infty(X)$ denotes the graded Fréchet space of vector fields on X of class C^∞ with grading given by the componentwise norms (C^r or Sobolev). These we also denote by $\|\cdot\|_r$.
3. $\mathrm{Vect}^\infty(X)^d$ and $\mathrm{Vect}^\infty(X)^{d\times d}$ denote d-tuples, and $d\times d$ matrices of vector fields on X, respectively. The grading on this space is given by norms which we again denote by $\|\cdot\|_r$ and which are again just maxima over the corresponding coordinate wise norms.

2.3. **Smoothing operators and some norm inequalities.** The space X is compact, thus the spaces of smooth functions and smooth vector fields are graded Fréchet spaces. There exists a collection of smoothing operators $S_t\colon C^\infty(X) \to C^\infty(X)$, $t > 0$, such that the following holds:

$$\|S_t F\|_{s+s'} \le C_{s,s'}\, t^{s'} \|F\|_s$$
$$\|(I - S_t)F\|_{s-s'} \le C_{s,s'}\, t^{-s'} \|F\|_s.$$

For the construction of smoothing operators see the following in [9]: Example 1.1.2.(2), Definition 1.3.2, Theorem 1.3.6, and Corollary 1.4.2.

Smoothing operators on $C^\infty(X)$ induce smoothing operators on $\mathrm{Vect}^\infty(X)^d$, $\mathrm{Vect}^\infty(X)^{d\times d}$ via smoothing operators applied to coordinate maps.

It is easy to see that averages of F with respect to the Haar measure on X, in various directions in the tangent space do not affect the properties of smoothing operators listed above, so without loss of generality we may assume that S_t are such that averages of $S_t F$ in directions of U_k^+, U_k^-, U_k^0 are the same as those of F.

The following inequality implies the statement of Proposition 1.1. It will also be used later in Section 6. Let $F, G \in \mathrm{Vect}^\infty(X)$. Then, for $r \ge 0$

$$\|[F,G]\|_r \le C_r(\|F\|_0 \|G\|_{r+1} + \|F\|_{r+1}\|G\|_0) \tag{2.1}$$

Also, existence of smoothing operators on a graded space induces the usual interpolation inequalities for the norms which make up the grading. This has been demonstrated in several papers, for example [16] and in a more general context in [9].

3. FIRST COHOMOLOGY AND TAME SPLITTING FOR FUNCTIONS

In this section we show that the exact sequence of maps

$$C^\infty(X) \xrightarrow{\mathscr{L}} C^\infty(X)^2 \cap \ker \mathscr{A} \xrightarrow{\mathscr{M}} C^\infty(X)$$

where

$$\mathscr{L}(f) = (\mathscr{L}_{U_1} f, \mathscr{L}_{U_2} f), \qquad \mathscr{M}(f,g) = \mathscr{L}_{U_1} g - \mathscr{L}_{U_2} f,$$

and $\mathscr{A}(f,g) = \left(\int_X f\, d\nu, \int_X g\, d\nu\right)$, has an r_0-tame splitting. The constant r_0 depends on spectral gap of Casimir operators $\Box_1$ and $\Box_2$ on $SL(2,\mathbb{R}) \times SL(2,\mathbb{R})/\Gamma$.

In [12], Mieczkowski proved tame estimates for solutions of the scalar cohomological equation, *i.e.,* showed that the sequence $\mathscr{M} \to \mathscr{L}$ is r_0-tamely exact where r_0 depends only on the lattice Γ. In the process he uses explicit obstructions to the solutions of cohomology problem for a single action generator U_k, $k = 1,2$. The obstructions come from work of Flaminio and Forni [8]. It is important for our application to remark that the space of cohomological obstructions for each U_k is found to be *finite*-dimensional in each irreducible representation (this follows immediately from [8]). More precisely, depending on the kind of irreducible representation the space of obstructions is either one or two-dimensional.

Below, we give a general construction of a tame splitting for the product examples, in the situation where for one action generator there is a finite-dimensional space of obstructions for the cohomology problem in each irreducible representation.

Irreducible unitary representations of $SL(2,\mathbb{R}) \times SL(2,\mathbb{R})$ are tensor products

$$\mathscr{H}_\mu \otimes \mathscr{H}_\theta$$

of the irreducible unitary representations $\mathscr{H}_\mu$ and $\mathscr{H}_\theta$ for $SL(2,\mathbb{R})$. For $i = 1$ or 2, μ (resp., θ) is the value which the Casimir operator $\Box_1$ in the first component (resp., the Casimir operator $\Box_2$ in the second component) of the product $SL(2,\mathbb{R}) \times SL(2,\mathbb{R})$ takes on $\mathscr{H}_\mu$ (resp., $\mathscr{H}_\theta$).

The space $L^2(SL(2,\mathbb{R}) \times SL(2,\mathbb{R})/\Gamma)$ decomposes as the direct integral of nontrivial irreducible unitary representations $\mathscr{H}_{\mu,\theta}$ with respect to Stieltjes measure $ds(\mu,\theta)$. Similarly, spaces $W^s(\mathscr{H})$ decompose for every $s \geq 0$ which implies that every $f \in W^s(\mathscr{H})$ has a decomposition

$$f = \int_\oplus m(\mu_1,\mu_2) f_{\mu_1,\mu_2}\, ds(\mu_1,\mu_2),$$

where $f_{\mu_1,\mu_2} \in W^s(\mathscr{H}_{\mu_1,\mu_2})$ and $m(\mu_1,\mu_2)$ is multiplicity with which $\mathscr{H}_{\mu_1,\mu_2}$ occurs in $\mathscr{H}$.

REMARK 3.1. *Only* in this section the notation $\|\cdot\|_r$ is used for *Sobolev* norms. In the rest of the paper $\|\cdot\|_r$ stands for C^r norms.

The Sobolev norm of $f \in W^s(\mathcal{H})$ is then given by

$$\|f\|_s^2 = \int m(\mu,\theta)\|f_{\mu,\theta}\|_s^2\, ds(\mu,\theta).$$

This implies that for $f, g \in W^s(\mathcal{H})$, the uniform Sobolev estimates

$$\|f_{\mu,\theta}\|_t \leq C_{s,t}\|g_{\mu,\theta}\|_s \qquad \text{for every } (\mu,\theta),$$

imply the global Sobolev estimate

$$\|f\|_t \leq C_{s,t}\|g\|_s.$$

This reduces the proof of the existence of tame splitting in Corollary 3.7 below to proving the existence of a tame splitting in each irreducible $\mathcal{H}_{\mu,\theta}$ in such a way that the constants involved in tame estimates can be chosen independently of (μ,θ).

When $SL(2,\mathbb{R}) \times SL(2,\mathbb{R})/\Gamma$ is compact, the above direct integral decompositions into irreducible representations reduce to infinite discrete sums.

We expect that the theorem below will be used for other actions, in particular for $\alpha_{\mathbb{R}\mathbb{C}}$ from Example 2, so we formulate it and prove it in some generality.

THEOREM 3.2. *Let G_1 and G_2 be Lie groups, let Γ be an irreducible lattice in $G_1 \times G_2$ and let $\mathcal{H} = \mathcal{H}_\mu \otimes \mathcal{H}_\theta$ be an irreducible representation in the decomposition of $L^2(G_1 \times G_2/\Gamma)$. Assume $U_1 \in \mathfrak{g}_1$ and $U_2 \in \mathfrak{g}_2$ commute. Assume that the space of U_1 invariant distributions $\mathcal{I}_1^s(\mathcal{H}_\mu)$ of Sobolev order s is finite-dimensional, spanned by distributions $D_1,\dots,D_m$. Let $\tilde{D}_i(f_\mu \otimes g_\theta) := D_i(f_\mu)g_\theta$ for $i = 1,\dots,m$. Further, assume that there exists $s_0 \geq 0$, $\sigma \geq 0$ independent of the pair (μ,θ), such that the following hold for $k = 1,2$ and any s sufficiently large:*

1. *If $f = U_1 h \in W^s(\mathcal{H})$ then $\tilde{D}_i(f) = 0$ for all $i = 1,\dots,m$.*
2. *If $\tilde{D}_i(f) = 0$ for all $i = 1,\dots,m$, then the equation $U_1 h = f$ has a solution $h \in W^t(\mathcal{H})$, $t < s - \sigma$, such that $\|h\|_t \leq C_{t,s}\|f\|_s$.*
3. *There exists s_0 such that if $U_k h = 0$ with $h \in W^{s_0}(\mathcal{H})$, then $h = 0$.*
4. *If $U_2 h = f \in W^s(\mathcal{H})$ and if $h \in W^{s_0}(\mathcal{H})$, then $h \in W^t(\mathcal{H})$, $t < s - \sigma$ and $\|h\|_t \leq C_{t,s}\|f\|_s$,*

where constants $C_{t,s}$ are independent of the pair (μ,θ).

Then if $U_2 f - U_1 g = \phi$ for $f, g, \phi \in W^s(\mathcal{H})$, there exists $h \in W^t(\mathcal{H})$ and there exists a constant C such that

$$\|f - U_1 h\|_t \leq C_{t,s}\|\phi\|_s$$
$$\|g - U_2 h\|_t \leq C_{t,s}\|\phi\|_s$$
$$\|h\|_t \leq C_{t,s}\|f\|_s.$$

The constant $C_{t,s}$ depends only on the constants which appear in the estimates of the assumptions (2) and (4) and can be chosen independently of the pair (μ,θ).

REMARK 3.3. Our original proof of Theorem 3.2 for the case of the action $\alpha_{\mathbb{RR}}$ from Example 1 was based on a specific and explicit construction of projections in each irreducible representation which used detailed descriptions of the representations. This proof completed and corrected attempted proof of Theorem 18 in [12] that in fact contains an error because it uses only one of the two obstructions that appear in any irreducible representation of $SL(2,\mathbb{R})$ of the principal series.

The shorter and more general proof presented below was inspired by our discussions with Livio Flaminio.

Proof. For each $i = 1,\dots,m$ there exists $\gamma_i \in W^s(\mathcal{H}_\mu)$ such that $D_i(\gamma_i) = 1$ and $\|\gamma_i\|_s \le 2$. By applying a linear map on the m-tuple $(\gamma_1,\dots,\gamma_m)$ we can obtain an m-tuple of maps $(\tilde\gamma_1,\dots,\tilde\gamma_m)$ satisfying $D_j(\tilde\gamma_i) = \delta_{ij}$. The norm of the linear map is bounded above by the maximum of the norms of D_i, $i = 1,\dots,m$.

Thus without loss of generality, we may assume that we have $\gamma_1,\dots,\gamma_m \in W^s(\mathcal{H}_\mu)$ such that for any $i = 1,\dots,m$, $\|\gamma_i\|_s \le C$, where C is a constant multiple of the maximum of the norms of D_i's on $W^s(\mathcal{H})$, such that $D_i(\gamma_i) = 1$ and $D_j(\gamma_i) = 0$ for $i \ne j$. We will use this property of γ_i in an essential way in the proof of the lemma below, in order to define a "good" projection map $\mathcal{R}$.

Let V_1^s be the intersection of kernels of the $\tilde D_i$'s in $W^s(\mathcal{H})$. The following is the central ingredient in the proof of the theorem.

LEMMA 3.4. *There exists a linear map* $\mathcal{R}\colon W^s(\mathcal{H}) \to W^t(\mathcal{H})$ *such that*

(i) $\mathcal{R}$ *is trivial on* V_1^s,
(ii) $\|\mathcal{R}f\|_s \le C\|f\|_s$,
(iii) $\tilde D_j(\mathcal{R}f) = \tilde D_j(f)$ *for all* $j = 1,\dots,m$,
(iv) $\mathcal{R}$ *commutes with* U_2.

Proof. For $f = f_\mu \otimes g_\theta$, let

$$\mathcal{R}f = \sum_{i=1}^{m} \gamma_i \otimes \tilde D_i(f).$$

Then since $D_i(f)$ is bounded, and each γ_i has the $W^s(\mathcal{H})$ norm bounded by a constant, we get $\|\mathcal{R}f\|_s \le C\|f\|_s$. It is obvious by the definition that $\mathcal{R}$ is trivial on V_1^s. Further, from $D_j(\gamma_i) = \delta_{ij}$, we have

$$\tilde D_j(\mathcal{R}f) = \sum_i D_j(\gamma_i)\tilde D_i(f) = \tilde D_j(f).$$

To check commutativity, we use the fact that U_2 commutes with $\tilde D_i$:

$$\mathcal{R}(U_2 f) = \sum_{i=1}^{m} \gamma_i \otimes \tilde D_i(U_2 f) = \sum_{i=1}^{m} \gamma_i \otimes U_2 \tilde D_i(f) = U_2(\mathcal{R}f).$$

This concludes the proof of Lemma 3.4. □

Now for $f \in W^s(\mathcal{H})$, we may define $\mathcal{P}f := f - \mathcal{R}f$. Then from the properties of $\mathcal{R}$ obtained in Lemma 3.4 $\mathcal{P}\colon W^s(\mathcal{H}) \to V_1^s$ and $\|\mathcal{P}f\|_s \le C\|f\|_s$, and $\mathcal{P}$ commutes with U_2 as well.

Given $U_2 f - U_1 g = \phi$, by assumption (1) and the properties of $\mathscr{R}$ we get $U_2\mathscr{R}f = \mathscr{R}\phi$. Then by assumptions (3) and (4), we have the estimate $\|\mathscr{R}f\|_t \le C\|\mathscr{R}\phi\|_s$ and by the bound for $\mathscr{R}$ we get $\|\mathscr{R}f\|_t \le C\|\phi\|_s$.

Now by assumption (2) we have that $\mathscr{P}f = U_1 h$ for some $h \in W^t(\mathscr{H})$ with $\|h\|_t \le C\|\mathscr{P}f\|_s \le C\|f\|_f$. On the other hand, we have from $U_2 f - U_1 g = \phi$ that $U_2(\mathscr{P}f) = U_1 g + \mathscr{P}\phi$ since $\mathscr{P}$ is identity on V_1^s. Substituting $\mathscr{P}f = U_1 h$ we get $U_1(U_2 h - g) = \mathscr{P}\phi$. Since for sufficiently large s, $U_2 h - g \in W^{s_0}$, by assumptions (1), (2), and (3), we have that $\|U_2 h - g\|_t \le C\|\mathscr{P}\phi\|_s \le C\|\phi\|_s$. □

The assumptions of Theorem 3.2 are checked using the results of [8] similarly to the way the same results were used by Mieczkowski in his proof of vanishing of the first cohomology over the action $\alpha_{\mathbb{R}\mathbb{R}}$ modulo averages [12]. Below we summarize conclusions of [8] adapted to the product case in the form which is sufficient for the purpose of the current paper.

THEOREM 3.5. *Let $\mathscr{H} = \mathscr{H}_\mu \times \mathscr{H}_\theta$ be an irreducible representation as in Theorem 3.2. Let $U_1 = U_+^1$, $U_2 = U_+^2$, and X be as in Example 1. Then the space of U_1-invariant distributions $\mathscr{I}_1^s(\mathscr{H}_\mu)$ is either one- or two-dimensional, and conditions (1)–(4) of Theorem 3.2 are satisfied in such a way that constants $C_{s,t} > 0$, σ and s_0 can be chosen independently of μ and θ.*

REMARK 3.6. All three constants obtained in the theorem above depend only on Γ, that is on spectral gap of Casimir operators $\Box_1$ and $\Box_2$ in the first and second component.

The tame splitting of the $\mathscr{L} \to \mathscr{M}$ sequence is now an easy corollary of the previous two theorems.

COROLLARY 3.7. *Let $U_1 = U_+^1$, $U_2 = U_+^2$, and X be as in Example 1. For $h, f, g \in C^\infty(X)$, let $\mathscr{L}(h) = (\mathscr{L}_{U_1}h, \mathscr{L}_{U_2}h)$ and let $\mathscr{M}(f,g) = \mathscr{L}_{U_2}f - \mathscr{L}_{U_1}g$. Then there exist $r_0 > 0$ and there exist r_0-tame operators (with respect to C^r norms) $\mathscr{L}'$ and $\mathscr{M}'$ such that*

$$\mathscr{L}\mathscr{L}' + \mathscr{M}'\mathscr{M} = \mathrm{Id}$$

on $C^\infty(X)^2 \cap \ker\mathscr{A}$, where $\mathscr{A}(f,g) = \left(\int_X f\,d\nu, \int_X g\,d\nu\right)$ and ν is the Haar measure on X.

Proof. For $\mathscr{H} = \mathscr{H}_\mu \otimes \mathscr{H}_\theta$ and $f, g \in W^s(\mathscr{H})$, Theorems 3.2 and 3.5 imply that there exists $h \in W^t(\mathscr{H})$ such that $\|h\|_t \le C_{t,s}\|f\|_s$, $\|f - U_1 h\|_t \le C_{t,s}\|\phi\|_s$ and $\|g - U_2 h\|_t \le C_{t,s}\|\phi\|_s$. Thus we define $\mathscr{L}'(f,g) = h$ and $\mathscr{M}'(\mathscr{M}(f,g)) = (f - U_1 h, g - U_2 h)$. Since by Theorem 3.5 the constants in the estimates do not depend on the representation $\mathscr{H}_\mu \times \mathscr{H}_\theta$, it is clear that after defining $\mathscr{L}'$ and $\mathscr{M}'$ on all of $\bigoplus \mathscr{H}_\mu \otimes \mathscr{H}_\theta$ by gluing $\mathscr{L}'$ and $\mathscr{M}'$ in each irreducible representation, the tame estimates in Sobolev norms for $\mathscr{L}'$ and $\mathscr{M}'$ remain true. From these the tame estimates in the C^r norms follow in a standard way with the loss of r_0 derivatives where r_0 depends only on the dimension of X and the lattice Γ. □

4. TAME SPLITTING FOR VECTOR FIELDS

As mentioned in the introduction, due to Proposition 1.1 the existence of tame splitting of the $L \to M$-sequence in (1.4) implies the existence of an *approximate* right inverse for the linearization L of our initial conjugacy problem and this is the necessary ingredient in any attempt to solve a conjugacy problem by linearization.

In this section we show the existence of a tame splitting $LL' + M'M = \mathrm{Id}$ on $\mathrm{Vect}^\infty(X)^d \cap \ker A$ for the sequence $L \to M$ for the action $\alpha_{\mathbb{R}\mathbb{R}}$ from Example 1, where A is a specifically defined functional on $\mathrm{Vect}^\infty(X)^d$. To do this we choose a basis in the Lie algebra $\mathfrak{g}$ and by looking at the equation (1.4) in this basis we reduce the splitting problem to several simpler problems. Those simpler problems inductively reduce to existence of the tame splitting for functions, *i.e.,* for the sequence $\mathscr{L} \to \mathscr{M}$ to which the results of the previous section apply.

We give a unified proof for all three examples described in Section 1.3 that splitting of functions implies splitting for vector fields, in anticipation of forthcoming tame splitting results for functions for actions $\alpha_{\mathbb{R}\mathbb{C}}$ and $\alpha_{\mathbb{C}}$ from Examples 2 and 3. Furthermore, the scheme of this proof can be adapted to even more general situations of unipotent homogeneous actions.

The reader should be warned though that uniformity comes at a price of certain clarity since in fact in the formulas written in a generic way m equals either one or two, and d equals two or three. Rewriting the calculations for each specific case will make them more compact and easier to visualize.

Let G^j, $j = 1,2$, be either $SL(2,\mathbb{R})$ or $SL(2,\mathbb{C})$ and consider one of the actions described in Section 1.3. By U_0^j, U_-^j, U_+^j, $(U_-^j)'$, and $(U_+^j)'$ we denote the vector fields that are equal corresponding to U_0, U_-, U_+, U'_-, and U'_+ in the j-th component of the product $G^1 \times G^2$, $j = 1,2$, and zero elsewhere.

We will use generic notation α for the actions $\alpha_{\mathbb{R}\mathbb{R}}$, $\alpha_{\mathbb{R}\mathbb{C}}$ and $\alpha_{\mathbb{C}}$. Generators of α are as follows:

1. Example 1: $G = SL(2,\mathbb{R}) \times SL(2,\mathbb{R})$; $U_1 = U_+^1$, $U_2 = U_+^2$.
2. Example 2: $G = SL(2,\mathbb{R}) \times SL(2,\mathbb{C})$; $U_1 = U_+^1$, $U_2 = U_+^2$, $U_3 = (U_+^2)'$.
3. Example 3: $SL(2,\mathbb{C})$; $U_1 = U_+^1$, $U_2 = (U_+^1)'$.

For $H \in \mathrm{Vect}^\infty(X)$, let $H = \sum_{j=1}^m h_0^j U_0^j + h_+^j U_+^j + h_-^j U_-^j$ and $* \in \{0,+,-\}$ in the following.

- In case (1), $m = 2$ and h_*^j are real-valued functions.
- In case (2), $m = 2$ and h_*^1 are real-valued and h_*^2 are complex-valued.
- In case (3), $m = 1$ and h_*^1 are complex-valued.

Now we use both notations for the Lie derivative, $\mathscr{L}_X Y = [X,Y]$, simultaneously. We have

$$\mathscr{L}_{U_k} H = \sum_{j=1}^m \Big(\mathscr{L}_{U_k} h_0^j U_0^j + h_0^j [U_k, U_0^j] + \mathscr{L}_{U_k} h_+^j U_+^j + h_+^j [U_k, U_+^j] + \mathscr{L}_{U_k} h_-^j U_-^j + h_-^j [U_k, U_-^j] \Big). \tag{4.1}$$

From the definition of U_k it follows that $[U_k, U_+^j] = 0$ for any j, k, and

$$[U_k, U_0^j] = \begin{cases} -iU_+^j, & U_k = (U_+^j)' \\ -U_+^j, & k = j,\ U_k = U_+^j \\ 0, & \text{otherwise,} \end{cases} \quad \text{and} \quad [U_k, U_-^j] = \begin{cases} 2iU_0^j, & U_k = (U_+^j)' \\ 2U_0^j, & U_k = U_+^j \\ 0, & \text{otherwise.} \end{cases}$$

Let

$$\psi_{kj} = \begin{cases} i, & U_k = (U_+^j)' \\ 1, & U_k = U_+^j \\ 0, & \text{otherwise.} \end{cases} \tag{4.2}$$

Then, from (4.1)

$$\mathcal{L}_{U_k} H = \sum_{j=1}^{m} \left(\mathcal{L}_{U_k} h_0^j + 2\psi_{kj} h_-^j\right) U_0^j + \left(\mathcal{L}_{U_k} h_+^j - \psi_{kj} h_0^j\right) U_+^j + \mathcal{L}_{U_k} h_-^j U_-^j.$$

This completely describes the linear operator $L\colon H \mapsto (\mathcal{L}_{U_1} H, \dots, \mathcal{L}_{U_d} H)$ with respect to the basis given by U_+^j, U_0^j, U_-^j, $j = 1, m$. It is clear from this description that j-th block of L_k is given by

$$L_{kj} = \begin{pmatrix} \mathcal{L}_{U_k} & -\psi_{kj} & 0 \\ 0 & \mathcal{L}_{U_k} & 2\psi_{kj} \\ 0 & 0 & \mathcal{L}_{U_k} \end{pmatrix}.$$

For a C^∞ function h define $\mathcal{C}_j(h)$ by

$$\mathcal{C}_j(h) = (\psi_{1j} h, \psi_{2j} h, \dots, \psi_{dj} h),$$

and let $\mathcal{L}$ be the following map: $h \in C^\infty(X) \mapsto (\mathcal{L}_{U_1} h, \dots, \mathcal{L}_{U_d} h) \in C^\infty(X)^d$. Then the j-th block of L is given by

$$L_j = \begin{pmatrix} \mathcal{L} & -\mathcal{C}_j & 0 \\ 0 & \mathcal{L} & 2\mathcal{C}_j \\ 0 & 0 & \mathcal{L} \end{pmatrix}. \tag{4.3}$$

By denoting

$$C_j = \begin{pmatrix} 0 & -\mathcal{C}_j & 0 \\ 0 & 0 & 2\mathcal{C}_j \\ 0 & 0 & 0 \end{pmatrix}, \tag{4.4}$$

and using the notation $\overline{\mathcal{L}}$ for the operator $\left(\begin{smallmatrix} \mathcal{L} & 0 & 0 \\ 0 & \mathcal{L} & 0 \\ 0 & 0 & \mathcal{L} \end{smallmatrix}\right)$, it follows that

$$L_j = \overline{\mathcal{L}} + C_j. \tag{4.5}$$

For operators M_{lk}, that are coordinates of the operator M defined in (1.3) and (1.4), it follows from (4.1) that for $W_l = \sum_{j=1}^m w_{l0}^j U_0^j + w_{l+}^j U_+^j + w_{l-}^j U_-^j$ and $k \neq l$

expression $\mathscr{L}_{U_k} W_l - \mathscr{L}_{U_l} W_k$, has the form

$$\begin{aligned}(4.6)\quad \mathscr{L}_{U_k} W_l - \mathscr{L}_{U_l} W_k = \sum_{j=1}^{m} \Big[(\mathscr{L}_{U_k} w_{l+}^{j} - \psi_{kj} w_{l0}^{j}) U_+^j \\ + (\mathscr{L}_{U_k} w_{l0}^{j} + 2\psi_{kj} w_{l-}^{j}) U_0^j + \mathscr{L}_{U_k} w_{l-}^{j} U_-^j \Big] \\ - \sum_{j=1}^{m} \Big[(\mathscr{L}_{U_l} w_{k+}^{j} - \psi_{lj} w_{k0}^{j}) U_+^j \\ + (\mathscr{L}_{U_l} w_{k0}^{j} + 2\psi_{lj} w_{k-}^{j}) U_0^j + \mathscr{L}_{U_l} w_{k-}^{j} U_-^j \Big].\end{aligned}$$

For a d-tuple of C^∞ functions $(h_1, \dots, h_d)$ define $\mathscr{D}_j(h_1, \dots, h_d)$ by

$$\mathscr{D}_j(h_1, \dots, h_d) = (\psi_{kj} h_l - \psi_{lj} h_k)_{k,l=1,\dots,d}.$$

Let us denote by $\mathscr{M}$ the map which takes a d-tuple of C^∞ functions $(h_1, \dots, h_d)$ to $(\mathscr{L}_{U_l} h_k - \mathscr{L}_{U_k} h_l)_{k,l=1,\dots,d} \in C^\infty(X)^{d\times d}$. Then it is easy to see just directly from the definitions of $\mathscr{M}$, $\mathscr{D}_j$, $\mathscr{C}_j$, and $\mathscr{L}$ that

$$(4.7)\qquad \mathscr{M}\mathscr{C}_j + \mathscr{D}_j \mathscr{L} = 0.$$

From the definition of the ψ_{kj}'s it follows that $\mathscr{D}_j \mathscr{C}_j = 0$. It follows from (4.6) that the operator M acts in the following way on $(W_1^j, \dots, W_d^j)$

$$(4.8)\qquad M_j = \begin{pmatrix} \mathscr{M} & -\mathscr{D}_j & 0 \\ 0 & \mathscr{M} & 2\mathscr{D}_j \\ 0 & 0 & \mathscr{M} \end{pmatrix}$$

where $W_l^j = w_{l0}^{j} U_0^j + w_{l+}^{j} U_+^j + w_{l-}^{j} U_-^j$ for $l = 1, \dots, d$. The operator M is defined to be the operator which acts by M_j in each direction g^j. Let

$$D_j = \begin{pmatrix} 0 & -\mathscr{D}_j & 0 \\ 0 & 0 & 2\mathscr{D}_j \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \overline{\mathscr{M}} = \begin{pmatrix} \mathscr{M} & 0 & 0 \\ 0 & \mathscr{M} & 0 \\ 0 & 0 & \mathscr{M} \end{pmatrix}.$$

Then,

$$(4.9)\qquad M_j = \overline{\mathscr{M}} + D_j.$$

We define here few operators which appear in the proof of the subsequent proposition and will be used in its proof. Let $\mathscr{C}'_j \colon (h_1, \dots, h_d) \mapsto \sum_{k=1}^{d} \psi_{kj} h_k$ and $\mathscr{D}'_j \colon (a_{ij})_{i,j=1,\dots,d} \mapsto \sum_{k=1}^{d} \psi_{kj} (a_{k1}, \dots, a_{k_d})$. Also, let

$$(4.10)\qquad D'_j = \begin{pmatrix} 0 & 0 & 0 \\ -\mathscr{D}'_j & 0 & 0 \\ 0 & \frac{1}{2}\mathscr{D}'_j & 0 \end{pmatrix}, \qquad C'_j = \begin{pmatrix} 0 & 0 & 0 \\ -\mathscr{C}'_j & 0 & 0 \\ 0 & \frac{1}{2}\mathscr{C}'_j & 0 \end{pmatrix},$$

and define C' and D' to act by C'_j and D'_j on each $\mathfrak{g}^j$. Since $\mathscr{M}$ and $\mathscr{L}$ are 0-tame it is immediate that both M and L are 0-tame. Now we are ready to state and prove the main splitting result.

PROPOSITION 4.1. *Assume that for an action α generated by U_k, $k = 1,\dots,d$ as described in the examples in Section 1.3,* i.e., *one of the actions $\alpha_{\mathbb{R}\mathbb{R}}$, $\alpha_{\mathbb{R}\mathbb{C}}$, $\alpha_{\mathbb{C}}$, there exists an r-tame splitting for the sequence of s-tame linear operators $\mathscr{L} \to \mathscr{M}$,* i.e., *there exist r-tame linear operators $\mathscr{L}'$ and $\mathscr{M}'$ such that*

$$\mathscr{L}\mathscr{L}' + \mathscr{M}'\mathscr{M} = \mathrm{Id}$$

on $C^\infty(X)^d \cap \ker\mathscr{A}$, where $\mathscr{A}$ is the average map.

Then the exact sequence of linear operators $M \to L$ is $3s$-tame and it admits $3r$-splitting, i.e., *there exist $3r$-tame linear operators L' and M' such that*

$$LL' + M'M = \mathrm{Id}$$

on $\mathrm{Vect}^\infty(X)^d \cap \ker A$, *where*

$$A = \overline{\mathscr{A}} - CC'\overline{\mathscr{A}} - D'D\overline{\mathscr{A}}$$

and $\overline{\mathscr{A}}$ is the coordinatewise average map.

Proof. First notice that it follows from (4.5) and (4.9) and the assumptions that the sequence $M \to L$ is exact and $3s$-tame. By assumption, there exist operators $\mathscr{L}'$ and $\mathscr{M}'$ such that

$$\mathscr{L}\mathscr{L}' + \mathscr{M}'\mathscr{M} = \mathrm{Id}$$

on $C^\infty(X)^d \cap \ker\mathscr{A}$.

Recall that for a C^∞ function h, $\mathscr{C}_j(h)$ is defined by

$$\mathscr{C}_j(h) = (\psi_{1j}h, \psi_{2j}h, \dots, \psi_{dj}h)$$

and for a d-tuple of $\mathscr{C}^\infty$ functions $(h_1,\dots,h_d)$, $\mathscr{D}_j(h_1,\dots,h_d)$ is defined by

$$\mathscr{D}_j(h_1,\dots,h_d) = (\psi_{kj}h_l - \psi_{lj}h_k)_{k,l=1,\dots,d}.$$

Below we use the notation $\overline{\mathscr{L}}$, $\overline{\mathscr{M}}$, $\overline{\mathscr{L}}'$, and $\overline{\mathscr{M}}'$ to denote operators which act on $\mathfrak{g}^j$-valued maps by $\mathscr{L}$, $\mathscr{M}$, $\mathscr{L}'$, and $\mathscr{M}'$, respectively, on each component in U_+^j, U_0^j, and U_-^j direction. Recall that $L_j = \overline{\mathscr{L}} + C_j$ from (4.5), and that $M_j = \overline{\mathscr{M}} + D_j$ from (4.9). Now define operators L_j' by

$$L_j' = \begin{pmatrix} \mathscr{L}' & \mathscr{L}'\mathscr{C}_j\mathscr{L}' & -2\mathscr{L}'\mathscr{C}_j\mathscr{L}'\mathscr{C}_j\mathscr{L}' \\ 0 & \mathscr{L}' & -2\mathscr{L}'\mathscr{C}_j\mathscr{L}' \\ 0 & 0 & \mathscr{L}' \end{pmatrix}.$$

Also, define the operators M_j' for $j = 2,\dots,m$ by

$$M_j' = \begin{pmatrix} \mathscr{M}' & \mathscr{M}'\mathscr{D}_j\mathscr{M}' & -2\mathscr{M}'\mathscr{D}_j\mathscr{M}'\mathscr{D}_j\mathscr{M}' \\ 0 & \mathscr{M}' & -2\mathscr{M}'\mathscr{D}_j\mathscr{M}' \\ 0 & 0 & \mathscr{M}' \end{pmatrix}.$$

Operator M_j' makes sense on the image of M, which consists of $\mathfrak{g}^j$-valued $d \times d$ skew-symmetric matrices.

Since $\mathscr{L}'$ and $\mathscr{M}'$ are assumed to be r-tame, and C_j and D_j are obviously 0-tame, it follows immediately that L' and M' are $3r$-tame. Now from the definitions of the operators L_j, L'_j, M_j, and M'_j, we have

$$L_j L'_j + M'_j M_j = \begin{pmatrix} \mathscr{L}_j \mathscr{L}'_j + \mathscr{M}'_j \mathscr{M}_j & -G_1 & 2G_2 \\ 0 & \mathscr{L}_j \mathscr{L}'_j + \mathscr{M}'_j \mathscr{M}_j & 2G_1 \\ 0 & 0 & \mathscr{L}_j \mathscr{L}'_j + \mathscr{M}'_j \mathscr{M}_j \end{pmatrix},$$

where

$$G_1 = (\mathrm{Id} - \mathscr{L}_j \mathscr{L}'_j) C_j \mathscr{L}'_j + \mathscr{M}'_j \mathscr{D}_j (\mathrm{Id} - \mathscr{M}'_j \mathscr{M}_j)$$
$$G_2 = (\mathrm{Id} - \mathscr{L}_j \mathscr{L}'_j) C_j \mathscr{L}'_j C_j \mathscr{L}'_j + \mathscr{M}'_j \mathscr{D}_j \mathscr{M}'_j \mathscr{D}_j (\mathrm{Id} - \mathscr{M}'_j \mathscr{M}_j).$$

It is easy to see that $\mathscr{C}_j \mathscr{C}'_j + \mathscr{D}'_j \mathscr{D}_j = \mathrm{Id}$. Thus, we may apply $L_j L'_j + M'_j M_j$ above to a vector field which is coordinatewise first composed with the trivial operator $\mathscr{C}_j \mathscr{C}'_j + \mathscr{D}'_j \mathscr{D}_j$. We see that we may use the assumption $\mathscr{L}_j \mathscr{L}'_j + \mathscr{M}'_j \mathscr{M}_j = \mathrm{Id}$ for any maps whose average map is in the image of $\mathscr{C}_j$ (*i.e.*, on maps for which $\mathscr{D}'_j \mathscr{D}_j$ part is trivial). Under this condition simple computation using (4.7) gives that G_1 and G_2 are trivial. On the other hand on vector fields this condition precisely means that one needs to correct the average map by the value of the operator

$$\begin{pmatrix} \mathscr{C}_j \mathscr{C}'_j & 0 & 0 \\ 0 & \mathrm{Id} & 0 \\ 0 & 0 & \mathscr{D}'_j \mathscr{D}_j \end{pmatrix}$$

applied to the average map of the vector field in order to have $L_j L'_j + M'_j M_j$ equal to the identity. The above operator can also be described via D' and C' as $C_j C'_j + D'_j D_j$. Thus,

$$L_j L'_j + M'_j M_j = \mathrm{Id}$$

on the kernel of the map $\overline{\mathscr{A}}_j - C_j C'_j \overline{\mathscr{A}}_j - D'_j D_j \overline{\mathscr{A}}_j$, where $\overline{\mathscr{A}}_j$ denotes the map which assigns to a given d-tuple of vector fields its coordinatewise averages in the component $\mathfrak{g}^j$, *i.e.*, to each vector field in the d-tuple it assigns the triple average in the directions U^j_+, U^j_0 and U^j_-. Maps C'_j and D'_j are defined in (4.10). Thus, by denoting

$$A_j = \overline{\mathscr{A}}_j - C_j C'_j \overline{\mathscr{A}}_j - D'_j D_j \overline{\mathscr{A}}_j,$$

we have the required splitting on $\ker A$,

$$L_j L'_j + M'_j M_j = \mathrm{Id}$$

where A is defined to be equal to A_j in each direction $\mathfrak{g}^j$, $j = 1, \dots, m$, just as L' and M' are defined to be L'_j and M'_j in each $\mathfrak{g}^j$. □

Corollary 3.7 asserts that assumptions of Proposition 4.1 hold for the action $\alpha_{\mathbb{R}\mathbb{R}}$ of Example 1. This gives the principal ingredient for carrying out the iteration scheme.

COROLLARY 4.2. *The sequence* $L \to M$ *in* (1.4) *of* 1*-tame operators for the action* $\alpha_{\mathbb{R}\mathbb{R}}$ *of Example 1 has an* r_0*-tame splitting on the space* $\mathrm{Vect}^\infty(X)^d \cap \ker A$*, where* $A = \overline{\mathscr{A}} - CC'\overline{\mathscr{A}} - D'D\overline{\mathscr{A}}$ *and* r_0 *depends only on the lattice* Γ.

REMARK 4.3. While it is obvious from the definition that $\ker A$ has finite codimension, on the surface it looks that this codimension is greater than d. In the next section we will show how to reduce the required number of parameters.

5. DETAILED STUDY OF THE AVERAGE MAP A: CONSTRUCTION OF COORDINATE CHANGES

Now we will carry out Step 4(i) and (ii) of the general scheme. Notice that part (ii) of this step in our setting involves conjugating the unipotent action inside the group G since unipotent homogeneous actions in our cases are all conjugate to the standard ones. Thus standard perturbations produce actions isomorphic to the original ones. Part (iii) corresponds to taking parametric families in the Main Theorem 1.2 with the number of parameters exactly equal to the codimension of unipotent homogeneous actions among all homogeneous actions of $\mathbb{R}^d$.

Recall that

$$\mathscr{C}'_j : (h_1, \dots, h_d) \mapsto \sum_{k=1}^{d} \psi_{kj} h_k$$

and that

$$\mathscr{D}'_j : (a_{ij})_{i,j=1,\dots,d} \mapsto \sum_{k=1}^{d} \psi_{kj}(a_{k1}, \dots, a_{kd}).$$

It is immediate from definitions that $\mathscr{C}_j\mathscr{C}'_j + \mathscr{D}'_j\mathscr{D}_j = \mathrm{Id}$. This implies

$$C_j C'_j + D'_j D_j = \begin{pmatrix} \mathscr{C}_j\mathscr{C}'_j & 0 & 0 \\ 0 & \mathrm{Id} & 0 \\ 0 & 0 & \mathscr{D}'_j\mathscr{D}_j \end{pmatrix}.$$

Thus, $\ker A$ in $\mathrm{Vect}^\infty(X)^d$ coincides with

$$\begin{aligned} &\{W = (W_1, \dots, W_d) \mid (\mathscr{A} - \mathscr{C}\mathscr{C}'\mathscr{A})(W_+) = 0,\ (\mathscr{A} - \mathscr{D}'\mathscr{D}\mathscr{A})(W_-) = 0\} \\ &\qquad = \{W = (W_1, \dots, W_d) \mid \mathscr{D}'\mathscr{D}\mathscr{A}(W_+) = 0,\ \mathscr{C}\mathscr{C}'\mathscr{A}(W_-) = 0\}. \end{aligned}$$

Here, W_+ denotes the d-tuple of functions which represent components of the d-tuple of vector fields W in the direction U_+^j and W_- denotes the d-tuple of components of W in direction U_-^j.

In fact, because of the definition of ψ_{kl} (4.2) it is not difficult to see that $\mathscr{D}'\mathscr{D}\mathscr{A}(W_+)$ involves the averages of W in all the directions U_l for which $\psi_{kl} = 0$ and that $\mathscr{C}\mathscr{C}'\mathscr{A}(W_-)$ involves all the averages of W in the direction of V_l whenever $\psi_{kl} \neq 0$. In particular, in the notation of Theorem 1.2 we may write

$$\Phi(W)(\lambda) \cdot V := \mathscr{C}\mathscr{C}'\mathscr{A}(W_-).$$

On the other hand, the following proposition shows that $\mathcal{D}'\mathcal{D}\mathcal{A}(W_+)$ can be assumed to be quadratically small by making a good choice of generating vector fields for the action.

PROPOSITION 5.1. *Let $\bar{U} = (\bar{U}_1, \dots, \bar{U}_d) \in \mathrm{Vect}^\infty(X)^d$, $\bar{U} = U + \overline{W}$, generate an abelian action $\tilde{\alpha}$. Then there exists a linear automorphism T of $\mathbb{R}^d$ such that $\tilde{U} := T\bar{U} = U + W$ satisfies*

$$W = W_A + \mathcal{C}'\mathcal{C}\mathcal{A}(\overline{W}_-) + W_e,$$

where $W_A \in \ker A$, W_e is a constant, and the following estimates hold:

(i) $\|T - I\| \le \|\overline{W}\|_0$,
(ii) $\|W\|_r \le C(1 + \|\overline{W}\|_0)\|\overline{W}\|_r$, $r \ge 0$,
(iii) $\|W_A\|_r \le C(1 + \|\overline{W}\|_0)\|\overline{W}\|_r$, $r \ge 0$,
(iv) $\|W_e\| \le C\|\overline{W}\|_0^2$.

REMARK 5.2. The d-tuple of vector fields $\tilde{U}$ represents a simple change of basis for the action $\tilde{\alpha}$.

Proof. Let $k \ne l$ and a_{kl} denote the average of the component of $\overline{W}_k$ in the direction of U_l.

Let

$$B = \begin{pmatrix} 1 & a_{12} \\ a_{21} & 1 \end{pmatrix} \qquad \text{or} \qquad B = \begin{pmatrix} 1 & a_{12} & a_{13} \\ a_{21} & 1 & a_{23} \\ a_{31} & a_{32} & 1 \end{pmatrix}$$

accordingly for $d = 2$ and $d = 3$. Then $\bar{U} = U + \overline{W} = BU + W'$, where $W' = \overline{W} + (I - B)U$. By definition of B, $\mathcal{D}'\mathcal{D}\mathcal{A}(W'_+) = 0$, $\mathcal{C}'\mathcal{C}\mathcal{A}(W'_-) = \mathcal{C}'\mathcal{C}\mathcal{A}(\overline{W}_-)$ and $\|W'\|_r \le C\|\overline{W}\|_r$ for $r \ge 0$.

Let $T = B^{-1}$. Then it is obvious from the definition of B that the estimate (i) holds for the norm of $T - I$.

Let $\tilde{U} = T\bar{U} = U + TW'$, let $W = TW'$ and $W'' = (T - I)W'$. Then $W = W' + W''$. This implies that $\|W''\|_r = \|(T - I)W'\|_r \le C\|\overline{W}\|_0\|\overline{W}\|_r$ for $r \ge 0$, from (i). This implies that

$$\|W\|_r \le \|W'\|_r + \|(T - I)W'\|_r \le C(1 + \|\overline{W}\|_0)\|\overline{W}\|_r.$$

As for the operators $\mathcal{D}'\mathcal{D}\mathcal{A}$ and $\mathcal{C}'\mathcal{C}\mathcal{A}$, the following are immediate from the definition of W:

$$\mathcal{D}'\mathcal{D}\mathcal{A}(W_+) = \mathcal{D}'\mathcal{D}\mathcal{A}(W'_+) + \mathcal{D}'\mathcal{D}\mathcal{A}(W''_+) = \mathcal{D}'\mathcal{D}\mathcal{A}(W''_+)$$

and

$$\mathcal{C}'\mathcal{C}\mathcal{A}(W_-) = \mathcal{C}'\mathcal{C}\mathcal{A}(W'_-) + \mathcal{C}'\mathcal{C}\mathcal{A}(W''_+) = \mathcal{C}'\mathcal{C}\mathcal{A}(\overline{W}_-) + \mathcal{C}'\mathcal{C}\mathcal{A}(W''_+).$$

Let $W_A := W - \mathcal{D}'\mathcal{D}\mathcal{A}(W_+) - \mathcal{C}'\mathcal{C}\mathcal{A}(W_+)$. Then clearly $W_A \in \ker A$. Moreover, since

$$\mathcal{D}'\mathcal{D}\mathcal{A}(W_+) + \mathcal{C}'\mathcal{C}\mathcal{A}(W_+) = \mathcal{D}'\mathcal{D}\mathcal{A}(W''_+) + \mathcal{C}'\mathcal{C}\mathcal{A}(\overline{W}_-) + \mathcal{C}'\mathcal{C}\mathcal{A}(W''_+),$$

by denoting $W_e = \mathcal{D}'\mathcal{D}\mathcal{A}(W''_+) + \mathcal{C}'\mathcal{C}\mathcal{A}(W''_+)$, we obtain

$$W = W_A + \mathcal{C}'\mathcal{C}\mathcal{A}(\overline{W}_-) + W_e$$

where W_e is constant and satisfies $\|W_e\| \le C\|\overline{W}\|_0^2$. □

6. ITERATIVE STEP

The following is an immediate corollary of the classical Implicit-Function Theorem.

LEMMA 6.1. *There exists an open ball $\mathcal{O} = \mathcal{O}(\mathrm{Id}, R)$ of radius R with $R < 1$ centered at the identity in the space $C^1(\mathbb{R}^d, \mathbb{R}^d)$, there exists a neighborhood $\mathcal{U}$ of $0 \in \mathbb{R}^d$ and a C^1 map $\Psi\colon \mathcal{O} \to \mathcal{U}$ such that for every $G \in \mathcal{O}$, $G(\Psi(G)) = 0$. Moreover, Ψ is Lipschitz on $\mathcal{O}$ and every $G \in \mathcal{O}$ is a diffeomorphism.*

To simplify notations in this section we will denote the action $\alpha_{\mathbb{R}\mathbb{R}}$ of Example 1 simply by α. Recall that Φ denotes the mapping which takes a 2-parameter family of perturbations of α generated by $U_1 + W_1(\lambda)$ and $U_2 + W_2(\lambda)$ where $U_i := U_+^i$, $i = 1,2$, to the pair of real numbers (μ_1, μ_2) where μ_i is the average of $W_i(\lambda)$ in the U_-^i direction with respect to the Haar measure, for $i = 1$ or 2.

Recall that $\|\Phi\|_{(r)}$ denotes the usual C^r norm of Φ as a map from $\mathbb{R}^2$ to $\mathbb{R}^2$, and that $\|W\|_{0,(r)}$ stands the supremum of the C^r norms of W in the λ variable over X. As before, we reserve the notation $\|W(\lambda)\|_r$ for the usual C^r norm on X of the vector field $W(\lambda)$ for a fixed parameter λ.

In the following proposition we use indices n and $n+1$ pertaining to the iterative step of the construction of conjugacy. This is done for convenience of the reader since these same notations are used in the convergence proof in the next section. What is in fact proved here is that, given a family of perturbations of the action α satisfying a certain set of conditions, one constructs a conjugacy such that the new family of actions satisfies another set of conditions. The letter C is used generically to denote various constants that depend only on the action α; indices such as r indicate additional dependence of the corresponding parameter.

PROPOSITION 6.2. *There exist constants r_0 and $\bar{C}$ such that the following holds:*

For a family $\tilde{U}^n(\lambda) = U + W^n(\lambda)$, $\lambda \in \mathbb{R}^2$, of perturbations of U, for all λ in a fixed closed ball B, for a natural number r and $t > 0$, assume

(1) $\|W^n(\lambda)\|_0 \le \varepsilon_n < 1$,
(2) $\Phi^n := \Phi(W^n) \in \mathcal{O}$ *has a zero at* λ_n,
(3) $W^n(\lambda)$ *is* C^2 *in* λ *and* $\|W^n(\lambda)\|_{0,(2)} \le K_n$,
(4) $\|W^n(\lambda)\|_{r_0+r} \le \delta_{r,n}$,
(5) $t^{r_0} \varepsilon_n^{1-1/(r+r_0)} \delta_{r,n}^{1/(r+r_0)} < \bar{C}$.

Then there exist a linear map $T_n\colon \mathbb{R}^2 \to \mathbb{R}^2$ and a C^∞ vector field H^n on X, such that for $h^n := \exp H^n$, $\tilde{U}^{n+1}(\lambda) := (h^n)_*(T_n \tilde{U}^n(\lambda))$, $\lambda \in B$, and $W^{n+1}(\lambda) := \tilde{U}^{n+1}(\lambda) - U$:

(a) $\|H^n\|_r \le C_r t^{2r_0} \|W^n(\lambda_n)\|_r$

(b) $\|W^{n+1}(\lambda)\|_0 \le K_n\|\lambda-\lambda_n\| + \mathrm{Err}_{n+1}(t,r)$ *where*

$$\mathrm{Err}_{n+1}(t,r) := C\varepsilon_n^2 + C\,\delta_{r,n}^{(r_0+1)/(r_0+r)}\,\varepsilon_n^{2-(r_0+1)/(r_0+r)} + C_r\,t^{-r}\,\delta_{r,n} + C_{r_0}\,t^{2r_0}\,\varepsilon_n^{2-1/(r_0+r)}\,\delta_{r,n}^{1/(r_0+r)} + C_{r_0}\,t^{2r_0}\,\varepsilon_n^{3-1/(r_0+r)}\,\delta_{r,n}^{2/(r_0+r)}$$

(c) $\|W^{n+1}(\lambda)\|_{r_0+r} \le C_r t^{2r_0}\delta_{r,n} =: \delta_{r,n+1}$

(d) $\Phi^{n+1} := \Phi(W^{n+1})$ *satisfies*

$$\|\Phi^{n+1}-\Phi^n\|_{(0)} \le \mathrm{Err}_{n+1}(t,r)$$
$$\|\Phi^{n+1}-\Phi^n\|_{(1)} \le K_n t^{r_0}\varepsilon_n + \mathrm{Err}_{n+1}(t,r).$$

If Φ^{n+1} is in $\mathscr{O}$, then it has a zero at $\lambda_{n+1}\in B$ which satisfies

$$\|\lambda_{n+1}-\lambda_n\| \le C\,\mathrm{Err}_{n+1}(t,r) + CK_n\big(K_n t^{r_0}\varepsilon_n + \mathrm{Err}_{n+1}(t,r)\big)^2$$

(e) $W^{n+1}(\lambda)$ *is C^2 in λ and*

$$\|W^{n+1}(\lambda)\|_{0,(2)} \le \big(1 + Ct^{r_0}\varepsilon_n^{1-1/(r+r_0)}\,\delta_{r,n}^{1/(r+r_0)}\big)K_n =: K_{n+1}(t,r)$$

(f) $\|T_n - \mathrm{Id}\| \le \varepsilon_n$.

REMARK 6.3. The proof of this proposition relies only on Proposition 5.1 and Corollary 4.2 which hold for Example 1, and is otherwise general.

Proof. Within this proof we use the notation $\overline{W}^n(\lambda) := \tilde{U}^n(\lambda) - U$ (instead of $W^n(\lambda)$) for the perturbation in the n-th step, in order to be consistent with the notation of Proposition 5.1.

By Proposition 5.1 and assumption (1): given $\tilde{U}^n(\lambda) = U + \overline{W}^n(\lambda)$ there exists $T_n(\lambda)$ such that $T_n(\lambda)\tilde{U}^n(\lambda) = U + W^n(\lambda)$ and

$$W^n(\lambda) = W_A^n(\lambda) + \Phi^n(\lambda)\cdot V + W_e^n(\lambda),$$

where we used the fact that $\mathscr{C}'\mathscr{C}\mathscr{A}(W(\lambda)_-) = \Phi(W)(\lambda)\cdot V$ and we used the notation $\Phi^n := \Phi(\tilde{U}^n)$.

From Proposition 5.1, it follows that for all λ and all $s \ge 0$ we have

$$\|W^n(\lambda)\|_s \le C\|\overline{W}^n(\lambda)\|_s \tag{6.1}$$

$$\|W_A^n(\lambda)\|_s \le C\|\overline{W}^n(\lambda)\|_s. \tag{6.2}$$

Due to loss of regularity which appears in estimates of Corollary 4.2 (recall that the operators which appear in Corollary 4.2 are r_0-tame) it is customary to use the smoothing operators described in Section 2.3. Then

$$W^n(\lambda) = S_t W_A^n(\lambda) + (I-S_t)W_A^n(\lambda) + \Phi^n(\lambda)\cdot V + W_e^n(\lambda),$$

so $W_A^n \in \ker A$.

By the choice of smoothing operators if $W_A^n(\lambda) \in \ker A$, then $S_t^n W_A(\lambda) \in \ker A$. Thus, we can apply Corollary 4.2 to $S_t W_A^n(\lambda)$ to obtain

$$H^n(\lambda) := L' S_t W_A^n(\lambda),$$

which satisfies

$$W^n(\lambda) = L(H^n(\lambda)) + M'M(S_t W_A^n(\lambda)) + (I-S_t)W_A^n(\lambda) + \Phi^n(\lambda)\cdot V + W_e^n(\lambda). \tag{6.3}$$

From Corollary 4.2, for all $r \geq 0$ and all $s \geq 0$ there exist constants C_r and $C = C_{r,s}$ such that

$$\|H^n(\lambda)\|_r \leq C_r \|S_t W_A^n(\lambda)\|_{r_0+r} \leq C_r t^{r_0} \|W_A^n(\lambda)\|_r \leq C_r t^{r_0} \|\overline{W}^n(\lambda)\|_r \tag{6.4}$$

and

$$\|(I-S_t)W_A^n(\lambda)\|_s \leq Ct^{-r}\|W_A^n(\lambda)\|_{s+r} \leq Ct^{-r}\|\overline{W}^n(\lambda)\|_{s+r}.$$

From Proposition 5.1, $W_e^n(\lambda)$ is a constant vector field depending (only) on λ, and

$$\|W_e^n(\lambda)\| \leq C\|\overline{W}^n(\lambda)\|_0^2.$$

Now by the assumption (2), there exists λ_n such that $\Phi^n(\lambda_n) = 0$. Then, at λ_n we have

$$M(W_A^n(\lambda_n)) = M(W^n(\lambda_n)) + M(W_e^n(\lambda_n)).$$

From Corollary 4.2, the operator M' is r_0-tame and M is 0-tame. Therefore, using Lemma 1.1, for any $r > 0$, and estimates (6.1) and (6.2), the following estimates hold:

$$\begin{aligned}
\|M'M(S_t W_A^n(\lambda_n))\|_0 &\leq \|M(S_t W_A^n(\lambda_n))\|_{r_0} \\
&\leq \|M(W_A^n(\lambda_n))\|_{r_0} + \|M((I-S_t)W_A^n(\lambda_n))\|_{r_0} \\
&\leq \|M(W^n(\lambda_n))\|_{r_0} + \|M(W_e^n(\lambda_n))\|_0 + \|M((I-S_t)W_A^n(\lambda_n))\|_{r_0} \\
&\leq C\|W^n(\lambda_n)\|_{r_0+1}\|W^n(\lambda_n)\|_0 + C\|W_e^n(\lambda_n)\|_0 + \|(I-S_t)W_A^n(\lambda_n)\|_{r_0} \\
&\leq C\|W^n(\lambda_n)\|_{r_0+1}\|W^n(\lambda_n)\|_0 + C\|\overline{W}^n(\lambda_n)\|_0^2 + C_r t^{-r}\|W_A^n(\lambda_n)\|_{r_0+r} \\
&\leq C\|\overline{W}^n(\lambda_n)\|_{r_0+1}\|\overline{W}^n(\lambda_n)\|_0 + C\|\overline{W}^n(\lambda_n)\|_0^2 + C_r t^{-r}\|\overline{W}^n(\lambda_n)\|_{r_0+r}.
\end{aligned}$$

Set

$$E_1^n(\lambda_n) = W^n(\lambda_n) - L(H^n(\lambda_n)) = M'M(S_t W_A^n(\lambda_n)) + (I-S_t)W_A^n(\lambda_n) + W_e^n(\lambda_n).$$

Then,

$$\begin{aligned}
\|E_1^n(\lambda_n)\|_0 \leq Ct^{-r}\big(1+\|\overline{W}^n(\lambda_n)\|_0\big)\|\overline{W}^n(\lambda_n)\|_r + 2C\|\overline{W}^n(\lambda_n)\|_0^2 \\
+ C\big(1+\|\overline{W}^n\|_0\big)^2\|\overline{W}^n(\lambda_n)\|_{r_0+1}\|\overline{W}^n(\lambda_n)\|_0 \\
+ C_r t^{-r}\|\overline{W}^n(\lambda_n)\|_{r_0+r}.
\end{aligned} \tag{6.5}$$

Define $H^n := H^n(\lambda_n)$. Let $h^n = \exp H^n$. Due to (6.4), the map h^n is close to the identity to the order of the initial perturbation. So by choosing sufficiently small perturbations initially, one can make sure that h^n is a diffeomorphism of X. For this purpose, by the estimate (6.4) and interpolation estimates we have

$$\|H^n\|_1 \leq C\, t^{r_0}\, \varepsilon_n^{1-1/(r+r_0)}\, \delta_{r,n}^{1/(r+r_0)}.$$

So, it suffices to assume $t^{r_0}\, \varepsilon_n^{1-1/(r+r_0)}\, \delta_{r,n}^{1/(r+r_0)} < \bar{C}$, where $\bar{C}$ is the radius of the C^1 neighborhood of Id in which all the maps are invertible. This is assumption (5). Thus, we may define

$$\tilde{U}^{n+1}(\lambda) := h_*^n\big(T_n(\lambda)\tilde{U}^n(\lambda)\big) = h_*^n\big(T_n(\lambda)(U+\overline{W}^n(\lambda)\big) = h_*^n\big(U+W^n(\lambda)\big),$$

and define $\alpha^{n+1}(\lambda)$ to be the family of actions generated by the commuting d-tuple $\tilde{U}^{n+1}(\lambda)$. We show now that this family of actions, with parameter λ satisfies the statements of the theorem.

Part (a) is immediate from (6.4). To prove the rest, we first linearize $h_*^n(U + W^n(\lambda))$. After linearization,

$$\tilde{U}^{n+1}(\lambda) := h_*^n(U + W^n(\lambda)) = U + W^n(\lambda) - L(H^n) + E_0^n(\lambda) + E_L^n(\lambda), \tag{6.6}$$

where $L(H^n) = (\mathscr{L}_{U_1} H^n, \dots, \mathscr{L}_{U_d} H^n)$ as before, $E_0^n(\lambda) = [H, W^n(\lambda)]$, and $E_L^n(\lambda)$ is the error of linearization. Thus, the new error is

$$W^{n+1}(\lambda) = W^n(\lambda) - W^n(\lambda_n) + E_1^n(\lambda_n) + E_0^n(\lambda) + E_L^n(\lambda). \tag{6.7}$$

To estimate norms of the new error we will make use of the interpolation estimates,

$$\|\overline{W}^n(\lambda)\|_p \le C \|\overline{W}^n(\lambda)\|_0^{1-p/s} \|\overline{W}^n(\lambda)\|_s^{p/s}$$

for $p, s > 0$. By (6.4), the interpolation estimates, and assumptions (1) and (4), we have

$$\begin{aligned}
\|E_0^n(\lambda)\|_0 &= \|[H^n, W^n(\lambda)]\|_0 \\
&\le C \|H^n\|_0 \|W^n(\lambda)\|_1 + C \|H^n\|_1 \|W^n(\lambda)\|_0 \\
&\le C_{r_0} t^{r_0} \big(\|W^n(\lambda_n)\|_0 \|W^n(\lambda)\|_1 + \|W^n(\lambda_n)\|_1 \|W^n(\lambda)\|_0 \big) \\
&\le C_{r_0} t^{r_0} \big(\|\overline{W}^n(\lambda_n)\|_0 \|\overline{W}^n(\lambda)\|_1 + \|\overline{W}^n(\lambda_n)\|_1 \|\overline{W}^n(\lambda)\|_0 \big) \\
&\le C_{r_0} t^{r_0} \Big(\|\overline{W}^n(\lambda_n)\|_0 \|\overline{W}^n(\lambda)\|_0^{1-1/(r_0+r)} \|\overline{W}^n(\lambda)\|_{r_0+r}^{1/(r_0+r)} \\
&\qquad\qquad + \|\overline{W}^n(\lambda_n)\|_{r_0+r}^{1/(r_0+r)} \|\overline{W}^n(\lambda_n)\|_0^{1-1/(r_0+r)} \|\overline{W}^n(\lambda)\|_0 \Big) \\
&\le C_{r_0} t^{r_0} \varepsilon_n^{2-1/(r_0+r)} \delta_{r,n}^{1/(r_0+r)}.
\end{aligned} \tag{6.8}$$

Using the assumption that $\|\overline{W}^n\|_0 \le \varepsilon_n < 1$, equations (6.1) and (6.4), the interpolation estimates, and assumptions (1) and (4), we have

$$\begin{aligned}
\|E_L^n(\lambda)\|_0 &= \big\| h_*^n(U + W^n(\lambda)) - (U + W^n(\lambda)) - [H^n, U + W^n(\lambda)] \big\|_0 \\
&\le C_r \|H^n\|_0 \|H^n\|_1 \|W^n(\lambda)\|_1 \\
&\le C_{r_0} t^{2r_0} \|W^n(\lambda_n)\|_0 \|W^n(\lambda_n)\|_1 \|W^n(\lambda)\|_1 \\
&\le C_{r_0} t^{2r_0} \|\overline{W}^n(\lambda_n)\|_0 \|\overline{W}^n(\lambda_n)\|_1 \|\overline{W}^n(\lambda)\|_1 \\
&\le C_{r_0} t^{2r_0} \|\overline{W}^n(\lambda)\|_0^{3-1/(r_0+r)} \|\overline{W}^n(\lambda)\|_{r_0+r}^{2/(r_0+r)} \\
&\le C_{r_0} t^{2r_0} \varepsilon_n^{3-1/(r_0+r)} \delta_{r,n}^{2/(r_0+r)}.
\end{aligned} \tag{6.9}$$

By applying interpolation estimates to the estimate (6.5) for $E_1(\lambda_n)$, assumptions (1) and (4), and putting all the estimates together in (6.7), we obtain

$$\begin{aligned}
\|W^{n+1}(\lambda)\|_0 \le K_n \|\lambda - \lambda_n\| &+ C\varepsilon_n^2 + C\delta_{r,n}^{(r_0+1)/(r_0+r)} \varepsilon_n^{2-(r_0+1)/(r_0+r)} \\
&+ C_r t^{-r} \delta_{r,n} + C_{r_0} t^{2r_0} \varepsilon_n^{2-1/(r_0+r)} \delta_{r,n}^{1/(r_0+r)} + C_{r_0} t^{2r_0} \varepsilon_n^{3-1/(r_0+r)} \delta_{r,n}^{2/(r_0+r)}
\end{aligned}$$

where we used the fact that $\|W^n(\lambda) - W^n(\lambda_n)\|_0 \le K_n \|\lambda - \lambda_n\|$ by the bound in assumption (3).

For the C^{r_0+r} norm of the new error we only need to have a "linear" bound with respect to the corresponding norm of the old error. Therefore we may use (6.6) for this purpose, and by (6.4), (6.9), (6.8), and (6.1), for any $s \geq 0$ it follows that

$$\|W^{n+1}(\lambda)\|_s = \|W^n(\lambda) - L(H^n) + E_0^n(\lambda) + E_L^n(\lambda)\|_s \leq C_s t^{2r_0} \|W^n(\lambda)\|_s.$$

Therefore,

$$\|W^{n+1}(\lambda)\|_{r_0+r} \leq C_r t^{2r_0} \delta_{r,n}.$$

The statement (e) follows from (6.4) and

$$\begin{aligned}\left\|\partial^i_{\lambda_j} h^n_* \tilde{U}^n\right\|_0 \leq \left\|h^n_* \partial^i_{\lambda_j} \tilde{U}^n\right\|_0 \leq (1 + C\|H^n\|_1)\left\|\partial^i_{\lambda_j} \tilde{U}^n\right\|_0 \\ \leq \left(1 + C\, t^{r_0}\, \varepsilon_n^{1-1/(r+r_0)}\, \delta_{r,n}^{1/(r+r_0)}\right) K_n\end{aligned}$$

for any $i, j \in \{1,2\}$.

Proof of (d): by (6.7)

$$\Phi^{n+1}(\lambda) - \Phi^n(\lambda) = \Phi(E_0^n(\lambda) + E_L^n(\lambda)).$$

This implies, by (6.8) and (6.9), that $\|\Phi^{n+1} - \Phi^n\|_0 \leq \mathrm{Err}_{n+1}(t,r)$. Also by (6.4) and assumption (3), we have

$$\begin{aligned}\|\partial_{\lambda_j} \Phi^{n+1} - \partial_{\lambda_j} \Phi^n\|_{(0)} \leq \|\partial_{\lambda_j} \Phi((Dh^n - \mathrm{Id}) W^n(\lambda))\|_{(0)} \\ \leq \|\partial_{\lambda_j} W^n(\lambda)\|_0 \|H^n\|_0 \leq K_n t^{r_0} \varepsilon_n\end{aligned}$$

for any $j \in \{1,2\}$. This implies that

$$\|\Phi^{n+1} - \Phi^n\|_{(1)} \leq K_n t^{r_0} \varepsilon_n + \mathrm{Err}_{n+1}(t,r).$$

By Lemma 6.1 if Φ^{n+1} is in $\mathcal{O}$ then it has a zero λ_{n+1} and

$$\|\lambda_{n+1} - \lambda_n\| \leq \|\Phi^{n+1} - \Phi^n\|_{(1)}.$$

Thus, we have $\|\lambda_{n+1} - \lambda_n\| \leq K_n t^{r_0} \varepsilon_n + \mathrm{Err}_{n+1}(t,r)$. But, we can obtain a better estimate for $\|\lambda_{n+1} - \lambda_n\|$. Namely, since $\|\Phi^{n+1} - \Phi^n\|_0 \leq \mathrm{Err}_{n+1}(t,r)$, it easily follows that $\|\Phi^n(\lambda_{n+1})\| \leq \mathrm{Err}_{n+1}(t,r)$. Now by using Taylor expansion of Φ^n about λ_n we have

$$\begin{aligned}\|\lambda_{n+1} - \lambda_n\| &\leq C\|\Phi^n(\lambda_{n+1})\| + CK_n\left(K_n t^{r_0} \varepsilon_n + \mathrm{Err}_{n+1}(t,r)\right)^2 \\ &\leq C\,\mathrm{Err}_{n+1}(t,r) + CK_n\left(K_n t^{r_0} \varepsilon_n + \mathrm{Err}_{n+1}(t,r)\right)^2.\end{aligned}$$

Here we used the assumption that $\Phi^n \in \mathcal{O}$, which is a small neighborhood of Id in the C^1 norm, so both the derivative of Φ^n and its inverse have a universal bound (depending only on the constant R of the Lemma 6.1), we labeled this constant here C. □

7. CONVERGENCE

In this section we assume α to be any of the actions $\alpha_{\mathbb{R}\mathbb{R}}$, $\alpha_{\mathbb{R}\mathbb{C}}$ or $\alpha_{\mathbb{C}}$ in the three examples in Section 1.3. Assume that Proposition 6.2 holds for α, with the obvious replacement of $\mathbb{R}^2$ by $\mathbb{R}^3$ in the case of $\alpha_{\mathbb{R}\mathbb{C}}$. In what follows we set an iterative scheme and show the convergence of the process to a smooth conjugacy between the initial perturbation and α up to a coordinate change.

Let R be the constant from Lemma 6.1, and let C' denote the Lipschitz constant for the map Ψ from Lemma 6.1. Recall that we denoted by $\bar{C}$ the constant for which differentiable maps in the C^1 neighborhood of Id of radius $\bar{C}$ are diffeomorphisms.

For a start, let $\varepsilon_0 < \min\{R/2, RC', \bar{C}^2, 1\}$. Later in the estimates, the initial bound for ε_0 will be made even smaller depending on the constants which appear in the estimates of the Proposition 6.2.

Recall that $r_0 \geq 1$ in Proposition 6.2 is a fixed positive number which depends only on the dimension of X and spectral gap, *i.e.*, on the lattice Γ. Set $\bar{r} = 8r_0 + 4$.

Let $\tilde{\alpha}^0(\lambda)$ be a family of perturbations of α, $\lambda \in B(0, \varepsilon_0) = \{x \in \mathbb{R}^2 | \|x\| \leq \varepsilon_0\}$, such that for its generating commuting d-tuple of vector fields $\tilde{U}^0(\lambda) = U + W^0(\lambda)$ the following hold:

1. $\|W^0(\lambda)\|_0 \leq \varepsilon_0$
2. If $\Phi^0 := \Phi(W^0)$, then $\|\Phi^0 - \mathrm{Id}\|_{(1)} < \varepsilon_0/(2C')$. Then $\Phi^0 \in \mathscr{O}$ and by Lemma 6.1 it has a zero λ_0 such that $\|\lambda_0\| < \varepsilon_0/2$.
3. $W^0(\lambda)$ is C^2 in λ and $\|W^0(\lambda)\|_{0,(2)} \leq K_0$.
4. $\|W^0(\lambda)\|_{r_0+\bar{r}} \leq \delta_{\bar{r},0} := \varepsilon_0^{-1}$
5. By defining $t_0 = \varepsilon_0^{-1/(3r_0+1)}$, it is now easy to check that

$$t_0^{r_0} \varepsilon_0^{1-1/(\bar{r}+r_0)} \delta_{\bar{r},0}^{1/(\bar{r}+r_0)} < \bar{C}.$$

By applying Proposition 6.2, one performs the first step of iteration and constructs the new family of perturbations generated by the d-tuple of commuting vector fields $\tilde{U}^1(\lambda) = U + W^1(\lambda)$.

Assume that at the n-th step of iteration, $\tilde{U}^n(\lambda) = U + W^n(\lambda)$ is a generating d-tuple of vector fields for a perturbation $\tilde{\alpha}^n(\lambda)$ of α satisfying the assumptions of Proposition 6.2. Let $\tilde{U}^{n+1}(\lambda) = U + W^{n+1}(\lambda)$ be the perturbation obtained after Proposition 6.2 is applied to $\tilde{U}^n(\lambda) = U + W^n(\lambda)$. This gives an infinite sequence of perturbations providing that Proposition 6.2 can be applied at every step.

Now we show that by choosing ε_0 possibly even smaller, and by choosing t_n and $\delta_{\bar{r},n}$ at every step in a convenient way, the Proposition 6.2 indeed applies at every step, and moreover the infinite sequence of actions obtained in this way converges as well as the sequence of conjugacies h^n obtained by successive application of Proposition 6.2. Naturally, the proof proceeds by induction. We will fix the value of parameter r at $\bar{r} = 8r_0 + 4$ throughout the induction process. The value of t will depend on n as follows:

$$\varepsilon_n = \varepsilon_0^{(4/3)^n}, \qquad t_n = \varepsilon_n^{-1/(3r_0+1)}, \qquad \delta_{\bar{r},n} = \varepsilon_n^{-1}.$$

Condition (4) of Proposition 6.2 is the easiest to check: for all n we have

$$\|W^{n+1}\|_{r_0+\bar r} \le C_{\bar r}\, t_n^{2r_0}\,\delta_{\bar r,n} = C_{\bar r}\,\varepsilon_n^{-(2+3r_0)/(3r_0+1)} < \varepsilon_n^{-4/3} = \delta_{\bar r,n+1}.$$

Also, condition (5) follows immediately given the assumptions on $\bar r$ and ε_0:

$$t_n^{r_0}\,\varepsilon_n^{1-1/(\bar r+r_0)}\,\delta_{\bar r,n}^{1/(\bar r+r_0)} \le \varepsilon_n^{1-1/(3r_0+1)-2/(\bar r+r_0)} \le \bar C.$$

Next is the estimate for $\mathrm{Err}_{n+1}(t_n,\bar r)$:

$$\mathrm{Err}_{n+1}(t_n,\bar r) \le C\Big(\varepsilon_n^2 + \varepsilon_n^{\frac43+(\frac23-2\frac{r_0+1}{r_0+\bar r})} + \varepsilon_n^{\frac43+(-\frac73+2\frac{\bar r}{3r_0+1})} + \varepsilon_n^{\frac43+(\frac23-\frac{2r_0}{3r_0+1}-\frac{2}{r_0+\bar r})} + \varepsilon_n^{\frac43+(\frac53-\frac{2r_0}{3r_0+1}-\frac{3}{r_0+\bar r})}\Big).$$

From this it easily follows that the assumption $\bar r = 8r_0+4$ on $\bar r$ suffices for

$$\mathrm{Err}_{n+1}(t_n,\bar r) < \varepsilon_n^{4/3} = \varepsilon_{n+1}. \tag{7.1}$$

This will hold for any constant in place of C, the only difference would be a possibly smaller upper bound for ε_0.

Now by (e) of Proposition 6.2

$$\begin{aligned} K_{n+1}(t_n,\bar r) &= \Big(1 + C\,t_n^{r_0}\,\varepsilon_n^{1-1/(\bar r+r_0)}\,\delta_{\bar r,n}^{1/(\bar r+r_0)}\Big)K_n \\ &= \Big(1 + C\,\varepsilon_n^{1-r_0/(3r_0+1)-2/(\bar r+r_0)}\Big)K_n < \big(1+\varepsilon_n^{1/2}\big)K_n \end{aligned}$$

by our choice of $\bar r$ and for ε_0 sufficiently small. Then due to rapid convergence of ε_n there exists a constant $\bar K$, such that for each n,

$$K_n = K_0 \prod_{k=1}^{n-1}\big(1+(\sqrt{\varepsilon_0})^{k^{4/3}}\big) \le \bar K.$$

The constant $\bar K$ varies with the upper bound for ε_0, but a large upper bound is sufficient to make $\bar K$ independent of all other constants in the scheme, for example $\varepsilon_0 < 1/2$ suffices.

Thus, by (e) of Proposition 6.2, for all n,

$$\|W^n(\lambda)\|_{0,(2)} \le K_n < \bar K.$$

By (d) of the Proposition 6.2,

$$\|\Phi^{n+1}-\Phi^n\|_{(1)} \le K_n\varepsilon_n^{1-r_0/(3r_0+1)} + \mathrm{Err}_{n+1}(t_n,\bar r) < \varepsilon_n^{2/3}.$$

This implies that $\|\Phi^n - \mathrm{Id}\|_{(1)} \le \varepsilon_0$ for every n. In particular, if $\varepsilon_0 < R/2$ where R is the constant in Lemma 6.1, then $\Phi^n \in \mathscr{O}$ so the condition (2) of Proposition 6.2 holds for all n.

Now from (d) of Proposition 6.2

$$\begin{aligned} 2\bar K\|\lambda_{n+1}-\lambda_n\| \le{}& C\bar K\,\mathrm{Err}_{n+1}(t_n,\bar r) + C\bar K^4\varepsilon_n^{2-2r_0/(3r_0+1)} \\ &+ 2C\bar K^2\varepsilon_n^{7/3-r_0/(3r_0+1)} + C\bar K^2(\mathrm{Err}_{n+1}(t_n,\bar r))^2 < \varepsilon_n^{4/3}. \end{aligned}$$

Consequently, λ_n converges to some $\bar\lambda \in B$, such that for all n

$$2\bar K\|\bar\lambda - \lambda_n\| \le 2\varepsilon_{n+1}.$$

Combining this with (7.1), for all n

$$\|W^n(\bar\lambda)\|_0 \le \varepsilon_n. \tag{7.2}$$

Now, for H we have an estimate in the C^1 norm

$$\begin{aligned}\|H^n\|_1 \le C t_n^{2r_0} \|W^n(\lambda_n)\|_1 &\le C\varepsilon_n^{-1/(3r_0+1)}\,\varepsilon_n^{1-1/(\bar r+r_0)}\,\delta_{\bar r,n}^{1/(\bar r+r_0)}\\ &\le C\varepsilon_n^{1-(\bar r+7r_0+2)/[(\bar r+r_0)(3r_0+1)]}.\end{aligned}$$

Thus, for all n

$$\|H^n\|_1 \le \varepsilon_n^{1/2}. \tag{7.3}$$

Recall that $h^n := \exp H^n$. Let $\mathscr{H}_n = h^n \circ \cdots \circ h^0$. Since the generators of the action $\tilde\alpha^{n+1}(\bar\lambda) = \mathscr{H}_n^{-1} \circ \tilde\alpha^0(\bar\lambda) \circ \mathscr{H}_n$ are by the construction $U + W^{n+1}(\bar\lambda)$ and they satisfy the estimate (7.2) for all n, it follows that the sequence of actions $\tilde\alpha^n(\bar\lambda)$ converges to the action α in C^0 topology.

The maps T_n acting on the generators $U+W^n(\bar\lambda)$ are only coordinate changes which due to (f) of Proposition 6.2 accumulate to an ε_0-small coordinate change of α but do not change the action α. This cumulative coordinate change is the map π which appears in the statement of Theorem 1.2.

By (7.3), $\mathscr{H}_n$ converges in C^1 topology to a C^1 conjugacy $\mathscr{H}$ between $\tilde\alpha^0(\bar\lambda)$ and α. To see that the constructed conjugacy $\mathscr{H}$ is of class C^∞, interpolation inequalities are applied exactly as in [6, end of Section 5.4] and [14].

This completes the proof of Theorem 1.2.

8. FEW WORDS ON POSSIBILITIES OF OBTAINING THEOREM 1.2 VIA SOME VARIANT OF THE IMPLICIT-FUNCTION THEOREM

The sequences of operators described in Section 1.2 are similar to the sequences considered in [7] for discrete group actions. In the case of a smooth action of a finitely generated finitely presented group, tame splitting of the linearized sequence at one point implies tame splitting of the linearized sequence in a neighborhood of that point (as in [1], for example) and this implies that the nonlinear sequence is tamely exact by a general theorem of Hamilton [9], which is labeled in [9] as Nash–Moser theorem for exact sequences. This then implies full local rigidity for the action. Now for continuous group actions this does not help since the first cohomology over the action (even when it is well understood) is not trivial, so one cannot have a tame splitting for the linearized sequence, except on a possibly smaller space, namely on the trivial cohomology class. But there is no reason why this class would be invariant under conjugation (in fact, it is not), therefore Hamilton's theorem does not apply.

Another possibility is to try to use one of the (generalized) Implicit-Function Theorems. For example, the one due to Zehnder in [16]. The context in this direction would be as follows.

Let $U = (U_1,\dots,U_d)$ be the d-tuple of commuting generating vector fields for the action α in one of the cases described in Section 1.3. Let $V = (V_1,\dots,V_d)$ be the d-tuple of "opposite" directions (see Section 1.3). Let $W\colon \mathbb{R}^d \to \mathrm{Vect}^\infty(X)^d$

be a continuously differentiable family such that for each λ, d-tuple of vector fields $U+W(\lambda)$ are pairwise commutative, let $\lambda \in \mathbb{R}^d$, $H \in \mathrm{Vect}^\infty(X)$, $\pi \in M^{d\times d}$ ($d\times d$ matrices) and let U^T denote the transpose of the d-tuple U. Let $W_0\colon \mathbb{R}^d \to \mathrm{Vect}^\infty(X)^d$ be the initial family defined by $W_0(\lambda) = (U_1+\lambda_1 V_1, \dots, U_d+\lambda_d V_d)$. Define the following map

$$\mathscr{F}(W,(\lambda,H,\pi)) = (\exp H)_*(U+W(\lambda)-\pi U^T)-U. \tag{8.1}$$

This map is well defined for all H in a sufficiently small neighborhood of 0. Clearly, $\mathscr{F}(W_0,(0,0,0)) = 0$.

IMPLICIT-FUNCTION STATEMENT. *There exist a neighborhood $\mathscr{O}$ of U and a neighborhood $\mathscr{U}$ of $(0,0,0)$ in $\mathbb{R}^d \times \mathrm{Vect}^\infty(X)\times M^{d\times d}$ and a continuous map $f\colon \mathscr{O}\to\mathscr{U}$ such that $\mathscr{F}(W,f(W))=0$ for all $W\in\mathscr{O}$.*

If this Implicit-Function Statement holds, then this implies that for the action $\tilde{\alpha}$ of $\mathbb{R}^d$ on X which is generated by d-tuple of vector fields $U+W(\lambda)-\pi U$ with W sufficiently small, there exists a transformation π there exists a parameter λ and the diffeomorphism $h=\exp H$ which conjugates the action generated by $U+W(\lambda)-\pi U$ to U.

The main condition for obtaining this Implicit-Function Statement via the generalized Implicit-Function Theorem of Zehnder is that the derivative of $\mathscr{F}$ at $(W_0,(0,0,0))$ with respect to the second variable is approximately invertible in a neighborhood of $(W_0,(0,0,0))$ and exactly invertible at $(W_0,(0,0,0))$. This derivative operator which needs to be inverted is $[H,U]+\lambda\cdot V+\pi U^T$. The contents of the Section 4 and Section 5 show that this derivative *restricted to the space* $\lambda=0$ has an approximate right inverse at $(W_0,(0,0,0))$, which is in fact never exact (due to nontrivial relations in the acting group), and this approximate right inverse can be constructed only for the data which has zero averages in V directions. If there were no extra relations in the acting group, *i.e.,* if there were no commutativity condition for the data, there would be no problem: the averages of the data in the V directions would be precisely the value of λ and after subtracting the averages from the data one is in the situation when $\lambda=0$ as well as the averages in the V directions for the data, so by our results there would be an approximate inverse. This would be similar to the approach for small perturbations of constant Diophantine vector fields on tori of dimension greater than 2. Small constant modifications of the data accumulate then to a constant modification of the initial perturbations. Modulo this constant modification, there is a smooth conjugacy to the initial action. But there is a catch in the higher-rank case. The catch is that such a modification of a *commuting* d-tuple of vector fields by a constant for the examples in this paper gives a *noncommuting* d-tuple of vector fields, and for a noncommuting data our results give no approximate inverse for the derivative operator. Commutativity is crucial for the existence of the approximate solution. Thus, the classical approach in this situation would lead us inevitably out of the space of $\mathbb{R}^d$ actions. This is the main reason why Zehnder's generalized Implicit-Function Theorem does not apply in this case.

9. COMMENTS ON UNIPOTENT AND OTHER ACTIONS

9.1. **Actions on $SL(2,\mathbb{C})/\Gamma$ and $SL(2,\mathbb{R})\times SL(2,\mathbb{C})/\Gamma$.** An immediate next step in our program is proving rigidity for the unipotent actions $\alpha_{\mathbb{R}\mathbb{C}}$ and $\alpha_{\mathbb{C}}$ from Examples 2 and 3 from Section 1.3. As we pointed out in that section, the remaining steps are Step 5(b) (tame estimates) for the scalar equation and Step 6 (tame splitting).

9.1.1. $SL(2,\mathbb{C})$ *case.* In Mieczkowski's approach, the cohomological equation for Example 3 is solved using analytic description of irreducible representation of $SL(2,\mathbb{C})$ in certain spaces of C^∞ functions on the complex plane and identifying obstructions within each irreducible subspace. For an individual one-parameter subgroup of the action $\alpha_{\mathbb{C}}$ in addition to the obstruction "at infinity" there is a whole continuous family of invariant distributions and their simultaneous vanishing for one-cocycles is achieved by a proper version of the higher-rank trick.

Mieczkowski does not obtain tame estimates for solutions in full Sobolev or C^r norms. Instead he considers two embeddings of $SL(2,\mathbb{R})$ into $SL(2,\mathbb{C})$ whose Lie algebras generate full Lie algebra $sl(2,\mathbb{C})$ via brackets, but not linearly. For Sobolev norms associated to those embeddings, he obtains tame estimates that are sufficient to conclude that the solutions are C^∞ via elliptic regularity but with the loss of half of the derivatives. The one missing direction generates a compact subgroup that commutes with the $\mathbb{R}^2$ unipotent homogeneous action $\alpha_{\mathbb{C}}$. There is a general argument suggested by L. Flaminio that allows to obtain tame estimates in the centralizer direction *a priori.*

Splitting however represents an additional problem since the space of invariant distributions for one-parameter subgroups of $\alpha_{\mathbb{C}}$ is infinite-dimensional within each irreducible space for $SL(2,\mathbb{C})$. Those obstructions have a form of integrals over lines parallel to a certain direction for a function of two variables with a certain behavior at infinity that guarantees existence of such integrals. While there is no natural projection to the space of cocycles (that in this case correspond to the functions whose integrals along the lines in question vanish) there is a natural strategy for constructing a tame splitting by distributing the value of the obstructions along certain intervals using smooth kernels. This is a work in progress.

It is also worth pointing out a peculiar feature of the $SL(2,\mathbb{C})$ case. For exactly one of the principal series representations (with parameter values $(0,2)$) the higher-rank trick does not eliminate the obstruction at infinity for vanishing of a cocycle. The multiplicity d of this representation in the decomposition of $L^2(SL(2,\mathbb{C})/\Gamma)$ into irreducible subspaces is equal to the first Betti number of the group Γ. If this number is not zero, there are two possibilities: (i) to consider $2+d$-parametric families of actions near the unipotent action $\alpha_{\mathbb{C}}$ and find an action conjugate to $\alpha_{\mathbb{C}}$ up to an automorphism of $\mathbb{R}^2$ or (ii) to start with a standard d-dimensional family of time changes for the action $\alpha_{\mathbb{C}}$ generated by harmonic forms with varying cohomology, and then consider a 2-parametric family of perturbations of $\alpha_{\mathbb{C}}$ and look for a parameter value that produces an action conjugate to an element of this d-dimensional family. Once the problem of tame splitting is resolved, both of these approaches can be carried out.

9.1.2. *$SR(2,\mathbb{R})\times SL(2,\mathbb{C})$ case.* For the $\mathbb{R}^3$ action $\alpha_{\mathbb{R}\mathbb{C}}$ in Example 2, Mieczkowski proves in [12] that the sequence $\mathscr{L}\to\mathscr{M}$ is tamely exact, so Step 5(b) follows in this case. Once tame splitting for $SL(2,\mathbb{C})$ is proved Theorem 3.2 will imply tame splitting for this case.

To summarize, the only ingredient left to prove the counterpart of Theorem 1.2 for the actions in Examples 2 and 3 is a splitting of a pair of scalar functions (f,g) within an irreducible representation of $SL(2,\mathbb{C})$ into a coboundary $(\mathscr{L}_{U_1}h,\mathscr{L}_{U_2}h)$, and the remainder that can be estimated tamely through the norms of $\mathscr{L}_{U_2}f-\mathscr{L}_{U_1}g$, where

$$U_1=\begin{pmatrix}0&1\\0&0\end{pmatrix},\qquad U_2=\begin{pmatrix}0&i\\0&0\end{pmatrix}.$$

9.2. **Other unipotent actions.** In a recent paper, Felipe Ramirez [15] considered the following generalization of our Example 1: he assumes that $SL(2,\mathbb{R})\times SL(2,\mathbb{R})$ (or, more generally, its finite cover) embeds into a noncompact simple Lie group G with finite and considers the action of the upper-diagonal unipotents in $SL(2,\mathbb{R})\times SL(2,\mathbb{R})$ on G/Γ. Theorem B' in [15] asserts that any C^∞ cocycle over that action is C^∞ cohomologous to a constant cocycle. If one can add tame estimates and tame splitting to Ramirez's result our scheme becomes applicable. Tame estimates in the $SL(2,\mathbb{R})\times SL(2,\mathbb{R})$ direction follow the same way as in the present paper. One can also obtain tame estimates in the centralizer direction by the argument suggested by L. Flaminio. Once tame estimates are obtained tame splitting for functions follows from Theorem 3.2. However, deduction of the splitting for vector fields from splitting for functions does not directly follow from Proposition 4.1 and needs to be reworked. These observations lead to new cases of rigidity that will be discussed in detail in a subsequent paper joint with Flaminio.

Not surprisingly, not all higher-rank abelian unipotent actions on homogeneous spaces of semisimple groups can be treated with the use of representation theory for $SL(2,\mathbb{R})$ or $SL(2,\mathbb{C})$ where we have good insights into the structure of invariant distributions and solutions of cohomological equations. A simple example, where cohomological equations are not well understood, is the action by the $\mathbb{R}^2$ subgroup of $SL(3,\mathbb{R})$ generated by $\mathrm{Id}+e_{12}$ and $\mathrm{Id}+e_{13}$, where e_{ij} is the matrix with 1 in the place (i,j) and zeroes elsewhere.

9.3. **Some partially hyperbolic actions.** While partial hyperbolicity in dynamics in its most general form simply means that the linearized system is uniformly hyperbolic in some directions and nonhyperbolic in others, the genuine partially hyperbolic paradigm strives to explore situations where hyperbolicity is somehow prevalent. As examples of this approach to local differentiable rigidity one may consider the results based on the "geometric method" and applications of algebraic K-theory started in our work [3, 4, 2] and recently extended by Zhenqi Wang [17, 18].

Approach developed in this paper is applicable to certain actions where partial hyperbolicity is present but could not be considered prevalent since in particular some elements of the action are nonhyperbolic. We postpone detailed description of the relevant situations to a future paper and now mention only another "model" example.

Mixed example on $SL(2,\mathbb{R}) \times SL(2,\mathbb{R})$. Consider the $\mathbb{R}^2$ action β on $SL(2,\mathbb{R}) \times SL(2,\mathbb{R})/\Gamma$ generated by commuting vector fields U_+^1 and U_0^2 as in Example 1. The first generator is unipotent, hence the corresponding one-parameter subgroup acts parabolically. The second generator is partially hyperbolic with isometric action in the neutral directions. Mieczkowski showed vanishing of the obstructions and solved the scalar cohomological equation in this case, see [12, Theorem 2]; in his thesis [13] he obtained necessary tame estimates for the diagonal action on $SL(2,\mathbb{R})$ parallel to the estimates in [8] for unipotent actions.

The only difference with our model unipotent action $\alpha_{\mathbb{R}\mathbb{R}}$ in Example 1 is in the structure of linearized conjugacy equations. Due to the presence of a partially hyperbolic generator some of the equations are twisted (compare with the discrete time partially hyperbolic case in [6]). Accordingly, considerations of Section 4 do not apply directly and have to be modified. This modifications is fairly routine, albeit somewhat tedious. The twisted equation related to the partially hyperbolic generator always has a solution but obstructions appear in the bootstrap regularity and their vanishing again involves invoking a version of the higher-rank trick. Those obstructions are not exactly invariant distributions but have similar nature (again compare with [6]). An interesting difference with our model unipotent case is that one only needs one-parametric families of perturbations since homogeneous actions conjugate to β up to a an automorphism of the acting group have codimension one among homogeneous $\mathbb{R}^2$ actions on $SL(2,\mathbb{R}) \times SL(2,\mathbb{R})/\Gamma$.

Once tame estimates are obtained for the twisted equation, tame splitting is obtained by the construction from the proof of Theorem 3.2. After that considerations from Sections 6 and 7 apply directly.

Detailed discussion of this and other partially hyperbolic cases that can be treated with variations of our method will appear in a subsequent paper.

REFERENCES

[1] E. J. Benveniste, *Rigidity of isometric lattice actions on compact Riemannian manifolds*, Geom. Funct. Anal., **10** (2000), 516–542.

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

[3] D. Damjanović and A. Katok, *Local rigidity of restrictions of Weyl chamber flows*, C. R. Math. Acad. Sci. Paris, **344** (2007), 503–508.

[4] D. Damjanović and A. Katok, *Local rigidity of partially hyperbolic actions. II. The geometric method and restrictions of Weyl chamber flows on $SL(n,R)/\Gamma$*, Int. Math. Res. Notes, 2010.

[5] D. Damjanović and A. Katok,*Local rigidity of actions of higher-rank abelian groups and KAM method*, Electron. Res. Announc. Amer. Math. Soc., **10** (2004), 142–154.

[6] D. Damjanović and A. Katok, *Local rigidity of partially hyperbolic actions I. KAM method and $\mathbb{Z}^k$ actions on the torus*, Ann. of Math. (2), **172** (2010), 1805–1858.

[7] D. Fisher, *First cohomology and local rigidity of group actions*, Ann. of Math., to appear.

[8] L. Flaminio and G. Forni, *Invariant distributions and time averages for horocycle flows*, Duke Math. J., **119** (2003), 465–526.

[9] R. S. Hamilton, *The inverse function theorem of Nash and Moser*, Bull. Amer. Math. Soc. (N.S.), **7** (1982), 65–222.

[10] B. Hasselblatt and A. Katok, *Principal structures*, in "Handbook in Dynamical Systems, **1A**, North-Holland, Amsterdam, (2002), 1–203.

[11] A. Katok and R. Spatzier, *Differential rigidity of Anosov actions of higher-rank abelian groups and algebraic lattice actions*, Proc. Steklov Inst. Math., **216** (1997), 287–314.

[12] D. Mieczkowski, *The first cohomology of parabolic actions for some higher-rank abelian groups and representation theory*, J. Mod. Dyn., **1** (2007), 61–92.

[13] D. Mieczkowski, "The Cohomological Equation and Representation Theory," Ph.D thesis, The Pennsylvania State University, 2006.

[14] J. Moser, *On commuting circle mappings and simoultaneous Diophantine approximations*, Math. Z., **205** (1990), 105–121.

[15] F. Ramirez, *Cocycles over higher-rank abelian actions on quotients of semisimple Lie groups*, J. Mod. Dyn., **3** (2009), 335–357.

[16] E. Zehnder, *Generalized implicit-function theorems with applications to some small divisor problems. I*, Comm. Pure Appl. Math., **28** (1975), 91–140.

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

[18] Z. Wang, *New cases of differentiable rigidity for partially hyperbolic actions: Symplectic groups and resonance directions*, J. Mod. Dyn., **4** (2010), 585–608.

DANIJELA DAMJANOVIC <dani@rice.edu>: Department of Mathematics, Rice University, 6100 Main St., Houston, TX 77005, USA

ANATOLE KATOK <katok_a@math.psu.edu>: Department of Mathematics, The Pennsylvania State University, University Park, PA 16802, USA